bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Complexity of evolutionary equilibria in static fitness landscapes

Artem Kaznatcheev[1,2]

[1]Department of Computer Science, University of Oxford, Oxford, UK
[2]Department of Translational Hematology & Oncology Research, Cleveland Clinic, Cleveland, OH, USA

Experiments show that evolutionary fitness landscapes can have a rich combinatorial structure due to epistasis and yet theory assumes that local peaks can be reached quickly. I introduce a distinction between easy landscapes where local fitness peaks can be found in a moderate number of steps and hard landscapes where finding evolutionary equilibria requires an infeasible amount of time. Hard examples exist even among landscapes with no reciprocal sign epistasis; on these, strong selection weak mutation dynamics cannot find the unique peak in polynomial time. On hard rugged fitness landscapes, no evolutionary dynamics – even ones that do not follow adaptive paths – can find a local fitness peak quickly; and the fitness advantage of nearby mutants cannot drop off exponentially fast but has to follow a power-law that long term evolution experiments have associated with unbounded growth in fitness. I present candidates for hard landscapes at scales from singles genes, to microbes, to complex organisms with costly learning (Baldwin effect). Even though hard landscapes are static and finite, local evolutionary equilibrium cannot be assumed.

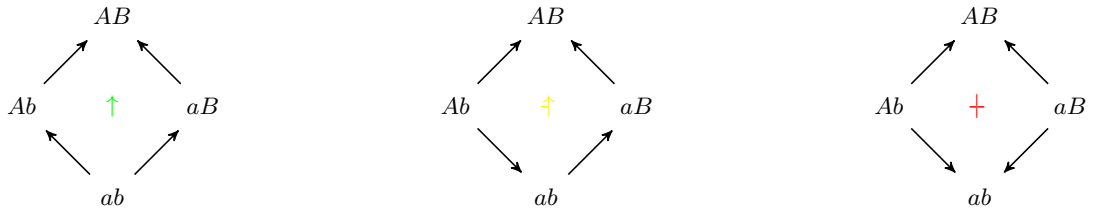
Fitness landscapes combine numeric fitnesses and a mutation-graph into a combinatorially structured genetic space where each vertex is a possible genotype (or phenotype). The numeric structure is given by a function that maps each genotype to a fitness; typically represented as a non-negative real number and having different physical operationalizations in different experimental systems. The mutation-graph specifies which genotypes are similar, typically as edges between any two genotypes that differ in a single mutation. We usually imagine fitness landscapes as hills or mountain ranges, and continue to assume – as Wright [1] originally did – that on an arbitrary landscape "selection will easily carry the species to the nearest peak". A local peak might not be the tallest in the mountain range, so reaching it can prevent us from walking uphill to the tallest peak. This has directed much of the work on fitness landscapes toward how to avoid these sub-optimal peaks or how a population might move from one peak to another [1, 2]. But we seldom consider that even reaching that local peak might be impossible in any reasonable amount of time. Instead, a careful analysis (formal mathematical proofs for all statements are available in the supplementary appendix (SA)) shows that it is more accurate to imagine fitness landscapes as mazes with the evolutionary equilibria as exits. Evolution cannot see far in the maze and must rely only on local information. In hard mazes, we can end up following exponentially long winding paths to the exit because we cannot spot the shortcuts. Or, worse yet, the hardest mazes might not have any shortcuts and even the most clever and farsighted navigator won't know how to reach the exit in a feasible amount of time.

Epistasis & semi-smooth landscapes

What makes some fitness landscapes difficult to navigate is that the effects of mutations at different loci can interact with each other. Epistasis is a measure of the kind and amount of inter-loci interactions. If the fitness effect of a mutation $a \to A$ can have a different sign depending on the genetic background b or B of another locus then these two loci are said to have sign epistasis (Figure 1b and SA Definition 4). If both mutations have one sign on their own, but the opposite sign together – either bad + bad = good or good + good = bad – then the landscape has reciprocal sign epistasis ([3, 4]; Figure 1c and SA Definiton 5). A classic example of reciprocal sign epistasis is a lock-and-key, changing just one of the lock or the key breaks the mechanism, but changing both can be beneficial. Finally, magnitude epistasis (positive and negative; SA Definition 3) are inter-loci interactions that deviate from additivity, but do not change the sign of fitness effects. This type of epistasis does not change the combinatorial structure of the landscape or the computational difficult of finding equilibra. As such, I treat it simply as a lack of sign-epistasis.

A landscape without sign epistasis – like the *Escherichia coli* β-lactamase fitness landscape measured by Chou *et al.* [5] in figure 2a – is called *smooth* ([4, 6] and SA B), so let's call a fitness landscape *semi-smooth* if it has no reciprocal sign epistasis. The fitness graphs ([4] and SA A) of semi-smooth fitness landscapes are equivalent to acyclic unique sink orientations previously defined by Szabó & Welzl [7] for the analysis of simplex algorithms (SA Definition 10 and Proposition 11). Since reciprocal sign epistasis is a necessary condition for multiple peaks (SA Corollary 9 and Poelwijk *et al.* [3]), smooth and semi-smooth fitness landscapes have a single peak x^*. Further, there are short adaptive paths in both: from any genotype x there always exists some adaptive path to x^* of length equal to the number of loci on which x and x^* differ (SA Theorem 12). This means that an omniscient mutator could always find a short adaptive path to the peak. But unlike smooth landscapes, in a semi-smooth

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



(a) No epistasis (b) Sign epistasis (c) Reciprocal sign epistasis

Figure 1: **Three different kinds of epistasis possible in fitness graphs.** Arrows are directed from lower fitness genotypes towards mutationally adjacent higher fitness genotypes. Genes a, A and b, B are labeled such that fitness $w(AB) > w(ab)$. In the center of each graph is a marker for the type of epistasis, the marker's various rotations & reflections cover the cases where AB does not have the highest fitness. For this more exhaustive classification and discussion see SA Figure 4 and SA A.1

landscape not every shortest path is adaptive and not every adaptive path is short. And since evolution does not have the foresight of an omniscient mutator, it is important to check which adaptive path evolutionary dynamics will follow.

When selection is strong and mutation is weak (SSWM dynamics; introduced by Gillespie [8]), the population can be represented as a single point on the fitness graph with an evolutionary step corresponding to a selective sweep that moves the population to a neighbouring genotype with higher fitness. A number of rules have been suggested for which fitter neighbour will take over the population [9]; the two most common rules are to select a fitter mutant uniformly at random, or to select the fittest mutant. All SSWM rules will quickly find the fitness optimum in a smooth fitness landscape. But there exist semi-smooth fitness landscapes such that when starting from a random initial genotype, an exponential number of evolutionary steps will be required for either the random fitter mutant ([10]; SA Theorem 13) or fittest mutant (SA Theorems 18 and 22; see Figure 3 for a small example landscape on six loci) dynamics to find evolutionary equilibrium. In other words, even when there is a single peak and adaptive paths of minimal length to it, SSWM dynamics can take exponential time to find that peak.

These results show that combinatorial complexity, without the need for suboptimal local peaks, can be enough to stop evolution from reaching equilibrium within a reasonable timescale. Although evolutionary time is long, it is not reasonable to think of it as exponentially long. For example, exponential search over biallelic genotypes on just 60 loci at a rate of 1 per second would require more seconds than the time since the Big Bang. With 360 loci, we wouldn't finish even if all the atoms in the universe were searching at that rate in parallel. Thus, I introduce a distinction between *easy* and *hard* families of fitness landscapes. If we can guarantee for any landscape in the family that an evolutionary equilibria can be found in a time the scales as a polynomial in the number of loci – as is the case for smooth fitness landscapes – then that is an easy family of landscapes. We'll call a family of landscapes hard if we can show that the family contains landscapes where finding an evolutionary equilibrium requires a super-polynomial amount of time – as I showed above for semi-smooth fitness landscapes.

