INFERRING GENETIC INTERACTIONS FROM COMPARATIVE FITNESS DATA

KRISTINA CRONA*[1], ALEX GAVRYUSHKIN*[2], DEVIN GREENE*[1], AND NIKO BEERENWINKEL[2]

Abstract. Darwinian fitness is a central concept in evolutionary biology. In practice, however, it is hardly possible to measure fitness for all genotypes in a natural population. Here, we present quantitative tools to make inferences about epistatic gene interactions when the fitness landscape is only incompletely determined due to imprecise measurements or missing observations. We demonstrate that genetic interactions can often be inferred from fitness rank orders, where all genotypes are ordered according to fitness, and even from partial fitness orders. We provide a complete characterization of rank orders that imply higher order epistasis. Our theory applies to all common types of gene interactions and facilitates comprehensive investigations of diverse genetic interactions. We analyzed various genetic systems comprising HIV-1, the malaria-causing parasite *Plasmodium vivax*, the fungus *Aspergillus niger*, and the TEM-family of β-lactamase associated with antibiotic resistance. For all systems, our approach revealed higher order interactions among mutations.

Keywords: Fitness landscape, epistasis, gene interactions, fitness graph, rank order, partial order

[1]Department of Mathematics and Statistics, American University, Washington D.C.
[2]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland; SIB Swiss Institute of Bioinformatics, Basel, Switzerland
E-mail addresses: kcrona@american.edu, alex@gavruskin.com, deving@american.edu, niko.beerenwinkel@bsse.ethz.ch.
Date: May 12, 2017.
*Alphabetic order. These authors contributed equally.

1. Introduction

The fitness of an individual with a particular genotype is a measure of its expected contribution to the next generation of the population. The collection of all fitness values for all genotypes, referred to as the fitness landscape, is a central concept in evolutionary biology (Wright 1932; Orr 2009). The fitness landscape can have a strong impact on the fate of the evolving population, such as, for example, the risk of a pathogen population to develop drug resistance and to survive under drug treatment.

Genetic interactions, or epistasis, are abundant in nature. They can have many causes and occur at various scales, for instance, among mutations of a protein-coding sequence or between sequences coding for different genes. Unless there are genetic interactions, we assume that fitness is additive, i.e., the fitness effects of individual mutations sum. An additive fitness landscape is determined by the wild-type and single-mutant fitness values.

If the fitness landscape is determined by the wild-type, single-mutant, and double-mutant fitness values, then we say that it has no higher order epistasis. Intuitively, higher order epistasis means that the fitness of a multiple mutant is unexpected given the fitness of the wild type and all single and double mutants. For example, Weinreich, Delaney, et al. (2006) showed that five mutations jointly increase antibiotic resistance considerably more than expected.

Measuring fitness experimentally is challenging. Fitness measurements tend to come with high uncertainty and they are often obtained only for a subset of genotypes. Moreover, fitness can sometimes not be measured directly at all. Instead, phenotypes are considered that can be measured and are believed to approximate fitness well. For instance, antimicrobial drug resistance is the dominating survival factor for a bacterial population under drug exposure, so that the degree of resistance is a good substitute measure of fitness. Several such fitness proxies are used in microbiology, including survival as measured by disc

diffusion tests. Fitness proxies cannot in general be used for determining epistasis. Although it is possible to study epistasis of the proxy phenotype, additive proxy data does not in general imply absence of epistasis with respect to fitness.

Experimentally, epistatic interactions have been measured in several genetic systems, including *E. coli* (Khan et al. 2011; Weinreich, Delaney, et al. 2006; Poelwijk, Kiviet, et al. 2007), HIV-1 (Silva et al. 2010; Segal, Barbour, and Grant 2004), and other viruses (Wylie and Shakhnovich 2011; Sanjuán 2010). These and similar studies involve the analysis of standing genetic variation or spontaneous mutations (Bonhoeffer et al. 2004; Bershtein et al. 2006), engineered site-directed mutations (Sanjuán, Moya, and Elena 2004; Weinreich, Delaney, et al. 2006), and combinations of both (Sanjuán, Cuevas, et al. 2005; Poon and Chao 2006). Competition experiments are also frequently employed to learn mutational fitness effects. For example, Sanjuán, Moya, and Elena (2004) studied the distribution of deleterious mutational effects in RNA viruses using this approach. Such experiments are typically run on single-nucleotide substitution mutants produced by site-directed mutagenesis. However, little is known about whether or not it is possible to learn higher order genetic interactions from the data produced in competition experiments.

Due to the rapid growth of the number of possible interactions with the number of loci, all interactions can exhaustively be studied only for a small number of loci. At the human genome scale, for example, a complete study of only pairwise gene interactions would already require hundreds of millions of experiments. On the other hand, for smaller organisms, such as yeast, all pairwise and several three-way gene interactions have been measured experimentally (Costanzo et al. 2010). Only when restricting to a small set of preselected loci, can one assess all combinations of mutations. This approach has been pursued, for example, by Weinreich, Delaney, et al. (2006) for a five-locus system associated with bacterial drug resistance.

Historically, the study of genetic interactions was mostly restricted to pairwise epistasis. According to Crow and Kimura (p. 224, 1970)

higher order interactions were generally believed not to be significant in nature, with references to Fisher, Haldane, and Wright. More recent arguments for the same view have been stated in the context of protein folding (Gupta and Adami 2016). On the other hand, empirical findings suggest that the opposite is true for many other systems. For example, Weinreich, Lan, et al. (2013) argue that three-way and four-way interactions can be as strong as pairwise epistasis and refer to various empirical fitness studies. A similar view is expressed in Sailer and Harms (2017). Although the significance of higher order interactions may vary between systems, the topic has not been thoroughly investigated. This is partly due to lack of adequate methodology to quantitatively assess the interactions underlying an observed empirical fitness landscape. Improved mathematical and statistical tools for detecting higher order interactions, as well as more empirical results, are necessary for more conclusive answers regarding the importance of higher order interactions.

In this paper, we consider fitness data that comes in the form of pairwise comparisons. Such data are frequent in practice and can arise in different ways. First, some assays rely on comparing the fitness of two genotypes, for example, by letting them grow in direct competition. Each competition experiment is informative about which of the two genotypes has higher fitness, without estimating the fitness values themselves. Second, direct but uncertain fitness measurements are also often summarized as pairwise fitness relations by recording only whether two genotypes displayed significantly different fitness values or not. Third, rather than fitness itself, a fitness proxy, i.e., a phenotype closely related to fitness, may be considered. Fitness proxies cannot be used directly to measure epistasis, because they generally do not preserve fitness linearity (Gong, Suchard, and Bloom 2013), but if proxy data preserves pairwise comparisons, they may be used instead. Lists of mutants found in a new environment, such as, for example, a new host for a pathogen or a drug environment can be utilized similarly. Assuming that the capability to transition to and survive in the new environment is an indication of higher fitness, this type of observational

data also provides pairwise fitness comparisons. Similarly, the population frequency of genotypes can sometimes be considered to draw conclusions about fitness. For example, by employing a specific model of viral evolution, fitness was inferred computationally from deep sequencing data of an HIV-1 population, and pairwise credible fitness differences were reported (Seifert et al. 2015).

Irrespective of how they were obtained, any consistent set of pairwise fitness relations can be regarded as a partial order of the genotypes with respect to fitness. Two specific types of partial orders play important roles for fitness landscapes. First, if comparisons are available for all pairs of genotypes, then the partial order is a total order, or rank order. In this case, all genotypes are ordered according to fitness. Second, several studies compare fitness only between mutational neighbors, i.e., genotypes which differ at exactly one locus. The resulting partial orders are referred to as fitness graphs and have recently been used extensively (Ogbunugafor, Wylie, et al. 2016; Wu et al. 2016; Smith and Cobey 2016; Mira et al. 2015).

The question addressed in the present study is whether higher order interactions can be inferred from rank orders, fitness graphs, and general partial orders. Connections between rank orders and fitness graphs to epistasis and global properties of fitness landscapes have been observed repeatedly (Greene and Crona 2014; Poelwijk, Tănase-Nicola, et al. 2011; Weinreich, Delaney, et al. 2006; Weinreich, Watson, and Chao 2005). Most recently, Wu et al. (2016) discussed an example of a fitness graph that implies higher order epistasis. It is known that some fitness graphs imply epistasis and an exact condition was given in Crona, Greene, and Barlow (2013). The significance of rank orders of genotypes for epistasis was recognized by Weinreich, Watson, and Chao (2005). The authors introduced the concept of sign epistasis, which refers to the two-locus system. Here, we develop a more general approach that applies to any number of loci and can be regraded as a higher order analogue to sign epistasis. In addition to the theoretical

work mentioned above, rank order arguments have been used for developing antimicrobial treatment strategies (e.g Smith and Cobey 2016; Nichol et al. 2015; Mira et al. 2015).

However, the full potential of rank order consideration for the comprehensive analysis of epistatic gene interactions in general n-locus genetic systems has not been exploited. Furthermore, to the best of our knowledge the general case of arbitrary partial fitness orders has yet to be considered.

Here, we develop quantitative tools to detect virtually any type of gene interaction described in the literature, including epistasis as described by Fourier coefficients, Walsh coefficients, and circuits (Beerenwinkel, Pachter, and Sturmfels 2007; Weinreich, Lan, et al. 2013). In particular, our approach applies to total n-way epistasis, conditional epistasis, and marginal epistasis. We used our approach to analyze genetic interactions in HIV-1, the parasite *Plasmodium vivax*, the fungus *Asbergillus niger*, and β-lactamase antibiotic resistance systems. In all cases, we detect higher order interactions based only on partial information about the fitness order of genotypes, without knowing or estimating the actual fitness values.

2. Results

We consider genetic systems consisting of n biallelic loci. A genotype can then be represented as a binary string of zeros and ones of length n, where 0 denotes the wild-type allele and 1 the alternative allele. We assume that fitness is additive in the absence of epistasis. The fitness of a genotype g is denoted by w_g, and we assume that the fitness landscape w is generic in the sense that no two genotypes have exactly the same fitness.

A complete analysis of all epistatic interactions would require fitness measurements of all 2^n genotypes. However, this level of completeness is rarely available in empirical data sets due to experimental design or an infeasible number of genotypes. To address this limitation, we developed methods that are applicable to partial orders of genotypes according to fitness. For example, the two fitness relations $w_{01} > w_{00}$

and $w_{10} > w_{11}$ together define a partial order. One can always extend a partial order to a rank order, i.e., a total order of the genotypes in the system from highest to lowest fitness. For example, the total order $w_{10} > w_{11} > w_{01} > w_{00}$ extends the partial order above. Our goal is to understand what fitness rank orders and more generally partial fitness orders of genotypes reveal about gene interactions.

Two-locus case. We first consider epistasis for a biallelic two-locus population consisting of the unmutated genotype, or wild type, 00, the two single mutants 01 and 10, and the double mutant 11. In this case, epistasis is denoted by ε_2, where the index 2 refers to the number of loci. It is defined as the deviation from additivity,

$$\varepsilon_2 = (w_{00} + w_{11}) - (w_{01} + w_{10}). \tag{1}$$

The system has no epistasis if $\varepsilon_2 = 0$, positive epistasis if $\varepsilon_2 > 0$ and negative epistasis if $\varepsilon_2 < 0$.

