Mathematical Models and Methods in Applied Sciences
Vol. 8, No. 2 (1998) 327–358

A STOCHASTIC MODEL FOR THE COOPERATIVE RELAXATION OF PROTEINS, BASED ON A HIERARCHY OF INTERACTIONS BETWEEN AMINO ACIDIC RESIDUES

M. ABUNDO*, L. ACCARDI and L. STELLA
Centro V. Volterra, Dipartimento di Matematica and Dipartimento di Scienze e Tecnologie Chimiche, Università Tor Vergata, 00133 Roma, Italy

N. ROSATO
Centro V. Volterra, Dipartimento di Medicina Sperimentale e Scienze Biochimiche, Università Tor Vergata, 00133 Roma, Italy

Communicated by P. Auger
Received 7 February 1996
Revised 4 March 1997

In this paper, a stochastic model for the cooperative relaxation of proteins, based on a hierarchic structure of the interactions between amino acids is proposed. It relies on the arbitrary splitting of interactions into two classes, *strong* and *weak*, and tests the preponderance of one class over the other. The presented model generalizes a first one valid for homogeneous interactions in the protein molecules previously studied by the authors. The time evolution of the system is studied as a function of five parameters, three of which are related to the cooperativity. Moreover, different approximations of the discrete system to a diffusion process, and to a Poisson process are considered, according to the magnitude of the parameters. A method for estimating the parameters from real data is proposed. Finally, numerical simulations and a comparison with the molecular dynamics of a real protein (Barnase) are reported.

1. Introduction

Proteins are essential elements of living organisms: they catalyze all the biochemical reactions that constitute life and perform several other functions. Their chemical structure is a chain of monomers (amino acids) covalently bonded. To be biologically active (e.g. as a catalyst) this chain must fold in a specific globular shape. The three-dimensional structure is determined by the sequence of different amino acids along the chain, but is realized through a complex network of non-covalent interactions (electrostatic, van der Waals, etc.) between monomers that are close in space, but that can be distant in the sequence.

*E-mail: abundo@mat.utovrm.it

The problem of protein folding is extremely complex for two reasons: (a) the protein chain is highly flexible, and has an astronomically high number of possible conformations; (b) the non-covalent interactions are cooperative, in the sense that the energy of interaction between the monomers depends on all the other interactions.

For these reasons the prevision of the three-dimensional structure of a protein, starting from its amino acidic sequence, is at present impossible and is one of the most ambitious goals of contemporary molecular biophysics. Solving this puzzle would give us the ability to understand the functional mechanism of a newly discovered protein just by a simple chemical analysis and, more importantly, would make protein engineering easy and reliable, allowing the improvement of the activity and stability of proteins, the cure of genetic diseases, etc. just by changing a few amino acids along the sequence.

More generally, a change in the three-dimensional structure of a protein can happen for several reasons (binding of a ligand, change of external macroscopic parameters, etc.). The energy relaxation process that leads from the starting to the final equilibrium conformation is a phenomenon similar to protein folding, since it involves a rearrangement of the nonbonded interactions between amino acids.

In the previous papers,[1,2] we have proposed a model of these processes of protein structure relaxation, based on the following simplifying assumptions:

(1) amino acids are considered as a whole unit, neglecting the details of their atomic composition;
(2) all amino acids are equivalent (their chemical structure is disregarded);
(3) every amino acid can interact with all the other monomers, regardless of their position along the chain;
(4) the interaction between two monomers is modelled as a two-state process. Two amino acids are either *interacting* or not; (therefore, we can speak of *bonds* even if we are treating non-covalent interactions);
(5) there is only one type of interaction;
(6) the formation or *breaking* of interactions between amino acids is a stochastic process;
(7) the probability of *forming* an additional interaction increases linearly with the number of interactions already *formed* (in this sense we speak of a cooperative phenomenon).

Assumption (3) derives from the flexibility of the protein chain and its ability to assume an enormous number of conformations. Assumption (6) is a result of the continuous thermal fluctuation of protein structure. Assumption (7) is a simple representation of the cooperativity described above. All the other hypotheses are just for the sake of simplicity.

The aim of such a crude model is to give a simple estimate of the degree of cooperativity involved in the folding of a specific protein, in order to compare different molecules and investigate possible correlations of their cooperativity with other characteristics (shape, flexibility, function, etc.). Applicability of the model

needs the knowledge of the behavior, during relaxation, of an experimental or simulated parameter that can be taken as a measure of the degree of *active* interactions in the protein (e.g. the potential energy). The simplified time behavior of the model can be used to fit the data, obtaining in this way a value for the two parameters defined in Ref. 1. One of them can be assumed as a measure of the *cooperativity* of that specific protein. When this was done for a molecular dynamics simulation of the denaturation of a protein (BPTI), a significant value of cooperativity was detected.[1]

In this paper we try to improve our model by removing simplification (5). In fact, the various interactions that contribute to the folding of a protein have different physical causes, different potential energies and different radii of action.

For the sake of simplicity we consider here only two types of interactions, *strong* and *weak*, with different extents of potential energy difference between separate and coupled amino acids. We assume that the cooperative relaxation of proteins is based on a *hierarchy* of interactions between amino acids, with a predominant influence of the *strong* interactions on the *weak* ones. Once a *strong* interaction forms, several amino acids get closer and then can interact *weakly* more easily. If just a *weak* interaction is formed, non-appreciable difference for the formation of a *strong* one is produced. This is described by a generalization of the model proposed in Ref. 1, which introduces a hierarchy of interactions: the *strong* ones (denoted by "s") that can influence the *weak* ones (denoted by "w"), but not vice versa. In a practical example, we will try to identify these two types of interactions of the model with electrostatic and van der Waals forces, respectively. While the first one is long ranged (due to the $1/r$ distance dependence of its potential), and can bring together amino acids that are far apart, the second one is weaker and short ranged (the attractive part of the potential has a $1/r^6$ distance dependence) and has only a local influence.

The present model still gives a simple description of the cooperativity of the relaxation of a protein, with only five parameters. Two of them (p_s and p_w, see Sec. 2 for their definition) are related to the average probability of forming an interaction (*strong* and *weak*, respectively), which, in turn, is a function of the energy involved in the interaction. The other parameters ($\Delta p_s, \Delta p_w, \Delta p_{sw}$, see Sec. 2 for definitions) express, respectively, the degree of cooperativity between the interactions of s-type, of w-type, and the influence of the number of s-type interactions on the formation of w-type ones.

In Sec. 2 we show how to describe the hierarchic model by a two-dimensional Markov chain (MC), in an analogous way as done in Ref. 1, where a one-dimensional MC was constructed for the model of cooperative interactions in proteins without hierarchic structure.

Section 3 is devoted to the study of the properties of the Markov chain. Moreover, generalizing the result obtained in Ref. 3, concerning the diffusion approximation of the one-dimensional MC for cooperative interactions without hierarchy, we obtain, in some range of values of the parameters, the continuous

approximation of the two-dimensional MC to a diffusion process in $[0,1]^2$, as the number of allowed interactions goes to infinity. In another range of values of the parameters, we obtain an approximation of the MC to a Poisson process.

In the range of values of the parameters where the diffusion approximation holds, we get an estimate for the stationary probabilities and also information about the attainability of boundary states.

In Sec. 4 we show how the parameters of the model can be estimated from data representing the time dependence of the number of *active strong* and *weak* interactions, using the continuous approximation introduced in Sec. 3.

In Sec. 5 we report the results of the computer simulation and we discuss the applicability of our model to a real situation. We compare our model to the data relative to a molecular dynamics simulation of the thermal denaturation of the protein Barnase, appeared in Ref. 5, and from these data we estimate the parameters of the model.

The results of this comparison are reported at the end of Sec. 5 and they show that our model, even with its crude simplifications, is satisfactorily applicable in the sense that it captures the qualitative behavior of cooperativity in the presence of a hierarchy of bonds.

2. The Model

Let there be a given number of particles (amino acidic residues) each of which can interact with one another by forming a bond. Suppose that two types of bonds are available: the *strong* ones s and the *weak* ones w. Denote N_{tot} the total number of permitted pairings among particles, we divide these pairings into two groups. The pairings linked by a bond of s-type belong to the first group; in the second group there are the pairings linked by w-type bonds. Let us suppose that the number of bonding of the first group is at most N_s, while that of the second group is at most N_w; then $N_s + N_w = N_{\text{tot}}$. The model described below is a generalization of the one studied in Ref. 1, where the interactions between residues were supposed to be all of the same type.

A configuration of the system is defined by a sequence of two-dimensional random variables (each component being a binary variable): $(\xi_k^{(j)}, \eta_h^{(j)})_{k=1,\dots,N_s;h=1,\dots,N_w}$ such that $\xi_k^{(j)} = +1$ if the kth *strong* bond is activated at the discrete time j, $\xi_k^{(j)} = -1$, otherwise. Analogously, $\eta_h^{(j)} = +1$ if the hth *weak* bond is activated at the discrete time j, $\eta_h^{(j)} = -1$, otherwise.

The evolution of the system, starting from an initial configuration $(\xi^{(0)}, \eta^{(0)})$ is described in terms of the *subtotal "energies"* of the system at time j:

$$S_{N_s}(\xi^{(j)}) = \sum_{k=1}^{N_s} \xi_k^{(j)}, \quad S_{N_w}(\eta^{(j)}) = \sum_{h=1}^{N_w} \eta_h^{(j)} \tag{2.1}$$

the total energy of the system at time j being

$$S_N(\xi^{(j)},\eta^{(j)}) = S_{N_s}(\xi^{(j)}) + S_{N_w}(\eta^{(j)})\,.$$

Now, we introduce five parameters $p_s, p_w, \Delta p_s, \Delta p_w, \Delta p_{sw}$; our main assumption is that the r.v. $\xi_1^{(j)},\dots,\xi_{N_s}^{(j)}$ and $\eta_1^{(j)},\dots,\eta_{N_w}^{(j)}$ are independent and $\xi_k^{(j+1)}$, $k = 1,\dots,N_s$, $\eta_h^{(j+1)}$, $h = 1,\dots,N_w$ are Bernoullian with distribution:

$$Pr\left(\xi_k^{(j+1)} = +1 | S_N(\xi^{(j)},\eta^{(j)})\right) = p_s + \frac{\Delta p_s}{N_s} S_{N_s}(\xi^{(j)})\,, \tag{2.2}$$

$$Pr\left(\eta_h^{(j+1)} = +1 | S_N(\xi^{(j)},\eta^{(j)})\right) = (p_w + \Delta p_{sw}) + \frac{\Delta p_w}{N_w} S_{N_w}(\eta^{(j)}) + \frac{\Delta p_{sw}}{N_s} S_{N_s}(\xi^{(j)})\,. \tag{2.3}$$

The equality (2.2) implies that the probability to form a bond of s-type at time $j+1$, depends only on the *relative* number of bonds of s-type already present at time j, through the coefficients $p_s, \Delta p_s$. Then, if Δp_s is *large* we have a great probability to create bonds of s-type at the new instant of time. Moreover, for fixed p_s and Δp_s, the greater the energy S_{N_s}, at time j, the greater the probability that bonds of s-type may be formed at time $j+1$. On the contrary, (2.3) means that the probability of forming a bond of w-type at time $j+1$, depends both on the *relative* number of bonds of w-type and the *relative* number of bonds of s-type, already present at time j, through the coefficients $p_w, \Delta p_w, \Delta p_{sw}$. Moreover, for a fixed situation of the w-type bonds, the greater the energy S_{N_s} of s-type bonds at time j, the greater the probability that bonds of w-type may be formed at time $j+1$, i.e. the presence of *strong* bonds among particles makes the formation of *weak* ones easier.

The parameters p_s and p_w may be called *the mean, or ground probabilities* for bonds of s-type and w-type, respectively; in fact, they represent the probability of forming a bond of s-type (respectively w-type) at zero energy of the *strong* bonds (resp. at zero energy of the *weak* bonds and with $\Delta p_{sw} = 0$). Physically, p_s and p_w are decreasing functions of the bonding energy for the formation of a bond of s-type and w-type, respectively.