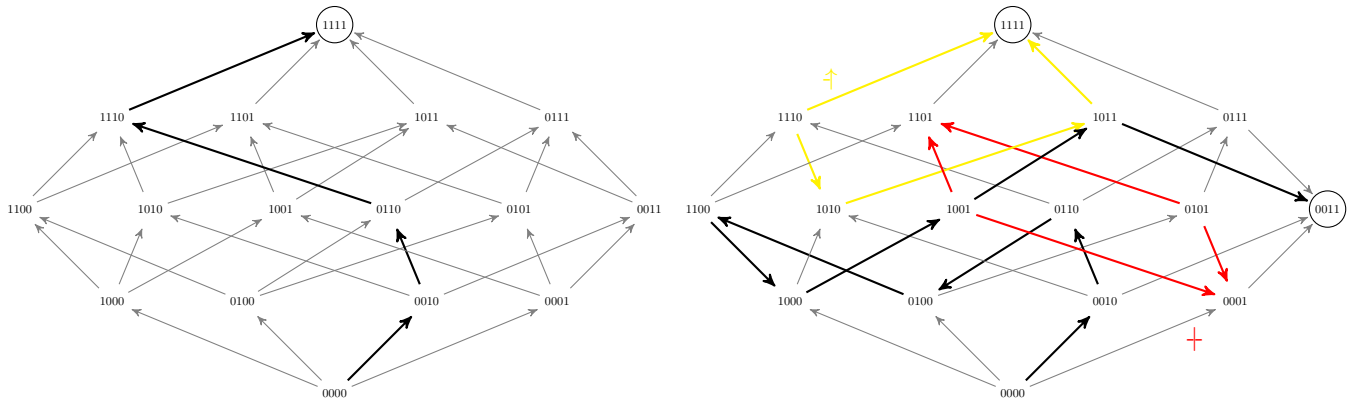
Given their exponential size, it is impossible to completely measure whole fitness landscapes on more than a few nucleotides. But with improvements in high-throughput second-generation DNA sequencing there is hope to measure local fitness landscapes of a few mutations away from a wildtype. Puchta *et al.* [11] estimated the fitness of 981 single-step mutations of a 333-nucleotide small nucleolar RNA (snoRNA) gene in yeast. They found no neighbours fitter than the wild type gene, this suggests that the snoRNA gene's fitness landscape is easy. In constant, Li *et al.* [12] estimated the fitness of 207 single-step mutants of a 72-nucleotide transfer RNA (tRNA) gene, also in yeast, finding two neighbours that are significantly fitter than the wildtype and a number that are fitter but only within experimental noise. Thus, the wildtype tRNA gene is not at a local fitness peak, and suggests this system as a candidate for hard fitness landscapes. Both studies also looked at many 2- and 3-step mutants, and the hard landscape of tRNA was measured to have more than 160 cases of significant sign epistasis [12], with none in the snoRNA landscape [11] mirroring the difference between hard semi-smooth fitness landscapes and easy smooth landscapes.

Rugged landscapes & approximate equilibria

But there exist natural fitness landscapes that are even more complicated than semi-smooth ones. For example, we know that they can contain reciprocal sign epistasis like the Lozovsky *et al.* [13] *Plasmodium falciparum* dihydrofolate reductase fitness landscape in figure 2b. Although there is not enough data to justify probability distributions over large landscapes (for discussion, see SA D.1), the biological intuition is that natural landscapes are a little rough and have multiple peaks. The NK-model is a family of fitness landscapes [15, 16] that was introduced to study this ruggedness. It allows tuning epsitasis: the fitness contribution of each of the n loci depends not only on its gene, but also on the genes at up to K other loci (SA Definition 23).

For $K \geq 2$, the NK-model has hard fitness landscapes where from some initial genotypes, any adaptive walk to a local peak is exponentially long (SA Corollary 26). On such landscapes, any adaptive evolutionary dynamic – including, but not limited to, all the SSWM dynamics we've considered

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



(a) *E. coli* β-lactamase

(b) *P. falciparum* dihydrofolate reductase

Figure 2: **Two examples of empirical biallelic fitness landscapes on four loci.** Arrows are directed from lower fitness genotypes to higher and evolutionary equilibria are circled. Examples of adaptive dynamics are highlighted with thick black arrows. Figure 2a, based on the *E. coli* β-lactamase data of Chou *et al.* [5], is a smooth landscape with no sign epistasis. Thus, it contains a single optimum (1111). Figure 2b is based on Lozovsky *et al.* [13]'a *P. falciparum* dihydrofolate reductase growth rate data in the absence of pyrimethamine. It has two peaks (0011 & 1111) and both single sign (an example in yellow; $\nearrow$) and reciprocal sign epistasis (example in red; $+$). Based on Szendro *et al.* [14]'s Figure 1.

so far – requires an exponential number of steps to reach evolutionary equilibrium. Even if an omniscient mutator could always choose the most clever adaptive single mutation to arise, the adaptive path would be unbounded over polynomial timescales. The most notable natural candidate for such a hard case might be the landscapes with unbounded growth in fitness observed in the *E. Coli* long-term evolutionary experiment [17] and similar model systems [18]. Current theoretical accounts of this rely on models where unbounded growth is directly built-in and the higher-order combinatorial structure of the mutation graph is ignored. Instead, we can acknowledge combinatorial complexity and explain these results by recognizing the existence of hard families among classic models of finite static fitness landscapes.

To better integrate the numeric structure of fitness, consider a genotype x to be at an s-approximate equilibrium [19] if each of x's mutational neighbours y have fitness $w(y) \leq (1+s)w(x)$ (SA Definition 28). On the hard rough fitness landscapes above – and also the winding semi-smooth landscape dynamics inset of Figure 3 and SA C.2 – this selection coefficient [20] drops off at the slow rate of $s(t) \approx 1/t$ for fittest mutant dynamics. In fact, if a landscape family's smallest possible adaptive fitness step ϵ_w is at least an exponential fraction of the max possible fitness $w_{\max}$ (i.e. if $\epsilon_w \geq w_{\max}/2^{\text{poly}(n)}$) then fittest mutant dynamics will encounter an s-approximate evolutionary equilibrium with moderately small s in a moderate number of mutational steps (polynomial in n and $1/s$; SA Theorem 30). Thus, on even the hardest family of landscapes, $s(t)$ can decay as fast as a power law. However, on hard fitness landscapes, it is not possible for $s(t)$ to decrease faster. In particular, on hard fitness landscapes, it is not possible to find an s-approximate evolutionary equilibrium for very small s in a feasible amount of time (i.e. not possible in time polynomial in n and $\ln 1/s$; SA Corollary 32) or to have $s(t) \approx e^{-t}$ (SA Theorem 33).

On hard landscapes, the selection coefficient decreases as a power law and not at the exponential rate that is typical of equilibration in non-biological systems. Given that s is defined in the same way as the selection coefficient of population genetics [20], these results allow us to link the distinction between easy and hard fitness landscapes to the rich empirical literature on fitness traces and declining fitness gains in microbial evolution experiments [18]. In particular, this slow decay in selection coefficient is consistent with the rule of declining adaptability observed in various microbial long-term evolution experiments [17, 18, 21], suggesting that at least some naturally occurring microbial fitness landscapes might be hard.

Baldwin effect & any evolutionary dynamics

As we move from single genes [11, 12], to microbes [17, 18], and on to large organisms, a richer space of possible evolutionary dynamics opens up. To capture this rich space of possibilities, we need to abstract beyond adaptive dynamics by considering arbitrary mutation operators, demographies, population structures and selection functions – even ones that can cross fitness valleys and distribute the population over many genotypes. To make sure that we've considered all possibilities, I'll model arbitrary evolutionary dynamics as a subset of all polynomial time algorithms. This takes us into the realm of the computational complexity class of polynomial local search (PLS; Johnson et al. [22], Roughgarden [23] and SA D). But even for these most permissive population updating procedures, evolution will sometimes require an infeasible amount of time to find a local fitness peak in the NK-model with $K \geq 2$ (SA Theorem 25 and Corollary 27), or to find an s-approximate equilibrium for very small s (SA Corollary 32). Evolution will be trapped in the maze of hard fitness landscapes and not reach anywhere near the 'exit' of evolutionary

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

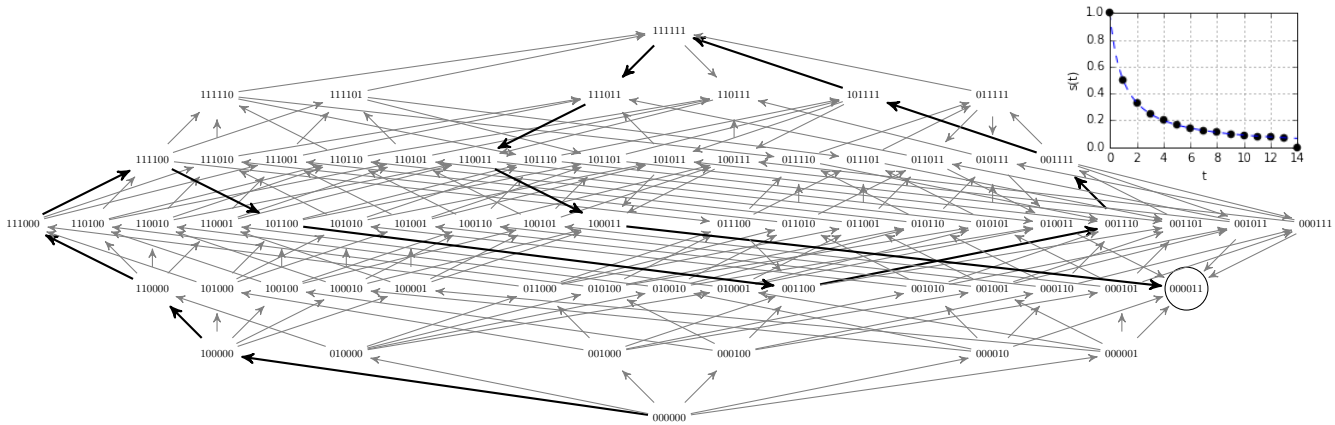


Figure 3: **Winding semi-smooth fitness landscape.** An example on six loci of the winding semi-smooth fitness landscapes from SA C.2 on which the length of the path followed by fittest-mutant SSWM dynamics scales exponentially with the number of loci. Here the black arrows are the fittest available mutation, and the adaptive path takes 14 steps to reach the fitness peak at 000011. For the generalization of this landscape to $2n$-loci, it would take $2^{n+1} - 2$ steps for fittest mutant dynamics to reach the fitness peak at $(00)^{n-1}11$ (SA Theorem 18). Inset is the selection coefficient ($s(t) = \max_{y \in N(x_t) \cup \{x_t\}} \frac{w(y)-w(x)}{w(x)}$; SA D.2) versus mutation step number (t) for the fittest mutant adaptive path.

equilibrium.

