OCCURRENCES OF RECIPROCAL SIGN EPISTASIS IN SINGLE- AND MULTI-PEAKED THEORETICAL FITNESS LANDSCAPES

A PREPRINT

Manda Riehl
Department of Mathematics
Rose-Hulman Institute of Technology
Terre Haute, IN
riehl@rose-hulman.edu

Reed Phillips
Department of Mathematics
Rose-Hulman Institute of Technology
Terre Haute, IN

Lara Pudwell
Department of Mathematics and Statistics
Valparaiso University
Valparaiso, IN

Nate Chenette
Department of Computer Science and Software Engineering
Rose-Hulman Institute of Technology
Terre Haute, IN

October 11, 2022

ABSTRACT

Fitness landscapes help model the theory of adaption. We consider genetic fitness landscapes abstractly as acyclic orientations of Boolean lattices under the assumptions laid out by Crona et al. We focus on occurrences of reciprocal sign epistasis (RSE) on the faces of the lattice. We computationally study which combinations of numbers of peaks and RSE faces are possible, and we determine limits on occurrences of RSE faces in both single-peaked and multi-peaked landscapes. Our main theorem extends a theorem of Poelwijk to show that any landscape with k peaks contains at least $k-1$ RSE faces.

***K*eywords** Fitness Landscapes · Acyclic orientations · Epistasis

MSC: 92D15, 05C30

1 Introduction

Fitness landscapes represent the reproductive and survival success of a collection of related genotypes, and have been a central concept in evolutionary biology for nearly 100 years [34]. Many approaches have been used to measure fitness [19], including laboratory experimental results as well as genomic data. Fitness landscapes enable a direct link between evolution, molecular biology, and systems biology [9]. The utility of investigating fitness landscapes has been shown with, for example, the development of antibiotic cycling regimens to combat antibiotic resistance [12]. Recent research has shown that in fact parallel populations of bacteria acquiring antibiotic resistance evolved similar mutations and acquired them in a similar order, further reinforcing the importance of fitness landscapes to determine evolutionary pathways [29]. Analysis of these pathways even has the potential to predict currently unknown evolutionary trajectories [30] including in antibiotic resistance [17] and the repeatability of evolution in general [1]. There are several reviews available of the field [16] [30].

1.1 Definitions

This paper will consider genetic fitness landscapes abstractly as Boolean lattices under the assumptions laid out by Crona et al. in their 2013 paper [3]. Crona and their collaborators have a long history of work in the area including [4] [5] [6] [7]. Restricting ourselves to haploid genomes, we will call the physical site of each gene under investigation a *locus*. Each locus in a genotype is assumed to have two possible configurations: the wild type, denoted with (0), and a mutated type denoted by (1). In other words, we are assuming each locus is biallelic. The portions of a genotype under consideration can then be expressed as a bitstring describing which of the two options for each locus is present. Each genotype has a particular fitness value associated with it, where larger fitness values are better-adapted. The wild type, represented by the string $000\cdots0$, is assumed to have the lowest fitness unless otherwise stated.

We will be using the Strong Selection Weak Mutation (SSWM) evolutionary regime [11], where the time between occurrences of new mutations is assumed to be much longer than the time it takes for a mutated genotype to take over the population. The latter is sometimes referred to as the time required for the mutation to *fix* in the population. As such, mutational paths do not exist directly between genotypes with differences at two or more loci; rather the population is assumed to add or remove one mutation at a time. Consequently a population is assumed to only be able to travel from one genotype to an adjacent genotype which differs in only one locus, which we will call a *neighbor*. Thus we will think of the Boolean lattice Q_n as the graph whose vertex set is labeled with binary words of length n and where two vertices are adjacent if and only if their labels differ in exactly one bit.

While some [3] have worked with these lattices using the fitness values, we use the values to assign each edge a direction (adjacent genotypes are assumed to have different fitness values) and consider the directed Boolean lattice (hypercube). In an experimental setting, not all of these combinations would exist as viable genomes, however we can consider those as occurring in the landscape but having fitness so low that they would never fix in a population (but nonzero, so that edges can still be oriented). Note that these orientations of the Boolean lattice must be acyclic, or a genotype would have higher fitness than itself. Accordingly we define a *peak* to be a vertex which has no edges directed away from itself, and a *valley* to be a vertex with all of its adjacent edges directed away from itself. (In other fields those would be called sinks and sources respectively.)

Sign epistasis at a locus means that a mutation at that locus may be beneficial or deleterious, depending on the state (mutated or wild) of other loci. In other words, the sign of the fitness effect of a mutation is under epistatic control [31]. Naively, one might assume that nearly all mutations that are beneficial under some genotype backgrounds are beneficial under all backgrounds, but in fact sign epistasis has been demonstrated in many genetic systems, including *E.coli* [32] [21], as well as with viruses including HIV-1 [26] [36] [24]. Sign epistasis has also been shown to limit the number of available evolutionary pathways [31] [15].

In traditional recombination studies, typically only two mutation sites and pairwise gene interactions are considered. However higher-dimensional epistasis (in other words across more than two dimensions) has been shown to occur frequently in fitness landscapes [33]. Complex interactions occur between mutations, and whether a mutation is beneficial or not depends dramatically on whether other mutations have occurred. One notable experimental example on seven loci appears in [20]. However in this work we will only consider epistatic effects pairwise and any subsequent use of the term “epistasis" refers to the interaction between exactly two loci. However we are not assuming that higher-dimensional epistasis does not exist, since we make no assumption that edges on parallel faces of our lattice point in the same direction. We address this issue somewhat further in Section 6.

Reciprocal Sign Epistasis is an extreme form of sign epistasis wherein multiple mutations that are advantageous when they occur individually are unfavorable when they occur jointly (or vice versa). Again, as with sign epistasis, from this point forward we will only be investigating reciprocal sign epistasis (RSE) between two loci though we do not preclude its existence among more loci in the larger landscape. Lastly, from this point forward, when we use the term “sign epistasis", we will mean “non-reciprocal sign epistasis", since our main interest lies in those faces exhibiting RSE.

1.2 Reciprocal Sign Epistasis in the Boolean Lattice

We can use the directed edges to examine whether sign or reciprocal sign epistasis exist on a face of the lattice, without considering the strength of that epistasis. Figure 1 shows the three different categories of faces, with both directed edges and vertices labelled with hypothetical fitness values. If we imagine that the lowest vertex in each diagram is wild type 00, and the topmost vertex is 11, then we see that no sign epistasis corresponds to two beneficial mutations resulting in more benefit than either single mutation. As an example of sign epistasis of the non-reciprocal type, the center diagram is when both mutations are beneficial, but exactly one of the single mutations is more beneficial than having both mutations occur. Reciprocal sign epistasis (RSE), the rightmost diagram, is as explained above when both single mutants are more fit than the double mutant. Note however, that these examples assume that the single mutants are both beneficial. We can also have no sign epistasis with deleterious (negative impact on fitness) mutations, or a combination of beneficial and deleterious mutations. In these cases, it is often easier to just consider the maximal paths in the faces in order to determine whether there is sign or reciprocal sign epistasis. No sign epistasis faces have two paths of length 2. Faces with sign epistasis but no reciprocal sign epistasis have one path of length 1 and one path of length 3. Faces with reciprocal sign epistasis have four paths of length 1.

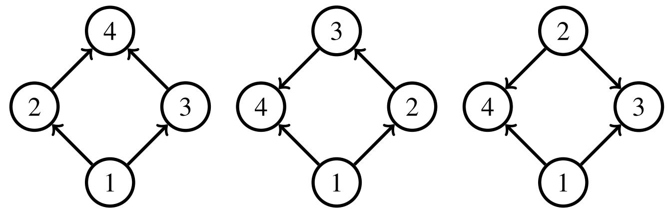

Figure 1: From left to right: no sign epistasis, sign epistasis, reciprocal sign epistasis; vertices are labeled with fitness values.

A fitness landscape with many local peaks surrounded by deep valleys is called *rugged*. Experimental fitness landscapes have been shown to be rugged [12], and multiple authors have proved that reciprocal sign epistasis (RSE) is required in order to have a rugged landscape [3, 22]. However this relationship is not bidirectional: one can have reciprocal sign epistasis without the landscape containing multiple peaks. We seek to highlight this unidirectionality by describing the relationship between occurrences of RSE faces and multiple peaks. How many RSE faces can a landscape have and still have only one peak? Are there bounds on how many RSE faces a landscape with k peaks will have?

Our primary focus is on which combinations of peak counts and RSE face counts are possible. For example, the 3-lattice in Figure 2 has one RSE face and one peak. Many of our arguments rely on building larger lattices from smaller lattices with known numbers of RSE faces and peaks, and we will next discuss those building methods. Throughout this paper, while Q_n was originally introduced as the (undirected) Boolean lattice, we are interested in various orientations of Q_n where directed edges point to vertices with higher fitness. For the remainder of the paper, Q_n is understood to be an acyclic orientation of the n-dimensional Boolean lattice. It will also be useful later for us to define the *alternating landscape* A_n as the acyclically oriented lattice of dimension n where every face is an RSE face and $0 \cdots 0$ is a valley. It has been long known that the alternating landscape has the maximal number of peaks [14]. In some places, this is referred to as the "Egg Box" landscape [8], for the familiar shape one would see if the peaks and valleys respectively had their identical fitness value heights plotted in a third dimension.

