# Hierarchical Distribution of Ascending Slopes, Highlands, Nearly Neutral Networks, and Local Optima at the $d$th order in an NK Fitness Landscape

Takuyo Aita[1,*]

[1] *Graduate School of Science and Engineering, Saitama University, Saitama 338-8570 Japan*

[*] corresponding author; e-mail address: taita@mail.saitama-u.ac.jp

June 6, 2008

## Abstract

We obtained several structural features of an NK fitness landscape by analytical approach. Particularly, we focused on spatial distributions of "ascending slopes", "highlands", "nearly neutral networks" and "local optima" along the fitness coordinate $W$, from the view point of adaptive walks with step-width $d$, where $d$ is the number of mutated sites (Hamming distance) after a generation. The parameter $k$ governs the degree of the ruggedness on the NK landscape, and we handled cases where $k$ is moderate against the sequence length. From the foot up to the middle region on the landscape, many ascending slopes exist (high evolvability) and these slopes extend up near the "highland", which is mathematically defined as the specific region $W = W_d^*$ where the expectation of the fitness increment becomes zero. Denoting the standard deviation of the fitness change at $W = W_d^*$ by $SD^*$, we considered the existence of "nearly neutral networks", which percolate in the fitness band between $W - SD^*$ and $W + SD^*$. Our results suggest that the highland corresponds to a phase-transition threshold of the formation of the nearly neutral networks. Near or over the highland, "local optima at the $d$th order" appear drastically (low evolvability), where $d$ means the radius of their basins. The value of $W_d^*$ increases with $d$ increasing. Then, as the fitness (=altitude) becomes higher, the basin size of the local optima increases. This leads to a conclusion that it is very hard or impossible for walkers with step-width $d$ to reach near the global peak when $d$ is a realistic large value: $d = 1 \sim 6$, and suggests that the region over the middle in real landscapes may be considerably smooth with small $k$-values to maintain high evolvability.

## Keywords

Fitness landscape; Sequence space; Adaptive walk; NK model; *In vitro* evolution; Evolvability

# 1 Introduction

Biological evolution process is comprehended as the "adaptive walk" or "hill-climbing" on a fitness landscape in genotype space (sequence space). The concept of which was first proposed by Sewall Wright (Wright,1932), and has been developed in evolutionary biology (Maynard-Smith,1970; Eigen,1992; Kauffman, 1993; Gavrilets, 2004). In the field of *in vitro* molecular evolution, that handles artificial evolution of protein or DNA sequences, the concept of "fitness" is extended to a molecular physicochemical property such as propagation rate (Eigen,1985), enzymatic activity, binding affinity or thermostability (Arnold, 2000; Matsuura & Yomo, 2006). Thus, the fitness landscape is regarded as the "evolutionary attribute" of biopolymers. Many theoretical studies have been done based on various mathematical models of the landscapes (Voigt *et al.*, 2000; Gavrilets, 2004). One of the most familiar models is the NK model (Kauffman & Weinberger, 1989; Kauffman, 1993). The NK model is a mathematical model describing a complex system in which an arbitrary element is affected by other $k$ elements. For a protein, an amino acid site corresponds to the element in the NK model. If the $k$ sites that cooperatively affect the $j$th site are located near the $j$th site, this model is called the "adjacent neighbor (NK) model", whereas, if the $k$ sites are located randomly through the sequence, this model is called the "random neighbor (NK) model" (Kauffman, 1993). The fitness landscape constructed by the NK model is called the "NK (fitness) landscape". The NK landscape with $k = 0$ has a single peak, whereas the landscape becomes more rugged and has more local peaks with $k$ increasing.

The original NK model was proposed by Kauffman & Levin (1987) and many slightly modified models have been studied (Barnett, 1998; Newman & Engelhardt, 1998; Iguchi *et al.*, 2005). In almost all cases, they have adopted a binary sequence space, where the number of available letters is two (0 or 1). We denote that by $\lambda = 2$. Many theorists have been interested in such issues as: the height of the global peak and local peaks, the number of local peaks, and the walk length from a random point. These issues were investigated numerically by computer

simulation in most studies (e.g. Kauffman,1993). The first analytical study was conducted by Weinberger (1991). Recently, several rigorous analytical studies were conducted by Evans & Steinsaltz (2002), Durrett & Limic (2003), and Limic & Pemantle (2004), in which they adopted the adjacent neighbor model with $\lambda = 2$. Iguchi *et al.* (2005) investigated the effect of a scale-free network of inter-sites interaction on the properties of the NK landscape. The NK model has also been adopted to study the neutrality on the evolution or neutral evolution (Barnett, 1998; Ohta, 1998; Newman & Engelhardt, 1998).

Meanwhile, several studies tried to estimate the $k$-value in real landscapes by fitting the NK model to experimental data. The $k$ value seems inherent in the physicochemical property of individual biopolymers. Kauffman and Weinberger applied the NK model to affinity maturation of the V region in immunoglobulin and estimated that $k$ is about 40, from the view point of the number of steps of adaptive walk up to the local optima (Kauffman & Weinberger, 1989; Kauffman, 1993). Fontana *et al.* (1993) examined the RNA free energy landscape by computer experiments, and estimated $k = 7 - 8$ for this landscape in terms of autocorrelation on the landscape. In our previous paper (Hayashi *et al.*, 2006; Aita *et al.*, 2007), we demonstrated that the experimental data in an *in vitro* molecular evolution can be explained quantitatively by our adaptive-walk theory on an NK landscape, which was defined by the random neighbor (NK) model. Hayashi and coworkers carried out *in vitro* molecular evolution beginning with a defective fd phage carrying a random polypeptide (139 a.a.) in place of the g3p minor coat protein D2 domain, which is essential for phage infection. Through 20 mutation-selection cycles, the random polypeptide evolved gradually up to the middle region of the fd-phage infectivity landscape in the free energy scale (Hayashi *et al.*, 2006). Our analysis of the experimental data suggested $k \approx 27$ for this landscape (Aita *et al.*, 2007).

Therefore, in light of the great importance of the NK landscapes, we examined analytically structural features of the NK landscape ($\lambda = 20$, random neighbor model) from the view point of an adaptive walker with step-width $d$. Particularly, we were concerned with spatial

distributions of the local optima at the $d$th order, where $d$ means the radius of their basins. Almost previous studies considered cases where $d = 1$, because they handled natural evolution in which mutation occurs infrequently. We want to make a note that our interest is not in natural evolution but in the *in vitro* molecular evolution, which can control the number of mutated sites (Hamming distance) $d$ after a generation and population size. In addition to this, our originality lies in that we analytically obtained the local fitness distribution over all possible $d$-fold point mutants generated from a reference sequence with fitness $W$ (Aita *et al.*, 2007). Many statistical properties of the landscape were derived from the local fitness distribution. In this paper, we will refer to the spatial distributions of "ascending slopes", "highlands", "nearly neutral networks", and "local optima" along the fitness coordinate.

## 2 Model of the NK fitness landscape

We consider all conceivable amino acid sequences with a chain length of $\nu$, and $\lambda$ letters are available at every site, where $\lambda$ is large enough to satisfy $(\lambda - 1)/\lambda \approx 1$. Then, each sequence is mapped into the corresponding point in the $\lambda$-valued $\nu$-dimensional sequence space. The fitness $W$ for a given sequence "$\mathrm{A}_1\mathrm{A}_2\cdots\mathrm{A}_\nu$" is defined by

$$W = \sum_{j=1}^{\nu} w_j(\mathrm{A}_j|\mathrm{A}_{j_1}, \mathrm{A}_{j_2}, \cdots, \mathrm{A}_{j_k}), \tag{1}$$

where $w_j(\mathrm{A}_j|\mathrm{A}_{j_1}, \mathrm{A}_{j_2}, \cdots, \mathrm{A}_{j_k})$ is the "site-fitness", *i.e.* a fitness contribution from a particular letter $\mathrm{A}_j$ at the $j$th site when the $k$ sites $\{j_1, j_2, \cdots, j_k\}$ are occupied by the particular letters $\{\mathrm{A}_{j_1}, \mathrm{A}_{j_2}, \cdots, \mathrm{A}_{j_k}\}$. The $k$ sites $\{j_1, j_2, \cdots, j_k\}$ are randomly chosen from all $\nu - 1$ sites except the $j$th site ("random neighbor model"). The assignment of site-fitness values is modeled as follows: with a set of letters $\{\mathrm{A}_{j_1}, \mathrm{A}_{j_2}, \cdots, \mathrm{A}_{j_k}\}$ given, a site-fitness value of an arbitrary letter $a$ (e.g., $a$ =Ala,Cys,$\cdots$,Tyr) for each site is randomly once assigned from the following set of $\lambda$

values ("quenched model"):

$$w_j(a|\mathrm{A}_{j_1}, \mathrm{A}_{j_2}, \cdots, \mathrm{A}_{j_k}) \in \left\{ \epsilon \left( 1 - \frac{2i}{\lambda - 1} \right) \middle| i = 0, 1, 2, \cdots, \lambda - 1 \right\}, \tag{2}$$

where $\epsilon$ is a positive constant: $\epsilon > 0$. We do not allow the degeneracy of assignment, that is, $w_j(a|\cdots) \neq w_j(a'|\cdots)$ for $a \neq a'$. Therefore, the underlying density function of site-fitness $w$ at each site is given by the comb function:

$$\frac{1}{\lambda} \sum_{i=0}^{\lambda-1} \delta \left( w - \epsilon \left( 1 - \frac{2i}{\lambda - 1} \right) \right), \tag{3}$$

where $\delta(x)$ is the Dirac's delta function. The reason why we fix the site-fitness distribution as shown in eqn. (3) lies in analytical tractability demonstrated in Appendix C [1]. We note that our theoretical conclusion is robust to the shape of the site-fitness distribution (Limic & Pemantle, 2004) and also robust to the site-dependence of $\epsilon$-values (Aita *et al.*, 2004). From eqn. (3), we can see that the mean of the site-fitness values over $\lambda$ letters is equal to zero, while the variance, denoted by $\sigma^2$, is given by

$$\sigma^2 \approx \frac{\epsilon^2}{3}, \tag{4}$$

which is a well-known property of the discrete uniform distribution.

The fitness landscape resulting from this model is called the "NK landscape", although there are several differences from the original NK landscape (Kauffman & Weinberger, 1989; Weinberger, 1991; Kauffman, 1993). In the case of $k = 0$, the fitness landscape has a single peak. As the $k$-value increases, the fitness landscape becomes more rugged. In spite of the ruggedness, there are so many slopes from the foot to the middle region, whereas, local peaks with a large basin size are likely to appear at high altitudes on the landscape. In this paper, we focus on the cases where $k$ is moderate or large, and describe these features quantitatively from an analytical approach.

[1] In more familiar variants of the NK model, site-fitness values are assigned from a continuous (uniform or normal) distribution.

In this paper, we describe several landscape properties, such as the occurrence probability of local optima, along one-dimensional fitness coordinate $W$. The probability density of the fitness $W$ over all possible sequences is given approximately as the following normal distribution with the mean 0 and variance $\mathscr{V}$:

$$\mathscr{N}(W|\,0,\,\mathscr{V}) \qquad \text{for } -H \leqq W \leqq H, \tag{5}$$

where

$$H \equiv \epsilon\,\nu \qquad \text{and} \qquad \mathscr{V} = \sigma^2\,\nu \approx \frac{\epsilon^2\nu}{3} \tag{6}$$

(derivation is shown in Appendix A). The mean of fitness over the whole sequence space corresponds to the "foot" of the landscape, while regions where $W < 0$ corresponds to an "undersea" and is negligible for the adaptive walks that start from random points, which are likely to be at the foot. Since the fitness at the global peak takes about $H(=\epsilon\nu)$ [2], then $H$ corresponds to the height of the landscape from the foot to the global peak. In this paper, we focus on the regions from the foot to the global peak: $0 \leqq W \leqq H$.

In the NK model mentioned above, an arbitrary single-point mutation causes the changes in site-fitness at about $1+k$ sites, that include both of the mutated site and other $k$ sites affected by the mutation. That is, the values of $w_j$'s at about $1+k$ sites are randomly re-assigned from eqn. (2) as a result of the mutation. Random $d$-fold point mutations cause the changes in site-fitness at about $d(1+k)$ sites. Let $\mathscr{D}$ be the mean number of sites that change their site-fitnesses as a result of random $d$-fold point mutations. $\mathscr{D}$ is given by

$$\mathscr{D} = \nu - (\nu - d)\left(1 - \frac{k}{\nu - 1}\right)^d \tag{7}$$

$$\approx d(1+k) \qquad \text{for } d\,k \ll \nu. \tag{8}$$

The derivation of eqn. (7) is shown in Appendix A in Aita *et al.*, (2007)[3]. $\mathscr{D}$ is plotted as a function of $d$ in Fig.1(a). $\mathscr{D}$ is one of the important parameters in the theory described below.

[2] This is not necessarily guaranteed for $k \geqq 1$.

[3] In this paper, we wrote "$\overline{d_{\rm eff}}$" as $\mathscr{D}$.