We first assume that the available information on fitness is a rank order of the genotypes (Figure 1). The rank order is sometimes sufficient for determining that the system has epistasis. For instance, the rank order $w_{11} > w_{00} > w_{10} > w_{01}$ (Figure 1 rank order 3), implies $w_{00} + w_{11} > w_{01} + w_{10}$, so $\varepsilon_2 > 0$. It follows that the rank order alone allows one to detect positive epistasis without knowledge of the actual fitness values. There are 24 rank orders of the biallelic two-locus system. Among these, eight imply positive epistasis, eight imply negative epistasis, and eight do not permit any inference regarding epistasis. In total two thirds of the rank orders imply epistasis. Each rank order which implies epistasis also determines the sign of ε_2 (Figure 1).

Sometimes even a partial order of the genotypes is sufficient for determining that the system has epistasis. For instance, if we know that $w_{01} > w_{00}$ and $w_{10} > w_{11}$, then we can infer that the system has negative epistasis (Figure 2 (a)). To see this, we consider all rank orders that extend the partial order. There are six such total extensions, namely rank orders 9, 10, 12, 13, 14, and 16 in Figure 1, and all imply negative epistasis. We conclude that the partial order implies epistasis, based

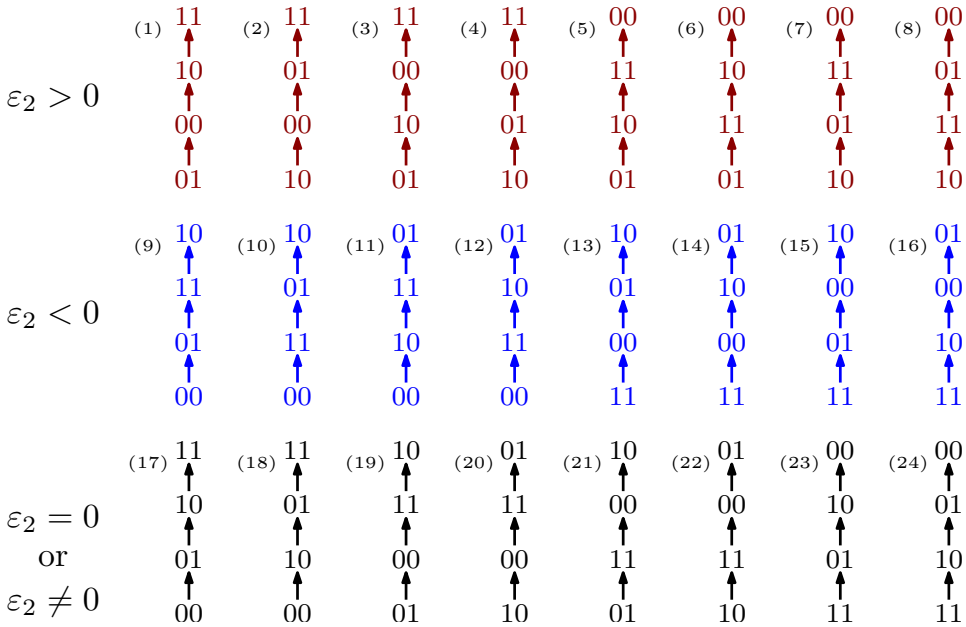


FIGURE 1. All 24 rank orders of the biallelic two-locus system, where the 16 colored rank orders imply epistasis. Red (top row) indicates positive epistasis and blue (middle row) negative epistasis.

on only two fitness comparisons and without knowing any of the actual fitness values. This observation holds in general: If all total extensions imply epistasis then the same is true for the partial order. We will use this argument repeatedly.

A partial order can also be compatible with several rank orders, some of which might imply epistasis while others do not. In this case, the information is not sufficient to detect epistasis from the partial order alone. For example, the partial order $w_{00} > w_{01} > w_{10},\ w_{11}$ is compatible with the two rank orders $w_{00} > w_{01} > w_{11} > w_{10}$ and $w_{00} > w_{01} > w_{10} > w_{11}$ (Figure 2(b)). The first rank order implies positive epistasis, but the other one does not. Consequently, the partial order does not reveal whether or not the system has epistasis, and further comparisons are needed for a conclusion.

Fitness graphs constitute an important subclass of partial orders, as they often are the reported result of experiments, and because of their relevance for evolutionary processes (Figure 3). Briefly, the nodes

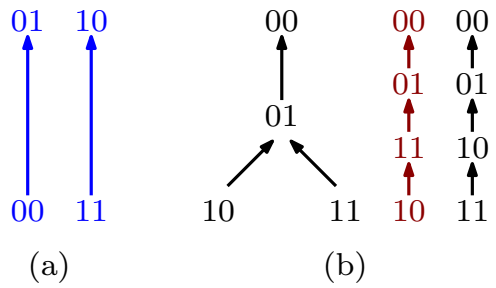


FIGURE 2. (a) A partial fitness order of genotypes. The rank orders that extend this partial order are orders (9), (10), (12), (13), (14), and (16) in Figure 1. All of them imply negative epistasis ($\varepsilon_2 < 0$). (b) A partial order of genotypes with all its total extensions shown on the right. The first extension shown in red implies positive epistasis ($\varepsilon_2 > 0$), while the second one in black does not.

of a fitness graph represent genotypes and for each pair of mutational neighbors, i.e., genotypes which differ at exactly one locus, an arrow points toward the genotype of higher fitness (SectionA.3).

A fitness graph implies epistasis exactly when all rank orders compatible with the graph do, as is the case for partial orders in general. For example, Figure 3 shows the four fitness graphs where genotype 00 has lowest fitness in the system. The graphs (b), (c) and (d) imply epistasis, whereas (a) is compatible with additive fitness.

A couple of observations from Figure 3 are useful for determining if a system is compatible with additive fitness. First, any rank order compatible with the graph (a) has the following property: For each genotype, replacing 0 by 1 results in a genotype of higher fitness. If the genotype 00 has minimal fitness in the system, then rank orders are compatible with additive fitness exactly if they satisfy the property. The second observation is that a fitness graph such that 00 has minimal fitness is compatible with additive fitness exactly if all arrows point up. Both observation generalize to any number of loci, and can be phrased in full generality (see Section A.3).

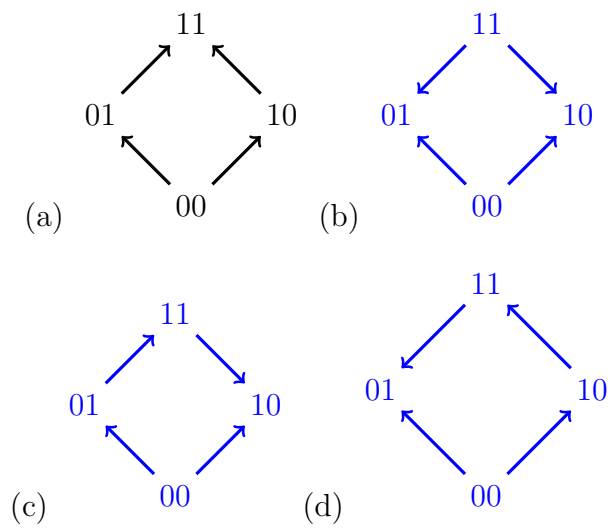


FIGURE 3. For a biallelic two-locus system where the genotype 00 has the lowest fitness, there are four fitness graphs. The graph (a) is compatible with additive fitness, whereas the remaining graphs imply negative epistasis.

Three-locus case. The biallelic three-locus system consists of the eight genotypes 000, 001, 010, 011, 100, 101, 110, and 111. The system has total three-way epistasis if

$$\varepsilon_3 = (w_{000} + w_{011} + w_{101} + w_{110}) - (w_{001} + w_{010} + w_{100} + w_{111}) \neq 0. \quad (2)$$

For the three-locus system, we distinguish between fitness landscapes with no epistasis (fitness is additive), with pairwise but not higher order epistasis (fitness is not additive but $\varepsilon_3 = 0$), and with three-way epistasis ($\varepsilon_3 \neq 0$)

Some rank orders imply three-way epistasis, similar to our observation of epistasis in the two-locus case. The condition for when a rank order implies three-way epistasis is remarkably simple, and we demonstrate it with an example. Consider the rank order

$$w_{110} > w_{111} > w_{101} > w_{011} > w_{100} > w_{010} > w_{000} > w_{001}. \quad (3)$$

We can represent this rank order by a word in the letters e and o using the following procedure. The genotype 110 with the highest fitness is

represented by e because it has an even number of 1's, the genotype 111 with the second highest fitness is represented by o because it has an odd number of 1's, and so forth. Working from highest to lowest fitness, we obtain the word

$$e\,o\,e\,e\,o\,o\,e\,o \tag{4}$$

If one reads the word letter by letter from left to right, then one has never encountered more o's than e's. This property means that $eoeeooeo$ is a Dyck word (Stanley 1999).

For a biallelic three-locus system, a rank order implies three-way epistasis exactly if its associated word (where the role of e and o can be interchanged) is a Dyck word (Supplementary Information, Proposition 1). This simple rule allows us to conclude that an empirical system has three-way epistasis. As in the two-locus case, a landscape may have three-way epistasis even if the rank order does not imply it. For biallelic three-locus systems, there are in total 40,320 rank orders, of which 16,128 (40%) imply three-way epistasis (Supplementary Information, Proposition 1).

Fitness graphs can be analyzed by using our results on rank orders as in the two-locus case. Figure 4 shows three fitness graphs for three-locus systems. The fitness graph (a) implies three-way epistasis, the graph (b) pairwise but not higher order epistasis, and the graph (c) is compatible with additive fitness.

There are in total $1,862$ fitness graphs for the biallelic three-locus system, of which 698 graphs (37%) imply three-way epistasis. In principle one can check a particular three-locus system for higher order epistasis using this result. However, it is not convenient to work with a list of over one thousand graphs. In order to make the problem more tractable, we can utilize the fact that some fitness graphs are isomorphic (Supplementary Figure 8). There are 54 distinct isomorphism classes of graphs for the three-locus system, of which 20 imply higher order epistasis (Supplementary Information, Section A.3). Consequently, to detect three-way epistasis, one can find the isomorphism

class and then check 54 graphs, namely one for each isomorphism class (Supplementary Information, Figure 10).

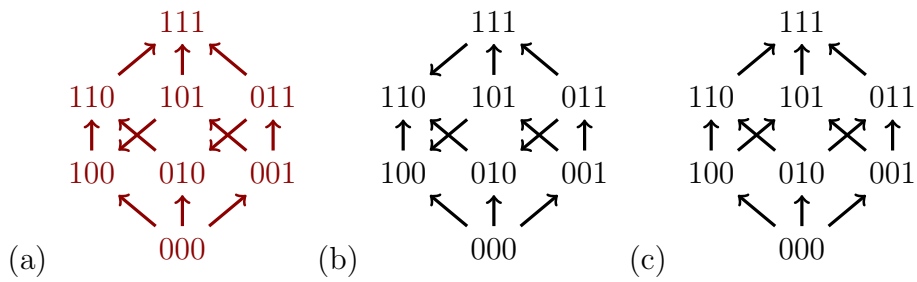


FIGURE 4. The fitness graph (a) implies three way epistasis, the graph (b) implies epistasis, but not higher order epistasis, and (c) does not imply epistasis, since all arrows point up.

We complete the consideration of the three-locus case by analyzing partial orders. Again, in favorable cases one can infer three-way epistasis. Indeed, if there exists a partition of all eight genotypes into four pairs (g_e, g_o), where e and o are as above, and $w_{g_e} > w_{g_o}$ for each pair, then one can conclude three-way epistasis (Supplementary Information, Proposition 7).