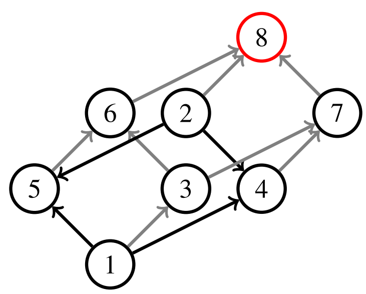

Figure 2: A 3-cube with one peak (red) and one RSE face (edges highlighted in black); vertices are labeled with fitness values.

2 Method: Lattice Gluing

Looking at fitness landscapes purely as Boolean lattices suggests several ways to build larger lattices from smaller ones. When we consider the Boolean lattice in n dimensions, half of the vertices are labeled with bitstrings that start with 0, while the other half are labeled with bitstrings that start with 1. Each of these halves of the n-dimensional oriented Boolean lattice can be viewed as an $(n-1)$-dimensional oriented Boolean lattice. From this point of view, the n-dimensional oriented lattice consists of (a) the two smaller $(n-1)$-dimensional oriented lattices, and (b) additional directed edges that connect pairs of vertices (one from each smaller lattice) whose length $n-1$ bitstrings agree. This allows for ways to glue two smaller lattices (Q_{n-1}) together to obtain a larger one (Q_n) with some desired properties. The only thing that has to be chosen, besides which smaller lattice orientations to use, is how to orient the edges connecting them. The only restriction is that the gluing method cannot introduce a cycle, where following a chain of edges to supposedly higher fitness values leads back to the original vertex. One way to check for cycles is to see if there is still a valid assignment of fitness values that would produce the lattice as a whole. In this section we define two specific gluings that can be used to construct an acyclic Q_n lattice from *any* two acyclic Q_{n-1} lattices.

The easiest way to guarantee valid fitness values in the higher-degree acyclic lattice is to simply orient every edge connecting the first acyclic lattice to the second in the same direction. One simple way this could be achieved via fitness values is by adding the highest fitness value from the first oriented lattice to all fitness values in the second acyclic lattice, as seen in Figure 3. We will call this a *basic gluing*.

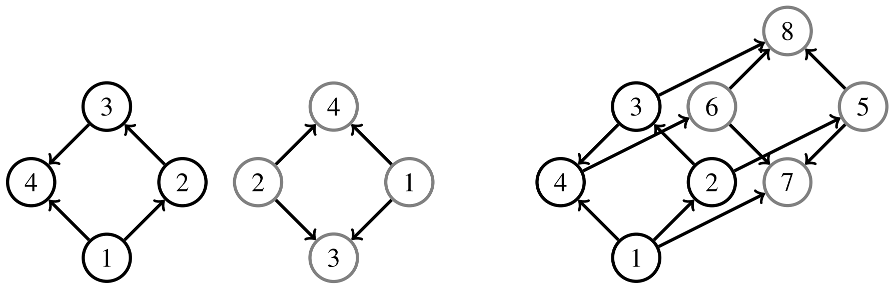

Figure 3: A basic gluing; vertices are labeled with fitness values. The first lattice (with black vertices) retains the same fitness values. The second lattice (with gray vertices) has the same orientation among its vertices, but its fitness values are all increased. Every edge between the two lattices is oriented from black to gray.

Another way to connect two lattices is with a *peak-preserving gluing*. Rather than having every edge between lattices directed towards the second lattice, the edges between lattices involving peaks in the first lattice are reversed. This allows for each peak in the first lattice to continue to be a peak in the new larger lattice, and also preserves some (and perhaps all) of the peaks in the second lattice. Further this orientation is acyclic as long as both smaller lattice orientations are acyclic. Any edge pointing from the second lattice to the first points to a peak, which cannot be part of a cycle by definition. All other edges between lattices point from the first to the second lattice, and since they are oriented in the same direction, they also cannot participate in cycles. To associate valid fitness values with these new oriented edges, increase the fitness in each vertex in the second lattice and a make a larger increase to the peaks in the first lattice. An example of a peak-preserving gluing is seen in Figure 4.

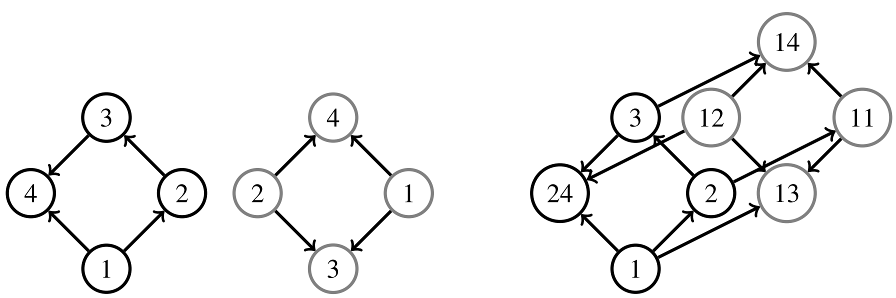

Figure 4: A peak-preserving gluing; vertices are labeled with fitness values. The first lattice (with black vertices) and the second lattice (with gray vertices) retain the same orientation among their vertices in the glued version; edges between lattices point from gray to black if the vertex was a peak in the first lattice and from black to gray otherwise.

It is conceivable that any desired combination of number of peaks and number of RSE faces in Q_n could be made by gluing together two copies of Q_{n-1} using simple methods like these. However, this does not appear to be the case, as can be seen in Tables 7 and 8 in the Appendix. These tables show which gluing methods, if any, could produce different combinations of peaks and RSE faces in the same lattice. The data was gathered by producing a large number of random lattices by various methods and seeing which methods produced which combinations. In particular, various combinations of few peaks and many RSE faces were found to be possible in general but not achieved by either of the above gluing methods. It is possible that such combinations are merely extremely unlikely and could be achieved with gluing; theoretical work and more focused experimental work has shown that more combinations are possible than are listed in Table 8. But as a starting point, the data shows the vast majority of possibilities can be achieved by gluing smaller lattices together.

This paper will discuss methods to determine the achievable combinations of peaks and RSE faces and find some limits on which are possible. In Section 3, we examine RSE face counts in single-peaked landscapes. In Section 4, we examine RSE face counts in multi-peaked landscapes.

3 RSE faces in single peak landscapes

By construction, Q_n has 2^n vertices and $n \cdot 2^{n-1}$ edges (OEIS 2021) [18] (A001787). This means that there are $2^{n \cdot 2^{n-1}}$ total orientations (OEIS 2021) [18] (A061301) of Q_n. Even when we restrict to acyclic orientations, Stanley showed there are $\chi(Q_n)(-1)$ acyclic orientations of Q_n, where $\chi(Q_n)(x)$ is the chromatic polynomial of Q_n [28]. While $\chi(Q_n)(x)$ is complicated to write down in a general form, there are recursive algorithms to compute it for any given dimension [2]. From the chromatic polynomial, there are over 193 million acyclic orientations of Q_4 and over $4.7 \times 10^1 9$ acyclic orientations of Q_5. Even though there are algorithms to generate all acyclic orientations of a graph [27], the number of acyclic hypercube orientations is so large for dimension $n \geq 4$ that random sampling is not an effective way to detect extreme behavior of combinatorial statistics on these orientations. Consequently we used several different approaches to computations, which we detail in Section 5. These computations provided intuition for many of the theoretical results of the current section.

Poelwijk et al. in 2011 [22] showed that having at least one RSE face is a necessary, but not sufficient, condition for having multiple peaks. In this section we consider bounds on the insufficiency of that condition. How many or few RSE faces can a landscape have while still having only one peak?

The minimum number of RSE faces is quickly addressed by example. By having every edge in a lattice directed upwards, towards the vertex whose bitstring has more 1s, one can obtain a single peak and zero RSE faces in any dimension.

The maximum number of RSE faces in a single peak landscape is much more interesting. While we have not found an exact formula for this maximum, we have bounds on it and a conjecture for what the true value is, as a function of the dimension n. First we present a lower bound on the maximum number of RSE faces.

Theorem 1. *For $n \geq 3$, single-peaked n-dimensional lattices exist with r_n RSE faces, where*

$$r_n = 2^{n-3}(n^2 - 5n + 8) - 1. \tag{1}$$

Notably, the most significant term in this expression is $2^{n-3}n^2$. The number of faces in a n-dimensional lattice is given by $\binom{n}{2}2^{n-2}$ since any face has vertices whose bitstring labels agree in $n-2$ positions and vary in the other two positions. There are $\binom{n}{2}$ ways to choose the positions of the two bits that will vary, and then 2 choices for the value of each of the remaining bits that remain constant. This quantity can be written as $2^{n-3}(n^2 - n)$, which has the same most significant term as the expression in Theorem 1. This means that, in high enough dimensions, an arbitrarily large proportion of the faces in a lattice can be RSE faces while still having only one peak.

Proof. Consider the following sequence L_i of single-peaked binary lattices. As a base case, L_2 is a single face with each edge pointed towards the vertex with more 1s, and with a unique peak at 11. Then L_{n+1} can be constructed by taking a basic gluing of lattices A_n (the n-dimensional alternating lattice defined above) and L_n, i.e. by drawing an edge from each vertex in A_n to the vertex with the same label in L_n. The resulting lattice has exactly one peak because no vertex in the A_n sublattice can be a peak and the L_n sublattice has exactly one peak.

