Solution of the Crow-Kimura and Eigen Models for Alphabets of Arbitrary Size by Schwinger Spin Coherent States

Enrique Muñoz · Jeong-Man Park · Michael W. Deem

Abstract To represent the evolution of nucleic acid and protein sequence, we express the parallel and Eigen models for molecular evolution in terms of a functional integral representation with an h-letter alphabet, lifting the two-state, purine/pyrimidine assumption often made in quasi-species theory. For arbitrary h and a general mutation scheme, we obtain the solution of this model in terms of a maximum principle. Euler's theorem for homogeneous functions is used to derive this 'thermodynamic' formulation of evolution. The general result for the parallel model reduces to known results for the purine/pyrimidine $h = 2$ alphabet and the nucleic acid $h = 4$ alphabet for the Kimura 3 ST mutation scheme. Examples are presented for the $h = 4$ and $h = 20$ cases. We also derive the maximum principle for the Eigen model for general h. The general result for the Eigen model reduces to a known result for $h = 2$. Examples are presented for the nucleic acid $h = 4$ and the amino acid $h = 20$ alphabet. An error catastrophe phase transition occurs in these models, and the order of the phase transition changes from second to first order for smooth fitness functions when the alphabet size is increased beyond two letters to the generic case. As examples, we analyze the general analytic solution for sharp peak, linear, quadratic, and quartic fitness functions.

Keywords Molecular evolution · Quasispecies theory · Evolution alphabet size · Error catastrophe

1 Introduction

There are two classical physical models of molecular evolution: the Eigen model [1–3] and the Parallel or Crow-Kimura model [4]. These models were originally formulated in the language of chemical kinetics [1], by a large system of differential equations representing the time evolution of the relative frequencies of each sequence type. Quasi-species models capture the basic microscopic processes of mutation and replication, for an infinite population of binary sequences. The most remarkable feature of these models is the existence of a

E. Muñoz · J.-M. Park · M.W. Deem (✉)
Department of Physics & Astronomy, Rice University, 6100 Main St., Houston, TX 77005-1892, USA
e-mail: mwdeem@rice.edu

phase transition, termed the "error threshold" [1, 5], when the mutation rate is below a critical value, separating a disordered non-selective phase from an organized or "quasi-species" phase. The quasi-species is characterized by a population of closely related mutants, rather than by identical sequences [1–3], and its emergence is related to the auto-catalytic character of the replication process [1, 5], which exponentially enriches the proportion of the fittest mutants in the population. Experimental studies provide support for quasi-species theory in the evolution of RNA viruses [6, 7].

The choice of a binary alphabet, which simplifies the mathematical and numerical analysis of the theory, represents a coarse graining of the four-letter alphabet of the nucleic acids DNA/RNA (A, C, G, T/U), by considering the two basic chemical structures of nitrogenated bases, purines (A, G) and pyrimidines (C, T/U). The choice of a four-letter alphabet represents nucleic acids. A 20-letter alphabet represents amino acids and protein structure and permits a close connection between sequence and fitness.

Most numerical and analytical studies on quasi-species models consider the binary alphabet simplification [1–4]. In particular, the assumption of a binary alphabet allows for an exact mapping of the quasi-species models into a 2D Ising model [8, 9], or into a quantum spin chain [10–14]. An exception is [15], where a four-letter alphabet was studied by a quantum spin chain representation of the parallel model. Other approaches to the nucleic acid evolution problem have been presented in [16, 17]. In all these studies it has been shown, through the application of different methodologies, that the steady state mean fitness of the population can be expressed in terms of a maximum principle, in the limit of infinite sequence length ($N \to \infty$). The Frobenius-Perrone theorem guarantees that there is a unique steady-state population distribution. It has been shown [18] that for a general family of linear (or effectively linear) models that evolve according to a matrix $H = M + R$, with M a Markov generator (typically representing mutations) and R a diagonal matrix (usually representing replication or degradation) of dimension $N \to \infty$, the Frobenius-Perrone largest eigenvalue can be expressed in terms of a Raleigh-Ritz variational problem. The high dimensional variational problem can be reduced to a low-dimensional maximum principle with an error $O(1/N)$, when basic symmetries can be assumed in the evolution matrix, such as permutation invariance of the replication rate, or symmetric mutation rates, which allows for a lumping [18] of the large sequence space into sequence types or classes. This analysis was applied in [17] to obtain a variational expression for the mean fitness in the Kimura model with a four letters alphabet. We note that the Eigen model, where replication and (multiple) mutations are correlated, possesses a different algebraic structure than the general family of models studied in [18]. In the Eigen model, the evolution matrix is of the form $H = Q \times R$, with Q representing the mutation matrix, and R a diagonal replication matrix.

In this article, we present exact analytical solutions of the h-alphabet Crow-Kimura and Eigen models by means of a quantum field theory. Our method generalizes the Schwinger spin coherent field theory for the binary alphabet in [19] to an alphabet of arbitrary size h. This method has also been recently applied in the solution of a model that includes transfer of genetic material between sequences in quasi-species theory [20, 21], and two-parent recombination [21]. For the parallel model, we present exact analytical solutions of this field theory, in terms of a maximum principle, for the steady state mean fitness of the population and average composition, (44). We present as examples, results for the Kimura 3 ST mutation scheme [22], (54). We develop in detail the result for the symmetric mutation rate scheme, (55), for four different examples of microscopic fitness functions: sharp peak (57) and (58), Fujiyama landscape (60)–(61), a quadratic landscape (65)–(69), and a quartic landscape (71) and (72). In Sect. 2.9, we apply our general formula to derive the mean fitness for the symmetric, general h case and discuss the $h = 20$ amino acid alphabet, (74). For the

symmetric case, we present results for the sharp peak, (75)–(77), and the quadratic case, (77).

For the Eigen model, we present the exact expression for arbitrary alphabet size h, (106). As an example, we apply the general expression to a mutation scheme analogous to the Kimura 3 ST [22], (112). We analyze in detail the solution for the symmetric mutations rate, (113), for four different examples of microscopic fitness functions: sharp peak (114) and (115), Fujiyama landscape (116)–(119), quadratic fitness landscape (121)–(123), and quartic fitness landscape (125) and (126). In Sect. 3.8, we apply our general solution to derive the mean fitness for the symmetric, general h case and discuss the $h = 20$ amino acid alphabet, (127). For the symmetric case, we present results for the sharp peak, (128)–(129), and the quadratic case.

These results bring quasi-species theory closer to the real microscopic evolutionary dynamics that occurs in the natural four-letter alphabet of nucleic acids or the 20-letter alphabet of amino acids.

2 The Parallel Model for an Alphabet of Size h

The parallel model [4] describes the continuous time evolution of an infinite size population of viral genetic sequences. The evolutionary dynamics is driven by point mutations and selection, with mutations occurring in parallel and independently of viral replication. Each viral genome is represented as a sequence of N letters, from an alphabet of size h, and therefore the total number of different viral genomes in the population is h^N. If we describe a viral genetic sequence in the alphabet of nucleic acids (DNA or RNA), the natural choice would be $h = 4$, and explicitly the alphabet corresponds to (A, C, G, T or U). It is common, to simplify the theoretical analysis, to choose instead a coarse grained alphabet of size $h = 2$, by 'lumping' together purines (A, T or U) and pyrimidines (C, G). Alternatively, to describe evolution at the scale of protein sequences, the natural choice is to consider the $h = 20$ amino acid alphabet. We here consider the case of general h.

The probability p_i for a virus to have a genetic sequence S_i, $1 \leq i \leq h^N$, evolves according to the following system of nonlinear differential equations

$$\frac{dp_i}{dt} = p_i \left(r_i - \sum_{j=1}^{h^N} r_j p_j \right) + \sum_{j=1}^{h^N} \mu_{ij} p_j \tag{1}$$

Here r_i is replication rate of sequence S_i, and μ_{ij} is the mutation rate from sequence S_j into sequence S_i. The nonlinear term in (1) represents the average replication rate in the population, or mean fitness. This non-linear term enforces the conservation of probability, $\sum_i p_i = 1$. This term can be removed through a simple exponential transformation, to obtain the linear system of differential equations

$$\frac{dq_i}{dt} = r_i q_i + \sum_{j=1}^{h^N} \mu_{ij} q_j \tag{2}$$

where $p_i(t) = q_i(t) / \sum_j q_j(t)$.

2.1 Fitness Landscape

We will assume that the replication rate (fitness) of an individual in the population depends on its relative composition with respect to a wild-type sequence S_w. We define the relative composition variables, x^α, to be the number of letters of type α, divided by N. The number of different letters is h, and the set of labels α refers to the set of chemical possibilities, such as {purine,pyrimidine}, {A, C, G, T}, or the 20 amino acids. For an alphabet of size h, at each site along the sequence, there are $h - 1$ independent compositions $0 \leq x^\alpha \leq 1$, for $1 \leq \alpha \leq h - 1$. Alternatively, these compositions may be interpreted as normalized Hamming distances from a reference wild-type sequence. Therefore, the replication rate for a sequence S_i in the parallel model (1) is defined by the fitness function

$$r_i = Nf(x^1, x^2, \ldots, x^{h-1})$$ (3)

where the x^α are defined within the simplex $\sum_{\alpha=1}^{h-1} x^\alpha \leq 1$.

2.2 Schwinger Spin Coherent States Representation of the Parallel Model

We can express the parallel model in operator form, by generalizing the method presented in [19]. We define h kinds of creation and annihilation operators: $\hat{a}_\alpha^\dagger(j), \hat{a}_\alpha(j)$, $1 \leq \alpha \leq h$ and $1 \leq j \leq N$. These operators satisfy the commutation relations

$$\begin{aligned} \left[\hat{a}_\alpha(i), \hat{a}_\beta^\dagger(j)\right] &= \delta_{\alpha\beta}\delta_{ij} \\ \left[\hat{a}_\alpha(i), \hat{a}_\beta(j)\right] &= \left[\hat{a}_\alpha^\dagger(i), \hat{a}_\beta^\dagger(j)\right] = 0 \end{aligned}$$ (4)

These operators create/annihilate a sequence letter state $1 \leq \alpha \leq h$, at position $1 \leq j \leq N$ in the sequence. Since at each site there is a single letter, we enforce the constraint

$$\sum_{\alpha=1}^{h} \hat{a}_\alpha^\dagger(j)\hat{a}_\alpha(j) = 1$$ (5)

for all $1 \leq j \leq N$.

We define $n_\alpha^i(j)$ as the power on $\hat{a}_\alpha^\dagger(j)$ for the sequence state S_i, $1 \leq i \leq h^N$, defined by the vectors

$$|S_i\rangle = \prod_{j=1}^{N} |\vec{n}_j^i\rangle$$ (6)

where $|\vec{n}_j^i\rangle = \prod_{\alpha=1}^{h}[\hat{a}_\alpha^\dagger(j)]^{n_\alpha^i(j)}|0\rangle$. The constraint in (5) ensures that the condition $\sum_{\alpha=1}^{h} n_\alpha^i(j) = 1$ for all i, j.

We introduce the unnormalized population state vector

$$|\psi\rangle = \sum_{i=1}^{h^N} q(S_i)|S_i\rangle$$ (7)

which evolves in time according to the equation

$$\frac{d}{dt}|\psi\rangle = -\hat{H}|\psi\rangle$$ (8)

Here, the Hamiltonian operator, to highest order in N, is given by

$$-\hat{H} = \hat{m} + Nf(\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^{h-1}) \tag{9}$$

where $\hat{m}$ represents the mutation operator, and $\hat{x}^\alpha$ represents the compositions in operator form.

Let us first discuss the mutation operator $\hat{m}$. In the most general case, $h(h-1)$ possible different substitutions can occur at each site in the sequence, i.e. $\beta \to \alpha$, with mutation rate $\mu_{\alpha\beta}$ that need not be symmetric.

Each individual process can be written in the operator form

$$\vec{a}^\dagger(j)\tau^{\alpha\beta}\vec{a}(j) = \hat{a}^\dagger_\alpha(j)\hat{a}_\beta(j) + \sum_{\gamma \neq \alpha} \hat{a}^\dagger_\gamma(j)\hat{a}_\gamma(j) \tag{10}$$

which represents the creation of letter α by annihilation of letter β. Here, the matrices $\tau^{\alpha\beta}$ are explicitly defined by

$$\left[\tau^{\alpha\beta}\right]_{\rho\gamma} = \delta_{\rho\alpha}\delta_{\gamma\beta} + \delta_{\rho\gamma}(1 - \delta_{\rho\alpha}) \tag{11}$$

After these definitions, the more general expression for the mutation operator is

$$\hat{m} = \sum_{j=1}^{N} \sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta}\left[\vec{a}^\dagger(j)\tau^{\alpha\beta}\vec{a}(j) - \vec{a}^\dagger(j) \cdot \vec{a}(j)\right] \tag{12}$$

Let us now consider the Schwinger spin coherent state representation of the average base composition terms,

$$\hat{x}^\alpha = \frac{1}{N}\sum_{j=1}^{N} \hat{a}^\dagger_\alpha(j)\hat{a}_\alpha(j) \equiv \frac{1}{N}\sum_{j=1}^{N} \vec{a}^\dagger(j)\Theta^\alpha\vec{a}(j) \quad 1 \leq \alpha \leq h-1 \tag{13}$$

where we defined the matrices

$$\left[\Theta^\alpha\right]_{\rho\gamma} = \delta_{\rho\alpha}\delta_{\gamma\alpha} \tag{14}$$

We introduce the vector notation

$$\vec{\hat{x}} = (\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^{h-1}) \tag{15}$$

$$\boldsymbol{\Theta} = (\Theta^1, \Theta^2, \ldots, \Theta^{h-1}) \tag{16}$$

Considering the previous expressions, the Hamiltonian operator becomes

$$-\hat{H} = Nf\left[\frac{1}{N}\sum_{j=1}^{N} \vec{a}^\dagger(j)\boldsymbol{\Theta}\vec{a}(j)\right] + \sum_{j=1}^{N} \sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta}\left[\vec{a}^\dagger(j)\tau^{\alpha\beta}\vec{a}(j) - \vec{a}^\dagger(j) \cdot \vec{a}(j)\right] \tag{17}$$

2.3 Functional Integral Representation of the Parallel Model

We convert the operator representation of the parallel model into a functional integral form by introducing Schwinger spin coherent states [19]. We define a coherent state by

$$
|\vec{z}(j)\rangle = e^{\vec{a}^\dagger(j)\cdot\vec{z}(j)-\vec{z}^*(j)\cdot\vec{a}(j)}|0\rangle
$$

$$
= e^{-\frac{1}{2}\vec{z}^*(j)\cdot\vec{z}(j)} \sum_{m_1,m_2,\ldots,m_h=0}^{\infty} \prod_{\alpha=1}^{h} \frac{[z_\alpha(j)]^{m_\alpha}}{\sqrt{m_\alpha!}} |(m_1,m_2,\ldots,m_h)_j\rangle \tag{18}
$$

Coherent states satisfy the completeness relation

$$
I = \int \prod_{j=1}^{N} \frac{d\vec{z}^*(j)d\vec{z}(j)}{\pi^h} |\{\vec{z}\}\rangle\langle\{\vec{z}\}| \tag{19}
$$

The overlap between a pair of coherent states is given by

$$
\langle\vec{z}'(j)|\vec{z}(j)\rangle = e^{-\frac{1}{2}\{\vec{z}'^*(j)\cdot[\vec{z}'(j)-\vec{z}(j)]-[\vec{z}'^*(j)-\vec{z}^*(j)]\cdot\vec{z}(j)\}} \tag{20}
$$

To enforce the constraint (5), we introduce the projector

$$
\hat{P} = \prod_{j=1}^{N} \hat{P}(j) = \prod_{j=1}^{N} \Delta[\vec{a}^\dagger(j)\cdot\vec{a}(j)-1]
$$

$$
= \int_0^{2\pi} \prod_{j=1}^{N} \frac{d\lambda_j}{2\pi} e^{i\lambda_j[\vec{a}^\dagger(j)\cdot\vec{a}(j)-1]} \tag{21}
$$

At long times, due to the Perrone-Frobenius theorem, we find that the system evolution is dominated by the unique largest eigenvalue, f_m, of $-\hat{H}$ and its corresponding eigenvector $|\psi^*\rangle$, such that $e^{-\hat{H}t}|\{\vec{n}_0\}\rangle \sim e^{f_m t}|\psi^*\rangle$.

