Full Randomness in the Higher Difference Structure of Two-State Markov Chains

Ashot Yu. Shahverdian
Institute for Informatics and Automation Problems of NAS RA
Yerevan, Armenia
e-mail: svrdn@yerphi.am

ABSTRACT

The paper studies the higher-order absolute differences taken from progressive terms of time-homogeneous binary Markov chains. Two theorems presented are the limiting theorems for these differences, when their order k converges to infinity. Theorems 1 and 2 assert that there exist some infinite subsets E of natural series such that kth order differences of every such chain converge to the equi-distributed random binary process as k growth to infinity remaining on E. The chains are classified into two types, and E depends only on the type of the given chain. Two kinds of discrete capacities for subsets of natural series are defined, and in their terms such sets E are described.

Keywords

Markov chain, higher-order absolute difference, discrete capacity, randomness

1. INTRODUCTION

In this paper an application of the suggested in [1]-[8] difference analysis to studying binary Markov chains is presented. In difference analysis we are interested in the following question: which is the higher-order difference structure of a given process and how this structure can characterize the process.

The paper studies the time-homogenous binary Markov chains $\boldsymbol{\xi} = (\xi_n)_{n\geq 0}$ where every ξ_n is binary variable describing the state of the cain $\boldsymbol{\xi}$ at the moment n. The main results, Theorem 1 and Theorem 2, are limiting theorems for such chains. These theorems concern infinite sets $E \subseteq \mathbb{N}$ which posses such a property: for arbitrary chain $\boldsymbol{\xi}$, the E permits the existence of the limit of kth order absolute differences $\boldsymbol{\xi}^{(k)}$, when k converges to ∞ remaining on E. The existence of such E is claimed, and their description in capacity terms is given. The chains are classified into two types, and E depends only on the type of the given chain. The limiting process is the equi-distributed random binary sequence, denoted $\boldsymbol{\theta}$ (see Eq. (3)).

Theorems 1 and 2 improve our previous results from [7, 8]; some details on this matter in Section 3 (points *(a) - (c)*) are given. The limiting process, which is the equi-distributed sequence, should be recognized as the most random binary sequence. Therefore, theorems presented state the existence of full randomness in the higher difference structure of an arbitrary time-homogeneous binary Markov chain.

Let us explain our statement in more detail. Let

$$\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_n, \ldots)$$

be some random sequence whose components ξ_n take binary values $x \in X$, $X = \{0, 1\}$ with some positive probabilities, $P(\xi_n = x) = p_n(x)$. Then kth order ($k \geq 0$) absolute differences $\xi_n^{(k)}$, which are defined recurrently,

$$\xi_n^{(0)} \equiv \xi_n \quad \text{and} \quad \xi_n^{(k)} = |\xi_{n+1}^{(k-1)} - \xi_n^{(k-1)}| \quad (n \geq 0),$$

also take binary values with some probabilities $P(\xi_n^{(k)} = x) = p_n^{(k)}(x)$, and hence, one can consider kth order difference random binary sequence

$$\boldsymbol{\xi}^{(k)} = (\xi_0^{(k)}, \xi_1^{(k)}, \ldots, \xi_n^{(k)}, \ldots).$$

Our interest is the limits of $\boldsymbol{\xi}^{(k)}$ when k goes to infinity. Let some infinite $E \subseteq \mathbb{N}$ be given. We say that $\boldsymbol{\xi}^{(k)}$ converge on E to a random binary sequence $\boldsymbol{\xi}_E^\infty$, and denote this

$$\boldsymbol{\xi}_E^\infty = \lim_{\substack{k\to\infty \\ k\in E}} \boldsymbol{\xi}^{(k)},$$

if for $n \in \mathbb{N}$ and $x \in X$ the probabilities $p_n^{(k)}(x)$ tend to some numbers $p_n^{(\infty)}(x)$ as $k \to \infty$ and $k \in E$,

$$\lim_{\substack{k\to\infty \\ k\in E}} p_n^{(k)}(x) = p_n^{(\infty)}(x)$$

(*convergence by probability* on E and *partial limits*). Therefore, $\boldsymbol{\xi}_E^\infty$ is a random binary sequence,

$$\boldsymbol{\xi}_E^\infty = (\xi_0^{(\infty)}, \xi_1^{(\infty)}, \ldots, \xi_n^{(\infty)}, \ldots)$$

whose components $\xi_n^{(\infty)}$ take the values $x \in X$ with probabilities $P(\xi_n^{(\infty)} = x) = p_n^{(\infty)}(x)$ (which depend on E).

