A statistical-mechanical study of evolution of robustness: An approach from two-temperature models

Ayaka Sakata¹, Koji Hukushima¹ and Kunihiko Kaneko¹,²

¹Graduate school of Arts and Sciences, The University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan
²Complex Systems Biology Project, Exploratory Research for Advanced Technology (ERATO), Japan Science and Technology Agency (JST), Tokyo, Japan

E-mail: ayaka@huku.c.u-tokyo.ac.jp, hukusima@phys.c.u-tokyo.ac.jp, kaneko@complex.c.u-tokyo.ac.jp

Abstract. Recently we proposed a simple statistical-mechanical model of interacting spins for studying evolutionary process in biological system. In our model, a configuration of spins and their interactions represent a phenotype and genotype, respectively. The phenotype dynamics are given by a stochastic process with temperature $T_S$ under a fixed Hamiltonian with $J$. Meanwhile, the evolution of $J$ is also stochastic with temperature $T_J$ and follows mutations introduced into $J$ and selection based on a fitness given by equilibrium spin configurations of a given set of target spins. Using Monte Carlo simulations, it is found that frustration around the target spins is strongly suppressed for the interactions $J$ evolved in the intermediate $T_S$ temperature region. The evolved $J$’s give the funnel-like dynamics and show robustness to mutations. This model can be regarded as an extension of partial annealing model, in which the interactions between spins are dynamical variables with characteristic time scale widely separated from that of the spins. We also discuss the partially annealed Sherrington-Kirkpatrick model by particularly paying attention to the frustration.

1. Introduction

The function of a biological unit is generally determined by a biological state, phenotype, that is shaped by a dynamical process that yields a specific pattern or structure. The dynamics themselves are governed by a genetic sequence. In evolution governed by a fixed condition, a phenotype that gives a higher value for the fitness function has a high chance of survival. The genes that produce such a phenotype are transferred to the next generation, and thus, specific genetic sequences are selected. However, it should be noted that the dynamical process producing a phenotype is subject to noise. For example, genetic information determines the amino acid sequence for a protein, while the tertiary structure responsible for its functions is generated only by folding dynamics. The genotype-phenotype mapping is formed by this folding process, and this mapping is stochastic because of the thermal noise in the folding process[1]. Related folding dynamics also occur in t-RNA[2]. Hence, phenotypes shaped from a genotype through the dynamical process are not necessarily identical, and therefore, they form a distribution.

The phenotype for the function is expected to be insensitive to noise encountered during the dynamical process[3]. However, such robustness is not a general property of dynamics. For
example, the complex folding process of a hetero-polymer from an arbitrary random sequence might not have such robustness. In this sense, the robustness could be a result of evolution. To have robustness to noise in such dynamics, it is ideal to adopt dynamics in which the target phenotype is reached smoothly and globally from a variety of initial configurations and is maintained thereafter. In fact, the existence of such global attraction in the protein folding process was proposed as a consistency principle by Go [4] and as “funnel” landscape by Onuchic et al. [5, 6], while similar global attraction dynamics have been discovered recently in gene regulatory networks [7] and developmental dynamics [8]. In spite of the ubiquity of such funnel-like structures for phenotype dynamics, little is understood about how these structures are shaped by the evolutionary process [1].

In addition to being robust to noise, a biological system has to remain relatively robust to mutations in the genetic sequence that occur through evolution; the phenotype has to be rather insensitive to changes in the genetic sequence. Are these two types of robustness correlated? Indeed, a possible relationship between robustness to noise and robustness to mutation has been discussed [9, 10, 11, 12, 13, 14, 15], following the pioneering study by Waddington [16] on the evolution-development relationship, which is referred to as canalization and genetic assimilation. However, a theoretical understanding of the evolution of robustness is still insufficient.

In order to understand the concepts in the evolution of robustness, we introduce a simple statistical-mechanical spin model that consists of $N$ Ising spins interacting globally [17, 18]. Indeed, this type of model has also been adopted by Saito et al. [1], who utilized it in the study of the evolution of protein folding dynamics. Even though the spin model is abstract, it can account for the basic structures required to study the evolution of genotype and phenotype. In comparison with the gene transcription network model utilized in the study of the evolution of robustness [9, 10, 13, 14], the present spin model is computationally efficient in Monte Carlo simulations and the methods developed in statistical mechanics of spin systems can be applied to study the general questions on evolution. Our model mainly discussed in this paper can be regarded as a variant of partial annealing models with two temperatures [19, 20]: one for the spins and the other for the interactions. The original partial annealing models can be analyzed by replica method with a replica number fixed to be ratio of two temperatures, while the corresponding replica analysis for our model is straightforward unfortunately. We thus reveal novel features of our model by using Monte Carlo simulations and compare the results with those of a partial annealing model.