We call Δp_s and Δp_w the *s-coupling capacity* and *w-coupling capacity,* respectively. Δp_s represents the maximum allowed increment of the probability to form a bond of s-type for given p_s; Δp_w represents the maximum allowed increment of the probability to form a bond of w-type, for given p_w, under the condition $S_{N_s} = 0$, and $\Delta p_{sw} = 0$. Finally, Δp_{sw} represents the *cross-coupling capacity.* When $\Delta p_{sw} = 0$, the processes $\{\xi_k^{(j)}\}$ and $\{\eta_k^{(j)}\}$ are independent and the number of s-type bonds does not influence the formation of w-type bonds. When $\Delta p_{sw} \neq 0$, the two processes are correlated and an increase of Δp_{sw} results in a more effective influence of the number of s-type bonds on the formation of w-type bonds.

Modeling the protein system in this way renders well enough the picture of *auto* and *cross-cooperativity* among different types of interactions (bonds) between amino acidic residues. In the following, in order to make (2.2), (2.3) well-defined we shall impose the following consistency conditions:

$$\begin{cases} p_s \geq \Delta p_s \geq 0\,, \quad p_s + \Delta p_s \leq 1\,, \\ p_w \geq \Delta p_w \geq 0\,, \quad p_w + \Delta p_w \leq 1\,, \\ 0 \leq p_w + \Delta p_w + 2\Delta p_{sw} \leq 1\,, \quad -\dfrac{1}{2} \leq \Delta p_{sw} \leq \dfrac{1}{2}\,. \end{cases} \tag{2.4}$$

Proceeding in analogy to Ref. 1, by a suitable change of variables, we look at the two-dimensional process $(X_n, Y_n)\,, n = 1, 2, \ldots$ whose components represent the actual number of bonds of s-type and w-type (at discrete time n). Conditions (2.2), (2.3) imply that the events $\{X_{n+1} = x\}$ and $\{Y_{n+1} = y\}$ are conditionally independent given the event $\{X_n = x', Y_n = y'\}$, i.e.:

$$\begin{aligned} &Pr(X_{n+1} = x, Y_{n+1} = y | X_n = x', Y_n = y') \\ &\quad = Pr(X_{n+1} = x | X_n = x', Y_n = y') \cdot Pr(Y_{n+1} = y | X_n = x', Y_n = y')\,. \end{aligned} \tag{2.5}$$

Moreover, we suppose that:

$$Pr(X_{n+1} = x | X_n = x', Y_n = y') = Pr(X_{n+1} = x | X_n = x')\,, \tag{2.6}$$

i.e. the event $\{X_{n+1} = x'\}$ is independent of $\{Y_n = y'\}$. All that leads to the construction of a two-dimensional Markov chain (X_n, Y_n) with state space $\Omega = \{(i_s, i_w) : i_s \in \{0, 1, \ldots, N_s\}, i_w \in \{0, 1, \ldots, N_w\}\}$ and transition probabilities given by[1]:

$$\begin{aligned} P_{(i_s,i_w)(j_s,j_w)} &= Pr\Big((X_{n+1} = j_s, Y_{n+1} = j_w) | (X_n = i_s, Y_n = j_w)\Big) \\ &= \binom{N_s}{j_s} \left(p_s - \Delta p_s + 2\Delta p_s \frac{i_s}{N_s}\right)^{j_s} \left(1 - p_s + \Delta p_s - 2\Delta p_s \frac{i_s}{N_s}\right)^{N_s - j_s} \\ &\quad \times \binom{N_w}{j_w} \left(p_w - \Delta p_w + 2\Delta p_w \frac{i_w}{N_w} + 2\Delta p_{sw} \frac{i_s}{N_s}\right)^{j_w} \\ &\quad \times \left(1 - p_w + \Delta p_w - 2\Delta p_w \frac{i_w}{N_w} - 2\Delta p_{sw} \frac{i_s}{N_s}\right)^{N_w - j_w}\,. \end{aligned} \tag{2.7}$$

It is easy to verify that $\sum_{(j_s,j_w)} P_{(i_s,i_w)(j_s,j_w)} = 1$, so that the quantities in (2.7) are indeed transition probabilities.

3. The Properties of the Markov Chain

3.1. *Absorption problems*

(i) *The case when* $p_s = \Delta p_s \neq 1/2$ *and* $p_w = \Delta p_w \neq 1/2$:
From (2.7) we obtain that $P_{(0,0)(0,0)} = 1$ and $P_{(0,0)(j_s,j_w)} = 0\,,\ \forall\ (j_s, j_w) \neq (0, 0)$. Therefore the state $(0, 0)$ is absorbing. In terms of the variable ξ_i, η_i, the state $(0, 0)$ corresponds to the subtotal energies $S_{N_s}(\xi) = -N_s$, $S_{N_w}(\eta) = -N_w$, that is to the case of *full uncoupling* for both types of bonds.

(ii) *The case when* $p_s = \Delta p_s \neq 1/2$ *and* $p_w + \Delta p_w = 1$:
From (2.7), we obtain that $P_{(0,N_w)(0,N_w)} = 1$ and $P_{(0,N_w)(j_s,j_w)} = 0 \; \forall \; (j_s, j_w) \neq (0, N_w)$. The state $(0, N_w)$ is absorbing; it corresponds to a situation of *full coupling* for w-type bonds and *full uncoupling* for s-type bonds.

(iii) *The case when* $p_s + \Delta p_s = 1$ *and* $\Delta p_w = p_w + 2\Delta p_{sw}$:
This is compatible with conditions (2.4) only if $\Delta p_{sw} \leq 0$. From (2.7), we obtain that $P_{(N_s,0)(N_s,0)} = 1$ and $P_{(N_s,0)(j_s,j_w)} = 0 \; \forall \; (j_s, j_w) \neq (N_s, 0)$. The state $(N_s, 0)$ is absorbing; it corresponds to a situation of *full coupling* for bonds of s-type and *full uncoupling* for w-type bonds. Thus, the bonds of s-type do not *cooperate* to increase the number of bonds of w-type, but vice versa. In fact, the *cross-coupling capacity* is negative. The case (iii) is perhaps only a mathematical feature of the model.

(iv) *The case when* $p_s + \Delta p_s = 1$ *and* $p_w + \Delta p_w + 2\Delta p_{sw} = 1$:
From (2.7), we obtain that $P_{(N_s,N_w)(N_s,N_w)} = 1$ and $P_{(N_s,N_w)(j_s,j_w)} = 0 \; \forall \; (j_s, j_w) \neq (N_s, N_w)$. The state (N_s, N_w) is absorbing; it corresponds to a situation of *full coupling* for both types of bonds.

(v) *The case when* $p_s > \Delta p_s > 0$, $p_s + \Delta p_s < 1$, $p_w > \Delta p_w > 0$, $p_w + \Delta p_w < 1$, $p_w + \Delta p_w + 2\Delta p_{sw} < 1$:
It is the most interesting case from the physical point of view. As easily seen, $P_{(i_s,i_w)(j_s,j_w)} > 0 \; \forall \; (i_s, i_w), (j_s, j_w)$ which implies that the chain is irreducible. Indeed, the second and fourth factors in (2.7) are evidently positive, as well the third and the fifth ones, because

$$\text{(a)} \; 1 - p_s + \Delta p_s - 2\Delta p_s \frac{i_s}{N_s} \geq 1 - p_s + \Delta p_s - 2\Delta p_s = 1 - p_s - \Delta p_s > 0 \,,$$

$$\text{(b)} \; 1 - p_w + \Delta p_w - 2\Delta p_w \frac{i_w}{N_w} - 2\Delta p_{sw} \frac{i_s}{N_s} \geq 1 - p_w + \Delta p_w - 2\Delta p_w - 2\Delta p_{sw}$$
$$= 1 - p_w - \Delta p_w - 2\Delta p_{sw} > 0 \,.$$

3.2. *The stationary probabilities*

In the case (v) above, since the MC is irreducible, the stationary probabilities $\Pi_{\mathbf{i}} = \Pi_{(i_s,i_w)}, i_s = 0, 1, \ldots, N_s, {}_1w = 0, 1, \ldots, N_w$ exist such that $\Pi_{\mathbf{i}} = \lim_{n\to\infty} P^{(n)}_{\mathbf{ki}}$, irrespective of the initial state $\mathbf{k} = (k_s, k_w)$, where $P^{(n)}_{\mathbf{ki}}$ is the probability of transition from the state $\mathbf{k} = (k_s, k_w)$ into the state $\mathbf{i} = (i_s, i_w)$ in n steps. Moreover, $\lim_{n\to\infty} Pr((X_n, Y_n) = \mathbf{i}) = \Pi_{\mathbf{i}}$.

The stationary probability $\Pi_{\mathbf{i}}$ represents the probability that the system is in the state $\mathbf{i}$ after an infinite time, irrespective of the initial state. Thus, the knowledge of the stationary probabilities gives information about the fluctuations of the system at the equilibrium.

The stationary probabilities $\Pi_{\mathbf{i}} = \Pi_{(i_s,i_w)}$ can be computed by solving the equation:

$$\Pi_{\mathbf{i}} = \sum_{\mathbf{j}} \Pi_{\mathbf{j}} P_{\mathbf{ji}} \,. \tag{3.1}$$

This equation is nothing but the equation which yields the left eigenvector of the matrix $\mathbf{P}$ with eigenvalue 1. In practice, since the index $\mathbf{i}$ is two-dimensional, when N_s, N_w are large, Eq. (3.1) becomes computationally intractable. The stationary probabilities $\Pi_{\mathbf{i}}$ and the *marginal* stationary probabilities $\pi_{i_s} = \sum_{i_w=0}^{N_w} \Pi_{(i_s,i_w)}$ and $\pi_{i_w} = \sum_{i_s=0}^{N_s} \Pi_{(i_s,i_w)}$ can be numerically found by computing the frequencies of visits to each state $i_s \in \{0,1,\ldots,N_s\}$, $j_s \in \{0,1,\ldots,N_w\}$ following a long enough trajectory of the system.

In Sec. 5, we shall present some numerical simulations and estimations of stationary probabilities. Since Eq. (3.1) is theoretically intractable, to find an analytical estimate of stationary probabilities for N_s, N_w large, we can proceed as was done in Ref. 1 for the analogous one-dimensional Markov chain. Indeed, a heuristic application of the law of large numbers (cf. also Ref. 8) suggests that the following limits hold in probability:

$$\frac{1}{N_s} \cdot \lim_{j\to\infty} X_j \to^P a_s = \frac{p_s - \Delta p_s}{1 - 2\Delta p_s}, \text{as } N_s \to \infty\,, \tag{3.2}$$

$$\frac{1}{N_w} \cdot \lim_{j\to\infty} Y_j \to^P a_w = \frac{(1-2\Delta p_s)(p_w - \Delta p_w) + 2\Delta p_{sw}(p_s - \Delta p_s)}{(1-2\Delta p_s)(1-2\Delta p_w)}, \text{ as } N_w \to \infty\,. \tag{3.3}$$

Then, we obtain the following approximation in probability of the stationary probabilities, for N_s, N_w large:

$$\begin{aligned}\Pi_{(n_s,n_w)} &= \lim_{k\to\infty} Pr(X_k = n_s, Y_k = n_w) \\ &\sim \binom{N_s}{n_s} a_s^{n_s}(1-a_s)^{N_s-n_s} \binom{N_w}{n_w} a_w^{n_w}(1-a_w)^{N_w-n_w}\,. \end{aligned} \tag{3.4}$$

However, this is a rather rough estimate, since in the limit $N_s, N_w \to \infty$, the processes X_k, Y_k appear to be independent and binomial distributed with parameters a_s and a_w, while they are known to be correlated.

3.3. *The diffusion approximation (for p_i, Δp_i close to 1/2 and Δp_{sw} close to zero)*

Another way to obtain information about the stationary probabilities is to consider the continuous approximation of the two-dimensional MC with transition probabilities matrix (2.7), to a diffusion process in $[0,1]^2$, in analogous way, as was done in Ref. 3, where in the case of protein interaction without hierarchic structure, we have found a continuous approximation of the MC to a diffusion process in $[0,1]$. In fact, once the stationary distribution of the diffusion process is known, by discretization, the stationary probabilities can be approximately obtained.

We observe that the diffusion approximation also represents a convenient tool for fitting experimental or simulated data with our model; indeed this approach will be followed in Secs. 4 and 5 in estimating the parameters.

Now we state the diffusion approximation in the case of a hierarchic structure between bonds.

