

# On the Continuous Diffusion Approximation of Some Discrete Markov Chains

Mario Abundo\*

Dipartimento di Matematica, Università “Tor Vergata”, via della Ricerca Scientifica, I-00133 Rome, Italy

**Abstract:** The diffusion approximation of some Markov chains with Binomial-like transition probabilities is considered and their qualitative behavior is analyzed. The stationary distribution, attainability of boundary and estimation of parameters are investigated. Some examples are reported.

**Keywords:** Markov chains, diffusion process, stationary distribution.

## 1. INTRODUCTION

A lot of phenomena of interest in Physics and Biology can be mathematically described by means of discrete Markov chains (MCs). In this paper, we will consider the following class of MCs: let  $X_k$  be the size of a population at time  $k$  (here, the term population is meant in generalized sense, i.e. it not only denotes a group of cells or individuals, but even a collection of particles, molecules, etc.), where  $X_k$  takes values  $n \in \{0, 1, \dots, N\}$ , and  $N$  is the maximum allowed size of the population. Then, we can construct the MC with states  $\{0, 1, \dots, N\}$  and binomial-like transition probabilities:

$$P(X_{k+1} = m | X_k = n) = \binom{N}{m} \theta \left(\frac{n}{N}\right)^m \left(1 - \theta\left(\frac{n}{N}\right)\right)^{N-m} \quad (1.1)$$

where  $\binom{N}{m} = \frac{N!}{m!(N-m)!}$  and the parameter  $\theta = \theta\left(\frac{n}{N}\right)$

is a function of the fraction  $\frac{n}{N}$  of individuals at time  $k$  over the maximum size of the population.

In particular, we suppose that  $\theta$  is a polynomial function of its argument:

$$\theta(x) = a_0 + a_1x + \dots + a_r x^r \quad (1.2)$$

where  $a_i$  are real constants,  $i = 0, 1, \dots, r$ . Examples of such MCs from biology will be given in the Section 4.

We observe that, because of the computational complexity of the formula (1.1) for  $N$  large, it is impractical to study the MC directly, so it is convenient to approximate it by a diffusion process; in fact many biological systems are well described in terms of continuous diffusion processes (see e.g. [1]). In our case, the advantages of a description in terms of diffusion processes are: (i) the shorter computer time required to simulate the evolution of the population

(especially for large  $N$ ); (ii) more convenience to get the stationary probabilities of the MC and to study conditions for attainability/unattainability of the boundary states; (iii) more convenience to estimate the parameters  $a_i$ . Thus, our goal is to study the diffusion approximation of the MC, that allows us to get easily information about the qualitative behavior of the system. Actually, by considering a sequence  $X_k^{(N)}$  of MCs where the coefficients  $a_i = a_i(N)$  in (1.2) depend on  $N$ , we study the problem of the convergence, as  $N \rightarrow \infty$ , of the rescaled MC  $\frac{X_k^{(N)}}{N}$  to a continuous diffusion process  $X(t)$  in  $[0,1]$ , in such a way, for  $N$  large:

$$X(t) \approx \frac{1}{N} X_{[Nt]}^{(N)} \quad (1.3)$$

( $[\cdot]$  denotes the integer part).

The sense of the convergence in (1.3) will be precised later.

Once the description in terms of diffusion process is available, the evolution of the system can be more easily studied by considering the associated stochastic differential equation (SDE) for the process  $X(t) \in [0, 1]$ . Of course, in order to make the SDE well defined, it needs that  $X(t)$  never exits from the interval  $[0, 1]$ , and in fact conditions on the coefficients  $a_i$  in (1.2) will be imposed to guarantee the above property.

While the drift coefficient of the SDE turns out to be Lipschitz-continuous and bounded, the diffusion term is not (in fact it is Hölder-continuous), so the usual conditions for the uniqueness of the solution, starting from  $X(0) \in [0, 1]$ , are not satisfied. However, the existence and uniqueness of the solution with given initial datum follow by a result of Yamada and Watanabe [2]; the continuity of the solution from initial value (i.e. the Feller Property) can be also proved (see e.g. [3]).

It is useful for the discrete process  $X_k$  to study conditions on the parameters  $a_i$  in (1.2) in order that the boundary states 0 and  $N$  are *accessible* or *unaccessible*. This, translated in terms of the limit diffusion process, results into studying the nature of the boundaries of the interval  $[0, 1]$ , following the

\*Address correspondence to this author at the Dipartimento di Matematica, Università “Tor Vergata”, via della Ricerca Scientifica, I-00133 Rome, Italy; E-mail: abundo@mat.uniroma2.it

Feller classification (see e.g. [1]). Another interesting problem for the discrete process  $X_k$  is to find the stationary probabilities  $\pi_i$  of the MC, for those values of the parameters  $a_i$  for which the chain is irreducible (for instance, this happens if  $p_{ij} > 0$ ). Unfortunately, we are not able to find the exact formula for  $\pi_i$  (see e.g. the discussion in [4]); it is more convenient to consider the forward (Fokker-Plank) equation for the diffusion process (see e.g. [1])

$$\frac{\partial u}{\partial t} = L^* u \tag{1.4}$$

where  $L^* u$  is the formal adjoint of the generator  $L$  of the diffusion process, and look at a stationary solution  $u$ , i.e.