This paper is organized as follows. In Sec. 2, we explain the model setup that captures the essential features of the evolution. In Sec. 3, the numerical results for the model are reviewed. This part is largely based on our previous papers [17, 18]. Then, we show that a robust system is realized at an intermediate temperature. A key for the robustness is frustration of the interactions. In Sec. 4, we discuss the frustration of partially annealed Sherrington-Kirkpatrick (SK) model in comparison with our findings. Finally, Sec. 5 is devoted to the conclusions and prospects for further development.

### 2. A spin model for evolution of robustness

We introduce a statistical-mechanical spin model in which the phenotype and genotype are represented by configurations of spin variables $S_i$ and an interaction matrix $J_{ij}$, respectively, with $i, j = 1, \cdots, N$. The spins $S_i$ and $J_{ij}$ can take one of only two values $\pm 1$, and the interaction matrix is assumed to be symmetric, i.e., $J_{ij} = J_{ji}$. A set of configurations is denoted by $S$ for the phenotype and by $J$ for the genotype. The dynamics of the phenotype are given by a flip-flop update of each spin with an energy function, which is defined by the Hamiltonian for a given $J$,

$$H(S|J) = -\frac{1}{\sqrt{N}} \sum_{i<j} J_{ij} S_i S_j.$$  (1)
We adopt the Glauber dynamics as an update rule of \( S \), where the \( N \) spins are in contact with their own heat bath at temperature \( T_S \). The Glauber dynamics, satisfying the detailed balance conditions, yields an equilibrium distribution for a given \( J \):

\[
P(S|J, T_S) = \frac{e^{-\beta_S H(S|J)}}{Z_S(J, T_S)},
\]

where \( Z_S(J, T_S) = \text{Tr}_S e^{-\beta_S H(S|J)} \) and \( \beta_S = T_S^{-1} \). After a relaxation process, the phenotype \( S \) follows from the equilibrium distribution, and it is not determined uniquely from the genotype \( J \); rather, it is distributed, except at zero temperature. The phenotype fluctuation is computed from the resulting equilibrium probability distribution. Thus, the degree of fluctuation is characterized by the temperature \( T_S \). When the interactions \( J \) are quenched variables given by Gaussian distribution, the model is identical to the SK model. In our model, however, they are supposed to be dynamical variables.

Next, we introduce evolutionary dynamics for the genotype \( J \). The genotype is transmitted to the next generation with some variation, while genotypes that produce a phenotype with higher fitness are selected. The time scale for genotype change is generally much larger than that of the genotype dynamics. We assume that the two time scales for the phenotype expression dynamics and the genotypic evolutionary dynamics are widely separated, so that the variables \( S \) are well equilibrating within the unit time scale of the slow variable \( T \). Then, the fitness should be expressed by a function of the phenotype \( S \) that is averaged with respect to the distribution. Here, we define the fitness as

\[
\Psi(J|T_S) = \left\langle \prod_{i<j \in t} \delta(S_i - S_j) \right\rangle \equiv \langle \psi \rangle,
\]

where \( \langle \cdots \rangle \) denotes the expectation value with respect to the equilibrium probability distribution of eq.(2). The set \( t \) denotes a subset of \( S \) with size \( t \); the members of \( t \) are termed as target spins. We refer to configurations such that all target spins are aligned in parallel as target configurations, which are assumed to give a requested appropriate function. By a gauge transformation on the target spin and the corresponding elements of \( J \), a choice of any other form of spin alignment for the fitness function, instead of the “ferromagnetic” configuration, yields the same result [21]. The fitness can be interpreted as the average frequency of finding the target configurations in equilibrium for a given \( J \). It should be noted that in our model, only the target spins contribute explicitly to the fitness and the remaining spins have no direct influence on the fitness and the selection of genes. Hence, the spin configuration for a given fitness has redundancy.