We consider binary Markov chains $\boldsymbol{\xi} = (\xi_n)_{n=0}^\infty$ whose state space X consists of two binary symbols, $X = \{0, 1\}$. We assume that the chains $\boldsymbol{\xi}$ are time-homogeneous, that is, for x, x_i, $y \in X$

$$P(\xi_n = y|\xi_{n-1} = x, \xi_{n-2} = x_1, \ldots, \xi_0 = x_{n-1}) \\ = P(\xi_n = y|\xi_{n-1} = x)$$

(Markov property) and there is some function $\pi(x, y)$ on $X \times X$ such that for $n \geq 1$ and $x, y \in X$

$$P(\xi_n = y|\xi_{n-1} = x) = \pi(x, y)$$

(homogeneity: one-step transition probabilities $P(\xi_n = y|\xi_{n-1} = x)$ do not depend on time n). It is also assumed that some initial distribution of probabilities $P(\xi_0 = x)$ on X is given. In what follows it is always assumed that $\boldsymbol{\xi}$ denotes the time-homogeneous binary Markov chain.

Some simple computations testify, that if for the given $\boldsymbol{\xi}$ an infinite $E \subseteq \mathbb{N}$ is chosen arbitrarily, then the limiting process $\boldsymbol{\xi}_E^\infty$ may not exist. On the other hand, it follows from [7, 8] that for $E = \{2^m - 1 : m \geq 0\}$ and large collection of $\boldsymbol{\xi}$, the limit $\boldsymbol{\xi}_E^\infty$ exists and it is the equi-distributed random sequence. The problem studied relates to the following question: how can the sets $E \subseteq \mathbb{N}$, which for arbitrary Markov chain $\boldsymbol{\xi}$ permit the existence of $\boldsymbol{\xi}_E^\infty$, be described?

The main results of this paper, Theorems 1 and 2, are limiting theorems for such chains. Two discrete capacities for subsets of $\mathbb{N}$ are defined, and in their terms such sets E are described. The limiting process $\boldsymbol{\xi}_E^\infty$, the existence of which assert these theorems, is the equi-distributed random binary sequence.

The transition matrix of any time-homogeneous binary Markov chain $\boldsymbol{\xi}$ can be written as

$$Q_{\boldsymbol{\xi}} = \begin{bmatrix} s & 1-s \\ 1-p & p \end{bmatrix} \qquad (0 < s, p < 1)$$

where $s = \pi(0,0)$ and $p = \pi(1,1)$. Theorems 1 and 2 consider two types of chains $\boldsymbol{\xi}$, depending on which of the next two relationships (I) and (II) between s and p

$$\textbf{(I)}\ \ s \neq p \ \text{ and } \ s \neq 1-p, \quad \textbf{(II)}\ \ s = p \ \text{ or } \ s = 1-p \quad (1)$$

holds: we say that the chain $\boldsymbol{\xi}$ is of I-st or II-nd type whenever for s and p the relations (I) or (II) (respectively) from Eq. (1) are satisfied.

The paper consists of four sections. The next Section 2 contains definitions of discrete capacities that we use. In Section 3 the formulations of main Theorems 1 and 2 are presented, and last Section 4 contains some additional comments.

2. SOME DEFINITIONS

To proceed to formulation of our Theorems 1 and 2, we need to present two discrete capacities $\mathcal{C}$ and $\boldsymbol{c}$ defined for subsets of natural series $\mathbb{N}$. Their definition is given by means of binary representation of natural numbers and binary version of Pascal triangle. The binary Pascal triangle $\mathbb{P}$ and its kth line $\boldsymbol{\ell}_k$ are defined as

$$\mathbb{P} = \{\alpha_{k,i} : k \geq 0, 0 \leq i \leq k\}, \ \boldsymbol{\ell}_k = (\alpha_{k,0}, \alpha_{k,1}, \ldots, \alpha_{k,k})$$