## 3 Local fitness distribution around a reference sequence

We will begin by considering the set of all sequences at distance $d$ from a reference sequence in the sequence space (Fig.2). That is, the set consists of all conceivable $d$-fold point mutants generated from the reference sequence. According to Fontana and Schuster (1998), we designate the set as the "$d$-boundary" of the reference sequence. The size of the $d$-boundary is denoted by $N_d^\dagger$:

$$N_d^\dagger \equiv \binom{\nu}{d}(\lambda-1)^d, \tag{9}$$

while we use $N$ as a parameter that represents the number of offsprings generated in the $d$-boundary: $N \leqq N_d^\dagger$. Let $W$ be the fitness of the reference sequence, and $\Delta W$ be the change in fitness from the reference sequence to its arbitrary mutant sequence in the $d$-boundary. The conditional probability density of $\Delta W$ with a given $W$, denoted by $\mathcal{P}_d(\Delta W|W)$, is approximately given by the following normal distribution ($\mathbf{E}_d[\Delta W|W]$ and $\mathbf{V}_d[\Delta W|W]$ represent respectively the conditional expectation and conditional variance of $\Delta W$ with a given $W$):

$$\begin{aligned}
\mathcal{P}_d(\Delta W|W) &= \mathscr{N}\left(\Delta W \,|\, \mathbf{E}_d[\Delta W|W], \mathbf{V}_d[\Delta W|W]\right) && (10)\\
\mathbf{E}_d[\Delta W|W] &= -\overline{w}\mathscr{D} && (11)\\
&= -\frac{W\mathscr{D}}{\nu} && (12)\\
\mathbf{V}_d[\Delta W|W] &\approx \left(\sigma^2 + \widetilde{\sigma}^2\left(1-\frac{\mathscr{D}-1}{\nu-1}\right)\right)\mathscr{D} && (13)\\
&= \left(\mathscr{V} + \widetilde{\sigma}^2\nu\left(1-\frac{\mathscr{D}-1}{\nu-1}\right)\right)\frac{\mathscr{D}}{\nu}, && (14)
\end{aligned}$$

where, for eqn. (14), $\mathscr{V}$ is the variance of fitness over all possible sequences in the sequence space (see eqn. (6)), and for equations (11) and (13),

$$\begin{aligned}
\overline{w} &\equiv \frac{1}{\nu}\sum_{j=1}^{\nu} w_j = \frac{W}{\nu} && (15)\\
\widetilde{\sigma}^2 &\equiv \frac{1}{\nu}\sum_{j=1}^{\nu} w_j^2 - \overline{w}^2, && (16)
\end{aligned}$$

where $w_j$ represents the site-fitness value at the $j$th site of the reference sequence: $\sum_{j=1}^{\nu} w_j = W$. The derivation of eqns (10)- (14) is shown in Appendix B.

Exactly, $\mathbf{V}_d[\Delta W|W]$ in eqn. (14) depends on $\widetilde{\sigma}^2$, that is the variance of site-fitness values over all sites for the reference sequence. In Appendix C, we deduced the most probable distribution of site-fitness values over all sites when mean site-fitness $\overline{w}$ is fixed, using the maximal entropy assumption. In conclusion, the most probable distribution is exponential distribution, which was confirmed by computer simulation. Therefore, we obtained the most probable $\widetilde{\sigma}^2$ as a function of $\overline{w}$ (Fig.3). Unfortunately, this function is implicit and does not allow us to make a further analysis. Then, we adopt the following empirical function for a further analysis:

$$\widetilde{\sigma}^2 \approx \sigma^2 \left(1 - \left(\frac{\overline{w}}{\epsilon}\right)^2\right) \tag{17}$$

(Fig.3). Then, $\mathbf{V}_d[\Delta W|W]$ (eqn. (13) and eqn. (14)) is uniquely determined as a function of fitness $W$ (or mean site-fitness $\overline{w}$). Fig.4 shows values of $\mathbf{E}_d[\Delta W|W]$ and $\mathbf{V}_d[\Delta W|W]$ against fitness $W$ of a reference sequence. The observed values through computer simulation agree well with the analytically derived ones. There is, however, a discrepancy between the analytical standard deviation and simulated one when fitness $W$ takes high values. This discrepancy seems to originate from the approximation shown in eqn. (17).

In the case where $k$ is large, $\mathcal{P}_d(\Delta W|W)$ in eqn. (10) is approximately an "independent identical distribution (i.i.d.)" for assignment of $\Delta W$ value to any sequence in the $d$-boundary, and is not i.i.d. when $k$ equal to zero (Appendix B).

Next, we consider the density function of $\Delta W$ over the $d$-boundary when $W$ is fixed, denoted by $\psi_d(\Delta W|W)$, assuming that a value of $\Delta W$ for any sequence in the $d$-boundary is randomly and independently assigned from the probability density $\mathcal{P}_d(\Delta W|W)$. Therefore, the density $\psi_d(\Delta W|W)$ can be approximately given by

$$\psi_d(\Delta W|W) \approx \begin{cases} \mathcal{P}_d(\Delta W|W) & \text{for } \Delta W_{\min} \leqq \Delta W \leqq \Delta W_{\max}, \\ 0 & \text{otherwise,} \end{cases} , \tag{18}$$

where $\Delta W_{\max}$ (or $\Delta W_{\min}$) is the maximal (or minimal) value of $\Delta W$ over the $d$-boundary, that is, $\Delta W_{\max}$ is the change in fitness from the reference sequence with fitness $W$ to the fittest sequence in its $d$-boundary. Several examples of $\psi_d(\Delta W|W)$ are shown schematically in Fig.2.

We are interested in the probability density of $\Delta W_{\max}$, because the $\Delta W_{\max}$ value determines whether the reference sequence is a local optimum or not. Note that $\Delta W_{\max}$ for the $d$-boundary should be written as $\Delta W_{\max}^{(d)}$. We write simply $\Delta W_{\max}$, but write $\Delta W_{\max}^{(d)}$ when necessary, hereafter. Based on the theory of extreme value density, the probability density of $\Delta W_{\max}$ is given by:

$$N_d^\dagger \, \mathcal{P}_d(\Delta W_{\max}|W) \left( \int_{-\infty}^{\Delta W_{\max}} \mathcal{P}_d(x|W) \, \mathrm{d}x \right)^{N_d^\dagger - 1}, \tag{19}$$

where $N_d^\dagger$ is defined in eqn. (9). The expectation and variance of $\Delta W_{\max}$ based on this density function are respectively given by

$$\mathbf{E}_d[\Delta W_{\max}|W] = \mathbf{E}_d[\Delta W|W] + \sqrt{\mathbf{V}_d[\Delta W|W]} \times \zeta_{\max} \tag{20}$$

$$\mathbf{V}_d[\Delta W_{\max}|W] \approx \frac{\mathbf{V}_d[\Delta W|W]}{{\zeta_{\max}}^2 + 1}, \tag{21}$$

where $\zeta_{\max}$ is defined as follows. Consider that $N$ random numbers are chosen according to the standard Gaussian probability density $\mathcal{N}(x|0,1)$, and let $x_m$ be the $m$th greatest number among the $N$ numbers. We denote the expectation of $x_m$ by $\zeta_m(N)$ $(\equiv \mathbf{E}[x_m])$ as a function of $N$. The $\zeta_m(N)$ is approximately given by transforming $N$ via:

$$\zeta_m \exp\left(\frac{\zeta_m^2}{2}\right) \int_{-\infty}^{\zeta_m} \exp(-\frac{s^2}{2}) \mathrm{d}s = \frac{N}{m} - 1. \tag{22}$$

$\zeta_{\max}$ is defined as $\zeta_{\max} \equiv \zeta_1(N_d^\dagger)$. Eqns (20)-(22) were derived in a manner similar to those shown in Appendix B in Aita *et al.*, (2007). For many cases, $N_d^\dagger$ takes a huge value, then $\zeta_{\max}$ is roughly given by the following explicit form:

$$\zeta_{\max} \approx \sqrt{2 \ln \frac{N_d^\dagger}{\sqrt{2\pi}}}. \tag{23}$$

$\zeta_{\max}$ is plotted as a function of $d$ in Fig.1(b). Fig.4 shows values of $\mathbf{E}_d[\Delta W_{\max}|W]$ and $\mathbf{V}_d[\Delta W_{\max}|W]$ against fitness $W$ of a reference sequence. The observed values through com-

puter simulation agree well with the analytically derived ones where fitness $W$ takes low values, while the discrepancy between analytical one and simulated one becomes larger as $W$ becomes higher. Obviously, this discrepancy stems from the discrepancy for $\mathbf{V}_d[\Delta W|W]$.

# 4 Structural features of the NK fitness landscape

In this section, we consider the structural features of the NK fitness landscape from the view point of an adaptive walker with step-width $d$. In this paper, the walker on the landscape corresponds to the reference sequence, and the walker moves toward the highest point in the $d$-boundary. We refer to the following three elements: ascending slopes, highland and local optima.

## 4.1 Autocorrelation function

Autocorrelation function of a fitness landscape is defined by

$$\mathscr{A}(d) \equiv \frac{\langle W(\mathbf{A})W(\mathbf{B})\rangle_{d(\mathbf{A},\mathbf{B})=d} - \langle W\rangle^2}{\mathscr{V}}, \tag{24}$$

where $W(\mathbf{A})$ is the fitness of an arbitrary sequence $\mathbf{A}$ (Fontana *et al.*,1993). $\mathbf{A}$ and $\mathbf{B}$ are variable under the restriction that the Hamming distance between them is $d$ ($d(\mathbf{A},\mathbf{B}) = d$). The average $\langle\cdots\rangle$ is conducted over the whole sequence space. Applying the definition to the NK landscape, we obtain

$$\begin{aligned}
\mathscr{A}(d) &= \frac{\langle W \int_{-\infty}^{\infty}(\Delta W + W)\,\psi_d(\Delta W|W)\;\mathrm{d}\Delta W\rangle}{\mathscr{V}} && (25)\\
&= \frac{\langle W\,\mathbf{E}_d[\Delta W|W] + W^2\rangle}{\mathscr{V}} && (26)\\
&= 1 - \frac{\mathscr{D}}{\nu} && (27)\\
&= (1-\frac{d}{\nu})\left(1-\frac{k}{\nu-1}\right)^d. && (28)
\end{aligned}$$

Eqn. (27) is derived by substituting eqn. (12) into eqn. (26). Eqn. (28) is the same form as that shown in Table I of Fontana *et al.* (1993) and that in Weinberger (1991), in which they

handled binary ($\lambda = 2$) sequence space. Our result implies that autocorrelation function is independent of the $\lambda$ value.

### 4.2 Ascending slopes and highland

Consider a reference sequence with fitness $W$ and its $d$-boundary in sequence space. If at least one sequence in the 1-boundary has a fitness value larger than $W$, namely, if $\Delta W^{(1)}_{\max} > 0$, then we say that the reference sequence is located on an "ascending slope". When substituting eqn. (12) and (14) (with eqns (15) and (17)) into eqn. (20), we can see that eqn. (20) is the monotonically decreasing function of $W$ for $0 \leqq W \leqq H$ (see Fig.4 hereafter). In this paper, $W$ is considered as the fitness of a walker. When $W$ takes a low value around 0 (as the mean fitness over the sequence space), that is the case where the walker is located on the foot of the landscape, $\mathbf{E}_d[\Delta W_{\max}|W]$ takes a large positive value. This means there are many steep ascending slopes in this region. As the walker climbs the landscape with $W$ increasing, $\mathbf{E}_d[\Delta W_{\max}|W]$ decreases gradually and finally the walker reaches a specific region $W = W^*_d$, where $W^*_d$ is defined as a specific $W$-value satisfying

$$\mathbf{E}_d[\Delta W_{\max}|W^*_d] = 0. \tag{29}$$

By substituting $\mathbf{E}_d[\Delta W_{\max}|W] = 0$ into eqn. (20), in which $\mathbf{E}_d[\Delta W|W]$ is given in eqn. (12) and $\mathbf{V}_d[\Delta W|W]$ is given in eqn. (14) with eqns (15) and (17), and solving the resulting quadratic equation for $W$, we can approximately determine the value of $W^*_d$ as follows:

$$W^*_d = \frac{\sqrt{\mathscr{V}\,\nu}}{\kappa(\tau^\dagger)\,\tau^\dagger} \tag{30}$$

$$\tau^\dagger \equiv \frac{\sqrt{\mathscr{D}/2}}{\zeta_{\max}} \tag{31}$$

$$\kappa(x) \equiv \sqrt{\frac{2 + (1 - \frac{\mathscr{D}-1}{\nu-1})(\frac{\sigma}{\epsilon})^2/x^2}{2 - \frac{\mathscr{D}-1}{\nu-1}}} \tag{32}$$

$$\approx \begin{cases} \sqrt{1 + (\frac{\sigma}{\epsilon})^2/2x^2} & \text{for } \mathscr{D} \ll \nu \\ \sqrt{2} & \text{for } \mathscr{D} \approx \nu\ . \end{cases} \tag{33}$$

Fig.5 shows analytical values of $W_d^*$ $(d = 1, 2, \cdots, 6)$ against various $k$-values. We can confirm that the value of $W_d^*$ becomes larger with $d$ increasing for any $k$ value shown in the figure. However, this has not been proven exactly.

We designate the region where $W = W_d^*$ as the "highland" at the $d$th order. The ascending slopes lead to the highland. Roughly, it is almost guaranteed that adaptive walkers can climb up near highland without getting trapped in local optima. The highland is a mathematically characteristic region on the fitness landscape. For actual fitness landscapes of real biopolymers, it is meaningful for the highland to have a width of fitness. Then, we redefine the "highland" (at the $d$th order) as the region in which the fitness $W$ satisfies the conditions

$$W_d^* - SD^* \leqq W \leqq W_d^* + SD^*, \tag{34}$$

where

$$SD^* \equiv \sqrt{\mathbf{V}_d[\Delta W_{\max}|W_d^*]} \approx \frac{2\sigma\tau^\dagger}{\kappa(\tau^\dagger)} \tag{35}$$

(eqn. (35) is derived by combining eqn. (12), (20)-(21), (29)-(31) and with ${\zeta_{\max}}^2 + 1 \approx {\zeta_{\max}}^2$).