General n-locus case. The results on rank orders and higher order epistasis for $n = 3$ generalize to any number of loci. The definition of n-way epistasis in an n-locus system is analogous to the three-locus case, as is the condition for when rank orders imply n-way epistasis. Accordingly, a characterization of rank orders that imply n-way epistasis can be phrased in terms of Dyck words (Supplementary Information, Proposition 1). From this result it follows that the fraction of rank orders that imply n-way epistasis is $2/(2^{n-1}+1)$ (Supplementary Information, Corollary 2) and that it can be determined in a computationally efficient manner whether or not a rank order implies n-way epistasis.

Rank order methods are useful for analyzing the total n-way epistasis for an n-locus system, as demonstrated. However, a single quantity cannot capture all possible gene interactions. Rank order approaches

have the capacity to reveal finer interactions as well. We start with a description of gene interactions based on the approach in (Beerenwinkel, Pachter, and Sturmfels 2007), before exploring what rank orders can reveal about these interactions. For simplicity, we restrict our analysis to the three-locus system, although the arguments used are readily extendable to any number of loci n.

First, we consider gene interactions that directly correspond to the two-locus case by fixing one allele at the third locus. For example, if we fix the third locus at 0, then

$$a = w_{000} + w_{110} - w_{010} - w_{100}$$

measures pairwise epistasis between the first and second locus. Similarly, if we set the third locus to 1, then

$$b = w_{001} + w_{111} - w_{011} - w_{101}$$

measures pairwise epistasis between the first and the second locus. The interactions described by a and b are referred to as conditional epistasis.

An example of an interaction with no correspondence in the two-locus setting is

$$m = w_{001} + w_{010} + w_{100} - w_{111} - 2w_{000},$$

which compares the fitness of the triple mutant 111 to the three single mutants. This expression is negative if the triple mutant has higher fitness than one would expect based on the fitness effects of the three single mutations. Expressions of these kinds are known as circuits, or minimal dependence relations, and for the biallelic three-locus system, there are twenty of them in total, including a, b, and m (Supplementary Information, Section A.2).

If interactions a and b differ substantially, it may be important to consider both of them. However, sometimes we are rather interested in the average interaction for two loci, i.e., epistasis averaged over all genetic backgrounds. We refer to such average effects as interaction coordinates. The four interaction coordinates differ from Walsh coefficients only by a constant scaling factor.

Arguments based on Dyck words can be used for analyzing rank orders and finer gene interactions, including interaction coordinates and circuits, exactly as for three-way epistasis. For each circuit and interaction coordinate, we identify all rank orders that determine its sign. The characterization is given in terms of general Dyck word conditions. We found that for each interaction coordinate, 2/5 of all rank orders determine its sign; for each circuit corresponding to either conditional two-way interaction or marginal epistasis between two pairs of loci, 2/3 of all rank orders determine its sign; and for each circuit relating the three-way interaction to the total two-way epistasis, 1/2 of all rank orders determine its sign (Supplementary Information, Corollary 4).

The signs of all twenty circuits determines the polyhedral shape of the fitness landscape (Beerenwinkel, Pachter, and Sturmfels 2007). The shape combines the circuit information into a more manageable object. However, no rank order determines a shape for $n = 3$ (Supplementary Information, Section A.2).

Our tools for detecting gene interactions work for total n-way epistasis, interaction coordinates and circuits. Moreover, our approach applies to any type of gene interaction that can be expressed by a linear form (Supplementary Information, Theorem 3), such as Fourier coefficients (Beerenwinkel, Pachter, and Sturmfels 2007) and Walsh coefficients (Weinreich, Lan, et al. 2013). We have implemented algorithms for detecting the gene interactions described in this section, both for rank orders and partial orders for three locus systems (`https://github.com/gavruskin/fitlands#fitlands`). In particular, algorithms for three-way epistasis, interaction coordinates and circuits are provided (Supplementary Information, Section B).

Analysis of empirical fitness data. As proof of principle, we applied our tools to fitness data from a diverse set of biological systems, ranging from HIV-1 (Segal, Barbour, and Grant 2004), malaria (Ogbunugafor and Hartl 2016), antibiotic resistance (Mira et al. 2015), to the fungus *Aspergillus Niger* (Franke et al. 2011). Our approach reveals higher order epistasis for all of these systems, only by considering

rank orders and partial orders of genotypes, without the need to access direct fitness measurements or estimates.

Our first application is to the HIV-1 data published by Segal, Barbour, and Grant (2004). Following Beerenwinkel, Pachter, and Sturmfels (2007), we consider the three-locus biallelic system that consists of the mutation L90M in the protease and mutations M184V and T215Y in the reverse transcriptase of HIV-1. Fitness was measured as the number of offspring in a single replication cycle of the virus in the original study, and was reported relative to the wild-type strain NL4-3 on a logarithmic scale. The data consist of 288 fitness measurements, including between 5 and 214 replicates per genotype.

The following rank order was obtained by comparing the mean fitness of the eight genotypes:

$$w_{000} > w_{100} > w_{011} > w_{110} > w_{101} > w_{001} > w_{010} > w_{111},$$

where 000 corresponds to the sequence of amino acids LMT and 111 to MVY comprising the three selected loci. This rank order implies positive three-way epistasis because the associated word

$$e\,o\,e\,e\,e\,o\,o\,o$$

is a Dyck word. It follows that the three mutations under consideration together have a stronger effect on fitness than one would predict from single and double mutants. A closer inspection of the word reveals more information. If we swap any two adjacent letters in the word, then we still have a Dyck word, with the single exception of the first two letters. In other words, only one pair of adjacent genotypes in the rank order, namely 000 and 100, could violate the conclusion if transposed.

If the experiment was to continue, our analysis could be used to direct the data collection process. Indeed, the argument above suggests that the position of the genotype 100 may violate the conclusion of positive three-way epistasis. To quantify the uncertainty in the ranking of 100 with respect to the wild type 000, we employed the Wilcoxon rank sum test on the replicate fitness measurements. The p-value of the test is 0.47 for the relation $w_{000} > w_{100}$, which implies considerable

uncertainty and justifies our recommendation of further experiments to clarify the position of 100. Importantly, the suggested experiment reduces the number of measurements required to make a more robust conclusion about epistasis considerably, namely to one out of 28 possible comparisons.

We proceeded by considering other types of gene interactions in this data set. When considering all 20 circuits, the rank order implies interactions for 55% of the circuits, with positive sign for 30% and negative for 25% of the circuits. This result is consistent with the conventional statistical approaches that use direct fitness measurements. Indeed, since the empirical study of Segal, Barbour, and Grant (2004) provided multiple fitness measurements of each genotype, it was possible to compare the conclusions of our analysis with statistical tests based on fitness estimates (Supplementary Information, Section B). Testing for interactions directly based on the replicate fitness measurements of each genotype confirmed our results based only on the rank order. Specifically, both approaches detected three-way epistasis. Furthermore, 11 of the 20 circuit interactions have been detected by our method and confirmed by t-test. In addition to the 11 interactions, the t-test found 6 circuit interactions as significant (Supplementary Information, Table 2). We emphasize that the rank order approach required much less information to arrive at the same conclusions, thus demonstrating the power of the method.

We conclude that the three sites in the HIV-1 genome under consideration are prone to a diverse set of interactions. Specifically, the strong support for the three-way epistasis, along with the 55% of informative circuit interactions, imply that the three loci together interact in a complex manner, meaning that the interactions cannot be explained using pairwise interactions alone. Thus, in this data set, higher order interactions have a strong impact on the fitness landscape.

Our second application is to a study of antimicrobial drug resistance in malaria (Ogbunugafor and Hartl 2016). The authors measured growth rates for several mutants of *Plasmodium vivax* under exposure to the antimalarial drug pyrimethamine. We identified higher order

epistasis by analyzing rank orders. More precisely, we considered a three-locus sub-system of the study that consists of mutations N50I, S58R, and S117N, in the context of T173L, a fixed mutation, under nine different concentrations of pyrimethamine. The genotypes comprising positions 50, 58, and 117 are labeled 000 (NSS), 100, 010, 001, 110, 101, 011, and 111 (IRN). The three highest concentrations of the drug resulted in the following rank orders:

$$w_{111} > w_{011} > w_{001} > w_{101} > w_{010} > w_{100} > w_{110} > w_{000}$$

$$w_{111} > w_{011} > w_{001} > w_{010} > w_{100} > w_{101} > w_{110} > w_{000}$$

$$w_{111} > w_{011} > w_{010} > w_{001} > w_{100} > w_{110} > w_{101} > w_{000}.$$

The corresponding words are *oeoeooee* for the first rank order, obtained under the highest concentration of the drug, and *oeoooeee* for the second and third rank orders. Since we obtain Dyck words in all cases, the system has negative three-way epistasis for the three highest concentrations of the drug. This consistency shows that the result is robust.

Next, we applied our tools to a study of the TEM-family of β-lactamase, associated with antibiotic resistance (Mira et al. 2015), The study measured growth rates for 16 genotypes exposed to 15 different antibiotics. Specifically, all 16 genotypes that combine subsets of the four amino acid substitutions M69L, E104K, G238S, N276D found in TEM-50, including eight known enzymes, were created using site-directed mutagenesis. We considered the fitness graph obtained when the system was exposed to the antibiotic FEP Cefepime at a concentration of 0.0156 μg/ml (Figure 5). The fitness graph implies higher order epistasis (Supplementary Information, Proposition 7), that is, the fitness of TEM-50 cannot be predicted even with complete knowledge of the fitness values of the remaining genotypes in the system. We conclude that even though some of the single and triple mutants confer low antibiotic resistance, a large population of triple mutants alone is more prone to become antibiotic resistant due to the epistatic fitness

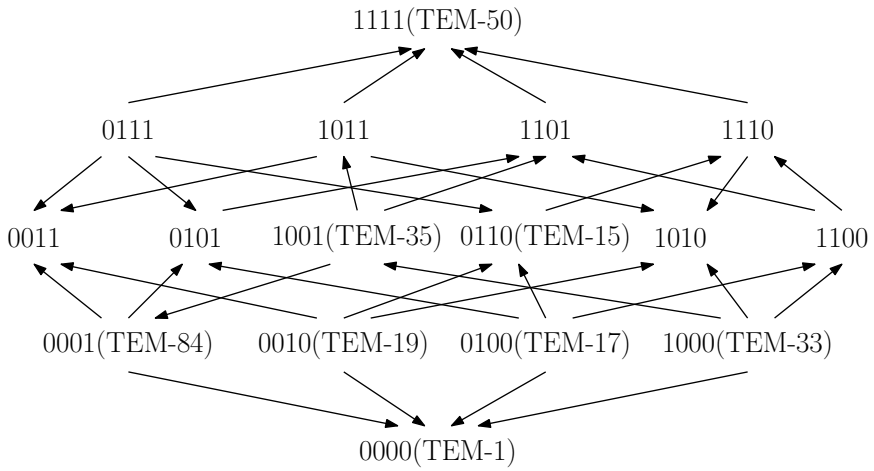


FIGURE 5. The TEM-family of β-lactamase contributes to antibiotic resistance problems in hospitals. The fitness graph shows a four-locus system consisting of the wild type, TEM-1, the quadruple mutant, TEM-50, and all intermediate mutants, including six clinically found mutants in the TEM family. The mutation M69L corresponds to 1000, E104K to 0100, G238S to 0010, and N276D to 0001. Growth rates were measured for the 16 genotypes under exposure to the antibiotic FEP Cefepime, and the fitness graph was determined accordingly (Mira et al. 2015). The graph reveals higher order epistasis.

advantage of TEM-50, as compared to a setting with no higher order epistasis.