that is, $\mathbb{P} = \bigcup_{k=0}^\infty \boldsymbol{\ell}_k$; here, $\alpha_{k,i} \in \{0,1\}$ are the following: $\alpha_{0,0} = 1$ (the vertex of $\mathbb{P}$ and the line $\boldsymbol{\ell}_0$), $\alpha_{1,0} = \alpha_{1,1} = 1$ (the line $\boldsymbol{\ell}_1$), and for $k \geq 2$ the line $\boldsymbol{\ell}_k$ consists of such $\alpha_{k,i}$,

$$\alpha_{k,i} = \begin{cases} 0, & \binom{k}{i} \text{ is even} \\ 1, & \binom{k}{i} \text{ is odd} \end{cases} \qquad (0 \leq i \leq k).$$

One can see that this is the same as if for $k \geq 1$ one defines: $\alpha_{k,0} = \alpha_{k,k} = 1$, and

$$\alpha_{k,i} = |\alpha_{k-1,i-1} - \alpha_{k-1,i}| \qquad (1 \leq i \leq k-1).$$

The capacities $\mathcal{C}$ and $\boldsymbol{c}$ are defined by means of some quantities related to binary expansion of natural numbers. For $k \geq 1$ its binary representation is given as

$$k = \sum_{i=0}^{p} \varepsilon_i 2^i \quad \text{where } \ p \geq 0, \ \ \varepsilon_i \in \{0,1\}, \ \ \varepsilon_p = 1; \quad (2)$$

we denote

$$b(k) = \sum_{i=0}^{p} \varepsilon_i, \quad \beta(k) = \sum_{i=0}^{k} \alpha_{k,i}.$$

For natural k we use the following notations: $\nu(k)$ denotes the maximal of such m, $0 \leq m \leq p$ for which all the coefficients ε_i, $0 \leq i \leq m$ in expansion (2) are equal to 1,

$$\nu(k) = \max\{m : \varepsilon_0 = \varepsilon_1 = \cdots = \varepsilon_m = 1\};$$

$\mu(k)$ denotes the maximal of such m, $0 \leq m \leq k$ for which all the $\alpha_{k,i}$, $0 \leq i \leq m$ (first m entries of the line $\boldsymbol{\ell}_k$ of the triangle $\mathbb{P}$) are equal to 1,

$$\mu(k) = \max\{m : \alpha_{k,0} = \alpha_{k,1} = \cdots = \alpha_{k,m} = 1\}.$$

The capacities $\mathcal{C}$ and $\boldsymbol{c}$ are assigned on the collection $2^{\mathbb{N}}$ of subsets of natural series and defined as follows:

Definition 1. *For $e \subseteq \mathbb{N}$ we define*

$$\mathcal{C}(e) = \sum_{k \in e} \nu(k), \quad \boldsymbol{c}(e) = \sum_{k \in e} b(k).$$

The $\mathcal{C}(e)$ and $\boldsymbol{c}(e)$ can be expressed by the entries of the Pascal triangle $\mathbb{P}$: one can prove that $\mu(k) = 2^{\nu(k)}$ and $\beta(k) = 2^{b(k)}$, and, therefore,

$$\mathcal{C}(e) = \sum_{k \in e} \log_2 \mu(k), \quad \boldsymbol{c}(e) = \sum_{k \in e} \log_2 \beta(k).$$

Both $\mathcal{C}$ and $\boldsymbol{c}$ differ from discrete capacity, considered in denumerable Markov chains and random walk (e.g., [9]); for details on $\mathcal{C}$ and $\boldsymbol{c}$ see [4] and [8]. We denote $\mathcal{C}(k) = \mathcal{C}(\{k\})$ and $\boldsymbol{c}(k) = \boldsymbol{c}(\{k\})$.

Let us present an example of computation of these capacities. We denote $L_p = \{k \in \mathbb{N} : 2^{p-1} \leq k < 2^p\}$ and for $p \geq 2$ and $0 \leq s \leq p$ consider the sets $B_p(s)$ and $b_p(s)$:

$$B_p(s) = \{k \in L_p : \nu_k \geq s\}, \ b_p(s) = \{k \in L_p : b(k) \geq s\}.$$

The complement of $b_p(s)$ is the Hamming ball of radius s ([4]; there is a misprint in [4] on computation of capacity of these balls).