### 4.3 Nearly neutral networks

According to Gavrilets (2004), the region in which the fitness $W$ satisfies the conditions $W_1 < W < W_2$ is called the "$(W_1, W_2)$-fitness band". Regarding all the sequences that belong to the fitness band as the "nearly neutral sequences", he referred to the existence of "percolating nearly neutral networks" (or "giant components") in the fitness band, where the nearly neutral network is defined as a contiguous set of the nearly neutral sequences (Gavrilets, 2004). The highland we defined is a specific fitness band in terms of the percolation in sequence space. We will mention the reason in the following.

We carried out a computer simulation of random walks with step-width $d$ in the following protocol. Giving a reference sequence with fitness $W$ as a starting point, we conducted random walks with step-width $d$ (we set $d = 1$ or 2) in the region in which fitnesses take values between

$W - z \times SD^*$ and $W + z \times SD^*$ until the maximal Hamming distance from the initial sequence to the walker cannot be updated, where parameter $z$ controls the width of the fitness band. Let $HD_{\max}$ be the maximal Hamming distance. If $HD_{\max} = \nu$, independent of any initial sequence, we can guess that there are percolating nearly neutral networks in this region, in which sequences are connected at intervals of the Hamming distance $d$. In this paper, we designate the network as the nearly neutral network at the $d$th order. Fig.6 shows observed values of $HD_{\max}$ against $W$ values, for the case of $d = 1$. There seems to be two phases, $HD_{\max}/\nu = 1$ and $HD_{\max}/\nu = 0 \sim 0.2$, and the average of $HD_{\max}/\nu$ was about 0.5 at $W = W_1^*$ for the case where $z = 1$. This means that the phase transition for the case where $z = 1$ occurred at around $W = W_1^*$. The percolation threshold was lower (or higher) when $z$ decreased (or increased). We also observed similar features for the case where $d = 2$ (data not shown).

These results from the simulation study were supported theoretically as follows. Let us consider a reference sequence that takes the stationary fitness value $W_d^*$, and its $d$-boundary. Concentrating on the number of nearly neutral sequences (in the $d$-boundary) that belong to the fitness band defined as $W_d^* - z \times SD^* \leqq W \leqq W_d^* + z \times SD^*$ (see eqn.(34)), we calculated the average number analytically (Appendix D). Particularly, this average number for $d = 1$ is called "the average degree of (nearly) neutrality" (Huynen *et al.*, 1996; Barnett, 1998). As a result, the average number was 1.4 for $z = 0.5$, 2.4 for $z = 1.0$, 4.0 for $z = 1.5$ and 6.7 for $z = 2.0$. Interestingly, it turned out that this average number is sensitive to $z$ but almost insensitive to other parameters, such as $d$, under our conditions (the details are in Appendix D). These analytical values shown above almost agree with those obtained by another computer simulation (Table 1). The critical average number that governs the percolation threshold is expected to be $2 \sim 3$ (see Fig. 4.2 in Gavrilets, (2004)). These numbers are almost equal to 2.4 (for $z = 1.0$) shown above, suggesting that the phase transition occurs at around the highland defined in eqn. (34), that is, the highland is approximately the upper limit of the existence of

nearly neutral networks for $z = 1.0$, or more precisely the collapse point for the networks.

The neutrality on the evolution and (nearly) neutral networks have been studied by using modified NK model (Ohta, 1998; Newman & Engelhardt, 1998; Barnett, 1998; Gravner *et al.*, 2007). In the "Discussion" section, we refer to the relation between these studies and ours.

### 4.4 Local optima and local mountains

Let $p_d(W)$ be the probability that the fitnesses of all the sequences in the $d$-boundary are less than the fitness $W$ of the reference sequence. $p_d(W)$ is a function of $W$ and is given by

$$p_d(W) = \left( \int_{-\infty}^{0} \mathcal{P}_d(\Delta W|W)\, \mathrm{d}\Delta W \right)^{N_d^\dagger} \tag{36}$$

$$= \left( \int_{-\infty}^{-\mathbf{E}_d[\Delta W|W]/\sqrt{\mathbf{V}_d[\Delta W|W]}} \mathscr{N}(x|0,1)\mathrm{d}x \right)^{N_d^\dagger}. \tag{37}$$

Let $W_d(p)$ be the inverse function of $p_d(W)$. It is more convenient for the mathematical analysis of local optima to handle $W_d(p)$ rather than $p_d(W)$. Thus, to derive $W_d(p)$, we introduce the following function of the probability $p$, $\xi(p) = \xi$:

$$\left( \int_{-\infty}^{\xi} \mathscr{N}(x|0,1)\mathrm{d}x \right)^{N_d^\dagger} \equiv p \qquad (0 \leqq p \leqq 1). \tag{38}$$

By comparing eqn. (38) with (37), the function $W_d(p)$ is given by the following implicit form:

$$\frac{-\mathbf{E}_d[\Delta W|W_d(p)]}{\sqrt{\mathbf{V}_d[\Delta W|W_d(p)]}} \equiv \xi. \tag{39}$$

Interestingly, eqn. (39) is the same form as eqn. (20) by replacing $\zeta_{\max}$ in eqn. (20) with $\xi$ and setting the right-hand side in eqn. (20) to be zero. Therefore, by the same manner of derivation of eqn. (30), $W_d(p)$ is determined as follows:

$$W_d(p) = \frac{\sqrt{\mathscr{V}\,\nu}}{\kappa(\tau^\ddagger)\,\tau^\ddagger} \tag{40}$$

$$\tau^\ddagger \equiv \frac{\sqrt{\mathscr{D}/2}}{\xi(p)}, \tag{41}$$

where $\kappa(x)$ is defined in eqn. (32). $\xi(p)$ is approximately given by

$$\xi(p) \approx \sqrt{2 \ln \frac{N_d^\dagger/(-\ln p)}{\sqrt{2\pi}}}, \tag{42}$$

for large $\xi$. Eqn. (42) is derived by substituting $\int_{-\infty}^{\xi} \mathscr{N}(x|0,1)\mathrm{d}x = 1 - \int_{\xi}^{\infty} \mathscr{N}(x|0,1)\mathrm{d}x$ into eqn. (38), and using $\ln(1-\delta) \approx -\delta$ and $\int_{y}^{\infty} \mathscr{N}(x|0,1)\mathrm{d}x \approx \mathscr{N}(y|0,1)$ for $y \gg 2$. The difference between eqn. (40) and eqn. (30) stems from the difference between eqn. (42) and eqn. (23). Then, we can easily see that $W_d(1/e) = W_d^*$. As a result, we found $p_d(W)$ to be a sigmoidal function of $W$, as shown in Fig.7. The $p_d(W)$ represents the probability that the adaptive walker with step-width $d$ gets trapped at altitude $W$, when the walker is put on the altitude. By performing an additional computer simulation of the adaptive walks starting from random sequences, we obtained the frequency distribution of altitude (=fitness) where the walkers got trapped. The frequency distribution was approximately of the Gaussian-type with the following mean $\pm$ standard deviation of fitness $W$: when $\nu = 20$ and $K = 10$, $W = 0.60 \pm 0.05$ for $d = 1$ and $W = 0.67 \pm 0.03$ for $d = 2$; when $\nu = 120$ and $K = 27$, $W = 0.47 \pm 0.03$ for $d = 1$ and $W = 0.51 \pm 0.02$ for $d = 2$. These values almost meet with the raising phase of $p_d(W)$ shown in Fig.7.

Consider a reference sequence with fitness $W$, and the set of all sequences in the 1-, 2-, $\cdots$, $r$-boundary of the reference sequence, that is, the set is mapped to the inner region of the hyper-ball with the radius of $r$ centering at the reference sequence. We designate the hyper-ball as the "$r$-ball" of the reference sequence. If all the sequences (except the reference sequence) in the $r$-ball have fitness values less than $W$ and there is, in the $(r+1)$-boundary, at least one sequence with fitness greater than $W$, we designate the reference sequence as the "local optimum at the $r$th order", and designate the $r$-ball as the "local mountain at the $r$th order". By an analogy to the energy landscape theory, the $r$ should be designated as the "basin size". According to van Nimwegen & Crutchfield (2000), the $r+1$ corresponds to the width of the "fitness barrier". The probability that a certain sequence with fitness $W$ is the local optimum at the $r$th order is approximately

$$P_r(W) \approx (1 - p_{r+1}(W)) \prod_{d=1}^{r} p_d(W)$$

$$= \prod_{d=1}^{r} p_d(W) - \prod_{d=1}^{r+1} p_d(W), \tag{43}$$

where we adopt a mean field approximation, that is we assume that $p_1(W), p_2(W), \cdots, p_{r+1}(W)$ are independent. As special cases where $r = 0$ and $r = \nu$, we define $P_0(W) \equiv 1 - p_1(W)$ and $P_\nu(W) \equiv \prod_{d=1}^{\nu} p_d(W)$. $P_0(W)$ represents the probability that the reference sequence is located on ascending slopes, and $P_\nu(W)$ represents the probability that the reference sequence is the global optimum. It is easy to show that $\sum_{r=0}^{\nu} P_r(W) = 1$. As a result, we found $P_r(W)$ to be a Gaussian-like function of $W$, as shown in Fig.7. To confirm the validity of our analytical prediction, we carried out computer experiments to obtain numerical values of $W_d^*$, $p_d(W)$ and $P_r(W)$ by random sampling. The result is shown in Fig.7(bottom). As a whole, the observed values by computer simulation agree almost exactly with the analytically derived ones. There is, however, a considerable discrepancy between the analytical and simulated ones for $\nu = 120, k = 27$.

A statistical view of the local mountains at the $r$th order is as follows. Fitness distribution over the $d$-boundary of the central point (=reference sequence) obeys eqn. (18). Therefore, a majority of sequences belonging to the local mountain is located at around $W + \mathbf{E}_r[\Delta W|W]$, where $\mathbf{E}_r[\Delta W|W]$ is given in eqn. (12) with $d = r$. Intuitively, the absolute value of $\mathbf{E}_r[\Delta W|W]$ can be interpreted as the "height" of the local mountain:

$$\text{height} = |\mathbf{E}_r[\Delta W|W]| \approx \frac{Wr(1+k)}{\nu}, \qquad \text{for } r(1+k) \ll \nu, \tag{44}$$

where $W$ is the fitness of the local peak. This means, given that $k$ is larger and the local mountain stands at higher altitudes on the global landscape, the local mountain is higher and steeper.

### 4.5 An extensive view of NK fitness landscape

In this subsection, we take an extensive view of the NK landscape, based on concepts of $\mathbf{E}_d[\Delta W_{\max}|W]$ shown in eqn. (20) and $W_d^*$ defined in eqn. (29). In the left corner of Fig.7(top

and middle), several types of the local structural features are shown schematically. For example, Type A represents the ascending slopes, and Type E represents the local mountain at the second order. Thus, Fig.7 shows a hierarchical distribution of ascending slopes, highlands, and local optima at the $d$th order. The dashed vertical lines indicate the values of $W_d^*$ for $d = 1, 2, \cdots$, calculated by eqn. (30). Type X (X=A,B,$\cdots$,F) appears mainly in Region X indicated at the top of each figure. The features of each Region are mentioned below. We refer to the existence of nearly neutral networks, which percolate in the fitness band between $W - SD^*$ and $W + SD^*$, where $SD^*$ is defined in eqn. (35).

**Region A** ($W \ll W_1^*$) : $0 < \mathbf{E}_1[\Delta W_{\max}] < \mathbf{E}_2[\Delta W_{\max}] < \mathbf{E}_3[\Delta W_{\max}] \cdots$. The ascending slopes exist everywhere in this region, and then the walker can continue climbing upward. In addition, there are nearly neutral networks at the $d$th order ($d = 1, 2, 3, \cdots$).

**Region B** ($W = W_1^*$) : $\mathbf{E}_1[\Delta W_{\max}] = 0 < \mathbf{E}_2[\Delta W_{\max}] < \cdots$. We can expect that the highest point in the 1-boundary of a reference sequence takes almost the same fitness value as the reference $W$. Adaptive walkers with step-width $d = 1$ can climb up near this region without getting trapped in local optima. Then, we designate this region as the highland at the first order. The nearly neutral network at the first order becomes broken. Sequences potentially become local optima at the first order with the probability of $P_1(W)$.

**Region C** : $\mathbf{E}_1[\Delta W_{\max}] < 0 < \mathbf{E}_2[\Delta W_{\max}] < \cdots$. The probability $P_1(W)$ takes the maximal value.

**Region D** ($W = W_2^*$) : $\mathbf{E}_1[\Delta W_{\max}] < \mathbf{E}_2[\Delta W_{\max}] = 0 < \mathbf{E}_3[\Delta W_{\max}] \cdots$. This region corresponds to the highland at the second order. Adaptive walkers with step-width $d = 2$ can climb up near this region without getting trapped in local optima. The nearly neutral network at the second order becomes broken. Sequences potentially become local optima at the second order with the probability of $P_2(W)$.

