Functional bottlenecks can emerge from non-epistatic underlying traits

Anna Ottavia Schulte,[1] Samar Alqatari,[2] Saverio Rossi,[1] and Francesco Zamponi[1]

[1] *Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy*
[2] *Department of Physics and The James Franck and Enrico Fermi Institutes, The University of Chicago, Chicago, IL 60637, USA*

Protein fitness landscapes frequently exhibit epistasis, where the effect of a mutation depends on the genetic context in which it occurs, *i.e.*, the rest of the protein sequence. Epistasis increases landscape complexity, often resulting in multiple fitness peaks. In its simplest form, known as global epistasis, fitness is modeled as a non-linear function of an underlying additive trait. In contrast, more complex epistasis arises from a network of (pairwise or many-body) interactions between residues, which cannot be removed by a single non-linear transformation. Recent studies have explored how global and network epistasis contribute to the emergence of functional bottlenecks - fitness landscape topologies where two broad high-fitness basins, representing distinct phenotypes, are separated by a bottleneck that can only be crossed via one or a few mutational paths. Here, we introduce and analyze a simple model of global epistasis with an additive underlying trait. We demonstrate that functional bottlenecks arise with high probability if the model is properly calibrated. Our results underscore the necessity of sufficient heterogeneity in the mutational effects selected by evolution for the emergence of functional bottlenecks. Moreover, we show that the model agrees with experimental findings, at least in small enough combinatorial mutational spaces.

I. INTRODUCTION

Understanding how genetic variation translates into differences in reproductive success is a central challenge in evolutionary biology. An organism's fitness — defined as the number of offspring it produces — depends on how its phenotypic traits interact with the environment. These traits, in turn, are shaped by the organism's genotype, suggesting a hierarchical mapping from genotype to phenotype to fitness. By assigning a fitness value to each genotype, one obtains a *fitness landscape* [1–5], a conceptual tool that captures the topology of the space of evolutionary possibilities.

Epistasis is a distinctive property of fitness landscapes. Broadly speaking, it is defined as the context-dependence of mutational effects [6–38]. More precisely, it is defined as follows. Consider a genotype $\mathbf{a} = (a_1, a_2, \cdots, a_L)$ with an associated fitness $F(\mathbf{a})$. The genotype $\mathbf{a}$ could be a protein sequence of length L, with a_i representing the amino acid at site i, or a nucleotide sequence, or a sequence of zeros and ones representing the presence/absence of a given mutation, gene, etc. Epistasis occurs when the fitness change $\Delta F_i(a \to b)$ due to substituting $a_i = a$ with $a_i = b$ at site i depends on the amino acids that are present at other sites in the sequence. As such, epistasis can be of two conceptually distinct origins.

What is sometimes called 'global' epistasis is based on introducing the simplest possible non-linearity in the genotype-phenotype-fitness mapping [2, 3, 5, 19, 23, 24, 27, 31, 32, 36, 37, 39]. More precisely, one assumes the existence of an underlying additive phenotype (or 'trait') $E(\mathbf{a}) = \sum_{i=1}^{L} h_i(a_i)$ associated with each genotype $\mathbf{a}$. Under this assumption, the variation of phenotype $E(\mathbf{a})$ associated with substituting a with b at site i, $\Delta E_i(a \to b) = h_i(b) - h_i(a)$, is independent of the rest of the genotype. Next, one assumes that the fitness $F(\mathbf{a}) = \phi(E(\mathbf{a}))$ is a non-linear function of $E(\mathbf{a})$, for instance a sigmoid (Fig. 1a). Under these assumptions, while the variation of the underlying trait is independent of the rest of the sequence, the variation of fitness depends on it due to the non-linearity of ϕ. Yet, such non-linearity is relatively easy to handle, because one can deduce the function $\phi(E)$ by 'deconvolution' of experimental fitness measurements [23, 24, 27, 32, 37], and thus describe the fitness function with a limited number of parameters. In the example above, two parameters enter in the definition of $\phi(E)$, and of the order of L parameters enter in the definition of E. Furthermore, such a globally epistatic fitness landscape only features a single maximum, because $F = \phi(E)$ is an increasing function of E, and E itself has a single maximum, obtained by choosing at each site the value a_i that maximizes the additive contribution $h_i(a_i)$ [5]. Multi-peaked landscapes can be obtained by choosing a non-monotonic function $\phi(E)$, as in Fisher's Geometric Model [2, 3, 5, 19].

A more complex behavior is obtained in presence of what is sometimes called 'network' epistasis [8, 11, 12, 14, 15, 20, 22, 28–30, 33–35], which implies that no function $E = \phi^{-1}(F)$ can transform fitness into a purely additive function of the genotype [27, 32, 37]. While a function ϕ can be chosen to make E as close as possible to an additive form, higher-order interactions will still play a role in the presence of network epistasis. More formally, E can always be decomposed as a sum of additive, two-site, three-site, and higher-order interactions [10, 14, 17, 18, 21, 27, 28, 32, 37], i.e.,

$$E(\mathbf{a}) = \sum_i h_i(a_i) + \sum_{i,j} J_{ij}(a_i, a_j) + \sum_{i,j,k} K_{ijk}(a_i, a_j, a_k) + \cdots \tag{1}$$

however, due to network epistasis, no choice of ϕ can eliminate the higher-order terms J_{ij}, K_{ijk}, and beyond. Let us focus for simplicity on the case in which only the two-site interactions are present. The number of param-

eters required to fully specify the function E now scales as the number of pairs, i.e., proportionally to L^2, which makes inferring the fitness function from a few measurements significantly more difficult. Even more importantly, a function like (1) can have many local maxima, corresponding to many fitness peaks, leading to a rough, complex fitness landscape.

The relative importance of global and network epistasis has been a subject of study in recent years. A particular focus of the debate has been the role played by epistasis in shaping the topology of the fitness landscape in the proximity of functional switches. Poelwijk et al. [27] studied experimentally the fitness landscape separating two phenotypically distinct variants of the *Entacmaea quadricolor* fluorescent protein, one fluorescent in the red and the other in the blue, separated by 13 mutations. Assaying all 2^{13} intermediate variants, they observed, in addition to global epistasis, significant network contributions (both pairwise and higher order) to fitness. They also identified a functional bottleneck separating the 'blue' and 'red' space of fluorescent variants – a narrow region of accessible evolutionary paths between the two phenotypes – which they attributed to epistatic constraints. These functional bottlenecks illustrate how fitness landscapes can remain navigable despite their ruggedness. While epistasis can constrain viable evolutionary paths, it can also facilitate adaptation by carving out narrow high-fitness pathways that connect otherwise isolated functional states – typically via critical "switch" mutations [40–42].

In a more recent study, Alqatari and Nagel [38] used similar methods to investigate function switch in an elastic network, considered as a model for allosteric behavior in proteins. Analyzing ensembles of fitness landscapes, they observed different topologies at the threshold for viable evolution, and found that functional bottlenecks – often involving a single critical mutation – emerged generically, similar to the observations in the experimental study.

Here, we identify the minimal conditions under which such topological features emerge. Motivated by these works, we investigate whether network epistasis is a *necessary* condition for the existence of functional bottlenecks in fitness landscapes. We thus introduce a simple toy model of global epistasis, based on a random underlying additive trait and a non-linear fitness function. We investigate the random fitness landscapes generated by our toy model after proper calibration, and show that there is a wide range of parameters such that the resulting random topologies have a high probability to display a functional bottleneck. This result, by itself, shows that global epistasis alone is sufficient to generate functional bottlenecks. Furthermore, we studied the role of the distribution of the single mutational effects (SMEs). More specifically, we find that when the model is properly calibrated, this distribution of SMEs selected during evolution displays strong heterogeneity.

Our model thus sheds light on some fundamental constraints that shape fitness landscapes and possibly evolution. Finally, the model is analytically solvable to some extent, which might help making further progress in the investigation of its behavior for larger mutational spaces and more than two phenotypic variants.

II. RESULTS

A. Reanalyzing experimental data

To obtain inspiration for our toy model, we begin by reanalyzing experimental data from Poelwijk et al. [27]. In this work, the whole space of intermediates between two variants of the *Entacmea Quadricolor* fluorescent protein *eqFP611* were studied. The two reference wild types are *mTagBFP2*, which is fluorescent in the red, and *mKate2* which is fluorescent in the blue. Their sequences differ by 13 amino acids, resulting in a total of $2^{13} = 8192$ different combinations of intermediate amino acid substitutions, whose fluorescence was measured in the experiment.

We downloaded the original data from Ref. [27] and represented the genotypes with binary vectors of 13 variables, with $\mathbf{a}_B = (0, 0, \cdots, 0)$ being the blue *mKate2* and $\mathbf{a}_R = (1, 1, \cdots, 1)$ being the red *mTagBFP2*. In the experiment, *E.Coli* cells expressing the mutant genotypes were encapsulated into droplets and sorted by a microfluidic device able to push droplets into separate channels according to their fluorescence. From the raw experimental measurement, we assign to each genotype a red fitness, $F_R(\mathbf{a})$, and a blue fitness, $F_B(\mathbf{a})$, based on the enrichment in the two channels, each normalized to its corresponding wild type value such that $F_R(\mathbf{a}_R) = F_B(\mathbf{a}_B) = 1$ (see the SI for details). This normalization provides a consistent reference scale, with each wild type serving as the baseline functional state for its phenotype, enabling a more direct quantitative comparison between the red and blue fitness values. Note that here, fitness refers to fluorescence intensity, which is unrelated to reproductive success. Rather, fluorescence measures how well the protein performs its function. Under the assumption that better-functioning proteins offer a selective advantage in their relevant environments, fluorescence serves as a reasonable proxy for fitness.