This general result is most relevant to paradoxes like the Baldwin effect [24, 25]. As Simpson [25] noted: "[c]haracters individually acquired by members of a group of organisms may eventually, under the influence of selection, be reinforced or replaced by similar hereditary character". For Simpson [25] this possibility constituted a paradox: if learning does not enhance individual fitness at equilibrium and would thus be replaced by simpler non-learning strategies then why do we observe the costly mechanism and associated errors of individual learning? Most resolutions of this paradox, explain the Baldwin effect by focusing on non-static fitness in rapidly fluctuating environments that are compatible with the speed of learning but not with evolutionary adaptation. But these dynamic landscapes are not necessary if we acknowledge the existence of hard static fitness landscapes. Since learning is just another polynomial time algorithm then – even if it can help guide or speed-up evolution – on these hard fitness landscapes the population will still not be able to find an evolutionary equilibrium. Without evolutionary equilibrium, the paradox of costly learning dissolves even in static fitness landscapes. This suggests that if we want a family of natural examples of evolution on hard fitness landscapes among more complex organisms then populations with persistent costly learning might be good candidates.

It is standard to frame adaptationism as "the claim that natural selection is the only important cause of the evolution of most nonmolecular traits and that these traits are locally optimal" [26]. Here, I showed that these are two independent claims. Even if we assume that (1) natural selection is the dominant cause of evolution then – on hard fitness landscapes – it does not follow that (2) traits will be locally optimal. Given the popularity of equilibrium assumptions in evolutionary biology, I expect that a number of other paradoxes in addition to the Baldwin effect could be eased by recognizing the independence of these two claims.

Currently, finding a species away from a local fitness peak is taken as motivation for further questions on what mechanisms cause this discrepancy. In this context, these results provide a general answer: hard landscapes allow adaptationist accounts for the absence of evolutionary equilibrium even in experimental models with static environments – like the tRNA gene in yeast [12] or the long-term evolutionary experiment in *E. coli* [17]. By treating evolution as an algorithm, we see that time is a limiting resource even on evolutionary timescales. These hard landscapes can be finite and deceptively simple – having only limited local epistasis or not having reciprocal sign-epistasis – and yet allow for unbounded fitness growth.

But a system found at a local fitness peak – like the snoRNA gene in yeast [11] – currently merits no further questions. The results in this report show that establishing evolutionary equilibrium should not be the end of the story. We need to also explain what mechanisms make the relevant fitness landscapes easy. The tools of theoretical computer science allow us to refine our logical characterization of such fitness landscapes to guarantee that local peaks can be found in polynomial time. For example, we could consider constraints on the gene-interaction network (SA D.1), or the type of interaction possible between genes [27] to separate easy from hard landscapes. This opens new avenues for both empirical and theoretical work.

On easy landscapes, it is reasonable to assume that evolution finds locally well adapted genotypes or phenotypes.

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We can continue to reason from evolutionary equilibria, debate questions of crossing fitness valleys, and seek solutions to Wright [1]'s problem of "a mechanism by which the species may continually find its way from lower to higher [local] peaks". But with hard landscapes, it is better to think of evolution as open ended and unbounded. We will have to switch to a language of "adapting" rather than "adapted", reason from disequilibrium, and seek mechanisms by which the species selects which unbounded adaptive path to follow. Theoretical computer science and combinatorial optimization offer us the tools to make rigorous this distinction between easy and hard fitness landscapes.

Acknowledgements

I am grateful to Julian Xue for introducing me to the NK-model, and for many helpful comments and discussions; and to Peter Jeavons for extensive feedback and encouragement. This paper benefited from comments by Eric Bolo, Frederic Guichard, Marc Harper, Sergey Kryazhimskiy, Andriy Marusyk, Daniel Nichol, Prakash Panangaden, Joshua Plotkin, Jacob Scott, and participants of the 2014 Computational Theories of Evolution workshop at the Simons Institute for the Theory of Computing. This work was started while I was at the School of Computer Science, McGill University.

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendices

In these appendices, I formally define the concepts introduced within the body of the report and prove the theorems on which the conclusions are based. The structure of the appendices is below:

A Formal definitions of fitness functions, fitness landscapes, fitness graphs, and adaptive. I focus specific attention on epistasis (A.1) because it can be used to define broad families of landscapes, like:

B *Smooth fitness landscapes*: these are the source of a lot of intuition and early model of fitness landscapes. So, I briefly remind the reader of important properties of smooth landscapes.

C *Semi-smooth fitness landscapes*: these share many properties in common with fitness landscapes, and have a characterization theorem that is structured in a similar was to smooth landscapes. However, computationally semi-smooth landscapes, unlike smooth ones, can be hard. In subsection C.1, I use the equivalence of semi-smooth fitness landscapes and acyclic unique sink orientations of cubes to adapt hardness results from the analysis of simplex algorithms. This provides hard landscapes for fitter mutant SSWM dynamics. In the subsequent subsections, I show how to construct hard fitness landscapes for fittest mutant SSWM dynamics from specific start (C.2) and random start (C.3).

D *NK-model of fitness landscapes*: this is a tunable rugged fitness landscape model that – unlike the previous two – can have many peaks. To analyze this model of landscapes, I review the complexity class PLS, show that the NK-model is PLS complete for $K \geq 2$, and discuss the generality of the results. In subsection D.1, I focus on easy instances of the NK-model and provide an intuition for why assuming simple distributions on fitness landscapes is unreasonable and might have made the existence of hard families difficult to spot.

A Fitness landscapes, graphs, and adaptive paths

In 1932, Wright introduced the metaphor of a fitness landscape [1]. The landscape is a genetic space where each vertex is a possible genotype and an edge exists between two vertices if a single mutation transforms the genotype of one vertex into the other. In the case of a biallelic system we have n loci (positions), at each of which it is possible to have one of two alleles, thus our space is the n-bit binary strings $\{0,1\}^n$. We could also look at spaces over larger alphabets; for, example 4 letters for sequence space of DNA, or 20 letters for amino acids; but the biallelic system is sufficiently general for us. A mutation can flip any loci from one allele to the other, thus two strings $x, y \in \{0,1\}^n$ are adjacent if they differ in exactly one bit. Thus, the landscape is an n-dimensional hypercube with genotypes as vertexes. The last ingredient, fitness, is given by a function that maps each string to a non-negative real number. For the purposes of this report, the exact fitness values or their physical interpretations do not matter. Only their rank-ordering matters.

Individual organisms can be thought of as inhabiting the vertexes of the landscape corresponding to their genotype. And we imagine evolution as generally trying to 'climb uphill' on the landscape by moving to vertexes of higher fitness.

Definition 1. In a fitness landscape with fitness f, a path $v_1 ... v_t$ is called *adaptive* if each v_{i+1} differs from v_i by one bit and $f(v_{i+1}) > f(v_i)$.

For the most general evolutionary dynamics, the paths taken don't have to be strictly increasing in fitness; i.e. they don't have to necessarily be adaptive. If any evolutionary dynamic produces only adaptive paths, though, then it is called an *adaptive dynamic*.

Sometimes it is useful to represent a fitness landscape as a *fitness graph* by replacing the fitness function by a flow: for adjacent genotypes in the mutation-graph, direct the edges from the lower to the higher fitness genotype. This results in a characterization of fitness landscapes of a biallelic system as directed acyclic graphs on $\{0,1\}^n$. Fitness peaks would correspond to sinks, and adaptive paths to paths that follow the edge directions. I will consider a population at evolutionary equilibrium if it finds a local peak in the fitness landscape; i.e. sink in the fitness graph. Crona *et al.* [4] introduced this representation into theoretical biology, but fitness graphs have been used implicitly in earlier empirical studies of fitness landscapes [14, 28–30]. Using fitness graphs is particularly useful empirically because it is difficult to quantitatively compare fitnesses across experiments. Although if pairwise competitions are used to build an empricial fitness graph, it is important to verify that the graph is transitive (acyclic) [31]. In theoretical work, the fitness graph approach has made the proofs of some classical theorems relating local structure to global properties easier and shifts our attention to global algorithmic properties of evolution instead of specific numeric properties.

A.1 Epistasis

Epistasis is a measure of the kind and amount of inter-loci interactions. Consider two loci with the first having alleles a or A, and the second b or B. Assume that the upper-case combination is more fit: i.e. $f(ab) < f(AB)$.