Let us introduce some notations. Set $N_1 = N_s = N, N_2 = N_w$, $\mathbf{N} = (N_1, N_2), i_1 = i_s$, $i_2 = i_w$, and $K_N = K_{N_1,N_2} = \{(u_1, u_2) = (i_1/N_1, i_2/N_2), i_1 \in \{0, 1, \ldots, N_1\}, i_2 \in \{0, 1, \ldots, N_2\}\}$ and let

$$\chi^{\mathbf{N}}(t) = (\chi_1^{\mathbf{N}}(t), \chi_2^{\mathbf{N}}(t)) = \left(\frac{X^{N_1}}{N_1}([Nt]), \frac{Y^{N_2}}{N_2}([Nt])\right) \tag{3.5}$$

be the two-dimensional rescaled process with values in $K_{\mathbf{N}}$, where $[z]$ denotes the integer part of z, i.e. the largest integer $\leq z$.

Here, $X^{N_1}(k), Y^{N_2}(k')$ are indeed the same processes considered in (2.5), (2.7) where, for simplicity, the superscripts N_1, N_2 are omitted.

We are interested in the behavior of the process $\chi^{\mathbf{N}}(t)$ in the limit $(N_1, N_2) \to \infty$ in such a way that $\lim_{N_1 \to \infty} N_2/N_1 = m$, for some $m > 0$. The following theorem holds:

Theorem 3.1. *Given* $\mathbf{N} = (N_1, N_2)$, *we can consider the MC with transition probabilities* (2.7), *depending on the parameters* $p_k = p_k(N_k), \Delta p_k = \Delta p_k(N_k)$, $k = 1, 2$ *and* $\Delta p_{12} = \Delta p_{12}(\mathbf{N})$. *Let us assume that* $\lim_{N_1 \to \infty} N_2/N_1 = m$, *for some* $m > 0$ *and suppose that the following limits exist*:

$$\begin{cases} \beta_k = \lim_{N_k \to \infty} N_k(p_k - \Delta p_k), \\ \alpha_k = \lim_{N_k \to \infty} \frac{1}{2} N_k(1 - 2\Delta p_k), \quad k = 1, 2 \\ \gamma = \lim_{N_2 \to \infty} 2N_2 \Delta p_{12}. \end{cases} \tag{3.6}$$

Then, the process $\chi^{\mathbf{N}}(t)$ *with values in* $K_{\mathbf{N}}$, *defined by* (3.5), *converges as* $N(= N_1)$ *goes to infinity to the two-dimensional diffusion process* $\chi(t)$ *on* $[0, 1]^2$, *whose associated backward differential equation has the form*:

$$\begin{aligned} \frac{\partial q}{\partial t}(x_1, x_2, t) &= \sum_{i=1}^{2} b_i(x_1, x_2) \frac{\partial q}{\partial x_i}(x_1, x_2, t) + \frac{1}{2} \sum_{i,j}^{2} a_{ij}(x_1, x_2) \frac{\partial^2 q}{\partial x_i \partial x_j}(x_1, x_2, t) \\ &= (Lq)(x_1, x_2, t), \ (x_1, x_2, t) \in [0, 1]^2 \times R^+, \end{aligned} \tag{3.7}$$

where

$$\begin{cases} b_1(x_1, x_2) = b_1(x_1) = \beta_1 - 2\alpha_1 x_1, \\ b_2(x_1, x_2) = \frac{\beta_2}{m} + \frac{\gamma}{m} x_1 - 2\frac{\alpha_2}{m} x_2, \\ a_{ij}(x_1, x_2) = 0, i \neq j, \\ a_{11}(x_1, x_2) = x_1(1 - x_1), \\ a_{22}(x_1, x_2) = \frac{1}{m} x_2(1 - x_2), \end{cases} \tag{3.8}$$

with conditions

$$\begin{cases} 0 < \beta_1 < 2\alpha_1 \\ \beta_2 + \gamma < 2\alpha_2\,,\ \gamma \geq 0\,. \end{cases} \tag{3.8'}$$

Roughly speaking, the above theorem says that, if we take:

$$\begin{cases} p_k = p_k(N_k) = \dfrac{1}{2} + \dfrac{\beta_k - \alpha_k}{N_k} + o\left(\dfrac{1}{N_k}\right) \\ \Delta p_k = \Delta p_k(N_k) = \dfrac{1}{2} - \dfrac{\alpha_k}{N_k} + o\left(\dfrac{1}{N_k}\right), \quad k = 1,2 \\ \Delta p_{12} = \dfrac{\gamma}{2N_2} + o\left(\dfrac{1}{N}\right) \end{cases} \tag{3.9}$$

for N large, the rescaled MC approximates a continuous diffusion process on $[0,1]^2$.

Before giving the proof of Theorem 3.1, we need to introduce further notations and some lemmas whose proofs will be given later. Indeed, the proofs of Theorem 3.1 and lemmas are very similar to those given in Ref. 3 for the diffusion approximation of the MC relative to the model without hierarchy.

The rescaled two-dimensional MC $\chi^{\mathbf{N}}$ takes on values in $K_{\mathbf{N}}$ which we rewrite as $K_{\mathbf{N}} = \{(u_1,u_2) = ([x_1N_1]/N_1, [x_2N_2]/N_2), (x_1,x_2) \in [0,1]^2\}$ ($[z]$ denotes the integer part of z); its marginal distributions conditioned to the values at time n, that is to say, conditioned to the event

$$\mathcal{G}_n = \{(1/N_1)X^{N_1}(n) = u_1, (1/N_2)Y^{N_2}(n) = u_2\}$$

are such that:

$$\begin{cases} \dfrac{1}{N_1}X^{N_1}(n+1)\Big|\mathcal{G}_n \sim \dfrac{1}{N_1}Z_{N_1}\,, \\ \dfrac{1}{N_2}Y^{N_2}(n+1)\Big|\mathcal{G}_n \sim \dfrac{1}{N_2}Z_{N_2}\,, \end{cases} \tag{3.10}$$

where $Z_{N_1} \sim B(N_1, p_1 - \Delta p_1 + 2\Delta p_1 u_1)$ and $Z_{N_2} \sim B(N_2, p_2 - \Delta p_2 + 2\Delta p_2 u_2 + 2\Delta p_{12} u_1)$, $p_\cdot = p_\cdot(N_\cdot), \Delta p_\cdot = \Delta p_\cdot(N_\cdot)$, where $B(n,\alpha)$ denotes the binomial distribution with parameters n and α.

Lemma 3.2. *With the above notations, positive constants c_1, c_2 exist, such that:*

$$E_{x_1,x_2}(|Z_{N_1} - u_1N_1|^4) \sim c_1x_1^2N_1^2 \tag{3.11}$$

and

$$E_{x_1,x_2}(|Z_{N_2} - u_2N_2|^4) \sim c_2x_2^2N_2^2 \tag{3.12}$$

where E_{x_1,x_2} denotes the conditional expectation given $\chi^{\mathbf{N}}(n) = (x_1,x_2)$. By the relation $U(n) \sim rn^2$, we mean $\lim_{n\to\infty} U(n)/n^2 = r$.

Lemma 3.3. *Denote $P_{\mathbf{N}}(\cdot,\cdot)$ the transition probability function of $\chi^{\mathbf{N}}$ and set:*

$$b_{i,N}(x_1,x_2) = b_{i,N}(u_1,u_2) = N\int_{\|\mathbf{v}-\mathbf{u}\|\le r}(v_i-u_i)P_{\mathbf{N}}(\mathbf{u},\mathbf{dv})\,, \tag{3.13}$$

$$a_{ij,N}(x_1,x_2) = a_{ij,N}(u_1,u_2) = N\int_{\|\mathbf{v}-\mathbf{u}\|\le r}(v_i-u_i)(v_j-u_j)P_{\mathbf{N}}(\mathbf{u},\mathbf{dv})\,, \tag{3.14}$$

where $u_1 = (1/N_1)[xN_1], u_2 = (1/N_2)[yN_2]$, $\mathbf{u} = (u_1,u_2), \mathbf{v} = (v_1,v_2) \in K_{\mathbf{N}}$.

Then $\forall\ \delta > 0, \forall\ r > 0$:

$$\text{(i)}\ \sup_{x_1^2+x_2^2\le r^2} NP_{\mathbf{N}}(\mathbf{u},\{\mathbf{v}:\|\mathbf{v}-\mathbf{u}\|>\delta\}) \longrightarrow 0,\ \textit{as}\ N\to\infty\,, \tag{3.15}$$

$$\text{(ii)}\ \sup_{x_1^2+x_2^2\le r^2} |b_{i,N}(x_1,x_2) - b_i(x_1,x_2)| \longrightarrow 0\ \textit{as}\ N\to\infty\,, \tag{3.16}$$

$$\text{(iii)}\ \sup_{x_1^2+x_2^2\le r^2} |a_{ij,N}(x_1,x_2) - a_{ij}(x_1,x_2)| \longrightarrow 0\ \textit{as}\ N\to\infty\,. \tag{3.17}$$

Proof of Theorem 3.1. The proof is quite analogous to that of Theorem 3.1 in Ref. 3. Indeed, as easily seen (cf. e.g. Ref. 9) (i), (ii), (iii) imply:

$$\lim_{N\to\infty}\sup_{\mathbf{u}\in K_{\mathbf{N}}} |N(T^{\mathbf{N}}-I)f(u_1,u_2) - Lf(u_1,u_2)| = 0\ \ \forall\ f\in C^2([0,1]^2)\,, \tag{3.18}$$

where L is the differential operator defined in (3.7), and $T^{\mathbf{N}}$ is the associated transition semigroup, i.e. $T^{\mathbf{N}}f(\mathbf{u}) = \int f(\mathbf{v})P_{\mathbf{N}}(\mathbf{u},\mathbf{dv})$. By using (3.18) and standard results,[7] Theorem 3.1 follows. □

Proof of Lemma 3.2. We omit the details of the proof, since it is exactly the same as that of Lemma 3.2 of Ref. 3; it suffices to replace N and E_x in Ref. 3 with N_1 and E_{x_1,x_2}, respectively. □

Proof of Lemma 3.3.

(i) We have:

$$\begin{aligned} &P_N(u,\{v:\|v-u\|>\delta\}) \\ &\quad\le P^1_{N_1}(u_1,\{|v_1-u_1|>\delta/\sqrt{2}\}) + P^2_{N_2}(u_2,\{|v_2-u_2|>\delta/\sqrt{2}\})\,, \end{aligned} \tag{3.19}$$

where $P^i_{N_i}(\cdot,\cdot)$ are the marginal transition probabilities. By the Chebyshev inequality:

$$\begin{aligned} NP^i_{N_i}(u_i,|v_i-u_i|>\delta/\sqrt{2}) &= NP\left(\left|\frac{1}{N_i}Z_{N_i}-u_i\right|>\delta/\sqrt{2}\right) \\ &\le N\frac{E(|Z_{N_i}/N_i-u_i|^4)}{(\delta/\sqrt{2})^4} = 4N/(\delta^4N_i^4)E(|Z_{N_i}-N_iu_i|^4)\,. \end{aligned}$$

By Lemma 3.2, the last quantity is of order $1/N$, so by using (3.19), Lemma 3.3 follows.

(ii) The estimate of $b_{i,N}(x_1, x_2)$, $i = 1, 2$ as $N \to \infty$, can be obtained exactly in the same way as was done in the proof of (ii) of Lemma 3.3 in Ref. 3, with the obvious modifications. Then, (3.16) follows easily.

(iii) Also the estimate of $a_{ij,N}(x_1, x_2), i, j = 1, 2$ as $N \to \infty$, can be obtained as in Lemma 3.3 of Ref. 3. The only thing to be remarked is that, for $i \neq j$, $a_{i,j,N}(x_1, x_2) \to 0$ as $N \to \infty$, i.e. the diffusion matrix is diagonal; this follows from the fact that the infinitesimal covariance is zero, thanks to the conditional independence of the processes $X_1(n) = X_n$ and $X_2(n) = Y_n$. Thus, (3.17) follows easily. □

To the diffusion process, $\chi(t)$ is associated to the stochastic differential equation in $[0, 1]^2$, having $b_i(x_1, x_2)$ as drift terms and a_{ij} as diffusion matrix, that is $\chi(t)$ is the solution of

$$\begin{cases} dX_1 = (\beta_1 - 2\alpha_1)\, dt + \sqrt{X_1(1 - X_1)}\, dW_1(t), \\ dX_2 = \dfrac{1}{m}(\beta_2 - 2\alpha_2 X_2 + \gamma X_1)\, dt + \sqrt{X_2(1 - X_2)}\, dW_2(t), \\ X_1(0) = \bar{x}_1,\ X_2(0) = \bar{x}_2, \end{cases} \tag{3.20}$$

where $W_1(t)$ and $W_2(t)$ are independent standard Brownian motions; the parameters are required to satisfy the following conditions:

$$\beta_i > 0,\quad i = 1, 2,\quad \gamma \geq 0,\quad 2\alpha_1 - \beta_1 > 0,\quad 2\alpha_2 - \beta_2 - \gamma > 0 \tag{3.21}$$

(3.21) are only technical conditions to assure that the process never exits from $[0, 1]^2$. By Theorem 3.1, one can obtain information about the behavior of the MC; in particular, by means of the stationary measure of the continuous diffusion process, one can study the behavior of the MC, at infinite time.