$$L^* u = 0 \tag{1.5}$$

If such a function  $u$  exists, it represents the probability density function of the invariant measure. Once the density  $u$  is known, one has the approximation:

$$\pi_i \approx \frac{1}{N} u\left(\frac{i}{N}\right), \quad N \text{ large} \tag{1.6}$$

Indeed, the eq. (1.5) cannot be solved, in general, and only in some cases we are able to exhibit the explicit formula. Thus, we shall consider the problem of proving the existence and uniqueness of the invariant measure and of its density w.r.t. Lebesgue measure. Finally, we shall be concerned with the estimation of the parameters  $a_i$  from experimental data, by means of the maximum likelihood method. Once again, the diffusion approximation turns out to be very useful, because the maximization of the likelihood function in the discrete case appears to be theoretically intractable.

Only some results of this paper will be proved, while others will be stated without proof, referring to the cited papers.

## 2. A DIFFUSION APPROXIMATION SCHEME

Here we show how one can obtain the diffusion approximation of the MC with transition probabilities (1.1) with  $\theta$  depending on  $a_i(N)$ . First, we present a non-rigorous, heuristic argument; then, we state precisely the result.

Let us consider the limit process  $X_t \in [0, 1]$ , then  $\forall x \in [0, 1]$  the infinitesimal drift  $b(x)$  must satisfy:

$$E(X(t + dt) - X(t) \mid X(t) = x) = b(x)dt \tag{2.1}$$

Now, if we take  $k = [Nt]$ ,  $dt = \frac{1}{N}$ , approximately we can write, for  $N \rightarrow \infty$ :

$$E(X(t + dt) - X(t) \mid X(t) = x) \approx \tag{2.2}$$

$$E(Y_{k+1} - Y_k \mid Y_k = \frac{[Nx]}{N})$$

where  $Y_k = Y_k^{(N)} \doteq \frac{X_k^{(N)}}{N}$  is the rescaled MC. Recalling the expectation value of a binomial r.v., by (1.1) the last quantity is:

$$\theta \left( \frac{[Nx]}{N} \right) - \frac{[Nx]}{N}$$

which implies by (2.1):

$$b(x) \approx N \left[ \theta \left( \frac{[Nx]}{N} \right) - \frac{[Nx]}{N} \right] \approx N(\theta(x) - x), \quad N \rightarrow \infty \tag{2.3}$$

(Note that, since  $\frac{Nx-1}{N} < \frac{[Nx]}{N} \leq \frac{Nx}{N}$ , then  $\frac{[Nx]}{N} \rightarrow x \in [0, 1]$  as  $N \rightarrow \infty$ ). Now, we search for  $b(x)$  in a polynomial form:

$$b(x) = \lambda_0 + \lambda_1 x + \dots + \lambda_r x^r \tag{2.4}$$

From (2.3) we must have, for  $N$  large:

$$\sum_{i=0}^r \lambda_i \left( \frac{[Nx]}{N} \right)^i \approx N \left[ \sum_{i=0}^r a_i^{(N)} \left( \frac{[Nx]}{N} \right)^i - \frac{[Nx]}{N} \right] \tag{2.5}$$

$$= N \left[ a_0^{(N)} + (a_1^{(N)} - 1) \frac{[Nx]}{N} + a_2^{(N)} \left( \frac{[Nx]}{N} \right)^2 + \dots + a_r^{(N)} \left( \frac{[Nx]}{N} \right)^r \right]$$

Equating the coefficients of the terms of same degree, we obtain in the limit  $N \rightarrow \infty$ :

$$\begin{cases} \lambda_0 = \lim N \cdot a_0(N) \\ \lambda_1 = \lim N \cdot a_1(N) - 1 \\ \lambda_i = \lim N \cdot a_i(N), \quad i = 2, \dots, r \end{cases} \tag{2.6}$$

Of course, (2.6) are consistent only if, as  $N \rightarrow \infty$ :

$$\begin{cases} a_0(N) \rightarrow 0 \\ a_1(N) \rightarrow 1 \\ a_i(N) \rightarrow 0, \quad i = 2, \dots, r \end{cases} \tag{2.7}$$

Set  $\bar{x} = E\left(\frac{X_{k+1}}{N} \mid \frac{X_k}{N} = \frac{[Nx]}{N}\right) = \theta\left(\frac{[Nx]}{N}\right)$ . From (1.2) and

$$(2.7), \text{ we obtain} \tag{2.8}$$

$$\bar{x} \rightarrow x \text{ as } N \rightarrow \infty$$

$$\text{(Precisely } \theta\left(\frac{[Nx]}{N}\right) - \frac{[Nx]}{N} = O(1/N)\text{)}$$