The genotype dynamics are a result of mutations and selection, i.e., changes according to the fitness function following random flip-flops of genes. Hence, for a genetic dynamics, we once again adopt the Glauber dynamics by using the fitness instead of the Hamiltonian in the phenotype dynamics, where the genotype \( J \) is in contact with a heat bath whose temperature \( T_J \) is different from \( T_S \). The dynamics for the genotype are given by a stochastic Markov process with the following stationary distribution:

\[
P(J|T_S, T_J) = \frac{e^{\beta_J \Psi(J|T_S)}}{Z_J(T_S, T_J)},
\]

where \( Z_J(T_S, T_J) = \text{Tr}_J e^{\beta_J \Psi(J|T_S)} \) and \( \beta_J = T_J^{-1} \). According to the dynamics, genotypes are selected rather uniformly at high values of the temperature \( T_J \), irrespective of the fitness, whereas at low values of \( T_J \), the genotypes with higher fitness values are preferentially selected. In this sense, the temperature \( T_J \) represents the selection pressure among mutated genotypes. In actual
Figure 1. Two-temperature dependence of the fitness $\Psi(J)$ in (a) and the energy in (b), respectively, with $N = 15$ and $t = 3$. Figure is quoted from [18].

simulations, a candidate $J'$ for the next generation is set by some flips of a randomly chosen $J_{ij}$ from the current $J$, while the transition probability from $J$ to $J'$ is given by Metropolis rules, $\min(1, \exp(\beta(J) - \Psi(J'))).

3. Numerical Results

3.1. Fitness and Energy

We have carried out MC simulations of the model discussed above and studied the dependence of the fitness and energy on $T_S$ and $T_J$. They are given by

$$\Psi(T_S, T_J) = \langle \Psi(J|T_S) \rangle_J, \quad E(T_S, T_J) = \langle \langle H(S|J) \rangle \rangle_J,$$

respectively, where $\langle \cdots \rangle_J$ denotes the average with respect to the equilibrium probability distribution, $P(J|T_S, T_J)$. MC sampling with temperature $T_S$ under the Hamiltonian $H$ and the stochastic selection process governed by the fitness are carried out alternately. In our simulations of the spin dynamics, the exchange Monte Carlo simulation (EMC) [22] is introduced to accelerate the relaxation time to equilibrium and obtain the equilibrium spin distribution efficiently. In this section, we concentrate on the analysis of the equilibrium state. Figure 1 (a) and (b) show the dependence of the fitness and the energy on $T_S$ and $T_J$, respectively, for $N = 15$ and $t = 3$. For each generation of the genotype dynamics, the fitness and energy are averaged with respect to the equilibrium distribution over 1500 MC steps after discarding the first 1500 MC steps; this number of steps is sufficient for equilibration. The data are averaged over the last 1000 generations. For any $T_S$, the fitness decreases monotonically with $T_J$, but the rate of decrease is affected significantly by $T_S$. The fitness for sufficiently low $T_S$ remains at a high level and decreases only slightly with an increase in $T_J$, while for a medium value of $T_S$, the fitness gradually decreases to a lower level as a function of $T_J$, and eventually, for a sufficiently high value of $T_S$, it never reaches a high level. This result implies that the structure of the fitness landscape depends on $T_S$, the temperature at which the system has evolved.

The energy function, on the other hand, shows a significant dependence on $T_S$. While the energy is represented by a monotonically increasing function of $T_S$ for high $T_J$, it exhibits non-monotonic behavior for low $T_J$ and has a minimum at $T_S \approx 2.0$. The configurations that include the target pattern at the energy minimum are obtained around $T_S \approx 2.0$. This non-monotonicity of the energy corresponds to a negative specific heat in the sense of standard thermodynamics. The configurations of $J$ giving rise to the highest fitness value generally have a huge redundancy. Using a fluctuation induced by $T_S$, a specific subset of the configurations of $J$ with lower energy is selected among the redundant configurations at around $T_S \approx 2.0$.  


Figure 2. Dependence of local frustrations on \( T_S \); \( \Phi_1(\square), \Phi_2(\bigcirc), \Phi_3(\triangle) \) represents the left axis and \( \mu_c(\ast) \) represents the right axis by fixing \( T_J \) at \( 0.5 \times 10^{-3} \). The data are computed by taking averages over 150 genotypes \( \mathbf{J} \) evolved at a given temperature \( T_S \). \( \mu_c \) is a threshold value, beyond which the fitness of the mutated genotype begins to decrease. The transition points \( T_S^{c1} \) and \( T_S^{c2} \) are estimated as a temperature at which \( \Phi_2 \) deviates from 1. Figure is quoted from [17].