When the cross-coupling capacity Δp_{sw} is zero (this corresponding to $\gamma = 0$, in the diffusion limit), the density of the stationary distribution for the continuous diffusion process can be explicitly found[4] in the form of a product of two beta functions:

$$u_{\alpha,\beta,0}(x, y) = \text{const} \cdot x^{2\beta_1 - 1}(1 - x)^{4\alpha_1 - 2\beta_1 - 1} y^{2\beta_2 - 1}(1 - y)^{4\alpha_2 - 2\beta_2 - 1},\quad (x, y) \in [0, 1]^2 . \tag{3.22}$$

For *small*, positive values of γ, the density $u_{\alpha,\beta,\gamma}(x, y)$ of the stationary measure is close to $u_{\alpha,\beta,0}(x, y)$; by increasing γ, $u_{\alpha,\beta,\gamma}(x, y)$ gradually deviates from $u_{\alpha,\beta,0}(x, y)$ (see Sec. 5 for more details).

From the diffusion approximation, the estimate for the stationary probabilities follows, for N_1 *large*:

$$\Pi_{(i,j)} \approx \frac{1}{N_1 N_2} u_{\alpha,\beta,\gamma}\left(\frac{i}{N_1}, \frac{j}{N_2}\right) . \tag{3.23}$$

3.4. *Attainability and unattainability of the boundary states*

Here, we discuss the possibility that the boundary states are assumed by the MC, in the dependence of the parameters. Indeed, the MC takes values $(n_s, n_w) \in \{0,1,\ldots,N_s\} \times \{0,1,\ldots,N_w\}$, or, in other words, on the rectangle $\mathcal{R} = Z^2 \cap [0, N_s] \times [0, N_w]$ of the two-dimensional integer lattice. We are interested in stating conditions under which the boundary of $\mathcal{R}$ may (or may not) be reachable by the process.

Let l_0 be the portion of the boundary of $\mathcal{R}$ consisting of the states of type $(0, n_w)$ (situation of full uncoupling of *strong* bonds) and let l_1 be the portion of the boundary of $\mathcal{R}$ consisting of the states of type (N_s, n_w) (situation of full coupling of *strong* bonds).

Moreover, let $\mathcal{L}_0$ and $\mathcal{L}_1$ be, respectively, the portion of boundary of $\mathcal{R}$ consisting of the states of type $(n_s, 0)$ (full uncoupling of *weak* bonds), and of type $(n_s, 1)$ (full coupling of *weak* bonds).

We recall that, for a process $\mathbf{X}$, a boundary point $\mathbf{x}$ is called *attainable* (or accessible) if $\mathbf{X}$ reaches $\mathbf{x}$ at a finite time, with positive probability, and *unattainable* (or inaccessible) otherwise. More generally, $\mathbf{x}$ is called *attractive* if $\mathbf{X}$ reaches $\mathbf{x}$ with positive probability, after a time which can be either finite or infinite; $\mathbf{x}$ is called *repelling* (or natural), otherwise.

Then, in the range of the parameters $p_i, \Delta p_i$ close to $1/2$ and Δp_{sw} close to zero, by using the continuous diffusion approximation for $N_1 = N_s \to \infty$, sufficient conditions can be found, so that each of the above portions of the boundary of $\mathcal{R}$ may (or may not) be reached, in the approximation N_s *large*. Indeed, translating some results of Ref. 4 in terms of parameters $p_i = p_i(N_i), \Delta p_i = \Delta p_i(N_i), i = s, w$ and $\Delta p_{sw} = \Delta p_{sw}(N_s, N_w)$, in the approximation $N_i \to \infty$, we obtain the following properties.

Proposition 3.2. *Let us assume that the parameters are taken in the range where the diffusion approximation holds; then for N_s and N_w large:*

(i) *l_0 is accessible and attractive if $2N_s(p_s - \Delta p_s) < 1$, inaccessible and repelling otherwise;*

(ii) *l_1 is accessible and attractive if $2N_s(1 - p_s - \Delta p_s) < 1$, inaccessible and repelling otherwise;*

(iii) *$\mathcal{L}_0$ is accessible and attractive if $2N_w(p_w - \Delta p_w + 2\Delta p_{sw}) < 1$, and it is inaccessible and repelling if $2N_w(p_w - \Delta p_w) \geq 1$;*

(iv) *$\mathcal{L}_1$ is accessible and attractive if $2N_w(1 - p_w - \Delta p_w) < 1$, and it is inaccessible and repelling if $2N_w(1 - p_w - \Delta p_w - 2\Delta p_{sw}) \geq 1$.*

These results satisfactorily agree with the numerical simulations (see Sec. 5).

3.5. *The Poisson approximation (for p_k, Δp_k not necessarily close to 1/2)*

The diffusion limit obtained in Sec. 3.4 concerns the approximation of the MC with transition probabilities (2.7) to a continuous diffusion process in $[0,1]^2$, when

$N_k \to \infty$, in the case when $p_k, \Delta p_k$ are close to $1/2$ and Δp_{sw} is close to zero. Now, we look for an asymptotic expression of the transition probabilities for $N_w/N_s \sim m$ and N_s large, in the case when all the parameters can assume any value far from $1/2$. In this situation, the MC is approximated by a Poisson process as $N_s \to \infty$.

This approximation is suitable for low cooperativity proteins, where Δp_k are close to zero and therefore it is complementary to the diffusion approximation. In this situation, the MC is approximated, as $N_s \to \infty$, by a Poisson process.

Let us consider a sequence of MCs, where $p_k = p_k(N), \Delta p_k = \Delta p_k(N), \Delta p_{sw} = \Delta p_{sw}(N)$, $k = s, w$, and let us suppose that there exist constants λ_k, μ_k and ρ such that, as $N_s \to \infty$,

$$\begin{cases} N_k(p_k - \Delta p_k) \to \mu_k\,, \\ 2\Delta p_k \to \lambda_k\,, \\ 2m\Delta p_{sw} \to \rho\,. \end{cases}$$

Then, let us consider the first three factors in (2.7) (for simplicity we write p, Δp, i, j, N in place of p_s, Δp_s, i_s, j_s, N_s, respectively):

$$\binom{N}{j}\left(p - \Delta p + \frac{2\Delta p}{N}\,i\right)^j \left[1 - \left(p - \Delta p + \frac{2\Delta p}{N}\,i\right)\right]^{N-j}$$

$$= \frac{N(N-1)\cdots(N-j+1)}{j!}\left(p - \Delta p + \frac{2\Delta p}{N}\,i\right)^j \left[1 - \frac{N(p-\Delta p) + 2\Delta p i}{N}\right]^{N-j}.$$

Now, as $N \to \infty$

(i) $$\frac{N(N-1)\cdots(N-j+1)}{N^j} \to 1\,,$$

(ii) $$\frac{(2\Delta p i + N(p-\Delta p))^j}{\left(1 - \frac{N(p-\Delta p)+2\Delta p i}{N}\right)^j} \to (2\Delta p i + \mu_s)^j\,,$$

(iii) $$\left[1 - \frac{N(p-\Delta p) + 2\Delta p i}{N}\right]^N \to e^{-(\mu_s + 2\Delta p i)}\,.$$

Therefore

$$I_{i,j} := \binom{N}{j}\left((p-\Delta p) + \frac{2\Delta p}{N}\,i\right)^j \left[1 - \left((p-\Delta p) + \frac{2\Delta p}{N}\,i\right)\right]^{N-j}$$

$$\to \frac{1}{j!}\,(2\Delta p i + \mu_s)^j e^{-(2\Delta p i + \mu_s)}\,.$$

Since as $N_s \to \infty$, $2\Delta p \longrightarrow \lambda_s$, then

$$I_{i,j} \longrightarrow \frac{1}{j!}\,(\lambda_s i + \mu_s)^j e^{-(\lambda_s i + \mu_s)}\,.$$

Thus, for the first component of the process, we have:

$$\begin{aligned}
&\mathbf{P}\big(S_{[Nt]}(\xi^{k+1}) = j | S_{[Nt]}(\xi^k) = i\big) \\
&\quad = \binom{[Nt]}{j} \left((p-\Delta p) + \frac{2\Delta p}{[Nt]}\, i\right)^j \left[1 - \left((p-\Delta p) + \frac{2\Delta p}{[Nt]}\, i\right)\right]^{[Nt]-j} \\
&\quad \to \frac{1}{j!}\,(\lambda_s i + \mu_s t)^j e^{-(\lambda_s i + \mu_s t)}\,.
\end{aligned}$$

Let us consider a Poisson process $(X_t)_{t\geq 0}$ with density $\lambda_s i + \mu_s t$, i.e.

$$\mathbf{P}(X_t = j | X_0 = i) = \frac{1}{j!}\,(\lambda_s i + \mu_s t)^j e^{-(\lambda_s i + \mu_s t)}\,,$$

$$\mathbf{P}_{t,N,i}(A) := \mathbf{P}(A | S_{[N,t]}(\xi^k) = i)\,,$$

$$\tilde{\mathbf{P}}_{t,N,i} := \mathbf{P}_{t,N,i} \circ (S_{[Nt]}(\xi^{k+1}))\,.$$

Then, as $N = N_s \to \infty$

$$\tilde{\mathbf{P}}_{t,N,i} \to \mathbf{P} \circ X_t^{-1}$$

or equivalently $\forall\, r \in \mathbf{N}$

$$\tilde{\mathbf{E}}_{t,N,i}[(S_{[Nt]}(\xi^{k+1}))^r] \to \mathbf{E}(X_t^r)\,.$$

Turning to the asympotic expression for the transition probabilities, in an analogous way, we can estimate the other factors in (2.7). Thus, finally we obtain, as $N_s \to \infty$

$$\begin{aligned}
P_{(i_s,i_w)(j_s,j_w)} \sim &\frac{1}{j_s!}(\lambda_s i_s + \mu_s)^{j_s} e^{-(\lambda_s i_s + \mu_s)} \\
&\times \frac{1}{j_w!}(\lambda_w i_w + \mu_w + \rho i_s)^{j_w} e^{-(\lambda_w i_w + \mu_w + \rho i_s)}\,. \qquad (3.24)
\end{aligned}$$

We remark that (3.24) holds in the approximation

$$\begin{cases}
\Delta p_k \sim \dfrac{\lambda_k}{2}\,, \\[2ex]
p_k \sim \dfrac{\lambda_k}{2} + \dfrac{\mu_k}{N_k}\,, \\[2ex]
\Delta p_{sw} \sim \dfrac{\rho}{2m}\,.
\end{cases}$$

4. Parameters Estimation

In the range of parameters where the diffusion approximation of the process (X_n, Y_n) holds (see Sec. 3.3), $p_i, \Delta p_i$, $i = s, w$, Δp_{sw} can be estimated by using the maximum likelihood method. Indeed, the maximum likelihood method can be directly

applied to the MC to estimate parameters, but the point at which the likelihood function, obtained by the transition probabilities (2.7), takes its maximum cannot be analytically found. For the same reason, the Poisson approximation is not convenient to estimate parameters either. Instead, the likelihood function obtained by using the diffusion approximation is analytically tractable, thanks to its exponential form.

We recall from Sec. 3.3 that the evolution of the diffusion process is described by the following stochastic differential equation:

$$\begin{cases} dX = (\beta_1 - 2\alpha_1)\, dt + \sqrt{X(1-X)}\, dW_1(t)\,, \\ dY = \dfrac{1}{m}(\beta_2 - 2\alpha_2 Y + \gamma X)\, dt + \sqrt{Y(1-Y)}\, dW_2(t)\,, \\ X(0) = \bar{X},\ Y(0) = \bar{Y}\,, \end{cases} \tag{4.1}$$

where $W_1(t)$ and $W_2(t)$ are independent standard Brownian motions.