Now, we consider the diffusion term  $a(x)$ . We have:

$$a(x)dt = E((X(t + dt) - X(t))^2 \mid X(t) = x) \tag{2.9}$$

Approximately, for  $N \rightarrow \infty$ :

$$a(x)dt \approx E\left[\left(\frac{X_{k+1}}{N} - \frac{X_k}{N}\right)^2 \mid \frac{X_k}{N} = \frac{[Nx]}{N}\right] = \tag{2.10}$$

$$= E\left(\frac{X_{k+1}}{N} - \bar{x}\right)^2 + 2E\left(\frac{X_{k+1}}{N} - \bar{x}\right)\left(\bar{x} - \frac{X_k}{N}\right) + E\left(\bar{x} - \frac{X_k}{N}\right)^2$$

This quantity is approximately equal to the infinitesimal variance. Indeed, by (2.8) the last two terms go to zero as  $N \rightarrow \infty$ , so we obtain:

$$E\left[\left(\frac{X_{k+1}}{N} - \frac{X_k}{N}\right)^2 \mid \frac{X_k}{N} = \frac{[Nx]}{N}\right] \approx \tag{2.11}$$

$$\text{Var}\left(\frac{X_{k+1}}{N} \mid \frac{X_k}{N} = \frac{[Nx]}{N}\right) = \frac{1}{N} \theta\left(\frac{[Nx]}{N}\right) \left(1 - \theta\left(\frac{[Nx]}{N}\right)\right)$$

$$\frac{1}{N} \left(\frac{[Nx]}{N}\right) \left(1 - \frac{[Nx]}{N}\right)$$

Finally, from (2.11), (2.10) and (2.9):

$$a(x)dt \approx a\left(\frac{[Nx]}{N}\right)\frac{1}{N} \approx \frac{1}{N}\left(\frac{[Nx]}{N}\right)\left(1 - \frac{[Nx]}{N}\right) \quad (2.12)$$

from which it follows

$$a(x) = x(1 - x) \quad (2.13)$$

The heuristic considerations above can be summarized into the following rigorous result:

**Theorem 2.1.** For any positive integer  $N$ , let us consider the MC with transition probabilities (1.1) depending on the parameters  $a_i(N)$  and suppose that the limits in (2.6) exist.

Then, as  $N \rightarrow \infty$ , the process  $\frac{1}{N}X_{[Nt]}^{(N)}$  with values in  $K_N =$

$\left\{\frac{i}{N}, i = 0, 1, \dots, N\right\}$  weakly converges in the Skorohod space  $D(\mathbb{R}^+, [0, 1])$ , to the diffusion process  $X_t \in [0, 1]$ , whose associated backward differential equation has the form:

$$\frac{\partial u}{\partial t} = b(x)\frac{\partial u}{\partial x} + \frac{1}{2}x(1-x)\frac{\partial^2 u}{\partial x^2}, u \in C^2[0,1] \quad (2.14)$$

where  $b(x) = \sum_{i=0}^r \lambda_i x^i$  and the coefficients  $\lambda_i$  are given by (2.6).

We omit here the proof, since it is quite analogous to that of theorem 3.1 in [5].

The Itô's stochastic differential equation (SDE) associated to our diffusion process in  $[0,1]$  is the following:

$$dX(t) = b(X)dt + \sqrt{[X(1-X)]^+} dB_t \quad (2.15)$$

where  $a^+ = \max(a, 0)$ , and  $B_t$  is a standard Brownian motion.

**Theorem 2.2.** The SDE (2.15) has a unique strong solution  $X(t)$ , for any initial condition  $X(0) \in [0, 1]$ . Moreover, the continuity with respect to initial conditions holds for the transition function (Feller property)

(For the proof see e.g. [3]).

In order that (2.15) describes the evolution of a process which never leaves the interval  $[0, 1]$ , we have to impose some further conditions on the function  $b(x)$ , i.e. on the coefficients  $\lambda_i$ . We obtain:

**Theorem 2.3.** If  $b(0) = \lambda_0 \geq 0$  and  $b(1) = \sum_{i=0}^r \lambda_i \leq 0$ , then the solution of the SDE (2.15) with initial condition  $X(0) \in [0, 1]$ , remains in the interval  $[0, 1]$ , for all times  $t \geq 0$ .

*Proof.* We shall prove that for all  $t \geq 0$  (i)  $X(t) \geq 0$ ; (ii)  $X(t) \leq 1$ .

(i) First, we consider the case  $\lambda_0 = 0$ . If  $X(0) = 0$ , then  $X(t) = 0$  for every  $t$ , because of the uniqueness of the solution. If  $X(0) > 0$ , let be  $\tau = \inf\{t : X(t)=0\}$ , i.e.  $X(s) > 0$  for  $0 < s < \tau$ . Then,  $\tilde{X}(t) = X(t + \tau)$  is the solution of (2.15) starting from  $\tilde{X}(0) = 0$ , and, for the uniqueness of the solution,  $\tilde{X}(t) = 0 \forall t$ , i.e.  $X(t + \tau) = 0 \forall t$  which implies  $X(t) = X(\min(t, \tau)) \geq 0$ .