Proposition 1. *For $p \geq 2$ and $0 \leq s \leq p$ the relations*

$$\mathcal{C}(B_p(s)) = \sum_{i=0}^{s} i 2^{p-i}, \quad \boldsymbol{c}(b_p(s)) = \sum_{i=s}^{p} i \binom{p}{i}$$

are true.

3. MAIN THEOREMS

In this section we formulate our main results, Theorems 1 and 2. They define some sets $E \subseteq \mathbb{N}$ and state the convergence of kth order difference processes $\boldsymbol{\xi}^{(k)}$ (as $k \to \infty$ and $k \in E$) to the equi-distributed random binary sequence $\boldsymbol{\theta}$; the $\boldsymbol{\theta}$ is defined as

$$\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_n, \ldots) \quad \text{where} \quad P(\theta_n = x) = \frac{1}{2} \quad (3)$$

for all $n \geq 0$ and $x \in \{0,1\}$. In next formulations $o_k(1)$ denotes the Landau symbol: it is some numerical quantity which tends to 0 as k converges to ∞.

Theorem 1 and Theorem 2, formulated in the next subsections, improve some our results from [7, 8]. If compared with [7, 8], the improvement is due to the following three features of Theorems 1 and 2: *(a)* the sets E in formulations of these theorems depend only on the type (I-st or II-nd type) of the chain $\boldsymbol{\xi}$ and do not depend on other details uniquely determining the given chain; *(b)* the theorems estimate the rate (exponential) of the convergence; *(c)* a different description of sets E (Eqs. (4) and (8)) is given.

In the next Sections 3.1 and 3.2 we present some examples of such sets E (Eqs. (6) and (10)). These examples appear to be quite general and connect us (Propositions 2 and 4) with another, considered in [8], description of these sets. In addition, Remarks 1 and 2 state that the sets E from these examples are the 'largest' ones, satisfying the assumptions (4) and (8) in these theorems. This allows us to derive some conclusions (Propositions 3 and 5) on densities of sets E from Theorems 1 and 2.

3.1 Chains of I-st type

Let us formulate our Theorem 1 which concerns Markov chains of the I-st type (defined by Eq. (1)). This theorem describes infinite sets $E \subseteq \mathbb{N}$ which possess the property that the limiting processes $\boldsymbol{\xi}_E^\infty$ exists for arbitrary Markov chain $\boldsymbol{\xi}$ of the I-st type: the theorem asserts the convergence of kth order difference processes $\boldsymbol{\xi}^{(k)}$ (as $k \to \infty$ and $k \in E$) to the equi-distributed process $\boldsymbol{\theta}$ (defined by Eq. (3)).

Theorem 1. *Let a set $E \subseteq \mathbb{N}$ be such that*

$$\lim_{\substack{k\to\infty \\ k\in E}} \mathcal{C}(k) = \infty \tag{4}$$

and $\boldsymbol{\xi}$ be Markov chain of the I-st type. Then the limiting process $\boldsymbol{\xi}_E^\infty$ exists and $\boldsymbol{\xi}_E^\infty = \boldsymbol{\theta}$, that is,

$$\lim_{\substack{k\to\infty \\ k\in E}} \boldsymbol{\xi}^{(k)} = \boldsymbol{\theta}. \tag{5}$$

The convergence in Eq. (5) is exponential: given $\boldsymbol{\xi}$ there is some δ, $|\delta| < 1$ which depends only on transition matrix of $\boldsymbol{\xi}$, such that for $n \geq 1$, $k \in E$ and $\lambda \in \{0,1\}$ the relation

$$P(\xi_n^{(k)} = \lambda) = \frac{1}{2} + o_k(1)\delta^k$$

holds.

Let us present some examples of sets $E \subseteq \mathbb{N}$ satisfying Eq. (4). With this aim we consider the unions of $B_p(s_p)$,

$$E = \bigcup_{p=1}^{\infty} B_p(s_p). \tag{6}$$

Proposition 2. *Let $E \subseteq \mathbb{N}$ be defined by Eq. (6). Then E satisfies Eq. (4) if and only if the conditions*

$$\lim_{p\to\infty} s_p = \infty \quad \textit{and} \quad \sum_{p=1}^{\infty} 2^{-p}\mathcal{C}(B_p(s_p)) = \infty \tag{7}$$

hold.