By discretization of Eq. (4.1), one obtains:

$$\begin{cases} x_{n+1} = x_n + (\beta_1 - 2\alpha_1 x_n)h + \sqrt{x_n(1-x_n)}\Delta W_1\,, \\ y_{n+1} = y_n + \dfrac{1}{m}(\beta_2 - 2\alpha_2 y_n + \gamma x_n)h + \sqrt{y_n(1-y_n)}\Delta W_2\,, \\ x_0 = \bar{x}, \quad y_0 = \bar{y}\,, \end{cases} \tag{4.2}$$

where x_n and y_n, $n = 0, 1, \ldots, M$ denote, respectively, the two processes evaluated at the time $t_n = nh$, and $\Delta W_i = W_i(t_{n+1}) - W_i(t_n), i = 1, 2$.

Equation (4.2) means that the two-dimensional random variable $\begin{pmatrix} X_{n+1} \\ Y_{n+1} \end{pmatrix}$, conditionally to $(X_n = x_n,\ Y_n = y_n)$ is distributed according to a bivariate Gaussian with expectation vector

$$\begin{pmatrix} x_n + (\beta_1 - 2\alpha_1 x_n)h \\ y_n + \dfrac{1}{m}(\beta_2 - 2\alpha_2 y_n + \gamma x_n)h \end{pmatrix} \tag{4.3}$$

and covariance matrix

$$\begin{pmatrix} x_n(1-x_n)h & 0 \\ 0 & y_n(1-y_n)h \end{pmatrix}. \tag{4.4}$$

From (4.4), given the finite sequence $(x_n, y_n)_{n=0,1,\ldots,M}$, we obtain the likelihood function:

$$\begin{aligned} &L(\alpha_i, \beta_i, \gamma) \\ &\quad = \prod_{n=1}^{M} \frac{1}{2\pi h\sqrt{x_n y_n (1-x_n)(1-y_n)}} \exp - \frac{[x_{n+1} - x_n - (\beta_1 - 2\alpha_1 x_n)h]^2}{2hx_n(1-x_n)} \\ &\qquad \times \exp - \frac{[y_{n+1} - y_n - \frac{1}{m}(\beta_2 - 2\alpha_2 y_n + \gamma x_n)h]^2}{2hy_n(1-y_n)} \end{aligned} \tag{4.5}$$

as well as the logarithm of the likelihood function:

$$V(\alpha_i,\beta_i,\gamma) = -M\log(2\pi h) - \frac{1}{2}\sum_{n=1}^{M}\{\log(x_n(1-x_n)) + \log(y_n(1-y_n))\}$$
$$- \frac{1}{2}\sum_n \left\{ \frac{(x_{n+1}-x_n-\beta_1 h+2\alpha_1 x_n h^2}{hx_n(1-x_n)} \right.$$
$$\left. + \frac{(y_{n+1}-y_n-\beta_2 h+2\alpha_{2y_n}h-\gamma x_n h)^2}{hy_n(1-y_n)} \right\}. \tag{4.6}$$

The maximum likelihood estimates $\hat{\alpha}_i,\hat{\beta}_i,\hat{\gamma}$ of the parameters $\alpha_i,\beta_i, i=1,2$ and γ are obtained by setting to zero the partial derivatives of the function $V(\cdot)$ with respect to its arguments. In this way, we obtain the system:

$$\begin{cases} \dfrac{\partial V}{\partial \alpha_1} = -2\displaystyle\sum_{n=1}^{M} \frac{x_{n+1}-x_n-\beta_1 h+2\alpha_1 x_n h}{1-x_n} = 0, \\ \dfrac{\partial V}{\partial \beta_1} = \displaystyle\sum_{n=1}^{M} \frac{x_{n+1}-x_n-\beta_1 h+2\alpha_1 x_n h}{x_n(1-x_n)} = 0, \\ \dfrac{\partial V}{\partial \alpha_2} = -2\displaystyle\sum_{n=1}^{M} \frac{y_{n+1}-y_n-\beta_2 h+2\alpha_2 y_n h-\gamma x_n h}{1-y_n} = 0, \\ \dfrac{\partial V}{\partial \beta_2} = \displaystyle\sum_{n=1}^{M} \frac{y_{n+1}-y_n-\beta_2 h+2\alpha_2 y_n h-\gamma x_n h}{y_n(1-y_n)} = 0, \\ \dfrac{\partial V}{\partial \gamma} = \displaystyle\sum_{n=1}^{M} \frac{(y_{n+1}-y_n-\beta_2 h+2\alpha_2 y_n h-\gamma x_n h)x_n}{y_n(1-y_n)} = 0. \end{cases}$$

A very long, but straightforward calculation shows that the solution of the above system is given by:

$$\begin{cases} \hat{\alpha}_1 = \dfrac{1}{2h}\dfrac{CD-EA}{A^2-CB}, \\ \hat{\beta}_1 = \dfrac{1}{hA}\left(D + \dfrac{B(CD-EA)}{A^2-CB}\right), \\ \hat{\alpha}_2 = \dfrac{1}{2h\Delta}\left(-E'A'\Delta_3 - D'\Delta_2^2 - E''C'\Delta_1 + E'\Delta_1\Delta_2 + E''\Delta_2 A' + \Delta_3 C'D'\right), \\ \hat{\beta}_2 = -\dfrac{1}{h\Delta}\left(-D'A'\Delta_3 - E''B'\Delta_2 - E'\Delta_1^2 + E''A'\Delta_1 + D'\Delta_1\Delta_2 + E'B'\Delta_3\right), \\ \hat{\gamma} = -\dfrac{1}{h\Delta}\left(-A'^2E'' - E'B'\Delta_2 - D'C'\Delta_1 + D'A'\Delta_2 + E'\Delta_1 A' + E''B'C'\right), \end{cases} \tag{4.7}$$

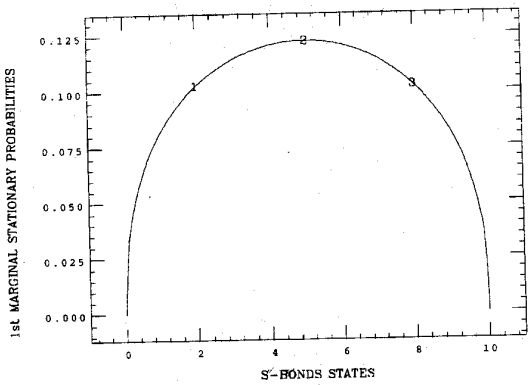
where

$$
\begin{cases}
A = \sum_n \dfrac{1}{1-x_n}\,, \\
B = \sum_n \dfrac{x_n}{1-x_n}\,, \\
C = \sum_n \dfrac{1}{x_n(1-x_n)}\,, \\
D = \sum_n \dfrac{x_{n+1}-x_n}{1-x_n}\,, \\
E = \sum_n \dfrac{x_{n+1}-x_n}{x_n(1-x_n)}\,, \\
A' = \sum_n \dfrac{1}{1-y_n}\,, \\
B' = \sum_n \dfrac{y_n}{1-y_n}\,, \\
C' = \sum_n \dfrac{1}{y_n(1-y_n)}\,, \\
D' = \sum_n \dfrac{y_{n+1}-y_n}{1-y_n}\,, \\
E' = \sum_n \dfrac{y_{n+1}-y_n}{y_n(1-y_n)}\,, \\
E'' = \sum_n \dfrac{x_n(y_{n+1}-y_n)}{y_n(1-y_n)}\,, \\
\Delta_1 = \sum_n \dfrac{x_n}{1-y_n}\,, \\
\Delta_2 = \sum_n \dfrac{x_n}{y_n(1-y_n)}\,, \\
\Delta_3 = \sum_n \dfrac{x_n^2}{y_n(1-y_n)}\,, \\
\Delta = A'^2\Delta_3 + B'\Delta_2^2 + C'\Delta_1^2 - 2A'\Delta_1\Delta_2 - C'B'\Delta_3\,.
\end{cases}
\tag{4.8}
$$

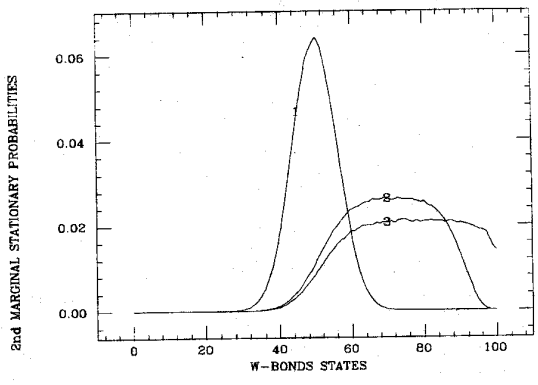
Finally, from $\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}$, the estimates $\hat{p}_i, \hat{\Delta}p_i$ and $\hat{\Delta}p_{12}$ of the parameters $p_i, \Delta p_i$, $i = 1,2$ and Δp_{12} are easily obtained by using (3.9), in the approximation $N_1 = N_s$, $N_2 = N_w$ *large*.

5. Numerical Results and Computer Simulations

In this section, we deal with numerical results obtained via computer simulations. A Monte–Carlo method was used to simulate the evolution of the system. The joint stationary probabilities were found numerically, by computing the frequencies of visit to each state $(i,j) \in \{0,1,\ldots,N_s\} \times \{0,1,\ldots,N_w\}$ in a very long simulation run (from 10^5 to 10^6 time iteration steps in some cases), for several values of the

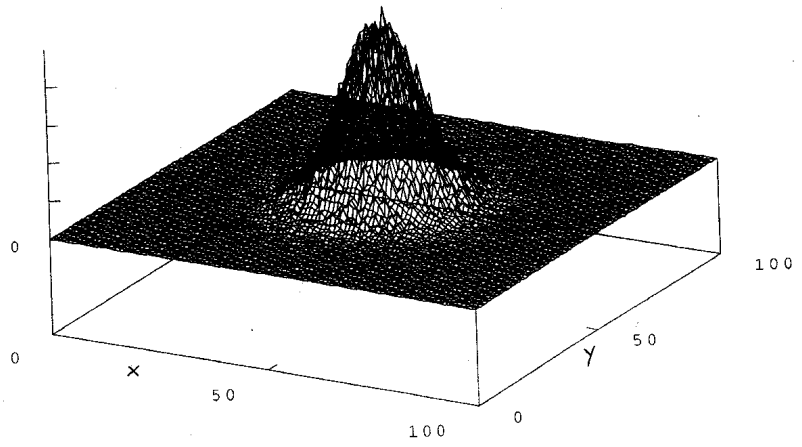


(a)

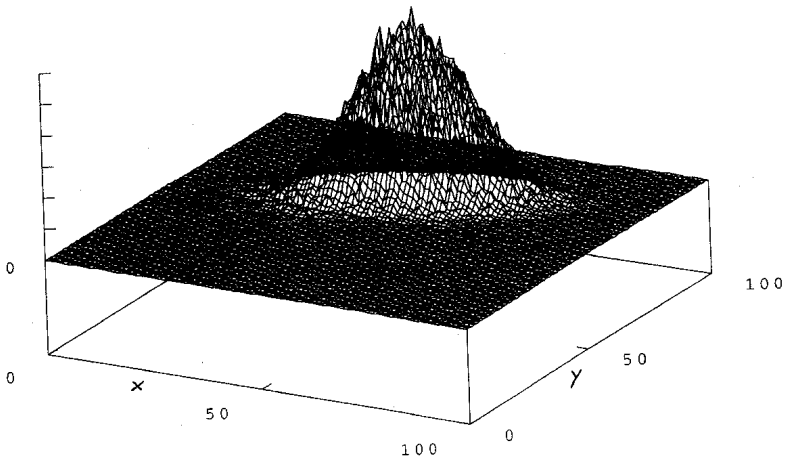


(b)

Fig. 1. (a) Plots of first marginal stationary distributions as a function of the s-bonds state (i.e. the number of *strong* bonds which are formed at infinite time) relative to the following three cases: $p_s = p_w = 0.5, \Delta p_s = 0.45, \Delta p_w = 0.3$, and $\Delta p_{sw} = 0$, 0.08 and 0.1 for (1), (2) and (3) respectively. Notice that the curves overlap, since the change in the cross-coupling capacity does not affect the bonds of *strong* type. The computation has required a simulation run of one million iteration steps. (b) Comparison of plots of second marginal stationary distributions versus the w-bonds state (i.e. the number of *weak* bonds which are formed at infinite time) in the three cases relative to (a).