### 3.2. Frustration

In the medium-temperature range, both a lower energy and a higher fitness are achieved. What is the structure in \( \mathbf{J} \) configurations that helps to achieve this? The statistical physics of spin systems tells us that a decrease in energy implies a decrease of the frustration in spin configurations. Triplets of interactions that satisfy \( J_{ij}J_{jk}J_{ki} < 0 \) are known to yield frustration, which is an obstacle to attaining the unique global energy minimum[21, 23]. Meanwhile, all the interactions satisfying \( J_{ij}J_{jk}J_{ki} > 0 \) do not have frustration, and the energy of the spin states attains the minimum value. However, the energetically favorable spin configuration cannot be uniquely determined only by the condition \( J_{ij}J_{jk}J_{ki} > 0 \). The spin configurations that have low energy should be the target configurations when both the decrease in energy and the increase in fitness are simultaneously achieved. In our case, the target spins play a distinct role, and therefore, we need to quantify the frustration while distinguishing between target and non-target spins; this is in contrast to the standard spin-glass study.

The interactions are divided into three categories: those between target spins, \( J_{tt} (\{J_{ij} \mid i, j \in t\}) \), those between target and non-target spins, \( J_{to} (\{J_{ij} \mid i \in t, j \notin t\}) \), and those between non-target spins, \( J_{oo} (\{J_{ij} \mid i, j \notin t\}) \). It can be assumed that the frustration of all categories decreases at intermediate \( T_S \). To confirm this, we should define the conditional frustration for each category of spins. We first define \( \Phi_1 \) as the frequency of positive coupling among target spins, i.e.,

\[
\Phi_1(T_S, T_J) = \frac{2}{t(t-1)} \left[ \sum_{i<j \in t} J_{ij} \right] J,
\]

(6)

The target configurations are energetically preferred under ferromagnetic coupling, i.e., \( \Phi_1 = 1 \), for which no frustration exists among the target spins. Second, we define \( \Phi_2 \) as

\[
\Phi_2(T_S, T_J) = \frac{2}{t(t-1)(N-t)} \left[ \sum_{i<j \in t} \sum_{k \notin t} J_{ik}J_{jk} \right] J,
\]

(7)

where \( t(t-1)(N-t)/2 \) is the total number of possible spins: two target spins and one non-target spin. Here, \( \Phi_2 = 1 \) implies that no frustration exists in the interactions between target and non-target spins, and thus, the target configuration is at an energy minimum even when these interactions are included. Lastly, as a measure of the frustration among non-target spins, \( \Phi_3 \) is defined as

\[
\Phi_3(T_S, T_J) = \frac{1}{C_2^{N-t}} \left[ \sum_{k<l \notin t} \left( \frac{1}{t} \sum_{i \in t} J_{ik} \right) \left( \frac{1}{t} \sum_{j \in t} J_{jl} \right) \right] J,
\]

(8)
where \( C_2^{N-k} \) is the total number of possible pairs of non-target spins. If \( \Phi_3 \) is equal to 1, the frustration is not introduced by the interactions between non-target spins; in other words, there is no frustration globally. Hence, the system with \( \Phi_1 = \Phi_2 = \Phi_3 = 1 \) is in the Mattis state [24], which can be transformed to ferromagnetic interaction by a gauge transformation.

For the interaction \( J \) with evolution under an environment with temperature \( T_S \), we have computed \( \Phi_1, \Phi_2, \) and \( \Phi_3 \) by performing MC simulations. In Figure 2, we present the \( T_S \) dependence of \( \Phi_1(T_S, T_J), \Phi_2(T_S, T_J), \) and \( \Phi_3(T_S, T_J) \) at a fixed \( T_J = 0.5 \times 10^{-3} \). At sufficiently low \( T_J \), the frustration parameters attain the maximum value 1 at the intermediate \( T_S \), where the frustrations are extensively eliminated, while they remain finite at low \( T_S \) and high \( T_S \). We define the intermediate temperature region as \( T_S^1 < T_S < T_S^2 \), where the frustration parameter \( \Phi_2 \) equals 1. These temperatures depend on \( T_J \), and we express them as \( T_S^1(T_J) \) and \( T_S^2(T_J) \). We see that for low \( T_J \) (\( \lesssim 0.05 \)), the phase diagram is split into three phases. The first one is frustrated and adapted phases for \( T_S < T_S^1 \) (\( T_J \)). For \( T_S < T_S^1(T_J) \), all \( \Phi_i \) (\( i = 1, 2, 3 \)) are less than unity, and hence, the frustration remains for target and non-target spins.