**Region E** : $\mathbf{E}_1[\Delta W_{\max}], \mathbf{E}_2[\Delta W_{\max}] < 0 < \mathbf{E}_3[\Delta W_{\max}] \cdots$. The probability $P_2(W)$ takes the maximal value.

In higher regions, we can explain the emergence of the highlands, nearly neutral networks and local optima at higher orders by expanding the explanation mentioned above.

Returning to Fig.5, we can see that there are no local optima from the foot up to the middle regions on the landscapes for moderate $k$ against $\nu$. This suggests that evolvability is high up to the middle regions on the landscapes, while evolvability becomes drastically lower over the middle. More considerations are given in the "Discussion" section.

## 5 Discussion

We obtained several structural features of an NK fitness landscape by analytical approach. Particularly, we focused on spatial distributions of "ascending slopes", "highlands", "nearly neutral networks" and "local optima" along the fitness coordinate $W$, from the view point of adaptive walks with step-width $d$, where $d$ is the number of mutated sites (Hamming distance) after a generation. The parameter $k$ governs the degree of the ruggedness on the NK landscape, and we handled cases where $k$ is moderate against the sequence length. From the foot up to the middle region on the landscape, many ascending slopes exist (high evolvability) and these slopes extend up near the "highland", which is mathematically defined as the specific region $W = W_d^*$ where the expectation of the fitness increment becomes zero. Denoting the standard deviation of the fitness change at $W = W_d^*$ by $SD^*$, we considered the existence of "nearly neutral networks", which percolate in the fitness band between $W - SD^*$ and $W + SD^*$. Our results suggest that the highland corresponds to a phase-transition threshold of the formation of the nearly neutral networks. Near or over the highland, "local optima at the $d$th order" appear drastically (low evolvability), where $d$ means the radius of their basins. The value of $W_d^*$ increases with $d$ increasing. Then, as the fitness (=altitude) becomes higher, the basin size

of the local optima increases.

The neutrality on the evolution and (nearly) neutral networks have been studied by using modified NK models. Newman & Engelhardt (1998) and Gravner *et al.* (2007) investigated some properties of neutral networks by using a NK model with the discrete uniform distribution (just like eqn.(2)) for site-fitness assignment. A major difference between their model and ours lies in the following. They considered cases where the number of available letters ($\lambda$, in our paper) is small, particularly $\lambda = 2$ in almost cases, whereas we considered cases where $\lambda$ is large enough to satisfy $(\lambda - 1)/\lambda \approx 1$. Thus, in their NK model, the effect of discrete fitness on its landscape structure becomes strong and then the landscape have neutral regions. In our model, the resulting landscape is considered as the approximately continuous one and then we introduced "nearly" neutral regions as a fitness band (Gavrilets, 2004). The reason we handled the discrete fitness model (eqn.(2)) lies in the analytical tractability, and then we do not emphasize the effect of discrete fitness. Essentially, their strong discretization is almost equivalent to our coarse-graining for the nearly neutrality. Intuitively, the width of the fitness band (the width is controled by $z$ in this paper) has the almost same effect on the existence of nearly neutral networks as the discretization of the fitness (the discretization is controled by $\lambda$ in this paper, and $F$ in Newman & Engelhardt (1998)) does. Fig.3 of Newman & Engelhardt (1998) shows that the fitness of the most fit percolating neutral network becomes higher as the $F$-value becomes smaller, whereas Fig.6 in our paper shows that the percolating neutral networks exist up to higher regions as the $z$-value becomes larger. Maybe both tell the almost same things as each other. However, a question remains what criterion of the nearly neutrality is.

A part of this view is compatible with our experimental study we mentioned in our "Introduction" (Hayashi *et al.*, 2006). Hayashi and coworkers carried out *in vitro* molecular evolution beginning with a defective fd phage carrying a random polypeptide (139 a.a.) in place of the g3p minor coat protein D2 domain, which is essential for phage infection. Through 20

mutation-selection cycles cycles, the random polypeptide evolved gradually up to the middle region of the fitness landscape. Our analysis of the experimental data suggested $k \approx 27$ for the fd-phage infectivity landscape in the free energy scale (Aita *et al.*, 2007). According to our results for $\nu = 120$ and $k = 27$ in this paper, adaptive walkers with step-width $d = 1$ can easily climb up to the middle region on the landscape, but gets trapped in a local optimum (at the first order) at around $W = 0.47$ (the global peak is $W = 1$). By enlarging the step-width to be $d = 2$, the walker can escape from it and step upward, but after a few steps, he will get trapped in another local optimum (at the second order) at around $W = 0.51$. As the fitness becomes higher, the basin size of the local optima becomes larger. Our rough calculation suggested that it is impossible for the walkers with a realistic step-width (ex. $d = 1 \sim 6$) to reach the global optimum, for such a large $k$ (data not shown). Recombination is also not expected to surmount them because there are not schema in this random neighbor model (Kauffman,1993; Perelson & Macken,1995).

Several theoretical studies proposed the use of diffusion on neutral networks to reach higher regions (Barnett, 1998; Newman & Engelhardt, 1998; van Nimwegen & Crutchfield, 2000). Barnett (1998) incorporated the neutralily into the original NK landscape. The resulting landscape is called the "NKp landscape" and he examined the structural features of the NKp landscapes. They predict the punctuated emergence of novel sequences with higher fitness through the diffusion. This may be one of the reasons how proteins in the origin of life could evolve to obtain high biological function and stability. Considering that an evolving population surmounts a "fitness barrier" or "entropy barrier" toward a higher region, van Nimwegen & Crutchfield (2000) gave the scaling function for the required time to cross the barrier. However, as Voigt *et al.* (2000) pointed out, the required time scale seems not feasible, with current methods in *in vitro* molecular evolution.

Another speculation is that even if the region under the middle is rugged with large $k$-values, the region over the middle may be considerably smooth with small $k$-values (Aita &

Husimi, in press). Saito *et al.* proposed a scenario of the emergence of the native conformation from a random peptide (Saito *et al.*, 1997). According to them, a random peptide has some functional conformations and innumerable nonfunctional ones. The selection on the local configurations at the active sites locks the lowest-energy conformation in one of the functional conformations. Intuitively, it seems possible that the $k$ value may change before and after the "lock" of conformational framework.

## 6 Acknowledgements

This work was performed as part of the Rational Evolutionary Design of Advanced Biomolecules (REDS) Project, Saitama Prefecture Collaboration of Regional Entities for the Advancement of Technological Excellence supported by JST. I thank Prof. Yuzuru Husimi and Prof. Koichi Nishigaki (Saitama Univ., Japan) for their encouragement and helpful comments.

## References

[1] Aita, T., Morinaga, S. & Husimi, Y. (2004). Thermodynamical Interpretation of Evolutionary Dynamics on a Fitness Landscape in an Evolution Reactor, I. *Bull. Math. Biol.* **66**, 1371-1403.

[2] Aita, T., Hayashi, Y., Toyota, H., Husimi, Y., Urabe, I., Yomo, T, (2007). Extracting characteristic properties of fitness landscape from in vitro molecular evolution: A case study on infectivity of fd phage to E.coli. *J.theor.Biol.* **246**, 538-550.

[3] Aita, T. & Husimi, Y. (in press). Fitting Protein-Folding Free Energy Landscape for a Certain Conformation to an NK Fitness Landscape. *J.theor.Biol.*

[4] Arnold, F.H. (ed), (2000) *Evolutionary Protein Design*, *Advances in Protein Chemistry*, **55**, Academic Press.

[5] Barnett, L. (1998). Ruggedness and neutrality - the NKp family of fitness landscapes. In Adami, C., Belew, R. K., Kitano, H., & Taylor, C. (Eds.), *ALIFE VI, Proceedings of the Sixth International Conference on Artificial Life*, 18-27. ALIFE, The MIT Press.

[6] Durrett, R. & Limic, V. (2003). Rigorous results for the $NK$ model. *Annals of Probability* **31**, 1713-1753.

[7] Eigen, M. (1985). Macromolecular evolution: dynamical ordering in sequence space. *Ber. Bunsenges. Phys. Chem.* **89**, 658-667.

[8] Eigen, M. (1992). *Steps towards life*, Oxford University press.

[9] Evans, S.N. & Steinsaltz, D. (2002). Estimating some features of $NK$ fitness landscapes. *Annals of Probability* **12**, 1299-1321.

[10] Fontana, W. & Schuster, P. (1998). Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. *J.theor.Biol.* **194**, 491-515.

[11] Fontana, W., Stadler, P.F., Bornberg-Bauer, E.G., Griesmacher, T., Hofacker, I.L., Tacker, M., Tarazona, P., Weinberger, E.D. & Schuster, P. (1993). RNA folding and combinatory landscapes. *Phys. Rev. E* **47**, 2083-2099.

[12] Gavrilets, S. (2004). *Fitness landscapes and the origin of species*, Princeton University Press.

[13] Gravner, J., Pitman, D. & Gavrilets, S. (2007) Percolation on fitness landscapes: effects of correlation, phenotype, and incompatibilities. *J.theor.Biol.* **248**, 627-645.

[14] Hayashi, Y., Aita, T., Toyota, H., Husimi, Y., Urabe, I., Yomo, T, (2006) Experimental Rugged Fitness Landscape in Protein Sequence Space. *PLoS ONE* **1**, e96.

[15] Huynen, M.A., Stadler, P.F. & Fontana, W. (1996) Smoothness within ruggedness: the role of neutrality in adaptation. *Proc.Natl.Acad.Sci.USA* **93** 397-401.

[16] Iguchi, K., Kinoshita, S. & Yamada, H. (2005) Rugged fitness landscapes of Kauffman models with a scale-free network. *Phys. Rev. E* **72**, 061901.

Kauffman, S.A. & Levin, S.A. (1987). Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.* **128**, 11-45.

[17] Kauffman, S.A. & Weinberger, E.D. (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. *J.theor.Biol.* **141**, 211-245.

[18] Kauffman, S.A. (1993). *The Origin of Order.* Oxford: Oxford University Press.

[19] Limic, V. & Pemantle, R. (2004). More rigorous results on the Kauffman-Levin model of evolution. *Annals of Probability* **32**, 2149-2178.

[20] Matsuura, T. & Yomo, T. (2006). In vitro evolution of proteins. *J. Biosci. Bioeng.* **101**, 449-456.

[21] Maynard-Smith, J. (1970). Natural selection and the concept of a protein space. *Nature* **225**, 563-564.

[22] Newman, M. E. J. & Engelhardt, R. (1998). Effects of selective neutrality on the evolution of molecular species. *Proc. R. Soc. Lond. B* **265**, 1333-1338.

[23] Ohta, T. (1998). Evolution by nearly-neutral mutations. *Genetica* **102**, 83-90.

[24] Perelson, A.S., & Macken, C.A. (1995). Protein evolution on partially correlated landscapes. *Proc.Natl.Acad.Sci.USA* **92**, 9657-9661.

[25] Saito, S. Sasai, M. & Yomo, T. (1997). Evolution of the folding ability of proteins through functional selection. *Proc.Natl.Acad.Sci.USA* **94**, 11324-11328.

[26] Schuster, P. (1995). How to search for RNA structures: theoretical concepts in evolutionary biotechnology. *J.Biotech.* **41**, 239-257.

[27] Voigt, C.A., Kauffman, S. & Wang, Z.G. (2000). Rational evolutionary design: the theory of *in vitro* protein evolution. *Adv Protein Chem.* **55**, 79-160.

[28] van Nimwegen, E. & Crutchfield, J.P. (2000). Metastable evolutionary dynamics: crossing fitness barriers or escaping via neutral paths? *Bull. Math. Biol.* **62**, 799-848.

[29] Weinberger, E.D. (1991). Local properties of Kauffman's N-k model: A tunably rugged energy landscape. *Phys.Rev.A* **44**, 6399-6413.

[30] Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics* **1**, 356-366.