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

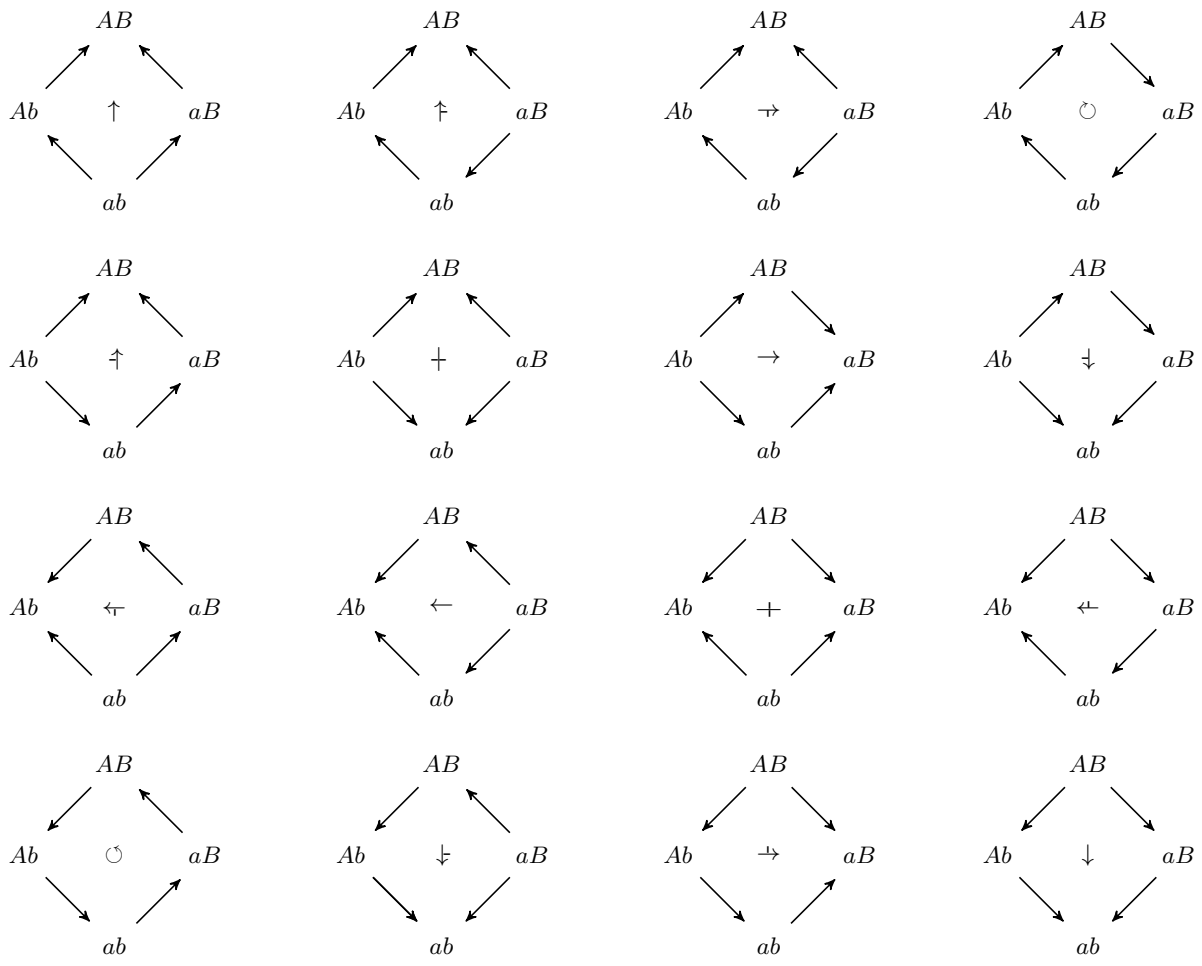


Figure 4: Three different kinds of epistasis possible in fitness graphs: no epistasis (↑), sign epistasis (↰,↱), and reciprocal sign epistasis (┼). Arrows in the fitness graph are directed from lower fitness genotypes towards mutationally adjacent higher fitness genotypes. In the middle of each fitness graph is a symbol showing the kind (and orientation) of epistasis. Note that the bottom left (⟲) and top right (⟳) fitness graphs violate transitivity.

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Definition 2. Two alleles are *non-interacting* if the fitness effects are additive and independent of background: $f(AB) - f(aB) = f(Ab) - f(ab)$, $f(AB) - f(Ab) = f(aB) - f(ab)$.

In *magnitude epistasis* this additivity is broken, but the signs remain: $f(AB) > f(aB) > f(ab)$ and $f(AB) > f(Ab) > f(ab)$. The difference between non-interacting alleles and magnitude epistasis is not invariant under rank-order preserving transformation of the fitness function, thus I will not distinguish between the two types. Throughout the paper, I will use no epistasis to mean both non-interacting alleles and magnitude epistasis.

Definition 3. If $f(AB) > f(aB) > f(ab)$ and $f(AB) > f(Ab) > f(ab)$ then we will say that there is *no epistasis* between those alleles.

A system has sign epistasis if it violates one of the two conditions for magnitude epistasis. For example, if the second locus is b then the mutation from a to A is not adaptive, but if the second locus is B then the mutation from a to A is adaptive.

Definition 4. Given two loci, if $f(AB) > f(aB) > f(ab) > f(Ab)$ then there is *sign epistasis* at the first locus.

Finally, a system has reciprocal sign epistasis if both conditions of magnitude epistasis are broken, or if we have sign epistasis on both loci [32].

Definition 5. Given two loci, if $f(AB) \geq f(ab)$ but $f(ab) > f(Ab)$ and $f(ab) > f(aB)$ then there is *reciprocal sign epistasis* between those two loci.

Figure 4 visualizes all the fitness graphs on two alleles and categorizes the type of epistasis present.

B Smooth fitness landscapes

If a fitness landscape has no sign epistasis then it is a smooth landscape and has a single peak x^* [4, 6]. Every shortest path from an arbitrary x to x^* in the mutation-graph is an adaptive path – a flow in the fitness graph – and every adaptive path in the fitness graph is a shortest path in the mutation graph [4]. Thus, evolution can quickly find the global optimum in a smooth fitness landscape, with an adaptive path taking at most n steps: all smooth fitness landscapes are easy. For an example, see the smooth *Escherichia coli* β-lactamase fitness landscape measured by Chou *et al.* [5] in Figure 2a.

Proposition 6 ([4, 6])**.** *If there is no sign epistasis in a fitness landscape, then it is called a* smooth landscape *and has a single peak x^*. Every shortest path (ignoring edge directions) from an arbitrary genotype x to x^* is an adaptive path, and every adaptive path from x to x^* is a shortest path (ignoring edge directions).*

C Semi-smooth fitness landscapes

Since a smooth landscape is always easy, let's introduce the minimal amount of epistasis: sign epistasis, without any reciprocal sign epistasis.

Definition 7. A *semi-smooth fitness landscape* on $\{0,1\}^n$ with fitness function f is a fitness landscape that has no reciprocal sign epistasis. Such a fitness function f is also called semi-smooth.

Given a landscape on n bits, I will use sublandscape spanned by $S \subseteq [n]$ to mean a landscape on $\{0,1\}^S$ where the indexes in S can vary but the indexes in $[n] - S$ are fixed according to some string $u \in \{0,1\}^{[n]-S}$. Note that the whole landscape is a sublandscape of itself (taking $S = [n]$). Reciprocal sign epistasis between bits i and j corresponds to a sublandscape on $\{i, j\}$ that has two distinct peaks. Now note a couple of important properties of semi-smooth landscapes:

Proposition 8. *If a fitness landscape on $\{0,1\}^n$ has some sublandscape with more than one distinct peak then it has reciprocal sign epistasis.*

The proof will show that a minimal multi-peak sublandscape must have size 2. I will do this by considering longest walks in a sublandscape.

Proof. Let's consider a minimal sublandscape L that has more than one distinct peak: that means that if this sublandscape is spanned by S (i.e. $\{0,1\}^S$) then no sublandscape spanned by $T \subset S$ has multiple peaks.

Since L is minimal, its peaks must differ from each other on each bit in S, for if there was a bit $i \in S$ on which two peaks agreed then that bit could be fixed to that value and eliminated from S to make a smaller sublandscape spanned by $S - \{i\}$ with two peaks. Thus, the minimal multipeak sublandscape has precisely two peaks.

Claim: In a minimal multipeak sublandscape, from each non-peak vertex, there must be a path to each peak.

Let's prove the claim by contradiction: Consider an arbitrary non-peak vertex x, and suppose it has no path to the x^* peak. Since any path from x in L must terminate at some peak, take the longest path from x to the peak y^* that it reaches, and let y be the last step in that path before the peak. Notice that y must only have one beneficial mutation (on bit i), the one to the peak. For if it had more than one beneficial mutation, it could take the non-peak step to y' and then proceed

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from y' to y^* (x^* is not an option by assumption, and there are only two peaks in L) and thus provide a longer path to the peak. Now consider the landscape on $S - \{i\}$, with the ith bit fixed to y_i. Since y_i is the same as x_i^* (both are opposite of y_i^*), x^* is still a peak over $S - \{i\}$, but so is y (since it's only beneficial mutation was eliminated by fixing i to y_i). But this contradicts minimality, so no such x exists.