Finally, we investigated a study of the filamentous fungus *Aspergillus Niger* (Franke et al. 2011). We considered a system consisting of the wild type and all combinations of the four individually deleterious mutations fwnA1, leuA1, oliC2 and crnB12 (Figure 6). Fitness was estimated with two-fold replication by measuring the linear mycelium growth rate in the original study. The fitness graph implies higher order epistasis (Supplementary Information, Proposition 7).

All four arrows incident to 0000 point towards the genotype, so that the genotype 0000 is a peak in the landscape. The same is true for the genotypes 1100, 0011, and 1001. Because of the four peaks, it is possible that the fungus population gets stranded at a suboptimal peak

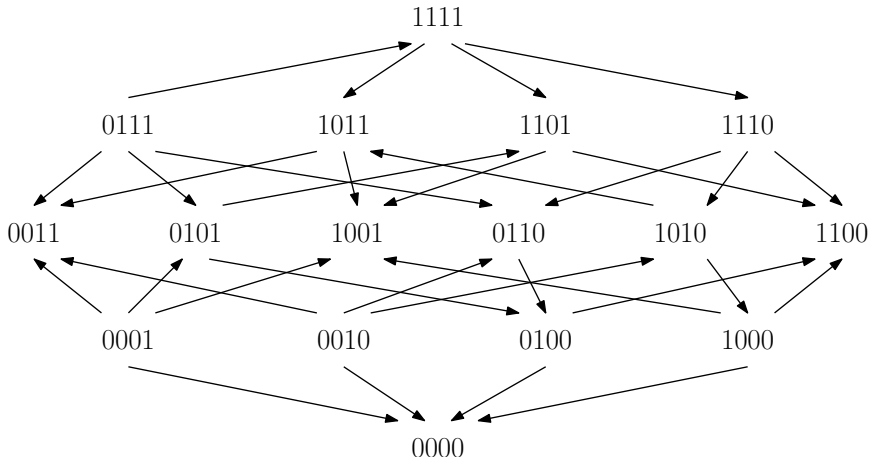


FIGURE 6. The fitness graph shows a four-locus system for the filamentous fungus *Aspergillus Niger*. The system consists of all combinations of the four individually deleteriouis mutations $fwnA1$, $leuA1$, $oliC2$ and $crnB12$. The landscape has in total four peaks, labeled $0000, 1100, 0011$ and 1001.

during the course of evolution (we do not necessarily assume that the starting point for an evolutionary process is at 0000). In contrast, an additive fitness landscape is single peaked. This example illustrates that epistasis may have an impact on the evolutionary dynamics. Several peaks can make the evolutionary process less predictable, depending also on other factors such as population size, mutation rate, etc. More generally, for three-locus fitness graphs, we analyzed the impact of higher order epistasis versus only pairwise epistasis systematically. We found that higher order epistasis correlates with more peaks as well as other features that can lead to involved evolutionary dynamics (Supplementary Information, Section A.4).

3. Discussion

Gene interactions play a critical role in evolutionary processes. Important features of fitness landscapes, such as the number of peaks, and accessible evolutionary trajectories, depend on epistastic gene interactions. The importance of higher order versus pairwise epistasis,

within and among genes or in non-coding regions, as well as the impact of higher order epistasis on evolutionary dynamics, has not been thoroughly investigated. Progress in all of these areas requires adequate mathematical and statistical approaches, in addition to empirical studies.

Here, we have developed new quantitative tools for detecting gene interactions from empirical data. The main advantage of our tools is that they can reveal gene interactions from the types of data frequently generated in empirical studies, specifically rank orders, fitness graphs, and general partial orders of genotypes. The reasons why, in practice, these types of data are available more often than precise fitness measurements for each genotype are manifold. They include restricted comparative experimental designs and known and unknown confounding factors in measuring fitness that can result in uncertain and biased estimates. The methods presented here allow for studying epistatic interactions even when direct fitness measurements are lacking or only a subset of pairwise fitness comparisons is available, either as the immediate outcome of the experiment or the reported summary.

We provide a complete characterization of rank orders that imply higher order epistasis, along with precise results for fitness graphs of three-locus systems. In principle, our approach applies to general partial orders as well, and we have implemented algorithms accordingly. However, because of the increasing complexity it would be desirable to have theoretical results for handling large systems. In particular, a characterization of fitness graphs that imply higher order epistasis is of independent mathematical interest.

We found that for biallelic three-locus systems, 40% of all possible rank orders and 37% of all possible fitness graphs imply higher order epistasis. These fractions suggest that our methods have a good capacity to detect higher order epistasis among three loci, even if exact and complete fitness measurements or estimates are not available. Moreover, the power of our methods was demonstrated for a diverse set of biological systems. We detected higher order epistasis for HIV,

malaria, the fungus *Aspergillus Niger*, and antibiotic resistance systems. Our findings suggest that genetic interactions beyond two-way epistasis shape the fitness landscapes of these genetic systems and may play an important role in determining their evolutionary trajectories. We also exhaustively investigated various types of higher order interactions in HIV-1 and discovered a complex pattern of interactions, confirming that our approach is powerful enough to detect finer gene interactions. Specifically, we identified over twenty interactions by conventional approaches, and rank order methods detected about half of them.

Another important application of our method is to experimental design. When the information available in the data does not contradict an interaction, but is not conclusive enough to claim the interaction, for example because the number of performed competition experiments is too small, then the method allows for prioritizing further experiments by suggesting additional comparisons of genotypes. This feature may prove useful in guiding fitness experiments that aim for testing specific interactions and allow for iteration.

For a more theoretical perspective, we emphasize the distinction between rank order-induced gene interactions, and interactions that do not change the rank order of genotypes. This distinction was pointed out for bialellic two-locus subsystems by Weinreich, Delaney, et al. (2006), who introduced the term "sign epistasis". In our terminology, sign epistasis is a special case of rank order-induced epistasis. Rank order induced epistasis can thus be considered a higher order analogue to sign epistasis, and this new concept is meaningful in a broader context. Indeed, there exist a number of possible ways to quantify and interpret higher order interactions (Weinreich, Lan, et al. 2013; Beerenwinkel, Pachter, and Sturmfels 2007; Hallgrímsdóttir and Yuster 2008), and our rank order approach applies to virtually any type of gene interaction. In particular, we can detect interactions as described by Fourier coefficients and Walsh coefficients. From our general argument based on Dyck words we investigated three-locus systems, and determined the number of rank orders that imply circuit interactions, including

conditional and marginal epistasis, and similarly for interactions coordinates. The method works equally well for for other interactions.

Further investigation of rank order-induced interaction has the potential to relate global and local properties of fitness landscapes, similarly to results on sign epistasis (Weinreich, Delaney, et al. 2006; Poelwijk, Tănase-Nicola, et al. 2011; Crona, Greene, and Barlow 2013). Global properties concern peaks and mutational trajectories in the fitness landscape, whereas local properties concern, for instance, fitness graphs for small system. The relation between global and local properties is important since only local properties can be easily observed in experiments or nature.

Although we have applied our method here only to fitness, any other continuous phenotype of interest can be analyzed in exactly the same manner. The fitness landscape w is then replaced by a more general genotype-to-phenotype map. For example, rather than using it as a fitness proxy, one may be concerned about the drug resistance phenotype itself and its genetic architecture.

In summary, rank order methods have potential for the interpretation of empirical data, as well as for relating higher order gene interactions and evolutionary dynamics. Our approach facilitates detecting higher order epistasis from a very broad range of empirical data, and will therefore contribute to enhancing our general understanding of empirical fitness landscapes and epistatic gene interactions.

Acknowledgments

This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing.

References

Beerenwinkel, Niko, Lior Pachter, and Bernd Sturmfels (2007). "Epistasis and Shapes of Fitness Landscapes". In: *Statistica Sinica* 17, pp. 1317–1342. URL: `http://arxiv.org/abs/q-bio/0603034`.

Bershtein, Shimon et al. (2006). "Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein". en. In: *Nature* 444.7121, pp. 929–932. URL: http://dx.doi.org/10.1038/nature05385.

Bonhoeffer, Sebastian et al. (2004). "Evidence for positive epistasis in HIV-1". en. In: *Science* 306.5701, pp. 1547–1550. URL: http://dx.doi.org/10.1126/science.1101786.

Costanzo, Michael et al. (2010). "The genetic landscape of a cell". en. In: *Science* 327.5964, pp. 425–431. URL: http://dx.doi.org/10.1126/science.1180823.

Coxeter, Harold Scott Macdonald (1973). *Regular polytopes*. Courier Corporation.

Crona, Kristina, Devin Greene, and Miriam Barlow (2013). "The peaks and geometry of fitness landscapes". In: *J. Theor. Biol.* 317, pp. 1–10. URL: http://dx.doi.org/10.1016/j.jtbi.2012.09.028.

Crow, J F and M Kimura (1970). *An introduction to population genetics theory*. Harper and Row, New York.

Franke, Jasper et al. (2011). "Evolutionary accessibility of mutational pathways". en. In: *PLoS Comput. Biol.* 7.8, e1002134. URL: http://dx.doi.org/10.1371/journal.pcbi.1002134.

Gärtner, B and V Kaibel (1998). "Abstract objective function graphs on the 3-cube". In: URL: http://e-collection.library.ethz.ch/view/eth:4337.

Gong, Lizhi Ian, Marc A Suchard, and Jesse D Bloom (2013). "Stability-mediated epistasis constrains the evolution of an influenza protein". en. In: *Elife* 2, e00631. URL: http://dx.doi.org/10.7554/eLife.00631.

Greene, Devin and Kristina Crona (2014). "The changing geometry of a fitness landscape along an adaptive walk". In: *PLOS Comput Biol* 10.5, e1003520.

Gupta, Aditi and Christoph Adami (2016). "Strong Selection Significantly Increases Epistatic Interactions in the Long-Term Evolution of a Protein". en. In: *PLoS Genet.* 12.3, e1005960. URL: http://dx.doi.org/10.1371/journal.pgen.1005960.

Hallgrímsdóttir, Ingileif B and Debbie S Yuster (2008). "A complete classification of epistatic two-locus models". In: *BMC genetics* 9.1, p. 17.

Khan, Aisha I et al. (2011). "Negative epistasis between beneficial mutations in an evolving bacterial population". en. In: *Science* 332.6034, pp. 1193–1196. URL: `http://dx.doi.org/10.1126/science.1203801`.

Mira, Portia M et al. (2015). "Rational design of antibiotic treatment plans: a treatment strategy for managing evolution and reversing resistance". en. In: *PLoS One* 10.5, e0122283. URL: `http://dx.doi.org/10.1371/journal.pone.0122283`.

Nichol, Daniel et al. (2015). "Steering evolution with sequential therapy to prevent the emergence of bacterial antibiotic resistance". In: *PLoS Comput Biol* 11.9, e1004493.

Ogbunugafor, C Brandon and Daniel Hartl (2016). "A pivot mutation impedes reverse evolution across an adaptive landscape for drug resistance in Plasmodium vivax". en. In: *Malar. J.* 15, p. 40. URL: `http://dx.doi.org/10.1186/s12936-016-1090-3`.

Ogbunugafor, C Brandon, C Scott Wylie, et al. (2016). "Adaptive landscape by environment interactions dictate evolutionary dynamics in models of drug resistance". In: *PLoS Comput Biol* 12.1, e1004710.

Orr, H Allen (2009). "Fitness and its role in evolutionary genetics". en. In: *Nat. Rev. Genet.* 10.8, pp. 531–539. URL: `http://dx.doi.org/10.1038/nrg2603`.