For each measured phenotype (we will use a suffix 'R' for red and 'B' for blue fluorescence), we can define an underlying trait and a non-linear mapping to fitness $F = \phi(E)$. Inverting the non-linear function, we can thus derive the underlying phenotype from the measured fitness,

$$E_R(\mathbf{a}) = \phi_R^{-1}\left[F_R(\mathbf{a})\right] \ , \quad E_B(\mathbf{a}) = \phi_B^{-1}\left[F_B(\mathbf{a})\right] \ , \quad (2)$$

where we used for simplicity the same non-linear function $\phi_R^{-1}(x) = \phi_B^{-1}(x) = x^{0.44}$ that was chosen in Ref. [27] to minimize the epistatic contributions. We note that when there are two phenotypes only, for simplicity, we can follow Alqatari and Nagel [38] and encode them both

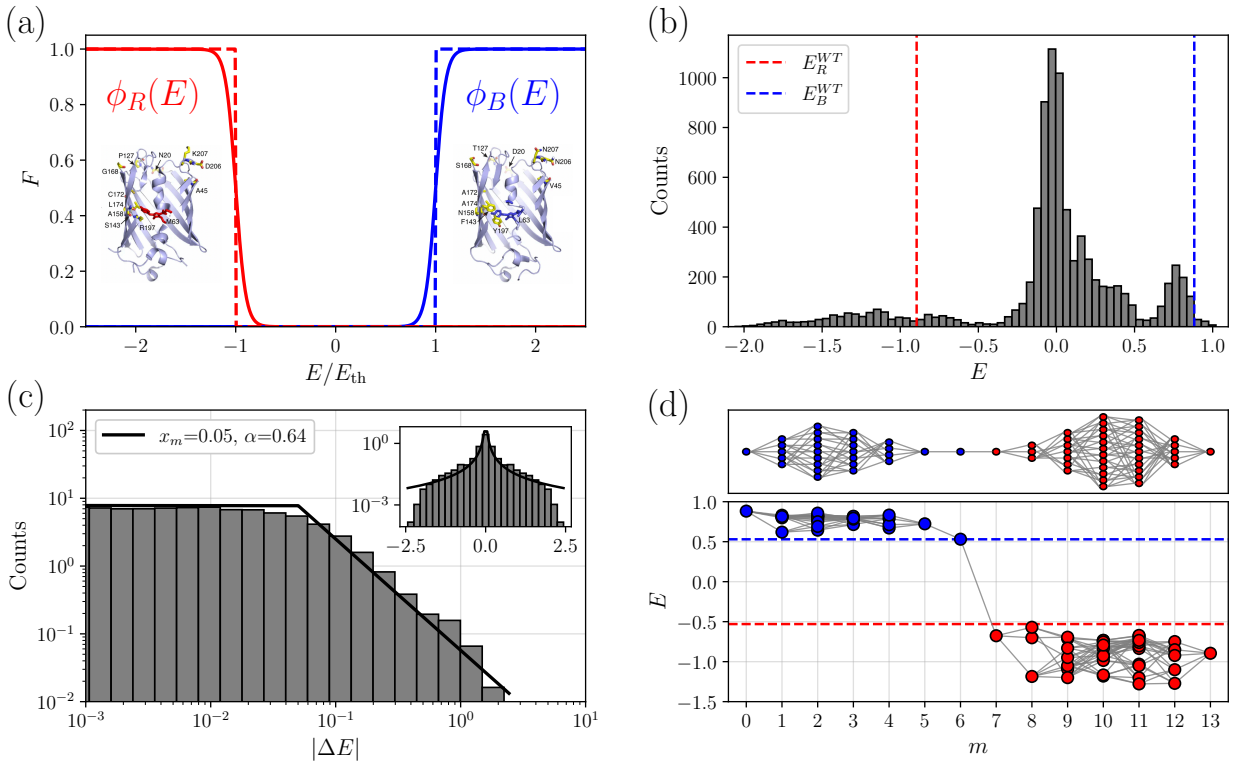

Figure 1. Reanalyzing experimental data from Ref. [27]. (a) Schematic shape of the fitness functions for the red and blue phenotypes as a function of the underlying trait E. Sigmoid functions have been used for illustrations. The two wild type structures with the corresponding mutations are also shown, from Ref. [27]. (b) Histogram of the value of E obtained from Eqs. (2,3) for each of the 2^{13} variants. The red and blue lines correspond respectively to the wild type values E_R^{WT} and E_B^{WT}. (c) Histogram of the absolute value $|\Delta E|$ of single mutational effects (SMEs) from (5) for all 13×2^{13} single mutations that can be obtained from the dataset, shown in log-log (main panel) and lin-log (inset) scales. The black curve is the fit obtained with the Pareto distribution in (6). (d) Topology of the space of paths obtained keeping only genotypes $\mathbf{a}$ with $|E(\mathbf{a})| > E_C$, and finding the largest possible value of $E_C = 0.53$ (dashed line) such that the red and blue wild types remain connected. The lower panel reports the values of $E(\mathbf{a})$ for each functional genotype (blue dots for $E > E_C$ and red dots for $E < -E_C$) as a function of the number of mutations m from the blue wild type, with the gray lines connecting pairs of genotypes that differ by a single mutation. The upper panel shows the resulting graph of connections.

in a single trait

$$E(\mathbf{a}) = E_B(\mathbf{a}) - E_R(\mathbf{a}) \ , \tag{3}$$

which is positive for the 'blue' phenotype and negative for the 'red' phenotype. This is only possible when the two phenotypes are exclusive, as in Ref. [27], which may not be true in other cases [32, 43]. We also define the wild type phenotype values as

$$E_B^{WT} = E(\mathbf{a}_B) \ , \qquad E_R^{WT} = E(\mathbf{a}_R) \ . \tag{4}$$

The structure of the two proteins and the 13 substitutions are illustrated in Fig. 1a, together with a schematic illustration of the non-linear mapping between $E = E_B - E_R$ and the two fitnesses. Note that other fitting functions could be used in the analysis, such as the sigmoid function ϕ (used for illustration in Fig. 1a), as suggested by more recent studies of global epistasis [24, 32, 37]. However, it would be even better to infer the non-linearity directly from the fitting procedure using the rank statistics [44].

Fig. 1b shows the histogram of $E(\mathbf{a})$ across all 2^{13} genotypes, with red and blue lines corresponding to the two wild types. Note that the wild type values are $E_B^{WT} \sim 0.88$ and $E_R^{WT} \sim -0.89$. Their absolute value slightly differs from one due to the contribution from the other phenotype. As observed in Ref. [27], most genotypes are non-functional with $E \sim 0$. There is a rather sharp peak of blue genotypes around $E \sim 1$, while red genotypes exhibit a much broader distribution, ranging from $E \sim -0.5$ down to large negative values of E. This indicates that some intermediates are significantly more fluorescent in the red than the red wild type itself.

Having defined $E(\mathbf{a})$, we can analyze the distribution of SMEs. For each of the 2^{13} backgrounds $\mathbf{a}$, we consider all 13 single mutations and we compute the SME,

$$\Delta E_i(\mathbf{a}) = E(a_i \to 1 - a_i \,|\mathbf{a}) - E(\mathbf{a}) \ . \tag{5}$$

The histogram of the absolute value of the 13×2^{13} SMEs is reported in Fig. 1c. (Note that the 13×2^{13} SMEs include each mutation and its reverse, hence the histogram

would be symmetric by construction, which is why we focus here on $|\Delta E_i|$.) Although we have limited data, we clearly observe that the distribution is fat-tailed. A good fit is achieved by a Pareto distribution,

$$p(x) = \frac{\alpha}{2x_m(\alpha+1)} \times \begin{cases} 1 & \text{if } |x| < x_m \ , \\ \left(\frac{x_m}{|x|}\right)^{\alpha+1} & \text{if } |x| > x_m \ , \end{cases} \quad (6)$$

here with $x = \Delta E$. The fit is just an indication, in particular due to the limited amount of data and the limited range of the experimental fitness, which both cutoff the distribution at large ΔE. While other distributions could also fit the data, the distribution tails are quite broad, indicating substantial heterogeneity among SMEs (in the SI we show similar results for another, independent dataset of SMEs [37]).

Finally, in Fig. 1d, we examine the topology of the resulting functional bottleneck. To do so, we first need a definition of functional genotype. As observed in Ref. [27] and evident from Fig. 1b, the threshold for defining a genotype as 'functional' is somewhat arbitrary, and playing with this threshold changes the topology substantially [38]. In this work, we consider that the fitness function is growing fast around a given threshold $E_{\rm th}$ (either on the positive or negative side, see Fig. 1a). Hence, genotypes with $|E(\mathbf{a})| < E_{\rm th}$ are considered dysfunctional and evolutionarily disallowed. Because the wild types are functional, they must of course be located above the functionality threshold, i.e. $|E^{WT}| \gtrsim E_{\rm th}$. However, Fig. 1b shows that no path of single mutations exists that connects the two wild types while always maintaining $|E(\mathbf{a})| \geq E^{WT}$, hence we need to consider $E_{\rm th} < E^{WT}$ in order to preserve the connection between the wild types. (Note that a different result was obtained in Ref. [27] when the two phenotypes were not normalized to the wild types. We discuss this point in the SI.) What is, then, the largest value of the functionality threshold $E_{\rm th}$ such that at least one viable mutational path exists between the two wild types?

In order to precisely answer this question, following Ref. [38], we characterize the space of the paths connecting the two wild types upon increasing the threshold $E_{\rm th}$. We consider the subset of paths that (i) are made of single mutations, (ii) connect the red and blue wild types, and (iii) are such that $|E(\mathbf{a})| > E_{\rm th}$ all along the path. We start from $E_{\rm th} = 0$, and gradually increase $E_{\rm th}$ to the maximum value such that at least one path remains. The resulting value, which we call E_C, quantifies how 'hard' it is for evolution to find a path that realizes the functionality switch. If $E_C \sim E^{WT}$, then one can find a path of single mutations, such that all intermediates are equally functional to the wild types. If instead $E_C \ll E^{WT}$, in order to connect the two wild types one has to accept a loss of fitness with respect to the wild types, which makes the transition less likely, but still possible.

For the experimental data of Ref. [27], this analysis yields a value of approximately $E_C \sim 0.53$, or $E_C/E^{WT} \sim 0.6$. The resulting space of paths, depicted in Fig. 1d, exhibits the characteristic bottleneck shape also reported in Ref. [38], where all paths traverse a single 'jumper' genotype with $E = E_C$. In Fig. 1d, the jumper genotype has a blue phenotype, after which a single mutation brings to the red phenotype. This functional switch is driven by a mutation located roughly at the midpoint. The number of functional intermediates determines the number of evolutionary paths that can reach the critical 'jumper' genotype. More paths imply greater evolutionary accessibility, making it easier for evolution to find the key mutation; while fewer paths indicate a more constrained transition requiring a specific mutational sequence. In contrast, the analysis in Ref. [27] found a less pronounced bottleneck, with multiple mutational paths surviving - a scenario that can be reproduced here by adopting a less stringent definition of E_C which allows for multiple paths, see Ref. [38].

The goal of the subsequent analysis is to determine whether this bottleneck structure arises solely from network epistasis or if it can also be explained by a model of global epistasis.

B. Toy model of global epistasis

To capture the key ingredients needed for bottleneck topologies, we introduce a minimal toy model featuring only global epistasis, defined as follows:

- The genotype $\mathbf{a} = (a_1, a_2, \cdots, a_L)$ is a binary sequence (i.e., $a_i \in \{0, 1\}$) of length L.
- An underlying additive phenotype $E(\mathbf{a}) = \sum_{i=1}^{L} h_i a_i$ has random SMEs h_i that are identically and independently distributed according to a symmetric input distribution $P(h) = P(-h)$. After having drawn the values h_i independently, we impose that $\sum_{i=1}^{L} h_i = 0$ by shifting their mean, i.e., $h_i \to h_i - \frac{1}{L}\sum_{i=1}^{L} h_i$.
- The two fitness functions that will be associated with the two 'wild types' are non-linear functions of the underlying random additive phenotype, $F_B = \phi_B(E) = \phi_0/(1 + e^{\beta(E_{\rm th}-E)})$ and $F_R = \phi_R(E) = \phi_0/(1 + e^{\beta(E_{\rm th}+E)})$, as illustrated in Fig. 1a. If $E(\mathbf{a}) > E_{\rm th}$ then the genotype $\mathbf{a}$ is functional 'blue', if $E(\mathbf{a}) < -E_{\rm th}$ it is functional 'red', and if $|E(\mathbf{a})| \ll E_{\rm th}$, the genotype is non-functional.