Let us consider now the case  $\lambda_0 = b(0) > 0$ ; if the process  $X(t)$  were allowed to become negative, there should be  $\varepsilon > 0$  and a time  $s > \tau$  such that, in the interval  $[\tau, s]$   $X(t)$  decreases from 0 to  $-\varepsilon$ ; we can suppose that  $\varepsilon$  is small enough so that, for  $t \in [\tau, s]$ , it holds:

$$b(X(t)) = \lambda_0 + \lambda_1 X(t) + \dots + \lambda_r X^r(t) > 0 \quad (2.16)$$

Then, by the equation

$$X(s) = X(\tau) + \int_{\tau}^s b(X(u))du + \int_{\tau}^s \sqrt{[X(u)(1-X(u)]^+} dB_u,$$

since the integrand of the stochastic integral vanishes for  $X(t) < 0$ , we obtain

$$-\varepsilon = \int_{\tau}^s b(X(u))du$$

which contradicts (2.16), because that equation implies that  $\int_{\tau}^s b(X(u))du$  is nonnegative; in conclusion it must be  $X(t) \geq 0$  for all  $t \geq 0$ .

(ii) To show that  $X(t) \leq 1 \forall t$ , we consider the process  $Y(t) = 1 - X(t)$  which solves the SDE  $dY = \bar{b}(Y)dt - \sqrt{[Y(1-Y)]^+} dB_t$ ,  $\bar{b}(Y) = -b(1-Y)$ . It suffices to proof that  $Y(t) \geq 0 \forall t$ ; we can proceed in analogous way as done in (i).

The proof of Theorem 2.3 is concluded.

Now, we look at the invariant measure of the diffusion process driven by the SDE (2.15). We recall that a measure  $\mu$  on  $[0, 1]$  is said to be invariant or stationary for the diffusion process  $X(t)$  if for every borel set  $E$ ,  $\forall t > 0$ :

$$\mu(E) = \int P(t, x, E)\mu(dx) \quad (2.17)$$

where  $P(t, x, \cdot) = P_x(X(t) \in \cdot)$ ,  $x \in [0, 1]$  is the transition probability of the diffusion process  $X(t)$ . Since  $[0, 1]$  is compact and the Feller property holds for the associated semigroup operator, one can easily prove the existence of a stationary measure  $\mu$  by a proof similar to the Markov-Kakutani theorem (see e.g. [6]). Thus, the problem is to show the uniqueness of  $\mu$  and to find its density w.r.t. the Lebesgue measure, if it exists. Let  $L$  be the infinitesimal generator of the diffusion driven by the SDE (2.15), acting on  $u \in C^2([0, 1])$ , that is:

$$Lu = b(x)\frac{du}{dx} + \frac{1}{2}x(1-x)\frac{d^2u}{dx^2} \quad (2.18)$$

As told in the Introduction, the density of the invariant measure, if it exists, is given by the solution of the equation  $L^*u = 0$  (see e.g. [1]), where  $L^*$  is the adjoint of the differential operator  $L$ , i.e.:

$$L^*u = \left(\frac{1}{2}x(1-x)u(x)\right)' - (b(x)u(x))' \quad (2.19)$$

### 3. ATTAINABILITY OF THE BOUNDARY OF [0, 1]

We recall some results about the attainability of the boundary of a scalar diffusion process, according to the Feller classification (see e.g. [1]). Let us consider the SDE in

$[0, 1] : dX(t) = b(X)dt + \sigma(X)dB_t$ , and let  $[x_1, x_2]$  be any subinterval of  $[0, 1]$ ; consider the functions:

$$\phi(x) = \exp\left(-\int_c^x \frac{2b(s)}{\sigma^2(s)} ds\right) \tag{3.1}$$

$$\xi(x) = \phi(x) \int_c^x 2[\sigma^2(s)\phi(s)]^{-1} ds \tag{3.2}$$

where  $c \in [x_1, x_2]$ .

Let  $\tau$  be the first exit time from the interval  $(x_1, x_2)$  of the process  $X(t)$  starting from an interior point  $x \in (x_1, x_2)$ . The boundary point  $x_i$  is called *attainable* (or accessible) if  $P_x(\lim_{t \rightarrow \tau} X(t) = x_i, \tau < \infty) > 0$  and *unattainable* (or unaccessible) otherwise. More generally, the boundary point  $x_i$  is called *attractive* if  $P(\lim_{t \rightarrow \tau} X(t) = x_i) > 0$  and *repelling* (or natural) otherwise. The classification above can be characterized in terms of integrability conditions of the functions  $\phi$  and  $\xi$  defined in (3.1) and (3.2). Precisely, we have:

- $x_i$  is attractive if  $\phi$  is integrable over a neighborhood of  $x_i$  and repelling otherwise;
- $x_i$  is attainable if  $\xi$  is integrable over a neighborhood of  $x_i$ , unattainable otherwise;
- $x_i$  is reflecting if the function  $[\phi\sigma^2]^{-1}$  is integrable over a neighborhood of  $x_i$ , absorbing otherwise.