To evaluate this eigenvalue, we perform a Trotter factorization, for $\epsilon = t/M$, with $M \to \infty$, and introduce resolutions of the identity as defined by (19) at each time slice [19]

$$
e^{-\hat{H}t} = \lim_{M\to\infty} \int \left[\prod_{k=1}^{M}\prod_{j=1}^{N} \frac{d\vec{z}_k^*(j)d\vec{z}_k(j)}{\pi^h}\right] |\{\vec{z}_M\}\rangle \prod_{k=1}^{M} \langle\{\vec{z}_k\}|e^{-\epsilon\hat{H}}|\{\vec{z}_{k-1}\}\rangle\langle\{\vec{z}_0\}| \tag{22}
$$

We define the partition function

$$
Z = \text{Tr}\, e^{-\hat{H}t}\hat{P}
$$

$$
= \int_0^{2\pi}\left[\prod_{j=1}^{N}\frac{d\lambda_j}{2\pi}\right]e^{-i\lambda_j}\lim_{M\to\infty}\int\left[\prod_{k=1}^{M}\prod_{j=1}^{N}\frac{d\vec{z}_k^*(j)d\vec{z}_k(j)}{\pi^h}\right]e^{-S[\vec{z}^*,\vec{z}]} \tag{23}
$$

Here, we defined

$$
e^{-S[\vec{z}^*,\vec{z}]} = \prod_{k=1}^{M} \langle\{\vec{z}_k\}|e^{-\epsilon\hat{H}}|\{\vec{z}_{k-1}\}\rangle \tag{24}
$$

with the boundary condition $\vec{z}_0(j) = e^{i\lambda_j}\vec{z}_M(j)$ [19]. An explicit expression for the matrix element in the coherent states representation is

$$
\langle \{\vec{z}_k\} | e^{-\epsilon \hat{H}} | \{\vec{z}_{k-1}\} \rangle
$$

$$
= \exp\left(-\frac{1}{2} \sum_{j=1}^{N} [\vec{z}_k^*(j) \cdot \vec{z}_k(j) - 2\vec{z}_k^*(j) \cdot \vec{z}_{k-1}(j) + \vec{z}_{k-1}^*(j) \cdot \vec{z}_{k-1}(j)] \right.
$$

$$
- \epsilon N \sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta} + \epsilon N f \left[\frac{1}{N} \sum_{j=1}^{N} \vec{z}_k^*(j) \mathbf{\Theta} \vec{z}_{k-1}(j) \right]
$$

$$
\left. + \epsilon \sum_{j=1}^{N} \vec{z}_k^*(j) \sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta} \tau^{\alpha\beta} \vec{z}_{k-1}(j) \right)
\tag{25}
$$

Let us now introduce an $(h-1)$-component vector field $\vec{x}_k = (x_k^1, x_k^2, \ldots, x_k^{h-1})$, with

$$
x_k^\alpha = \frac{1}{N} \sum_{j=1}^{N} \vec{z}_k^*(j) \Theta^\alpha \vec{z}_{k-1}(j)
\tag{26}
$$

We make this definition by introducing an integral representation of the corresponding delta function

$$
1 = \int \mathcal{D}[\vec{x}] \prod_{k=1}^{M} \delta^{(h-1)} \left[\vec{x}_k - \frac{1}{N} \sum_{j=1}^{N} \vec{z}_k^*(j) \mathbf{\Theta} \vec{z}_{k-1}(j) \right]
$$

$$
= \int \left[\prod_{k=1}^{M} \prod_{\alpha=1}^{h-1} \frac{d\bar{x}_k^\alpha dx_k^\alpha}{2\pi} \right] e^{i \sum_{k=1}^{M} \vec{\bar{x}}_k \cdot \vec{x}_k - \frac{i}{N} \sum_{k=1}^{M} \sum_{j=1}^{N} \vec{z}_k^*(j) \vec{\bar{x}}_k \cdot \mathbf{\Theta} \vec{z}_{k-1}(j)}
$$

$$
= \int \left[\prod_{k=1}^{M} \prod_{\alpha=1}^{h-1} \frac{i\epsilon N d\bar{x}_k^\alpha dx_k^\alpha}{2\pi} \right] e^{-\epsilon N \sum_{k=1}^{M} \vec{\bar{x}}_k \cdot \vec{x}_k + \epsilon \sum_{k=1}^{M} \sum_{j=1}^{N} \vec{z}_k^*(j) \vec{\bar{x}}_k \cdot \mathbf{\Theta} \vec{z}_{k-1}(j)}
\tag{27}
$$

Inserting this into the functional integral (23), we have [19]

$$
Z = \lim_{M\to\infty} \int \mathcal{D}[\vec{\bar{x}}] \mathcal{D}[\vec{x}] e^{N \sum_{k=1}^{M} [f(\vec{x}_k) - \vec{\bar{x}}_k \cdot \vec{x}_k - \sum_{\alpha\neq\beta=1}^{h} \mu_{\alpha\beta}]}
$$

$$
\times \int \mathcal{D}[\vec{z}^*] \mathcal{D}[\vec{z}] \mathcal{D}[\lambda] \prod_{j=1}^{N} e^{-i\lambda_j} e^{\epsilon \sum_{l,k=1}^{M} \vec{z}_k^*(j) S_{kl}(j) \vec{z}_l(j)} \Bigg|_{\{\vec{z}_0\} = \{e^{i\lambda_j}\vec{z}_M\}}
\tag{28}
$$

The matrix $S(j)$ has the structure

$$
S(j) = \begin{pmatrix}
I & 0 & 0 & \ldots & -e^{-i\lambda_j} A_1(j) \\
-A_2(j) & I & 0 & \ldots & 0 \\
0 & -A_3(j) & I & \ldots & 0 \\
& & & \ddots & \\
0 & & \ldots & -A_M(j) & I
\end{pmatrix}
\tag{29}
$$

where $A_k(j) = I + \epsilon[\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k \cdot \boldsymbol{\Theta}]$. After performing the Gaussian integration over the coherent state fields, we obtain

$$Z = \lim_{M \to \infty} \int \mathcal{D}[\vec{\bar{x}}]\mathcal{D}[\vec{x}]e^{\epsilon N \sum_{k=1}^M [f(\vec{x}_k) - \vec{\bar{x}}_k \cdot \vec{x}_k - \sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}]}$$
$$\times \int \mathcal{D}[\lambda]\prod_{j=1}^N e^{-i\lambda_j}\left[\det S(j)\right]^{-1} \tag{30}$$

Here,

$$\det S(j) = \det\left[I - e^{i\lambda_j}\prod_{k=1}^M A_k(j)\right]$$
$$= \det\left[I - e^{i\lambda_j}\hat{T}e^{\epsilon \sum_{k=1}^M [\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k \cdot \boldsymbol{\Theta}]}\right]$$
$$= e^{\text{Tr}\ln[I - e^{i\lambda_j}\hat{T}e^{\epsilon \sum_{k=1}^M [\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k \cdot \boldsymbol{\Theta}]}]} \tag{31}$$

where the operator $\hat{T}$ indicates time ordering. Substituting this result in the partition function, we obtain

$$Z = \lim_{M \to \infty} \int \mathcal{D}[\vec{\bar{x}}]\mathcal{D}[\vec{x}]e^{\epsilon N \sum_{k=1}^M [f(\vec{x}_k) - \vec{\bar{x}}_k \cdot \vec{x}_k - \sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}]}$$
$$\times \int \mathcal{D}[\lambda]\prod_{j=1}^N e^{-i\lambda_j}e^{-\text{Tr}\ln[I - e^{i\lambda_j}\hat{T}e^{\epsilon \sum_{k=1}^M [\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k \cdot \boldsymbol{\Theta}]}]}$$
$$= \lim_{M \to \infty} \int \mathcal{D}[\vec{\bar{x}}]\mathcal{D}[\vec{x}]e^{\epsilon N \sum_{k=1}^M [f(\vec{x}_k) - \vec{\bar{x}}_k \cdot \vec{x}_k - \sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}]}\prod_{j=1}^N Q \tag{32}$$

with

$$Q = \lim_{M \to \infty} \text{Tr}\,\hat{T}\prod_{k=1}^M\left[I + \epsilon\left(\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k \cdot \boldsymbol{\Theta}\right)\right]$$
$$= \lim_{M \to \infty} \text{Tr}\,\hat{T}e^{\epsilon \sum_{k=1}^M [\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k \cdot \boldsymbol{\Theta}]}$$
$$= \text{Tr}\,\hat{T}e^{\int_0^t dt'[\sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}\tau^{\alpha\beta} + \vec{x}(t') \cdot \boldsymbol{\Theta}]} \tag{33}$$

After taking the limit $M \to \infty$, (32) becomes

$$Z = \int \mathcal{D}[\vec{\bar{x}}]\mathcal{D}[\vec{x}]e^{-S[\vec{\bar{x}},\vec{x}]} \tag{34}$$

where the effective action is given by

$$S[\vec{\bar{x}},\vec{x}] = -N\int_0^t dt'[f(\vec{x}(t')) - \vec{\bar{x}}(t') \cdot \vec{x}(t') - \sum_{\alpha \neq \beta=1}^h \mu_{\alpha\beta}] - N\ln Q \tag{35}$$

Here, we have defined

$$Q = \text{Tr} \, \hat{T} \, e^{\int_0^t dt' [\sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} \tau^{\alpha\beta} + \sum_{\alpha=1}^{h-1} \bar{x}^\alpha(t') \Theta^\alpha]} \tag{36}$$

2.4 The Large N Limit of the Parallel Model is a Saddle Point

Considering that the sequence length N is very large, $N \to \infty$, we can evaluate the functional integral (34) for the partition function by looking for a saddle point. With the action defined in (36), we have

$$\left. \frac{\delta S}{\delta x^\alpha} \right|_{\bar{x}_c, \vec{\bar{x}}_c} = -N \left(\left. \frac{\partial f[\vec{x}]}{\partial x^\alpha} \right|_c - \bar{x}_c^\alpha \right) = 0$$
$$\left. \frac{\delta S}{\delta \bar{x}^\alpha} \right|_{\bar{x}_c, \vec{\bar{x}}_c} = -N \left(-x_c^\alpha + \frac{1}{Q} \left. \frac{\delta Q}{\delta \bar{x}^\alpha} \right|_{\bar{x}_c, \vec{\bar{x}}_c} \right) = 0 \tag{37}$$

We denote the value of the action at the saddle point by S_c. We have therefore the system of equations

$$\bar{x}_c^\alpha = \left. \frac{\partial f[\vec{x}]}{\partial x^\alpha} \right|_c \tag{38}$$

$$x_c^\alpha = \langle \Theta^\alpha \rangle, \, 1 \leq \alpha \leq h - 1 \tag{39}$$

where we defined

$$\langle (\cdot) \rangle = \frac{\text{Tr}(\cdot) e^{t[\sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} \tau^{\alpha\beta} + \sum_{\alpha=1}^{h-1} \bar{x}_c^\alpha \Theta^\alpha]}}{\text{Tr} \, e^{t[\sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} \tau^{\alpha\beta} + \sum_{\alpha=1}^{h-1} \bar{x}_c^\alpha \Theta^\alpha]}} \tag{40}$$

After this saddle-point analysis, we obtain a general expression for the mean fitness f_m of the population, for an arbitrary microscopic fitness function $f(\vec{x})$,

$$f_m = \lim_{N, t \to \infty} \frac{-S_c}{Nt} = \max_{\{\vec{x}_c, \vec{\bar{x}}_c\}} \left[f(\vec{x}_c) - \vec{\bar{x}}_c \cdot \vec{x}_c - \sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} + \lambda_{\max} \right] \tag{41}$$

with $\lambda_{\max}$ defined as

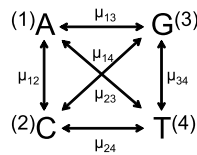$$\lambda_{\max} = \lim_{t \to \infty} \frac{\ln Q}{t} \tag{42}$$

and corresponding to the largest eigenvalue of the matrix

$$\mathbf{M}(\vec{\bar{x}}_c, \{\mu_{\alpha\beta}\}) = \sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} \tau^{\alpha\beta} + \sum_{\alpha=1}^{h-1} \bar{x}_c^\alpha \Theta^\alpha \tag{43}$$

As shown in detail in Appendix 1, the compositions $\bar{x}_c^\alpha$ can be eliminated to reduce (41) to the final expression

$$f_m^{(h)} = \max_{\{x_c^1, x_c^2, \dots, x_c^{h-1}\}} \left\{ f(x_c^1, x_c^2, \dots, x_c^{h-1}) + \sum_{\alpha \neq \beta = 1}^{h-1} \mu_{\alpha\beta} \left[\sqrt{x_c^\alpha x_c^\beta} - x_c^\alpha \right] \right.$$

Fig. 1 The generalized mutation scheme

$$
+ \sum_{\alpha=1}^{h-1} \mu_{\alpha h} \left[\sqrt{ x_c^\alpha \left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma \right) } - x_c^\alpha \right]
$$

$$
+ \sum_{\beta=1}^{h-1} \mu_{h\beta} \left[\sqrt{ \left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma \right) x_c^\beta + \sum_{\gamma=1}^{h-1} x_c^\gamma } - 1 \right] \Bigg\} \tag{44}
$$

Here, the compositions $0 \leq x_c^\alpha \leq 1$ are defined within the simplex $\sum_{\alpha=1}^{h-1} x_c^\alpha \leq 1$.

2.5 The Purine/Pyrimidine Alphabet $h = 2$

Most analytical and numerical studies of quasi-species models have been formulated in the past [12, 14, 19, 20] by using a coarse-grained alphabet of nucleotides, where the nucleotide bases are lumped into purines and pyrimidines, and hence $h = 2$. The maximum principle for this binary alphabet can be derived from the general expression (44), by assuming a symmetric mutation rate $\mu_{12} = \mu_{21} = \mu$, and by noticing that a single composition $x_c^1 \equiv x_c$ (or normalized Hamming distance) is required in this case,

$$
f_m^{(2)} = \max_{\{0 \leq x_c \leq 1\}} \left\{ f(x_c) + \mu[2\sqrt{x_c(1-x_c)} - 1] \right\} \tag{45}
$$

It is customary to use in this case a magnetization coordinate [12, 14, 19, 20] defined as $\xi_c = 1 - 2x_c$, and hence (45) becomes

$$
f_m^{(2)} = \max_{\{-1 \leq \xi_c \leq 1\}} \left\{ f(\xi_c) + \mu\sqrt{1 - \xi_c^2} - \mu \right\} \tag{46}
$$

Equation (46) is a well known result, which has been obtained in the past with different methods [12, 14], including a version of the Schwinger boson method employed in the present work [19]. It is a special case of (44).

2.6 The Nucleic Acid Alphabet $h = 4$

In the nucleic acids RNA or DNA, the alphabet is constituted by the monomers of these polymeric chains, which are $h = 4$ different nucleotides A, C, G, T/U. The general mutation scheme displayed in Fig. 1 is represented by setting $h = 4$ in our general solution (44), with 3 independent compositions (or normalized Hamming distances) x_c^1, x_c^2, x_c^3

$$
f_m = \max_{\{x_c^1, x_c^2, x_c^3\}} \left\{ f(x_c^1, x_c^2, x_c^3) + \sum_{\alpha \neq \beta = 1}^{3} \mu_{\alpha\beta} \left[\sqrt{x_c^\alpha x_c^\beta} - x_c^\alpha \right] \right.
$$

$$
+ \sum_{\alpha=1}^{3} \mu_{\alpha 4} \left[\sqrt{ x_c^\alpha \left(1 - \sum_{\gamma=1}^{3} x_c^\gamma \right) } - x_c^\alpha \right]
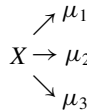$$

Fig. 2 The Kimura 3 ST mutation scheme
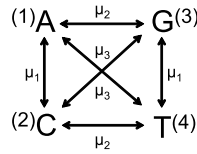
$$+ \sum_{\beta=1}^{3} \mu_{4\beta} \left[\sqrt{\left(1 - \sum_{\gamma=1}^{3} x_c^\gamma \right) x_c^\beta} + \sum_{\gamma=1}^{3} x_c^\gamma - 1 \right] \right\} \tag{47}$$

Here, the compositions x_c^α are defined within the simplex $x_c^1 + x_c^2 + x_c^3 \le 1$.