Poelwijk, Frank J, Daniel J Kiviet, et al. (2007). "Empirical fitness landscapes reveal accessible evolutionary paths". en. In: *Nature* 445.7126, pp. 383–386. URL: `http://dx.doi.org/10.1038/nature05451`.

Poelwijk, Frank J, Sorin Tănase-Nicola, et al. (2011). "Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes". In: *Journal of theoretical biology* 272.1, pp. 141–144.

Poon, Art F Y and Lin Chao (2006). "Functional origins of fitness effect-sizes of compensatory mutations in the DNA bacteriophage

ϕX174". In: *Evolution* 60.10, pp. 2032–2043. URL: http://dx.doi.org/10.1554/06-013.1.

Sailer, Zachary R and Michael J Harms (2017). "Detecting High-Order Epistasis in Nonlinear Genotype-Phenotype Maps". In: *Genetics* 205.3, pp. 1079–1088.

Sanjuán, Rafael (2010). "Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies". en. In: *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365.1548, pp. 1975–1982. URL: http://dx.doi.org/10.1098/rstb.2010.0063.

Sanjuán, Rafael, José M Cuevas, et al. (2005). "Epistasis and the adaptability of an RNA virus". en. In: *Genetics* 170.3, pp. 1001–1008. URL: http://dx.doi.org/10.1534/genetics.105.040741.

Sanjuán, Rafael, Andrés Moya, and Santiago F Elena (2004). "The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 101.22, pp. 8396–8401. URL: http://dx.doi.org/10.1073/pnas.0400146101.

Segal, Mark R, Jason D Barbour, and Robert M Grant (2004). "Relating HIV-1 sequence variation to replication capacity via trees and forests". en. In: *Stat. Appl. Genet. Mol. Biol.* 3, Article2, discussion article 7, article 9. URL: http://dx.doi.org/10.2202/1544-6115.1031.

Seifert, David et al. (2015). "A Framework for Inferring Fitness Landscapes of Patient-Derived Viruses Using Quasispecies Theory". In: *Genetics* 199.1, pp. 191–203. URL: http://www.genetics.org/content/199/1/191.abstract.

Silva, Jack da et al. (2010). "Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region". en. In: *Genetics* 185.1, pp. 293–303. URL: http://dx.doi.org/10.1534/genetics.109.112458.

Smith, Maria and Sarah Cobey (2016). "Fitness landscapes reveal simple strategies for steering evolution to minimize antibiotic resistance". In: *bioRxiv*, p. 093153.

Stanley, Richard P (1999). *Enumerative Combinatorics*. Vol. 2. Cambridge: Cambridge University Press. URL: http://ebooks.cambridge.org/ref/id/CBO9780511609589.

The On-Line Encyclopedia of Integer Sequences (2016). https://oeis.org/. Accessed: 2016-8-19. URL: https://oeis.org/.

Weinreich, Daniel M, Nigel F Delaney, et al. (2006). "Darwinian evolution can follow only very few mutational paths to fitter proteins". en. In: *Science* 312.5770, pp. 111–114. URL: http://dx.doi.org/10.1126/science.1123539.

Weinreich, Daniel M, Yinghong Lan, et al. (2013). "Should evolutionary geneticists worry about higher-order epistasis?" en. In: *Curr. Opin. Genet. Dev.* 23.6, pp. 700–707. URL: http://dx.doi.org/10.1016/j.gde.2013.10.007.

Weinreich, Daniel M, Richard A Watson, and Lin Chao (2005). "Perspective: Sign epistasis and genetic constraint on evolutionary trajectories". en. In: *Evolution* 59.6, pp. 1165–1174. URL: https://www.ncbi.nlm.nih.gov/pubmed/16050094.

Wright, S (1932). "The roles of mutation, inbreeding, crossbreeding, and selection in evolution". In: *Proc 6th Int. Cong. Genet.* 1, pp. 356–366. URL: http://www.esp.org/books/6th-congress/facsimile/contents/6th-cong-p356-wright.pdf.

Wu, Nicholas C et al. (2016). "Adaptation in protein fitness landscapes is facilitated by indirect paths". URL: http://biorxiv.org/content/early/2016/03/22/045096.

Wylie, C Scott and Eugene I Shakhnovich (2011). "A biophysical protein folding model accounts for most mutational fitness effects in viruses". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 108.24, pp. 9916–9921. URL: http://dx.doi.org/10.1073/pnas.1017572108.

Supplementary Information

A. Mathematical framework and proofs

Here we provide proofs for the results in the main text, and give a brief background on the discrete Fourier transform, Dyck words, and Catalan numbers. We start with rank orders and the total n-way epistasis, followed by more general results on rank orders, circuits and other linear forms. The next topic is epistasis and fitness graphs, including some related graph theory. Finally we provide a few observations on epistasis and partial orders.

Gene interactions for a biallelic n-locus system can be described in terms of the Fourier transform for $(\mathbb{Z}_2)^n$ defined as

$$u_{i_1 i_2 \ldots i_n} = \frac{1}{2^{n-1}} \cdot \sum_{j_1=0}^{1} \sum_{j_2=0}^{1} \cdots \sum_{j_n=0}^{1} (-1)^{i_1 j_1 + i_2 j_2 + \cdots + i_n j_n} \, w_{j_1 j_2 \ldots j_n}.$$

The interaction coordinate $u_{1\ldots1}$ measures the total n-way epistasis. By abuse of notation we will ignore the scaling factor $\frac{1}{2^{n-1}}$, so that

$$u_{1\ldots1} = \sum_{j_1=0}^{1} \sum_{j_2=0}^{1} \cdots \sum_{j_n=0}^{1} (-1)^{j_1 + j_2 + \cdots + j_n} \, w_{j_1 j_2 \ldots j_n}.$$

In particular,

$$u_{11} = w_{00} - w_{10} - w_{01} + w_{11} = \varepsilon_2$$

and

$$u_{111} = w_{000} - w_{100} - w_{010} - w_{001} + w_{110} + w_{101} + w_{011} - w_{111} = \varepsilon_3$$

as defined in Eqs. 1 and 2 in the main text. A biallelic three-locus system has three-way epistasis exactly if $u_{111} \neq 0$. Otherwise the system has only pairwise interactions. Similarly, a biallelic n-locus system has (total) n-way epistasis exactly if $u_{1\ldots1} \neq 0$.

A.1. **Rank orders.** We will determine the number of rank orders which imply n-way epistasis. The proof depends on Catalan numbers and Dyck words (Stanley 1999). Let C_i denote the ith Catalan number for $i \geq 0$, that is, $C_i = \frac{(2i)!}{(i+1)!i!}$. In particular, $C_0 = C_1 = 1, C_2 = 5$,

$C_3 = 14$ and $C_4 = 42$. A Dyck word of length $2n$ in the letters X and Y is a string consisting of n X's and n Y's such that no initial segment of the string has more Y's than X's. For instance, the Dyck words of length 4 are $XXYY$ and $XYXY$. The initial segments of $XXYY$ are X, XX, XXY, and $XXYY$.

Proposition 1. *Consider a biallelic n-locus system. The number of rank orders which imply n-way epistasis is*

$$\frac{(2^n)! \times 2}{2^{n-1} + 1}$$

Proof. There are $(2^n)!$ rank orders in total. Let e_i denote the fitnesses of genotypes with an even number of 1's in the subscripts ($w_{0\ldots0}$, $w_{110\ldots0}$, and so forth) and o_i the fitnesses of genotypes with an odd number of 1's, ordered in such a way that $e_i > e_{i+1}$ and $o_i > o_{i+1}$ for all i. We will refer to even and odd elements from now on. Let $u_{1\ldots1}$ denote the interaction coordinate as defined above. Notice that $u_{1\ldots1} = 0$ exactly if $\sum_i e_i - \sum_i o_i = 0$. Consequently a rank order implies positive n-way epistasis ($u_{1\ldots1} > 0$) when the sum $\sum_i (e_i - o_i)$ is positive for all fitness landscapes compatible with the rank order. It is therefore sufficient to count such rank orders.

We define a map from fitness rank orders to words in the alphabet $\{e, o\}$ as follows: $e_i \mapsto e, o_i \mapsto o$. For instance, the order $w_{00} > w_{11} > w_{10} > w_{01}$ is mapped to *eeoo*. We claim that a rank order satisfies $\sum_i (e_i - o_i) > 0$ exactly when it is mapped to a Dyck word (where e precedes o).

It is immediate that $\sum_i (e_i - o_i) > 0$ holds if the rank order is mapped to a Dyck word. Conversely, suppose that a rank order is not mapped to a Dyck word. Let s be the least number such that the number of o's exceeds the number of e's for an initial segment of length s (note that s has to be odd in this case) and let $j = \frac{s+1}{2}$. Clearly one can make the sum of $\sum_{i=1}^{j} (e_i - o_i)$ negative for a particular choice of e_i and o_i. By choosing the remaining numbers e_i, o_i sufficiently small, we get $\sum_i (e_i - o_i) < 0$, which proves the claim.

It remains to count the rank orders where $\sum_i (e_i - o_i) > 0$. Such rank orders are mapped to Dyck words (where e precedes o) consisting of 2^{n-1} e's and 2^{n-1} o's. There are $C_{2^{n-1}}$ such Dyck words (Stanley 1999). For each word there are $(2^{n-1})! \times (2^{n-1})!$ fitness rank orders which map to the word. Indeed, one can choose the ordering of even and odd elements each in $(2^{n-1})!$ different ways.

In total there are $C_{2^{n-1}} \times (2^{n-1})! \times (2^{n-1})!$ fitness rank orders such that $\sum_i (e_i - o_i) > 0$ for all landscapes. By symmetry, the same number of fitness rank orders satisfy the negative epistasis condition $\sum_i (e_i - o_i) < 0$. One verifies that

$$C_{2^{n-1}} \times (2^{n-1})! \times (2^{n-1})! \times 2 = \frac{(2^n)! \times 2}{(2^{n-1}+1)},$$

which completes the proof. □

A few observations in the proof of Proposition 1 are of interest. Importantly, the proof gives a computationally efficient method (linear in the number of genotypes) for checking if a rank order implies n-way epistasis. Indeed, the rank order implies higher order epistasis exactly if it is mapped to a Dyck word. Moreover, the proposition states that $\frac{(2^n)! \times 2}{(2^{n-1}+1)}$ orders imply n-way epistasis. From the proof it is clear that half of these orders imply positive n-way epistasis ($u_{1\ldots1} > 0$) and the other half negative n-way epistasis ($u_{1\ldots1} < 0$). Also, the proof points out some symmetries. If a rank order implies epistasis, then the same is true for rank orders obtained by (i) any permutation of the even elements, (ii) any permutation of the odd elements, and (iii) the flip obtained by replacing every "<" by ">" in the rank order. It follows that each rank order that implies three-way epistasis belongs to a class of 1152 ($= 4! \cdot 4! \cdot 2$) elements, which differ by the operations (i)–(iii) only.

Corollary 2. *The fraction of rank orders that imply n-way epistasis among all rank orders is*

$$\frac{2}{2^{n-1}+1}.$$

Proof. Since the number of all rank orders is $(2^n)!$, the result follows. □

The results on rank orders and epistasis for $2 \leq n \leq 4$ are summarized in Table 1.

Loci	Rank orders	Imply epistasis	Fraction
2	24	16	2/3
3	40,320	16,128	2/5
4	20,922,789,890,000	4,649,508,864,000	2/9

TABLE 1. Numbers and fractions of rank orders that imply n-way epistasis.

Interestingly, no integer sequence that starts with 16, 16 128, ... is available at *The On-Line Encyclopedia of Integer Sequences* (2016).