For \( T_S \geq T_S^1(T_J) \), \( \Phi_1 \) equals 1, so that a target configuration is embedded as an energetically favorable state. For a finite system with finite \( T_J \), \( \Phi_j \) cannot be exactly 1. However, as long as \( T_J \) is low, the deviation of \( \Phi_j \) from 1 at the intermediate temperature is negligible. For \( T_S^1 \leq T_S \leq T_S^2(T_J) \), \( \Phi_2 \) is also equal to 1, implying that the frustration among spins is not introduced via interactions with a non-target spin. In this temperature range, \( \Phi_3 \) is not always equal to 1, except for \( T_S \sim 2.0 \), where the Mattis state arises. When \( \Phi_2 = 1 \) and \( \Phi_3 \neq 1 \), the frustration is not completely eliminated from the non-target spin interactions \( J_{oo} \); this is in contrast to the Mattis state. We call such a \( J \) configuration “local Mattis state” (LMS), and it is characterized by \( \Phi_1 = \Phi_2 = 1 \) but \( \Phi_3 \neq 1 \). This implies that the interactions \( J \) have no frustration around the target spins, but there is some frustration between non-target spin interactions. The interactions \( J \) required to form such a local Mattis state are obtained as a consequence of the evolution around \( T_S^1 \leq T_S \leq T_S^2(T_J) \) for low \( T_J \), where both the fitted target configuration and lower energy are achieved. The \( T_S \) range in which the local Mattis state is stabilized becomes narrower with an increase in \( T_J \). The phase diagram of the model is shown in figure 3. For \( T_S > T_S^2(T_J) \), the frustration parameter \( \Phi_2 \) is less than 1, and consequently, the frustration remains, and the fitness \( \Psi \) starts to decrease and the energy increases. Thus, neither adaptation nor energy minimization is achieved.

**Figure 3.** Phase diagram of the evolved \( J \) at \( N = 15 \) and \( t = 3 \). Three types of evolved \( J \) are defined on the basis of the value of the fitness and \( \Phi_2 \).
3.3. Robustness to mutation

We examine the robustness that represents the stability of \( J \)’s fitness with respect to changes in the \( J \) configuration. From the evolved genotype under a given \( T_S \) and \( T_J \), denoted as \( J_{TS} \), mutations are imposed by flip-flopping the sign of a certain fraction of randomly chosen matrix elements in \( J_{TS} \). The value of the fraction represents the mutation rate \( \mu \). We evaluate the fitness of the mutated \( J_{TS}^{\text{adp}}(\mu) \) at \( T_S = 10^{-5} \) (\( \neq T_S \)), i.e., \( \Psi(J_{TS}^{\text{adp}}(\mu)|T_S = 10^{-5}) \), by taking an average over 150 samples of mutated \( J_{TS}^{\text{adp}}(\mu) \). Figure 4 shows \( \mu \) dependence of the fitness for \( T_S = 10^{-4} \) and \( T_S = 2.0 \). For low values of \( T_S \), the fitness of mutated \( J_{TS}^{\text{adp}}(\mu) \) exhibits a rapid decrease with respect to the mutation rate, but for \( T_S \) between \( T_S^1 \) and \( T_S^2 \), it does not decrease until the mutation rate reaches a specific value. We define \( \mu_c(T_S) \) as the threshold point in the mutation rate beyond which the fitness \( \Psi(J_{TS}^{\text{adp}}(\mu)|T_S) \) begins to decrease from unity. Figure 2 shows the dependence of \( \mu_c(T_S) \), which has a plateau at \( T_S^1 \leq T_S \leq T_S^2 \), where \( \Phi_2 \) is unity. This range of temperatures that exhibits mutational robustness agrees with the range that gives rise to the LMS. In other words, mutational robustness is realized for a genotype with no frustration around the target spins. Evolution in a mutationally robust genotype \( J \) is possible only when the phenotype dynamics is subjected to noise within the range \( T_S^1 \leq T_S \leq T_S^2 \). This mutational robustness is interpreted as a consequence that the fitness landscape becomes non-neutral for \( T_S \geq T_S^1 \) [18].

4. Frustration of partially annealed SK model

In our model discussed so far, the fitness is defined as a function of equilibrium configurations of the target spins. Instead, when the free energy \( \beta_S F(J, T_S) = -\log Z_S(J, T_S) \) is chosen as the fitness, the model system is equivalent to the partial annealing[19, 20] for the SK model, which enables us to study by the replica method. The frustration parameter \( \Phi_3 \) might be defined as

\[
\Phi_3 = \frac{1}{C_3^N} \sum_{i<j<k} J_{ij} J_{jk} J_{ki}.
\]  

(9)