## Appendix A: Derivation of eqn. (6)

Consider the following sum of fitness over all the sequences of $\lambda^\nu$:

$$S_1 \equiv \sum_{all\ sequences} W = \sum_{all\ sequences} \sum_{j=1}^{\nu} w_j(\mathrm{A}_j|\mathrm{A}_{j_1}, \mathrm{A}_{j_2}, \cdots, \mathrm{A}_{j_k}).$$

In the set of all the sequences, there are $\lambda^{\nu-(1+k)}$ sequences which include a set of letters, $\{\mathrm{A}_j, \mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}$. Then,

$$\begin{aligned} S_1 &= \sum_{j=1}^{\nu} \lambda^{\nu-(1+k)} \sum_{\{\mathrm{A}_j, \mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}} w_j(\mathrm{A}_j|\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}) \\ &= \sum_{j=1}^{\nu} \lambda^{\nu-(1+k)} \sum_{\{\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}} \sum_{a} w_j(a|\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}). \end{aligned}$$

Here, let us focus on

$$\sum_{\{\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}} \sum_{a} w_j(a|\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}).$$

With a set $\{\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}$ fixed, we can see

$$\sum_{a} w_j(a|\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}) = \langle w \rangle \lambda,$$

where $\langle w \rangle$ is the mean site-fitness over $\lambda$ letters and is set to be $\langle w \rangle = 0$ in this paper. This equation holds for all possible $\lambda^k$ states of $\{\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}$. Then

$$S_1 = \nu \lambda^{\nu-(1+k)} \langle w \rangle \lambda \lambda^k = \langle w \rangle \nu \lambda^\nu.$$

Therefore, the mean of fitness over all the sequences is given by

$$\langle W \rangle = \frac{S_1}{\lambda^\nu} = \langle w \rangle \nu = 0.$$

Next, consider the following sum of the square of fitness over all the sequences of $\lambda^\nu$:

$$\begin{aligned} \sum_{all\ sequences} W^2 &= S_2 + S_3 \\ S_2 &\equiv \sum_{all\ sequences} \sum_{j=1}^{\nu} w_j(\mathrm{A}_j|\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k})^2 \\ S_3 &\equiv \sum_{all\ sequences} 2 \sum_{i<j} w_i(\mathrm{A}_i|\mathrm{A}_{i_1}, \cdots, \mathrm{A}_{i_k}) w_j(\mathrm{A}_j|\mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}). \end{aligned}$$

We can easily calculate $S_2$ through the procedure similar to the calculation of $S_1$:

$$S_2 = \nu\lambda^{\nu-(1+k)}(\langle w\rangle^2 + \sigma^2)\lambda\lambda^k = (\langle w\rangle^2 + \sigma^2)\nu\lambda^\nu.$$

The calculation of $S_3$ is done in the following way. Let $q_{ij}$ be the number of sites included in $\{i_1, \cdots, i_k, j_1, \cdots, j_k\}$ except sites $i$ and $j$. For example, if there are no overlapped sites between $\{i_1, \cdots, i_k\}$ and $\{j_1, \cdots, j_k\}$, and sites $i$ and $j$ are not included in them, we get $q_{ij} = 2k$. The number of sites that affect the site-fitness of the $i$th and $j$th sites ($w_i$ and $w_j$) is then $2 + q_{ij}$. In the set of all the sequences, there are $\lambda^{\nu-(2+q_{ij})}$ sequences which include a set of $2 + q_{ij}$ letters, $\{\mathrm{A}_i, \mathrm{A}_{i_1}, \cdots, \mathrm{A}_{i_k}, \mathrm{A}_j, \mathrm{A}_{j_1}, \cdots, \mathrm{A}_{j_k}\}$. Then,

$$\begin{aligned} S_3 &= 2\sum_{i<j}\lambda^{\nu-(2+q_{ij})} \sum_{\{\mathrm{A}_i,\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k},\mathrm{A}_j,\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}\}} w_i(\mathrm{A}_i|\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k})w_j(\mathrm{A}_j|\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}) \\ &= 2\sum_{i<j}\lambda^{\nu-(2+q_{ij})} \sum_{\{\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k},\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}\}} \sum_a\sum_b w_i(a|\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k})w_j(b|\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}). \end{aligned}$$

Here, let us focus on

$$\sum_a\sum_b w_i(a|\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k})w_j(b|\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}).$$

Since the site-fitness values are assigned randomly from the discrete uniform distribution with mean $\langle w\rangle$ and variance $\sigma^2$ without degeneracy (see Section 2), then the covariance of site-fitnesses is zero for large $\lambda$:

$$\frac{1}{\lambda^2}\sum_a\sum_b w_i(a|\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k})w_j(b|\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}) - \langle w\rangle^2 = 0. \tag{45}$$

Therefore,

$$\sum_a\sum_b w_i(a|\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k})w_j(b|\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}) = (\langle w\rangle\lambda)^2.$$

This equation holds for all possible $\lambda^{q_{ij}}$ states of $\{\mathrm{A}_{i_1},\cdots,\mathrm{A}_{i_k},\mathrm{A}_{j_1},\cdots,\mathrm{A}_{j_k}\}$. Then

$$\begin{aligned} S_3 &= 2\sum_{i<j}\lambda^{\nu-(2+q_{ij})}(\langle w\rangle\lambda)^2\lambda^{q_{ij}} \\ &= 2\frac{\nu(\nu-1)}{2}\langle w\rangle^2\lambda^\nu = \langle w\rangle^2\nu(\nu-1)\lambda^\nu. \end{aligned}$$

Therefore, the variance of fitness over all the sequences is given by

$$\begin{aligned} \mathscr{V} &= \frac{S_2+S_3}{\lambda^\nu} - \langle W\rangle^2 = (\langle w\rangle^2+\sigma^2)\nu + \langle w\rangle^2\nu(\nu-1) - (\langle w\rangle\nu)^2 \\ &= \sigma^2\nu. \end{aligned}$$

From eqn. (2), the underlying probability density of site-fitness $w$ in each site is given by the comb function:

$$f(w) \equiv \frac{1}{\lambda}\sum_{i=0}^{\lambda-1}\delta\left(w-\epsilon\left(1-\frac{2i}{\lambda-1}\right)\right), \tag{46}$$

where $\delta(x)$ is the Dirac's delta function, and the site-fitnesses between arbitrary sites $i$ and $j$ are uncorrelated for large $\lambda$ (eqn. (45)). Then, the joint probability density of $w_1, w_2, \cdots, w_\nu$ is approximately given by

$$\prod_{j=1}^{\nu} f(w_j).$$

Then, the probability density of fitness $W$ over the whole sequence space is approximately given by

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{j=1}^{\nu} f(w_j)\ \delta\left(\sum_{j=1}^{\nu} w_j - W\right)\mathrm{d}w_1\mathrm{d}w_2\cdots\mathrm{d}w_\nu. \\ &\qquad = \overset{\nu}{*} f(W) \\ &\qquad \approx \mathscr{N}(W|\,0,\,\sigma^2\nu) \qquad\qquad \text{for large } \nu. \end{aligned}$$

## 7 Appendix B: The derivation of equations (10)- (14)

Consider a reference sequence with the fitness-value of $W$ and its $d$-boundary in $\lambda$-valued $\nu$-dimensional sequence space, and consider a change in fitness, $\Delta W$, from the reference sequence to an arbitrary sequence in the $d$-boundary. Note that an arbitrary sequence in the $d$-boundary results from random $d$-fold point mutations occurring in the reference sequence. Let $\mathscr{D}$ be the mean number of sites that change their site-fitnesses as a result of random $d$-fold point mutations, and consider that the arbitrary sequence also has the number $\mathscr{D}$. $\mathscr{D}$ is given by eqn. (7).

The fitness change $\Delta W$ is divided into the fitness income $\Delta W^{\rm in}$ and fitness outgo $\Delta W^{\rm out}$:

$$\Delta W = \Delta W^{\rm in} - \Delta W^{\rm out}. \tag{47}$$

From eqn. (2), the underlying probability density of site-fitness $w$ in each site is given by the comb function:

$$f(w) \equiv \frac{1}{\lambda}\sum_{i=0}^{\lambda-1}\delta\left(w - \epsilon\left(1 - \frac{2i}{\lambda-1}\right)\right), \tag{48}$$

where $\delta(x)$ is the Dirac's delta function. Therefore, the probability density of $\Delta W^{\rm in}$ is given by the $\mathscr{D}$-fold convolution of $f(w)$:

$$\begin{aligned}
\mathcal{P}^{\rm in}(\Delta W^{\rm in}) &= {}^{\mathscr{D}}_{*} f(\Delta W^{\rm in}) && (49)\\
&\approx \mathcal{N}(\Delta W^{\rm in}|\, 0,\, \sigma^2\mathscr{D}) \qquad \text{for large } \mathscr{D}, && (50)
\end{aligned}$$

which is a result from the central limit theorem. This is one of the reasons that our theoretical conclusion is robust to the shape of the site-fitness distribution (Limic & Pemantle, 2004).

Denote the site-fitness at the $j$-th site on the reference sequence by $w_j$: $W = \sum_{j=1}^{\nu} w_j$. The probability density of $\Delta W^{\rm out}$ is given by

$$\begin{aligned}
\mathcal{P}^{\rm out}(\Delta W^{\rm out}) &= \frac{1}{\binom{\nu}{\mathscr{D}}}\sum_{j_1<}\sum_{j_2<}\cdots\sum_{<j_{\mathscr{D}}}\delta(\Delta W^{\rm out} - \sum_{l=1}^{\mathscr{D}} w_{j_l})\\
&= \frac{1}{\binom{\nu}{\mathscr{D}}}\sum_{S} g_{S,\mathscr{D}}\,\delta(\Delta W^{\rm out} - S), && (51)
\end{aligned}$$

where $S \equiv \sum_{l=1}^{\mathscr{D}} w_{j_l}$ and $g_{S,\mathscr{D}}$ is defined as the coefficient of the following generating function:

$$\begin{aligned}
G(X,Z) &= (1+X^{w_1}Z)(1+X^{w_2}Z)(1+X^{w_3}Z)\cdots(1+X^{w_\nu}Z)\\
&= \sum_{\mathscr{D}=0}^{\nu}\Big(\sum_{S} g_{S,\mathscr{D}}X^S\Big)Z^{\mathscr{D}}. && (52)
\end{aligned}$$

The coefficient $g_{S,\mathscr{D}}$ satisfies

$$\frac{1}{\binom{\nu}{\mathscr{D}}}\sum_{S} g_{S,\mathscr{D}} = 1.$$

If the distribution of site-fitness values over all sites, $\{w_1, w_2, \cdots, w_\nu\}$, obeys an ordinary distribution, then the probability distribution of $S \equiv \sum_{l=1}^{\mathscr{D}} w_{j_l}$ for large $\mathscr{D}$, that is $g_{S,\mathscr{D}}/\binom{\nu}{\mathscr{D}}$,

tends to be a normal distribution. Then, the probability density $\mathcal{P}^{\rm out}(\Delta W^{\rm out})$ (eqn. (51)) is approximately given by

$$\mathcal{P}^{\rm out}(\Delta W^{\rm out}) \approx \mathscr{N}\left(\Delta W^{\rm out}\Big|\overline{w}\mathscr{D}, \widetilde{\sigma}^2\mathscr{D}\left(1-\frac{\mathscr{D}-1}{\nu-1}\right)\right) \qquad \text{for large } \mathscr{D}, \tag{53}$$

where $\overline{w} \equiv \frac{1}{\nu}\sum_{j=1}^{\nu} w_j = W/\nu$ and $\widetilde{\sigma}^2 \equiv \frac{1}{\nu}\sum_{j=1}^{\nu} w_j^2 - \overline{w}^2$. The probability density of $\Delta W$ $(= \Delta W^{\rm in} - \Delta W^{\rm out})$ is given by

$$\begin{aligned}\mathcal{P}(\Delta W) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathcal{P}^{\rm in}(x)\,\mathcal{P}^{\rm out}(y)\,\delta(\Delta W-(x-y))\,{\rm d}x{\rm d}y \\ &\approx \mathscr{N}(\Delta W|\mathbf{E}_d[\Delta W|\overline{w}], \mathbf{V}_d[\Delta W|\widetilde{\sigma}^2]).\end{aligned} \tag{54}$$

The conditional mean and variance of $\Delta W$ with $\overline{w}$ and $\widetilde{\sigma}^2$ given are respectively obtained as follows:

$$\begin{aligned}\mathbf{E}_d[\Delta W|\overline{w}] &\equiv \int_{-\infty}^{\infty}\Delta W\,\mathcal{P}(\Delta W)\,{\rm d}\Delta W \\ &= -\overline{w}\mathscr{D}\end{aligned} \tag{55}$$

$$\begin{aligned}\mathbf{V}_d[\Delta W|\widetilde{\sigma}^2] &\equiv \int_{-\infty}^{\infty}\Delta W^2\,\mathcal{P}(\Delta W)\,{\rm d}\Delta W - \mathbf{E}_d[\Delta W|\overline{w}]^2 \\ &\approx \left(\sigma^2+\widetilde{\sigma}^2\left(1-\frac{\mathscr{D}-1}{\nu-1}\right)\right)\mathscr{D}.\end{aligned} \tag{56}$$

$\mathcal{P}(\Delta W)$ corresponds to eqn. (10).

Here, we note that $\mathcal{P}^{\rm in}(\Delta W^{\rm in})$ and $\mathcal{P}^{\rm out}(\Delta W^{\rm out})$ include the information of all possible changes of site-fitnesses. The number of all possible patterns of site-fitness changes is $\binom{\nu}{\mathscr{D}}(\lambda-1)^{\mathscr{D}}$ $(= N_{\partial\mathscr{D}})$. However, actually the number of all possible $d$-fold point mutations is $\binom{\nu}{d}(\lambda-1)^d$ $(= N_d^\dagger)$. When $k$ is large, $N_d^\dagger \ll N_{\partial\mathscr{D}}$. This is the reason for considering $\mathcal{P}(\Delta W)$ in eqn. (54) or $\mathcal{P}_d(\Delta W|W)$ in eqn. (10) to be approximately an independent identical distribution (i.i.d.) for assignment of $\Delta W$ value to any sequence in the $d$-boundary, when $k$ is large. This approximation is more valid as $k$ increases. When $k = 0$, $\mathcal{P}(\Delta W)$ in eqn. (54) or $\mathcal{P}_d(\Delta W|W)$ in eqn. (10) is not i.i.d. but equivalent to the density of $\Delta W$ over the $d$-boundary: $\psi_d(\Delta W|W) = \mathcal{P}_d(\Delta W|W)$.