The fitness functions grow sharply around $\pm E_{\rm th}$, in a way controlled by the parameter β. The choice we will make implicitly, following Ref. [38], is to consider the limit $\beta \gg 1$, in which the fitness is close to a sharp threshold function at $E_{\rm th}$ (see Fig. 1a, dashed curves). Note that the values h_i can be multiplied by an arbitrary constant that can be absorbed into $E_{\rm th}$ and β, which allows us to fix the overall scale (e.g. the variance) of $P(h)$ without loss of generality.

Given the above definitions, both genotypes $\mathbf{a} = \mathbf{0} = (0, 0, \cdots, 0)$ and $\mathbf{a} = \mathbf{1} = (1, 1, \cdots, 1)$ have $E = 0$ and are thus non-functional. We stress that contrarily to the analysis of experimental data reported above, these two genotypes will not be the wild types in the toy model. Following Ref. [38], we thus introduce a 'tuning procedure' to generate two wild types, the 'red' $\mathbf{a}_R$ with $E_R^{WT} < -E_T$ and the 'blue' $\mathbf{a}_B$ with $E_B^{WT} > E_T$, for a chosen 'tuning' value $E_T > 0$. The procedure starts from the genotype $\mathbf{a} = \mathbf{0} = (0, 0, \cdots, 0)$, thus considered as a 'common ancestor', and sequentially introduces mutations as follows.

- With probability p, we perform a 'greedy' step. We scan all the h_i values that have not been introduced yet (those for which $a_i = 0$), and choose the one that has the maximum impact towards the desired goal, i.e., the largest positive h_i for the blue wild type or the smallest negative h_i for the red wild type. We introduce the corresponding mutation by switching $a_i = 1$.

- With probability $1-p$, we perform a 'random' step. We choose at random one site a_i such that $a_i = 0$, and mutate it to $a_i = 1$.

The tuning procedure is performed independently for the two wild types starting from $\mathbf{a} = \mathbf{0}$, and continues until for the first time $E_R^{WT} < -E_T$ to construct the red wild type, or $E_B^{WT} > E_T$ to construct the blue wild type. The motivation for this procedure is to mimic somehow the tuning of elastic networks performed in Ref. [38], in which the network is tuned based on a local proxy for the global fitness; we hereby assume that this local proxy can indeed provide the best choice (with probability p) or miss it (with probability $1 - p$).

The parameters that define the model are the sequence length L, the *a priori* distribution of SMEs $P(h)$, the probability p and the tuning parameter E_T that controls the wild type construction. We found that many choices of $P(h)$ lead to similar outcomes if the model is properly calibrated. For simplicity, we thus restrict to two natural choices.

1. **Gaussian**: A Gaussian distribution with unit variance.

2. **Pareto cutoff**: Motivated by the analysis of experimental data in Fig. 1c, a Pareto distribution, i.e., (6) with $x = h$, where we fix $\alpha = 0.7$ and $x_m = 0.1$. We also include an upper cutoff as in the data, i.e. we set $P(h) = 0$ for $|h| > 2$ and adjust the normalization constant accordingly.

For a given choice of parameters, an 'instance' of our random model is thus defined by the set of the L SMEs, h_i, and by the two wild types $\mathbf{a}_R$ and $\mathbf{a}_B$. For a given instance, we can define the number of mutations M separating $\mathbf{a}_R$ and $\mathbf{a}_B$ (i.e., the number of a_i that differ in the two sequences), and we can then investigate the $M!$ directed paths that connect them, obtained by introducing these mutations in all possible orders. Following Ref. [38] and as discussed above, we can find the maximum value of $E_{\rm th}$ such that at least one single-mutational path connecting the two wild types, with all intermediate states having $|E| > E_{\rm th}$, survives, and we call this value E_C. The closer E_C is to E^{WT}, the easier it is to perform the functional switch.

Given (L, p, E_T) and $P(h)$, we wish to characterize the statistical properties of the quantities M, E^{WT}, and E_C, which characterize the overall topology of the fitness landscape. We will also define $E_{\rm max}^{WT} = \max(E_B^{WT}, |E_R^{WT}|)$. In the following, we will use brackets, $\langle A \rangle$, to denote the statistical average of an observable A over an ensemble of random instances generated by the model. Unless otherwise specified, the average is taken over 20000 independent instances of the model for each value of the parameters.

Our model is similar in spirit to Fisher's geometric model (FGM) as discussed in Refs. [2, 5, 19], from which it however differs in two important aspects. First, the choice of non-linearity, which reflects here the existence of two distinct fitness functions associated to each wild type. Second, the fact that we choose two reference wild types using the stochastic procedure described above, and we then restrict to the space of intermediates between the selected wild types. As we show in the following, these two ingredients, together with a proper calibration of the model, give rise to the desired bottleneck structure.

Before discussing the details of the calibration procedure, a typical example of the space of mutational paths obtained from a calibrated model is shown in Fig. 2. The similarity with Fig. 1d, modulo the (irrelevant) change of scale of E, is striking.

C. Calibration of the model

The first step is to calibrate the parameters (L, p, E_T) in such a way that our toy model reproduces the basic phenomenology observed in Refs. [27, 38]. The requirements are the following.

- The phenotype of the two wild types should be comparable to E_T, i.e. $|E^{WT}|/E_T \simeq 1$, in order for E_T to be a meaningful fitness scale.

- We want to reproduce a 'bottlenecked' structure of the space of evolutionary paths connecting the two wild types, as shown in Figs. 1d, 2 and in Ref. [38]. This is characterized by a single 'jumper' genotype through which all paths must pass. The mutation responsible for the functionality switch occurs just after (or just before) this genotype. We note that the jumper genotype can be connected by a single mutation to only one, or more than one, genotypes carrying the other phenotype.

- The 'jumper' genotype should occur approximately at half distance along the evolutionary trajectory,

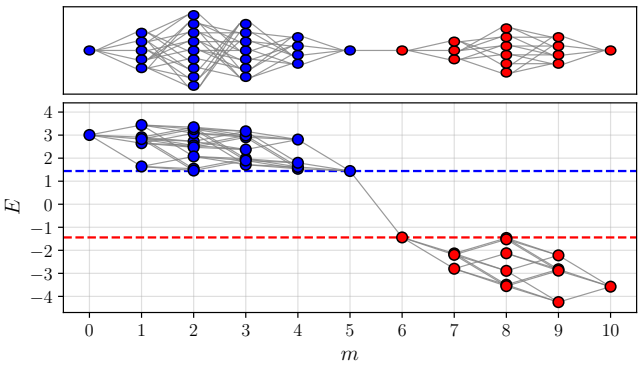

Figure 2. Same representation as in Fig. 1d, using an instance of the calibrated model with Gaussian $P(h)$ instead of the experimental data. The lower panel reports the value of E for each functional variant as a function of the distance from the blue wild type (here with $E_B^{WT}/E_T \approx 1.50$ and $E_R^{WT}/E_T \approx -1.79$). The threshold value for which at least a mutational path remains is $E_C/E_T \approx 0.72$ (dashed lines) and when normalized with respect to the largest of the two wild-types it reads $E_C/E_{\max}^{WT} \approx 0.40$.

as it is observed in experimental data (Fig. 1d). Furthermore, we want the largest possible number of functional intermediate variants to survive before and after the jump, thus maximizing the number of functional evolutionary paths connecting the two wild types.

- The fitness variation produced by the 'jumper' mutation(s) should be as large as possible, in such a way that E_C is as close as possible to E^{WT}, as discussed in Ref. [38]. This makes the resulting mutational paths more viable for evolution.

- Finally, for practical reasons, the number of mutations M connecting the two wild types should be small enough, otherwise enumerating all the $M!$ paths becomes computationally very expensive. To avoid this, we impose $\langle M \rangle \sim 8$, a value close to that of Ref. [38].

An important remark is in order about the expected value of E_C. Consider for illustration the case in which the tuning procedure reaches the target with a single greedy step for both wild types. We can thus expect $E_B^{WT} \sim \max(h_i)$ and $E_R^{WT} \sim \min(h_i)$. In absence of network epistasis, the largest jumper mutation that can happen along a path has $|\Delta E| \sim \max(|h_i|) \sim E_{\max}^{WT}$, but this mutation must bring the system from E_C to $-E_C$, hence $2E_C \sim |\Delta E| \sim E_{\max}^{WT}$. Based on this simple argument, the largest possible value of E_C we can expect is such that $E_C/E_{\max}^{WT} \sim 1/2$, as we observe numerically (Fig. 3c).

To fix the model parameters (L, p, E_T) we proceed as follows. First of all, we fix $L = 500$ as the length of a typical protein. We checked that the results are very weakly dependent on L (see the SI). Next, for each p, we progressively increase E_T until the average number of mutations $\langle M \rangle$ separating the two wild types is $\langle M \rangle \sim 8$. The resulting $E_T(p)$ is plotted as a function of p for both choices of $P(h)$ in Fig. 3a, and we find that it is approximately linear in p. The linear fit allows us to fix $E_T(p)$ for all p.

The next step is to calibrate p. For this, we consider the ratio $|E^{WT}|/E_T$ for both wild types, which we want to be close to one. The evolution of its statistical mean with p is shown in Fig. 3b. For small p, we mostly perform random steps in the training, and the requirement that M is small forces E_T to be small as well (Fig. 3a). As a result, it is very likely to 'overshoot' the target and end up with $|E^{WT}|/E_T \gg 1$. Hence, in order to have $|E^{WT}|/E_T \sim 1$, we need a large enough p.

Next, we consider the connectivity threshold E_C. We recall that E_C is obtained, for each instance, in such a way that at least one single-mutational path connecting the two wild types exists, with all intermediate states having $|E| > E_C$. Because both wild types must always be included in the space of allowed genotypes, increasing the value of E_C beyond $E_{\min}^{WT} = \min(E_B^{WT}, |E_R^{WT}|)$ is not meaningful [38]. This gives an upper bound for the largest value of E_C for each instance. We show in Fig. 3c its average value as a function of p, where the average is restricted to instances such that $E_C < 0.9E_{\min}^{WT}$. When p is too large, the tuning procedure requires too many greedy steps and as a consequence $\langle E_C/E_{\max}^{WT} \rangle$ tends to decrease, contrarily to what we want in order to have bottleneck topologies with large E_C. This suggests to lower the value of p as much as possible.

However, in Fig. 3d we report the probability of having a single jumper at E_C, which we estimate for numerical convenience by the probability that $E_C < 0.9E_{\min}^{WT}$, as a function of p. When p is too small, this probability drops quickly below one. This is because $|E^{WT}|/E_T$ starts increasing, the wild types are obtained by many random steps, and the bottleneck structure is lost.

Taken together, these results suggest to choose the lowest value of p such that the probability of a single jumper is still very close to one, which also gives both $E^{WT}/E_T \sim 1$ and $E_C/E_{\max}^{WT} \sim 0.5$. The value of $E_T(p)$ is fixed by the requirement that $\langle M \rangle \sim 8$. For both choices of $P(h)$, this optimal value is around $p = 0.25$, although slightly larger values could be considered as well.