Because every edge connecting the two halves of L_{n+1} is pointed in the same direction, there can be no RSE faces that use those edges. Therefore, the number of RSE faces in L_{n+1} is simply the sum of the numbers of RSE faces in A_n and L_n. Let r_n be the number of RSE faces in L_n. The number of RSE faces in A_n is the number of faces in an n-dimensional lattice, which is $2^{n-2}\binom{n}{2} = 2^{n-3}(n^2 - n)$. This gives a recurrence relation for r_n as

$$r_{n+1} = r_n + 2^{n-3}(n^2 - n), \tag{2}$$

which can be solved to give

$$r_n = 2^{n-3}(n^2 - 5n + 8) + C, \tag{3}$$

where C is a constant. Using the initial condition $r_2 = 0$ gives $C = -1$, which yields the desired expression (1). □

Corollary 1. *As $n \to \infty$, the proportion of 2-dimensional faces that can be RSE faces in a single-peak n-dimensional landscape approaches 1.*

This theorem and corollary show that the converse of the statement "If a landscape has multiple peaks, then it contains at least one RSE face" is not only false, but remarkably false, since with large enough dimension, we can have nearly all faces as RSE faces while still containing only a single peak.

Next, we determine an upper bound on the number of possible RSE faces in a single-peaked landscape.

Theorem 2. *A single-peaked n-dimensional lattice cannot have more than*

$$2^{n-3}(n^2 - n - 2\lfloor n/2 \rfloor) \tag{4}$$

RSE faces.

The sketch of the proof is as follows: we consider a single vertex v and the faces containing v. We show that it is impossible to have every face containing v be an RSE face and still have a single peak. Knowing this, we consider how many faces containing v can be RSE faces while still preserving a single peak. The upper bound is obtained by multiplying this number of maximum RSE faces containing v by the total number of vertices in the lattice.

Proof of Theorem 2. We first consider how many RSE faces can surround a single vertex v by considering the faces which contain v. Every edge touching v and every vertex one edge away from v is included in one of these faces. If all the faces containing v are RSE faces, the entire set can be oriented in two different ways: designate v as a valley with all edges pointing toward v's neighbors, which are all peaks, or designate v as a peak, where all of v's neighbors are valleys. We will show that each of these situations contradicts having a single-peaked lattice.

- If v is a peak and all edges involving v are directed toward v. If all faces involving v are RSE faces, the lattice is split into two parts: v and the rest of the lattice. Since v's neighbors are all valleys, if we follow an edge from one of those neighbors going away from v, we must eventually reach a peak other than v, and the landscape has at least two peaks.
- If v is a valley, then its neighbors are all peaks, and so it is impossible to have a single-peaked landscape in this case.

This prevents any vertex in a single-peaked lattice from being completely surrounded by RSE faces. To do better, each of these conditions can be relaxed a little.

For the first case, where v is a peak but not all faces containing v are RSE faces, one problem that arises comes from the opposite corners from v on the RSE faces. Call one of these vertices u. To prevent u from being a peak, u must have a lower fitness value than v and there must exist a chain of edges from u to v. Call the last vertex before v in this chain w. If the $v - w$ edge is part of an RSE face, then let u be the corner opposite v of that face instead and find a new w. Because the new u always has a higher fitness value than the previous one and there are finitely many choices for u, this process must end eventually. That means there is a vertex w, adjacent to v, which is not part of any RSE face. That means, when v is a peak, at least $n - 1$ faces (those containing v and w) must not be RSE faces to allow for a single peak, so the maximum number of RSE faces involving v is $\binom{n}{2} - (n-1)$.

For the second case, where v is a valley, all but one of v's neighbors must be prevented from being peaks. Each face containing v also contains two of v's neighbors, so at least $\lfloor n/2 \rfloor$ faces must be not be RSE faces in order to remove enough peaks. That makes the maximum number of RSE faces containing v $\binom{n}{2} - \lfloor n/2 \rfloor$. This is more faces than the previous case, so it is the upper bound. We will count every vertex in every RSE face, which will result in overcounting by a factor of 4, so we will first correct for that. Letting every vertex be contained in exactly this many RSE faces gives

$$\frac{1}{4}2^n\left(\binom{n}{2} - \lfloor n/2 \rfloor\right)2^{n-2}\left(\binom{n}{2} - \lfloor n/2 \rfloor\right) = 2^{n-2}\left(\frac{n^2 - n - 2\lfloor n/2 \rfloor}{2}\right) = 2^{n-3}(n^2 - n - 2\lfloor n/2 \rfloor) \tag{5}$$

as an upper bound on the maximum number of RSE faces in a single-peaked landscape. □

In addition to these results, we have a conjecture for the maximum number of RSE faces that matches the dominant term of our bounds and matches the existing computer-generated data, including more careful searches of single-peaked lattices.

Conjecture 1. *The maximum number of RSE faces in a single-peaked n-dimensional lattice is $2^{n-3}(n^2 - 4n + 4)$.*

n	2	3	4	5	6	7	8
Lower bound	0	1	7	31	111	351	1023
Conjectured maximum	0	1	8	36	128	400	1152
Upper bound	0	4	16	64	192	576	1536

Table 1: The values provided by Theorems 1 and 2 and Conjecture 1 for small n.

4 RSE faces in Multipeaked Landscapes

4.1 Lower bounds for RSE faces and k peaks

Another extension of Poelwijk et al.'s result that an RSE face is a necessary condition for multiple peaks is whether similar conditions can be found for more than two peaks, three peaks, etc. The smallest step along this path is the following theorem.

Theorem 3. *At least two RSE faces are required for a lattice to have three peaks.*

Proof. Consider a three-peaked lattice, labeled so that one of the peaks is at $0 \cdots 0$ for convenience. Define the connecting sublattice of two vertices as the sublattice with those vertices as its opposite corners. More formally, this is the set of points where, for each bit position:

- If the two vertices have the same value for that position, so does the point.
- If the two vertices have different values for that position, the point can have either bit value.

This is itself a binary lattice. Therefore, the connecting sublattice of two peaks must contain an RSE face. Let v be the vertex at $0 \cdots 0$ and let the other two peaks be u and w, which have coordinates $u_1 u_2 \cdots u_n$ and $w_1 w_2 \cdots w_n$. The connecting sublattice of u and v is the set of vertices that have a 0 everywhere u has a 0. The connecting sublattice of w and v is the set of vertices that have a 0 everywhere w has a 0. Assume for contradiction that there are not two RSE faces. Then the only RSE face must be in both connecting sublattices. The intersection of these is the set of vertices which have a 0 where either u or w has a 0. In other words, these are vertices whose only 1s are in the coordinates where u_i and w_i are both 1. This means exactly one vertex is in both this intersection and the connecting sublattice of u and w, since the vertices in that connecting sublattice must have 1 everywhere u and w both do. But in order for an RSE face to be in all three sublattices, we need at least four vertices. Therefore we may conclude that there must be more than one RSE face. □

We have now established the beginning of a pattern to address the question: what is the fewest number of RSE faces necessary for a landscape to have k peaks? The $k = 1$ case is trivial, (0 RSE faces are necessary), Poelwijk et al.'s results address the $k = 2$ case (1 RSE face is necessary), and Theorem 3 addresses the $k = 3$ case (2 RSE faces are necessary). Before proving the minimum number of RSE faces for general number of k peaks, we give an existence result.

Theorem 4. *Lattices with k peaks and $k - 1$ RSE faces exist for any $k \geq 1$.*

To prepare for the proof of Theorem 4, define an *almost peak* to be a vertex where all but one of its edges points towards it. For example, the 3-dimensional lattice in Figure 5 has three almost peaks.

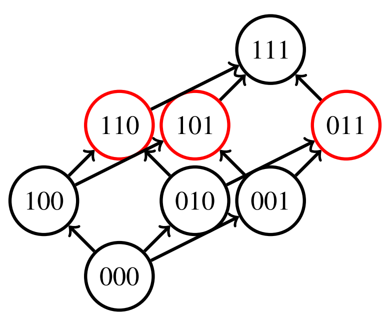

Figure 5: A lattice with three almost peaks (red)

Almost peaks are useful because they allow for peak-preserving gluings that add only one RSE face. For example, consider the gluing of part of two 4-dimensional lattices in Figure 6. Each of the faces involving the edge uv has the potential to be an RSE face because each of the edges $u_i v_i$ points in the opposite direction to uv and each u_i points towards u. To allow only a single RSE face, v must be an almost peak because any edge directed away from v will result in an RSE face. For instance, the vv_3u_3u face here is an RSE face while the others are not. Note that v cannot be a peak if the resulting lattice is to have more peaks than the separate halves since it would be below u in fitness value.

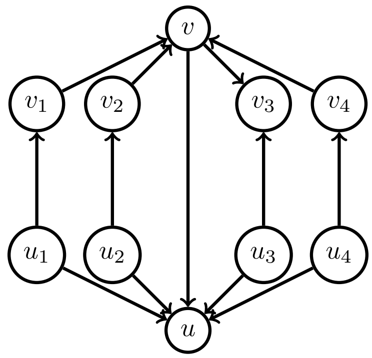

Figure 6: Part of a gluing of two four-dimensional lattices, where u is a peak in the first lattice and v is an almost peak in the second lattice.