The next Remark 1 asserts that the given by Proposition 2 example of sets E satisfying Eq. (4) is quite general.

Remark 1. *For a set $E \subseteq \mathbb{N}$ the condition (4) holds if and only if there is a set $E' \subseteq \mathbb{N}$ of the form (6) satisfying (4) and such that $E \subseteq E'$.*

We describe the density of sets E from Theorem 1. For $m \geq 1$ we denote $E_m = \{k \in E : 1 \leq k \leq m\}$, consider the ratio $\rho_m(E) = \dfrac{|E_m|}{m}$, where $|E_m|$ denotes the cardinality of E_m, and define

$$dens(E) = \lim_{m\to\infty} \rho_m(E).$$

Proposition 3. *If a set $E \subseteq \mathbb{N}$ satisfies Eq. (4), then $dens(E) = 0$. For a given $0 < \delta_m \leq 1$, $\delta_m \downarrow 0$ there is a set $E \subseteq \mathbb{N}$ which satisfies Eq. (4) and such that $\rho_m(E) \geq \delta_m$ for all $m \geq 1$.*

3.2 Chains of II-nd type

The next Theorem 2 concerns Markov chains of the II-nd type and describes infinite sets $E \subseteq \mathbb{N}$, which possess the property that the limiting processes $\boldsymbol{\xi}_E^\infty$ exist for arbitrary Markov chains $\boldsymbol{\xi}$ of the II-nd type; the limiting process is again the equi-distributed process $\boldsymbol{\theta}$.

Theorem 2. *Let a set $E \subseteq \mathbb{N}$ be such that*

$$\lim_{\substack{k\to\infty \\ k\in E}} \boldsymbol{c}(k) = \infty \tag{8}$$

and $\boldsymbol{\xi}$ be Markov chain of the II-nd type. Then the limiting process $\boldsymbol{\xi}_E^\infty$ exists and $\boldsymbol{\xi}_E^\infty = \boldsymbol{\theta}$, that is,

$$\lim_{\substack{k\to\infty \\ k\in E}} \boldsymbol{\xi}^{(k)} = \boldsymbol{\theta}. \tag{9}$$

The convergence in Eq. (9) is exponential: given $\boldsymbol{\xi}$ there is some δ, $|\delta| < 1$ which depends only on transition matrix of $\boldsymbol{\xi}$, such that for $n \geq 1$, $k \in E$ and $\lambda \in \{0,1\}$ the relation

$$P(\xi_n^{(k)} = \lambda) = \frac{1}{2} + o_k(1)\delta^k$$

holds.

As the examples of sets $E \subset \mathbb{N}$ satisfying Eq. (8) we consider the unions of $b_p(s_p)$,

$$E = \bigcup_{p=1}^{\infty} b_p(s_p). \tag{10}$$

Proposition 4. *Let $E \subseteq \mathbb{N}$ be defined by Eq. (10). Then E satisfies Eq. (8) if and only if the conditions*

$$\lim_{p\to\infty} s_p = \infty \quad \textit{and} \quad \sum_{p=1}^{\infty} 2^{-p}\boldsymbol{c}(b_p(s_p)) = \infty \tag{11}$$

hold.

Remark 2. *For a set $E \subseteq \mathbb{N}$ the condition (8) holds if and only if there is a set $E' \subseteq \mathbb{N}$ of the form (10) satisfying (8) and such that $E \subseteq E'$.*

We compute the density of sets E from Eq. (11):

Proposition 5. *If a set $E \subseteq \mathbb{N}$ defined by Eq. (10) satisfies Eq. (11), then $dens(E) = 1$.*

Particularly, Propositions 4 and 5 imply that the sets E from Theorem 2 can be as 'large', that their density equals 1.

4. SOME COMMENTS

The chains considered can be treated as two-state probabilistic automata.

In [4] independent random sequences have been studied (there is an unnecessary (and wrong) assumption in [4] on independence of $\boldsymbol{\xi}^{(k)}$). Theorem 2 remains valid also for arbitrary independent identically distributed binary sequences.