The crucial insight from the calibration is that one needs to tune p in order to achieve a good balance between greedy and random steps. Correspondingly, as we will discuss later, the distribution of the SMEs that are selected by the tuning procedure that generates the wild types is very heterogeneous, with a coexistence of large SMEs (generated by greedy steps) and small SMEs (generated by random steps).

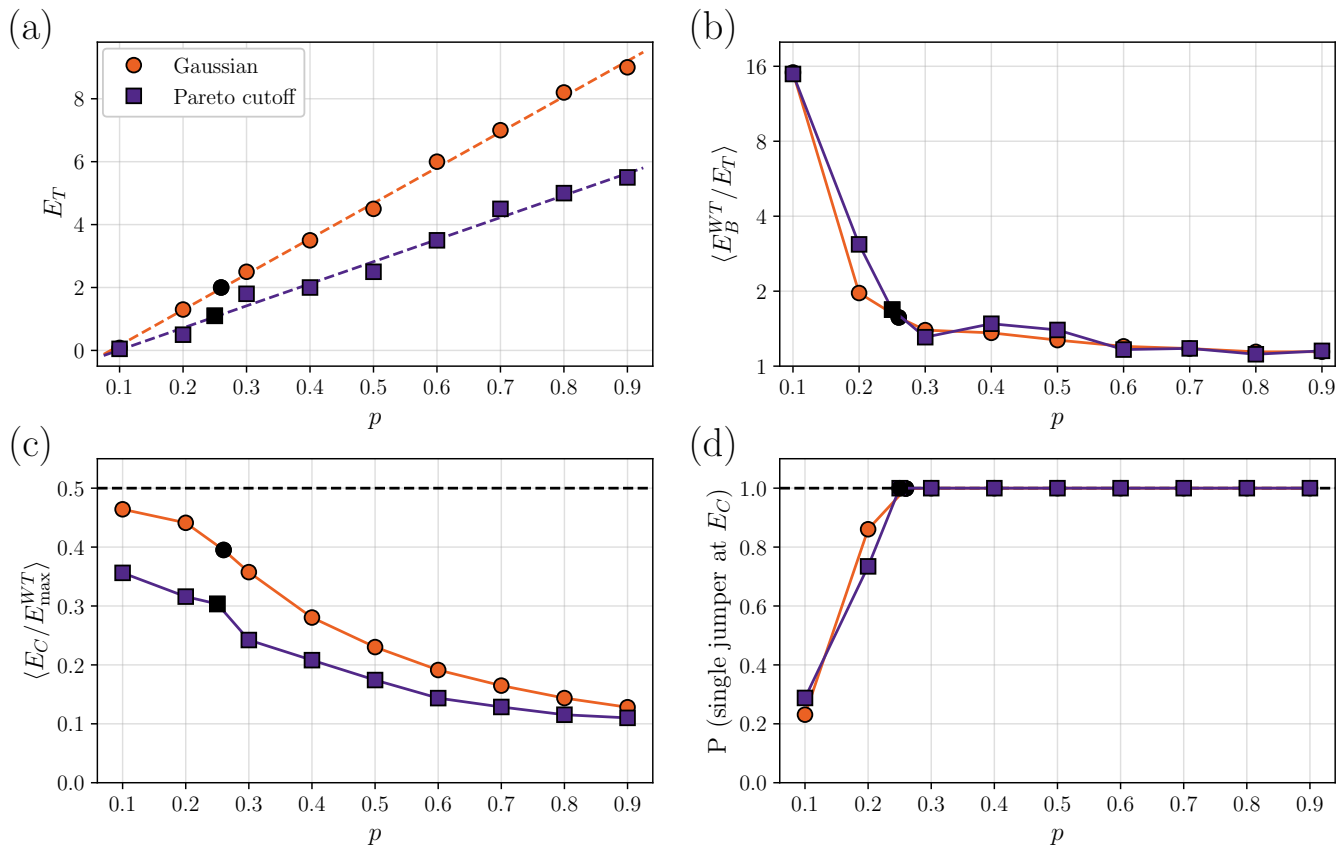

Figure 3. Calibration of the two models with different $P(h)$, a Gaussian distribution or a Pareto distribution with cutoff. The black symbols correspond to the choice of p after calibration. (a) Value of E_T for which $\langle M \rangle \approx 8$ as a function of p. Dashed lines are linear fits. (b) Average value $\langle E_B^{WT}/E_T \rangle$ as a function of p. (c) Average of $E_C/E_{\max}^{WT}$ as a function of p. (d) Probability of having a single jumper at E_C, or equivalently that $E_C < E_{\min}^{WT}$, as a function of p.

D. Results for the calibrated model

In the following, we thus choose $L = 500$ and the parameters corresponding to the black dots in Fig. 3 as the final parameters for which we report the results: Gaussian $P(h)$ with $p = 0.26$ and $E_T = 2.0$ for the tuning procedure, and Pareto cutoff $P(h)$ with $p = 0.25$ and $E_T = 1.1$. Keeping these two models fixed, we generated many (~ 20000) random instances of the model and studied the distribution of the relevant quantities that characterize the topology of the fitness landscape. An example of a 'good' topology, closely resembling the one found experimentally, has already been given in Fig. 2. We want to quantify the probability of generating such a topology in the toy model.

We first focus on the properties of the two wild types, shown in Fig. 4 for both models. More specifically, Figs. 4a,e show the histograms of M, the total number of mutations separating the two wild types. We note that, by construction, $M \geq 2$, and we fixed $\langle M \rangle = 8$. We see that the distribution is quite broad, leading to a significant fraction of instances with M as large as 20, and resembles that shown for elastic networks in Ref. [38]. However, when we investigate the mutational paths, we restrict the analysis to instances with $M \leq 14$ for computational tractability, as in Ref. [38]. Figs. 4b,f show the histograms of E_B^{WT}/E_T for the blue wild type (similar results are obtained for the red wild type due to the inherent symmetry of the model). We recall that, by construction of the training procedure, $E_B^{WT}/E_T \geq 1$. We observe that the histogram is peaked slightly above the minimal value, which confirms that the tuning procedure produces $E_B^{WT} \sim E_T$ and $E_R^{WT} \sim -E_T$. Figs. 4c,g show the histograms of the number of greedy steps in the tuning procedure, which is peaked around one, as expected. Finally, Figs. 4d,h show the histograms of the M SMEs, here called $\widetilde{h}_i$, that are selected by the tuning procedure, hence characterize *a posteriori* (after tuning) the space of intermediates between the wild type. We observe in both cases a strongly bimodal distribution, which reflects the choice of the tuning procedure: mutations selected by the random steps tend to be small, while mutations selected by the greedy steps tend to be much larger (in a way that can be quantified by extreme value statistics, see the SI).

Next, we analyze the statistical properties of the space of paths connecting the wild types (Fig. 5). We recall

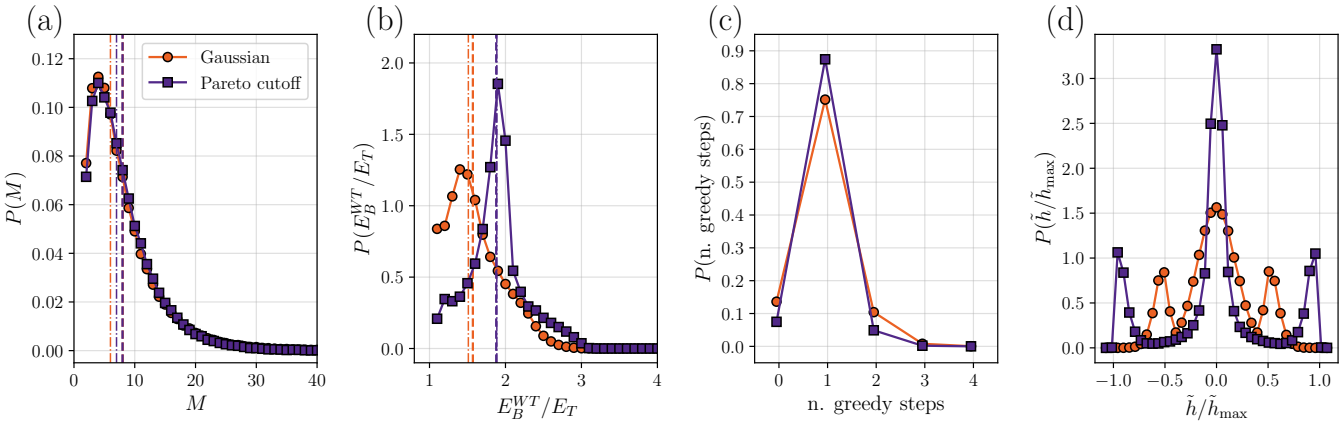

Figure 4. Distribution of some relevant quantities for two different calibrated model, the Gaussian model and the Pareto cutoff model. In the upper line, $L = 500$ SMEs are extracted from a Gaussian distribution with unit variance and the tuning procedure has $p = 0.26$, $E_T = 2.0$; in the lower line, $L = 500$ SMEs are extracted from a distribution with Pareto tails decaying with $\alpha = 0.7$ and a cutoff $|h| < 2.0$, and the tuning procedure has $p = 0.25$, $E_T = 1.1$. In each plot the dashed and dot-dashed vertical lines represent the mean and the median of the data, respectively. (a) Distribution of the number of mutations. (b) Distribution of the value of the positive (blue) wild type divided by E_T. (c) Distribution of the number of greedy steps needed to reach the target value E_T when generating the two wild types. (d) Distribution of SMEs $\tilde{h}_i$ selected by the tuning procedure to generate the two wild types.

that in the calibrated models we find $E_C < E_{\min}^{WT}$ with probability one. The space of paths thus contains a single genotype through which all paths must pass: raising E_C would remove this genotype from the set, and disconnect the wild types. A bottleneck is thus observed. We note that the jumper genotype, in most cases, is connected by a single mutation to only one genotype carrying the other phenotype, as in Fig. 2. However, in a few cases, we also observe situations where the jumper genotype is connected to a few genotypes carrying the other phenotype. In Fig. 5a,d we show the histograms of $E_C/E_{\max}^{WT}$, separately for instances with given M. We observe that the histogram is almost independent of M (with more variability observed for the Pareto cutoff model), and peaked around $E_C/E_{\max}^{WT} \sim 1/2$ as discussed above. Fig. 5b,e shows the histogram of the location of the jumper, called j, along the mutational path. We observe that the jumper is often located mid-way between the wild types, leading to a peaked histogram around $j/M \sim 1/2$. The peak (slowly) becomes sharper upon increasing M. The qualitative shape of the possible topologies produced by the model is indicated in Fig. 5g. Finally, we can compute the number of paths that remain viable at E_C. The average of the logarithm of this number is shown in Fig. 5c,f as a function of M. We find that it grows almost linearly at large M, indicating that the number of paths grows exponentially in M. This suggests that there are exponentially many ways to reach the bottleneck, which could thus maintain a good evolutionary accessibility from both wild types [38].