An interesting reduction of this general model is provided by the Kimura 3 ST mutation scheme [15–17, 22], Fig. 2. The Kimura 3 ST mutation scheme considers mutations in three independent directions, with rates μ_1, μ_2, μ_3

$$X \to \begin{matrix} \nearrow \mu_1 \\ \mu_2 \\ \searrow \mu_3 \end{matrix}$$

Accordingly, three components of the Hamming distance between a pair of sequences S_i and S_j are defined as follows

$$d_1(S_i, S_j) = \#_{A \leftrightarrow C}(S_i, S_j) + \#_{G \leftrightarrow T}(S_i, S_j)$$
$$d_2(S_i, S_j) = \#_{A \leftrightarrow G}(S_i, S_j) + \#_{C \leftrightarrow T}(S_i, S_j) \tag{48}$$
$$d_3(S_i, S_j) = \#_{A \leftrightarrow T}(S_i, S_j) + \#_{C \leftrightarrow G}(S_i, S_j)$$

Here, $\#_{X \leftrightarrow Y}(S_i, S_j)$ is the number of sites at which X and Y are exchanged between sequences S_i and S_j. The total Hamming distance between sequences S_i and S_j is given by

$$d(S_i, S_j) = d_1(S_i, S_j) + d_2(S_i, S_j) + d_3(S_i, S_j) \tag{49}$$

The mutation rate is therefore modeled by the function

$$\mu_{ij} = \begin{cases} \mu_\alpha, & d_\alpha(S_i, S_j) = d(S_i, S_j) = 1 \\ -N \sum_{\alpha=1}^{3} \mu_\alpha, & S_i = S_j \\ 0, & d(S_i, S_j) > 1 \end{cases} \tag{50}$$

To be consistent with existing notation in the quasispecies literature, we define 3 independent variables, which are simply transformations of the composition variables, and which are called 'surplus' variables in the literature

$$u_1(S_i) = 1 - \frac{2}{N}[d_1(S_i, S_w) + d_3(S_i, S_w)]$$

$$u_2(S_i) = 1 - \frac{2}{N}[d_2(S_i, S_w) + d_3(S_i, S_w)] \tag{51}$$

$$u_3(S_i) = 1 - \frac{2}{N}[d_1(S_i, S_w) + d_2(S_i, S_w)]$$

Notice that according to this definition, the maximum value of $u = (u_1 + u_2 + u_3)/3 = 1$ is reached when $d_1 = d_2 = d_3 = 0$, that is the sequence S_i being identical to the wild type S_w. The minimum value for the average base composition surplus is obtained when one of the Hamming distance components, say $d_i = N$, while the others are null $d_{j \neq i} = 0$. Then, $d = d_i = N$ and $u = -1/3$.

The Kimura 3 ST mutation scheme result is obtained from the general (47) if the following symmetries are assumed for the mutation rates

$$
\begin{aligned}
\mu_{12} &= \mu_{21} = \mu_{34} = \mu_{43} \equiv \mu_1 \\
\mu_{13} &= \mu_{31} = \mu_{24} = \mu_{42} \equiv \mu_2 \\
\mu_{14} &= \mu_{41} = \mu_{23} = \mu_{32} \equiv \mu_3
\end{aligned}
\tag{52}
$$

We follow the quasispecies literature convention and define the 3 independent 'ancestral distribution' coordinates ξ_c^α (the subscript c denoting the saddle-point limit), after (51)

$$
\begin{aligned}
\xi_c^1 &= 1 - 2(x_c^1 + x_c^3) \\
\xi_c^2 &= 1 - 2(x_c^2 + x_c^3) \\
\xi_c^3 &= 1 - 2(x_c^1 + x_c^2)
\end{aligned}
\tag{53}
$$

The ancestral distribution variable ξ is defined as the steady state analog of the 'surplus,' but for the time-reversed evolution process [23].

After some algebra, we obtain

$$
\begin{aligned}
f_m = \max_{\{\xi_c^1, \xi_c^2, \xi_c^3\}} \Bigg\{ & f(\xi_c^1, \xi_c^2, \xi_c^3) - (\mu_1 + \mu_2 + \mu_3) \\
& + \frac{\mu_1}{2} \Big[\sqrt{(1 + \xi_c^1 + \xi_c^2 + \xi_c^3)(1 - \xi_c^1 - \xi_c^2 + \xi_c^3)} \\
& + \sqrt{(1 + \xi_c^1 - \xi_c^2 - \xi_c^3)(1 - \xi_c^1 + \xi_c^2 - \xi_c^3)} \Big] \\
& + \frac{\mu_2}{2} \Big[\sqrt{(1 + \xi_c^1 + \xi_c^2 + \xi_c^3)(1 + \xi_c^1 - \xi_c^2 - \xi_c^3)} \\
& + \sqrt{(1 - \xi_c^1 + \xi_c^2 - \xi_c^3)(1 - \xi_c^1 - \xi_c^2 + \xi_c^3)} \Big] \\
& + \frac{\mu_3}{2} \Big[\sqrt{(1 + \xi_c^1 + \xi_c^2 + \xi_c^3)(1 - \xi_c^1 + \xi_c^2 - \xi_c^3)} \\
& + \sqrt{(1 + \xi_c^1 - \xi_c^2 - \xi_c^3)(1 - \xi_c^1 - \xi_c^2 + \xi_c^3)} \Big] \Bigg\}
\end{aligned}
\tag{54}
$$

From this general expression, the average composition 'surplus' $\vec{u} = (u_1, u_2, u_3)$ is obtained by applying the self-consistent condition $f(\vec{u}) = f_m$. This result is equivalent to that derived by [17].

2.7 Analytic Results for the Symmetric Mutational Scheme

For a symmetric mutational scheme, $\mu_1 = \mu_2 = \mu_3 \equiv \mu$, we specialize the general (54) by setting $\xi_c^1 = \xi_c^2 = \xi_c^3 \equiv \xi_c$, and $u_i = u$, and thus obtaining an expression for the mean fitness

$$f_m = \max_{-\frac{1}{3} \leq \xi_c \leq 1} \left\{ f(\xi_c) - \frac{3}{2}\mu(1 + \xi_c) + \frac{3}{2}\mu\sqrt{(1 - \xi_c)(1 + 3\xi_c)} \right\} \tag{55}$$

This result is equivalent to that derived by [17]. We remark that (55) represents an exact analytical expression for the mean fitness of the population, for any arbitrary microscopic fitness $f(u)$, with the assumption of symmetric mutation rates $\mu_1 = \mu_2 = \mu_3 = \mu$. From this exact expression, the average composition u is obtained by applying the self-consistency condition $f_m = f(u)$. In the following sections, we apply (55) to analyze in detail some examples of microscopic fitness functions: The sharp peak landscape, a Fujiyama landscape, a quadratic fitness landscape, and a quartic fitness landscape. We note that the $h = 2$ case contains a symmetry in the mutation terms about $\xi = 0$. In the general $h > 2$, this symmetry will be lost. As we will see, loss of this symmetry leads to a change in the order of the error catastrophe phase transition.

2.7.1 The Sharp Peak Landscape

We shall first consider the sharp peak landscape, which is described by the function

$$f(u) = A\delta_{u,1} \tag{56}$$

That is, only sequences identical to the wild-type replicate with a rate $A > 0$. From (55), we notice that this implies: $\xi_c = 1$, if $A > 3\mu$, or $\xi_c = 0$ otherwise. Therefore, we obtain for the mean replication rate

$$f_m = \begin{cases} A - 3\mu, & A > 3\mu \\ 0, & A < 3\mu \end{cases} \tag{57}$$

The fraction of the population at the wild-type p_w is obtained from the self-consistent condition $f_m = p_w A$,

$$p_w = \begin{cases} 1 - \frac{3\mu}{A}, & A > 3\mu \\ 0, & A < 3\mu \end{cases} \tag{58}$$

There exists an error threshold in this case, which is given by the critical value $A_{crit} = 3\mu$, as shown in (57), (58) and displayed in Fig. 3. The phase transition is first order as a function of A/μ.

One may compare this result with the error threshold observed in the binary alphabet case, which is [19] $A_{crit} = \mu$. This result is intuitive, because in the 4 letters alphabet, there exist 3 mutation channels to escape from the wild type instead of just one as in the binary alphabet, and therefore a stronger selection pressure is required to retain the wild-type features.

Fig. 3 Average composition u, magnetization ξ_c and fraction of the population at the wild-type sequence p_w, as a function of the parameter A/μ, for the sharp peak fitness
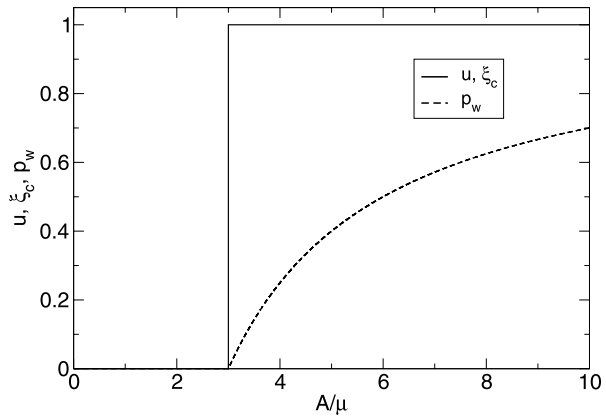
2.7.2 The Fujiyama Fitness Landscape

The Fujiyama landscape is obtained as a linear function of the composition

$$f[\vec{u}] = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 \tag{59}$$

We will present analytical results for the symmetric case $\alpha_i \equiv \alpha$, $\mu_i \equiv \mu$. Thus, $\xi_c^1 = \xi_c^2 = \xi_c^3 = \xi_c$. Substituting in (55), we have

$$f_m = \max_{-\frac{1}{3} \le \xi_c \le 1} \left\{ 3\alpha\xi_c - \frac{3}{2}\mu(1 + \xi_c) + \frac{3}{2}\mu\sqrt{(1 - \xi_c)(1 + 3\xi_c)} \right\} \tag{60}$$

We look for a maximum

$$\frac{\partial f_m}{\partial \xi_c} = 3\alpha - \frac{3}{2}\mu + \frac{3}{2}\mu\frac{2 - 6\xi_c}{2\sqrt{1 + 2\xi_c - 3\xi_c^2}} = 0 \tag{61}$$

From this equation, we obtain

$$\xi_c = \frac{1}{3}\left(1 + \frac{2\alpha - \mu}{\sqrt{\alpha^2 - \alpha\mu + \mu^2}}\right) \tag{62}$$

To obtain the average base composition u, we apply the self-consistent condition $f_m = f(u) = 3\alpha u$, to obtain

$$u = \frac{1}{3}\left(1 - 2\frac{\mu}{\alpha} + \frac{2}{\alpha}\sqrt{\alpha^2 - \mu\alpha + \mu^2}\right) \tag{63}$$

Clearly, no phase transition is observed in this fitness landscape, as $0 < u < 1$ for $0 < \alpha < \infty$. This result is in agreement with the analysis presented in [15], were a quantum spin chain formulation was employed.

2.7.3 Quadratic Fitness Landscape

The quadratic fitness landscape is given by the general quadratic form

$$f(\vec{u}) = \sum_{i=1}^{3} \left(\frac{\beta_i}{2} u_i^2 + \alpha_i u_i \right) \tag{64}$$

We will present the analytical solution for the symmetric case $\alpha_i \equiv \alpha$, $\beta_i \equiv \beta$, with the symmetric mutation scheme $\mu_i \equiv \mu$. Under these conditions, we have $\xi_c^1 = \xi_c^2 = \xi_c^3 \equiv \xi_c$, and from (55) we have for the mean fitness

$$f_m = \max_{-\frac{1}{3} \leq \xi_c \leq 1} \left\{ \frac{3}{2}\beta\xi_c^2 + 3\alpha\xi_c - \frac{3}{2}\mu(1+\xi_c) + \frac{3}{2}\mu\sqrt{(1-\xi_c)(1+3\xi_c)} \right\} \tag{65}$$

The maximum is obtained from the equation

$$\frac{\partial f_m}{\partial \xi_c} = 3\beta\xi_c + 3\alpha - \frac{3}{2}\mu + \frac{3}{2}\mu\frac{2-6\xi_c}{2\sqrt{1+2\xi_c-3\xi_c^2}} = 0 \tag{66}$$

From (66), we obtain

$$\beta\xi_c + \alpha - \frac{\mu}{2} = \frac{\mu}{2}\frac{3\xi_c-1}{\sqrt{1+2\xi_c-3\xi_c^2}} \tag{67}$$

As shown in Appendix 2, this equation can be cast in the form of a quartic equation, whose roots are the values of ξ_c. The average composition u is finally obtained through the self-consistency equation

$$f_m = f(u) = \frac{3}{2}\beta u^2 + 3\alpha u \tag{68}$$

We find that the error threshold transition towards a selective phase for $\alpha = 0$ is defined by $\xi_c > 0$, $u > 0$, at $\beta > 3/2\mu$. The value of u is continuous at the transition, as it is straightforward to check from (66) $f_m(\xi_c = 0) = f_m(\xi_c = 2/3, \beta = 3\mu/2) = 0$, which implies after (68) (for $\alpha = 0$) that $u = 0$ when approaching the critical point $\beta = 3\mu/2$ from both sides. However, ξ_c jumps from 0 [for $\beta \to (3\mu/2)^-$] to 2/3 [for $\beta \to (3\mu/2)^+$] (Appendix 2). To analyze the order of the transition, we expand (65) as a quadratic polynomial in ξ_c in a neighborhood of the critical point $\beta \simeq 3\mu/2$,

$$f_m(\xi_c) = \begin{cases} \frac{3}{2}(\beta - 3\mu/2)\xi_c^2, & \beta < 3\mu/2 \\ \frac{2}{3}\beta - \mu + (2\beta - 3\mu)(\xi_c - \frac{2}{3}) + (\frac{3}{2}\beta - 3\mu)(\xi_c - \frac{2}{3})^2, & \beta > 3\mu/2 \end{cases} \tag{69}$$

we find that the first derivative $df_m/d\beta[\beta \to (3\mu/2)^-] = 0$, while $df_m/d\beta[\beta \to (3\mu/2)^+] = 2/3$, and thus it has a discontinuous jump from 0 to 2/3. Therefore, the phase transition is first order as a function of β/μ. We notice that the order of the phase transition, for a similar quadratic fitness landscape, is found to be of second order for a binary alphabet [19].

When $0 \leq \alpha/\beta \leq \frac{1}{3}(\sqrt{\frac{4}{3}} - 1)$, as shown in Appendix 2, we find a finite jump in the magnetization from $\xi_{c,+}$ to $\xi_{c,-}$, with $\xi_{c,\pm} = 1/3(1 \pm \sqrt{1 - 18\alpha/\beta - 27(\alpha/\beta)^2})$. This result is in agreement with [15], where an alternative method of quantum spin chains was applied for the derivation. A complete analysis of the different possible cases other than this, is presented in Appendix 2.