## 8 Appendix C: The most probable distribution of site-fitness values over all sites when fitness $W$ is given

The probability density of fitness in the $d$-boundary of a reference sequence is dependent on the distribution of site-fitness values over all sites of the sequence. Our aim in this section is to deduce the most probable distribution of site-fitness values over all sites when fitness $W$ of the sequence is given. Let $\mathrm{w}_i$ be the site-fitness for the $i$th available letter among $\lambda$ ones, and let $p_i$ be the occurrence probability of the $i$th letter over all sites. Then, $\sum_{i=0}^{\lambda-1} p_i = 1$ and $\sum_{i=0}^{\lambda-1} \mathrm{w}_i p_i = \overline{w}$. From eqn. (2), the underlying site-fitness distribution function for each site is the discrete uniform distribution:

$$\mathrm{w}_i \equiv \epsilon \left(1 - \frac{2i}{\lambda - 1}\right), \qquad i = 0, 1, 2, \cdots, \lambda - 1. \tag{57}$$

Adopting the maximal entropy assumption, we deduced the most probable distribution by Lagrange's method of undetermined multipliers:

$$\{p_0, p_1, \cdots, p_{\lambda-1}\} = \arg\max \left\{ -\sum_{i=0}^{\lambda-1} p_i \, \ln p_i \right\} \tag{58}$$

with constraint condition

$$\begin{aligned} \sum_{i=0}^{\lambda-1} p_i &= 1 \\ \sum_{i=0}^{\lambda-1} \mathrm{w}_i \, p_i &= \overline{w}. \end{aligned} \tag{59}$$

Solving eqn. (58), we obtain

$$p_i = e^{\beta \mathrm{w}_i}/Z \quad \text{, where} \quad Z \equiv \sum_{i=0}^{\lambda-1} e^{\beta \mathrm{w}_i}. \tag{60}$$

Substituting eqn. (57) and eqn. (60) into eqn. (59), we obtain $p_i$ as the following exponential distribution:

$$p_i = x^i/Z(x) \quad \text{, where} \quad Z(x) \equiv \sum_{i=0}^{\lambda-1} x^i = \frac{1 - x^\lambda}{1 - x}, \tag{61}$$

where $x$ is determined by solving

$$\begin{aligned}\frac{\overline{w}}{\epsilon} &= 1-\frac{2xZ'(x)}{(\lambda-1)Z(x)} = 1-\frac{2x}{(\lambda-1)}\frac{\mathrm{d}\ln Z(x)}{\mathrm{d}x} \\ &= 1-\frac{2}{(\lambda-1)}\left(\frac{\lambda x^{\lambda}}{x^{\lambda}-1}+\frac{x}{1-x}\right). \end{aligned} \tag{62}$$

Fig.3 shows the following variance by using eqn. (61) with eqn. (62):

$$\widetilde{\sigma}^2 = \sum_{i=0}^{\lambda-1} \mathrm{w}_i^2\, p_i - \overline{w}^2. \tag{63}$$

From Fig.3, we can see that the analytically derived $\widetilde{\sigma}^2$ agrees well with the observed values by computer simulation.

# 9 Appendix D: Rough estimation of the average number of nearly neutral sequences in the $d$-boundary

In Aita *et al.* (2004), we referred to the probability density of the fitness for the $m$-th fittest mutant among $N$ mutants ($m \ll N$), which are randomly chosen from the $d$-boundary. Following the results, we estimate roughly the average number of nearly neutral sequences in the $d$-boundary. Let us consider a reference sequence with the stationary fitness $W_d^*$, and its $d$-boundary. Picks up $N$ sequences randomly from the $d$-boundary. Let $\Delta W_m^*$ be the change in fitness from the reference sequence to the $m$-th fittest sequence among the $N$ sequences ($m \ll N$). The probability density of $\Delta W_m^*$ is roughly given by

$$\varphi_m(\Delta W_m^*) \approx \mathscr{N}\left(\Delta W_m^*\Big| - a(\zeta_1-\zeta_m), \frac{a^2}{m\,\zeta_m^2}\right), \tag{64}$$

where $a \equiv \sqrt{2\mathscr{V}\mathscr{D}/\nu}$ and $\zeta_m \approx \sqrt{2\ln(N/\sqrt{2\pi}m)}$. Eqn.(64) is derived from eqns (20)-(24), (35), (36) and (C.5) in Aita *et al.*, (2004)[4]. The probability that $n$ sequences among the $N$ sequences belong to the fitness band defined as $W_d^* - z\times SD^* \leqq W \leqq W_d^* + z\times SD^*$ is roughly given by

$$\mathcal{P}(n) \approx C \times \int_{-z\times SD^*}^{z\times SD^*} \varphi_n(x)\mathrm{d}x \int_{-\infty}^{-z\times SD^*} \varphi_{n+1}(x)\mathrm{d}x, \tag{65}$$

[4]In this paper, we wrote "$\widetilde{N}_m$" as $\zeta_m$.

where $C$ is a normalization constant. Note that $\mathcal{P}(0) \approx C \times \int_{-\infty}^{-z \times SD^*} \varphi_1(x)\mathrm{d}x$.

We focus on the case where $N = N_d^\dagger$, that is a large value. Then $SD^* \approx a/\zeta_1$, which is easily derived from eqn.(35) with $\zeta_{\max} = \zeta_1$. By using $\zeta_1 - \zeta_m \approx (\ln m)/\zeta_1$ and $\zeta_m^{-1} \approx \zeta_1^{-1}$, eqn.(64) is approximated by

$$\varphi_m(\Delta W_m^*) \approx \mathscr{N}\left(\Delta W_m^* \middle| -\frac{a \ln m}{\zeta_1}, \frac{a^2}{m\,\zeta_1^2}\right). \tag{66}$$

Then, eqn.(65) is approximated by

$$\mathcal{P}(n) \approx C \times \int_{-z}^{z} \mathscr{N}(x| - \ln n, 1/n)\mathrm{d}x \int_{-\infty}^{-z} \mathscr{N}(x| - \ln(n+1), 1/(n+1))\mathrm{d}x. \tag{67}$$

It is remarkable that the probability $\mathcal{P}(n)$ shown in eqn.(67) is dependent on $z$ but independent of other parameters. The analytical numbers presented in Table 1 were calculated by $\mathbf{E}[n] = \sum_n n\,\mathcal{P}(n)$ with eqn.(67).

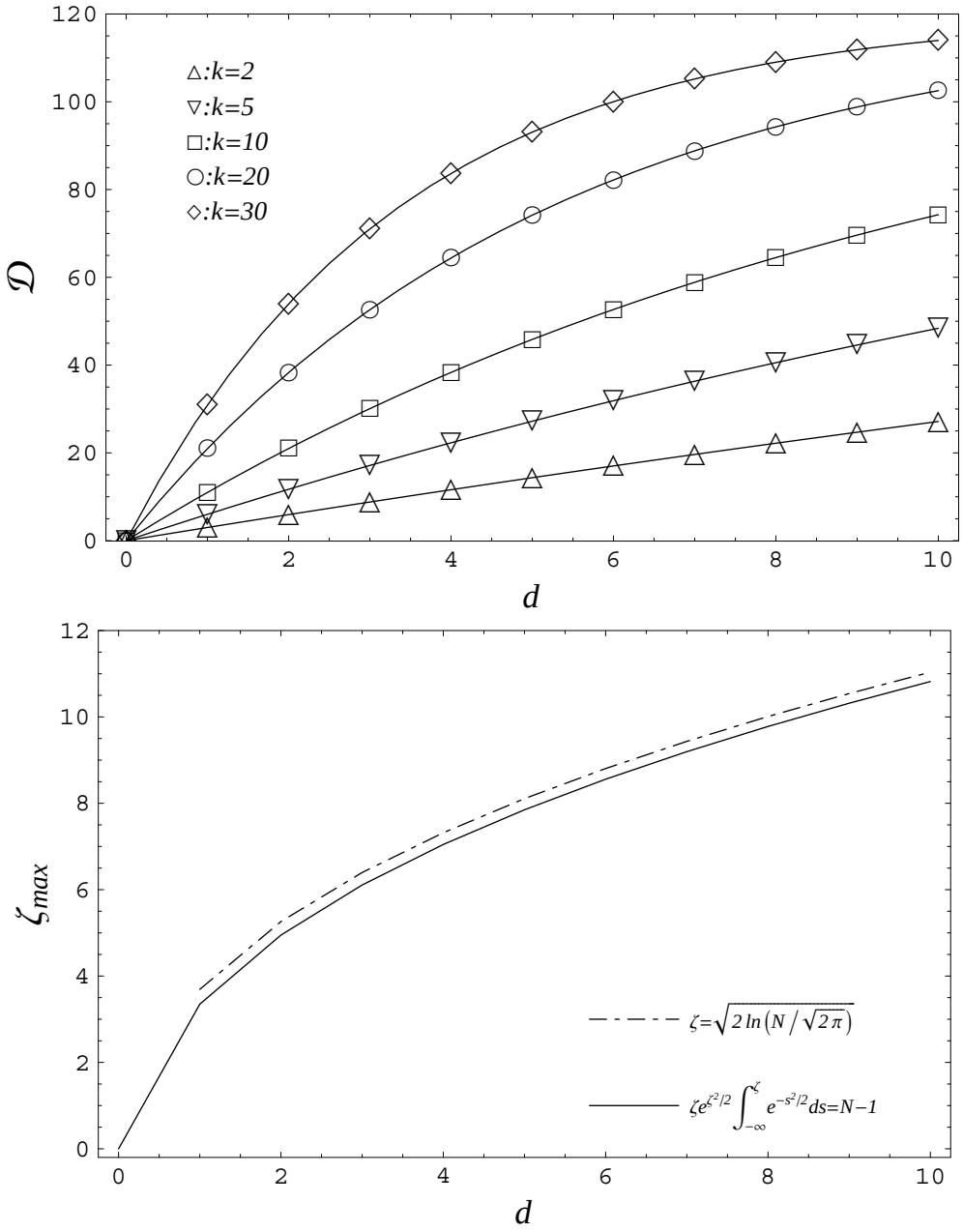



Figure 1: **Two important parameters $\mathscr{D}$ and $\zeta_{\max}$ as a function of Hamming distance $d$.** (Top) $\mathscr{D}$ values were calculated by eqn. (7). $\nu = 120$. The values of parameter $k$ were indicated in the figure. (Bottom) $\zeta_{\max}$ values were calculated by eqn. (22) and eqn. (23) with eqn. (9). $\lambda = 20, \nu = 120$.

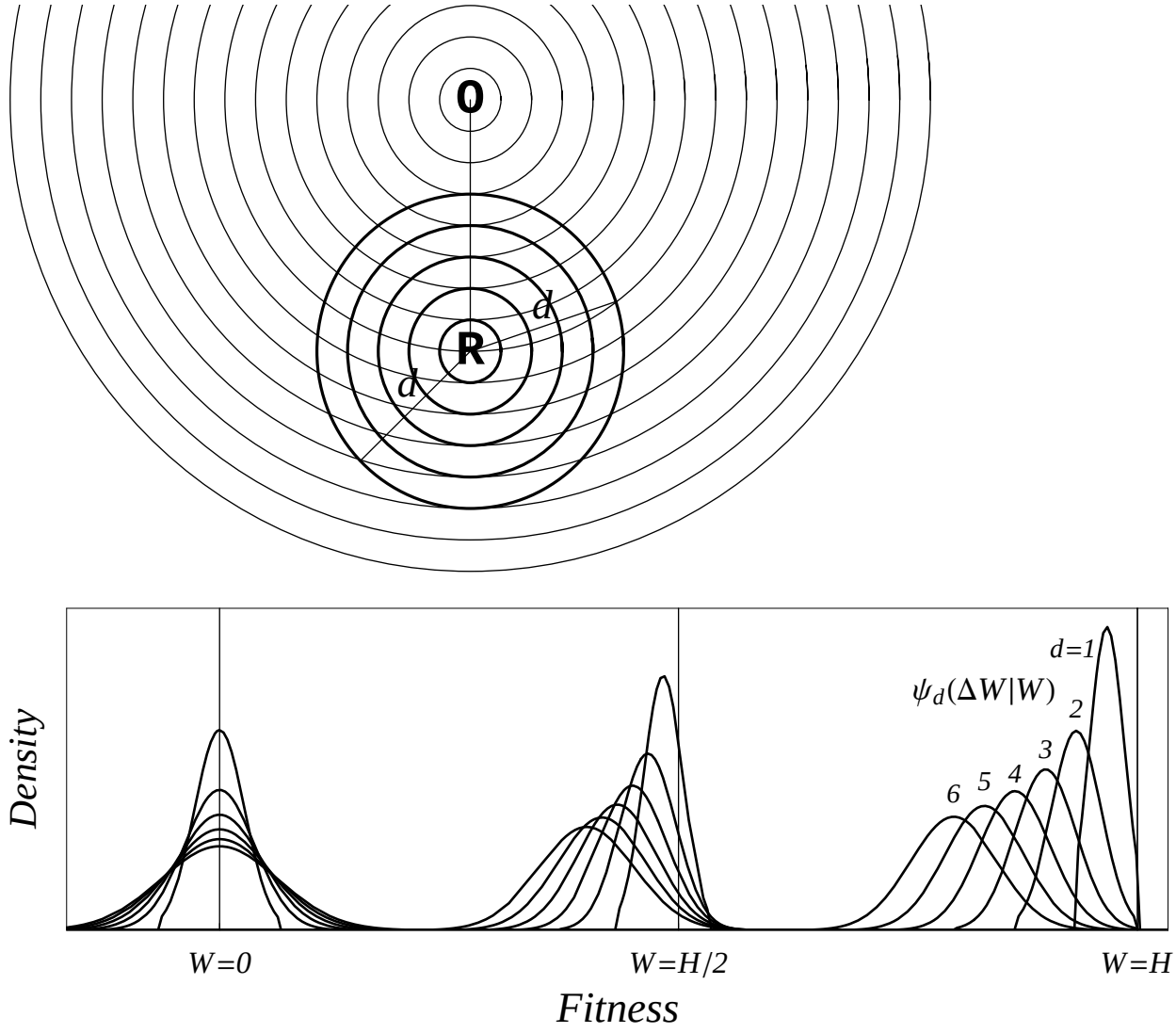



Figure 2: **The $d$-boundary and fitness distribution on it.** (Top) Global sequence space is schematically represented by concentric circles centering at the global peak "**O**". The radius of each concentric circle represents the Hamming distance. The thick concentric circle with radius $d$ centering at a reference sequence "**R**" represents the $d$-boundary of the reference sequence. (Bottom) Each Gaussian-like distribution represents the density function of fitnesses on the $d$-boundary (= $\psi_d(\Delta W|W)$). Three cases for the location of reference sequences are shown: $W = H$ (peak), $W = H/2$ (middle) and $\mathscr{W} = 0$ (foot).