Proof of Theorem 4. We can construct a lattice with one peak at $1 \cdots 1$ and no RSE faces in any dimension by having all edges point to the vertex with more 1s in its coordinates (as in Figure 5). Notice that this orientation of an n-dimensional lattice has n almost peaks. We will call this lattice U_n

To construct a lattice with k peaks and $k-1$ RSE faces, we will construct a sequence of lattices $L_0, L_1, \dots$ where lattice L_i has i RSE faces. Let L_0 be the $(k-1)$-dimensional lattice $L_0 = U_{k-1}$, which has one peak, zero RSE faces, and $k-1$ almost peaks.

Now, given lattice L_i, which has dimension $i+(k-1)$, construct a lattice L_i^* of the same dimension by taking the orientation of $U_{i+(k-1)}$ and relabeling the vertices so that its single peak has the same label as an almost peak in L_i. Glue L_i^* and L_i together with a peak-preserving gluing that preserves the single peak of L_i^*. As in Section 2, the resulting lattice L_{i+1}, of dimension $(i+1)+(k-1)$, is acyclic. Further, L_{i+1} has one more RSE face than L_i using the peak from L_i^* and the corresponding almost peak from L_i, as in the construction in Figure 5. The lattice L_{i+1} also has one more peak than L_i because all peaks of L_i are preserved, while the peak of L_i^* is an additional peak. L_{i+1} also has one fewer almost peak because the almost peak of L_i that shares an edge with L_i^*'s peak now has two edges directed away from it; however all other almost peaks in L_i remain almost peaks.

Following this procedure $k-1$ times results in $L_{2(k-1)}$, a lattice of dimension $2(k-1)$ with k peaks, $k-1$ RSE faces, and 0 almost peaks.

□

We now are ready for the general lower bound on the number of RSE faces in a k-peak lattice.

Theorem 5. *In any dimension, a lattice with k peaks contains at least $k-1$ RSE faces.*

Proof. We proceed by induction on k. For the base case, it is already known that a lattice with 2 peaks contains at least 1 RSE face.

Now, suppose that there are k peaks in some acyclic Boolean lattice Q_n where $k > 2$. Pick two of those peaks. These two peaks induce a Boolean sublattice Q^*, where the two peaks are as far apart within Q^* as possible. Since Q^* has 2 peaks, Q^* must have an RSE face. Suppose one of the edges of the RSE face goes from vertex v to vertex u where – if they were labeled with binary strings – they differ in bit i. Now decompose the original Boolean lattice Q_n (of dimension n) into two $(n-1)$-dimensional Boolean sublattices where the first sublattice $Q_{n-1}^{(0)}$ consists of all points with a 0 in bit i and the other sublattice $Q_{n-1}^{(1)}$ consists of all points with a 1 in bit i. Notice that the RSE face we found in Q^* is neither in $Q_{n-1}^{(0)}$ nor $Q_{n-1}^{(1)}$ since it has an endpoint in each of these two sublattices.

Together $Q_{n-1}^{(0)}$ and $Q_{n-1}^{(1)}$ must have k peaks. In fact, each of them must have at least one peak. Consider the two peaks we used to define the sublattice Q^*. Since the edge from u to v is part of this sublattice defined by two peaks it must be the case that one of these peaks has a 0 in bit i and is in lattice $Q_{n-1}^{(0)}$ while the other of these peaks has a 1 in bit i and is in lattice $Q_{n-1}^{(1)}$. Therefore, $Q_{n-1}^{(0)}$ has j peaks where $j > 0$ and $Q_{n-1}^{(1)}$ has $k-j$ peaks where $k-j > 0$. Since $Q_{n-1}^{(0)}$ has j peaks, it has at least $j-1$ RSE faces. Since $Q_{n-1}^{(1)}$ has $k-j$ peaks, it has at least $k-j-1$ RSE faces, and together with the RSE face that uses an edge between $Q_{n-1}^{(0)}$ and $Q_{n-1}^{(1)}$, we have at least $(j-1)+(k-j-1)+1 = k-1$ RSE faces in Q_n. $\square$

We note that the theorem above was proved independently using a different method (discrete Morse theory) in a paper which was published while this work was under peer review (Saona, 2022). In fact both proofs use induction, and rely on the existence of an RSE face between peaks, but otherwise the techniques are distinct.

4.2 Impossibility of certain RSE/Peak Combinations

It would be reasonable to assume that finding the upper and lower bounds for the number of RSE faces that can exist in a lattice with a given number of peaks would be enough, as every combination between them would be possible. Surprisingly, this is not the case. At least one of the gaps visible in Table 7 can be proven to exist. Namely, it is possible to have a 4-dimensional lattice with 5 peaks and 18 RSE faces, but it is not possible to have a 4-dimensional lattice with 5 peaks and 17 RSE faces. Moreover, a similar gap exists in every dimension $n \geq 4$. We find the existence of these "holes" in Table 7 fascinating. However, we recognize that the proof of Theorem 6 is the longest in this paper and the result may not have as broad appeal, so we have moved the proof to an appendix.

Theorem 6. *For $n \geq 4$, an n-dimensional lattice with $2^{n-1} - (n-1)$ peaks can have $2^{n-2}\binom{n}{2} - \binom{n}{2}$ RSE faces but not $2^{n-2}\binom{n}{2} - \binom{n}{2} - 1$ RSE faces.*

4.3 Many peaks, many RSE faces

The possibilities for numbers RSE faces and peak combinations are very limited when many peaks or RSE faces are involved. This is because, as the proof of the previous theorem demonstrates, when almost every face is an RSE face, the lattice is very close to the alternating lattice. This can be used to exclude entire rows or columns of Tables 7 and 8.

Theorem 7. *The only n-dimensional lattice with either the maximum number of peaks (2^{n-1}) or RSE faces ($2^{n-2}\binom{n}{2}$) is the alternating lattice.*

Proof. This maximum number of RSE faces corresponds to every face being an RSE face. If every face of a lattice is an RSE face, then by definition it is the alternating lattice and has the maximum number of peaks as well.

This maximum number of peaks corresponds to half the vertices being peaks, since two peaks cannot share an edge. The only way to have half the vertices be peaks (and have $00\cdots0$ be a valley) is for every vertex with an odd number of 1s in its coordinate be a peak and the others be valleys. This is just the alternating lattice. $\square$

This theorem excludes every combination of number of peaks and number of RSE faces where only one of them is at maximum.

Theorem 8. *For $n \geq 4$, if an n-dimensional lattice has at least $2^{n-2}\binom{n}{2} - (n-1) - (n-2)$ RSE faces, then it must have exactly:*

- $2^{n-2}\binom{n}{2}$ *(every face),*
- $2^{n-2}\binom{n}{2} - (n-1)$, *or*
- $2^{n-2}\binom{n}{2} - (n-1) - (n-2)$

RSE faces.

Proof. Lattices with this many RSE faces must be constructed by starting from the alternating lattice and flipping the orientation of a few edges. By Lemma 1 (proof can be found in appendix), flipping one edge removes at least $(n-1)$ RSE faces. No edge flip can remove more than $(n-1)$ RSE faces because each edge touches $(n-1)$ faces, so the first edge flip will remove exactly $(n-1)$ RSE faces. Since flipping more edges just removes more RSE faces, fewer than $(n-1)$ RSE faces is impossible. Similarly, the next edge flip removes at least $(n-2)$ RSE faces, for a minimum of $(n-1)+(n-2)$ non-RSE faces. No value between this and $(n-1)$ is possible because flipping more edges removes even more RSE faces and flipping fewer cannot flip that many. $\square$

The pattern described in the proof of Theorem 8 breaks down here as the first two edge flips could both remove $(n-1)$ RSE faces, so the next lower value would be $2^{n-2}\binom{n}{2} - (n-1) - (n-1)$. A similar argument would show that one fewer RSE face still is impossible in 6 or more dimensions, and that in large enough dimensions there are many such gaps.

5 Computational Strategies

As mentioned above, the number of acyclic orientations of a Boolean lattice is tremendously large as n increases, and so we took a multi-pronged approach to generating computational evidence. We used Python, Maple, and C, and used different construction techniques. Each had their advantages and disadvantages. These are detailed below.

5.1 Gluing Constructions

We made extensive use of the basic gluing and peak-preserving gluing ideas described in Section 2. These were executed in Python, and are presented in the Appendix in Tables 7 and 8.