Now that we know that we can reach each peak from any vertex x, let us again consider the longest path from x to y^* with y as the last step in that path before the peak, and i as the position of the last beneficial mutation. Since all non-peak vertexes must reach both peaks, there must be some other beneficial mutation j from y to x' that eventually leads to x^*. But if x' is not a peak then it must also have a way to reach y^*, but then we could make a longer path, contradicting the construction of y. Thus x' must be the peak x^*.

This means that x^* and y^* differ in only the two bits i and j. But in a minimal multipeaked sublandscape they must differ in all bits, so $S = \{i, j\}$; i.e. this sublandscape is an example of reciprocal sign epistasis. □

Corollary 9. *A fitness landscape without reciprocal sign epistasis has a unique single peak.*

Proof. This follows from the contrapositive of Proposition 8, since the whole landscape is a sublandscape of itself. □

The above theorem can also be restated in the terminology used to analyze simplex algorithms.[7, 10]

Definition 10. A directed acyclic orientation of a hypercube $\{0,1\}^n$ is called an *acyclic unique sink orientation (AUSO)* if every subcube (face; including the whole cube) has a unique sink.

This makes the contrapositive of Proposition 8 into the following proposition:

Proposition 11. *A semi-smooth fitness landscape is an AUSO*

Theorem 12. *A semi-smooth fitness landscape has a unique fitness peak x^* and for any vertex x in the landscape, there exists a path of length $||x^* \oplus x||_1$ (Hamming distance to peak) from x to the peak.*

Proof. The unique peak x^* is just a restatement of Corollary 9. To show that there is always a path of Hamming distance to the peak, I will show that given an arbitrary x, we can always pick a mutation k that decreases the Hamming distance to x^* by 1.

Let S be the set of indexes that x and x^* disagree on, $|S| = ||x^* \oplus x||_1$. Consider the sub-landscape on S with the other bits fixed to what x and x^* agree on. In this sublandscape x^* is a peak, thus by Proposition 8 x isn't a peak and must have some beneficial mutation $k \in S$. This is the k we were looking for. □

Note that this proof specifies an algorithm for constructing a short adaptive walk to the fitness peak x^*. However, this algorithm requires knowing x^* ahead of time – i.e. seeing the peak in the distance. But evolution does not know ahead of time where peaks are, and so cannot carry out this algorithm. Even though a short path to the peak always exists, evolutionary dynamics might not follow it.

C.1 Hard landscapes for random fitter SSWM

The simplest evolutionary rule to consider is picking a mutation uniformly at random among ones that increase fitness. This can be restated as picking and following one of the out-edges in the fitness graph at random; i.e. this is equivalent to the random-edge simplex pivot rule [10]. Proposition 11 allows me to use the hard AUSOs constructed by Matousek & Szabo [10] as a family of hard semi-smooth landscapes.

Theorem 13 ([10] in biological terminology)**.** *There exist semi-smooth fitness landscapes on $\{0,1\}^n$ such that SSWM dynamics starting from a random vertex, with probability at least $1 - e^{-\Omega(n^{1/3})}$ follows an adaptive path of at least $e^{\Omega(n^{1/3})}$ steps to evolutionary equilibrium.*

In other words, multiple peaks – or even reciprocal sign-epistasis – are not required to make a complex fitness landscape. In fact, AUSOs were developed to capture the idea of a linear function on a polytope (although AUSOs are a slightly bigger class). It is not surprising to find the simplex algorithm in the context of semi-smooth landscapes, since we can regard it as a local search algorithm for linear programming where local optimality coincides with global optimality. Linear fitness functions are usually considered to be some of the simplest landscapes by theoretical biologists; showing that adaptation is hard on these landscapes (or ones like them) is a surprising result.

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C.2 Construction of hard semi-smooth landscapes for fittest SSWM

One might object to taking random fitter mutants because sometimes the selected mutations are only marginally fitter than the wildtype. It might seem natural to speed-up evolution by always selecting the fittest possible mutant. Here I show that, in general, this does not help.

Consider a fitness landscape on $\{0,1\}^m$ with semi-smooth fitness function f that if started at 0^m will take k steps to reach its evolutionary equilibrium at x^*. I will show how to grow this into a fitness landscape on $\{0,1\}^{m+2}$ with semi-smooth fitness function f' that if started at 0^{m+2} will take $2(k+1)$ steps to reach its evolutionary equilibrium at 0^m11.

For simplicity of analysis, let us define the following functions and variables for all points in $\{0,1\}^m$ that aren't an evolutionary equilibrium under f; i.e. all except x^*. Let

$$s^+(x) = \max_{y \in N(x) \text{ s.t. } f(y)>f(x)} f(y) - f(x) \tag{1}$$

and

$$s^-(x) = \min_{y \in N(x) \text{ s.t. } f(x)+s^+(x)>f(y)>f(x)} f(y) - f(x) \tag{2}$$

where $N(x)$ are the neighbours of x in the mutation graphs; i.e. genotypes that differ from x in one bit.

Now overload these into constants, as follows: define $s^+ = \min_x s^+(x)$ and $s^- = \min_x s^-(x)$. Suppose that f is such that $s^- < s^+$; otherwise set $s^- = s^+/2$ (do this also, if $N(x)$ s.t. $f(x)+s^+(x) > f(y) > f(x)$ is empty for some non-equilibrium x).

Let $x \oplus y$ mean the XOR between x and y. Consider the 'reflected' function $f(x \oplus x^*)$. Note that if $f(x)$ is semi-smooth then so is $f(x \oplus x^*)$, since it just relabels the directions of some dimensions. The reflected function preserves all the important structure. In particular, if under $f(x)$ it took k steps to go from 0^m to x^* then under $f(x \oplus x^*)$ it will take k steps to from from x^* to 0^m.

Now define $f' : \{0,1\}^{m+2} \to \mathbb{R}$ as:

$$f'(xab) = \begin{cases} f(x) & \text{if } a=b=0 \\ f(x)+s^- & \text{if } a \neq b \text{ and } x \neq x^* \\ f(x^*)+s^- & \text{if } a=0, b=1 \text{ and } x=x^* \\ f(x^*)+s^+ & \text{if } a=1, b=0 \text{ and } x=x^* \\ f(x \oplus x^*)+f(x^*)+2s^+ & \text{if } a=b=1 \end{cases} \tag{3}$$

Basically the $x00$ subcube is the original landscape, the $x10$ and $x01$ subcubes serve as 'buffers' to make sure that the walk doesn't leave the first subcube before reaching x^*00, and the $x11$ is the original landscape reflected around x^* that takes us from x^*11 to 0^m11.

Notice, that f' has the same s^+ and s^- as f.

Now we just need to establish some properties:

Proposition 14. *Fittest mutant SSWM dynamics will not leave the $\{0,1\}^m00$ subcube until reaching x^*00.*

Proof. By definition, the fittest mutant (i.e. neighbour over $\{0,1\}^m$) from each genotype $x \in \{0,1\}^m$ that isn't x^* in f, has a fitness advantage of s^+ or higher. Hence adding two extra edges from $x00$ to $x10$ and $x01$, each with fitness advantage $s^- < s^+$ will not change the edge that fittest-mutant SSWM picks. □

Proposition 15. *SSWM dynamics will not leave the $\{0,1\}^m11$ subcube after entering it.*

Proof. This is because f' has strictly greater fitness on the $\{0,1\}^m11$ subcube than on the other three subcubes. Confirming this, note that for every $x \in \{0,1\}^m$:

$$\begin{aligned} f(x \oplus x^*)+f(x^*)+2s^+ &\geq f(x^*)+2s^+ && \text{since } f \text{ is non-negative} && (4)\\ &\geq f(x^*)+s^+ && \text{since } s^+>0 && (5)\\ &\geq f(x^*)+s^- && \text{since } s^+>s^- && (6)\\ &\geq f(x)+s^- && \text{since } x^* \text{ is fitness peak of } f && (7)\\ &\geq f(x) && \text{since } s^->0 && (8) \end{aligned}$$

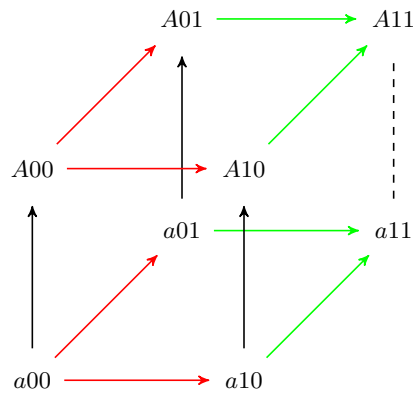
□

Proposition 16. *If f on $\{0,1\}^m$ has no reciprocal sign-epistasis then f' on $\{0,1\}^{m+2}$ has no reciprocal sign-epistasis.*

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Proof. Consider any pair of genes $i, j \in [m]$. Among these first m genes, depending the last two bits, we are looking at landscapes on $\{0,1\}^m 00$, $\{0,1\}^m 01$, $\{0,1\}^m 10$, or $\{0,1\}^m 11$, with the fitness given by $f(x)$,$f(x)+s^-$,$f(x)+s^-$, or $f(x \oplus x^*)+f(x^*)+2s^+$ (respectively). All these landscapes have isomorphic combinatorial structure to f and thus the same kinds of epistasis. Since f has no reciprocal sign-epistasis, all these subcubes lack it, too.