A.2. **Circuits.** The proof of Proposition 1 depends on the map defined from the rank orders to words in the alphabet $\{e, o\}$. We will use a generalization of the map in subsequent proofs. The starting point is a given linear form. The form determines a map from rank orders to words. Although the idea is closely related to the previous proof, we will work with positive and negative coefficients in the linear forms. For that reason, we will use P and N rather than e and o (even and odd is no longer meaningful).

We start with a clarifying example. Assume that a given linear form has integer coefficients and that the sum of its coefficients is zero. For instance, the form

$$m = w_{001} + w_{010} + w_{100} - w_{111} - 2w_{000}$$

defines a map φ^m as follows: Each of the variables $w_{001}, w_{010}, w_{100}$ corresponds to the letter P (for positive), and the variable w_{111} corresponds to N (for negative). The variable w_{000} corresponds to NN, because of the coefficient -2. In this case, the rank order

$$w_{111} > w_{001} > w_{000} > w_{100} > w_{010} > w_{110} > w_{101} > w_{011}$$

is mapped to $NPNNPP$ under φ^m. Specifically, starting from left w_{111} corresponds to N, w_{001} to P, w_{000} to NN, w_{100} to P, and w_{100} to P. The remaining variables $w_{110}, w_{101}, w_{011}$ do not impact the word, since their coefficients are zero for the form m.

Definition 1. Let f be a linear form with integer coefficients. Assume that the sum of its coefficients is zero. Let φ^f denote the map from a total order on the variables (a rank order) to words in the alphabet $\{P, N\}$ defined as follows: Each variable of f with a positive integer coefficient c corresponds to a substring of c letters P. Each variable in f with a negative integer coefficient c' corresponds to a substring of c' letters N. A rank order of the variables is mapped to the word consisting of the substrings obtained for each variable with non-zero coefficient in f. Specifically, the substrings (from left to right) of the word correspond to the variables in the rank order (from highest to lowest).

The proof of Proposition 1 uses e and o instead of P and N. However, notice that the map from rank orders to words in e's and o's is exactly $\varphi^{u_{1\cdots 1}}$ (modulo the labeling). The next result is a generalization of Proposition 1. The proof is similar in every step with the modification that $\varphi^{u_{1\cdots 1}}$ is replaced by φ^f for an arbitrary form f, so we omit the details.

Theorem 3. *Let f be a linear form with integer coefficients. Assume that the sum of its coefficients is zero. Then a rank order implies that f is not zero if and only if it is mapped to a Dyck word by φ^f.*

The theorem can be applied in different contexts, and we start with circuits for biallelic three-locus systems. Recall from the main text that $a = w_{000} - w_{010} - w_{100} + w_{110}$ is a circuit. Notice that $a = 0$ for all additive fitness landscapes. Moreover, a is minimal with this property, in the sense that no linear form in a proper subset of $\{w_{000}, w_{010}, w_{100}, w_{110}\}$ equals zero for all additive landscapes. The circuits for a system of genotypes are defined as such minimal dependence

relations. For a complete description of the gene interactions one can analyze the circuits.

There are 20 circuits $a, \ldots, t$ for three-locus systems (Beerenwinkel, Pachter, and Sturmfels 2007), namely

$$
\begin{aligned}
a &:= w_{000} - w_{010} - w_{100} + w_{110} \\
b &:= w_{001} - w_{011} - w_{101} + w_{111} \\
c &:= w_{000} - w_{001} - w_{100} + w_{101} \\
d &:= w_{010} - w_{011} - w_{110} + w_{111} \\
e &:= w_{000} - w_{001} - w_{010} + w_{011} \\
f &:= w_{100} - w_{101} - w_{110} + w_{111} \\
g &:= w_{000} - w_{011} - w_{100} + w_{111} \\
h &:= w_{001} - w_{010} - w_{101} + w_{110} \\
i &:= w_{000} - w_{010} - w_{101} + w_{111} \\
j &:= w_{001} - w_{011} - w_{100} + w_{110} \\
k &:= w_{000} - w_{001} - w_{110} + w_{111} \\
l &:= w_{010} - w_{011} - w_{100} + w_{101} \\
m &:= w_{001} + w_{010} + w_{100} - w_{111} - 2w_{000} \\
n &:= w_{011} + w_{101} + w_{110} - w_{000} - 2w_{111} \\
o &:= w_{010} + w_{100} + w_{111} - w_{001} - 2w_{110} \\
p &:= w_{000} + w_{011} + w_{101} - w_{110} - 2w_{001} \\
q &:= w_{001} + w_{100} + w_{111} - w_{010} - 2w_{101} \\
r &:= w_{000} + w_{011} + w_{110} - w_{101} - 2w_{010} \\
s &:= w_{000} + w_{101} + w_{110} - w_{011} - 2w_{100} \\
t &:= w_{001} + w_{010} + w_{111} - w_{100} - 2w_{011}
\end{aligned}
$$

The circuits $a, \ldots, f$ measure conditional two-way epistasis between two loci when the allele at the third locus is fixed, the circuits $g, \ldots, l$ relate marginal epistasis of two pairs of loci, and the circuits $m, \ldots, t$

relate the three-way interaction to the total two-way epistasis. (Beerenwinkel, Pachter, and Sturmfels 2007).

For a given circuit, some rank orders imply that the circuit is positive, i.e., the circuit is positive for all fitness values compatible with the rank order. Similarly, some rank orders imply that the circuit is negative, whereas the sign cannot be determined from other rank orders. We will use Theorem 3 to check whether a rank order determines the sign of a circuit or not.

Corollary 4. *For the circuits $a, \dots l$, two thirds of all possible rank orders determine the sign of the circuit. For the circuits $m, \dots, t$, one half of all possible rank orders determine the sign of the circuit.*

Proof. Fix one of the circuits from a to l and a rank order. The circuit has exactly four variables with non-zero coefficients (for instance, for the circuit a the variables are $w_{000}, w_{100}, w_{010}, w_{110}$, so that φ^a maps rank orders to four-letter words). By Theorem 3, the rank order implies that the circuit differs from zero when it is mapped to one of the Dyck words $PPNN$, $PNPN$, $NNPP$ or $NPNP$ under φ, whereas the sign of the circuit is not determined when the word is $PNNP$ or $NPPN$. One concludes that the sign of a given circuit from a to l is determined for $2/3$ of the rank orders.

Using a similar argument, we consider words of length 6 for the circuits labeled m to t. There are in total 20 words consisting of 3 P's and 3 N's. Ten of them are Dyck words. We conclude that the sign of a given circuit from m to t is determined for $1/2$ of the rank orders. $\square$

The gene interactions for a biallelic three-locus system can be classified in terms of shapes of the fitness landscape, or triangulations of the 3-cube (Beerenwinkel, Pachter, and Sturmfels 2007). There are 74 shapes for the 3-cube. The shape of the fitness landscapes is determined by the signs of the 20 circuits. It follows that rank orders provide some information about possible shapes. However, the following result shows that rank orders do not determine shapes.

Proposition 5. *Consider a three-locus biallelic system. No rank order determines the shape of a fitness landscape.*

Proof. The result follows from the characterization of shapes for the 3-cube in (Beerenwinkel, Pachter, and Sturmfels 2007), where each shape is described in terms of a circuit sign pattern. We verified computationally that no rank order implies that all the circuits have the signs which describe a particular shape (`https://github.com/gavruskin/fitlands#analysis-of-rank-orders`) . More precisely, for every circuit $a, \ldots, t$, we determined the set of all rank orders that imply that the circuit is positive or negative.

For every rank order, we then considered the circuit signs determined by the order. In no case did a rank order determine all the circuit signs necessary for describing a particular shape. □

The fact that rank orders do not determine the shape of a fitness landscape over a three-locus system is expected. Shapes reflect interactions in a very fine scaled way, whereas rank orders provide only coarse information.

In addition to u_{111}, there exist three other interaction coordinates u_{110}, u_{101}, u_{011}. These coordinates correspond to two-way epistasis marginalized over the loci marked by 0 (Beerenwinkel, Pachter, and Sturmfels 2007). We again apply Theorem 3 to identify rank orders that determine the sign of the interaction coordinates.

Corollary 6. *For each of the interaction coordinates u_{110}, u_{101}, and u_{011}, the number of rank orders which imply three-way epistasis is 16,128.*

Proof. The linear form for each interaction coordinate consists of 8 elements, 4 with positive signs and 4 with negative signs. Notice that $16,128$ rank orders imply three-way epistasis, by Proposition 1. By Theorem 3 the problem can be reduced to counting Dyck words of length 8. It follows that the number of rank orders is $16,128$ for each interaction coordinate. □

The approach to epistasis given in (Weinreich, Lan, et al. 2013) depends on Walsh coefficients. Notice that interaction coordinates and Walsh coefficients are closely related in this case. Specifically, the Walsh coefficients E_{110}, E_{101}, E_{011} differ from the interaction coordinates u_{110}, u_{101}, u_{011} only by a scalar, so that Corollary 6 applies to the coefficients as well.

A.3. **Partial orders and fitness graphs.** We now consider partial orders, for instance,

$$w_{111} > w_{110}, w_{100}, w_{010}, w_{001} > w_{000} > w_{101}, w_{011}$$

for a three-locus system. Arguing as in the proof of Proposition 1, the (unknown) total order is mapped to the word $oxxxxeee$ under $\varphi^{u_{1\cdots1}}$, where $xxxx$ is some permutation of $eooo$. For any such permutation we get a Dyck word. It follows that the system has three-way epistasis. This condition can be stated and proved in a more general form.

Proposition 7. *Consider an n-locus biallelic system. Let e_i and o_i be defined as in the proof of Proposition 1. If there exists a partition of the total set of fitness values into pairs (e_i, o_i), where $e_i > o_i$ for all i, then one can conclude n-way epistasis. By symmetry, the same is true for a partition where $e_i < o_i$ for each pair.*

Proof. We will verify that the existence of a partition as described is equivalent to the order being mapped to a Dyck word under the map $\varphi^{u_{1\cdots1}}$. It is immediate that the existence of such a partition implies that the order is mapped to a Dyck word. Conversely, if the rank order is mapped to a Dyck word under $\varphi^{u_{1\cdots1}}$, then one can construct a partition as follows. One pair in the partition corresponds to the first e and the first o in the Dyck word, a second pair corresponds to the second e and the second o in the word, and so forth. This partition has the desired property. □

A fitness graph is a directed acyclic graph where each node represents a genotype, and arrows connect each pair of mutational neighbors, directed toward the node representing the genotype of higher fitness.

Moreover, fitness graphs are structured so that the node labeled $0\ldots0$ is at the bottom, genotypes with exactly one 1 on the level above, and so forth (see Figure 10).

Some fitness graphs imply epistasis, whereas other fitness graphs are compatible with additive fitness. As illustrated in the two-locus case, a fitness graph is compatible with additive fitness if all arrows point up, that is toward a higher level. More generally, a fitness graph implies epistasis unless it is isomorphic to a graph with all arrows up (Crona, Greene, and Barlow 2013).

Similarly, we can easily characterize rank orders that are compatible with additive fitness. After relabeling genotypes, we can assume that the genotype $0\ldots0$ has the lowest fitness in the system. Then a rank order is compatible with additive fitness, exactly if for each genotype replacing 0 by 1 results in higher fitness. Otherwise the rank order implies epistasis. For instance, $w_{00} < w_{10} < w_{01} < w_{11}$ is compatible with additive fitness, whereas the rank order $w_{00} < w_{11} < w_{01} < w_{10}$ implies epistasis. The claims follows immediately from the result stated above for fitness graphs.