If the boundaries  $x_1$  and  $x_2$  are both reflecting, then the process  $X(t)$  has an invariant measure with density  $const \cdot [\phi\sigma^2]^{-1}$ . In particular, for the SDE (2.15) we have:

**Theorem 3.1.** *If  $2\lambda_0 < 1$ , then the boundary point  $x = 0$  is accessible and attractive, unaccessible and repelling otherwise; if  $2(\lambda_2 + \dots + \lambda_r - \lambda_0 - \lambda_1) < 1$ , then the boundary point  $x = 1$  is accessible and attractive, unaccessible and repelling otherwise. If  $\lambda_0 > 0$  and  $\lambda_0 + \lambda_1 - (\lambda_2 + \dots + \lambda_r) < 0$ , then the boundary points  $x = 0$  and  $x = 1$  are both reflecting and the process  $X(t)$  has an invariant measure with density*

$$const \cdot e^{-q(x)} x^{2\lambda_0-1} (1-x)^{2(\lambda_2+\dots+\lambda_r-\lambda_0-\lambda_1)-1} \tag{3.3}$$

where  $q(x)$  is a polynomial of degree  $r + 1$ .

*Proof.* It suffices to write explicitly the functions  $\phi$  and  $\xi$  in (3.1), (3.2) and to study the integrability conditions of them over neighborhoods of  $x = 0$  and  $x = 1$ .

#### 4. EXAMPLES

##### 4.1. Population Genetics (Fisher & Wright and Feller)

Take  $\theta(x) = x$  in (1.1), (1.2), i.e.  $a_0 = 0, a_1 = 1, a_i = 0, i > 1$ . One obtains the MC with transition probabilities:

$$P_{nm} = \binom{N}{m} \left(\frac{n}{N}\right)^m \left(1 - \frac{n}{N}\right)^{N-m} \tag{4.1}$$

It can be interpreted as follows. A number  $M$  of individuals are selected from a population in each generation. A particular gene assuming the form  $A$  and  $a$  has  $N = 2M$  representatives; if  $A$  occurs  $n$  times in the  $k$ -th generation, then  $a$  occurs  $N - n$  times. We assign to the population, at the  $k$ -th generation, the state  $n, 0 \leq n \leq N$ .

Assuming random mating, the composition of the successive generation is determined by  $N$  Bernoulli trials in which the

$A$ -gene has probability  $\frac{n}{N}$ . Then, (4.1) holds. The diffusion approximation of the MC is given by the process  $X(t) \in [0, 1]$  which is solution of the SDE: (by (2.6) the drift coefficient  $b(X)$  is identically zero)

$$dX(t) = \sqrt{[X(1-X)]^+} dB_t \tag{4.2}$$

The boundary point  $x = 0$  and  $x = 1$  are both attainable and absorbing. The density of the invariant measure does not exist.

##### 4.2. Cooperative Interaction in Proteins [4]

Choose two numbers  $p$  and  $\Delta p$  such that  $p \geq \Delta p > 0$  and  $0 \leq p \pm \Delta p \leq 1$  and take  $\theta(x) = (p - \Delta p) + 2x\Delta p$ , that is  $a_0 = p - \Delta p, a_1 = 2\Delta p, a_i = 0, i > 1$ . One obtains the MC having transition probabilities:

$$P_{nm} = \binom{N}{m} \left[ p - \Delta p + 2\Delta p \frac{n}{N} \right]^m \left[ 1 - (p - \Delta p) - 2\Delta p \frac{n}{N} \right]^{N-m} \tag{4.3}$$

The interpretation of the system is as follows. A certain number of particles (molecules) which can be connected by chemical bonds is considered. Let  $N$  be the total number of pairings (chemical bonds) among the particles of the system (amino acidic residues). A configuration is defined by a sequence of binary r.v.  $\{\xi_i^{(k)}\}_{i=1, \dots, N}$  such that  $\xi_i^{(k)} = +1$  if the  $i$ -th pair is linked at time  $k, \xi_i^{(k)} = -1$ , otherwise. The number  $p$  (mean probability) represents the probability of forming a chemical bond when the total number of existing bonds is exactly  $N/2$  (physically,  $p$  is inversely related to the activation energy for the formation of that chemical bond). The number  $\Delta p$  (coupling capacity) is the maximum increment of the probability to form a chemical bond, for a given  $p$ . It is supposed that the r.v.  $\xi_1^{(k)}, \dots, \xi_N^{(k)}$  are independent and Bernoullian with distribution:

$$P\left(\xi_i^{(k+1)} = +1 \mid \sum_{j=1}^N \xi_j^{(k)} = n\right) = p + \frac{\Delta p}{N} n \tag{4.4}$$