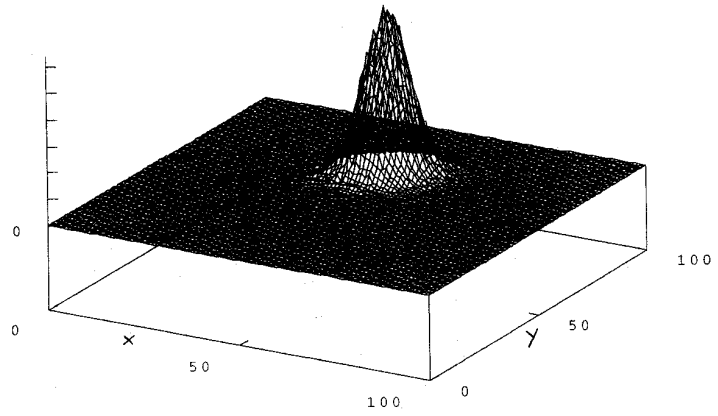


(a)



(b)

Fig. 2. (a) Three-dimensional plot of the joint stationary distribution for $N_s = N_w = 100$, $p_s = p_w = 0.5, \Delta p_s = \Delta p_w = 0.4$, $\Delta p_{sw} = 0$. The peak value is obtained exactly at the middle, $(x_M, y_M) = (50, 50)$. For these values of the parameters, with probability one, no portion of the boundary of $\mathcal{R}$ is accessible. The computation required a simulation run of 100,000 iteration steps. (b) Three-dimensional plot of the joint stationary distribution for $p_s = p_w = 0.5$, $\Delta p_s = 0.45$, $\Delta p_w = 0.2$, $\Delta p_{sw} = 0.1$. On the x-axes the s-bonds state is reported, from 0 to N_s; on the y-axes the w-bonds state is reported, from 0 to N_w; on the vertical axes z there is the stationary probability that the system stays at the state (x, y) at equilibrium, i.e. $z = \Pi_{(x,y)}$. Here, for the sake of diagram simplicity, we took $N_s = N_w = 100$ (as easily seen, the frequency of visit to a state does not depend on the maximum number of the allowed bonds). The peak value at (x_M, y_M), i.e. the most probable value, is not the middle state, but it is shifted in the direction of the y-axes, with $(x_M, y_M) \approx (50, 70)$. The computation required a simulation run of 100,000 iteration steps. (c) As in (b) for $N_s = N_w = 100$, $p_s = p_w = 0.5$, $\Delta p_s = 0.1$, $\Delta p_w = 0.22$, $\Delta p_{sw} = 0.1$. The peak point is obtained approximately at the same values of x, y of (b), but the peak is more narrow than in (b); the distribution is more concentrated around its mean value (x_M, y_M). The computation required a simulation run of 100,000 iteration steps.



(c)

Fig. 2. (*Continued*)

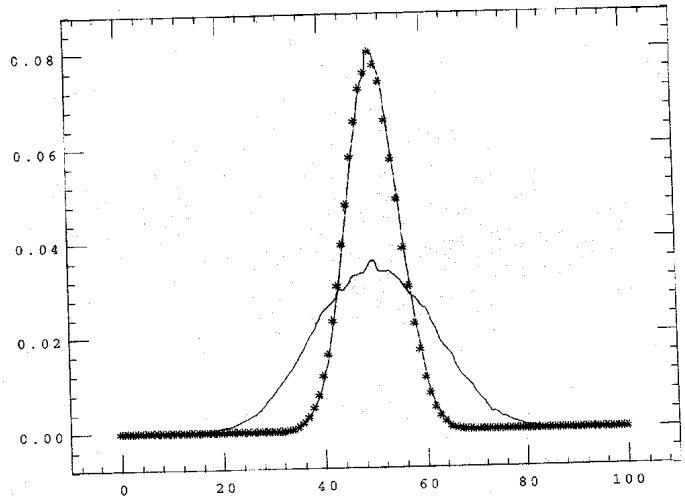
parameters. Also the marginal stationary probabilities were computed, for both processes X_n and Y_n.

In Figs. 1a–b the plots of the marginal stationary probabilities are reported as a function of the state $n_s \in \{0, 1, \ldots, N_s\}$ and $n_w \in \{0, 1, \ldots, N_w\}$, respectively, for a set of values of the parameters (see the figures' caption).

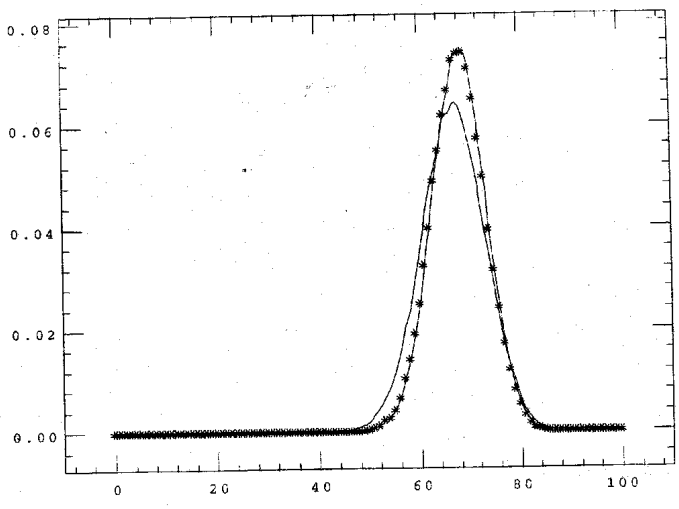
From these graphs it follows that, for zero values of the cross-coupling capacity Δp_{sw}, the curves relative to the marginal stationary probabilities behave like beta functions centered at the middle state. When Δp_{sw} increases from 0 to positive values, the curve relative to the first marginal stationary probabilities does not change, while the abscissa of the maximum of the curve relative to *weak* bonds (second marginal stationary distribution) shifts to the right. This means, as expected, that, for $\Delta p_{sw} > 0$, the influence of *strong* bonds on the formation of additional *weak* ones becomes non-negligible.

In fact, for $\Delta p_{sw} > 0$, not only the mean of the second marginal distribution is shifted to the right, but also there are very few values concentrated around the mean, implying that, with a large probability, a number of w-type bonds, belonging to an interval of length $30\%N_w$ around the mean, can be formed at equilibrium (see Fig. 1b). All that is detectable also by examination of three-dimensional plots of the joint stationary distribution of the process (X_n, Y_n) (see Figs. 2a–c). Moreover, the simulations show that, at parity of p_s, p_w and Δp_{sw} the two parameters Δp_s and Δp_w can be *interchanged,* in some sense, if one wants to obtain the same behavior in the formation of *weak* bonds. This means that the mean value of w-type bonds at equilibrium remains unchanged under the condition that a decrease of Δp_w is accompanied by an increase of Δp_s (see Figs. 3a–b).

Thus, we observe an effective picture of a cooperative phenomenon, with a hierarchic structure, where for $\Delta p_{sw} \neq 0$, the evolution of the system is



(a)



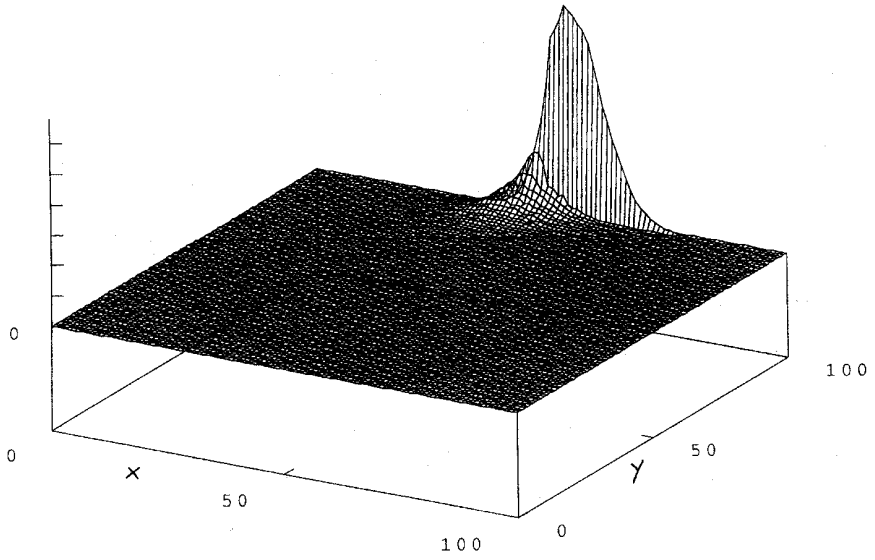
(b)

Fig. 3. (a) Comparison of plots of first marginal stationary distributions in the two cases relative to Figs. 2b (– – –) and 2c (* * *). (b) Comparison of plots of second marginal stationary distribution in the two cases relative to Figs. 2b (– – –) and 2c (* * *).

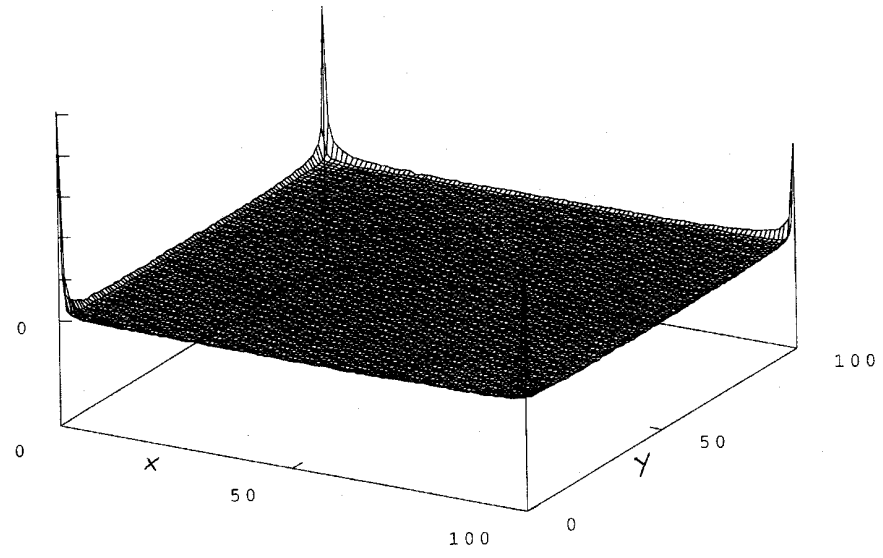
essentially driven by the capability to form *strong* bonds, since they also influence the formation of *weak* ones.

For what concerns attainability or unattainability of the boundary states (full coupling or uncoupling for one or both types of bonds), the results of Proposition 3.2 are amply confirmed by numerical computation.

The possibility or impossibility to reach the boundaries of $\mathcal{R}$ with a positive probability is shown by three-dimensional plots of the joint stationary distribution

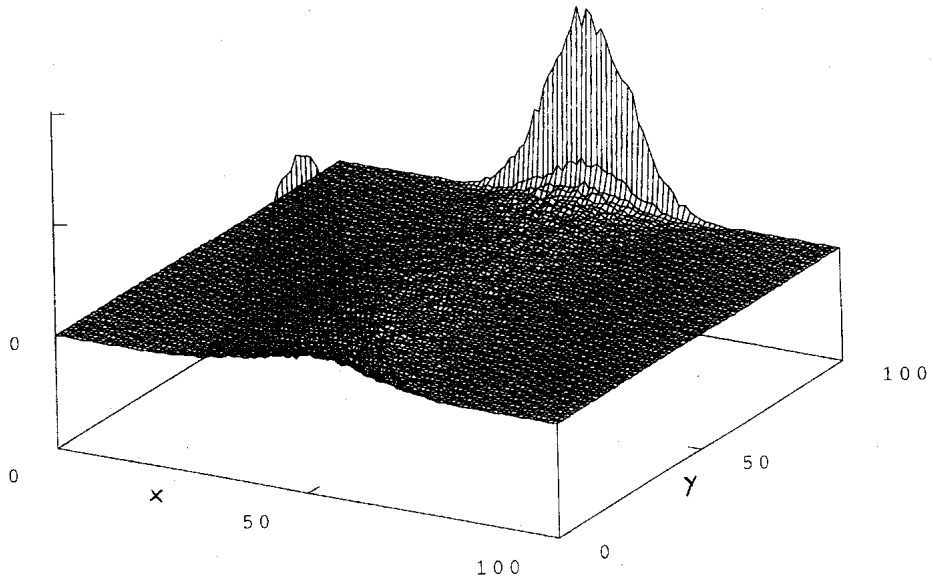


(a)