In the case of the SK model, the frustration \( [\Phi_3]_J \) can be calculated analytically, as shown in Dotsenko et al.[20]. We explain their method briefly and derive the expression of \( [\Phi_3]_J \) under the replica symmetry (RS) ansatz. The product of \( J \) can be described by the product of spins by using the fact that \( P_0(J) \) is Gaussian distribution, as follows,

\[
[\Phi_3]_J = \frac{1}{[\mathcal{Z}^n]_J} \int D J P_0(J) Z^n(J) \left( \frac{1}{C_3^N} \sum_{i<j<k} J_{ij} J_{jk} J_{ki} \right) = \beta_S \sqrt{N C_3^N} \sum_{\alpha, \beta, \gamma} \sum_{i<j<k} [\langle g^a_i g^b_j g^c_k S_i^a S_j^b S_k^c \rangle]_J.
\]
shows the spin-glass order parameter from top to bottom. The dashed line is \( n\beta_S^3 \) and the dotted line is \( \beta_j^3 = (n\beta_S)^3 \). The point \( \times \) shows the \([\Phi_3]_J\) for the 1 step RSB solution obtained in eq.(A.3).

(b): \( T_S \) dependence of the normalized frustration parameter, \( (|\Phi_3|_J - n\beta_S^3)/(\beta_j^3 - n\beta_S^3) \), for each \( n \) shown in (a).

Taking the leading term of \( N \) in the summation \( \sum_{i<j<k} \), \([\Phi_3]_J\) can be expressed as

\[
N^{3/2}[\Phi_3]_J = \left( \frac{\beta_S}{N} \right)^3 \sum_{\alpha,\beta,\gamma} \left[ \left( \sum_i S_i^\alpha S_i^\gamma \sum_j S_j^\beta \sum_k S_k^\gamma \right) \right]_J = \beta_S^3 \text{Tr}_n(I + Q)^3, \tag{10}
\]

where \( Q \) is \( n \times n \) matrix whose diagonal elements take 0 and off-diagonal ones are \( q^{\alpha \beta} (\alpha \neq \beta) \), and \( I \) is \( n \times n \) identical matrix. In the RS case, all of the non-diagonal components of the \( n \times n \) matrix \( Q \) take the same value \( q \). The spin-glass order parameter \( q \) under the RS ansatz with finite-\( n \) is given as

\[
q = \frac{\int Dz \tanh^2(\beta_S \sqrt{q} z) \cosh^3(\beta_S \sqrt{q} z)}{\int Dz \cosh^3(\beta_S \sqrt{q} z)}, \tag{12}
\]

where \( Dz = dz e^{-z^2/2}/\sqrt{2\pi} \). The expression of \([\Phi_3]_J\) for 1 step replica-symmetry breaking is shown in Appendix.

Figure 5(a) shows the dependence on \( T_S \) of \([\Phi_3]_J\) in the thermodynamic limit for \( n = 3, 2, 1, 0.5 \), and 0.3, beginning at the top, respectively. The dashed and dotted lines are \( n\beta_S^3 \) and \( \beta_j^3 \), respectively. The spin-glass transition temperature is \( T_S \approx 1.15 \) for \( n = 3 \) and \( T_S = 1.0 \) for other \( n \). At \( T_S > T_S^* \), \([\Phi_3]_J\) behaves as \( n\beta_S^3 \). At \( T_S^* \), the spin-glass order parameter \( q \) begins to take a finite value and \([\Phi_3]_J\) deviates from \( n\beta_S^3 \), except in the case of \( n = 1 \) where \([\Phi_3]_J\) always behaves as \( n\beta_S^3 \) for any \( q \). At sufficiently low \( T_S \), \([\Phi_3]_J\) scales as \( \beta_j^3 \) below a certain temperature that varies with \( n \). In this way, crossover of frustration from \( n\beta_S^3 \) to \( \beta_j^3 = (n\beta_S)^3 \) is observed except for \( n = 1 \). The fact that \([\Phi_3]_J\) is equal to \( \beta_j^3 \) means that the overlap \( q \) takes the value 1.

Then, \([\Phi_3]_J\) takes the possible maximum value, implying that the frustration is reduced as less as possible.

For \( n > 1 \), \( \beta_j^3 \) is larger than \( n\beta_S^3 \) for whole temperature region. Therefore, the spin-glass transition induces the increase of \([\Phi_3]_J\) that enhances the elimination of the frustration. On
the other hand, for \( n \leq 1 \) where \( n\beta_S^3 \) is less than \( \beta_J^3 \), \( [\Phi_J]_J \) increases slowly in comparison with the case for \( n > 1 \) and the crossover to less-frustrated state occurs at very low temperature. In figure 5(b), we show the normalized frustration parameter, \((\Phi_J - n\beta_S^3)/(\beta_J^3 - n\beta_S^3)\), which takes the value within \([0,1]\). It begins to increase at the spin-glass transition temperature and converges to 1 at low \( T_S \).