Thus, the greater the number of the chemical bonds already formed among amino acidic residues, the greater the probability that additional bonds can be formed. This is the distinctive feature of a cooperative phenomenon. From (4.4) a two-side Markov chain with state space  $\{-N, \dots, 0, \dots, N\}$  can be constructed; then the one-side MC with state space  $\{0, \dots, N\}$  and transition probabilities given by (4.3) is obtained by means of a variable change. The diffusion approximation of the MC (4.3) is given by the process  $X(t) \in [0, 1]$  which is solution of the SDE:

$$dX(t) = b(X)dt + \sqrt{[X(1-X)]^+} dB_t \tag{4.5}$$

The drift is  $b(X) = \beta - 2\alpha X$ , where  $\beta = \lambda_0 = \lim N(p(N) - \Delta p(N)) \geq 0; -2\alpha = \lambda_1 = \lim N(2\Delta p(N) - 1) < 0$ , with the condition  $\beta - 2\alpha \leq 0$ . (These conditions assure that the process never exit from  $[0, 1]$ , by theorem 2.3). By Theorem

3.1, if  $\beta < \frac{1}{2}$  the boundary point  $x = 0$  is accessible and attractive, unaccessible and repelling, otherwise; if  $4\alpha - 2\beta < 1$ , the boundary point  $x = 1$  is accessible and attractive, unaccessible and repelling, otherwise. If  $\beta < 1/2$  and  $4\alpha - 2\beta < 1$  the invariant measure of the process exists and it has a beta-type density:

$$u(x) = \frac{\Gamma(4\alpha)}{\Gamma(2\beta)\Gamma(4\alpha - 2\beta)} x^{2\beta-1}(1-x)^{4\alpha-2\beta-1} \quad (4.6)$$

(see also [5], for more details).

**4.3. Cooperative Growth with Overcrowding Effect**

Let be  $\alpha, \beta, \gamma > 0$  and take  $\theta(x) = \alpha + \beta x - \gamma x^2$  in (1.1), (1.2). One obtains the MC with state space  $\{0, \dots, N\}$  and transition probabilities:

$$p_{nm} = \binom{N}{m} \left( \alpha + \beta \frac{n}{N} - \gamma \left( \frac{n}{N} \right)^2 \right)^m \left( 1 - \alpha - \beta \frac{n}{N} + \gamma \left( \frac{n}{N} \right)^2 \right)^{N-m} \quad (4.7)$$

This MC is suitable to describe cooperative growth phenomena concerning populations of particles, where a positive cooperativity occurs when the size of the population is not too big. Otherwise, overcrowding inhibits the growth of the population. This is exactly the contrary of what happens for protein folding. The SDE associated to the diffusion approximation is:

$$dX(t) = (a + bX - cX^2)dt + \sqrt{[X(1-X)]^+} dB_t \quad (4.8)$$

where  $a = \lim N\alpha(N), b = \lim N(\beta(N) - 1), c = \lim N\gamma(N)$ . If  $a > 0$  and  $a + b - c < 0$ , then the process  $X(t)$  never exit from  $[0, 1]$ , by theorem 2.3.

**4.4. Cooperative Interaction in Proteins with a Hierarchic Structure Among Chemical Bonds [7]**

This example is somewhat more complicated than the previous ones, and indeed it is not included in the class of MCs considered. In fact, now a two-dimensional MC is considered having binomial-like transition probabilities. Let us suppose that two types of chemical bonds are available to link a pair of particles: the *strong* bonds and the *weak* ones. If  $N$  is the total number of permitted coupling,  $N = N_1 + N_2$ , where  $N_i$  is the total number of permitted bonds of strong and weak type, respectively. Moreover, we suppose that strong bonds make easier the formation of weak bonds, but not viceversa.

We look at the two-dimensional process  $(X_n, Y_n), n = 1, 2, \dots$  whose components represent the number of chemical bonds of strong and weak type, respectively, at time  $n$ . Let us suppose that the processes  $X_n$  and  $Y_n$  are conditionally independent, then  $(X_n, Y_n)$  turns out to be a MC with state space  $\{0, 1, \dots, N_1\} \times \{0, 1, \dots, N_2\}$  and transition probabilities: (see [7])

$$P_{(i_1, j_2) \times (j_1, i_2)} = \Pr((X_{n+1} = j_1, Y_{n+1} = i_2) | (X_n = i_1, Y_n = j_2)) =$$

$$= \binom{N_1}{j_1} \left( p_1 - \Delta p_1 + 2\Delta p_1 \frac{i_1}{N_1} \right)^{j_1} \times \left( 1 - p_1 + \Delta p_1 - 2\Delta p_1 \frac{i_1}{N_1} \right)^{N_1 - j_1} \times \binom{N_2}{j_2} \left( p_2 + \Delta p_2 - 2\Delta p_2 \frac{i_2}{N_2} + 2\Delta p_{12} \frac{i_1}{N_1} \right)^{j_2} \times \left( 1 - p_2 - \Delta p_2 + 2\Delta p_2 \frac{i_2}{N_2} - 2\Delta p_{12} \frac{i_1}{N_1} \right)^{N_2 - j_2} \quad (4.9)$$