5.2 Topological Orderings

In this subsection we consider topological orders as a strategy to maximize RSE faces in single-peak orientations. A *topological order* of an oriented graph is a list of all the vertices such that each edge is directed from an earlier vertex to a later vertex in the list. An example of a graph orientation and a corresponding topological order is given in Figure 7

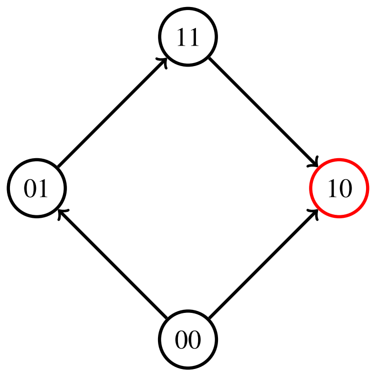

Figure 7: An orientation of Q_2 with topological order 00, 01, 11, 10

A graph orientation has a topological order if and only if the orientation is acyclic. Therefore, a topological order can be a helpful one-dimensional representation of the Boolean lattice orientations we are interested in. However topological orders are not unique. One topological order uniquely determines the direction of each edge in the lattice, but multiple topological orders may really correspond to the same oriented lattice. For example, the orientation of Q_2 in Figure 8 has four possible topological orders. Note that the idea of topological orders that we use here is a special case of the rank orders used by Crona in 2020 [6], where the partial order used is in fact a total order. So one could consider our topological ordering a "total order consistent with the fitness graph". However we chose to retain the term topological order since our purposes were computational, and in the computer science field, that is the terminology used.

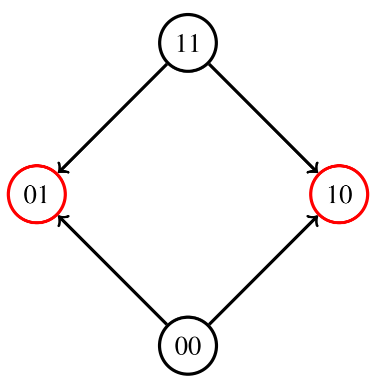

Figure 8: An orientation of Q_2 with multiple topological orders

Since we are interested in RSE faces and peaks, we consider how to detect these two features from a topological order.

First, define an *even (resp. odd) vertex* as a vertex of Q_n whose label has an even (resp. odd) number of 1s. Notice that by construction, every face in Q_n consists of two odd vertices and two even vertices. The face is an RSE if and only if the four edges are all directed from odd vertices to even vertices or all directed from even vertices to odd vertices. Taking this to the extreme, the alternating lattice A_n (i.e. where there is an arrow from u to v exactly when u is even and v is odd) has every face as an RSE. Any permutation that puts all even vertices before all odd vertices is a topological order of the alternating orientation.

Next, we consider conditions that guarantee a single peak.

Proposition 1. *A topological order $v_1, \dots, v_m$ corresponds to a single peak orientation if and only if for all v_i with $i < m$, there exists v_j with $j > i$ such that v_i is adjacent to v_j.*

Proof. Suppose that $v_1, \dots, v_m$ is a topological order that corresponds to a single peak. Then by definition v_m has all edges oriented toward it, so v_m is the peak. If $i < m$, then in order for v_i to not be a peak, it must have at least one outward edge. This outward edge corresponds to an adjacent vertex v_j with $j > i$ that appears later in the topological order.

For the converse, suppose that $v_1, \dots, v_m$ is a topological order that corresponds to multiple peaks. By definition, v_m must be a peak, so suppose v_p ($p \neq m$) is a second peak. If v_p is a peak, then all edges are oriented into it, so all adjacent vertices to v_p appear earlier in the topological order than v_p, which means there exists no v_j matching the conditions of the proposition. □

While the conditions of Proposition 1 technically require considering the set of all $m = 2^n$ vertices of the n-dimensional lattice (denoted $V(Q_n)$) in the topological order to determine whether there is a single peak, we can be more efficient. Suppose there is some $1 \leq c \leq m$ such that every vertex v_i where $i < c$ is adjacent to some vertex in $C = \{v_c, \dots, v_m\}$. Then the vertices of C guarantee the condition of Proposition 1 no matter how the vertices of $\{v_1, \dots, v_{c-1}\}$ are ordered.

We will call a set of vertices $C = \{v_c, \dots, v_m\}$ a *connected cover* when C meets the conditions of Proposition 1 and when for all $v \in V(Q_n) \setminus C$ there exists $u \in C$ such that v is adjacent to u .

We have now arrived at a technique to search for single-peak orientations of Q_n that have multiple RSE faces using topological ordering. We seek a topological ordering of the vertices of Q_n of the form (even vertices, odd vertices, connected cover). The connected cover guarantees a single peak, while the separation of even vertices vs. odd vertices outside the cover is a technique to create as many RSE faces as possible among the remaining vertices.

A connected cover can be determined via a depth first search: picking a desired peak vertex at random, and then adding additional vertices to a set C of vertices one at a time until C is a cover. As soon as a cover is located, we can quickly write a topological order of the form (even vertices, odd vertices, connected cover) and compute the corresponding number of RSE faces.

In practice, we can combine this topological order motivation with other constructions to generate single-peak landscapes with many RSE faces. If Q^1_{n-1} is an acyclic $(n-1)$-dimensional lattice with r_1 RSE faces and Q^2_{n-1} is a single-peak acyclic $(n-1)$-dimensional lattice with r_2 RSE faces, then we can use a basic gluing of Q^1_{n-1} and Q^2_{n-1} into a single-peak orientation of Q_n where each vertex of Q^1_{n-1} has an edge directed toward the corresponding vertex of Q^2_{n-1}. No vertices of Q^1_{n-1} will be peaks since each has an out-degree of at least one. The single peak of Q^2_{n-1} will remain a single peak. Further, the number of RSE faces in this new orientation is given by $r_1 + r_2$.

Since the number of peaks in Q^1_{n-1} is irrelevant to the number of peaks in the glued orientation, we can maximize r_2, and thus $r_1 + r_2$, by using the the alternating orientation as Q^1_{n-1}, whereas, we can use the maximum RSE single-peak orientation discovered by topological ordering search as Q^2_{n-1}. We know that there are $2^{n-3}n(n-1)$ 2-dimensional faces (OEIS A001788) in Q_n, so this gives $r_1 = 2^{n-4}(n-1)(n-2)$, while r_2 must be determined by search.

The data in Table 2 gives a window into this process. The smallest dimension of a Boolean lattice with a single peak and non-zero RSE faces is 3. An example of such an orientation is given in Figure 9. As predicted, we can find this orientation by searching for a connected cover, such as $(101, 010, 110, 111)$. The remaining even vertices are 000 and 011. The remaining odd vertices are 001 and 100, so a topological order is $(000, 011, 001, 100, 101, 010, 110, 111)$. Notice that both even vertices of the one RSE (000 and 011) appear before both odd vertices (001 and 010).

Now that we have a 1-RSE single peak orientation of Q_3, we can use this orientation as Q^2_3 and the alternating lattice A_3 with its 6 RSE faces as Q^1_3 to form a 7-RSE orientation of Q_4. However, a computer search for a topological order improves on this and found an 8 RSE orientation. Similarly, we could take this 8-RSE orientation of Q_4 glued with the alternating orientation of Q_4 (which has 24 RSE faces) to obtain a 32-RSE orientation of Q_5. However, searching for a

dimension	RSE faces by basic gluing (% of all 2d faces)	RSE faces by search (% of all 2d faces)
2	0 (0)	0 (0)
3	1 (16.7)	1 (16.7)
4	7 (29.2)	8 (33.3)
5	32 (40.0)	36 (45.0)
6	116 (48.3)	119 (49.6)
7	359 (53.4)	

Table 2: Results of a hybrid topological construction for RSE faces

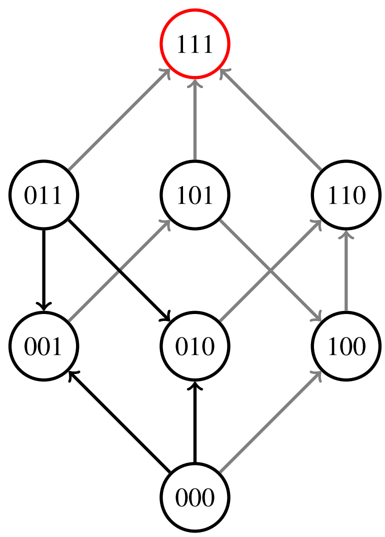

Figure 9: A single-peak orientation of Q_3 with one RSE (highlighted with black edges)

connected cover again revealed a single-peak Q_5 orientation with 36 RSE faces. This 36-RSE orientation can be glued with an alternating orientation of Q_5 (which has 80 RSE faces) to produce an orientation of Q_6 with 116 RSE faces. Searching via topological sort again revealed a single peak 119-RSE orientation of Q_6, which, when combined with an alternating orientation of Q_6 (which has 240 RSE faces), produces a single-peak orientation of Q_7 with 359 RSE faces, and so on.

The data in Table 2 is the result of one round of searching for connected covers, and guarantees there is a single peak orientation with *at least* this many RSE faces, but there could be other topological orders that lead to even higher numbers of RSE faces as the dimension of the lattice increases. However, two interesting trends are illustrated through this guided search process. First, finding a connected cover and producing a well-constructed topological order more efficiently builds an acyclic single-peak orientation than simpler gluing methods. Second, the percentage of RSE faces out of all 2-dimensional faces in Q_n is increasing over time.

5.3 Computation in C

We have also written code in C to analyze the RSE face count problem in low dimensions. The code is based on brute-force technique, but it is highly optimized in actual implementation.

5.3.1 Computational Strategy

The C function is invoked with one argument, the dimension n, and also enables an optional setting to force a single vertex (the wild type) to always be a valley. Subject to that setting, the code begins with the wild type and searches over all possible orientations of a n-dimensional Boolean lattice using reflected binary Gray codes, so that the exhaustive search is performed with each orientation obtained from the previous orientation by reversing a single edge (see [13, 23]) For each orientation, it first checks to ensure the orientation is acyclic using a depth-first search. If so, the code computes the number of RSE faces as well as the number of peaks, each from scratch using brute force. The output of the algorithm is a table that gives the number of acyclic orientations having exactly r RSE faces and k peaks.