Now, let's look at the case of where the gene pair goes outside the first m genes. Consider an arbitrary gene $i \in [m]$, let $u \in \{0,1\}^{i-1}$, $v \in \{0,1\}^{m-i}$ be arbitrary. Label $a, A \in \{0,1\}$ such that $f(uav) < f(uAv)$. look at the subcube $u\{0,1\}v\{0,1\}^2$:



The solid black edges have their directions from the definition of a and A. The red edges have their direction because $s^+ > s^- > 0$. The green edges have their direction because of Prop. 15. The direction of the dotted black edge will depend on if $x*$ contains 0 (point up) or 1 (point down) at position i, but regardless of the direction, no reciprocal sign epistasis is introduced.

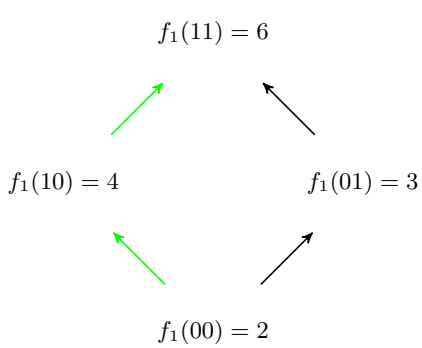
□

Corollary 17. *Given f' on $\{0,1\}^{m+2}$, the fittest mutant SSWM dynamics starting at 0^{m+2} will take $2(k+1)$ steps to reach its unique fitness peak at $0^m 11$.*

Proof. By Prop. 14, the walk will first proceed to x^*00 taking k steps. From x^*00, there are only two adaptive mutations x^*10 or x^*01, and the first is fitter. From x^*10 there is only a single adaptive mutation (to x^*11), taking us to $k+2$ steps. From x^*11, by Prop. 15, it will take us k more steps to reach $0^m 11$; totaling $2(k+1)$ steps. □

Theorem 18. *There exist semi-smooth fitness landscapes on $2n$ loci that take $2^{n+1}-2$ fittest mutant steps to reach their unique fitness peak at $0^{2(n-1)}11$ when starting from 0^{2n}.*

Proof. We will build the family of landscapes inductively using our construction, starting from an initial landscape:



The resulting path length T_n will be given by the recurrence equation: $T_{n+1} = 2T_n + 2$ with $T_1 = 2$. This recurrence is solved by $T_n = 2^{n+1} - 2$.

□

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Call the landscapes constructed as in the above proof, a *winding landscapes*. A visual example of the winding landscape construction on 6 loci ($n = 3$ in Theorem 18) is given in Figure 3. The winding landscapes construction is similar to Horn *et al.* [33]'s *Root2path* construction, except their approach introduced reciprocal sign epistasis despite having a single peak.

Of course, this is an arbitrary initial fitness landscape and any semi-smooth landscape can be used as a starting point; the walk would still scale exponentially, but there would be a different initial condition. Further, this winding product construction I showed above is just one example for building families. Many more could be considered.

In particular, if we are interested in larger mutation operators like k-point mutations instead of just 1-point mutations then it is relatively straightforward to modify the winding landscape construction. As written, equation 3 uses a buffer of 2 bits in $f'(xab)$ to transition from $f(x)$ to its reflection $f(x \oplus x^*)$. In the more general setting, we'd pad the buffer to be $k+1$ bits: define $f'(xy)$ where $|y| = k+1$ with a smooth landscape on the y portion of the input taking us from $f(x)$ to its reflection. Which leaves most of the above arguments unchanged, only modifying Theorem 18 to have the landscape to be on kn loci and the recurrence relation at the end of the proof to be $T_{n+1} = 2T_n + k + 1$.

C.3 Hard landscapes from random start

Unfortunately, one might not be impressed by a result that requires starting from a specific genotype like 0^m and ask for the expected length of the walk starting from a random vertex. Of course, if a genotype on this long walk is chosen as a starting point then the walk will still be long in most cases. However, there are only $2^{n+1} - 2$ vertexes in the walk, among 2^{2n} vertexes total, so the probability of landing on the walk is exponentially small. Instead, I will rely on direct sums of landscapes and Proposition 14 to get long expected walks.

Proposition 19. *With probability* $1/4$*, a winding landscape on* $2n$ *loci will take* 2^n *or more fittest mutant steps to reach the fitness peak from a starting genotype sampled uniformly at random.*

Proof. With probability $1/4$, the randomly sampled starting vertex has the form $x00$ (i.e. its last two bits are 0s). By prop. 14, the walk can't leave the $\{0,1\}^{2(n-1)}00$ landscape until reaching its peak at $0^{2(n-2)}1100$. This might happen quickly, or it might even already be at that peak. But after, it has to follow the two steps to $0^{2(n-2)}1111$ and then due to prop. 15 it will have to follow the normal long path, taking $2^n - 2$ more steps. □

Because of the constant probability of an exponentially long walk, we can get a big lower bound on the expected walk time:

Corollary 20. *Fittest mutant dynamics starting from a uniformly random genotype will have an expected walk length greater than* 2^{n-2} *on a* $2n$*-loci winding landscape.*

Proof. With probability $1/4$, the the walk takes 2^n or more steps, and with probability $3/4$ it takes 0 or more steps. Thus the expected walk length is greater than or equal to $(1/4) * 2^n + (3/4) * 0 = 2^{n-2}$. □

However, 75% of the time, we can't make a guarantee of long dynamics. We can overcome this limitation by taking direct sums of landscapes.

Definition 21. Given two fitness landscapes, one with fitness f_1 on $\{0,1\}^{n_1}$ and the other with fitness f_2 on $\{0,1\}^{n_2}$, the *direct sum* $(f_1 \oplus f_2)$ is a landscape with fitness f on $\{0,1\}^{n_1+n_2}$ where $f(xy) = f_1(x) + f_2(y)$.

Now, for any probability of failure $0 < \delta < 1$, let $m_\delta = \lceil \frac{\log \frac{1}{\delta}}{2 - \log 3} \rceil$ (where log is base 2).

Theorem 22. *There exist semi-smooth fitness landscapes on* $2nm_\delta$ *loci that with probability* $1 - \delta$*, take* 2^n *or more fittest mutant steps to reach their fitness peak from a starting genotype sampled uniformly at random.*

Proof. Consider a landscape that is the direct sum of m_δ separate $2n$-loci winding landscapes. Since each constituent is semi-smooth and since sums don't introduce epistasis, the resulting 'tensor sum' landscape is also semi-smooth. Further, to reach its single peak, the walk has to reach the peak of each of the m_δ independent winding sublandscapes. But as long as at least one sublandscape has a long walk, we are happy. By prop. 19, we know that for each sublandscape, we will have a short-walk starting genotype with probability less than $3/4$. The probability that none of them get a long walk then is less than $(3/4)^{m_\delta} \leq \delta$. □

D NK model with $K \geq 2$ is PLS-complete

Definition 23 ([15, 16, 34])**.** The *NK model* is a fitness landscape on $\{0,1\}^n$. The n loci are arranged in a *gene-interaction network* where each locus x_i is linked to K other loci $x_1^i, ..., x_K^i$ and has an associated fitness contribution function $f_i : \{0,1\}^{K+1} \to \mathbb{R}_+$ Given a vertex $v \in \{0,1\}^n$, we define the fitness $f(x) = \sum_{i=1}^n f_i(x_i x_1^i ... x_K^i)$.

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

By varying K we can control the amount of epistasis in the landscape. The model also provides an upper bound of $n\binom{K+1}{2}$ on the number of gene pairs that have epistatic interactions.

Weinberger [35] showed that checking if the global optimum in an NK model is greater than some input value V is NP-complete for $K \geq 3$. Although this implies that finding a global optimum is difficult, it says nothing about local optima. As such, it has generated little interest among biologists, although it spurred interest as a model in the evolutionary algorithms literature, leading to a refined proof of NP-completeness for $K \geq 2$ [36].

To understand the difficulty of finding items with some local property like being an equilibrium, Johnson, Papadimitrio & Yannakakis [22] defined the complexity class of polynomial local search (PLS). A problem is in PLS if it can be specified by three polynomial time algorithms [23]:

1. An algorithm I that accepts an instance (like a description of a fitness landscape) and outputs a first candidate to consider (the initial genotype).

2. An algorithm F that accepts an instance and a candidate and returns a objective function value (i.e. computes the fitness).

3. An algorithm M that accepts an instance and a candidate and returns an output with a strictly higher objective function value, or says that the candidate is a local maximum.