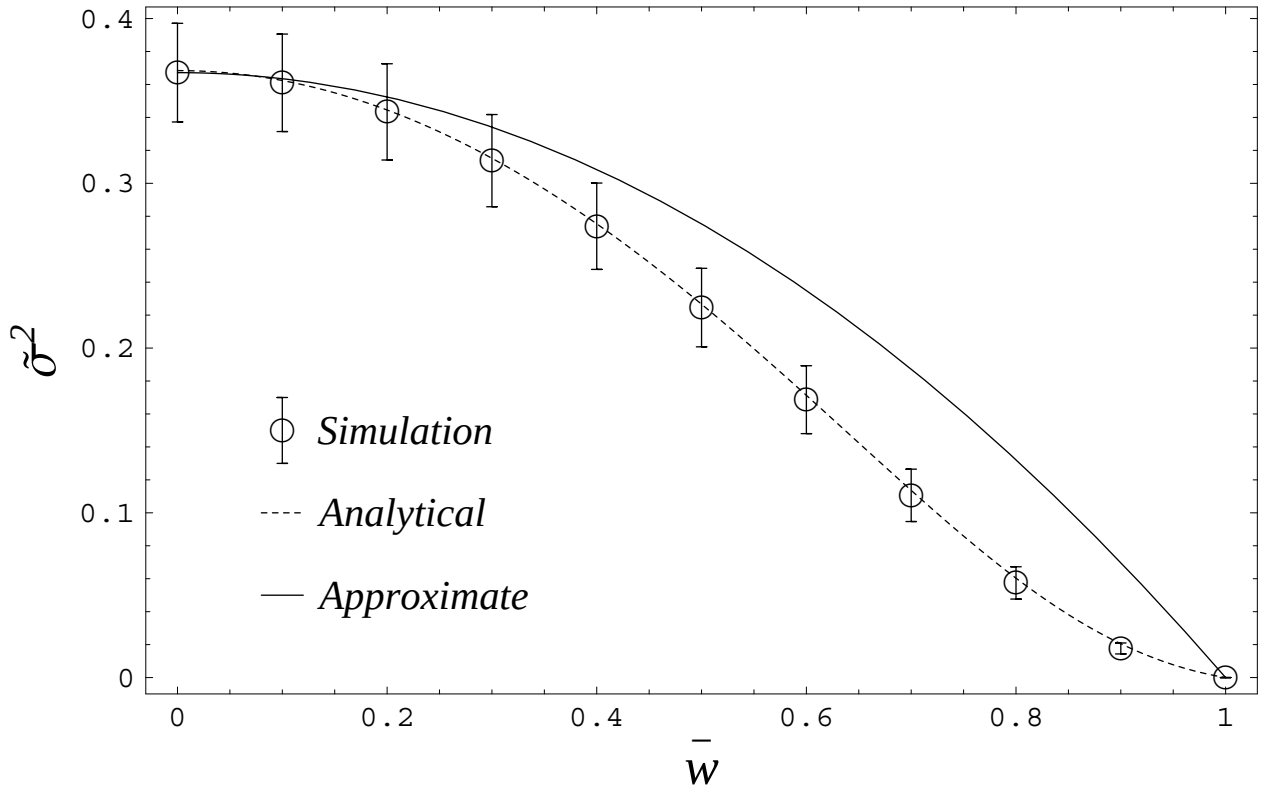



Figure 3: $\widetilde{\sigma}^2$ **against** $\overline{w}$**.** The result of the computer simulation is plotted with mean (circles) and standard deviation (error bar). The dashed line shows the analytically derived one (eqns (61)-(63)), which was obtained by using the maximal entropy assumption (details are in Appendix C). The solid line shows the parabola shown in eqn. (17). Parameters are $\lambda = 20$ and $\epsilon = 1$.

| z | 0.5 | 1.0 | 1.5 | 2.0 |
|---|---|---|---|---|
| Analytical number | 1.4 | 2.4 | 4.0 | 6.7 |
| Simulation ($N_d^\dagger = 10^2$) | 1.9 | 2.6 | 3.7 | 5.5 |
| Simulation ($N_d^\dagger = 10^3$) | 1.9 | 2.7 | 4.2 | 6.2 |
| Simulation ($N_d^\dagger = 10^4$) | 1.9 | 2.7 | 4.2 | 6.4 |

Table 1: **The average number of nearly neutral sequences in the** $d$**-boundary, in which the reference sequence takes the stationary fitness** $W_d^*$**.** The parameter $z$ governs the width of the fitness band, $W_d^* - z \times SD^* \leqq W \leqq W_d^* + z \times SD^*$. The analytical numbers were calculated based on eqn.(67) in Appendix D. $N_d^\dagger$ represents the number of all the sequences in the $d$-boundary.

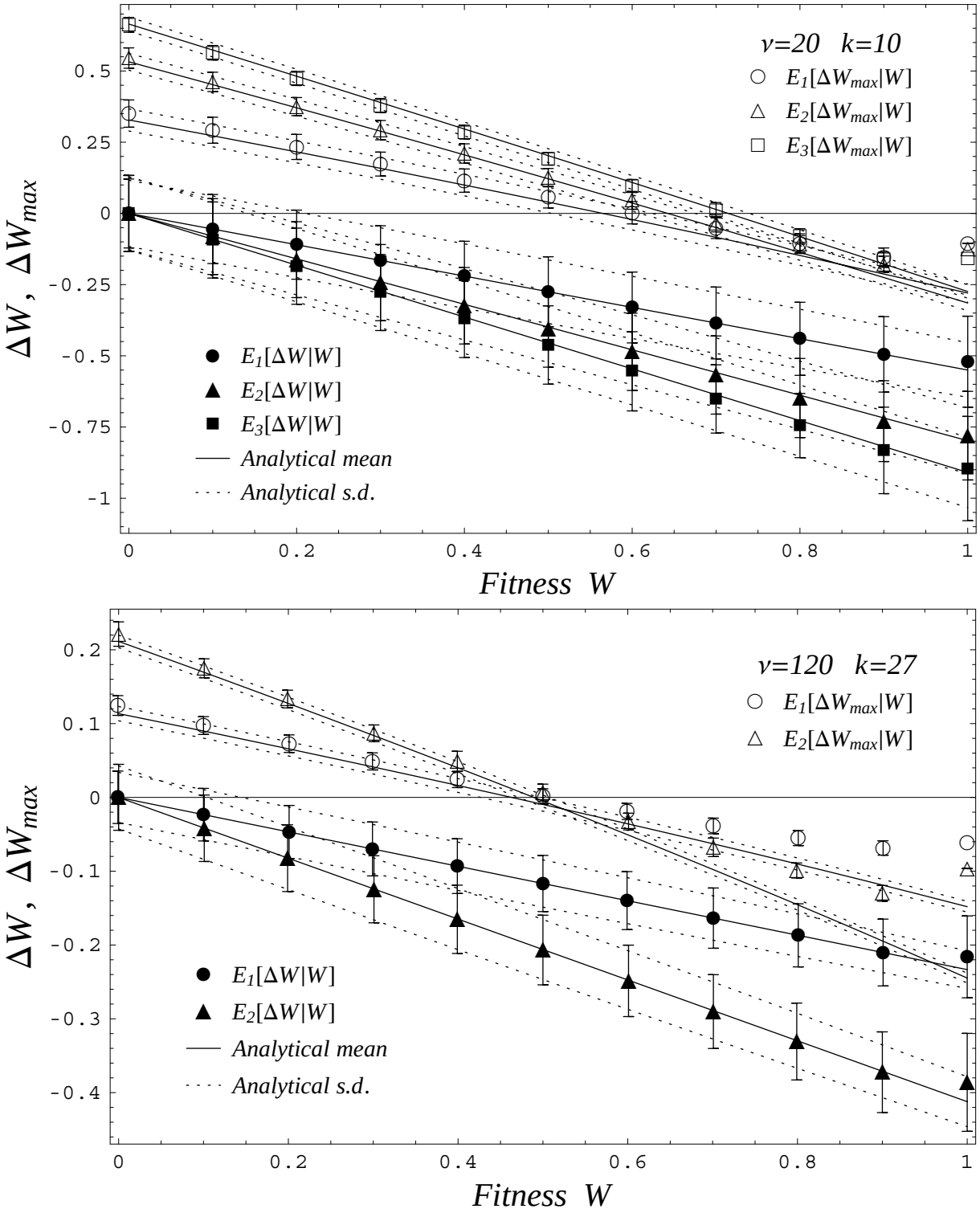



Figure 4: **The mean ($\mathbf{E}_d[*]$) and standard deviation ($\sqrt{\mathbf{V}_d[*]}$) of $\Delta W$ and $\Delta W_{\max}$ against fitness $W$ of a reference sequence.** The regions $W = 0$ and $W = 1$ correspond to the foot and global peak, respectively. The simulated mean and standard deviation are plotted with symbols and error bars: the circle ($d = 1$), triangle ($d = 2$), and square ($d = 3$). The solid and dotted lines show the analytical mean and standard deviation, respectively, according to eqns (12),(14),(20), and (21). Parameters are as follows: (top) $\lambda = 20, \nu = 20, k = 10$; (bottom) $\lambda = 20, \nu = 120, k = 27$; $\epsilon = 1/\nu$.

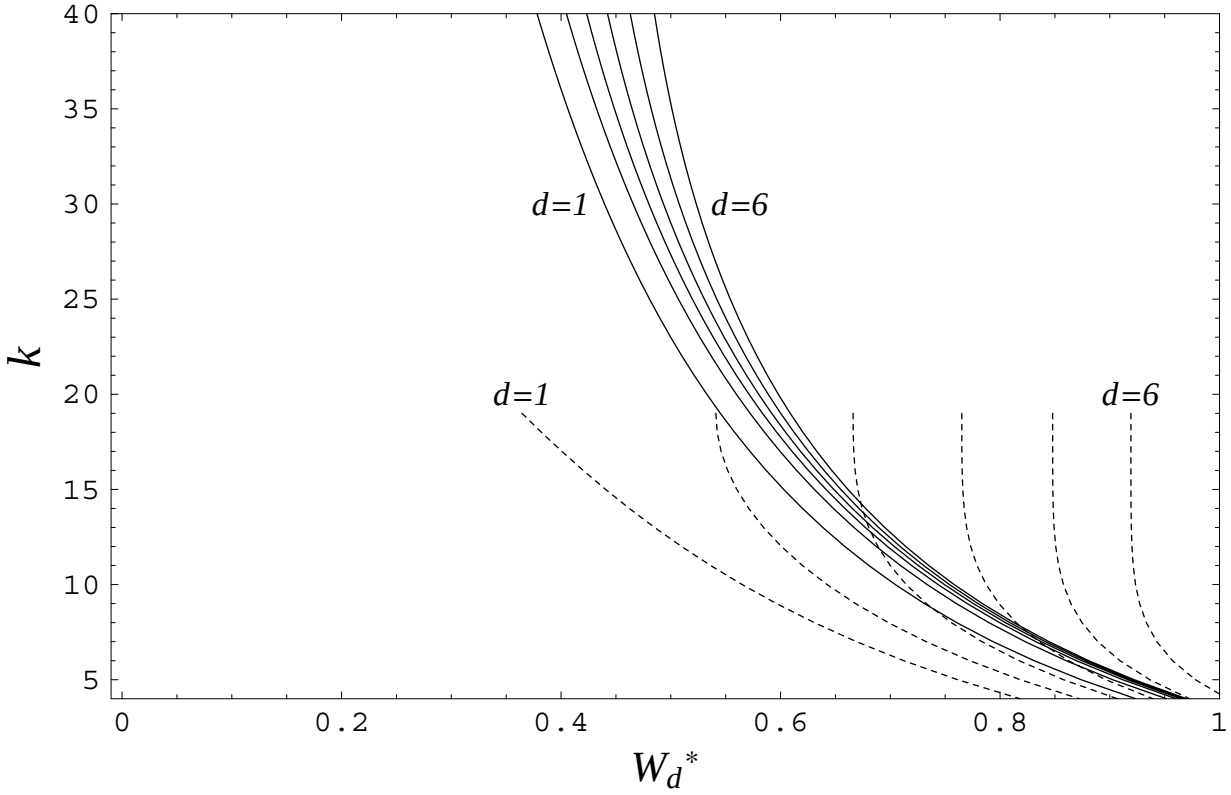



Figure 5: **The fitness of the highland, $W_d^*$, against $k$ values.** The regions $W_d^* = 0$ and $W_d^* = 1$ correspond to the foot and global peak, respectively. $W_d^*$ values were calculated by eqn. (30). The solid and dashed lines are for $\nu = 120$ and $\nu = 20$, respectively. From the leftmost line to the rightmost one, $d = 1, 2, \cdots, 6$. Parameters are $\lambda = 20$ and $\epsilon = 1/\nu$.