As we have already seen, rank orders have potential far beyond detecting whether or not there is epistasis in a system. The same is true for fitness graphs, and we proceed with higher order interactions. In order to analyze fitness graphs and three-way epistasis, we consider the set of rank orders compatible with a given fitness graphs. For instance, the fitness graph in Figure 7 is compatible with the following two rank orders:

$$w_{111} > w_{000} > w_{100} > w_{010} > w_{001} > w_{110} > w_{101} > w_{011}, \tag{5}$$

$$w_{000} > w_{111} > w_{100} > w_{010} > w_{001} > w_{110} > w_{101} > w_{011}. \tag{6}$$

The first order implies three-way epistasis (it is mapped to *oeoooeee* under $\varphi^{u_1,\ldots,1}$ and the second does not (it is mapped to *eooooeee* under $\varphi^{u_1,\ldots,1}$). We conclude that in this case, the fitness graph does not imply higher order epistasis. However, if every rank order compatible with the fitness graph implies higher order epistasis, then the fitness

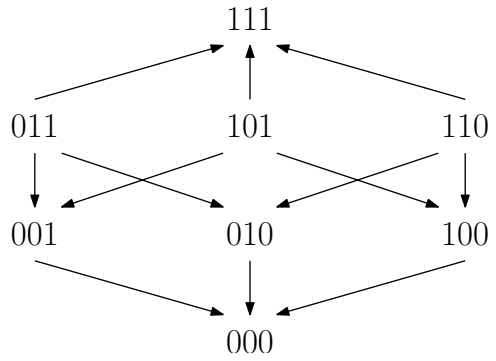


FIGURE 7. The fitness graph is compatible with the two rank orders (5) and (6).

graph itself does imply higher order epistasis. More generally, the same observation holds for any partial order.

Remark. A partial order implies higher order epistasis exactly if all its total extensions imply higher order epistasis.

Indeed, if all total extensions imply higher order epistasis, then in particular the (unknown) rank order does. The converse holds by definition.

Consequently, one can in principle determine if a fitness graph implies higher order epistasis by checking all of the compatible rank orders. For a systematic study of the three-locus case, it is convenient to reduce the problem to isomorphic graphs. In biological terms, isomorphic graphs differ by the labeling of the genotypes only (a relabeling needs to respect the adjacency structure so that mutational neighbors stay neighbors). Figure 8 shows 4 isomorphic fitness graphs.

As mentioned in the main text, there are in total $1,862$ fitness graphs for three-locus systems, and the number of fitness graphs that imply higher order epistasis is 698 (37 percent). Up to isomorphism there are in total 54 fitness graphs, and 20 graphs imply higher order epistasis. This result was verified by reducing the study of all 1862 graphs for three-locus systems, to a non-redundant list of 54 graphs, i.e., no two graphs in the list are isomorphic.

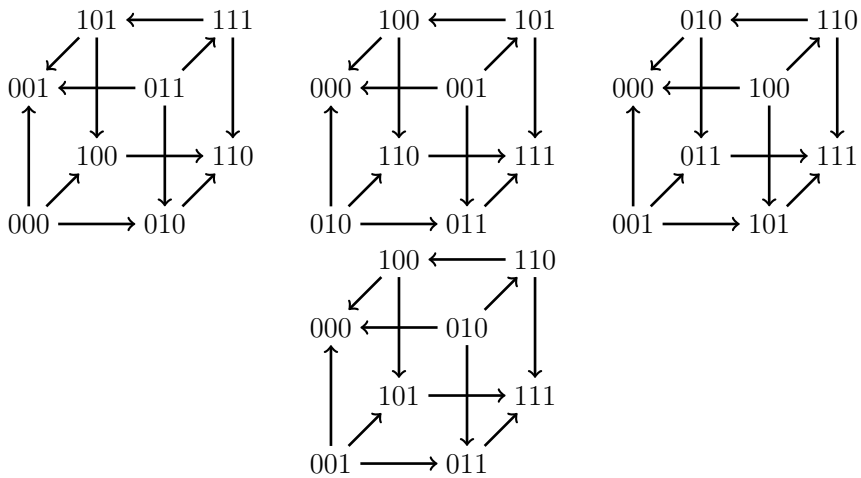


FIGURE 8. The four graphs are isomorphic (by abuse of notation, we do not follow the level convention for the fitness graphs). In biological terms, the fitness graphs differ only by the labeling of the genotypes. For instance, the second graph differs from the first by a 90 degree clockwise rotation of the label system.

Isomorphisms between three-locus systems have a geometric interpretion. The fitness graph can be regarded as a three-dimensional cube, with vertices corresponding to genotypes and edges corresponding to arrows. The group of isomorphisms then corresponds to the symmetry group of the three dimensional cube (Coxeter 1973). Indeed, it was by way of this equivalence that we carried out the enumeration described above.

A.4. **Graph theoretical aspects.** As mentioned in the main text, the 20 graphs which imply higher order epistasis (see Supplementary Figure 10) constitute a diverse category. We analyzed the category from a graph theoretical point of view, but could not see that the graphs have any property which singles them out.

Recall that a unique sink orientation is a graph where each face has no more than one sink. Equivalently, there is no subsystems with reciprocal sign epistasis (Crona, Greene, and Barlow 2013; Poelwijk, Kiviet, et al. 2007). The category of 20 graphs includes unique sink orientations (also called USO or AOF graphs), as well as non-USO's.

Moreover, in the terminology of Gärtner and Kaibel (1998), the category includes separable and non-separable graphs, as well as realizable and non-realizable graphs.

There were some indications of higher complexity for the category, but only in a statistical sense. Indeed, as can be verified from Figure 10, the graphs in the category have on average 1.8 sinks (a sink corresponds to a peak in the landscape), whereas the average number of sinks for all graphs is 1.6. Moreover, 5 out of the 20 graphs (25 percent) in the category are unique sink orientations, whereas in total 19 out of the 54 graphs (35 percent) are unique sink orientations.

Even though the category of fitness graphs which implies epistasis is diverse, it is still possible that a characterization exists. This is an open problem.

B. Software and HIV-1 study

In this section, we provide the details of our analysis of the HIV-1 data in Section 2, after a brief discussion of code related to this paper. We have implemented algorithms based on our theoretical results in an open source software package (`https://github.com/gavruskin/fitlands#fitlands`). The package provides algorithms for detecting gene interactions as described in the main text for two- and three-locus systems. In particular, algorithms for detecting total three-way epistasis, interaction coordinates and circuits have been implemented. The documentation also explains how to reproduce results for our application to HIV-1 data described in the main text.

For the HIV-1 study, we had the opportunity to compare our rank order based approach with conventional methods. To compare our rank order methods with conventional Wilcoxon rank sum tests, we used the replicate fitness measurements provided in the original study (Segal, Barbour, and Grant 2004). We computed the values of interaction coordinates and circuit interactions for the summary statistics reported in (Beerenwinkel, Pachter, and Sturmfels 2007). Figure 9 shows that our rank order methods detected the majority of circuit interactions identified by using the summary statistics.

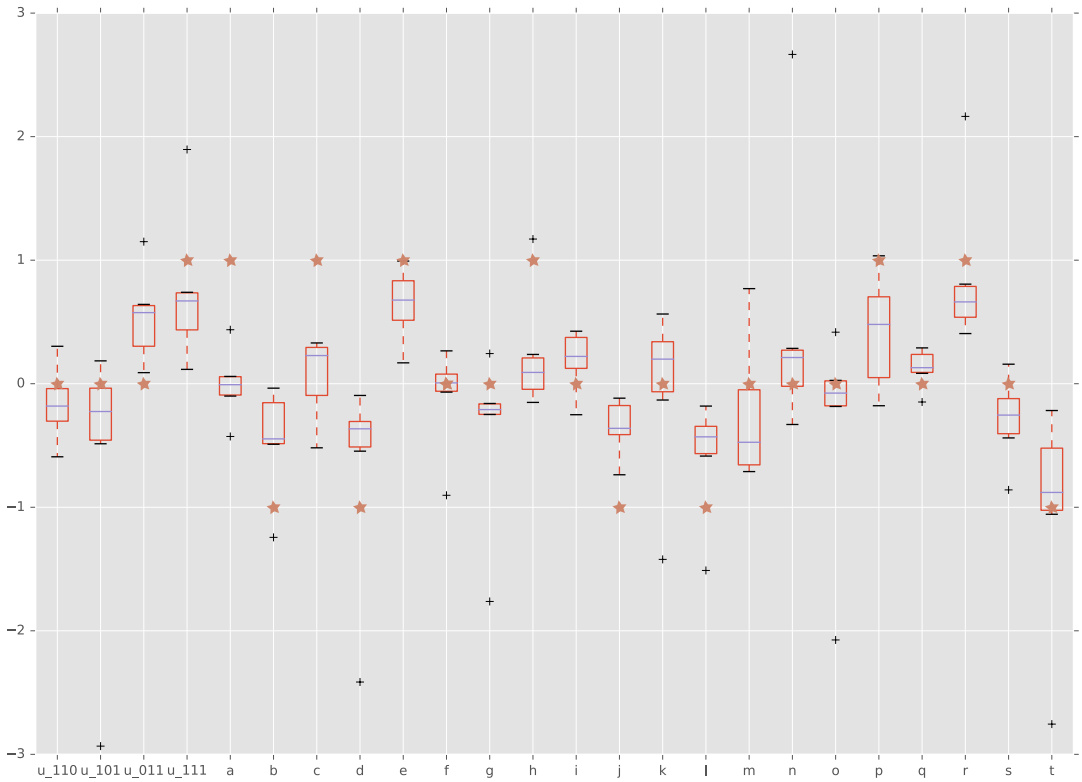


FIGURE 9. Interactions detected from fitness summary statistics and from rank orders. The horizontal axis is labeled by the four interaction coordinates $u_{110}, \ldots, u_{111}$ and twenty circuits $a, \ldots, t$. The boxplots show the distributions given by the fintess values and the stars indicate whether or not the interaction has been detected using our rank order methods. The star with Y-coordinate 1 means positive interaction, with Y-coordinate -1 negative interaction, and with Y-coordinate 0 means the rank order does not imply interaction.