To show a temperature region where the frustration sufficiently close to the scaling \( \beta_J^3 \), we plot a contour map on \( T_S \) and \( T_J \) of the normalized frustration parameter in figure 6. At low \( T_J \), the less-frustrated region is expanded, similar to that found in our model. However, the reduction of frustration is induced by the emergence of spin-glass order in this partially annealed SK model. In fact, the less-frustrated region appears at high \( T_J \) and low \( T_S \), where the spin-glass order exists, and this is sharp contrast to that found in our model. This might be due to the fact that the free energy is chosen as the fitness. In order to understand the local Mattis state found in our model, some local structure in the fitness is to be taken into account. This is an issue in the future work.

5. Summary and discussion
We have considered the evolution of a Hamiltonian system to generate a specific configuration for target spins that captures the basic features required to study the evolution. In this study, we adopted a Markov process, which is given by temperature \( T_J \) and fitness \( \Psi(J) \), for evolutional dynamics. By performing numerical simulation, we found that a specific subset of \( J \) with low energy and high fitness is evolved at an intermediate \( T_S \) and low \( T_J \). From the statistical-mechanical viewpoint, we focused on frustration and found that the interactions \( J \) that evolved at the intermediate \( T_S \) are less frustrated. We called these \( J \) the local Mattis states. In general, the less frustrated \( J \) states are robust to mutation. Hence, the robustness of evolving states to mutation is realized at intermediate temperatures \( T_S^1 \leq T_S \leq T_S^2 \). In other words, robustness to thermal noise introduces mutational robustness; this has also been recently discussed for gene regulation network models[13]. The relevance of thermal noise to robust evolution is thus demonstrated.

In this study, we observed transitions at \( T_S^1 \) and \( T_S^2 \). The phase below \( T_S^1 \) corresponds to the spin-glass phase and that above \( T_S^2 \) corresponds to the paramagnetic phase in the context of statistical physics; in contrast, the local Mattis phase between the two phases, which could correspond to the funnel landscape, is a novel discovery in this study. Since a framework of statistical physics of spin systems has been adopted in the present study, the theoretical concepts developed therein, such as replica symmetry breaking, may be applicable in the context
developed here to understand this transition. A promising direction would be the study of the partial annealing models with the replica theory. Some related systems based on the partial annealing have been studied in various areas; protein folding and its evolution [25], ligand-receptor reaction [26] and so on. In fact, when the fitness is set to the free energy, our model is reduced to the partial annealing version of the SK model. We can find a similarity between our model and its partial annealing model. However, some points, in particular the mechanism for the emergence of local Mattis state, are less clear and remains to be solved in the future.

Acknowledgments

We would like to thank C. H. Nakajima and T. Nakajima for helpful comments and discussions. This work was partially supported by a Grant-in-Aid for Scientific Research (No.18079004) from MEXT and JSPS Fellows (No.20-10778) from JSPS.

Appendix A. 1RSB expression of frustration $[\Phi_3]^J$ for the Sherrington-Kirkpatrick model with finite replica number $n$

In this appendix, we discuss how the frustration $[\Phi_3]^J$ is modified by taking into account replica symmetry breaking. At the level of one-step replica symmetry breaking (1RSB), the saddle-point equations for the order parameters $q_0$ and $q_1$ are given as

$$q_0 = \int D_u \left( \int Dv \cosh^m \Xi \right)^{n/m} \left( \frac{\int Dv \cosh^m \Xi \tanh \Xi}{\int Dv \cosh^m \Xi} \right)^2 / \int D_u \left( \int Dv \cosh^m \Xi \right)^{n/m},$$

$$q_1 = \int D_u \left( \int Dv \cosh^m \Xi \right)^{n/m} \left( \frac{\int Dv \cosh^m \Xi \tanh^2 \Xi}{\int Dv \cosh^m \Xi} \right) / \int D_u \left( \int Dv \cosh^m \Xi \right)^{n/m} \quad (A.1)$$

with $\Xi = \beta_S (\sqrt{q_0^2} + \sqrt{q_1 - q_0^2})$. Here $Du$ and $Dv$ are the Gaussian kernel, $DX = dX e^{-X^2/2} / \sqrt{2\pi}$. The value of replica-partition parameter $m$ is determined as a solution of

$$\frac{\beta_S^2}{4} (q_1^2 - q_0^2) m = \int \tilde{D}_u \left( \int Dv \cosh^m \Xi \log \cosh \Xi \right) / \int Dv \cosh^m \Xi - \frac{1}{m} \int \tilde{D}_u \log \left( \int Dv \cosh^m \Xi \right). \quad (A.2)$$