The capacities $\mathcal{C}$ and $\boldsymbol{c}$ are some 'discrete' instances of the Fuglede-Choquet capacities [10, 11], which are the abstract version of classical capacities (e.g., [12]). Another kind of discrete capacity, applied to a self-organized criticality model [13], is considered in [14].

The second relations in (7) and (11) are the analogs for the Wiener criterion from potential theory (e.g., [15, 16, 17]; the sets E from (6) and (10), satisfying these relations, are called thick sets. Apparently, the most known application of thick sets in classical theory is given by Keldysh theorem on the Dirichlet problem [17].

One of the basic concepts of ergodic theory is the notion of shift in probabilistic spaces [18]. E.g., independent sequences and Markov chains can be treated as consecutive iterates of some ergodic shifts (Bernoulli and Markov shifts [18]). In [4] we have defined the difference shift M; it is such, that kth order absolute difference $\boldsymbol{\xi}^{(k)}$ coincides with kth iterate M^k of the shift M applied to the random sequence $\boldsymbol{\xi}$, $\boldsymbol{\xi}^{(k)} = M^k\boldsymbol{\xi}$. Therefore, Theorems 1 and 2 can also be treated as some statements on iterates of the difference shift M.

REFERENCES

[1] A. Yu. Shahverdian, A. V. Apkarian, *On irregular behavior of neural spike trains*, Fractals, 7(1), 93-103, 1999.

[2] A. Yu. Shahverdian, *The finite-difference method for analyzing one-dimensional nonlinear systems*, Fractals, 8(1), 49-65, 2000.

[3] A. Yu. Shahverdian, A. V. Apkarian, *A difference characteristic for one-dimensional nonlinear systems*, Comm. Nonlin. Sci. & Comput. Simul., 12, 233-242, 2007.

[4] A. Yu. Shahverdian, *Minimal Lie algebra, fine limits, and dynamical systems*, Reports Armenian Natl. Acad. Sci., 112(2), 160-169, 2012.

[5] A. Yu. Shahverdian, A. Kilicman, R. B. Benosman *Higher difference structure of some discrete processes*, Adv. Difference Equations, 202, 1-10, 2012.

[6] A. Yu. Shahverdian, R. P. Agarwal, R. B. Benosman, *The bistability of higher-order differences of periodic signals*, Adv. Difference Equations, 60, 1-9, 2014.

[7] A. Yu. Shahverdian, *A theorem on higher-order differences of two-state Markov chains*, Proc. Intern. Conf. CSIT-2015. Yerevan, 251-252, 2015 (reprinted in: IEEE Conference Ser., CSIT-2015, 137-138, 2015).

[8] A. Yu. Shahverdian, *Discrete capacity and higher-order differences of two-state Markov chains*, Reports Armenian Natl. Acad. Sci., 116(3), 195-201, 2016.

[9] E. B. Dynkin, A. A. Yushkevich, *Markov processes: Theorems and Problems*, Plenum Press, New York, 1969.

[10] M. Brelot, *Eléments de la Théorie Classique du Potentiel*, Paris, 1969.

[11] G. Choquet, *Theory of Capacities*, Ann. Inst. Fourier, 5, 131-295, 1955.

[12] L. Carleson, *Selected Problems on Exceptional Sets*, Van Nostrand, Princeton, 1967.

[13] A. Yu. Shahverdian, A. V. Apkarian, *Avalanches in networks of weakly coupled phase shifting rotators*, Comm. Math. Sci., 6(1), 217-234, 2008.

[14] A. Yu. Shahverdian, *Avalanches and memory in rotator networks*, Reports Armenian Natl. Acad. Sci., 111(3), 240-249, 2011.

[15] N. Wiener, *The Dirichlet problem*, J. Math and Phys., 3, 123-146, 1924.

[16] A. Yu. Shahverdian, *Fine topology and estimates for potentials and subharmonic functions*, Computational Methods and Function Theory, 11(1), 71-121, 2011.

[17] M. V. Keldysh, *On the solvability and the stability of the Dirichlet problem*, Amer. Math. Soc. Transl. (2), 51, 1-73, 1966.

[18] P. Billingsley, *Ergodic Theory and Information*, Wiley, New York, 1965.