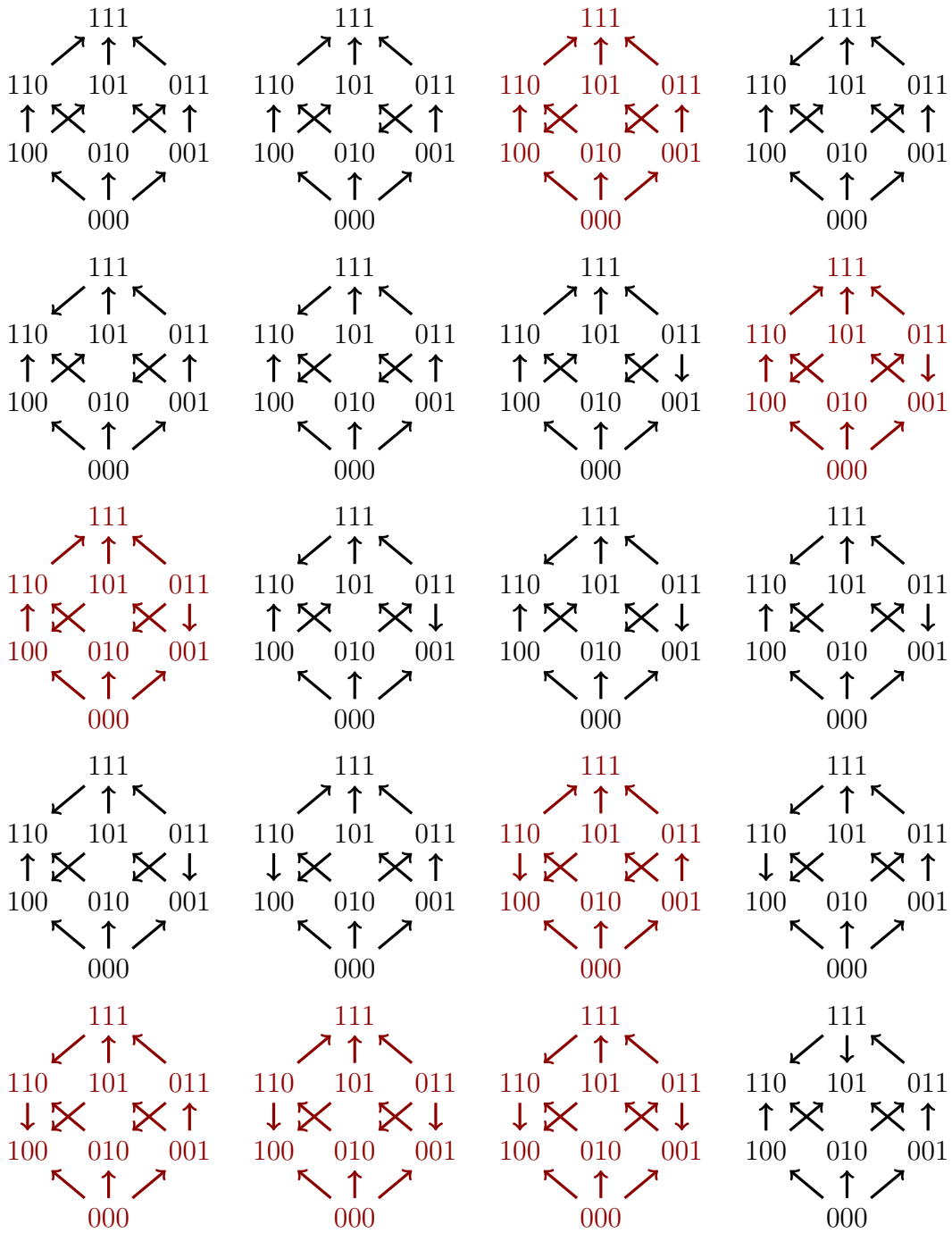
Second, we applied Student's t-test to detect interactions and quantify the significance of the estimates. Table 2 summarizes the results,

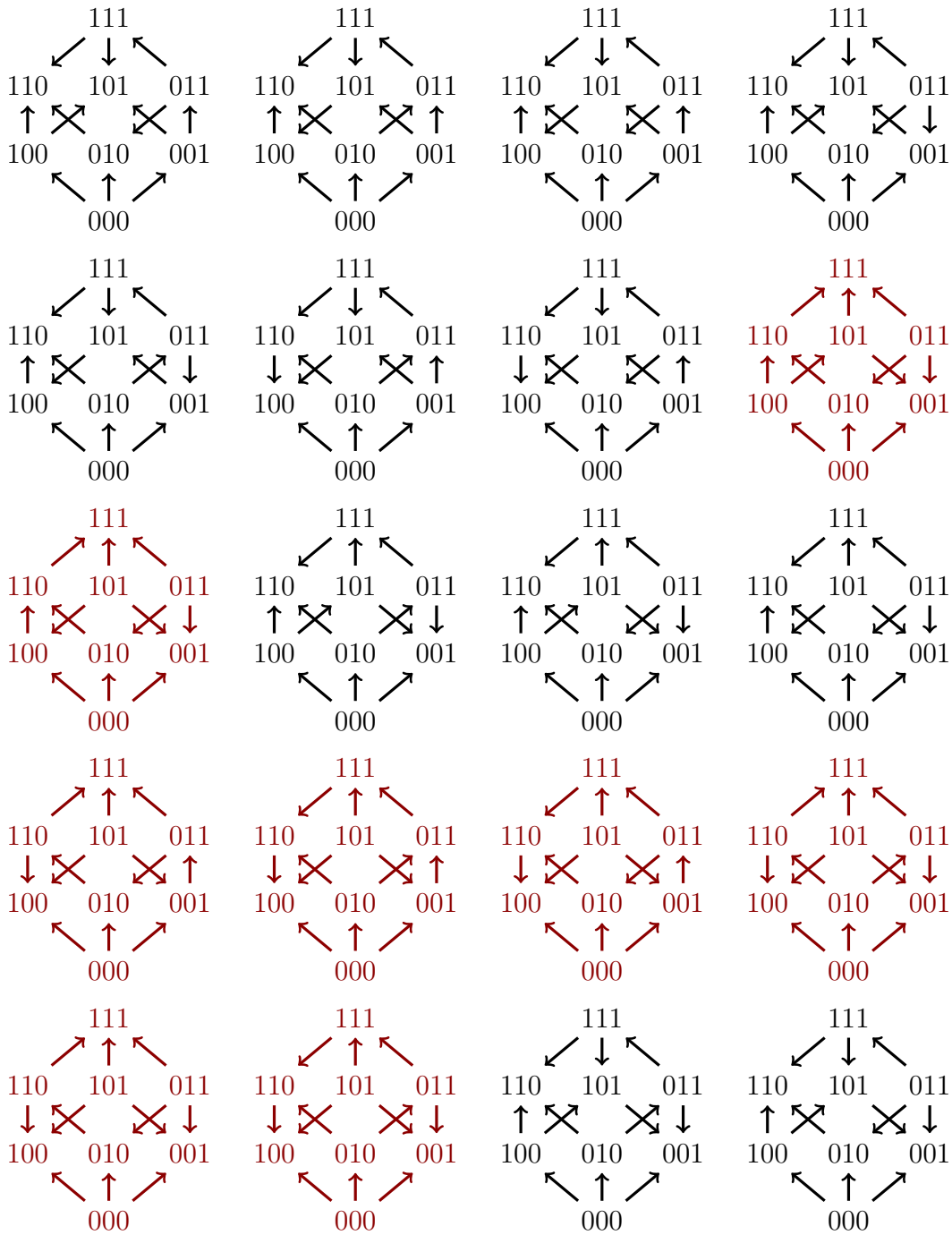
and for related code see `https://github.com/gavruskin/fitlands/blob/master/HIV_2007_conventional_analysis.ipynb`

Interaction	p-value	Result	From rank order
u_{011}	1.13e-31	+	0
u_{101}	2.67e-12	−	0
u_{110}	1.20e-24	−	0
u_{111}	1.50e-29	+	+
a	7.10e-16	−	+
b	5.23e-32	−	−
c	7.62e-04	+	+
d	8.36e-68	−	−
e	1.39e-38	+	+
f	2.59e-01	0	0
g	3.10e-59	−	0
h	2.22e-02	−	+
i	7.97e-05	+	0
j	2.20e-32	−	−
k	1.96e-05	+	0
l	7.50e-51	−	−
m	4.88e-07	−	0
n	9.87e-37	+	0
o	8.83e-03	+	0
p	7.18e-19	+	+
q	1.94e-01	0	0
r	5.02e-50	+	+
s	7.10e-27	−	0
t	8.49e-61	−	−

TABLE 2. The first column lists the four interaction coordinates and twenty circuits. The second column shows p-values returened by Student's t-test based on fitness measurements. The third column shows which interactions are significant based on 0.03 threshold and their signs. For comparison, the last column displays the results obtained from rank order methods.

C. Supplementary figures




111
110 101 011
100 010 001
000

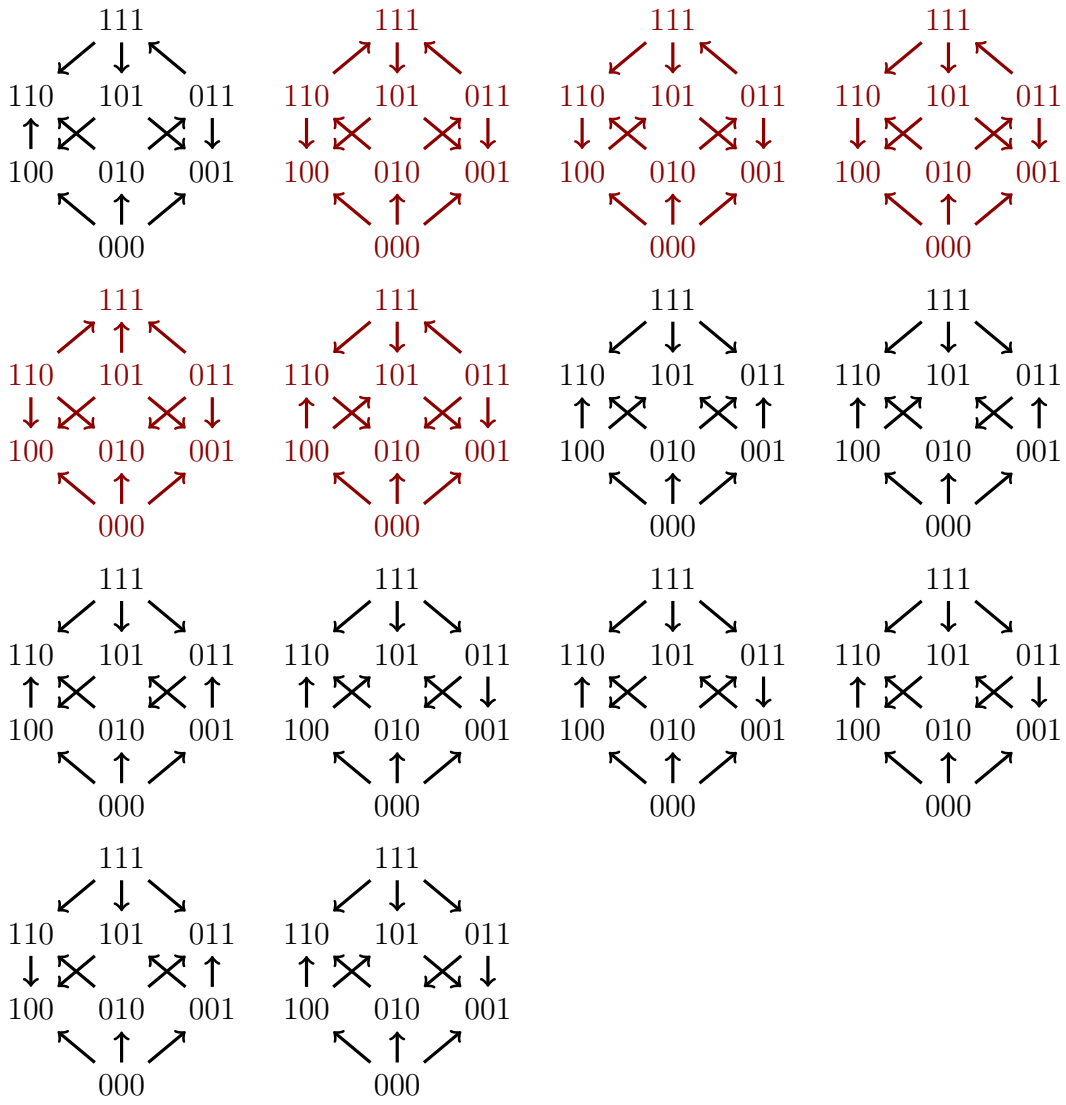


FIGURE 10. All 54 fitness graph types. Those depicted in red imply three-way epistasis.