We conclude that our toy model, which features a global non-linearity on top of an underlying additive phenotype, with properly calibrated parameters, generates topologies that are very close to those observed experimentally. An example is shown in Fig. 2, to be compared with the result obtained re-analyzing the experimental data from Ref. [27], shown in Fig. 1d. The model generates an ensemble of random fitness functions [5], whose topology is likely (with probability one) characterized by a bottleneck, i.e., by a genotype through which all paths connecting the two wild types must pass. Of course, the bottleneck can be made less pronounced if one allows for less stringent selection, i.e., reduces E_C [38]. The bottleneck is more likely found mid-way along the path. These features are essentially independent of M, i.e., the number of mutations separating the two wild types, at least in the limited range investigated here.

We note that there is quite a lot of flexibility in the choice of parameters. The initial size L can be varied in a broad interval without affecting much the results (see SI). The tuning parameter E_T has been chosen here in such a way to obtain the desired average of M, but other choices are possible. The value of p, i.e., the probability of making a greedy step when constructing the wild types, can be varied in a relatively wide range, provided it is neither too small (otherwise E_T is too small and the wild types end up overshooting the fitness threshold) nor too large (otherwise $\langle E_C/E_{\max}^{WT}\rangle$ decreases. Finally, the *a priori* distribution $P(h)$ that characterizes the input distribution of SMEs can be varied essentially at will in a broad range (see also the SI for other choices of the input distribution). The tuning procedure used to construct the wild types ensures that the SMEs selected *a posteriori* are broadly distributed (Fig. 4d,h), which is a necessary condition for the emergence of a bottleneck. We thus believe that our results are robust and do not

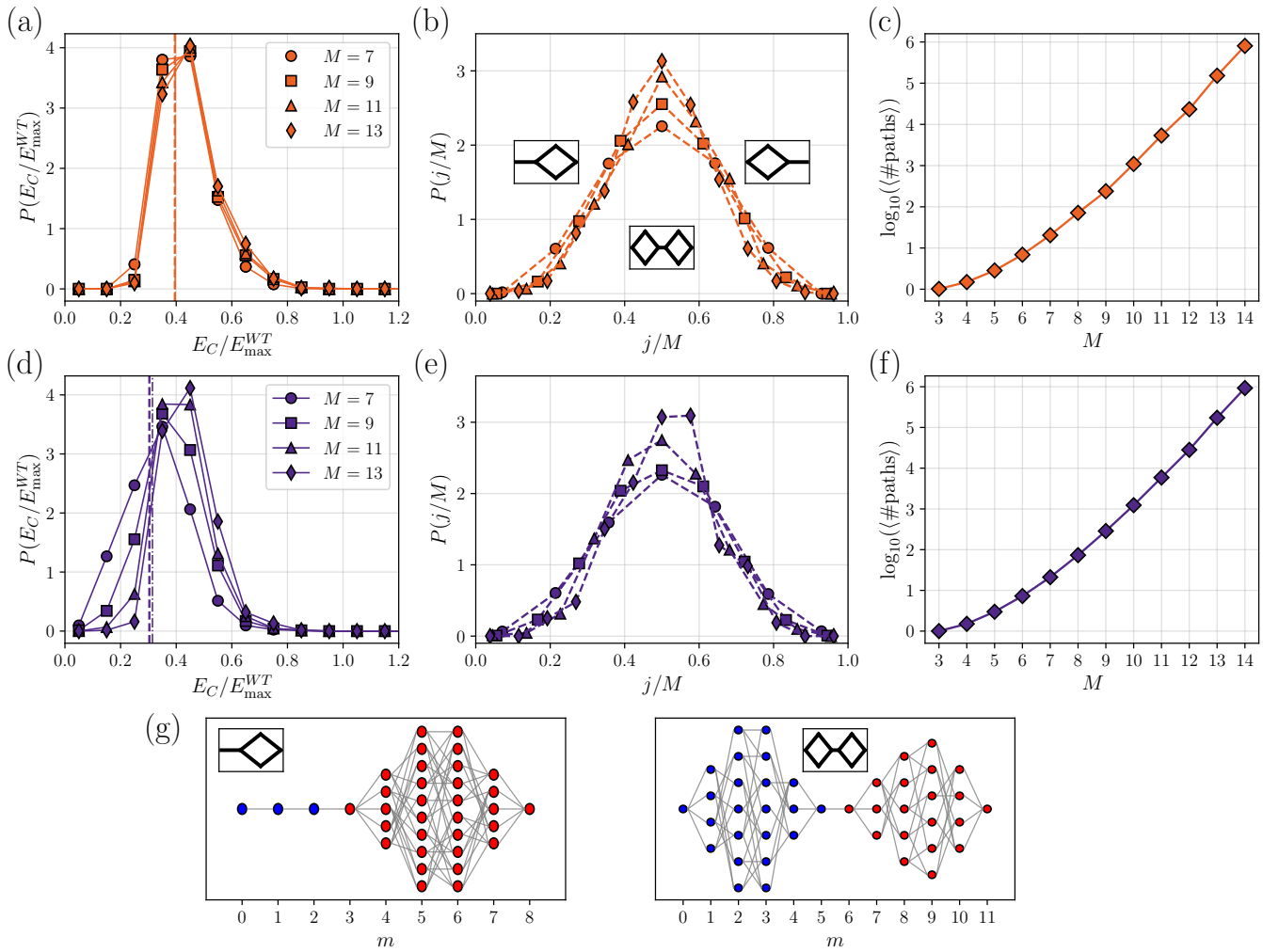

Figure 5. Statistical properties of the space of paths for two different calibrated model, the Gaussian model (a-c) and the Pareto cutoff model (d-f), with the same parameters as in Fig. 4. (a), (d) Distribution of values of $E_C/E_{\max}^{WT}$, separately for several values of M. The mean and the median (dashed and dot-dashed vertical lines, respectively) are computed with respect to the full distribution (with all M values). (b), (e) Probability distribution of the jumper position j along the path, for different values of M. (c), (f) Average of logarithm of the number of paths that remain viable at E_C, as a function of M. (g) Some examples of topologies, each associated to a schematic representation used in panel b.

depend crucially on the specific choices made during calibration.

III. DISCUSSION

In this work, we have demonstrated that functional bottlenecks in protein fitness landscapes can emerge even in the absence of network epistasis, solely due to a non-linear mapping between an underlying additive phenotype and fitness. By constructing and analyzing a simple model of global epistasis, we have shown that the emergence of bottlenecks is closely tied to the heterogeneity of the single-mutation effects selected during the construction of the wild types. When mutational effects are broadly distributed, evolutionary paths between distinct phenotypes are constrained, leading to the formation of narrow evolutionary corridors.

Our findings provide a novel perspective on how fitness landscape topologies arise and suggest that strong constraints on evolutionary accessibility can exist even in systems governed by simple, global epistatic interactions [2, 3, 5, 19, 23, 24, 27, 31, 32, 36, 37]. This has important implications for protein evolution, as it suggests that functional bottlenecks may not necessarily indicate complex, higher-order interactions but could instead be a natural consequence of non-linear selection on an underlying additive phenotype. Moreover, our model successfully captures key features observed in experimental data, supporting the idea that broad mutational effect distributions play a central role in shaping evolutionary trajectories.

The model is simple enough that it could possibly be solved analytically, thus eliminating the limitation on the number of mutations M separating the two wild types, using techniques from probability and statistical physics [2, 5, 19]. Future work could explore this possibility to extend our approach to larger mutational spaces, incorporate additional biological constraints, and investigate the robustness of these results across different fitness functions and evolutionary pressures, possibly with more than two phenotypes.

By highlighting the role of mutational heterogeneity in the emergence of functional bottlenecks, our study contributes to a deeper understanding of the constraints shaping protein evolution and hopefully opens new directions for theoretical and experimental exploration of evolutionary landscapes.

IV. METHODS

For convenience, we summarize here in closed form the main steps that have to be followed to generate an instance of our model.

A. Genotype representation and underlying additive phenotype

The first step consists in choosing a value of L and considering a genotype $\mathbf{a} = (a_1, a_2, \ldots, a_L)$ as a binary sequence of length L, where each site i can take values $a_i \in \{0, 1\}$. The genotype-dependent underlying additive phenotype is given by:

$$E(\mathbf{a}) = \sum_{i=1}^{L} h_i a_i, \tag{7}$$

where the single-mutation effects (SMEs) h_i are independent and identically distributed according to the *a priori* distribution $P(h)$. The two choices we used are a Gaussian distribution with zero mean and unit variance, and a Pareto distribution with cutoff:

$$P(h) = A \begin{cases} 1 \ , & |h| < 0.1 \ , \\ 1/|h|^{\alpha+1} \ , & 0.1 < |h| < 2 \ , \\ 0 \ , & |h| > 2 \ , \end{cases} \tag{8}$$

with $\alpha = 0.7$ and the constant A determined by normalization. After sampling, the SMEs are shifted to ensure $\sum_{i=1}^{L} h_i = 0$. These choices fix an additive constant and an overall multiplicative factor in $E(\mathbf{a})$, which can be absorbed by β and E_T in the fitness function, without loss of generality:

B. Fitness function and phenotypic classes

The global epistatic fitness function is a non-linear transformation of $E(a)$, defined separately for the blue (B) and red (R) phenotypes:

$$F_B(E) = \frac{\phi_0}{1 + e^{\beta(E_{\rm th} - E)}}, \quad F_R(E) = \frac{\phi_0}{1 + e^{\beta(E_{\rm th} + E)}}, \tag{9}$$

where $\phi_0 = 1$ without loss of generality, $E_{\rm th}$ is the functionality threshold parameter, and $\beta \gg 1$ controls the sharpness of the fitness transition. Accordingly, genotypes with $E(\mathbf{a}) > E_{\rm th}$ are functional for the "blue" phenotype, those with $E(\mathbf{a}) < -E_{\rm th}$ are functional for the "red" phenotype, and those with $|E(\mathbf{a})| \ll E_{\rm th}$ are non-functional.

C. Wild type construction

The red and blue wild types, $\mathbf{a}_R$ and $\mathbf{a}_B$, are independently generated starting from the ancestral genotype $\mathbf{a} = (0, 0, \ldots, 0)$ by sequentially introducing mutations. At each step, with probability p, the mutation with the largest contribution to E in the desired direction is chosen (greedy step), and with probability $1 - p$, a random site is mutated (random step). The process is iterated until $E_R^{WT} = E(\mathbf{a}_R) < -E_T$ and $E_B^{WT} = E(\mathbf{a}_B) > E_T$.

D. Mutational paths and bottleneck characterization

The mutational distance M between $\mathbf{a}_R$ and $\mathbf{a}_B$ is defined as the number of differing sites. All $M!$ possible evolutionary paths between the wild types are considered, and a value E_C is determined such that at least one path of single mutations exists, maintaining $|E(\mathbf{a})| > E_C$ at all intermediate steps. The distribution of E_C and the likelihood of a single critical 'jumper' mutation define the presence and severity of the bottleneck.

ACKNOWLEDGMENTS

We warmly thank Sid Nagel and Martin Weigt for providing very useful feedback and inspiration all along the development of this work. This research has been supported by first FIS (Italian Science Fund) 2021 funding scheme (FIS783 - SMaC - Statistical Mechanics and Complexity) from MUR, Italian Ministry of University and Research and from the PRIN funding scheme (2022LMHTET - Complexity, disorder and fluctuations: spin glass physics and beyond) from MUR, Italian Ministry of University and Research.