where the parameters have an analogous meaning as in Example 4.2; the further parameter  $\Delta p_{12}$  represents the *cross-coupling capacity*; it measures the influence of strong bonds on the formation of weak ones. The limit, as  $N_i \rightarrow \infty$ , of the rescaled process is a two-dimensional diffusion process in  $[0, 1]^2$ , which is driven by the SDE:

$$\begin{cases} dX = (\beta_1 - 2\alpha_1 X)dt + \sqrt{[X(1-X)]^+} dB_t^1 \\ dY = (\beta_2 - 2\alpha_2 Y + \gamma X)dt + \sqrt{[Y(1-Y)]^+} dB_t^2 \end{cases}$$

where  $\beta_i, \alpha_i$  are defined in a similar manner as in Example 4.2, and  $\gamma = \lim_{N \rightarrow \infty} 2N_2 \Delta p_{12}(N)$ .

When  $\gamma = 0, \beta_i < 1/2$  and  $4\alpha_i - 2\beta_i < 1$ , the density of the invariant measure can be explicitly found as the product of two beta functions:

$$u_{\alpha, \beta, \gamma}(x, y) = const \cdot x^{2\beta_1-1}(1-x)^{4\alpha_1-2\beta_1-1} y^{2\beta_2-1}(1-y)^{4\alpha_2-2\beta_2-1}$$

For  $\gamma \neq 0$ , the PDE  $L^*u = 0$  cannot be solved explicitly. In this case, many technical difficulties arise; however, under some conditions on the coefficients, it is possible to prove that the process  $(X_t, Y_t)$  never exit from  $[0, 1]^2$ , and the (unique) invariant measure has a density, even if its explicit form cannot be found (see [3]).

**5. ESTIMATION OF PARAMETERS**

The parameters  $a_i$  in (1.2) can be easily estimated, in the approximation *N large*, by using the diffusion approximation of the MC  $X_k$  and the maximum likelihood method. Here, we outline the method in the case of Example 4.2 (for the case of Example 4.4, see [7]). By discretization of eq. (4.5), one obtains:

$$\begin{cases} x_{n+1} = x_n + (\beta - 2\alpha x_n)h + \sqrt{x_n(1-x_n)}\Delta B_n \\ x_0 = X(0) \end{cases} \quad (5.1)$$

where, for  $n = 0, 1, \dots, x_n$  denotes the process  $X_t$  evaluated at the time  $t_n = nh$ , and  $\Delta B_n = B(t_{n+1}) - B(t_n)$  is the increment of a standard Brownian motion. The equation (5.1) means that the random variable  $X_{n+1}$  conditionally to  $(X_n = x_n)$  is distributed according to a Gaussian with expectation  $x_n + (\beta - 2\alpha x_n)h$  and variance  $x_n(1-x_n)h$ .

Then, given the sample of data  $(x_n)_{n=0,1,\dots,M}$ , we obtain the likelihood function:

$$L(\alpha, \beta) = \prod_{n=1}^M \frac{1}{2\pi h \sqrt{x_n(1-x_n)}} \times \exp\left(-\frac{[x_{n+1} - x_n - (\beta - 2\alpha x_n)h]^2}{2hx_n(1-x_n)}\right)$$

The maximum likelihood estimates  $\hat{\alpha}, \hat{\beta}$  of the parameters  $\alpha, \beta$  are obtained setting to zero the partial derivatives of the log-likelihood function with respect to its arguments. In this way, we obtain:

$$\begin{cases} \hat{\alpha} = \frac{1}{2h} \frac{CD - EA}{A^2 - CB} \\ \hat{\beta} = \frac{1}{hA} \left( D + \frac{B(CD - EA)}{A^2 - CB} \right) \end{cases}$$

where

$$\begin{cases} A = \sum_n \frac{1}{1-x_n} \\ B = \sum_n \frac{x_n}{1-x_n} \\ C = \sum_n \frac{1}{x_n(1-x_n)} \\ D = \sum_n \frac{x_{n+1} - x_n}{1-x_n} \\ E = \sum_n \frac{x_{n+1} - x_n}{x_n(1-x_n)} \end{cases}$$

Finally, from  $\hat{\alpha}, \hat{\beta}$ , the estimates  $\hat{p}, \hat{\Delta p}$  of the parameters  $p$  and  $\Delta p$  are easily recovered, in the approximation  $N$  large, by using the relations  $\beta = \lim N(p(N) - \Delta p(N))$  and  $2\alpha = -\lim N(2\Delta p(N) - 1)$ .

The goodness of the estimates has been checked by using simulated data obtained with given input parameters and, indeed, very satisfactory estimates have been found. Moreover, in the cases of examples 4.2 and 4.4, the method was successfully applied to data concerning real proteins (see [4, 7]).