Fig. 4 Average composition u and magnetization ξ_c as a function of the parameter β/μ, for the quadratic fitness when $\alpha = 0$
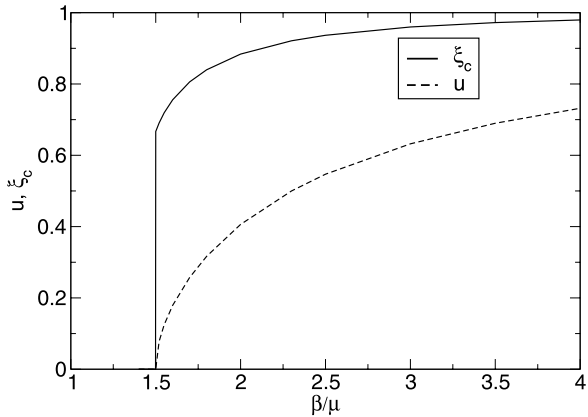
2.8 Quartic Fitness Landscape

As a final example, we consider a quartic fitness landscape

$$f(\vec{u}) = \sum_{i=1}^{3} \frac{\beta_i}{4} u_i^4 \tag{70}$$

As in the previous cases, we consider the symmetric mutation rates $\mu_i \equiv \mu$, $\beta_i \equiv \beta$, and hence $\xi_{c,i} \equiv \xi_c$. Considering this fitness function in the general equation (55), we have that the mean fitness is given by the analytical expression

$$f_m = \max_{\{-\frac{1}{3} \leq \xi_c \leq 1\}} \left\{ \frac{3}{4}\beta\xi_c^4 - \frac{3}{2}\mu(1+\xi_c) + \frac{3}{2}\mu\sqrt{(1-\xi_c)(1+3\xi_c)} \right\} \tag{71}$$

The average composition u is obtained by applying the self-consistent condition

$$f(u) = \frac{3}{4}\beta u^4 = f_m \tag{72}$$

In Fig. 5, we present the values of u and ξ_c, as obtained from (71), (72), as a function of the parameter β/μ. A discontinuous jump in the bulk magnetization from $\xi_c = 0$ to $\xi_c = 0.971618$ is observed at $\beta/\mu = 3.67653$. By expanding (71) near the critical point, after similar procedure as in the quadratic fitness case, we find a discontinuous jump in the derivative $df_m/d\beta$, from 0 to 0.66841101. Therefore, the phase transition is first order in β/μ. The average composition u, however, experiences a fast but smooth transition. This behavior is much alike the one observed in the sharp peak fitness landscape, (57) and Fig. 3, except for the fact that the average composition u is continuous at the transition. Indeed, from a purely mathematical perspective, a fitness function following a power law $f_n(u) = ku^n$, for $0 < u < 1$, will satisfy the limit

$$\lim_{n \to \infty} f_n(u) = k\delta_{u,1} \tag{73}$$

which is precisely the sharp peak landscape, (56).

Fig. 5 Average composition u and magnetization ξ_c as a function of the parameter β/μ, for the quartic fitness landscape
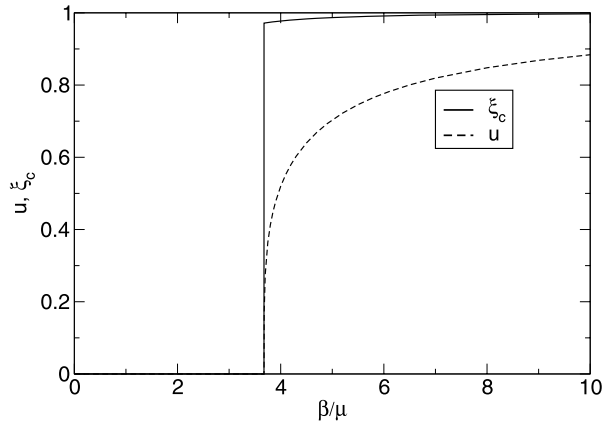
2.9 Symmetric Case, General h, with Application to the Amino Acid Alphabet $h = 20$

An alternative language to describe molecular evolution is in terms of mutation and selection acting over the translated protein sequence, which is drawn from an $h = 20$ amino acid alphabet. For the parallel model, the time evolution of an infinite population of protein sequences is described by the system of differential equations (2), with $h = 20$. Thus, we are lead to consider the $h = 20$ case. We first consider the symmetric case with general h.

For an alphabet of arbitrary size h, a symmetrical mutation scheme $\mu_{\alpha\beta} = \mu$, and a symmetrical fitness function that leads to $x_c^\alpha = x_c$, we define a magnetization coordinate $\xi_c = 1 - hx_c$, and (44) reduces to

$$f_m = \max_{\{-1/(h-1) \leq \xi_c \leq 1\}} \left\{ f(\xi_c) + (h-1)\mu \left[\frac{2}{h}\sqrt{(1-\xi_c)(1+(h-1)\xi_c)} \right. \right.$$
$$\left. \left. + \frac{h-2}{h}(1-\xi_c) - 1 \right] \right\} \tag{74}$$

As an example of application of (74), we consider the sharp peak fitness landscape $f(\xi_c) = A\delta_{\xi_c,1}$. Then, from (74) we obtain the mean fitness

$$f_m = \begin{cases} A - (h-1)\mu, & A > (h-1)\mu \\ 0, & A \leq (h-1)\mu \end{cases} \tag{75}$$

We obtain the fraction of sequences in the wild type, p_w, by applying the self-consistency condition $f_m = Ap_w$, which yields

$$p_w = \begin{cases} 1 - \frac{(h-1)\mu}{A}, & A > (h-1)\mu \\ 0, & A \leq (h-1)\mu \end{cases} \tag{76}$$

This result is intuitive, since there exists $h - 1$ independent mutation channels for the sequence to escape from the wild type. Moreover, for a general alphabet of size h, a first order phase transition occurs at the critical point $A_{crit}^h = (h-1)\mu$.

As a second example, we consider the quadratic fitness landscape for an alphabet of size h, $f(\xi_c) = (h-1)\beta\xi_c^2/2$. For the quadratic fitness function we can work out the order

of the phase transition for general h. We consider $f(\xi) = (h-1)\beta\xi^2/2$. There is a phase transition at $\beta_{crit} = 2\mu(h-1)/h$. The magnetization jumps from $\xi_c = 0$ at $\beta \to \beta_{crit}^-$ to $\xi_c = (h-2)/(h-1)$ at $\beta \to \beta_{crit}^+$. The first derivative at the critical point is:

$$\frac{df_m}{d\beta} = \begin{cases} 0, & \beta < 2\mu(h-1)/h \\ (h-2)^2/[2(h-1)], & \beta > 2\mu(h-1)/h \end{cases} \tag{77}$$

Thus, the jump in the first derivative is $(h-2)^2/[2(h-1)]$. Thus, the transition is second order for $h = 2$ and first order for $h > 2$.

3 The h-states Eigen Model

The Eigen model conceptually differs from the parallel or Kimura model because it is assumed that mutations arise as a consequence of errors in the replication process. For an alphabet of size h, the system of equations which describes the time evolution of the probabilities p_i, with $1 \le i \le h^N$, is

$$\frac{dp_i}{dt} = \sum_{j=1}^{N}[B_{ij}r_j - \delta_{ij}D_j]p_j - p_i\left[\sum_{j=1}^{N}(r_j - D_j)p_j\right] \tag{78}$$

Here, r_i is the replication rate of sequence S_i, and the components of the matrix

$$B_{ij} = (1-q)^{N-d_{ij}}\left(\frac{q}{h-1}\right)^{d_{ij}} \tag{79}$$

represent the transition rates from sequence S_j into S_i, where $1-q$ is the probability to copy a nucleotide without error, and q is the probability per site for a base substitution during the replication process.

We consider a generalized version of this, by considering that the base substitution probabilities are not necessarily identical nor symmetric. That is, for a base substitution $\beta \to \alpha$, we consider a probability $q_{\alpha\beta} \ne q_{\beta\alpha}$. Correspondingly, we define $h-1$ independent base compositions $0 \le x^\alpha \le 1$, which can also be interpreted as normalized Hamming distances with respect to the h-th reference species, $x^\alpha = d^\alpha/N$. For this generalized mutation scheme, the transition rate matrix components are defined by

$$B_{ij} = (1-q)^{N(1-\sum_{\alpha=1}^{h-1}x^\alpha)}\prod_{\alpha\ne\beta=1}^{h}\left(q_{\alpha\beta}\right)^{Nx^\alpha} \tag{80}$$

Here, $q = \sum_{\alpha\ne\beta=1}^{h}q_{\alpha\beta}$, and the x^α are as in Sect. 2.1.

3.1 The h-states Eigen Model in Operator Form

By similar arguments as in the parallel model, we formulate a Hamiltonian operator for the Eigen model

$$-\hat{H} = \prod_{j=1}^{N}\left[(1-q)\vec{\hat{a}}^\dagger(j)\cdot\vec{\hat{a}}(j) + \sum_{\alpha\ne\beta=1}^{h}q_{\alpha\beta}\vec{\hat{a}}^\dagger(j)\tau^{\alpha\beta}\vec{\hat{a}}(j)\right]$$

$$\times Nf\left[\frac{1}{N}\sum_{l=1}^{N}\vec{a}^{\dagger}(l)\mathbf{\Theta}\vec{a}(l)\right] - Nd\left[\frac{1}{N}\sum_{l=1}^{N}\vec{a}^{\dagger}(l)\mathbf{\Theta}\vec{a}(l)\right] \qquad (81)$$

Here, the matrices $\tau^{\alpha\beta}$ are defined as in the parallel model by (11) and in the matrix array $\mathbf{\Theta} = (\Theta^1, \Theta^2, \ldots, \Theta^{h-1})$, the matrices Θ^α are defined as in (14). Let us define the coefficients $\mu_{\alpha\beta} = Nq_{\alpha\beta}$. The degradation function is given by $D_i = Nd(x^1, x^2, \ldots, x^{h-1})$. Then, we have for $q \ll 1$

$$-N\ln(1-q) \simeq Nq = N\sum_{\alpha\neq\beta=1}^{h} q_{\alpha\beta} = \sum_{\alpha\neq\beta=1}^{h} \mu_{\alpha\beta} \qquad (82)$$

The Hamiltonian operator (81) is expressed, to $\mathcal{O}(1/N)$, by

$$-\hat{H} = Ne^{-\sum_{\alpha\neq\beta=1}^{h}\mu_{\alpha\beta}}e^{\sum_{j=1}^{N}[\sum_{\alpha\neq\beta=1}^{h}\frac{\mu_{\alpha\beta}}{N}\vec{a}^{\dagger}(j)\tau^{\alpha\beta}\vec{a}(j)]}$$

$$\times f\left[\frac{1}{N}\sum_{l=1}^{N}\vec{a}^{\dagger}(l)\mathbf{\Theta}\vec{a}(l)\right] - Nd\left[\frac{1}{N}\sum_{l=1}^{N}\vec{a}^{\dagger}(l)\mathbf{\Theta}\vec{a}(l)\right] \qquad (83)$$

To study the equilibrium properties of the system, as in the case of the parallel model, we calculate the partition function by performing a Trotter factorization

$$Z = \operatorname{Tr} e^{-\hat{H}t}\hat{P}$$

$$= \int_0^{2\pi}\left[\prod_{j=1}^{N}\frac{d\lambda_j}{2\pi}\right]e^{-i\lambda_j}\lim_{M\to\infty}\int\left[\prod_{k=1}^{M}\prod_{j=1}^{N}\frac{d^h\vec{z}_k^*(j)d^h\vec{z}_k(j)}{\pi^h}\right]e^{-S[\vec{z}^*,\vec{z}]} \qquad (84)$$

Here,

$$e^{-S[\vec{z}^*,\vec{z}]} = \prod_{k=1}^{M}\langle\{\vec{z}_k\}|e^{-\epsilon\hat{H}}|\{\vec{z}_{k-1}\}\rangle \qquad (85)$$

where the matrix elements in the coherent states basis are given by the expression

$$\langle\{\vec{z}_k\}|e^{-\epsilon\hat{H}}|\{\vec{z}_{k-1}\}\rangle$$

$$= e^{-\frac{1}{2}\sum_{j=1}^{N}(\vec{z}_k^*(j)\cdot\vec{z}_k(j) - 2\vec{z}_k^*(j)\cdot\vec{z}_{k-1}(j) + \vec{z}_{k-1}^*(j)\cdot\vec{z}_{k-1}(j))}$$

$$\times \exp\Bigg(\epsilon Ne^{-\sum_{\alpha\neq\beta=1}^{h}\mu_{\alpha\beta}}e^{\frac{1}{N}\sum_{j=1}^{N}\vec{z}_k^*(j)(\sum_{\alpha\neq\beta=1}^{h}\mu_{\alpha\beta}\tau^{\alpha\beta})\vec{z}_{k-1}(j)}$$

$$\times f\left[\frac{1}{N}\sum_{j=1}^{N}\vec{z}_k^*(j)\mathbf{\Theta}\vec{z}_{k-1}(j)\right] - \epsilon Nd\left[\frac{1}{N}\sum_{j=1}^{N}\vec{z}_k^*(j)\mathbf{\Theta}\vec{z}_{k-1}(j)\right]\Bigg) \qquad (86)$$

Let us introduce the $(h-1)$-component vector field $\vec{x}_k = (x_k^1, x_k^2, \ldots, x_k^{h-1})$, and an integral representation of the corresponding delta function

$$1 = \int \mathcal{D}[\vec{x}]\prod_{k=1}^{M}\delta^{(h-1)}\left[\vec{x}_k - \frac{1}{N}\sum_{j=1}^{N}\vec{z}_k^*(j)\mathbf{\Theta}\vec{z}_{k-1}(j)\right]$$

$$= \int \left[\prod_{k=1}^{M} \prod_{\alpha=1}^{h-1} \frac{i \epsilon N d \bar{x}_k^\alpha d x_k^\alpha}{2\pi} \right] e^{-\epsilon N \sum_{k=1}^{M} \vec{\bar{x}}_k \cdot \vec{x}_k + \epsilon \sum_{k=1}^{M} \sum_{j=1}^{N} \vec{z}_k^*(j) \vec{\bar{x}}_k \cdot \boldsymbol{\Theta} \vec{z}_{k-1}(j)} \tag{87}$$

Similarly, let us introduce a second set of fields $\eta^{\alpha\beta}$,

$$1 = \int \mathcal{D}[\eta^{\alpha\beta}] \prod_{k=1}^{M} \delta \left[\eta_k^{\alpha\beta} - \frac{1}{N} \sum_{j=1}^{N} \vec{z}_k^*(j) \tau^{\alpha\beta} \vec{z}_{k-1}(j) \right]$$

$$= \int \left[\prod_{k=1}^{M} \frac{i \epsilon N d \bar{\eta}_k^{\alpha\beta} d \eta_k^{\alpha\beta}}{2\pi} \right]$$

$$\times e^{-\epsilon N \sum_{k=1}^{M} \bar{\eta}_k^{\alpha\beta} \eta_k^{\alpha\beta} + \epsilon \sum_{k=1}^{M} \bar{\eta}_k^{\alpha\beta} \sum_{j=1}^{N} \vec{z}_k^*(j) \tau^{\alpha\beta} \vec{z}_{k-1}(j)} \tag{88}$$

Inserting both constraints (87) and (88) in the expression for the trace (84), we obtain

$$Z = \lim_{M \to \infty} \int \mathcal{D}[\vec{x}] \mathcal{D}[\vec{\bar{x}}] \prod_{\alpha \neq \beta} \mathcal{D}[\eta^{\alpha\beta}] \mathcal{D}[\bar{\eta}^{\alpha\beta}]$$

$$\times e^{\epsilon N \sum_{k=1}^{M} (-\vec{\bar{x}}_k \cdot \vec{x}_k - \sum_{\alpha \neq \beta=1}^{h} \bar{\eta}_k^{\alpha\beta} \eta_k^{\alpha\beta} + e^{\sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta}(\eta^{\alpha\beta}-1)} f[\vec{x}_k] - d[\vec{x}_k])}$$

$$\times \int \mathcal{D}[\vec{z}^*] \mathcal{D}[\vec{z}] \mathcal{D}[\lambda] \prod_{j=1}^{N} e^{-i\lambda_j} e^{\epsilon \sum_{k,l=1}^{M} \vec{z}_k^*(j) S_{kl} \vec{z}_l(j)} \Bigg|_{\{\vec{z}_0\} = \{e^{i\lambda_j} \vec{z}_M\}} \tag{89}$$