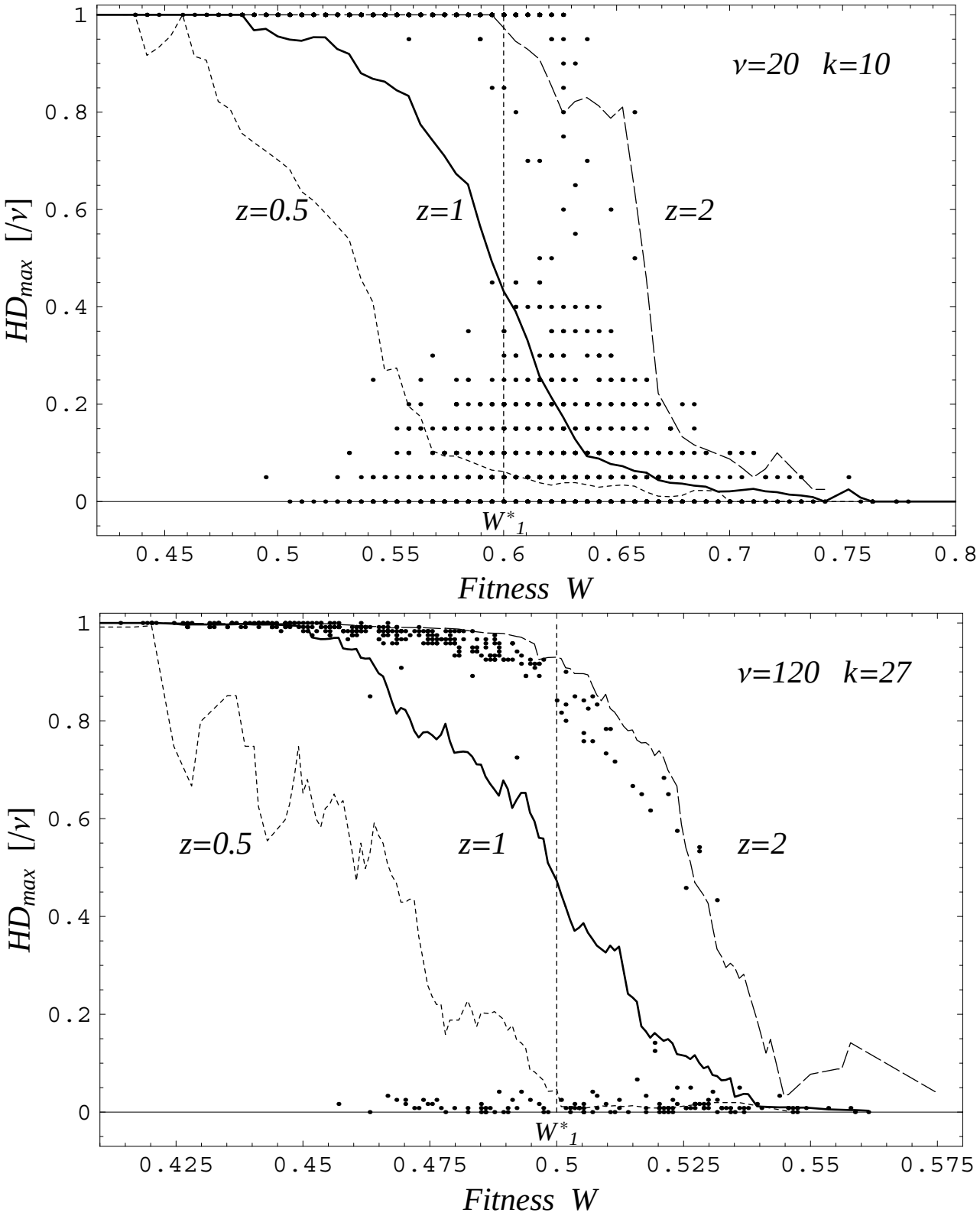
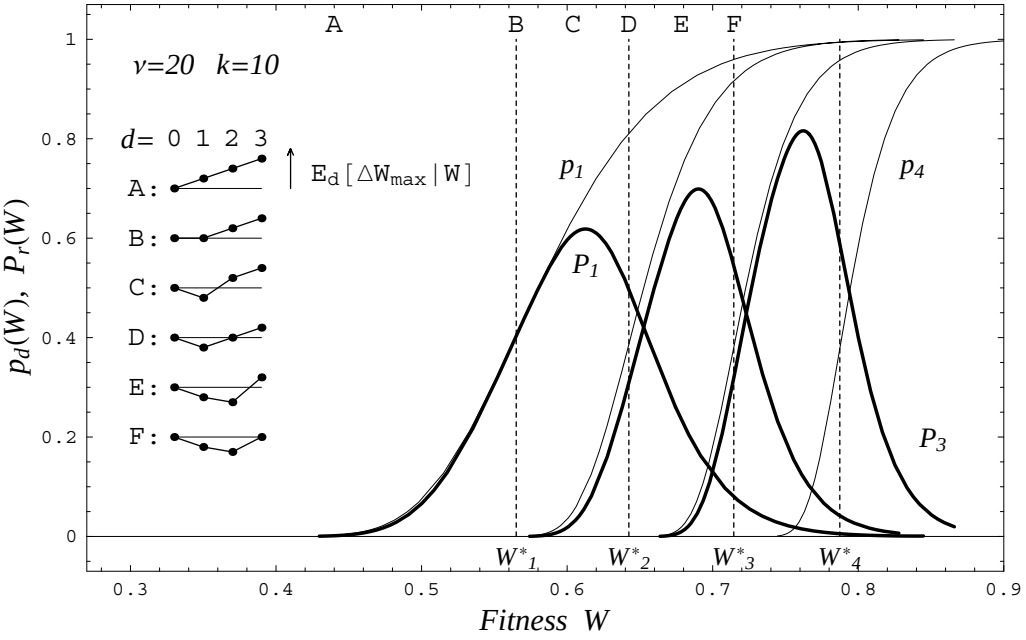
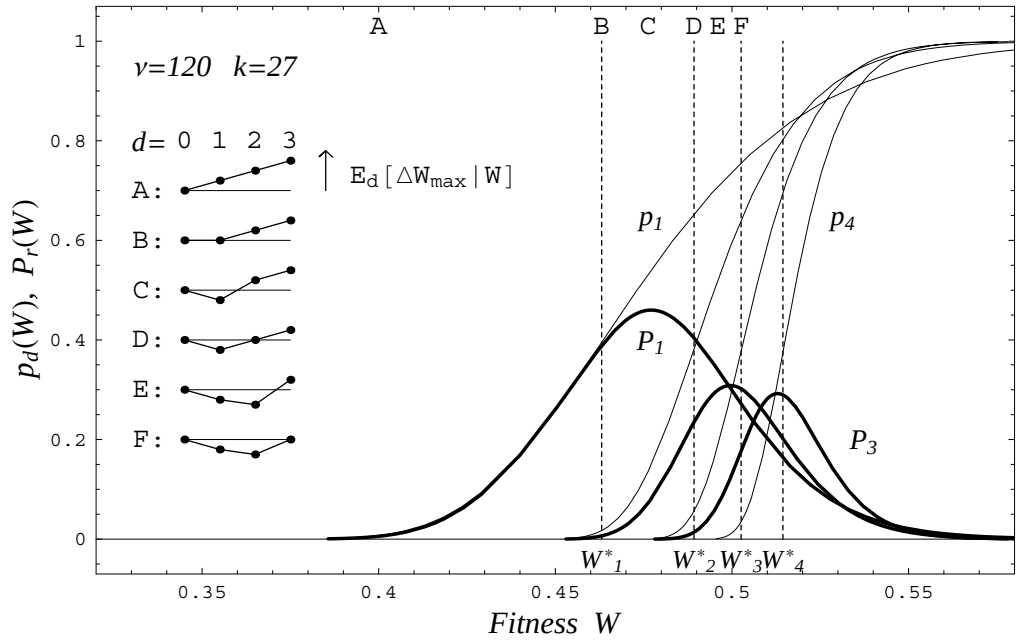



Figure 6: **The maximal Hamming distance, $HD_{\max}$, against fitness $W$.** The three descending lines represent the average of $HD_{\max}$ against $W$, for $z = 0.5, 1$, and 2. The dashed vertical line indicates the value of $W_1^*$ obtained by the simulation. The dots indicate the raw data for $z = 1$. Some of these data are degenerated because of discrete fitness. Common parameters are $\lambda = 20$ and $\epsilon = 1/\nu$.


A
B
C
D
E
F
ν=20 k=10
d= 0 1 2 3
A:
B:
C:
D:
E:
F:
E_d [ΔW_max | W]
p1
p4
P1
P3
W*1
W*2
W*3
W*4
p_d(W), P_r(W)
Fitness W


A
B
C
D E F
ν=120 k=27
d= 0 1 2 3
A:
B:
C:
D:
E:
F:
E_d [ΔW_max | W]
p1
p4
P1
P3
W*1
W*2 W*3 W*4
p_d(W), P_r(W)
Fitness W

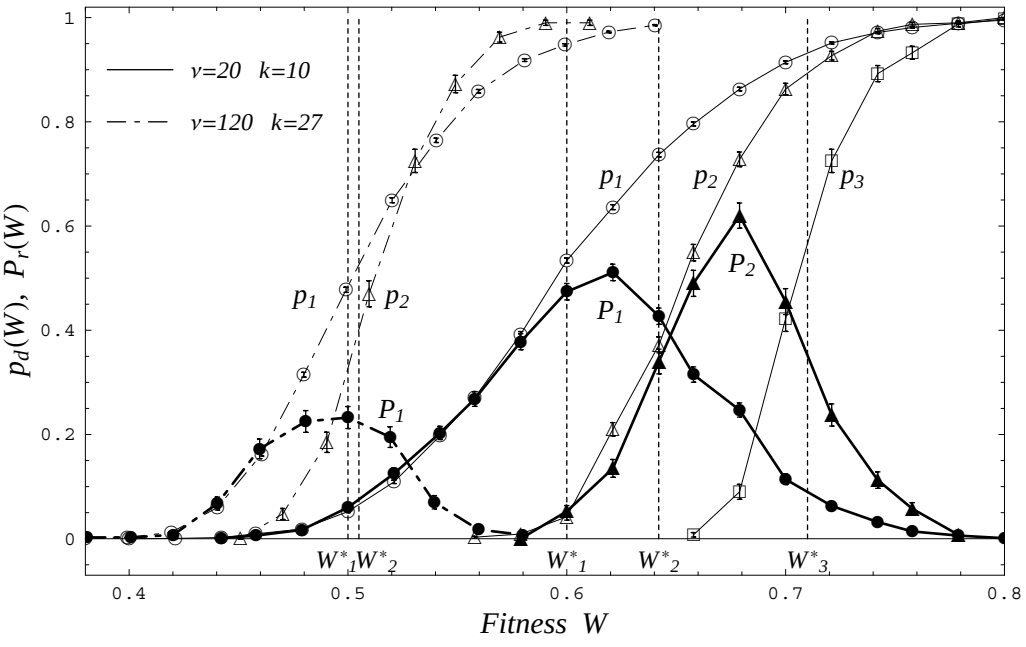

ν=20 k=10
ν=120 k=27
p1
p2
p3
P1
P2
W*1 W*2
W*1
W*2
W*3
p_d(W), P_r(W)
Fitness W

Figure 7: **Hierarchical distribution of ascending slopes, highlands, and local optima along the fitness coordinate $W$.** (Top and middle) the analytically predicted view. The dashed vertical lines indicate the values of $W_d^*$: (from left to right) $W_1^*, W_2^*, W_3^*, W_4^*$. The solid sigmoidal lines (thin lines) represent $p_d(W)$: (from left to right) $p_1, p_2, p_3, p_4$. The Gaussian-like lines (thick lines) represent $P_r(W)$: (from left to right) $P_1, P_2, P_3$. In the left corner, several types of local structural features (A,B,C,D,E,F) are shown schematically. The dot at $d = 0$ indicates the fitness of a reference point, and the three dots at $d = 1, 2, 3$ indicate their respective $\mathbf{E}_d[\Delta W_{\max}|W]$. For example, Type A represents an ascending slope, and Type E represents the local optimum (mountain) at the second order. Type X (X=A,B,$\cdots$,F) appears mainly in Region X (X=A,B,$\cdots$,F) on the fitness coordinate. A: ascending slopes exist everywhere. B: the highland at the first order. C: the probability $P_1(W)$ takes the maximal value. D: the highland at the second order. E: the probability $P_2(W)$ takes the maximal value. Parameters are as follows: (top) $\lambda = 20, \nu = 20, k = 10$; (middle) $\lambda = 20, \nu = 120, k = 27$; $\epsilon = 1/\nu$. Details are in the main text. (Bottom) a result of computer experiments to compare with the top and middle figures. Let $N$ and $n$ be the number of samples and the number of objects, respectively. The dot represents $n/N$ and the error bar represents $\sqrt{n(N-n)/N^3}$. The solid lines are for $\lambda = 20, \nu = 20, k = 10$, while the broken lines are for $\lambda = 20, \nu = 120, k = 27$.