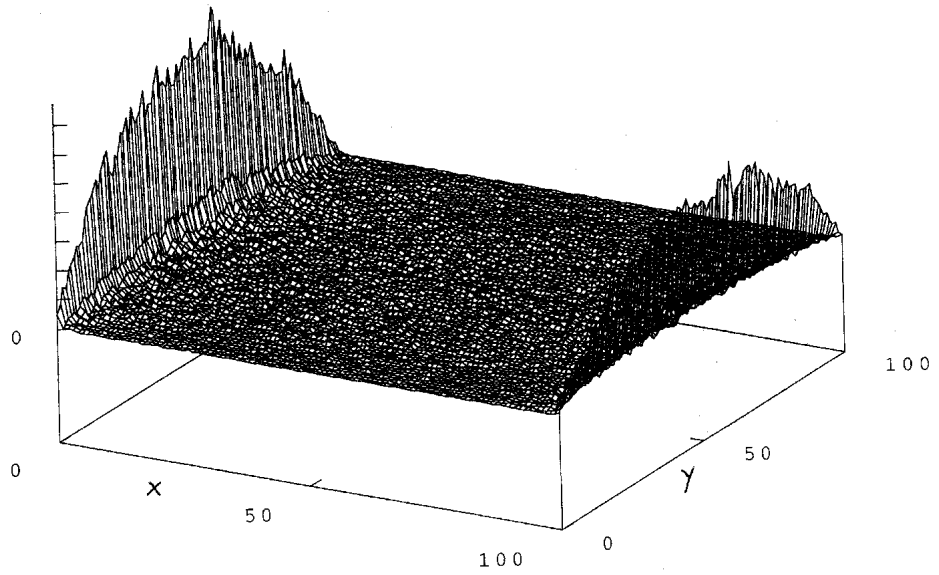


(b)

Fig. 4. (a) As in Fig. 2, for $p_s = p_w = 0.5, \Delta p_s = \Delta p_w = 0.3$, $\Delta p_{sw} = 0.1$. The y-coordinate of the peak value is largely shifted to the right. (b) As in Fig. 2, for $p_s = p_w = 0.5$, $\Delta p_s = \Delta p_w = 0.499$, $\Delta p_{sw} = 0$; the graph is a concave surface. For these values of the parameters, all the portions of the boundary of $\mathcal{R}$ are attainable.



(a)



(b)

Fig. 5. (a) As in Fig. 2, for $p_s = p_w = 0.5$, $\Delta p_s = 0.4$, $\Delta p_w = 0.499$, $\Delta p_{sw} = 0$. For these values of the parameters, the portions of boundary l_0 and l_1 are inaccessible, while $\mathcal{L}_0$ and $\mathcal{L}_1$ are accessible. (b) As in Fig. 2, for $p_s = p_w = 0.5$, $\Delta p_s = 0.499$, $\Delta p_w = 0.49$, $\Delta p_{sw} = 0$. For these values of the parameters, the portions of boundary l_0 and l_1 are accessible, while $\mathcal{L}_0$ and $\mathcal{L}_1$ are inaccessible.

of the process (X_n, Y_n). In the case of Figs. 2a–c and 4a, for instance, no portion of the boundary of $\mathcal{R}$ can be reached (roughly speaking, the *probability density calculated at the boundary is zero*).

Figure 4b illustrates the case in which each portion of the boundary of $\mathcal{R}$ can be reached (roughly speaking, the *probability density is infinite at the boundary*).

Figure 5a illustrates the case when the portions of the boundary l_0 and l_1 are inaccessible and $\mathcal{L}_0$ and $\mathcal{L}_1$ are accessible. Figure 5b refers to the converse case in which l_0 and l_1 are accessible and $\mathcal{L}_0$ and $\mathcal{L}_1$ are inaccessible (see the figures caption).

Now, we are concerned with numerical results about parameters estimation. The parameters of the model were estimated by means of the maximum likelihood method.

5.1. *Parameters estimation by using simulated data*

By using sequences of simulated data, obtained via a Monte–Carlo method by running a FORTRAN computer program with given input values of the parameters, we recovered the estimates of these parameters, as if they were unknown. We used sequences consisting of 2,000 two-dimensional data, but equally good estimates can be obtained by using shorter sequences. The obtained estimates are very satisfactory, in a lot of cases. In fact, we obtained a high accuracy in the estimates of all the parameters (the error was less than 0.1%) when the true (input) value of Δp_{sw} is *small* enough ($< 5 \times 10^{-2}$) (see Table 1). When the input value of Δp_{sw} is greater than 5×10^{-2}, the estimates of p_s and Δp_s, Δp_w are equally *precise,*

Table 1. Parameters estimation recovered by the maximum likelihood method, using several simulated data obtained from given input parameters. We used sequences consisting of 2,000 two-dimensional data $(x_n, y_n)_{n=1,\ldots,2000}$.

Input Values

p_s	Δ_{p_s}	p_w	Δ_{p_w}	$\Delta_{p_{sw}}$	$p_w + \Delta_{p_{sw}}$
0.5	0.	0.5	0.	0.	0.5
0.5	0.1	0.5	0.1	0.01	0.51
0.5	0.1	0.5	0.1	0.05	0.55
0.5	0.1	0.5	0.1	0.06	0.56
0.5	0.1	0.5	0.1	0.1	0.6
0.5	0.3	0.5	0.2	0.02	0.52
0.5	0.3	0.5	0.3	0.03	0.53
0.5	0.3	0.5	0.3	0.06	0.56
0.5	0.3	0.5	0.3	0.08	0.58
0.5	0.2	0.4	0.1	0.02	0.42
0.5	0.4	0.5	0.4	0.05	0.55

Table 1 (*Continued*)

Estimates

$\hat{p}_s$	$\hat{\Delta}_{p_s}$	$\hat{p}_w$	$\hat{\Delta}_{p_w}$	$\hat{\Delta}_{P_{sw}}$	$\hat{p}_w + \hat{\Delta}_{p_{sw}}$
0.500	0.00	0.494	0.00	0.00	0.5
0.500	0.10	0.495	0.099	0.015	0.51
0.500	0.10	0.52	0.10	0.024	0.549
0.500	0.10	0.528	0.102	0.0303	0.559
0.500	0.10	0.56	0.100	0.037	0.599
0.500	0.31	0.499	0.20	0.021	0.52
0.500	0.31	0.498	0.30	0.030	0.53
0.500	0.31	0.51	0.30	0.0412	0.557
0.500	0.31	0.52	0.30	0.0543	0.575
0.500	0.20	0.399	0.0889	0.0189	0.418
0.500	0.40	0.508	0.40	0.0413	0.55

while the error in the estimates of p_w and Δp_{sw} is bigger; however, the estimate of the sum $p_w + \Delta p_{sw}$, is *more precise* (the error is less than 0.01%) (see Table 1). This is satisfactory enough, since, on average, one can consider that the *ground* probability to form a *weak* bond is given by the sum of two contributions: the one coming from the *weak* bonds themselves (i.e. p_w), and that due to the *strong* bonds (i.e. Δp_{sw}). Moreover, some inaccuracy has to be ascribed to the fact that the method for estimating parameters makes use of the diffusion approximation in the limit $N_s, N_w \to \infty$, while N_s and N_w are taken finite in the simulation.

5.2. *Parameters estimation for a real protein*

To test our model on a real system, we analyzed the data relative to a molecular dynamics simulation of the thermal denaturation of the protein Barnase, appeared in Ref. 5. In that paper the authors simulated the unfolding of the protein following a temperature jump from 300 K to 600 K, at neutral pH. During the relaxation, the molecule rapidly expands and the total potential energy increases due to a progressive rupture of intermolecular interactions.

As a measure of the number of the *strong* and *weak* bonds of our model, we used the potential energy for the Coulomb (electrostatic) and van der Waals forces, respectively, considering intermolecular and protein–water interactions. To obtain an estimate of the number of *bonds*, an average energy associated to every *bond* was needed. We chose the values

$$\Delta E_{es} = -1 \text{ kcal/mole},$$

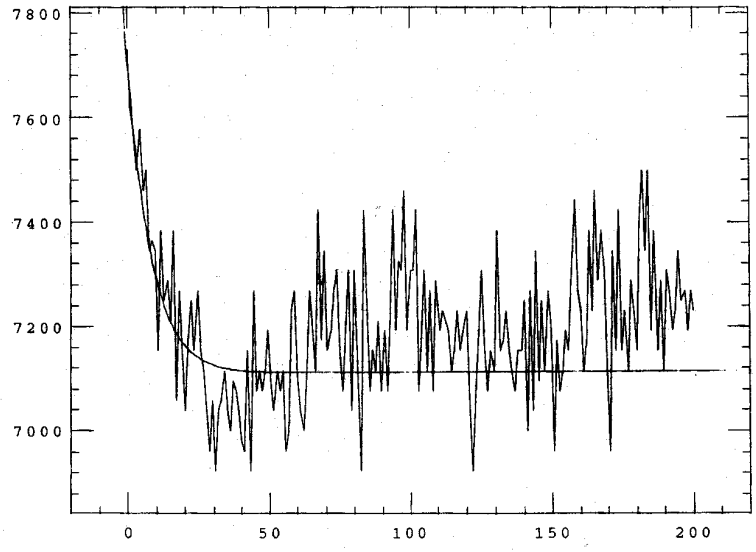
$$\Delta E_{vdw} = -0.1 \text{ kcal/mole}.$$

These values are somewhat arbitrary, but, reasonably, they do not vary more than a factor of 10; thus, the actual number of *bonds* should not be critical as long as it is large enough. Using the above values, we obtained the number of *weak* and *strong* bonds as a function of time (with a time-step of 1 ps), as shown in Figs 6a–b. We find $N_s \approx 7730$, $N_w \approx 10,350$, $N_{\text{tot}} \approx 18,080$.

The data corresponding to the total number of bonds (see Fig. 6c) were analyzed by the model for cooperative interactions without hierarchic structure.[1] By means of the one-dimensional diffusion approximation and the maximum likelihood method, we obtained the estimates:

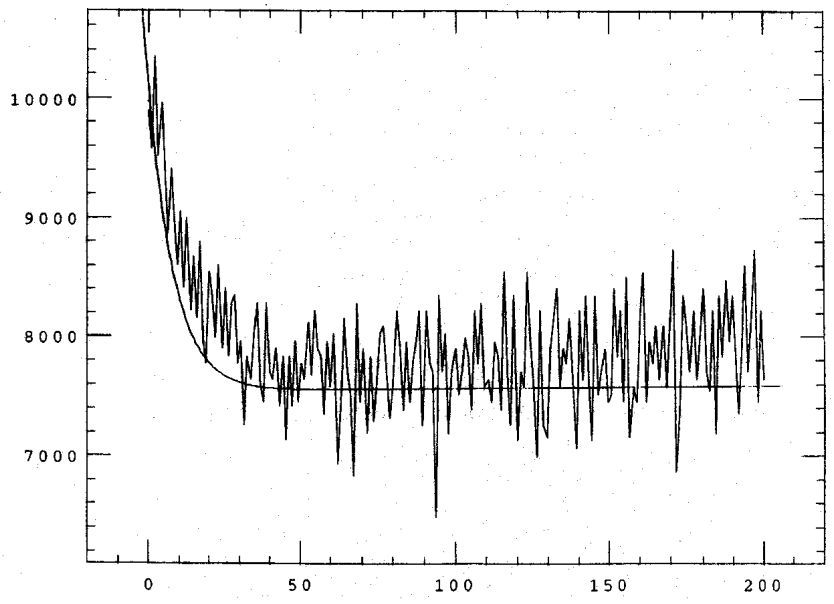
$$\bar{p}_{\text{tot}} = 0.60, \quad \bar{\Delta p}_{\text{tot}} = 0.37 .$$

In this case $\bar{\Delta p}_{\text{tot}}$ represents the *global cooperativity* of the protein.