[1] S. Wright, Proceedings of the Sixth International Congress on Genetics **1**, 355–366 (1932).
[2] J. A. G. De Visser and J. Krug, Nature Reviews Genetics **15**, 480 (2014).
[3] O. Tenaillon, Annual Review of Ecology, Evolution, and Systematics **45**, 179 (2014).
[4] S. Manrubia, J. A. Cuesta, J. Aguirre, S. E. Ahnert, L. Altenberg, A. V. Cano, P. Catalán, R. Diaz-Uriarte, S. F. Elena, J. A. García-Martín, P. Hogeweg, B. S. Khatri, J. Krug, A. A. Louis, N. S. Martin, J. L. Payne, M. J. Tarnowski, and M. Weiß, Physics of Life Reviews **38**, 55 (2021).
[5] S. Pahujani and J. Krug, arXiv:2502.05896 (2025).
[6] W. H. Bateson, Science **26 672**, 649 (1907).
[7] R. A. Fisher, Transactions of the Royal Society of Edinburgh **52**, 399–433 (1919).
[8] M. Lunzer, G. B. Golding, and A. M. Dean, PLoS genetics **6**, e1001162 (2010).
[9] M. J. Harms and J. W. Thornton, Nature Reviews Genetics **14**, 559 (2013).
[10] D. M. Weinreich, Y. Lan, C. S. Wylie, and R. B. Heckendorn, Current opinion in genetics & development **23**, 700 (2013).
[11] C. A. Olson, N. C. Wu, and R. Sun, Current biology **24**, 2643 (2014).
[12] M. Figliuzzi, H. Jacquier, A. Schug, O. Tenaillon, and M. Weigt, Molecular biology and evolution **33**, 268 (2016).
[13] T. N. Starr and J. W. Thornton, Protein science **25**, 1204 (2016).
[14] F. J. Poelwijk, V. Krishna, and R. Ranganathan, PLoS computational biology **12**, e1004771 (2016).
[15] O. Rivoire, K. A. Reynolds, and R. Ranganathan, PLoS computational biology **12**, e1004817 (2016).
[16] C. M. Miton and N. Tokuriki, Protein Science **25**, 1260 (2016).
[17] Z. R. Sailer and M. J. Harms, PLoS computational biology **13**, e1005541 (2017).
[18] Z. R. Sailer and M. J. Harms, Genetics **205**, 1079 (2017).
[19] S. Hwang, S.-C. Park, and J. Krug, Genetics **206**, 1049–1079 (2017).
[20] S. Cocco, C. Feinauer, M. Figliuzzi, R. Monasson, and M. Weigt, Reports on Progress in Physics **81**, 032601 (2018).
[21] J. Domingo, G. Diss, and B. Lehner, Nature **558**, 117 (2018).
[22] T. N. Starr, J. M. Flynn, P. Mishra, D. N. Bolon, and J. W. Thornton, Proceedings of the National Academy of Sciences **115**, 4453 (2018).
[23] J. Otwinowski, Molecular biology and evolution **35**, 2345 (2018).
[24] J. Otwinowski, D. M. McCandlish, and J. B. Plotkin, Proceedings of the National Academy of Sciences **115**, E7550 (2018).
[25] J. Domingo, P. Baeza-Centurion, and B. Lehner, Annual review of genomics and human genetics **20**, 433 (2019).
[26] A. Biswas, A. Haldane, E. Arnold, and R. M. Levy, eLife **8**, e50524 (2019).
[27] F. J. Poelwijk, M. Socolich, and R. Ranganathan, Nature communications **10**, 4213 (2019).
[28] A. Ballal, C. Laurendon, M. Salmon, M. Vardakou, J. Cheema, M. Defernez, P. E. O'Maille, and A. V. Morozov, Molecular biology and evolution **37**, 1907 (2020).
[29] D. M. Lyons, Z. Zou, H. Xu, and J. Zhang, Nature Ecology & Evolution **4**, 1685 (2020).
[30] C. M. Miton, K. Buda, and N. Tokuriki, Current opinion in structural biology **69**, 160 (2021).
[31] G. Reddy and M. M. Desai, Elife **10**, e64740 (2021).
[32] A. M. Phillips, K. R. Lawrence, A. Moulana, T. Dupic, J. Chang, M. S. Johnson, I. Cvijovic, T. Mora, A. M. Walczak, and M. M. Desai, Elife **10**, e71393 (2021).
[33] C. W. Bakerlee, A. N. Nguyen Ba, Y. Shulgina, J. I. Rojas Echenique, and M. M. Desai, Science **376**, 630 (2022).
[34] L. Vigué, G. Croce, M. Petitjean, E. Ruppé, O. Tenaillon, and M. Weigt, Nature Communications **13**, 4030 (2022).
[35] J. Chen, M. Bisardi, D. Lee, S. Cotogno, F. Zamponi, M. Weigt, and N. Tokuriki, Nature communications **15**, 8441 (2024).
[36] M. S. Johnson, G. Reddy, and M. M. Desai, BMC biology **21**, 120 (2023).
[37] K. Buda, C. M. Miton, and N. Tokuriki, Nature Communications **14**, 8508 (2023).
[38] S. Alqatari and S. Nagel, arXiv:2408.16926 (2024).
[39] K. Husain and A. Murugan, Molecular Biology and Evolution **37**, 2865 (2020).
[40] D. M. Weinreich, N. F. Delaney, M. A. DePristo, and D. L. Hartl, Science **312**, 111 (2006).
[41] N. C. Wu, L. Dai, C. A. Olson, J. O. Lloyd-Smith, and R. Sun, Elife **5**, e16965 (2016).
[42] P. Kantroo, G. P. Wagner, and B. B. Machta, bioRxiv:2024.11.13.623265 (2024).
[43] D. Bendixsen, J. Collet, B. Østman, and E. Hayden, PLoS Biol **17**, e3000300 (2019).
[44] M. O. Carlson, B. L. Andrews, and Y. B. Simons, bioRxiv:2025.04.08.647864 (2025).

SUPPLEMENTARY INFORMATION

S1. DETAILS ON THE ANALYSIS OF EXPERIMENTAL DATA

A. Data collection

The experimental measurements that were analyzed in the main text were obtained from the public repository associated with Ref. [27]. These are raw sequencing data capturing the genotype variants present in the red and blue channel after two-color Fluoresce Activated Cell Sorting. Following the procedure detailed in Ref. [27] (see the original paper for details), sequences with low quality scores and noisy measurements were filtered out. The enrichment scores for each genotype $\mathbf{a}$ were computed counting the occurrence of each variant in the two channels normalized to its abundance in the input channel, i.e. before sorting. We further normalized the enrichment scores with respect to the wild types, such that $F_R(\mathbf{a}_R) = F_B(\mathbf{a}_B) = 1$, and afterwards removed global epistasis via a non-linear function

$$E_f = \phi^{-1}(F_f) = F_f^{0.44} \ , \tag{S1}$$

as proposed by the authors of Ref. [27], separately for the blue ($f = B$) and red ($f = R$) fluorescence.

B. Results with non-normalized enrichment scores

In the main text, and as discussed in Sec. S1 A, we normalized the enrichment scores to the wild type values, a choice that was not made in the original Ref. [27]. Using normalized enrichment scores, we find that no path of single mutations exists that connects the two wild types while always maintaining $|E(\mathbf{a})| \geq E^{WT}$, hence we need to consider $E_{\rm th} < E^{WT}$ in order to preserve the connection between the wild types, as discussed in the main text. Instead, Ref. [27] shows a topology of the space of paths connecting genotypes with $|E(\mathbf{a})| \geq E_{\rm th}$ and $E_{\rm th} = E^{WT}_{\rm min}$. This discrepancy arises entirely from the choice of not normalizing the phenotypic trait.

The space of paths presented in Ref. [27] is constructed by defining a phenotype $E = \sqrt{\epsilon_B^2 + \epsilon_R^2}$, where ϵ_B and ϵ_R are the enrichment scores in the two channels, which are *not* normalized with respect to the enrichment scores of the red and blue wild types. However, they are rescaled to reflect the experimental brightness ratio between the parental blue and red genotypes, which is 0.172, and subsequently transformed using the non-linear function in Eq. S1 to remove global epistasis. To carry out a proper comparison with our analysis, we reproduce the space of paths of Ref. [27] modifying their definition of phenotype to $\widetilde{E} = \epsilon_B - \epsilon_R$. Fig. S1a shows that this is a minor change: because $\epsilon_R \sim 0$ when $\epsilon_B > 0$ and vice versa, the two phenotypes are essentially exclusive and as a consequence $E = \sqrt{\epsilon_B^2 + \epsilon_R^2} \sim \max(\epsilon_B, \epsilon_R) \sim |\widetilde{E}|$. Hence, $\widetilde{E}$ is numerically very close to the score used in the original Ref. [27], while keeping the negative sign for the red phenotype.

Compared to the normalized case shown in the main text, the resulting distribution of phenotypes, shown in Fig. S1b, is not symmetric around zero: the blue wildtype has $|\widetilde{E}_R^{WT}| = 0.56$ and the red one is more then twice as large, with $\widetilde{E}_B^{WT} = 1.49$. This asymmetry arises entirely from the absence of normalization. When looking for the largest value of the functionality threshold $\widetilde{E}_{\rm th}$ such that at least one viable mutational path exists between the two wild types, we find that $\widetilde{E}_C$ saturates to $\widetilde{E}^{WT}_{\rm min}$ and does not produce a topology with a single jump. The topology is the same as the one shown in Ref. [27], and we consistently find that $\widetilde{E}_C = \widetilde{E}^{WT}_{\rm min}$ is the critical value above which any connection between the two wild types is lost. This is exactly what we would expect given the imbalance between the two wild types. However, given the imbalance between wild types, we believe that choosing a symmetric threshold, i.e. choosing the same $\widetilde{E}_{\rm th}$ for both functions, is not appropriate.

We therefore believe that the choice of normalizing the enrichment scores to the wild types provides a better view of the global fitness landscape, with the distribution of phenotypic values centered around zero. This choice results in a critical threshold $\widetilde{E}_C < \widetilde{E}_{WT}$, as discussed in the main text.

C. Removing network epistasis from the data

From now on, we stick to the choice made in the main text, of normalizing the enrichment scores to the wild types. We note that even after the non-linear fitness function has been inverted by the fitting procedure we described, the effect of epistasis is still present in the data. This is due to the fact that the fitness cannot be simply expressed as a nonlinear function of an additive trait. An interesting test is then to try and remove this second source of epistasis (network epistasis) and check whether one can still obtain a topology similar to the one discussed in Refs. [27, 38] and in the main text. We do so by defining an additive phenotype $E_f^{(1)}$ obtained expanding the phenotypes E_f around their respective wild type $\mathbf{a}_f$ and cutting the higher order terms, namely

$$E_f^{(1)}(\mathbf{a}) = E_f(\mathbf{a}_f) + \sum_{i=1}^{13} h_i^{f,(1)}(a_{f,i} - a_i) \ . \tag{S2}$$

After linearization, we combine the two phenotypes into a single additive phenotype $E^{(1)}(\mathbf{a}) = E_B^{(1)}(\mathbf{a}) - E_R^{(1)}(\mathbf{a})$.