After performing the Gaussian integral over the fields $\vec{z}^*, \vec{z}$, we obtain

$$Z = \lim_{M \to \infty} \int \mathcal{D}[\vec{x}] \mathcal{D}[\vec{\bar{x}}] \prod_{\alpha \neq \beta} \mathcal{D}[\eta^{\alpha\beta}] \mathcal{D}[\bar{\eta}^{\alpha\beta}]$$

$$\times e^{\epsilon N \sum_{k=1}^{M} (-\vec{\bar{x}}_k \cdot \vec{x}_k - \sum_{\alpha \neq \beta=1}^{h} \bar{\eta}_k^{\alpha\beta} \eta_k^{\alpha\beta} + e^{-\sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta}(1-\eta_k^{\alpha\beta})} f[\vec{x}_k] - d[\vec{x}_k])}$$

$$\times \mathcal{D}[\lambda] \prod_{j=1}^{N} e^{-i\lambda_j} [\det S(j)]^{-1} \tag{90}$$

The matrix $S(j)$ has the structure

$$S(j) = \begin{pmatrix} I & 0 & 0 & \dots & -e^{-i\lambda_j} A_1(j) \\ -A_2(j) & I & 0 & \dots & 0 \\ 0 & -A_3(j) & I & \dots & 0 \\ & & & \ddots & \\ 0 & & \dots & -A_M(j) & I \end{pmatrix} \tag{91}$$

where $A_k(j) = I + \epsilon[\sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta} \bar{\eta}_k^{\alpha\beta} + \vec{\bar{x}}_k \cdot \boldsymbol{\Theta}]$. We obtain

$$\det S(j) = \det \left[I - e^{i\lambda_j} \prod_{k=1}^{M} A_k(j) \right]$$

$$= \det\left[I - e^{i\lambda_j}\hat{T}e^{\epsilon\sum_{k=1}^{M}[\sum_{\alpha\neq\beta=1}^{h}\bar{\eta}_k^{\alpha\beta}\tau^{\alpha\beta}+\vec{x}_k\cdot\Theta]} \right]$$

$$= e^{\mathrm{Tr}\ln[I - e^{i\lambda_j}\hat{T}e^{\epsilon\sum_{k=1}^{M}[\sum_{\alpha\neq\beta=1}^{h}\bar{\eta}_k^{\alpha\beta}\tau^{\alpha\beta}+\vec{x}_k\cdot\Theta]}]} \tag{92}$$

Substituting this last expression into the functional integral (90), and then performing the integrals over the λ fields, we obtain

$$Z = \lim_{M\to\infty}\int \mathcal{D}[\vec{\bar{x}}]\mathcal{D}[\vec{x}]\prod_{\alpha\neq\beta=1}^{h}\mathcal{D}[\bar{\eta}^{\alpha\beta}]\mathcal{D}[\eta^{\alpha\beta}]$$

$$\times e^{\epsilon N\sum_{k=1}^{M}(-\vec{\bar{x}}_k\cdot\vec{x}_k - \sum_{\alpha\neq\beta=1}^{h}\bar{\eta}_k^{\alpha\beta}\eta_k^{\alpha\beta} + e^{\sum_{\alpha\neq\beta=1}^{h}\mu_{\alpha\beta}(\eta^{\alpha\beta}-1)}f[\vec{x}_k]-d[\vec{x}_k])}\prod_{j=1}^{N}Q \tag{93}$$

Here,

$$Q = \lim_{M\to\infty}\mathrm{Tr}\,\hat{T}\prod_{k=1}^{M}\left[I + \epsilon\left(\sum_{\alpha\neq\beta=1}^{h}\bar{\eta}_k^{\alpha\beta}\tau^{\alpha\beta} + \vec{x}_k\cdot\Theta\right)\right]$$

$$= \lim_{M\to\infty}\mathrm{Tr}\,\hat{T}e^{\epsilon\sum_{k=1}^{M}[\sum_{\alpha\neq\beta=1}^{h}\bar{\eta}_k^{\alpha\beta}\tau^{\alpha\beta}+\vec{x}_k\cdot\Theta]}$$

$$= \mathrm{Tr}\,\hat{T}e^{\int_0^t dt'[\sum_{\alpha\neq\beta=1}^{h}\bar{\eta}^{\alpha\beta}(t')\tau^{\alpha\beta}+\vec{x}(t')\cdot\Theta]} \tag{94}$$

After taking the limit $M\to\infty$, we obtain

$$Z = \int \mathcal{D}[\vec{\bar{x}}]\mathcal{D}[\vec{x}]\prod_{\alpha\neq\beta}\mathcal{D}[\bar{\eta}^{\alpha\beta}]\mathcal{D}[\eta^{\alpha\beta}]e^{-S[\vec{\bar{x}},\vec{x},\{\bar{\eta}^{\alpha\beta}\},\{\eta^{\alpha\beta}\}]} \tag{95}$$

Here,

$$S[\vec{\bar{x}}, \vec{x}, \{\bar{\eta}^{\alpha\beta}\}, \{\eta^{\alpha\beta}\}]$$

$$= -N\int_0^t dt'\{-\vec{\bar{x}}(t')\cdot\vec{x}(t') - \sum_{\alpha\neq\beta=1}^{h}\bar{\eta}^{\alpha\beta}(t')\eta^{\alpha\beta}(t')$$

$$+ e^{\sum_{\alpha\neq\beta=1}^{h}\mu_{\alpha\beta}(\eta^{\alpha\beta}(t')-1)}f[\vec{x}(t')] - d[\vec{x}(t')]\} - N\ln Q \tag{96}$$

where

$$Q = \mathrm{Tr}\,\hat{T}e^{\int_0^t dt'[\sum_{\alpha\neq\beta=1}^{h}\bar{\eta}^{\alpha\beta}(t')\tau^{\alpha\beta}+\sum_{\alpha=1}^{h-1}\bar{x}^{\alpha}(t')\Theta^{\alpha}]} \tag{97}$$

With this last simplification, the effective action becomes

$$S[\vec{\bar{x}}, \vec{x}, \{\bar{\eta}^{\alpha\beta}\}, \{\eta^{\alpha\beta}\}]$$

$$= -N\int_0^t dt'(-\vec{\bar{x}}(t')\cdot\vec{x}(t') - \sum_{\alpha\neq\beta=1}^{h}\bar{\eta}^{\alpha\beta}(t')\eta^{\alpha\beta}(t')$$

$$+ e^{\sum_{\alpha\neq\beta=1}^{h}\mu_{\alpha\beta}(\eta^{\alpha\beta}-1)}f[\vec{x}(t')] - d[\vec{x}(t')]) - N\ln Q \tag{98}$$

3.2 The Large N Limit of the h-state Eigen Model is a Saddle Point

By assuming that the sequence length N is very large, $N \to \infty$, we can evaluate the functional integral (95) by a saddle point method. Considering the action defined in (98), we have

$$
\left.\frac{\delta S}{\delta x^\alpha}\right|_{\bar{\vec{x}}_c,\vec{x}_c,\{\bar{\eta}_c^{\alpha\beta}\},\{\eta_c^{\alpha\beta}\}} = N\left(\bar{x}_c^\alpha - e^{\sum_{\gamma\neq\rho=1}^h \mu_{\gamma\rho}(\eta_c^{\gamma\rho}-1)}\left.\frac{\partial f[\vec{x}_c]}{\partial x^\alpha}\right|_c + \left.\frac{\partial d[\vec{x}_c]}{\partial x^\alpha}\right|_c\right) = 0
$$

$$
\left.\frac{\delta S}{\delta \eta^{\alpha\beta}}\right|_{\bar{\vec{x}}_c,\vec{x}_c,\{\bar{\eta}_c^{\alpha\beta}\},\{\eta_c^{\alpha\beta}\}} = N\left(\bar{\eta}_c^{\alpha\beta} - \mu_{\alpha\beta}e^{\sum_{\gamma\neq\rho=1}^h \mu_{\gamma\rho}(\eta_c^{\gamma\rho}-1)}f[\vec{x}_c]\right) = 0
$$

$$
\left.\frac{\delta S}{\delta \bar{x}^\alpha}\right|_{\bar{\vec{x}}_c,\vec{x}_c,\{\bar{\eta}_c^{\alpha\beta}\},\{\eta_c^{\alpha\beta}\}} = N\left(x_c^\alpha - \frac{1}{Q}\left.\frac{\delta Q}{\delta \bar{x}^\alpha}\right|_c\right) = 0
$$

$$
\left.\frac{\delta S}{\delta \bar{\eta}^{\alpha\beta}}\right|_{\bar{\vec{x}}_c,\vec{x}_c,\{\bar{\eta}_c^{\alpha\beta}\},\{\eta_c^{\alpha\beta}\}} = N\left(\eta_c^{\alpha\beta} - \frac{1}{Q}\left.\frac{\delta Q}{\delta \bar{\eta}^{\alpha\beta}}\right|_c\right) = 0
$$

(99)

We have therefore the system of equations

$$
\bar{x}_c^\alpha = e^{\sum_{\gamma\neq\rho=1}^h \mu_{\gamma\rho}(\eta_c^{\gamma\rho}-1)}\left.\frac{\partial f[\vec{x}_c]}{\partial x^\alpha}\right|_c - \left.\frac{\partial d[\vec{x}_c]}{\partial x^\alpha}\right|_c
\tag{100}
$$

$$
\bar{\eta}_c^{\alpha\beta} = \mu_{\alpha\beta}e^{\sum_{\gamma\neq\rho=1}^h \mu_{\gamma\rho}(\eta_c^{\gamma\rho}-1)}f[\vec{x}_c]
\tag{101}
$$

$$
x_c^\alpha = \langle\Theta^\alpha\rangle
\tag{102}
$$

$$
\eta_c^{\alpha\beta} = \langle\tau^{\alpha\beta}\rangle
\tag{103}
$$

where we defined

$$
\langle(\cdot)\rangle = \frac{\operatorname{Tr}(\cdot)e^{t[\sum_{\alpha\neq\beta=1}^h \bar{\eta}_c^{\alpha\beta}\tau^{\alpha\beta}+\sum_{\alpha=1}^{h-1}\bar{x}_c^\alpha\Theta^\alpha]}}{\operatorname{Tr}e^{t[\sum_{\alpha\neq\beta=1}^h \bar{\eta}_c^{\alpha\beta}\tau^{\alpha\beta}+\sum_{\alpha=1}^{h-1}\bar{x}_c^\alpha\Theta^\alpha]}}
\tag{104}
$$

After the saddle-point analysis, we obtain an exact analytical expression for the mean fitness f_m of the population, in the limit of very large sequences $N \to \infty$, for an arbitrary microscopic fitness function $f(\vec{x})$ and degradation rate $d(\vec{x})$

$$
\begin{aligned}
f_m &= \lim_{N,t\to\infty}\frac{-S_c}{Nt}\\
&= \max_{\{\bar{x}_c\vec{x}_c,\{\eta_c^{\alpha\beta}\},\{\bar{\eta}_c^{\alpha\beta}\}\}}\left[e^{\sum_{\alpha\neq\beta=1}^h \mu_{\alpha\beta}(\eta_c^{\alpha\beta}-1)}f(\vec{x}_c) - d(\vec{x}_c)\right.\\
&\qquad\left. - \bar{\vec{x}}_c\cdot\vec{x}_c - \sum_{\alpha\neq\beta=1}^h \eta_c^{\alpha\beta}\bar{\eta}_c^{\alpha\beta} + \lambda_{\max}\right]
\end{aligned}
\tag{105}
$$

As shown in Appendix 3, the $\eta^{\alpha\beta}$ can be eliminated in terms of the compositions, to obtain the final expression

$$
f_m = \max_{\{x_c^1,x_c^2,\ldots,x_c^{h-1}\}}\left\{f(x_c^1,x_c^2,\ldots,x_c^{h-1})e^{\sum_{\alpha\neq\beta=1}^{h-1}\mu_{\alpha\beta}[\sqrt{x_c^\alpha x_c^\beta}-x_c^\alpha]}\right.
$$

$$\times \, e^{\sum_{\alpha=1}^{h-1} \mu_{\alpha h}[\sqrt{x_c^\alpha(1-\sum_{\gamma=1}^{h-1} x_c^\gamma)}-x_c^\alpha]}$$

$$\times \, e^{\sum_{\beta=1}^{h-1} \mu_{h\beta}[\sqrt{(1-\sum_{\gamma=1}^{h-1} x_c^\gamma)x_c^\beta}+\sum_{\gamma=1}^{h-1} x_c^\gamma - 1]} - d(x_c^1, x_c^2, \ldots, x_c^{h-1}) \Bigg\}$$ (106)

Here, the compositions $0 \le x_c^\alpha \le 1$ are defined within the simplex $\sum_{\alpha=1}^{h-1} x_c^\alpha \le 1$. We note that the mutation terms in (106) are the exponential of the mutation terms in (44), which is a result of the mutation terms in the Eigen Hamiltonian, (83), being the exponential of those in the parallel Hamiltonian, (17).

3.3 The Purine/Pyrimidine Alphabet $h = 2$

As an application of our general solution (106), we first consider the purine/pyrimidine alphabet with $h = 2$. The maximum principle for this binary alphabet can be derived from (106) by assuming a symmetric mutation rate $\mu_{12} = \mu_{21} \equiv \mu$, and by noticing that a single composition (or normalized Hamming distance) $x_c^1 \equiv x_c$ is required in this case,

$$f_m^{(2)} = \max_{0 \le x_c \le 1} \left\{ e^{\mu[2\sqrt{x_c(1-x_c)}-1]} - d(x_c) \right\}$$ (107)

It is customary to use in this case a magnetization coordinate, defined as $\xi_c = 1 - 2x_c$, and hence (107) becomes

$$f_m^{(2)} = \max_{\{-1 \le \xi_c \le 1\}} \left\{ e^{\mu[\sqrt{1-\xi_c^2}-1]} f(\xi_c) - d(\xi_c) \right\}$$ (108)

Equation (108) is a well known result [14, 19], and a special case of our general result (106).

3.4 The Nucleic Acid Alphabet $h = 4$

For a general, non-symmetric mutation scheme as in Fig. 1, by considering $h = 4$ in our general result (106) we obtain

$$f_m = \max_{\{x_c^1, x_c^2, x_c^3\}} \left\{ f(x_c^1, x_c^2, x_c^3) e^{\sum_{\alpha \ne \beta=1}^{3} \mu_{\alpha\beta}[\sqrt{x_c^\alpha x_c^\beta}-x_c^\alpha]} e^{\sum_{\alpha=1}^{3} \mu_{\alpha 4}[\sqrt{x_c^\alpha(1-\sum_{\gamma=1}^{3} x_c^\gamma)}-x_c^\alpha]} \right.$$

$$\left. \times \, e^{\sum_{\beta=1}^{3} \mu_{4\beta}[\sqrt{(1-\sum_{\gamma=1}^{3} x_c^\gamma)x_c^\beta}+\sum_{\gamma=1}^{3} x_c^\gamma - 1]} - d(x_c^1, x_c^2, x_c^3) \right\}$$ (109)

Here, the compositions x_c^α are defined within the simplex $x_c^1 + x_c^2 + x_c^3 \le 1$.