For dimension n, the Boolean lattice has 2^n vertices and $n2^{n-1}$ edges: the latter can be derived by deciding in which of the dimensions the edge lies and then enumerating the 2^{n-1} fixed possibilities for settings in the other $n-1$ dimensions. If we set one vertex (e.g., the wild type) as a valley, this fixes the orientation of its n edges, and there are $n(2^{n-1}-1)$ remaining edges to orient. Since each edge can be oriented in either direction, there are $2^{n(2^{n-1}-1)}$ orientations to consider with one vertex fixed as a valley. The relevant parameters for $n \in \{2, 3, 4, 5\}$ are given in Table 3.

Dimension	Vertex count	Edge count	Edge count (one valley)	Orientation count (one valley)
2	4	4	2	$2^2 = 4$
3	8	12	9	$2^9 = 512$
4	16	32	28	$2^{28} = 268435456$
5	32	80	75	2^{75}

Table 3: Relevant parameters for the Boolean lattice on small dimensions.

Currently, our code works for $n \in \{2, 3, 4, 5\}$; however, the astronomical runtime of the program on $n = 5$ makes it intractable in its current state. The number of orientations to exhaust over for $n = 5$ is 2^{75}; on a single personal computer the code is estimated to complete in tens of billions of years. It is unlikely that a supercomputer or cluster could improve this time sufficiently. Rather, the best way forward for the $n = 5$ case is likely to improve the algorithm. We provide three suggestions for improving the algorithm to avoid duplicate work in several ways: (1) store RSE face counts and peak counts and adjust them incrementally as we iterate from one orientation to the next using an edge reversal; (2) account for symmetries in orientations to avoid re-counting equivalent cases; (3) include branch-and-bound steps to “end early" on partial orientations that are already detected to have a cycle. If these three techniques were implemented, we conjecture that the $n = 5$ search could be made practical, in other words, that it could reduce the time required to months instead of billions of years. However, these steps greatly increase the complexity of the code, and we leave our conjecture as an open question, and one that is not easily answered.

5.3.2 Results

In all of our reported results, we force the wild type to be a valley, fixing the orientation of its n edges outward.

For dimension 3, there are 512 total orientations, of which 340 are acyclic. The count of acyclic orientations with r RSE faces and k peaks is given as entry (r, k) in the table of Table 4.

		Peaks				
		0	1	2	3	4
RSE faces	0	0	91	0	0	0
	1	0	84	42	0	0
	2	0	0	93	0	0
	3	0	0	12	8	0
	4	0	0	0	9	0
	5	0	0	0	0	0
	6	0	0	0	0	1

Table 4: For dimension 3, number of acyclic orientations with each (number of RSE faces, number of peaks).

For dimension 4, there are 268435456 total orientations, of which 24671134 (about 9.2%) are acyclic. The count of acyclic orientations with r RSE faces and k peaks is given as entry (r, k) in the table of Table 5. The code runs in approximately 1 minute on a modern personal computer.

6 Conclusion

We obtained a collection of bounds using a variety of gluing techniques and informed by a variety of computational techniques. We suspect our bounds could be improved with finer-grained analysis. We see several directions ripe for investigation, both in the theoretical and computational domains.

One such direction for investigation includes looking at the likelihood of occurrences of r RSE faces in a landscape with k peaks. In a landscape chosen uniformly from all acyclic orientations, what is the most likely number of RSE faces

		Peaks								
		0	1	2	3	4	5	6	7	8
	0	0	299511	0	0	0	0	0	0	0
	1	0	913656	227580	0	0	0	0	0	0
	2	0	1590669	1042032	11211	0	0	0	0	0
	3	0	1482852	2474108	153132	0	0	0	0	0
	4	0	974148	3355704	614796	0	0	0	0	0
	5	0	376440	2623086	1367388	12876	0	0	0	0
	6	0	127548	1459384	1523046	75708	0	0	0	0
	7	0	27936	524706	1211520	196788	0	0	0	0
	8	0	1485	192600	614094	248253	297	0	0	0
	9	0	0	22470	287724	231820	4828	0	0	0
	10	0	0	6180	72684	133764	12012	0	0	0
	11	0	0	0	19980	72144	15444	0	0	0
RSE faces	12	0	0	75	2430	21488	14361	25	0	0
	13	0	0	0	612	8670	9276	306	0	0
	14	0	0	0	0	1116	5220	744	0	0
	15	0	0	0	0	480	1696	650	0	0
	16	0	0	0	0	0	936	798	0	0
	17	0	0	0	0	0	0	216	0	0
	18	0	0	0	0	0	35	264	25	0
	19	0	0	0	0	0	0	42	36	0
	20	0	0	0	0	0	0	0	0	0
	21	0	0	0	0	0	0	0	28	0
	22	0	0	0	0	0	0	0	0	0
	23	0	0	0	0	0	0	0	0	0
	24	0	0	0	0	0	0	0	0	1

Table 5: For dimension 4, number of acyclic orientations with each (number of RSE faces, number of peaks).

given k peaks, and what is the most likely number of peaks given r RSE faces? In a similar vein, what is the expected number of RSE faces given k peaks? Likewise, what is the expected number of peaks given r RSE faces? Although there exist experimental results on fitness landscapes with reasonably large dimension n (for example, Palmer et.al.'s 7 loci [20] and Franke et.al.'s 8 loci [10]), the construction of complete experimental fitness landscapes with large dimension n is difficult, so these probabilistic results may inform predictions for as-yet unmeasured larger landscapes.

A second direction is to extend this work from biallelic to multiallelic landscapes. Mathematically, possible vertex labels would be j-ary strings instead of binary strings. Biologically this could correspond to one wild type and $j-1$ different possible mutations at each locus. There is experimental research that points to biologists' interest in such landscapes. For example, Wu et.al. in 2016 investigated 20 alleles on each of 4 loci for a staggering $160,000$ variants [35].

Another direction is to remove one of our simplifications: we restricted our consideration to RSE faces in 2 dimensions, but as we have mentioned, higher-order epistasis occurs frequently, and could be the source of the variability that we saw in the number of peaks possible with a fixed number of RSE faces. This paper demonstrates that though RSE faces are a requirement for the existence of multiple peaks, they are not sufficient to guarantee it. It is possible that there is a collection of conditions of higher-dimensional epistasis that would give a necessary and sufficient condition for multiple peaks. As a first step, could specifying k occurrences of epistasis among pairs of loci (what we here called RSE faces) and j occurrences of epistasis among collections of 3 loci give us a narrower range of possible peaks? We might even hope that finding the correct restriction on higher dimensional epistatic interactions could make the number of peaks determinable. The converse could also be interesting to explore. Even without the connection to peaks, the distribution of higher-order epistatic interactions given a fixed number of 2-dimensional RSE faces could prove useful in estimating quantities in incomplete fitness landscapes, which many experimental landscapes are.

Finally, we did not allow any adjacent genotypes to both have zero fitness, since it would not allow us to orient the edge between them. Another way to resolve this issue is to instead place a bidirectional edge between two neighboring genotypes that are incompatible with life. This opens up many interesting directions in loosening or tightening the graphical definition of an RSE face to account for such situations.

References

[1] Bauer B, Gokhale C (2015) Repeatability of evolution on epistatic landscapes. Sci Rep. 5:9607. https://doi.org/10.1038/srep09607

[2] Birkhoff GD, Lewis DC (1946) Chromatic Polynomials, Trans Amer Math Soc 60:355-451. https://doi.org/10.2307/1990348

[3] Crona K, Greene D, Barlow M (2013) The peaks and geometry of fitness landscapes. J Theor Biol 317:1-10. https://doi.org/10.1016/j.jtbi.2012.09.028

[4] Crona K, Gavryushkin A, Greene D, Beerenwinkel N (2017). Inferring genetic interactions from comparative fitness data. eLife 6:e28629 https://doi.org/10.7554/eLife.28629

[5] Crona, K (2018). Sign epistasis and the geometry of interactions. arXiv:1802.05228 https://doi.org/10.48550/arXiv.1802.05228

[6] Crona K (2020). Rank orders and signed interactions in evolutionary biology eLife 9:e51004 https://doi.org/10.7554/eLife.51004

[7] Crona K, Krug J, Srivastava M. (2021). Geometry of fitness landscapes: Peaks, shapes and universal positive epistasis. https://arxiv.org/abs/2105.08469

[8] Ferretti L, Weinreich D, Schmiegelt B, Yamauchi A, Kobayashi Y, Tajima F, and Achaz G. (2015). Epistasis and constraints in fitness landscapes. https://arxiv.org/pdf/1507.00041.pdf

[9] Fragata I, Blanckaert A, Dias Louro MA, Liberles DA, Bank C (2019). Evolution in the light of fitness landscape theory. Trends Ecol Evol. 34(1):69-82. doi: 10.1016/j.tree.2018.10.009. Epub 2018 Dec 21. PMID: 30583805.