Fig. S2a shows a scatter plot of the approximated additive phenotype $E^{(1)}(\mathbf{a})$ against $E(\mathbf{a})$. In the figure we highlight the wild types (blue and red stars) along with the variants that are at a small Hamming distance from the blue (d_B) and red (d_R) wild types. One notices that

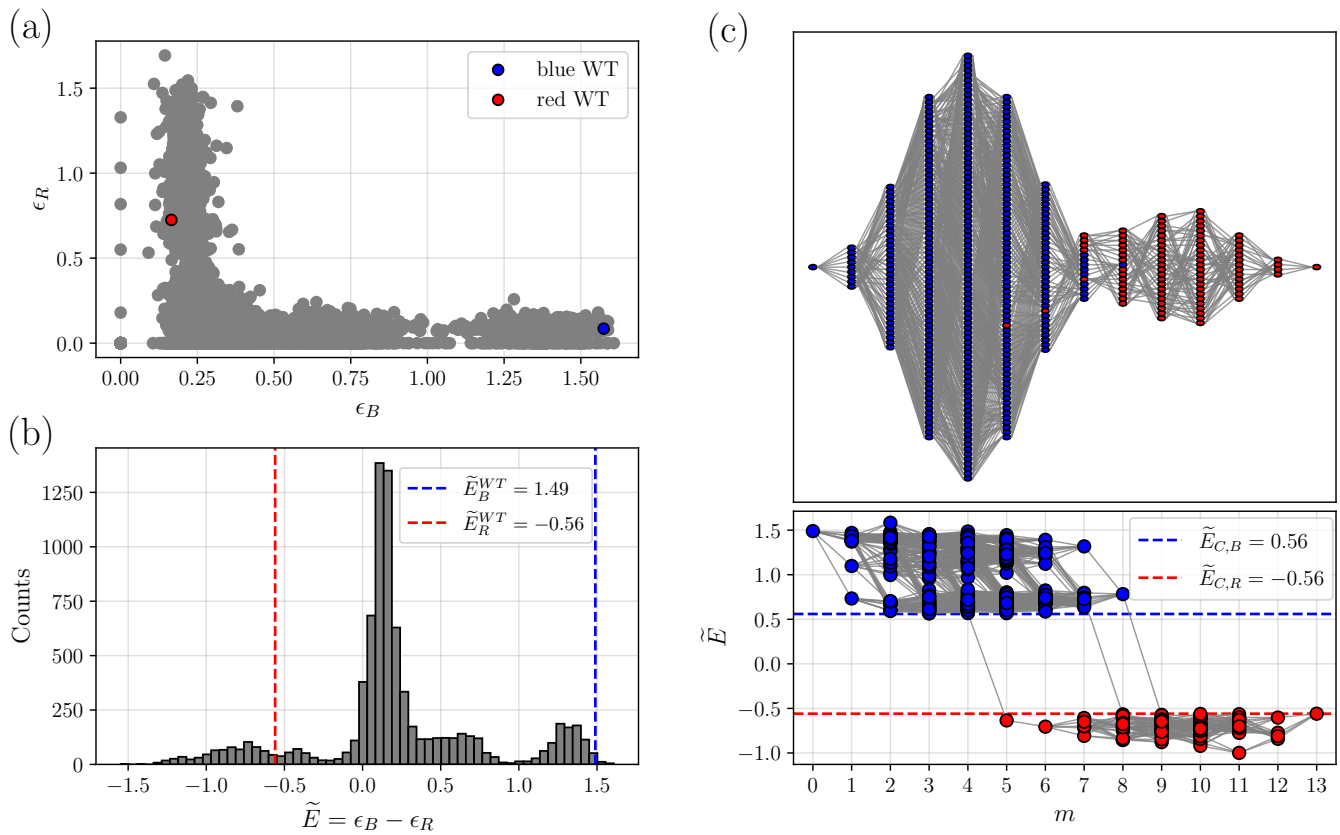

Figure S1. (a) Scatter plot of the enrichment scores in the blue and the red channel, ϵ_B and ϵ_R, not normalized with respect to the wildtypes. For the blue wildtype $\epsilon_B = 1.58$, $\epsilon_R = 0.09$. For the red wildtype $\epsilon_B = 0.17$, $\epsilon_R = 0.73$. (b) Histogram of the values of $\widetilde{E}$ obtained for each of the 2^{13} variants. The red and blue lines correspond respectively to the wild type values $\widetilde{E}_R^{WT}$ and $\widetilde{E}_B^{WT}$. (c) Topology of the space of paths obtained keeping only genotypes $\mathbf{a}$ with $|\widetilde{E}(\mathbf{a})| > \widetilde{E}_C$, where $\widetilde{E}_C = 0.56$ (dashed lines) is the largest possible value such that the red and blue wild types remain connected. The lower panel reports the values of $\widetilde{E}(\mathbf{a})$ for each functional genotype (blue dots for $\widetilde{E} > \widetilde{E}_C$ and red dots for $\widetilde{E} < -\widetilde{E}_C$) as a function of the number of mutations m from the blue wild type, with the gray lines connecting pairs of genotypes that differ by a single mutation. The upper panel shows the resulting graph of connections.

E and $E^{(1)}$ are not simply linearly related, even if they have a relatively high Spearman correlation coefficient of 0.7. This suggests that the first order expansion is not able to fully capture all mutational effects and that network epistasis is indeed present in the data [27]. We note that the variants at small Hamming distance from the wild types display a better correlation than the very distant ones, as expected.

Keeping in mind that the linearization is only an approximation, we analyzed the space of viable paths connecting the two wild types. Following the same approach described in the main text, we look for the largest value of $E_{\rm th}$ for which at least one path allowing for a functionality switch survives, and call this value E_C. However, because the two wild types in this case turn out to be quite asymmetric, with $E_B^{WT} \sim 0.49$ and $E_R^{WT} \sim 2.8$, we allowed for an asymmetric connectivity threshold [38] for the paths joining the two variants (see also the discussion in Sec. S1 B). Instead of increasing $E_{\rm th}$ symmetrically from 0 and impose $|E| > E_{\rm th}$ for every path, we impose $E > E_{\rm th}^B$ and $E < -E_{\rm th}^R$ separately for the two phenotypes. The largest values for which a path survives are $E_C^B = 0.21$ and $E_C^R = 1.64$, and we obtain the topology shown in Fig. S2b. Note that in the end, $E_C^B/E_B^{WT} \sim 0.42$ and $E_C^R/E_R^{WT} \sim 0.59$, hence we always end up with a connectivity threshold at about half the wild type phenotype. We observe a significant reduction in functional variants, approximately at half distance between the two wild types, and the emergence of a bottleneck topology.

S2. ANALYSIS OF ADDITIONAL DATA FOR SINGLE-MUTATION EFFECTS

In Ref. [37] an extensive review of combinatorial experiments similar to that of Ref. [27] is presented. The authors analyzed the results from ten studies (see references in [37]) of combinatorially complete fitness landscapes of seven different enzymes, totaling 1440 genotype-

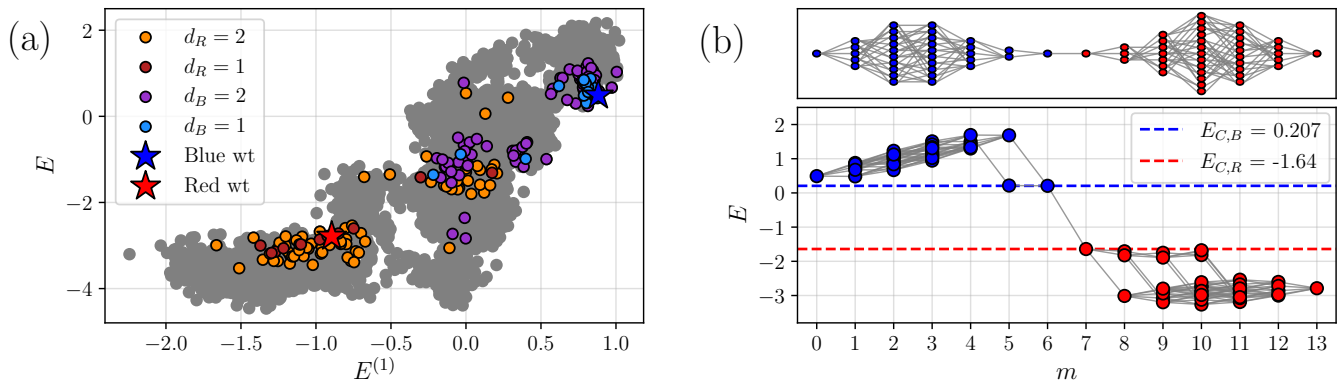

Figure S2. Linearization of the phenotypes according to the expansion described in Eq. (S2). (a) Scatter plot of the phenotype values shown in the main text, E, and after the expansion, $E^{(1)}$. The Spearman correlation coefficient is about 0.7. The red and blue stars correspond respectively to the red and blue wild types. (b) Topology of paths surviving at the connectivity thresholds E_C^B and E_C^R. Here E is the linearized additive phenotype.

phenotype data-points. The experimental results for fitness are then translated into an underlying phenotype by attempting at removing non-linearity (see Ref. [37] for details), and the authors provide the resulting distribution of SMEs that can be downloaded from the paper repository.

In Fig. S3, we show the same distribution as in Ref. [37], plotted similarly to the other figures in this paper. Once again the fit with the Pareto distribution seems to perform well, although with different values of x_m and α. Notice that the uniform behavior at small $|\Delta E|$, as well as the cutoff at large $|\Delta E|$, are not as sharp as in the data from Ref. [27]. This is probably due to the fact that Ref. [37] combined results for many different experiments, which have different sensitivity and a different range of measurable fitness.

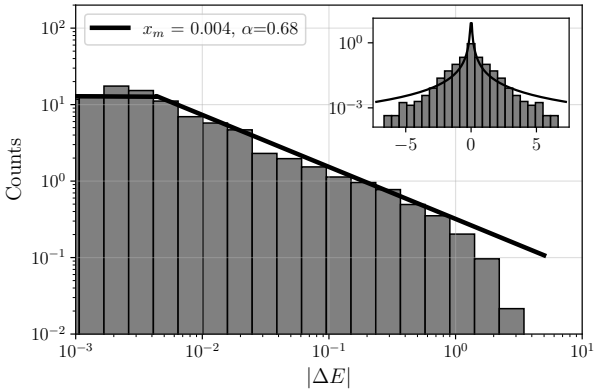

Figure S3. Histogram of the absolute value $|\Delta E|$ for all SMEs extracted from the dataset of Ref. [37].

S3. ADDITIONAL RESULTS ON THE TOY MODEL

A. Evolution of the space of paths with the number of mutations

In the main text we presented the statistics of $E_C/E_{\max}^{WT}$ for the calibrated models, i.e. for the parameter choices p and E_T that maximize both the connectivity threshold $E_C/E_{\max}^{WT}$ and the probability of observing bottleneck structures. These two quantities are shown in Fig. S4 as a function of the number of mutations M separating the wild types, for different values of p and E_T. As explained in the main text, the dotted line at $E_C = 0.5E_{\max}^{WT}$ corresponds to the maximum value of connectivity threshold that can be found when observing a bottleneck with a single functionality switch. At large p, it can be seen that although the fraction of bottlenecks remains equal to one, $\langle E_C/E_{\max}^{WT}\rangle$ decreases with increasing p. Low values of $E_C/E_{\max}^{WT}$ correspond to a connectivity threshold that lies deep within the region of non-functionality, making the corresponding evolutionary trajectories unlikely to occur. On the other hand, for small p, corresponding to a random selection of mutational effects, the fraction of observed bottlenecks decreases significantly.