An interesting reduction of this general model is provided by the Kimura 3 ST mutation scheme, introduced in Sect. 2.6, and represented in Fig. 2. We obtain the solution for the Kimura 3 ST mutation scheme by assuming the following symmetries in the mutation coefficients

$$\mu_{12} = \mu_{21} = \mu_{34} = \mu_{43} \equiv \mu_1$$

$$\mu_{13} = \mu_{31} = \mu_{24} = \mu_{42} \equiv \mu_2$$ (110)

$$\mu_{14} = \mu_{41} = \mu_{23} = \mu_{32} \equiv \mu_3$$

along with the magnetization coordinates ξ_c^α defined in agreement with (110),

$$\xi_c^1 = 1 - 2(x_c^1 + x_c^3)$$
$$\xi_c^2 = 1 - 2(x_c^2 + x_c^3)$$
$$\xi_c^3 = 1 - 2(x_c^1 + x_c^2)$$

(111)

With these assumptions, after some algebra, (106) reduces to the expression for the Kimura 3ST scheme,

$$f_m = \max_{\{\xi_c^1, \xi_c^2, \xi_c^3\}} \left\{ e^{\mu_1(\frac{1}{2}\sqrt{(1+\xi_c^1+\xi_c^2+\xi_c^3)(1-\xi_c^1-\xi_c^2+\xi_c^3)}+\frac{1}{2}\sqrt{(1+\xi_c^1-\xi_c^2-\xi_c^3)(1-\xi_c^1+\xi_c^2-\xi_c^3)}-1)} \right.$$

$$\times\, e^{\mu_2(\frac{1}{2}\sqrt{(1+\xi_c^1+\xi_c^2+\xi_c^3)(1+\xi_c^1-\xi_c^2-\xi_c^3)}+\frac{1}{2}\sqrt{(1-\xi_c^1+\xi_c^2-\xi_c^3)(1-\xi_c^1-\xi_c^2+\xi_c^3)}-1)}$$

$$\times\, e^{\mu_3(\frac{1}{2}\sqrt{(1+\xi_c^1+\xi_c^2+\xi_c^3)(1-\xi_c^1+\xi_c^2-\xi_c^3)}+\frac{1}{2}\sqrt{(1+\xi_c^1-\xi_c^2-\xi_c^3)(1-\xi_c^1-\xi_c^2+\xi_c^3)}-1)}$$

$$\left. \times\, f(\xi_c^1, \xi_c^2, \xi_c^3) - d(\xi_c^1, \xi_c^2, \xi_c^3) \right\}$$

(112)

3.5 Analytical Results for the Symmetric Mutation Scheme

If the mutation rates are identical $\mu_1 = \mu_2 = \mu_3 = \mu$, then we have the symmetric case $\xi_c^1 = \xi_c^2 = \xi_c^3 \equiv \xi_c$, and after (112) we obtain

$$f_m = \max_{-\frac{1}{3} \leq \xi_c \leq 1} \left\{ e^{\frac{3}{2}\mu[-1-\xi_c+\sqrt{(1+3\xi_c)(1-\xi_c)}]} f[\xi_c] - d[\xi_c] \right\}$$

(113)

3.5.1 The Sharp Peak Fitness Landscape

Let us first consider the sharp peak landscape $f(u) = (A - A_0)\delta_{u,1} + A_0$, with $A > A_0$. That is, the replication rate is $f(u = 1) = A$ for sequences identical to the wild type, and $f(u \neq 1) = A_0$, for all other sequences. With zero degradation rate, $d = 0$, we notice that this result implies: $\xi_c = 1$ if $A > A_0 e^{3\mu}$, or $\xi_c = 0$ otherwise. Therefore, we obtain the mean replication rate

$$f_m = \begin{cases} e^{-3\mu} A, & A > A_0 e^{3\mu} \\ A_0, & A < A_0 e^{3\mu} \end{cases}$$

(114)

The system experiences a phase transition which is first order in A. The steady-state probability for the wild-type is obtained from the self-consistent condition: $f_m = A p_w + A_0(1 - p_w)$,

$$p_w = \begin{cases} \frac{e^{-3\mu} A - A_0}{A - A_0}, & A > A_0 e^{3\mu} \\ 0, & A < A_0 e^{3\mu} \end{cases}$$

(115)

Notice that the error threshold is reached at the critical value $A_{crit} = A_0 e^{3\mu}$, as follows from (114), (115) and as displayed in Fig. 6. We notice that this result differs from the analytical value obtained for the binary alphabet [19], $A_{crit}^{(2)} = A_0 e^\mu$. The additional factor of three which is explicit in the exponent is clearly a consequence of the existence of three mutation channels into which evolving sequences can escape from the wild-type. This effect, which is purely entropic- and not fitness-like, is an explicit consequence of the larger alphabet size.

Fig. 6 The average composition u and magnetization ξ_c are represented as a function of the parameter A/A_0, for the sharp peak landscape. The mutation rate was set to $\mu = 1.0$. Also shown (*inset*) is the fraction of the population located at the peak, p_w
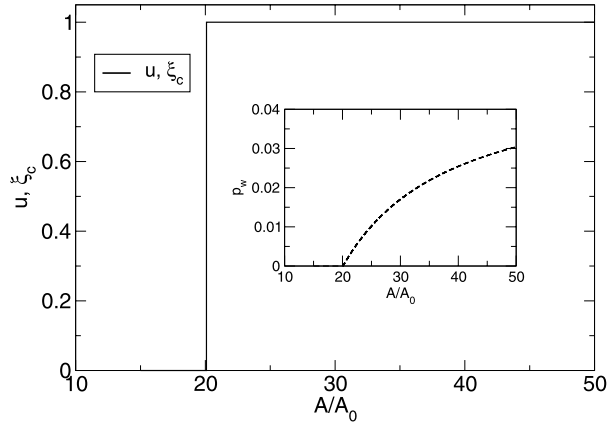
3.5.2 The Fujiyama Fitness Landscape

We will consider the Fujiyama fitness landscape, which is a linear function of the composition

$$f(\vec{u}) = \sum_{i=1}^{3}(\alpha_i u_i) + \alpha_0 \tag{116}$$

For the symmetric case, $\alpha_i = \alpha$, $\mu_i = \mu$. Therefore, we have $\xi_c^1 = \xi_c^2 = \xi_c^3 \equiv \xi_c$. The mean fitness, in the absence of degradation, from (113) becomes

$$f_m = \max_{-\frac{1}{3} \leq \xi_c \leq 1} \left\{ (3\alpha\xi_c + \alpha_0)e^{\frac{3}{2}\mu[-1-\xi_c+\sqrt{(1+3\xi_c)(1-\xi_c)}]} \right\} \tag{117}$$

By maximizing with respect to ξ_c, $\frac{\partial f_m}{\partial \xi_c} = 0$, we obtain the nonlinear equation

$$\alpha - \frac{\mu}{2}\alpha_0 - \frac{3}{2}\mu\alpha\xi_c = \frac{\mu}{2} \frac{(3\alpha\xi_c + \alpha_0)(3\xi_c - 1)}{\sqrt{(1 + 3\xi_c)(1 - \xi_c)}} \tag{118}$$

No error threshold is observed for this fitness landscape, except for the trivial limit $\alpha \to 0$, $\alpha_0 \geq 0$. The average surplus u is obtained by the self-consistent equation

$$f_m = 3\alpha u + \alpha_0 \tag{119}$$

3.6 The Quadratic Fitness Landscape

Next we consider the quadratic fitness landscape

$$f(\vec{u}) = \sum_{i=1}^{N}\left(\frac{\beta_i}{2}u_i^2 + \alpha_i u_i\right) + 1 \tag{120}$$

For the symmetric case, $\beta_i = \beta$, $\alpha_i = \alpha$, $\mu_i = \mu$, we have $\xi_c^i = \xi_c$ and $\bar{\xi}_c^i = \bar{\xi}_c$. Thus, the mean fitness, for a null degradation rate, after (113) is

$$f_m = \max_{-\frac{1}{3} \leq \xi_c \leq 1} \left\{ \left(\frac{3}{2}\beta\xi_c^2 + 3\alpha\xi_c + 1\right) e^{\frac{3}{2}\mu[-1-\xi_c+\sqrt{(1+3\xi_c)(1-\xi_c)}]} \right\} \tag{121}$$

Fig. 7 The average composition u and magnetization ξ_c are represented as a function of the parameter β/μ for the quadratic fitness, when $\alpha = 0$
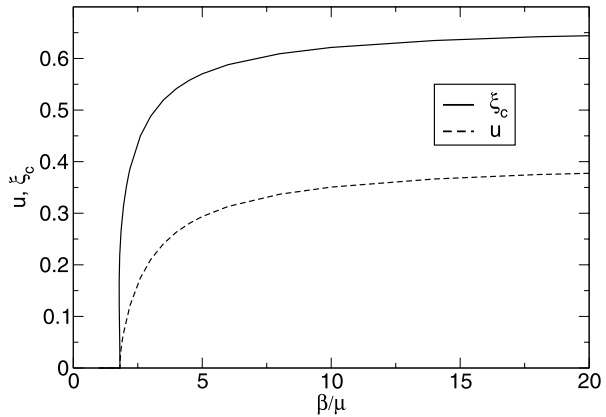
We maximize with respect to ξ_c, $\frac{\partial f_m}{\partial \xi_c} = 0$, to obtain

$$\beta \xi_c + \alpha = \frac{\mu}{2} \left(\frac{3}{2} \beta \xi_c^2 + 3\alpha \xi_c + 1 \right) \left[1 + \frac{3\xi_c - 1}{\sqrt{(1 + 3\xi_c)(1 - \xi_c)}} \right] \qquad (122)$$

The average base composition u is obtained from the self-consistent condition

$$f_m = f(u) = \frac{3}{2} \beta u^2 + 3\alpha u + 1 \qquad (123)$$

The selected phase, $\xi_c > 0$, $u > 0$, occurs for $\beta > 1.8096\mu$ when $\alpha = 0$. The value of u is continuous at the transition, as it can checked from (121) that $f_m(\xi_c = 0) = f_m(\xi_c = 0.2289, \beta = 1.8096, \mu = 1) = 1$, which implies after (123) (for $\alpha = 0$) that $u = 0$ when approaching the critical point $\beta = 1.8066\mu$ from both sides. However, ξ_c jumps from 0 (for $\beta \to 1.8066\mu^-$) to 0.2289 (for $\beta \to 1.8066\mu^+$). By expanding (121) near the critical point, after a similar procedure as in (69) for the parallel model, we find a discontinuous jump in $df_m/d\beta$ from 0 to 0.06883. Therefore, the phase transition is of first order in β. A graphical representation is displayed in Fig. 7.

3.7 The Quartic Fitness Landscape

As a final example, we consider the quartic fitness landscape,

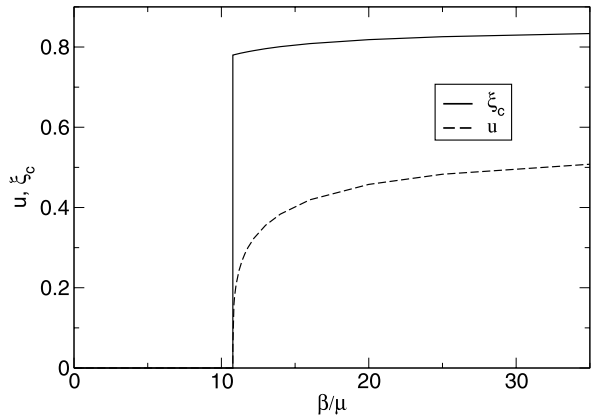$$f(\vec{u}) = \sum_{i=1}^{3} \frac{\beta_i}{4} u_i^4 + 1 \qquad (124)$$

We further consider the symmetric case $\mu_i \equiv \mu$, $\beta_i \equiv \beta$, and hence $\xi_c^i \equiv \xi_c$. From the general expression (113), we obtain an analytical expression for the mean fitness

$$f_m = \max_{\{-\frac{1}{3} \leq \xi_c \leq 1\}} \left\{ \left(\frac{3}{4} \beta \xi_c^4 + 1 \right) e^{\frac{3}{2} \mu [-1 - \xi_c + \sqrt{(1 + 3\xi_c)(1 - \xi_c)}]} \right\} \qquad (125)$$

The average composition of the population u is obtained from the self-consistent condition

$$f(u) = \frac{3}{4} \beta u^4 + 1 = f_m \qquad (126)$$

In Fig. 8, we present the values of u and ξ_c, as obtained from (125), (126), as a function of the parameter β/μ. We notice that a discontinuous jump in the bulk magnetization from $\xi_c = 0$ to $\xi_c = 0.779856$ is observed at $\beta/\mu = 10.776165$. By expanding (125) near the critical point, we find a discontinuous jump in $df_m/d\beta$, from 0 to 0.066213. Therefore, the phase transition is of first order in β. The average composition shows a fast but continuous transition. This behavior is much like the one observed in the sharp peak fitness landscape, (114), and in the corresponding example for the parallel model.

3.8 Symmetric Case, General h, with Application to the Amino Acid Alphabet $h = 20$

We consider the case of the amino acid alphabet, which is derived from our general solution (106) by setting $h = 20$. In particular, when a symmetric mutation scheme is assumed $\mu_{\alpha\beta} = \mu$, for all α, β, and $x_c^\alpha = x_c$. We first consider the symmetric case for general h.

For an alphabet of size h, we define a magnetization coordinate $\xi_c = 1 - hx_c$, and obtain that (106) reduces to

$$f_m = \max_{\{-1/(h-1) \le \xi_c \le 1\}} \left\{ f(\xi_c)e^{(h-1)\mu[\frac{2}{h}\sqrt{(1-\xi_c)(1+(h-1)\xi_c)}+\frac{h-2}{h}(1-\xi_c)-1]} - d(\xi_c) \right\} \tag{127}$$

As an example of application of (127), we consider the sharp peak fitness landscape $f(\xi_c) = (A - A_0)\delta_{\xi_c,1} + A_0$, and zero degradation function $d(\xi_c) = 0$. Then, from (127) we obtain the mean fitness

$$f_m = \begin{cases} e^{-(h-1)\mu}A, & A > A_0 e^{(h-1)\mu} \\ A_0, & A \le A_0 e^{(h-1)\mu} \end{cases} \tag{128}$$

The system experiences a first order phase transition at $A_{crit} = A_0 e^{(h-1)\mu}$. The steady-state probability for the wild-type is obtained from the self-consistency condition: $f_m = Ap_w + A_0(1 - p_w)$,

$$p_w = \begin{cases} \frac{Ae^{-(h-1)\mu}-A_0}{A-A_0}, & A > A_0 e^{(h-1)\mu} \\ 0, & A \le A_0 e^{(h-1)\mu} \end{cases} \tag{129}$$

The factor $(h - 1)\mu$ in the exponential is intuitive, since there exists $h - 1$ independent mutation channels into which evolving sequences can escape from the wild-type.

As a second example, we consider the quadratic fitness landscape $f(\xi_c) = (h - 1)\beta\xi_c^2/2 + 1$ for an alphabet of size h. We consider the case of the amino acid alphabet, with $h = 20$. We set $\mu = 1$. By a similar analysis as in the parallel model case, we find a phase transition at the critical point $\beta = 7.483\mu$. The magnetization parameter ξ_c has a finite jump from $\xi_c = 0$ (for $\beta \to 7.483\mu^-$) to $\xi_c = 0.0827$ (for $\beta \to 7.483\mu^+$). The mean fitness is continuous at the transition, since $f_m(\beta \to 7.483\mu^+) = f_m(\beta \to 7.483\mu^-) = 1$, which implies that the observable $u = 0$ at the critical point. We observe that the first derivative has a finite jump at the critical point, from $df_m/d\beta(\beta \to 7.483\mu^-) = 0$, to $df_m/d\beta(\beta \to 7.483\mu^+) = 0.0437$, and therefore the phase transition is of first order. A similar analysis shows that the transition is first order for $h = 3$ as well. As with the parallel model, we find that the transition for the quadratic fitness function is second order for $h = 2$ and first order for $h > 2$.

4 Conclusion

Using the quantum spin chain approach, the 2-state, purine/pyrimidine assumption for quasi-species theory has been lifted to arbitrary alphabet sizes h. We have here expressed the general result for the fitness of the evolved population as a maximization principle. We have derived the solution for a general fitness function using the Schwinger spin coherent states approach. We have presented analytic results for the sharp peak, as well as linear, quadratic, and quartic fitness functions. For the Kimura 3 ST mutation scheme, we have presented an explicit solution for a general fitness function, expressed as a maximization principle.