[10] Franke J, Klözer A, de Visser JAGM, Krug J (2011) Evolutionary Accessibility of Mutational Pathways. PLOS Computational Biology 7(8): e1002134. https://doi.org/10.1371/journal.pcbi.1002134

[11] Gillespie JH (1984). Molecular evolution over the mutational landscape. Evolution 38(5): 1116-1129.

[12] Goulart CP, Mahmudi M, Crona KA, Jacobs SD, Kallmann M, et al. (2013) Designing Antibiotic Cycling Strategies by Determining and Understanding Local Adaptive Landscapes. PLOS ONE 8:e56040. https://doi.org/10.1371/journal.pone.0056040

[13] Gray F, Pulse Code Communication, U.S. Patent 2,632,058, March 17, 1953.

[14] Haldane J (1931). A Mathematical Theory of Natural Selection. Part VIII. Metastable Populations. Mathematical Proceedings of the Cambridge Philosophical Society, 27(1), 137-142. doi:10.1017/S0305004100009439

[15] Hartl DL. (2014). What can we learn from fitness landscapes?. Current opinion in microbiology 21:51–57. https://doi.org/10.1016/j.mib.2014.08.001

[16] Kondrashov DA, Kondrashov FA (2015). Topological features of rugged fitness landscapes in sequence space, Trends in Genetics 31(1):24-33. ISSN 0168-9525, https://doi.org/10.1016/j.tig.2014.09.009. (https://www.sciencedirect.com/science/article/pii/S0168952514001577)

[17] Novais Â, Comas I, Baquero F, Cantón R, Coque TM, et al. (2010) Evolutionary Trajectories of Beta-Lactamase CTX-M-1 Cluster Enzymes: Predicting Antibiotic Resistance. PLOS Pathogens 6:e1000735. https://doi.org/10.1371/journal.ppat.1000735

[18] The On-Line Encyclopedia of Integer Sequences (2021), published electronically at `https://oeis.org`.

[19] Orr, H (2009). Fitness and its role in evolutionary genetics. Nat Rev Genet 10, 531–539. https://doi.org/10.1038/nrg2603

[20] Palmer AC, Toprak E, Baym M, Kim S, Veres A, Bershtein S, Kishony R (2015). Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. Nat Commun. 2015 Jun 10;6:7385. doi: 10.1038/ncomms8385.

[21] Poelwijk F, Kiviet D, Weinreich D, et al. (2007) Empirical fitness landscapes reveal accessible evolutionary paths. Nature 445:383–386. https://doi.org/10.1038/nature05451

[22] Poelwijk FJ, Tănase-Nicola S, Kiviet DJ, Tans SJ. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. J Theor Biol. 2011 Mar 7;272(1):141-4. doi: 10.1016/j.jtbi.2010.12.015. Epub 2010 Dec 16. PMID: 21167837.

[23] Roberts FS, Tesman B (2009) *A*pplied Combinatorics (Boca Raton: CRC Press) p. 87-88.

[24] Sanjuán R (2010). Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Philosophical Transactions of the Royal Society B: Biological Sciences 365:1975–1982.https://doi.org/10.1098/rstb.2010.0063

[25] Saona R, Kondrashov FA, Khudiakova KA (2022). Relation Between the Number of Peaks and the Number of Reciprocal Sign Epistatic Interactions. Bull Math Biol 84, 74 (2022). https://doi.org/10.1007/s11538-022-01029-z

[26] Segal MR, Barbour JD, Grant RM (2004). Relating HIV-1 Sequence Variation to Replication Capacity via Trees and Forests. Statistical Applications in Genetics and Molecular Biology 3:1–18. https://doi.org/10.2202/1544-6115.1031

[27] Squire MB (1998) Generating the Acyclic Orientations of a Graph, J Algorithms 26:275-290. https://doi.org/10.1006/jagm.1997.0891

[28] Stanley R (2006) Acyclic orientations of graphs, Discrete Math 306:905-909. https://doi.org/10.1016/j.disc.2006.03.010

[29] Toprak E, Veres A, Michel JB, Chait R, Hartl DL, Kishony R (2012) Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nature Genetics 44:101–105. https://doi.org/10.1038/ng.1034

[30] de Visser JA, Krug J. (2014) Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet. 15(7):480-90. doi: 10.1038/nrg3744. Epub 2014 Jun 10. PMID: 24913663.

[31] Weinreich DM, Watson RA, Chao L (2005). Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59:1165–1174. https://pubmed.ncbi.nlm.nih.gov/16050094/

[32] Weinrich DM, et al. (2006). Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. SCIENCE 312(5770): 111-114. DOI: 10.1126/science.1123539

[33] Weinreich DM, et al. (2013). Should Evolutionary Geneticists Worry about High Order Epistasis? Current Opinion in Development and Genetics 23: 700-707.

[34] Wright S (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of the Sixth International Congress of Genetics. D. F. Jones. Menasha, WI, Brooklyn Botanic Garden: 356-366

[35] Wu NC, Dai L, Olson CA, Lloyd-Smith, JO, Sun R. (2016) Adaptation in protein fitness landscapes is facilitated by indirect paths eLife 5:e16965. https://doi.org/10.7554/eLife.16965

[36] Wylie CS, Shakhnovich EI (2011). A biophysical protein folding model accounts for most mutational fitness effects in viruses. PNAS 108:9916–9921. https://doi.org/10.1073/pnas.1017572108

7 Appendix: Proof of Theorem 6

Theorem 6. *For $n \geq 4$, an n-dimensional lattice with $2^{n-1} - (n-1)$ peaks can have $2^{n-2}\binom{n}{2} - \binom{n}{2}$ RSE faces but not $2^{n-2}\binom{n}{2} - \binom{n}{2} - 1$ RSE faces.*

Note that since an n-dimensional lattice has 2^n vertices (of which half can be peaks at once) and $2^{n-2}\binom{n}{2}$ faces, this corresponds to lattices where all but $n-1$ possible peak vertices are peaks and all but $\binom{n}{2}$ (or $\binom{n}{2}+1$) faces are RSE faces.

Lemma 1. *Consider the set of faces in an n-dimensional lattice that touch one of j selected edges. Adding another edge to that selection adds at least $n-1-j$ new faces to the set.*

Proof. Two edges which share a vertex define a face. In an n-dimensional lattice, each vertex has n edges, so each of those edges touches $n-1$ faces (paired with every edge but themselves). Therefore, for a new edge to add fewer than $n-1$ faces it must border faces already in the set. Because two edges define a face, each already-selected edge can contribute at most one such face. Therefore, the new edge adds at least $n-1-j$ faces. □

Lemma 2. *Consider a set of edges in an n-dimensional lattice that all share a common vertex, and consider the set of faces that touch these edges. Adding an edge that does not share that vertex will add at least $n-2$ faces.*

Proof. Assume without loss of generality that the common vertex is at $11\cdots1$, and suppose the new edge goes between vertices v and w. Since neither is $11\cdots1$, they each have a 0 bit. Because they are the endpoints of an edge, their coordinates differ in exactly one place, so there has to be a position where both v and w have a 0. Assume without loss of generality that this is the first bit and the second bit varies between them.

Among the faces currently in the set, each vertex contains at most two 0s in their bitstring, so in order for this new edge to be on one of those faces all other bits must be 1. That means the endpoints of this edge are $0011\cdots1$ and $0111\cdots1$. Any faces touching this edge that were already affected must go through $11\cdots1$, which forces the fourth vertex to be at $1011\cdots1$ and uniquely defines the face. Since the edge borders $n-1$ faces, the only way fewer than $n-1$ faces can be added to the set is if some were already included. Since at most one face was already included, this edge adds at least $n-2$ faces. □

Proof of Theorem 6. Consider again the n-dimensional alternating lattice A_n where every face is an RSE face. Then choose one valley of this lattice and reverse all edges containing it, turning it into a peak. All of the n vertices involved in these edges, which used to be peaks, are no longer peaks. This operation has added one new peak but removed n other peaks, for a net loss of $n-1$.

This operation also affects all the faces which had the selected valley as a vertex. There are $\binom{n}{2}$ such faces. Each of these faces has had two of its edges flipped; therefore these faces are no longer RSE faces. The resulting lattice is exactly of the kind specified to be possible by our theorem statement.

Next we will attempt to construct a lattice with one fewer RSE face than this, starting from the alternating lattice and flipping a small number of edges. As long as the vast majority of the faces are RSE faces, they form a contiguous region where every edge matches up with the alternating lattice. In any directed lattice, an edge can only be incident to one peak (since its direction precludes both vertices being peaks.) So flipping each edge can only remove one peak. Then in order to remove $n-1$ peaks, we must flip at least $n-1$ edges. By Lemma 1, the first of these edges removes $n-1$ RSE faces, the second $n-2$, and so on, for a total of

$$(n-1) + (n-2) + (n-3) + \cdots + 1 = \sum_{i=1}^{n-1} i = \frac{n(n-1)}{2} = \binom{n}{2} \tag{6}$$

removed RSE faces. The arrangement we have just described has $2^{n-1} - (n-1)$ peaks and $2^{n-2}\binom{n}{2} - \binom{n}{2}$ RSE faces. In order to have exactly one fewer RSE, that is, $2^{n-2}\binom{n}{2} - \binom{n}{2} - 1$ RSE faces, exactly one additional RSE face would have to be removed.