We consider a PLS problem solved if an algorithm can output a locally optimal solution for every instance. This algorithm does not necessarily have to use I, F, or M or follow adaptive paths. For instance, it can try to uncover hidden structure from the description of the landscape. A classical example would be the ellipsoid method for linear programming. The hardest problems in PLS – i.e. ones for which a polynomial time solution could be converted to a solution for any other PLS problem – are called PLS-complete. It is believed that PLS-complete problems are not solvable in polynomial time, but – much like the famous $\mathsf{P} \neq \mathsf{NP}$ question – this conjecture remains open. Note that finding local optima on fitness landscapes is an example of a PLS problem, where I is your method for choosing the initial genotype, F is the fitness function, and M computes an individual adaptive step.

Definition 24 (Weighted 2SAT)**.** Consider n variables $x = x_1...x_n \in \{0,1\}^n$ and m clauses $C_1, ..., C_m$ and associated positive integer weights $c_1, ...c_m$. Each clause C_k contains two literals (a literal is a variable x_i or its negation $\bar{x_i}$), and contributes c_k to the fitness if at least one of the literals is satisfied, and nothing if neither literal is satisfied. The total fitness $c(x)$ is the sum of the individual contributions of the m clauses. Two assignments x and x' are adjacent if there is exactly one index i such that $x_i \neq x'_i$. We want to maximize fitness.

The Weighted 2SAT problem is PLS-complete [37]. To show that the NK model is also PLS-complete, I will show how to reduce any instance of Weighted 2SAT to an instance of the NK model.

Theorem 25. *Finding a local optimum in the NK fitness landscape with $K \geq 2$ is PLS-complete.*

Proof. Consider an instance of Weighted 2SAT with variables $x_1, ..., x_n$, clauses $C_1, ..., C_m$ and positive integer costs $c_1, ..., c_m$. We will build a landscape with $m + n$ loci, with the first m labeled $b_1, ..., b_m$ and the next n labeled $x_1, ..., x_n$. Each b_k will correspond to a clause C_k that uses the variables x_i and x_j (i.e., the first literal is either x_i or $\bar{x_i}$ and the second is x_j or $\bar{x_j}$; set $i < j$ to avoid ambiguity). Define the corresponding fitness effect of the locus as:

$$f_k(0x_ix_j) = \begin{cases} c_k & \text{if } C_k \text{ is satisfied} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$f_k(1x_ix_j) = f_k(0x_ix_j) + 1 \tag{10}$$

Link the x_i arbitrarily (say to $x_{(i \mod n)+1}$ and $x_{(i+1 \mod n)+1}$, or to nothing at all) with a fitness effect of zero, regardless of the values.

In any local maximum bx, we have $b = 11..1$ and $f(x) = m + c(x)$. On the subcube with $b = 11..1$ Weighted 2SAT and this NK model have the same exact fitness graph structure, and so there is a bijection between their local maxima. □

Assuming – as most computer scientists do – that there exists some problem in PLS not solvable in polynomial time, then Theorem 25 implies that no matter what mechanistic rule evolution follows (even ones we have not discovered, yet), be it as simple as SSWM or as complicated as any polynomial time algorithm, there will be NK landscapes with $K = 2$ such that evolution will not be able to find a fitness peak efficiently. But if we focus only on rules that follow adaptive paths then we can strengthen the result:

Corollary 26. *There is a constant $c > 0$ such that, for infinitely many n, there are instances of NK models (with $K \geq 2$) on $\{0,1\}^n$ and initial genotype v such that any adaptive path from v will have to take at least 2^{cn} steps before finding a fitness peak.*

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Proof. If the initial vertex has $s = 11...1$ then there is a bijection between adaptive paths in the fitness landscape and any weight-increasing path for optimizing the weighted 2SAT problem. Thus, Schaffer & Yannakakis [37]'s Theorem 5.15 applies. □

This result holds independent of any complexity theoretic assumptions about the relationship between polynomial-time and PLS. Hence, there are some landscapes and initial genotypes, such that any rule we use for adaptation that only considers fitter single-gene mutants will take an exponential number of steps to find the local optimum.

If we turn to larger mutational neighbourhoods than singe-gene mutants then – due to the large class of possible adaptive dynamics – a variant of Corollary 26 will have to be reproved (often using a buffer padding argument similar to the end of section C.2) but Theorem 25 is unaffected:

Corollary 27. *For any definition of local equilibrium with respect to a mutation neighbourhood that contains point-mutations as a subset (i.e. if $\forall x \ \{y \mid ||y - x||_1 = 1\} \subseteq N(x)$), the NK model with $K \geq 2$ is PLS-hard.*

Proof. Any mutation operator that is a superset of point-mutations will only decrease the number of evolutionary equilibria without introducing new ones. Thus, it will only make the task of finding that equilibrium (just as, or) more difficult. However, since the algorithms studied by PLS do not have to use the mutation operator during their execution, changing it does not give them any more computational resources. □

Finally, it is important to see the NK-model as an example model, albeit a simple and natural one. If we consider more complex models of fitness landscapes – say dynamic fitness landscapes – it is often the case that there is some parameter or limit that produces the special case of a static fitness landscape like the NK-model. In particular, static landscapes are often a sub-model of dynamic fitness landscapes and thus solving dynamic fitness landscapes can only be more difficult that static ones.

D.1 Easy instances of NK-model and random fitness landscapes

Note that this doesn't mean that all instances of the NK-model are hard. In fact, there are natural sub-families of the NK-model that are easy.

The simplest easy family is $K = 0$. In that case, the genes are non-interacting and we have a smooth fitness landscapes. And all smooth landscapes are easy. For $K = 1$, Wright *et al.* [36] presented a dynamic programming approach that can find the global fitness peak in polynomial time. Since we could use this as our algorithm I to pick the initial genotype, this means the model cannot be PLS-complete for $K \leq 1$ (unless PLS = P, in which case all local search problems are easy). This means that Theorem 25 is as tight as possible in terms of K.

Alternatively, instead of restricting K, we can restrict how the gene-interaction network is connected. It will come in useful to visualize these gene-interaction networks by drawing an edge directed from a focal locus to the K loci that affect its fitness contribution. For example, if the genes can be arranged in a circle and a focal gene can interact with only the next K genes in the circle then there is a polynomial time dynamic programming algorithm to find an evolutionary equilibrium [36]. Thus, this restricted model cannot be PLS complete for any constant K.

It is an open question if SSWM dynamics – or some other reasonable evolutionary dynamics – is sufficient in the cases of $K = 1$ and circular arrangements. I conjecture that adaptive dynamics are sufficient in these cases, but proof of this is left for future work.

The NK model is frequently studied through simulation, or statistical mechanics approaches. In a typical biological treatment, the gene-interaction network is assumed to be something simple like a generalized cycle (where x_i is linked to $x_{i+1}, ...x_{i+K}$) or a random K-regular graph. The fitness contributions f_i are usually sampled from some choice of distribution. As such, we can think of biologists as doing average case analysis of these fitness landscapes. Given that randomly sampling landscapes can introduce structure like short paths [38], the structure of this simple sampling led prior research to miss the possibility of exponentially long walks. There has been a disconnect between theory and data [9, 39]. Since there is no empirical or theoretically sound justification for the choice of distributions, I avoid relying on a simple generating distribution and instead reason from only the logical description of the model. That way we know that our results are features of the logic that characterizes a particular family of fitness landscapes and not artifacts of a simple sampling distribution.

If a single sampling distribution is required despite very little or no data on the distribution of fitness landscapes in nature then it is tempting to turn to Occam's razor and consider simpler landscapes as more likely. This can be done by sampling landscapes with negative log probability proportional to their minimum description length, i.e. according to the Kolmogorov universal distribution. If landscapes are sampled in this way then all the orders of magnitude for hardness results established herein are expected to hold [40].

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

D.2 Approximate equilibria

Finally, let us also consider relaxations of equilibrium, and being "close" to a peak instead of exactly at one. For this, we need to use the whole numeric structure of the fitness function f and not just the rank-ordering that was sufficient until this point. The following definitions and proofs are based on combinatorial optimization results by Orlin *et al.* [19].

Definition 28. A genotype x is at an s-approximate equilibrium if $\forall y \in N(x)\ f(y) \leq (1+s)f(x)$.

The question becomes how big does s have to be for evolution to find an s-approximate equilibrium. But since there is no absolute units of fitness, we will need to define $f_\delta = \min_x \min_{y \in N(x) \text{ s.t. } f(y)>f(x)}(f(y) - f(x))$ and $f_{\max} = \max_x f(x)$.

First, it is important to note that all landscapes where f_δ isn't small compared to $f_{\max}$ are easy.

Proposition 29. *If $f_{max}/f_\delta \in O(n^k)$ for some constant k then an exact equilibrium can be found in a polynomial in n number of mutations by any adaptive dynamic.*

Proof. Since each adaptive step increases fitness by at least f_δ then after t adaptive steps, we have $f(x_t) \geq f_\delta t$. Combine this with $f(x_t) \leq f_{\max}$ to get that $t \leq f_{\max}/f_\delta$. □

So, we need to focus on bigger gaps between f_δ and $f_{\max}$. If the gap is exponential then we can find approximate equilibrium for moderate sized s on any landscape.