We have also derived the general solution to the Eigen model of mutation and selection for arbitrary alphabet size and for a general mutation scheme. We have presented analytic results for the sharp peak, linear, quadratic, and quartic fitness functions.

These results bring quasi-species theory closer to the evolutionary dynamics that occurs at the genetic level.

Appendix 1

In what follows, we will need to apply Euler's theorem and three properties of the maximum eigenvalue $\lambda_{\max}$ of the matrix $\mathbf{M}$ defined by (43).

A.1 Euler's Theorem for Homogeneous Functions

A.1.1 Definition 1: Homogeneous Function of Degree k

A function $f(x_1, \ldots, x_n)$ of n variables is homogeneous of degree k if, for all $\alpha > 0$,

$$f(\alpha x_1, \alpha x_2, \ldots, \alpha x_n) = \alpha^k f(x_1, x_2, \ldots, x_n) \qquad (130)$$

A.1.2 Euler's Theorem

A differentiable function $f(x_1, \ldots, x_n)$ of n variables is homogeneous of degree k if and only if

$$\sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i} f(x_1, \ldots, x_n) = k f(x_1, \ldots, x_n) \tag{131}$$

The theorem and its proof is presented in most textbooks of mathematical analysis [24].

A.2 Property I

The maximum eigenvalue $\lambda_{\max}$ of the matrix $\mathbf{M}$ defined by (43) is a homogeneous function of degree $k = 1$ in the vector $(\bar{x}_c^1, \bar{x}_c^2, \ldots, \bar{x}_c^{h-1}, \{\mu_{\alpha\beta}\})$.

The proof of this proposition follows directly from Definition 1. Notice that after (43),

$$\mathbf{M}(\{\bar{x}_c^{\alpha}\}, \{\mu_{\alpha\beta}\}) = \sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} \tau^{\alpha\beta} + \sum_{\alpha=1}^{h-1} \bar{x}_c^{\alpha} \Theta^{\alpha} \tag{132}$$

Since $\mathbf{M}$ is a linear function of the vector $(\bar{x}_c^1, \ldots, \bar{x}_c^{h-1}, \mu_{\alpha\beta})$

$$\mathbf{M}(\alpha \bar{x}_c^1, \alpha \bar{x}_c^2, \ldots, \alpha \bar{x}_c^{h-1}, \alpha \mu_{12}, \alpha \mu_{21}, \ldots) = \alpha \mathbf{M}(\bar{x}_c^1, \bar{x}_c^2, \ldots, \bar{x}_c^{h-1}, \mu_{12}, \mu_{21}, \ldots) \tag{133}$$

Therefore, the maximum eigenvalue, as obtained from the long-time limit of the trace,

$$\lambda_{\max}(\alpha \bar{x}_c^1, \alpha \bar{x}_c^2, \ldots, \alpha \bar{x}_c^{h-1}, \alpha \mu_{12}, \alpha \mu_{21}, \ldots)$$

$$= \lim_{t \to \infty} \frac{1}{t} \ln \left[\mathrm{Tr}\, e^{t \mathbf{M}(\alpha \bar{x}_c^1, \alpha \bar{x}_c^2, \ldots, \alpha \bar{x}_c^{h-1}, \alpha \mu_{12}, \alpha \mu_{21}, \ldots)} \right]$$

$$= \lim_{t \to \infty} \frac{1}{t} \ln \left[\mathrm{Tr}\, e^{t \alpha \mathbf{M}(\bar{x}_c^1, \bar{x}_c^2, \ldots, \bar{x}_c^{h-1}, \mu_{12}, \mu_{21}, \ldots)} \right]$$

$$= \alpha \lambda_{\max}(\bar{x}_c^1, \bar{x}_c^2, \ldots, \bar{x}_c^{h-1}, \mu_{12}, \mu_{21}, \ldots) \tag{134}$$

is also a homogeneous function of degree $k = 1$, after Definition 1.

A.3 Property II

$\lambda_{\max}$ satisfies the identity

$$\lambda_{\max} = \sum_{\alpha=1}^{h-1} \bar{x}_c^{\alpha} \frac{\partial \lambda_{\max}}{\partial \bar{x}_c^{\alpha}} + \sum_{\alpha \neq \beta = 1}^{h} \mu_{\alpha\beta} \frac{\partial \lambda_{\max}}{\partial \mu_{\alpha\beta}} \tag{135}$$

The proof follows directly by application of Euler's theorem, for a homogeneous function of degree $k = 1$.

Notice that, after the saddle-point (37)–(40), we have the following identities

$$x_c^{\alpha} = \langle \Theta^{\alpha} \rangle = \lim_{t \to \infty} \frac{\mathrm{Tr}[\Theta^{\alpha} e^{t\mathbf{M}}]}{\mathrm{Tr}\, e^{t\mathbf{M}}}$$

$$= \lim_{t \to \infty} \frac{\partial}{\partial \bar{x}_c^{\alpha}} \frac{\ln[\text{Tr}\, e^{t\mathbf{M}}]}{t} = \frac{\partial}{\partial \bar{x}_c^{\alpha}} \lambda_{\max} \tag{136}$$

$$\langle \tau^{\alpha\beta} \rangle = \frac{\partial}{\partial \mu_{\alpha\beta}} \lambda_{\max} \tag{137}$$

Substituting (136)–(137) into (135), we obtain the identity

$$\lambda_{\max} - \sum_{\alpha=1}^{h-1} \bar{x}_c^{\alpha} x_c^{\alpha} = \sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta} \langle \tau^{\alpha\beta} \rangle \tag{138}$$

After (138), (41) becomes

$$f_m = \max_{\{\bar{x}_c^{\alpha}\}, \{\mu_{\alpha\beta}\}} \left[f(\{x_c^{\alpha}\}) - \sum_{\alpha \neq \beta=1}^{h} \mu_{\alpha\beta} (1 - \langle \tau^{\alpha\beta} \rangle) \right] \tag{139}$$

A.4 Property III

The 'average' $\langle \mathbf{A} \rangle$ of an arbitrary $h \times h$ matrix $\mathbf{A}$, satisfies the identity

$$\langle \mathbf{A} \rangle = \lim_{t \to \infty} \frac{\text{Tr}[\mathbf{A}\, e^{t\mathbf{M}}]}{\text{Tr}\, e^{t\mathbf{M}}} = \vec{y}_{\max}^T \mathbf{A} \vec{y}_{\max} \tag{140}$$

with $\vec{y}_{\max}$ the eigenvector corresponding to the maximum eigenvalue $\lambda_{\max}$ of the matrix $\mathbf{M}$ in (132).

The proof follows by considering the unitary $h \times h$ matrix $\mathbf{P} = [\vec{y}_1, \vec{y}_2, \vec{y}_3, \ldots, \vec{y}_h]$, $\mathbf{P}^{-1} = \mathbf{P}^T$ whose columns are formed by the h orthogonal eigenvectors $\vec{y}_{\alpha}$ of $\mathbf{M}$ which satisfy

$$\mathbf{M}\vec{y}_{\alpha} = \lambda_{\alpha} \vec{y}_{\alpha} \quad 1 \leq \alpha \leq h$$
$$\vec{y}_{\alpha} \cdot \vec{y}_{\beta} = \delta_{\alpha\beta} \tag{141}$$

From elementary linear algebra, the matrix $\mathbf{P}$ induces a similarity transformation which diagonalizes $\mathbf{M}$, that is $\mathbf{PMP}^{-1} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_h) \equiv \mathbf{D}$. Hence, it also diagonalizes the exponential of $\mathbf{M}$,

$$\mathbf{P}e^{t\mathbf{M}}\mathbf{P}^{-1} = \mathbf{P}\left(\mathbf{I} + t\mathbf{M} + \frac{t^2}{2!}\mathbf{M}\mathbf{M} + \cdots \right)\mathbf{P}^{-1}$$

$$= \mathbf{I} + t\mathbf{PMP}^{-1} + \frac{t^2}{2!}\mathbf{PMP}^{-1}\mathbf{PMP}^{-1} + \cdots$$

$$= \mathbf{I} + t\mathbf{D} + \frac{t^2}{2!}\mathbf{D}^2 + \cdots = e^{t\mathbf{D}} \tag{142}$$

The 'average' of an arbitrary $h \times h$ matrix $\mathbf{A}$, defined by (140), is calculated as

$$\langle \mathbf{A} \rangle = \lim_{t \to \infty} \frac{\text{Tr}[\mathbf{A}\, e^{t\mathbf{M}}]}{\text{Tr}\, e^{t\mathbf{M}}} = \lim_{t \to \infty} \frac{\text{Tr}[\mathbf{PAP}^{-1}\mathbf{P}e^{t\mathbf{M}}\mathbf{P}^{-1}]}{\text{Tr}[\mathbf{P}e^{t\mathbf{M}}\mathbf{P}^{-1}]}$$

$$= \lim_{t \to \infty} \frac{\text{Tr}[\mathbf{PAP}^{-1}e^{t\mathbf{D}}]}{\text{Tr}\, e^{t\mathbf{D}}} = \lim_{t \to \infty} \frac{\sum_{i=1}^{h} \vec{y}_i^T \mathbf{A} \vec{y}_i e^{t\lambda_i}}{\sum_{j=1}^{h} e^{t\lambda_j}} = \vec{y}_{\max}^T \mathbf{A} \vec{y}_{\max} \tag{143}$$

which proves the Property III.

We can express the eigenvector $\vec{y}_{\max} = (y^1, y^2, \ldots, y^h)^T$ in terms of the fields x_c^α, by combining the result in Property III, with the saddle-point equations (136), (137), as follows

$$x_c^\alpha = \langle \Theta^\alpha \rangle = \vec{y}_{\max}^T \Theta^\alpha \vec{y}_{\max} = (y^\alpha)^2 \tag{144}$$

These equations are inverted to obtain

$$y^\alpha = \begin{cases} \sqrt{x_c^\alpha}, & 1 \le \alpha \le h-1 \\ \sqrt{1 - \sum_{\gamma=1}^{h-1} x_c^\gamma}, & \alpha = h \end{cases} \tag{145}$$

Equipped with this result, we can now calculate the 'averages' $\langle \tau^{\alpha\beta} \rangle$ in (139),

$$\langle \tau^{\alpha\beta} \rangle = \vec{y}_{\max}^T \tau^{\alpha\beta} \vec{y}_{\max} = y^\alpha y^\beta + \sum_{\gamma \ne \alpha} (y^\gamma)^2 \tag{146}$$

Substituting (145) into (146), we obtain the result

$$\langle \tau^{\alpha\beta} \rangle = \begin{cases} \sqrt{x_c^\alpha x_c^\beta} + 1 - x_c^\alpha, & \alpha \ne \beta \ne h \\ \sqrt{x_c^\alpha \left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma\right)} + 1 - x_c^\alpha, & \beta = h, \alpha \ne h \\ \sqrt{\left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma\right) x_c^\beta} + \sum_{\gamma=1}^{h-1} x_c^\gamma, & \alpha = h, \beta \ne h \end{cases} \tag{147}$$

Substituting (147) into (139), we obtain the final solution for the mean fitness of the parallel model in an alphabet of size h, with an arbitrary mutation scheme,

$$\begin{aligned} f_m^{(h)} = \max_{\{x_c^1, x_c^2, \ldots, x_c^{h-1}\}} \Bigg\{ & f(x_c^1, x_c^2, \ldots, x_c^{h-1}) + \sum_{\alpha \ne \beta = 1}^{h-1} \mu_{\alpha\beta} \left[\sqrt{x_c^\alpha x_c^\beta} - x_c^\alpha \right] \\ & + \sum_{\alpha=1}^{h-1} \mu_{\alpha h} \left[\sqrt{x_c^\alpha \left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma\right)} - x_c^\alpha \right] \\ & + \sum_{\beta=1}^{h-1} \mu_{h\beta} \left[\sqrt{\left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma\right) x_c^\beta} + \sum_{\gamma=1}^{h-1} x_c^\gamma - 1 \right] \Bigg\} \end{aligned} \tag{148}$$

Appendix 2

By performing elementary algebraic manipulations (67)

$$3\beta\xi_c + 3\alpha - \frac{3}{2}\mu + \frac{3}{2} \frac{2 - 6\xi_c}{2\sqrt{1 + 2\xi_c - 3\xi_c^2}} = 0$$

can be cast into the standard form of a quartic equation

$$A\xi_c^4 + B\xi_c^3 + C\xi_c^2 + D\xi_c + E = 0 \tag{149}$$

where, by defining $\tilde{\mu} \equiv \mu/\beta$ and $\tilde{\alpha} \equiv \alpha/\beta$, the coefficients correspond to

$$A = 3 \tag{150}$$

$$B = 6\tilde{\alpha} - 3\tilde{\mu} - 2$$

$$C = 3\tilde{\mu}^2 - 3\tilde{\alpha}\tilde{\mu} + 3\tilde{\alpha}^2 + 2\tilde{\mu} - 4\tilde{\alpha} - 1$$

$$D = -2\tilde{\alpha} - 2\tilde{\alpha}^2 + \tilde{\mu} + 2\tilde{\alpha}\tilde{\mu} - 2\tilde{\mu}^2 \tag{151}$$

$$E = -\tilde{\alpha}^2 + \tilde{\alpha}\tilde{\mu}$$

We remark that this quartic equation introduces additional, unphysical solutions to the original (67). However, discarding these unphysical solutions whenever appropriate, the quartic (149) allows us to obtain explicit analytical expressions for ξ_c in the entire region of parameters. Following Ferrari's method [25], we define the parameters

$$a_1 = -\frac{3B^2}{8A^2} + \frac{C}{A} = -\frac{1}{2} - \frac{\tilde{\alpha}}{3} - \frac{\tilde{\alpha}^2}{2} + \frac{\tilde{\mu}}{6} + \frac{\tilde{\alpha}\tilde{\mu}}{2} + \frac{5\tilde{\mu}^2}{8} \tag{152}$$

$$a_2 = \frac{B^3}{8A^3} - \frac{BC}{2A^2} + \frac{D}{A} = -\frac{4}{27} - \frac{4\tilde{\alpha}}{9} + \frac{2\tilde{\mu}}{9} - \frac{\tilde{\mu}^2}{4} - \frac{3\tilde{\alpha}\tilde{\mu}^2}{4} + \frac{3\tilde{\mu}^3}{8} \tag{153}$$

$$a_3 = -\frac{3B^4}{256A^4} + \frac{CB^2}{16A^3} - \frac{BD}{4A^2} + \frac{E}{A}$$

$$= -\frac{5}{432} - \frac{7\tilde{\alpha}}{108} - \frac{5\tilde{\alpha}^2}{72} + \frac{\tilde{\alpha}^3}{12} + \frac{\tilde{\alpha}^4}{16} + \frac{7\tilde{\mu}}{216} + \frac{5\tilde{\alpha}\tilde{\mu}}{72} - \frac{\tilde{\alpha}^2\tilde{\mu}}{8} - \frac{\tilde{\alpha}^3\tilde{\mu}}{8}$$

$$+ \frac{\tilde{\mu}^2}{288} + \frac{3\tilde{\alpha}\tilde{\mu}^2}{16} + \frac{9\tilde{\alpha}^2\tilde{\mu}^2}{32} - \frac{7\tilde{\mu}^3}{96} - \frac{7\tilde{\alpha}\tilde{\mu}^3}{32} + \frac{13\tilde{\mu}^4}{256} \tag{154}$$

and solve the depressed quartic equation in the auxiliary variable $z = \xi_c + B/4A$,

$$z^4 + a_1 z^2 + a_2 z + a_3 = 0 \tag{155}$$

We analyze the different cases in the parameter space that defines the possible solutions of this equation.