In principle, achieving this total would be easier if some of the faces affected by flipping edges remained RSE faces. However, doing so would require flipping all four edges of a face, since there are only two possible orientations that result in RSE on a given face, and one is obtained from the other by flipping all four edges. This is inefficient, because a face of the alternating lattice only contains two peaks. The above lower bound for the number of flipped edges to remove $n-1$ peaks can only be achieved if every flipped edge removes a peak, but on this face four flipped edges remove just two peaks. Furthermore, the four edges on this face are poorly arranged to take advantage of Lemma 1.

Achieving Lemma 1's lower bound requires each new edge to share a unique face with each existing edge, but the third and fourth edges of this RSE face share the same face with the first two edges and each other. That necessarily affects three more faces, outweighing the one new RSE face it creates.

Because reintroducing RSE faces is not viable, the only way to remove $\binom{n}{2}+1$ RSE faces is to slightly adjust one of the numbers in (6) to get a sum just one more than the minimum. Suppose there were a list of flipped edges E which did remove exactly one more RSE face. Choose some vertex v which is part of the most edges in E (breaking ties arbitrarily), and arrange E so that the edges containing v are at the start. Consider several cases for how many edges in E contain v.

In the first case, there is only a single edge containing v. Then no pair of edges in E shares a common vertex. Under these restrictions, no pair of edges can even share a face, as flipping opposite edges of an RSE face makes it into a cycle. Then every edge in E affects $(n-1)$ new faces, so after just three edges (the minimum for $n=4$) the number of affected faces is already three more than the minimum.

In the second case, there are three or more edges containing v. Note that E must contain some edge that does not contain v, because the sum in (6) uses only the edges around a single vertex and thus changing it requires edges off that vertex. By Lemma 2, this off-v edge affects at least $(n-2)$ new faces. But because it is at least fourth in the list, the term in (6) it is replacing is at most $(n-4)$. Even assuming the subsequent terms can still be minimized, this difference adds two faces over the minimum, which is too many.

Finally, consider the case where exactly two edges contain v. Now the first off-v edge e_3 is replacing the $(n-3)$ term in (6), so in order to affect no more than one additional face this edge must affect either $(n-2)$ or $(n-3)$ faces. By Lemma 2, only $(n-2)$ is possible. In order to achieve that, e_3 must share a face with one of the on-v edges e_1 and e_2. Without loss of generality, assume e_3 is not opposite one of these edges on a face; such a choice for e_3 must exist, since flipping just the opposite edges on an RSE face makes it a cycle. Then e_3 shares a vertex with either e_1 or e_2; assume it is e_2. Then the three edges e_1, e_2, and e_3 form a chain through four vertices, which in the original alternating lattice can only contain two peaks. As with attempts to flip all four edges on a face, this mismatch between flipped edges and removed peaks means $(n-1)$ edges are no longer sufficient, and an additional edge is necessary. Flipping e_3 has already added the one new affected face, so all subsequent edges in E have to affect the minimum number of faces.

By Lemma 1's reasoning, that means all of the other edges in E must share a face with all previous edges. But this is impossible, as no other edge can share a face with e_2, the middle of the chain. Any attempt to do so will either:

- be on a face containing both e_2 and either e_1 or e_3, which means it does not share a different face with all of them;
- share a vertex with e_2 (and thus e_1 or e_3), which contradicts the assumption that at most two edges share any given vertex;
- or be opposite e_2 on a face, which (since no edge can share a vertex with e_2) turns that face into a cycle when the edges are flipped.

The above cases cover any choice of E, and so no choice of E can work. Therefore, it is impossible to remove exactly $(n-1)$ peaks while also removing exactly $\binom{n}{2}+1$ RSE faces from an alternating lattice. □

8 Appendix: Data Tables

Tables 7 and 8 use the key in Table 6 where each letter denotes which gluing(s) of two $n-1$ dimensional lattices can produce a specific combination of peaks and RSE faces in a lattice n dimensions. Here, the overlapping peak-preserving gluing refers to peak-preserving gluings where peaks in each half match up so not every peak is preserved; in such cases, peaks in the second lattice are given higher priority, which differs from the description in the main text.

Letter	Meaning
G	Basic and peak-preserving
g	Basic and overlapping peak-preserving
P	Peak-preserving
O	Overlapping peak-preserving
B	Basic
F	No gluing, but possible
X	Proven impossible
None	Neither found nor proven impossible

Table 6: Key to Tables 7 and 8

RSE faces \ Peaks	1	2	3	4	5	6	7	8
0	g	X	X	X	X	X	X	X
1	g	G	X	X	X	X	X	X
2	g	G	P	X	X	X	X	X
3	B	G	G		X	X	X	X
4	B	G	G			X	X	X
5	B	G	G	P			X	X
6	B	g	G	G				X
7	B	g	G	G				X
8	F	B	G	G	P			X
9		B	G	G	P			X
10		F	G	G	P			X
11			O	P	P			X
12		F	F	G	P	P		X
13			F	P	P	P		X
14				O	P	P		X
15				O	P	P		X
16					P	P		X
17	X				X	P		X
18	X				O	P	P	X
19	X					P	P	X
20	X	X	X	X	X	X	X	X
21	X						P	X
22	X	X	X	X	X	X	X	X
23	X	X	X	X	X	X	X	X
24	X	X	X	X	X	X	X	P

Table 7: Combinations of number of RSE faces and number of peaks in four dimensions, based on roughly $100,000$ samples.

RSE faces \ Peaks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	g	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1	g	G	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2	g	G	G	X	X	X	X	X	X	X	X	X	X	X	X	X
3	g	G	G		X	X	X	X	X	X	X	X	X	X	X	X
4	g	G	G	P		X	X	X	X	X	X	X	X	X	X	X
5	g	G	G	G			X	X	X	X	X	X	X	X	X	X
6	g	G	G	G				X	X	X	X	X	X	X	X	X
7	g	G	G	G	P				X	X	X	X	X	X	X	X
8	g	G	G	G	G					X	X	X	X	X	X	X
9	g	G	G	G	G						X	X	X	X	X	X
10	g	G	G	G	G	P						X	X	X	X	X
11	g	G	G	G	G	P							X	X	X	X
12	g	G	G	G	G	G								X	X	X
13	g	G	G	G	G	G									X	X
14	g	G	G	G	G	G	P									X
15	g	G	G	G	G	G	P									X
16	g	G	G	G	G	G	P									X
17	B	G	G	G	G	G	P									X
18	B	G	G	G	G	G	G	P								X
19	B	G	G	G	G	G	G	P								X
20	B	G	G	G	G	G	G	P								X
21	B	g	G	G	G	G	G	P								X
22	B	g	G	G	G	G	G	P								X
23	B	g	G	G	G	G	G	P	P							X
24	B	g	G	G	G	G	G	G	P							X
25	B	B	G	G	G	G	G	G	P							X
26	B	B	G	G	G	G	G	G	P							X
27	B	B	G	G	G	G	G	G	P							X
28	B	B	g	G	G	G	G	G	P							X
29	B	B	g	G	G	G	G	G	P	P						X
30	B	B	B	G	G	G	G	G	P	P						X
31	B	B	B	G	G	G	G	G	P	P						X
32	B	B	B	G	G	G	G	G	P	P						X
33	F	B	B	G	G	G	G	G	P	P						X
34		B	B	g	G	G	G	G	P	P						X
35		F	B	g	G	G	G	G	P	P	P					X
36		B	B	B	G	G	G	G	P	P	P					X
37		F	B	B	G	G	G	G	P	P	P					X
38		F	F	B	G	G	G	G	P	P	P					X
39		F	F	B	g	G	G	G	P	P	P					X
40		F	F	F	g	G	G	G	P	P	P					X
41			F	F	F	G	P	G	P	P	P					X
42		F	F	F	g	G	G	G	P	P	P	P				X
43			F	F	F	G	G	G	P	P	P	P				X
44			F	F	F	P	P	P	P	P	P	P				X
45				F	F	O	G	G	P	P	P	P				X
46				F	F	O	P	P	P	P	P	P				X
47				F	F	O	P	P	P	P	P	P				X
48					F	F	P	G	P	P	P	P				X
49					F	F	O	P	P	P	P	P				X
50						F	O	P	P	P	P	P				X
51					F		O	P	P	P	P	P	P			X
52						F	F	P	P	P	P	P	P			X
53						F	O	P	P	P	P	P	P			X
54							F	P	P	P	P	P	P			X
55								O	P	P	P	P	P			X
56							F	F	P	P	P	P	P			X
57								O	O	P	P	P	P			X
58									P	P	P	P	P			X
59									O	P	P	P	P			X
60								O		P	P	P	P	P		X
61									O		P	P	P	P		X
62										P	P	P	P	P		X
63										O	P	P	P	P		X
64											P	P	P	P		X
65	X											P	P	P		X
66	X										O	P	P	P		X

RSE faces \ Peaks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
67	X											P	P	P		X
68	X												P	P		X
69	X											X	P	P		X
70	X											O		P	P	X
71	X												P		P	X
72	X													P		X
73	X													P	P	X
74	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
75	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
76	X														P	X
77	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
78	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
79	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
80	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	P

Table 8: Combinations of number of RSE faces and number of peaks in five dimensions, based on roughly $2,000,000,000$ samples.