For the values of p and E_T chosen after calibration (black dots), we observe simular results to Ref. [38]. The probability of a bottleneck remains always equal to one, while the value of $E_C/E_{\max}^{WT}$ increases with M and saturates to the expected maximum.

B. Dependence on system size

All the results reported in the main text have been obtained by fixing the number of SMEs to $L = 500$. This would correspond to a protein of 500 amino acids or to

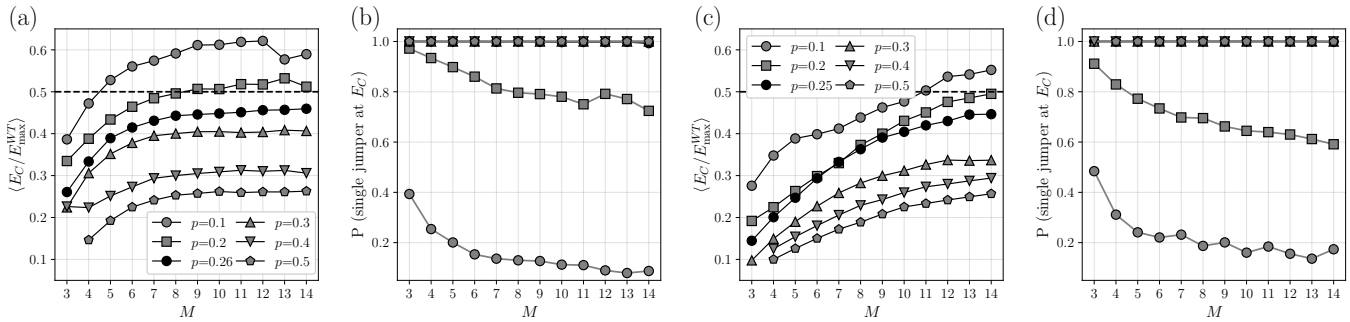

Figure S4. Statistical properties of the space of paths upon varying the number of mutations M, at fixed values of p, for the Gaussian (a),(b), and the Pareto cutoff (c),(d) distributions. (a),(c) The average value of $E_C/E_{\max}^{WT}$. (b),(d) The fraction of single jumpers present at E_C.

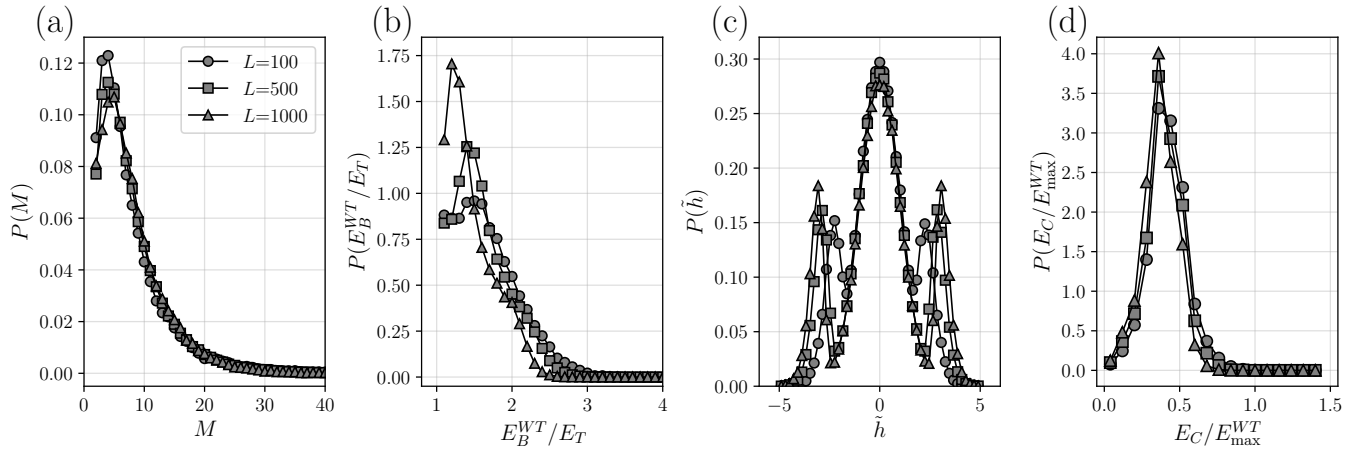

Figure S5. Statistics of a few relevant quantities for the model with Gaussian distribution, for $L = 100, 500, 1000$. (a) Distribution of the number of mutations. (b) Distribution of the value of the positive (blue) wildtype divided by E_T. (c) Distribution of the selected SMEs $\tilde{h}$. (d) Distribution of $E_C/E_{\max}^{WT}$.

an elastic network with 500 edges. To check the consistency of our results we also perform simulations varying the length L of the binary sequence. In particular we study $L = 100$ and $L = 1000$ as well. We focus on the same parameters used in the main text, namely a Gaussian with $\mu = 0$, $\sigma = 1.0$ and we determine p, E_T fixing $\langle M\rangle \approx 8$. As expected, the required value of E_T increases with L, because the maximum of the h_i increases proportionally to $\sqrt{\log L}$ as dictated by extreme value statistics. The growth is slow enough, however, and it is compensated by the tuning of E_T, in such a way that the results discussed in the main text are not significantly affected by the choice of the system size. In Fig. S5, panels (a) and (b), we plot the distribution of the number of mutations and of the value of E_B^{WT}/E_T, respectively. We see that, with an appropriate choice of E_T and p for each L, the results are almost indistinguishable for the three values of L. In panels (c) we plot the distribution of the values of h_i selected by the tuning procedure. We see that, as L increases, the peak of strongly beneficial (or strongly deleterious) mutations shifts towards larger absolute values, sharpening the separation between greedy and random steps. Finally, in panel (d) we show the distribution of $E_C/E_{\max}^{WT}$, which is once again unaffected by the choice of L.

C. Heavy tailed distribution of single-mutation effects

The analysis of experimental data for SMEs suggests that the initial part of the distribution is described by a heavy-tailed Pareto distribution, which is then cutoff at large $|\Delta E|$, in part because of biological reasons (the effect of a mutation on fitness cannot be too large), in part due to the limited range of fitness variations that can be measured in experiments.

Yet, it is interesting to check what would happen in our model if the heavy tail would be allowed to continue up to infinity (keeping in mind that the finite value of L will induce a cutoff anyways). Hence, we show here some results obtained for various choices of the input distri-

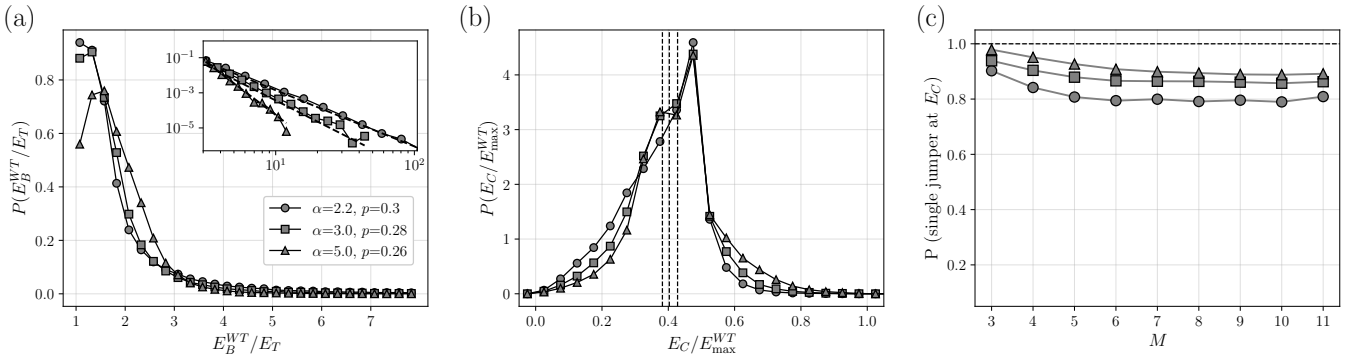

Figure S6. Distribution of the relevant quantities for the model with Pareto distributed h values. Results are shown for different exponents $\alpha = 2.2$, 3.0, and 5.0. (a) Distribution of E_B^{WT}/E_T. Inset: tails of the distributions in log-log scale. The dashed lines are obtained fitting with an exponent equal to α. (b) Distribution of the value of $E_C/E_{\max}^{WT}$. (c) Fraction of single jumpers at E_C as a function of the number of mutations M.

bution $P(h)$ with Pareto tails decaying according to an exponent α as in Eq.(5) of the main text, without cutoff. We analyze different values of $\alpha = 2.2, 3, 5$, for which the variance of the distribution is finite. We fix the system size to $L = 500$ as in the main text, to facilitate comparison. We then proceed as described in the main text and we characterize the statistics of E_f^{WT}/E_T, $E_C/E_{\max}^{WT}$, and the topologies that arise from this input distribution of SMEs. The results are shown in figure S6, for the optimal choice of p and E_T.

For this choice of the parameters we find $\langle E_C/E_{\max}^{WT}\rangle \simeq 0.5$ which is the largest possible value of E_C maximizing the fraction of bottleneck topologies in the ensemble. We note that if there is only one greedy step, the value of the wild type E_B^{WT} is given by the sum of $M-1$ values drawn from $P(h)$ and a value drawn from $P_{\max}^L(h_{\max})$, the probability distribution of the maximum value of the L variables h, which is easily derived from extreme value statistics. If $P(h)$ has heavy tails, we expect the sum to be dominated by the maximum and, as a consequence, $P(E_B^{WT}) \sim P_{\max}^L(h_{\max})$. The same reasoning works for the red wild type E_R^{WT}. In Fig. S6a we show the plot of E_B^{WT} for different values of the exponent α. In the inset of this panel one can see how the tails of such distribution decay following a power law with the same exponent α, as expected.

The distribution of $E_C/E_{\max}^{WT}$, shown in Fig. S6b, seems to be rather independent of the exponent α, and is centered around 0.5 as observed in the previous cases. What is different is the fraction of single jumpers at E_C, shown in Fig. S6c. At variance with the Gaussian and Pareto cutoff cases, here the fraction of bottleneck topologies is different from one even for the optimal values of p and E_T. This difference is due to the fat tails distribution of the wild type phenotype E_f^{WT}. Since the distribution is very broad, it is possible to have a very unbalanced situation in which one of the two wild types is much larger than the other (in particular, when $E_{\max}^{WT} \gtrsim 2E_{\min}^{WT}$), which allows for the saturation of E_C to $E_{\min}^{WT}$ and the absence of a bottleneck. As expected, this effect disappears for larger values of α, when the tails are less pronounced, and absent in the Gaussian and Pareto cutoff distributions.