With these parameters, the frustration $[\Phi_3]^J$ is written as

$$[\Phi_3]^J_{RSB} = (\beta_S N)^3 \left\{ \lambda_1^3 + (m-1) \lambda_2^3 + \left( \frac{n}{m} - 1 \right) \lambda_3^3 + (m-1) \left( \frac{n}{m} - 1 \right) \lambda_4^3 \right\} \quad (A.3)$$

with these eigen values

$$\lambda_1 = 1 + (m-1) q_1 + (n-m) q_0, \quad \lambda_2 = 1 - q_1, \quad \lambda_3 = 1 + (m-1) q_1 - m q_0, \quad \lambda_4 = 1 - q_1.$$

The frustration for RS solution, eq.(11) can be recovered by setting $q_1 = q_2 = q$.

The frustration parameter $[\Phi_3]^J$ of the 1 step RSB version is shown in Figure 5 for $n = 0.1$. The behavior of $[\Phi_3]^J_{RSB}$ is almost the same as that of $[\Phi_3]^J_{RS}$. At low $T_S$ with $n = T_S/T_J$ kept, the overlap parameters takes the value $q_1 \simeq q_0$ and actually $[\Phi_3]^J_{RS}$ is recovered. It is hard to detect the breaking of the replica symmetry through the behavior of the frustration.

References

[1] Saito S, Sasai M and Yomo T 1997 Proc. Natl. Acad. Sci. USA 94 11324
[2] Ancel L W and Fontana W 2000 J. Exp. Zool. (Mol. Dev. Evol.) 288 242
[3] Alon U, Surette M G, Barkai N and Leibler S 1999 Nature 397 168
[4] Go N 1983 Ann. Rev. Biophys. Bioeng. 12 183
[5] Bryngelson J D and Wolynes P G 1987 Proc. Natl. Acad. Sci. USA 84 84
[6] Onuchic J N and Wolynes P G 2004 Curr. Opin. Struc. Biol. 14 70
[7] Li F, Long T, Lu Y, Ouyang Q and Tang C 2004 Proc. Natl. Acad. Sci. USA 101 4781
[8] Kaneko K, Sato K, Michiue T, Okabayashi K, Ohnuma K, Danno H and Assashima M 2008 J. Exp. Zool. B 310 492
[9] Ciliberti S, Martin O C and Wagner A 2007 PLoS Comput. Biol. 3 e15
[10] Wagner A 2005 Robustness and Evolvability in Living Systems (New Jersey: Princeton University Press)
[11] Sun J and Deem M W 2007 Phys. Rev. Lett. 99 228107
[12] Kaneko K 2006 Life: An Introduction to Complex Systems Biology (Berlin, New York: Springer-Verlag)
[13] Kaneko K 2007 PLoS ONE 2 e434
[14] Kaneko K 2008 Chaos 18 026112
[15] Kaneko K and Furusawa C 2006 J. Theor. Biol. 240 78
[16] Waddington C H 1957 The Strategy of the Genes (Bristol: George Allen & Unwin LTD)
[17] Sakata A, Hukushima K and Kaneko K 2009 Phys. Rev. Lett. 102 148101
[18] Sakata A, Hukushima K and Kaneko K 2009 Preprint arXiv: 0906.0900 (to be published in Phys. Rev. E )
[19] Penney R W, Coolen A C C and Sherrington D 1993 J. Phys. A: Math. Gen. 26 3681
[20] Dotsenko V, Frantz S and Mézard M 1994 J. Phys. A: Math. Gen. 27 2351
[21] Nishimori H 2001 Statistical Physics of Spin Glasses and Information Processing: An Introduction (New York: Oxford University Press)
[22] Hukushima K and Nemoto K 1996 J. Phys. Soc. Japan 65 1604
[23] Mézard M, Parisi G and Virasoro M A 1987 Spin Glass Theory and Beyond (Singapore: World Scientific Publisher)
[24] Mattis D C 1976 Phys. Rev. 56 421
[25] Rabello S, Coolen A C C, Perez-Vicente C J and Fraternali F 2008 J. Phys. A: Math. Theor. 41 285004
[26] Behringer H, Degenhard A and Schmid F 2006 Phys. Rev. Lett. 97 128101