### 6. CONCLUSION AND FINAL REMARKS

In this paper we have studied the continuous diffusion approximation of certain discrete Markov chains, with state space  $\{0, 1, \dots, N\}$ , of interest in a variety of applications ranging from Physics, Biology and Engineering to Mathematical Finance. As told in the Introduction, because of the computational complexity to get the transition probabilities of such a Markov chain for  $N$  large, it is impractical to study the MC directly, so it is convenient to approximate it by a diffusion process. Indeed, the simplest diffusion process is Brownian motion; the analogue of Brownian motion process in discrete times is the Random Walk model. Brownian motion can be obtained as the limit of Random Walks, when the displacement size and the time step get smaller and

smaller. In analogous way, the limit as  $N \rightarrow \infty$  of the scaled MC here considered turns out to be a diffusion process.

A model for pure chance is given by an ideal coin tossed with equal probability for the heads and tails to come up. Let us consider a random variable  $\zeta$  taking value +1 (heads) and -1 (tails) with probability  $1/2$ . If the coin is tossed  $n$  times then the sequence of random variables  $\zeta_1, \zeta_2, \dots, \zeta_n$  describes this experiment. All  $\zeta_k$  have exactly the same distribution as  $\zeta_1$ , moreover they are all independent. Random walk is the process  $Z_n$  defined by  $Z_0 = 0$  and  $Z_n = \zeta_1 + \zeta_2 + \dots + \zeta_n$ . For instance,  $Z_n$  can represent the fortune of a player in a game of chance after  $n$  plays, in which a coin is tossed and one wins 1\$ if heads come up and loses 1\$ when tails come up. Random walk is the central model for stock price in which the standard assumption is that returns on stocks follow a random walk.

Since  $E(\zeta_k) = 0$ , and  $Var(\zeta_k) = E(\zeta_1^2) = 1$ , the mean value and the variance of the random walk are given by:

$$\begin{aligned} E(Z_n) &= E(\zeta_1 + \zeta_2 + \dots + \zeta_n) = E(\zeta_1) + E(\zeta_2) + \dots + E(\zeta_n) = 0 \\ Var(Z_n) &= Var(\zeta_1 + \zeta_2 + \dots + \zeta_n) = \\ &= Var(\zeta_1) + Var(\zeta_2) + \dots + Var(\zeta_n) = n Var(\zeta_1) = n \end{aligned}$$

since the variance of a sum of independent random variables equals the sum of variances.

By scaling the Random walk and taking the limit as the time step and the displacement size go to zero, one obtains Brownian motion. While in the random walk model, at any time step the system can pass from the position  $x$  to  $x+1$  or  $x-1$  with probability  $1/2$ , for the MCs considered in the present paper a transition from a state  $n \in \{0, 1, \dots, N\}$  to a state  $m \in \{0, 1, \dots, N\}$  is allowed with transition probability  $p_{nm}$  given by (1.1). The rescaled MC  $X_k/N$  converges, as  $N \rightarrow \infty$ , to a continuous process, slightly more complicated than Brownian motion, that is a diffusion with mean displacement (drift)  $b(x)$  and infinitesimal variance  $x(1-x)$ . Actually, the drift  $b(x)$  turns out to be a polynomial in the variable  $x$ .

By means of the diffusion approximation, we have analyzed the qualitative behavior of the MC. Moreover, we have investigated the stationary distribution, attainability of boundary and estimation of parameters which intervene in the description of the MC. We have also reported several examples of models for generalized growth phenomena in Biology and Biophysics, such as population genetics and cooperative evolution of protein macromolecules. However, the techniques and methods here used can be applied to any other field of investigation, when a description in terms of a continuous diffusion approximation of this kind, is appropriate.

### REFERENCES

- [1] Ricciardi LM. Diffusion processes and related topics in biology. Lecture Notes in Biomathematics. Springer-Verlag, Berlin-Heidelberg 1977; vol. 14.
- [2] Yamada T, Watanabe S. On the uniqueness of solutions of stochastic differential equations. J Math Kyoto Univ 1971; 11: 155-67.

- [3] Abundo M, Baldi P, Caramellino L. A diffusion approximation which models hierarchic interactions in cooperative biological systems. *Open Syst Inf Dyn* 1998; 5: 1-23.
- [4] Abundo M, Accardi L, Rosato N. A Markovian model for cooperative biological systems. *Math Models Meth Appl Sci* 1995; 6(5): 835-63.
- [5] Abundo M, Caramellino L. Some remarks about a Markov chain which models cooperative biological systems. *Open Syst Inf Dyn* 1995; 3 (3): 325-43.
- [6] Ethier SN, Kurtz TG. *Markov processes. Characterization and convergence.* New York, Wiley & Sons 1986.
- [7] Abundo M, Accardi L, Rosato N. A stochastic model for cooperative interactions in proteins, based on a hierarchy of bonds between aminoacidic residues. *Math Models Meth Appl Sci* 1998; 8 (2): 327-58.

---

Received: January 01, 2009

Revised: March 03, 2009

Accepted: March 03, 2009

© Mario Abundo; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.