Case 1: $a_2 = 0$. This situation arises at the critical value

$$\tilde{\mu}_c^{(1)} = \frac{2}{3} + 2\tilde{\alpha} \tag{156}$$

We obtain four possible roots, according to the general formula

$$\xi_c = -\frac{B}{4A} \pm \sqrt{\frac{-a_1 \pm \sqrt{a_1^2 - 4a_3}}{2}}$$

$$= \frac{1}{6}\left(2 \pm \sqrt{2}\sqrt{1 - 9\tilde{\alpha}(2 + 3\tilde{\alpha})} \pm |1 - 9\tilde{\alpha}(2 + 3\alpha)|\right) \tag{157}$$

Depending on the sign of the term in the square root, we have the following solutions

(i) If $1 - 18\tilde{\alpha} - 27\tilde{\alpha}^2 > 0$. This situation occurs when $-\frac{1}{3} \leq \tilde{\alpha} \leq \frac{1}{3}(\sqrt{\frac{4}{3}} - 1)$, and the solution is

$$\xi_{c,\pm} = \frac{1}{3}(1 \pm \sqrt{1 - 18\tilde{\alpha} - 27\tilde{\alpha}^2}), \qquad \xi_c = \frac{1}{3} \tag{158}$$

(ii) If $1 - 18\tilde{\alpha} - 27\tilde{\alpha}^2 \leq 0$. This situation occurs when $\tilde{\alpha} > \frac{1}{3}(\sqrt{\frac{4}{3}} - 1)$.

$$\xi_c = \frac{1}{3} \tag{159}$$

We shall consider $\tilde{\alpha} \geq 0$ in the region of physically meaningful parameters. When $\tilde{\alpha} = 0$, a non-selective phase is obtained, from (158), if $\beta < \frac{3}{2}\mu$. At $\beta = \frac{3}{2}\mu$, for $\alpha = 0$, a finite 'jump' in the value of ξ_c from 0 to 2/3 defines a phase transition, where the value of u varies continuously from 0 to a positive value.

When $0 \leq \tilde{\alpha} \leq \frac{1}{3}(\sqrt{\frac{4}{3}} - 1)$, a finite jump in the bulk magnetization from $\xi_{c,-}$ to $\xi_{c,+}$ is observed. This result is in agreement with [15].

Case 2: $a_3 = 0$, $a_2 \neq 0$. This situation occurs at the critical values

$$\tilde{\mu}_c^{(2)} = \frac{2}{39}\left(1 + 3\tilde{\alpha} + 2\sqrt{49 - 18\tilde{\alpha} - 27\tilde{\alpha}^2}\right), \quad 0 \leq \tilde{\alpha} \leq 1.054444 \tag{160}$$

$$\tilde{\mu}_c^{(3)} = \frac{2}{39}\left(1 + 3\tilde{\alpha} - 2\sqrt{49 - 18\tilde{\alpha} - 27\tilde{\alpha}^2}\right), \quad 1 \leq \tilde{\alpha} \leq 1.054444 \tag{161}$$

In this case, the quartic equation in z factorizes,

$$z(z^3 + a_1 z + a_2) = 0 \tag{162}$$

There is a solution $z = 0$ for (162). This is however not a solution of (67), but an artifact of introducing the algebraic transformation into the fourth order polynomial (149).

The solutions corresponding to the remaining cubic equation in (162) are analyzed as follows. Let us define the parameters,

$$s_1 = \left[-\frac{a_2}{2} + \left(\frac{a_1^3}{27} + \frac{a_2^2}{4}\right)^{1/2}\right]^{1/3}$$
$$s_2 = \left[-\frac{a_2}{2} - \left(\frac{a_1^3}{27} + \frac{a_2^2}{4}\right)^{1/2}\right]^{1/3} \tag{163}$$

Then, we have the following cases,

Case 2.a: Consider $\tilde{\mu} = \tilde{\mu}_c^{(2)}$, defined by (160). This situation is possible when $0 \leq \tilde{\alpha} \leq 1.054444$. Within this range of values for $\tilde{\alpha}$, the parameter $\frac{a_1^3}{27} + \frac{a_2^2}{4} \geq 0$. Then, we find a single real solution

$$\xi_c = \frac{1}{39}\left(7 - 18\tilde{\alpha} + \sqrt{49 - 18\tilde{\alpha} - 27\tilde{\alpha}^2}\right) + s_1 + s_2 \tag{164}$$

Case 2.b: Consider $\tilde{\mu} = \tilde{\mu}_c^{(3)}$, defined by (161). This situation is possible when $1 \leq \tilde{\alpha} \leq 1.054444$. Within this range of values for $\tilde{\alpha}$, the parameter $\frac{a_1^3}{27} + \frac{a_2^2}{4} \geq 0$. Then, we find a

single real solution

$$\xi_c = \frac{1}{39}\left(7 - 18\tilde{\alpha} - \sqrt{49 - 18\tilde{\alpha} - 27\tilde{\alpha}^2}\right) + s_1 + s_2 \tag{165}$$

Case 3: $a_3 \neq 0$, $a_2 \neq 0$. In this case, we consider again the general quartic (149). Following Ferrari's method [25], we find 4 possible roots

$$\xi_{c,(1,2)} = -\frac{B}{4A} + \frac{1}{2}P \pm \frac{1}{2}Q$$

$$\xi_{c,(3,4)} = -\frac{B}{4A} - \frac{1}{2}P \pm \frac{1}{2}U \tag{166}$$

Here, we defined

$$P = \sqrt{\frac{B^2}{4A^2} - \frac{C}{A} + y_1} \tag{167}$$

$$Q = \begin{cases} \sqrt{\frac{3}{4}\frac{B^2}{A^2} - P^2 - 2\frac{C}{A} + \frac{1}{4}(4\frac{BC}{A^2} - 8\frac{D}{A} - \frac{B^3}{A^3})P^{-1}}, & P \neq 0 \\ \sqrt{\frac{3}{4}\frac{B^2}{A^2} - 2\frac{C}{A} + 2\sqrt{y_1^2 - 4\frac{E}{A}}}, & P = 0 \end{cases} \tag{168}$$

$$U = \begin{cases} \sqrt{\frac{3}{4}\frac{B^2}{A^2} - P^2 - 2\frac{C}{A} - \frac{1}{4}(4\frac{BC}{A^2} - 8\frac{D}{A} - \frac{B^3}{A^3})P^{-1}}, & P \neq 0 \\ \sqrt{\frac{3}{4}\frac{B^2}{A^2} - 2\frac{C}{A} - 2\sqrt{y_1^2 - 4\frac{E}{A}}}, & P = 0 \end{cases} \tag{169}$$

From (166), the largest real root corresponds to the physical solution of (67).

The parameter y_1 in (167–169) is obtained as the real root of the auxiliary cubic equation

$$y^3 + \gamma_2 y^2 + \gamma_1 y + \gamma_0 = 0 \tag{170}$$

Here, we defined the parameters

$$\gamma_2 = -\frac{C}{A}$$

$$\gamma_1 = \frac{BD}{A^2} - 4\frac{E}{A} \tag{171}$$

$$\gamma_0 = 4\frac{CE}{A^2} - \frac{D^2}{A^2} - \frac{B^2E}{A^3}$$

Let us define

$$q = \frac{\gamma_1}{3} - \frac{\gamma_2^2}{9}$$

$$r = \frac{1}{6}(\gamma_1\gamma_2 - 3\gamma_0) - \frac{\gamma_2^3}{27} \tag{172}$$

$$\Delta = q^3 + r^2$$

We have three possible cases: $\Delta > 0$, $\Delta = 0$, and $\Delta < 0$.

Case 3.a: $\Delta > 0$. In this case, we have one real root y_1 for the auxiliary cubic (170), and two complex roots. The real root to be used in (166)–(169) is given by

$$y_1 = (r + \Delta^{1/2})^{1/3} + (r - \Delta^{1/2})^{1/3} - \frac{\gamma_2}{3} \tag{173}$$

Case 3.b: $\Delta = 0$. In this case, all roots of the auxiliary cubic (170) are real, with two of them identical, and given by

$$y_1 = 2r^{1/3} - \frac{\gamma_2}{3}$$
$$y_2 = y_3 = -r^{1/3} \tag{174}$$

In this case, we take the root y_1 in (174), to be used in the formulas (166)–(169).

Case 3.c: $\Delta < 0$. In this case, all three roots of the auxiliary cubic (170) are real and different.

$$y_1 = 2(r^2 - \Delta)^{1/6} \cos(\theta/3) - \frac{\gamma_2}{3}$$
$$y_2 = -2(r^2 - \Delta)^{1/6} \cos(\theta/3 + \pi/3) - \frac{\gamma_2}{3} \tag{175}$$
$$y_3 = -2(r^2 - \Delta)^{1/6} \cos(\theta/3 - \pi/3) - \frac{\gamma_2}{3}$$

Here, $\theta = \tan^{-1}(\frac{(-\Delta)^{1/2}}{r})$. We take the root y_1 to be used in (166)–(169).

Appendix 3

The same arguments based on the homogeneous property of $\lambda_{\max}$ and application of Euler's theorem can be repeated for the case of the Eigen model, to obtain a solution for the non-symmetric case, starting from (105). Now, the matrix to consider is

$$\mathbf{M}(\{\bar{x}_c^\alpha\}, \{\bar{\eta}^{\alpha\beta}\}) = \sum_{\alpha\neq\beta=1}^{h} \bar{\eta}_c^{\alpha\beta} \tau^{\alpha\beta} + \sum_{\alpha=1}^{h-1} \bar{x}_c^\alpha \Theta^\alpha \tag{176}$$

Since $\mathbf{M}$ is clearly a homogeneous function of degree $k = 1$ in the vector $(\bar{x}_c^1, \bar{x}_c^2, \ldots, \bar{x}_c^{h-1}, \{\eta_c^{\alpha\beta}\})$, $\lambda_{\max}$ is homogeneous of degree $k = 1$ as well (the proof is identical as in Property I, Appendix 1 for the parallel model). Therefore, after Euler's theorem (see Appendix 1), we have the identity

$$\lambda_{\max} - \sum_{\alpha=1}^{h-1} \bar{x}_c^\alpha x_c^\alpha - \sum_{\alpha\neq\beta=1}^{h} \bar{\eta}_c^{\alpha\beta} \eta_c^{\alpha\beta} = 0 \tag{177}$$

After the saddle-point (103)–(104), we obtain the same components for the eigenvalue $\vec{y}_{\max}$ as in the parallel case, (146),

$$y^\alpha = \begin{cases} \sqrt{x_c^\alpha}, & 1 \leq \alpha \leq h - 1 \\ \sqrt{1 - \sum_{\gamma=1}^{h-1} x_c^\gamma}, & \alpha = h \end{cases} \tag{178}$$

Thus, by applying Property III (Appendix 1) for the Eigen model, we obtain

$$\eta_c^{\alpha\beta} = \langle \tau^{\alpha\beta} \rangle = \vec{y}_{\max}^T \tau^{\alpha\beta} \vec{y}_{\max}$$

$$= \begin{cases} \sqrt{x_c^\alpha x_c^\beta} + 1 - x_c^\alpha, & \alpha \neq \beta \neq h \\ \sqrt{x_c^\alpha \left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma\right)} + 1 - x_c^\alpha, & \beta = h, \alpha \neq h \\ \sqrt{\left(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma\right) x_c^\beta} + \sum_{\gamma=1}^{h-1} x_c^\gamma, & \alpha = h, \beta \neq h \end{cases} \quad (179)$$

Thus, the mean fitness for the h-states Eigen model under arbitrary mutation scheme is given by the expression

$$f_m = \max_{\{x_c^1, x_c^2, \ldots, x_c^{h-1}\}} \left\{ f(x_c^1, x_c^2, \ldots, x_c^{h-1}) e^{\sum_{\alpha \neq \beta=1}^{h-1} \mu_{\alpha\beta}[\sqrt{x_c^\alpha x_c^\beta} - x_c^\alpha]} \right.$$

$$\times e^{\sum_{\alpha=1}^{h-1} \mu_{\alpha h}[\sqrt{x_c^\alpha(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma)} - x_c^\alpha] + \sum_{\beta=1}^{h-1} \mu_{h\beta}[\sqrt{(1 - \sum_{\gamma=1}^{h-1} x_c^\gamma) x_c^\beta} + \sum_{\gamma=1}^{h-1} x_c^\gamma - 1]}$$

$$\left. - d(x_c^1, x_c^2, \ldots, x_c^{h-1}) \right\} \quad (180)$$

References

1. Eigen, M., Schuster, P.: Selforganization of matter and evolution of biological macromolecules. Naturwissenschaften **58**, 465 (1971)
2. Eigen, M., McCaskill, J., Schuster, P.: Molecular quasi-species. J. Phys. Chem. **92**, 6881–16891 (1988)
3. Eigen, M., McCaskill, J., Schuster, P.: The molecular quasi-species. Adv. Chem. Phys. **75**, 149–263 (1989)
4. Crow, J., Kimura, M.: An Introduction to Population Genetics Theory. Harper and Row, New York (1970)
5. Biebricher, C.K., Eigen, M.: The error threshold. Virus Res. **107**, 117–127 (2005)
6. Domingo, E., Sabo, D., Taniguchi, T., Weissman, C.: Nucleotide sequence heterogeneity of an RNA phage population. Cell **13**, 735–744 (1978)
7. Domingo, E., Escarmis, C., Lazaro, E., Manrubia, S.C.: Quasispecies dynamics and RNA virus extinction. Virus Res. **107**, 129–139 (2005)
8. Leuthausser, I.: Statistical mechanics of Eigen's evolution model. J. Stat. Mech.: Theor. Exp. **48**, 343–360 (1987)
9. Tarazona, P.: Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin-glass models. Phys. Rev. A **45**, 6038–6050 (1992)
10. Baake, E., Baake, M., Wagner, H.: Ising quantum chain is equivalent to a model for biological evolution. Phys. Rev. Lett. **78**, 559–562 (1997)
11. Baake, E., Baake, M., Wagner, H.: Quantum mechanics versus classical probability in biological evolution. Phys. Rev. E **57**, 1191–1192 (1998)
12. Baake, E., Wagner, H.: Mutation-selection models solved exactly with methods of statistical mechanics. Genet. Res. Camb. **78**, 93–117 (2001)
13. Saakian, D.B., Hu, C.-K.: Eigen model as a quantum spin chain: Exact dynamics. Phys. Rev. E **70**, 021913 (2004)
14. Saakian, D.B., Muñoz, E., Hu, C.-K., Deem, M.W.: Quasispecies theory for multiple-peak fitness landscapes. Phys. Rev. E **73**, 041913 (2006)
15. Hermisson, J., Wagner, H., Baake, M.: Four-state quantum chain as a model for sequence evolution. J. Stat. Phys. **102**, 315–343 (2001)
16. Garske, T., Grimm, U.: A maximum principle for the mutation-selection equilibrium of nucleotide sequences. Bull. Math. Biol. **66**, 397–421 (2004)

17. Garske, T., Grimm, U: Maximum principle and mutation thresholds for four-letter sequence evolution. J. Stat. Mech.: Theor. Exp., P07007 (2004)
18. Baake, E., Baake, M., Bovier, A., Klein, M.: An asymptotic maximum principle for essentially linear evolution models. J. Math. Biol. **50**, 83–114 (2005)
19. Park, J.-M., Deem, M.W.: Schwinger boson formulation and solution of the Crow-Kimura and Eigen models of quasispecies theory. J. Stat. Phys. **125**, 975–1015 (2006)
20. Park, J.-M., Deem, M.W.: Phase diagrams of quasispecies theory with recombination and horizontal gene transfer. Phys. Rev. Lett. **98**, 05101 (2007)
21. Muñoz, E., Park, J.-M., Deem, M.W.: Quasispecies theory for horizontal gene transfer and recombination. Phys. Rev. E **78**, 061921 (2008)
22. Kimura, M.: Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl. Acad. Sci. USA **78**, 454–458 (1981)
23. Hermisson, J., Redner, O., Wagner, H., Baake, E.: Mutation-selection balance: Ancestry, load, and maximum principle. Theor. Popul. Biol. **62**, 9–42 (2002)
24. Fritz, J.: Partial Differential Equations, 4th edn. Springer, Berlin (1982)
25. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions. Dover, New York (1970)