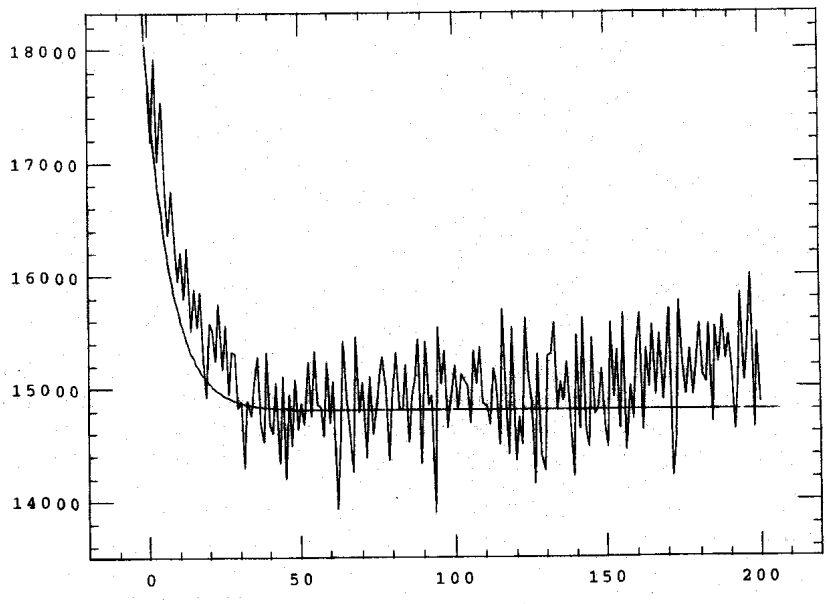


(a)

Fig. 6. (a) Plot of the time evolution (in ps) of the number of *strong* bonds obtained from the electrostatic energy for the thermal denaturation of the protein Barnase (data reported from Fig. 5g of Ref. 5). From these data the parameters regarding the *strong* bonds of the model have been estimated (see text for the numerical values). The superimposed curve represents the time evolution of the average process obtained by different simulation runs with input parameters equal to these estimates. (b) Plot of the time evolution (in ps) of the number of *weak* bonds obtained from the van der Waals energy for the thermal denaturation of the protein Barnase (data reported from Fig. 5f of Ref. 5). From these data and those of Fig. 6a, the parameters of the model regarding the *weak* bonds were estimated (see text for the numerical values). The superimposed curve represents the time evolution of the average process obtained by different simulation runs with input parameters equal to these estimates. (c) Plot of the time evolution (in ps) of the total number of bonds (*strong* + *weak*) for the protein Barnase. From these data, the parameters for the total number of bonds (*strong* + *weak*) were estimated (see text for the numerical values). The superimposed curve represents the time evolution of the average process.



(b)



(c)

Fig. 6. (*Continued*)

The two series of data corresponding to *strong* and *weak* bonds were analyzed by means of the hierarchic model; by using the two-dimensional diffusion approximation and the maximum likelihood method (see (4.7)), we obtained estimates for α_i, β_i and γ. From these, we recovered the estimates for p_i, Δp_i and Δp_{sw}:

$$\begin{cases} \bar{p}_s = 0.67\,, \\ \bar{p}_w = 0.33\,, \\ \bar{\Delta} p_s = 0.25\,, \\ \bar{\Delta} p_w = 0.16\,, \\ \bar{\Delta} p_{sw} = 0.18\,. \end{cases}$$

These values show a certain degree of cooperativity in both the *strong* and *weak* bonds and a dependence of the *weak* bonds on the *strong* ones.

As a further test of our model, we exchanged the roles of the van der Waals and electrostatic interactions, using the first ones as the *strong* bonds and the second ones as the *weak* ones. The fit obtained in this way is worse and Δp_{sw} is smaller (of order 10^{-3}) and even negative. This supports our choice of the electrostatic interactions as the *strong* bonds, and the significance of the value found for Δp_{sw} in that case.

We also calculated the likelihood function directly from the transition probabilities of the MC (2.7); then, we searched for its maximum, by using a routine for global optimization with constraints. Putting $\Delta p_{sw} = \Delta p_s$ (this hypothesis is justified by the fact that the *strong* bonds influence the *weak* ones in the same extent at least as they influence themselves), we obtained the following estimates of the parameters:

$$\begin{cases} \tilde{p}_s = 0.70\,, \\ \tilde{p}_w = 0.25\,, \\ \tilde{\Delta} p_s = 0.22\,, \\ \tilde{\Delta} p_w = 0.18\,, \\ \tilde{\Delta} p_{sw} = 0.22\,. \end{cases}$$

In this way, we removed the restriction on the magnitude of the parameter Δp_{sw}, since the diffusion approximation is adapted in principle to describe only a situation in which Δp_{sw} is rather small. However, we must observe that, also by using the diffusion approximation, the estimate of Δp_{sw} is not much smaller than Δp_s. Moreover, to find the maximum of the likelihood function without using the diffusion approximation, we spent a much longer computer time than in the case of the diffusion approximation (in that case, the point at which the likelihood function takes the maximum can be analytically found).

Although there are some differences between numerical values of the estimates obtained with the two methods, the shapes of the average processes (obtained by

different simulation runs with input parameters equal to the estimates respectively found) appear to be approximately the same. (See Fig. 6, where the data from Ref. 5 regarding the time-evolution of the number of bonds (*strong, weak, and strong + weak*) are reported together with the graphs of the three average processes, obtained by estimating the parameters indifferently with one of two methods.)

We remark that the estimates obtained by using the diffusion approximation in the case of protein interactions without hierarchic structure are still more satisfactory. This is because when the term Δp_{sw} disappears, the problem about the range of the cross-coupling parameter also disappears.

In fact, in the case of data regarding the unfolding of the bovine pancreatic trypsin inhibitor (BPTI), obtained by Daggett and Levitt through molecular dynamic simulations,[1,6] the estimates obtained by using the diffusion approximation are very close to those obtained numerically by us in our previous work,[1] where these data were analyzed by means of the model for cooperative interactions in proteins without hierarchic structure. Precisely, up to an error of order 10^{-5}, we recovered (by using the one-dimensional diffusion limit) the same values for $\bar{p}$ and $\bar{\Delta}p$ ($\bar{p} = 0.4956, \Delta p = 0.4575$; cf. Fig. 13 in Ref. 1 and the relative discussion).

6. Concluding Remarks

In this paper, we described a Markovian model for the cooperative interactions in proteins, based on a hierarchic structure of the interactions. It relies on the arbitrary splitting of interactions into two classes, *strong* and *weak*, and tests the preponderance of one class on the other (a greater number of different types of classes could result only in a heavy notation).

The model generalizes the one valid for homogeneous cooperative interactions in protein molecules previously studied by the authors.[1]

The time evolution of the system has been studied as a function of five parameters, three of which are related to the cooperativity. The main assumption is that, when the polypeptidic chain is relaxing toward the native protein conformation, the amino acidic residues can be bounded by *strong* and *weak* interactions, and the *strong* interactions can influence the *weak* ones, but not vice versa; thus, the presence of a *strong* bond between amino acidic residues could facilitate the local formation of *weak* ones.

The model depends on parameters p_s and p_w (which represent the mean probability to form *strong* and *weak* bonds, respectively), on parameters Δp_s and Δp_w (which represent the cooperativity capacities for *strong* and *weak* bonds, respectively), and finally on the parameter Δp_{sw} (which represents the cross coupling capacity). As in the case of interactions without structure,[1] the model presents a surprising wealth of qualitative behaviors, when the five parameters are varied. Moreover, this stochastic model, which is alternative to deterministic ones, presents some advantages: (i) its simplicity, (ii) the fact that it depends only on five parameters, (iii) the relatively short time required for numerical simulations.

The evolution of the system was described by a two-dimensional Markov chain with state space $\{(i,j) : i \in \{0,1,\ldots,N_s\}, j \in \{0,1,\ldots,N_w\}\}$, where N_s and N_w are, respectively, the total numbers of permitted couplings among residues of *strong* and *weak* type. Then, absorption problems were studied as a function of the parameters. When $p_i > \Delta p_i > 0$, $p_i + \Delta p_i < 1$, $(i = s, w)$ and $p_w + \Delta p_w + 2\Delta p_{sw} < 1$, the chain is irreducible and the stationary probabilities exist (they are the probabilities that the system stays in given states, at equilibrium). By simulating the evolution of the system with a Monte–Carlo method, and computing the frequency of visit to each state $(i,j) \in \{0,1,\ldots,N_s\} \times \{0,1,\ldots,N_w\}$ in a very long simulation run (about 10^6 time iteration steps), we numerically found the stationary probabilities, for several values of the parameters. Also the marginal stationary probabilities were computed for each component of the two-dimensional process.

The numerical results show the particular relevance of the cross-coupling capacity parameter Δp_{sw}. When it is set to zero, the two-dimensional process has independent components, each of them behaves as in the case of protein interactions without hierarchic structure. When Δp_{sw} increases from zero to a positive value, the first marginal stationary probabilities do not change, while the peak point of the joint stationary distribution is shifted to the right in the direction of the y-axes (number of *weak* bonds). This agrees with the fact that, as expected, for $\Delta p_{sw} > 0$, the influence of *strong* bonds on the formation of additional *weak* ones becomes non-negligible. Moreover, the simulations show that, for fixed p_s, p_w and Δp_{sw} the two parameters Δp_s and Δp_w can be *interchanged,* in some sense, if one wants to obtain the same behavior in the formation of *weak* bonds. This means that the mean value of w-type bonds at equilibrium remains unchanged under the condition that a decrease of Δp_w is accompanied by an increase of Δp_s.

For what concerns attainability or unattainability of the boundary states (full coupling or uncoupling for one or both types of bonds), the possibility or impossibility to reach the boundary states can be detected by the three-dimensional plot of the stationary probabilities: roughly speaking, a portion of the boundary is accessible or inaccessible, whether or not the stationary distribution is *infinite, or zero* at that portion.

Indeed, the qualitative results described above can be obtained by studying the diffusion approximation of the system. In fact, for p_i and Δp_i $(i = s, w)$ close to $1/2$ and Δp_{sw} close to zero, the discrete process can be approximated, for large N_s and N_w, by a continuous diffusion process. Once such a description of the system is available, its evolution could be more easily studied by considering the associated stochastic differential equation for the continuous process. Actually, it can be shown that the stationary distribution has a density and its shape is that of a perturbed product of two beta functions (the larger Δp_{sw}, the more it deviates from the product of beta functions), and this agrees with numerical results. The stationary probabilities can be approximately obtained by (3.11). We estimated the parameters of the model by using the maximum likelihood method

and the diffusion approximation. The estimates re-obtained from simulated data are rather precise, and excellent in a lot of cases. For the present model, the diffusion approximation represents (to our knowledge) the only analytical tool to obtain parameters estimation.

To test our model on a real system, we analyzed the data relative to a molecular dynamics simulation of the thermal denaturation of the protein Barnase, appeared in Ref. 5. In that paper, during the relaxation following a temperature jump from 300 K to 600 K, the total potential energy increases due to a progressive rupture of intermolecular interactions. By choosing the electrostatic interactions as the *strong* ones and the van der Waals interactions as the *weak* ones, we estimated the parameters of our model from these data (which represent the number of *strong* and *weak* bonds present at any time of the simulation). Numerical computations support our choice of the electrostatic interactions as the *strong* bonds, and the significance of the value found for Δp_{sw} in that protein.

Acknowledgments

We wish to thank A. Finazzi Agrò and P. Baldi for their valuable discussions on the argument. We would like to express particular thanks to the referee for useful comments and suggestions which have improved the presentation of the whole paper. This work was partially supported by a grant of CNR.

References

1. M. Abundo, L. Accardi and N. Rosato, *A Markovian model for cooperative interactions in proteins, Math. Models Methods Appl. Sci.* **5** (1995) 835–863.
2. M. Abundo, L. Accardi, A. Finazzi Agrò, G. Mei and N. Rosato, *A stochastic model for the sigmoidal behaviour of cooperative biological systems, Biophys. Chem.* **58** (1996) 313–323.
3. M. Abundo and L. Caramellino, *Some remarks about a Markov chain modelling cooperative biological systems, Open Systems Inform. Dynamics* **3** (1995) 325–343.
4. M. Abundo, P. Baldi and L. Caramellino, *A diffusion approximation which models hierarchic interactions in cooperative biological systems,* preprint Volterra No. 221 (1995), to appear in *Open Systems Inform. Dynamics* (1998).
5. A. Caflisch and M. Karplus, *Acid and thermal denaturation of Barnase investigated by molecular dynamics simulations, J. Mol. Biol.* **252** (1995) 672–708.
6. V. Daggett and M. Levitt, *Protein unfolding pathways explored through molecular dynamics simulations, J. Mol. Biol.* **232** (1993) 600–619.
7. S. N. Ethier and T. G. Kurtz, **Markov Processes. Characterization and Convergence** (Wiley, 1986).
8. D. Koroliuk and V. S. Koroliuk, *Diffusion approximation of Markov chains for cooperative interactions in proteins,* preprint Volterra (1994).
9. D. W. Strook and S. R. S. Varadhan, **Multidimensional Diffusion Processes** (Springer-Verlag, 1979).