Theorem 30. *If $\log(f_{max}/f_\delta) \in O(n^k)$ then fittest mutant SSWN dynamics can find a local s-approximate equilibrium in time polynomial in n and $\frac{1}{s}$.*

Proof. Let x_0 be the initial genotype, if it is an exact equilibrium then we are done. Otherwise, let x_1 be the next adaptive step, by definition of f_δ, we have that $f(x_1) \geq f(x_0) + f_\delta \geq f_\delta$. Now, consider an adaptive path $x_1 ... x_t$ that hasn't encountered an s-approximate equilibrium; i.e. a mutation was always available such that $f(x_{i+1}) > (1+s)f(x_i)$. Thus, we have that $f(x_t) \leq f_{\max}$ and that $f(x_t) \geq (1+s)^t f_1 \geq (1+s)^t f_\delta$. Putting these two together:

$$(1+s)^t f_\delta \leq f_{\max} \tag{11}$$

$$t \ln(1+s) \leq \ln \frac{f_{\max}}{f_\delta} \tag{12}$$

$$t \leq (\ln \frac{f_{\max}}{f_\delta}) / \ln(1+s) \tag{13}$$

$$\leq (1 + 1/s) \ln \frac{f_{\max}}{f_\delta} \tag{14}$$

Where I used $\ln(1+s) \geq \frac{s}{1+s}$ in the last step. Combining with the conditions on $\log f_{\max}/f_\delta$, we get: $t \in O(\frac{n^k}{s})$. □

But for very small s, finding approximate equilibrium is as hard as finding an exact equilibrium.

Proposition 31. *If $s \leq f_\delta / f_{max}$ then any s-approximate equilibrium is a (exact) local equilibrium.*

Proof. If an s-approximate equilibrium at x is not an exact equilibrium then there exists a $y \in N(x)$ such that $f(y) - f(x) \geq f_\delta$ but $f(y) < (1+s)f(x)$. Combining this with $f(x) \leq f_{\max}$, we get that $s > f_\delta / f_{\max}$. □

Thus, it isn't possible to find an s-approximate equilibrium for very small s on hard fitness landscapes:

Corollary 32. *If PLS $\neq$ P and $\log(f_{max}/f_\delta) \in O(n^k)$ then (for NK-model with $K \geq 2$) a local s-approximate equilibrium cannot be found in time polynomial in n and $\log \frac{1}{s}$.*

Proof. If such an algorithm existed then we'd run it with $s = f_\delta / f_{\max}$ and – by Proposition 31 – the approximate equilibrium it finds would be exact. Further, in this case $\log \frac{1}{s} = \log(f_{\max}/f_\delta) \in O(n^k)$ and thus the runtime would be polynomial in n. This is not possible for the NK-model with $K \geq 2$ by Theorem 25 (unless PLS = P). □

This also means that the selective coefficient of the fittest mutant $s(t) = \max_{y \in N(x_t) \cup \{x_t\}}(f(y) - f(x_t))/f(x_t)$ cannot decay exponentially quickly.

Corollary 33. *If PLS $\neq$ P then there are no evolutionary dynamics such that $s(t) \leq e^{-mt}$ for all instances of the NK-model with $K \geq 2$.*

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Main Text & Appendix References

1. Wright, S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress of Genetics,* 356–366 (1932).
2. Obolski, U., Ram, Y. & Hadany, L. Key Issues Review: Evolution on rugged adaptive landscapes. *bioRxiv,* 112177 (2017).
3. Poelwijk, F., Sorin, T.-N., Kiviet, D. & Tans, S. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *Journal of Theoretical Biology* **272,** 141 –144 (2011).
4. Crona, K., Greene, D. & Barlow, M. The peaks and geometry of fitness landscapes. *Journal of Theoretical Biology* **317,** 1–10 (2013).
5. Chou, H. H., Chiu, H. C., Delaney, N. F., Segr?, D. & Marx, C. J. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* **6034,** 1190–1192 (332 2011).
6. Weinreich, D., Watson, R. & Chan, L. Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59,** 1165–1174 (2005).
7. Szabó, T. & Welzl, E. *Unique sink orientations of cubes* in *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on* (2001), 547–555.
8. Gillespie, J. A simple stochastic gene substitution model. *Theor. Pop. Biol.* **23,** 202 (1983).
9. Orr, H. A. The genetic theory of adaptation: a brief history. *Nature Reviews. Genetics* **6,** 119–127 (2005).
10. Matousek, J. & Szabo, T. RANDOM EDGE can be exponential on abstract cubes. *Advances in Mathematics* **204,** 262–277 (1 2006).
11. Puchta, O. *et al.* Network of epistatic interactions within a yeast snoRNA. *Science* **352,** 840–844 (2016).
12. Li, C., Qian, W., Maclean, C. J. & Zhang, J. The fitness landscape of a tRNA gene. *Science* **352,** 837–840 (2016).
13. Lozovsky, E. R. *et al.* Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proc. Natl. Acad. Sci. USA* **106,** 12025–12030 (29 2009).
14. Szendro, I., Schenk, M., Franke, J., Krug, J. & de Visser, J. Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech.* p01005 (2013).
15. Kauffman, S. & Levin, S. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology* **128,** 11–45 (1 1987).
16. Kauffman, S. & Weinberger, E. The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology* **141,** 211–245 (2 1989).
17. Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-term dynamics of adaptation in asexual populations. *Science* **342,** 1364–1367 (2013).
18. Couce, A. & Tenaillon, O. A. The rule of declining adaptability in microbial evolution experiments. *Frontiers in genetics* **6,** 99 (2015).
19. Orlin, J., Punnen, A. & Schulz, A. Approximate local search in combinatorial optimization. *SIAM J. Comput.* **33,** 1201–1214 (5 2004).
20. Gillespie, J. H. *Population genetics: a concise guide* (John Hopkins University Press, 2010).
21. Lenski, R. E. *et al.* Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with Escherichia coli. *Proc. R. Soc. B* **282,** 20152292 (2015).
22. Johnson, D., Papadimitriou, C. & Yannakakis, M. How easy is local search? *Journal of Computer and System Sciences* **37,** 79–100 (1 1988).
23. Roughgarden, T. Computing equilibria: A computational complexity perspective. *Economic Theory* **42,** 193–236 (1 2010).
24. Baldwin, J. M. A new factor in evolution. *The american naturalist* **30,** 441–451 (1896).
25. Simpson, G. G. The Baldwin effect. *Evolution* **7,** 110–117 (1953).
26. Orzack, S. H. & Sober, E. *Adaptationism and optimality* (Cambridge University Press, 2001).
27. Chapdelaine, P. & Creignou, N. The complexity of Boolean constraint satisfaction local search problems. *Annals of Mathematics and Artificial Intelligence* **43,** 51–63 (2005).
28. de Visser, J., Park, S. & Krug, J. Exploring the effect of sex on empirical fitness landscapes. *The American Naturalist* (2009).

bioRxiv preprint doi: https://doi.org/10.1101/187682; this version posted September 12, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

29. Franke, J., Klozer, A., de Visser, J. & Krug, J. Evolutionary accessibility of mutation pathways. *PLoS Comp. Biol.* **7,** e1002134 (8 2011).

30. Goulart, C. *et al.* Designing antibiotic cycling strategies by determining and understanding local adaptive landscapes. *PLoS One* **8,** e56040 (2 2013).

31. De Visser, J. A. G. & Lenski, R. E. Long-term experimental evolution in Escherichia coli. XI. Rejection of non-transitive interactions as cause of declining rate of adaptation. *BMC Evolutionary Biology* **2,** 19 (2002).

32. Poelwijk, F., Kiviet, D., Weinreich, D. & Tans, S. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445,** 383–386 (2007).

33. Horn, J., Goldberg, D. E. & Deb, K. *Long path problems* in *International Conference on Parallel Problem Solving from Nature* (1994), 149–158.

34. Kauffman, S. *The origins of order: Self organization and selection in evolution.* (Oxford University Press, 1993).

35. Weinberger, E. NP completeness of Kauffman's N-k model, a tunably rugged fitness landscape. *Santa Fe Institute working paper,* 1996–02–003 (1996).

36. Wright, A., Thompson, R. & Zhang, J. The computational complexity of N-K fitness functions. *Evolutionary Computation, IEEE Transactions on* **4,** 373–379 (4 2000).

37. Schaffer, A. & Yannakakis, M. Simple local search problems that are hard to solve. *SIAM Journal on Computing* **20,** 56–87 (1991).

38. Tovey, C. A. Low order polynomial bounds on the expected performance of local improvement algorithms. *Mathematical Programming* **35,** 193–224 (1986).

39. Kryazhimskiy, S., Tkacik, G. & Plotkin, J. The dynamics of adaptation on correlated fitness landscapes. *Proc. Natl. Acad. Sci. USA* **106,** 18638–18643 (44 2009).

40. Li, M. & Vitányi, P. M. Average case complexity under the universal distribution equals worst-case complexity. *Information Processing Letters* **42,** 145–149 (1992).