

The Cut-Off Phenomenon for Random Walks on Hamming Graphs with Variable Growth Conditions

Dedicated to Professor Takeshi Hirai on his 60th birthday

By

Akihito HORA*

Abstract

The cut-off phenomenon is a remarkable critical phenomenon observed in the process of convergence to equilibrium in a wide variety of Markov chains. Diaconis-Graham-Morrison [3] established the first precise evaluation around the critical time for Ehrenfests' urn model concerning 2 urns and d balls, i.e. nearest neighbor random walks on hypercube \mathbf{Z}_2^d . They showed that deviation from the equilibrium state is described well by using the error function. In this article, we work out the evaluation around the critical time for simple random walks on Hamming graphs $H(d, n)$, which coincide with an extended Ehrenfests' urn model concerning n urns and d balls. In our case, not only d but also n can grow in several manners. If n/d tends to 0, the similar result to [3] remains valid and microscopic deviation from the equilibrium state is described by the error function. If n/d tends to a nonzero constant, however, it is shown that the error function has to be replaced by an expression involving Poisson distributions.

§1. Introduction

The *cut-off phenomenon* is observed in the process of convergence to equilibrium in a wide variety of Markov chains including certain shuffling and diffusions. Taking a viewpoint of statistical mechanics into account, we can regard it as a critical phenomenon owing to the huge cardinality of a state space. Lots of articles dealing with this interesting phenomenon are available at the present time. Among them we here refer to [2] as a concise summary with bibliographical information.

Let us consider a Markov chain on a finite state space X with transition matrix P and invariant probability π . Under mild conditions, we see the

Communicated by Y. Takahashi, January 20, 1997. Revised April 30, 1997.

1991 Mathematics Subject Classifications: 60J15, 60C05, 82C41, 05E30.

*Department of Environmental and Mathematical Sciences, Faculty of Environmental Science and Technology, Okayama University, Okayama 700, Japan.

convergence to equilibrium of the chain:

$$(P^k)_{x,y} \rightarrow \pi(y) \quad (k \rightarrow \infty) \quad \text{for } x, y \in X.$$

To analyze the manner of the convergence in more detail, we consider

$$(1) \quad \|(P^k)_{x,\cdot} - \pi\|_{\text{var}} = \frac{1}{2} \sum_{y \in X} |(P^k)_{x,y} - \pi(y)| \quad (\text{total variation distance})$$

as a quantity to measure closeness to the equilibrium at time k . In this article, we assume some spatial symmetry of the chain, which implies $\pi(y) = 1/|X|$ and (1) is independent of x . Thus we set

$$(2) \quad D(k) = \frac{1}{2} \sum_{y \in X} |(P^k)_{x,y} - \frac{1}{|X|}| = \frac{1}{2|X|} \sum_{x,y \in X} |(P^k)_{x,y} - (E_0)_{x,y}|$$

where E_0 is the matrix whose entries are all $1/|X|$. We are interested in the decay rate of the function $D(k)$. Like many other critical phenomena, the cut-off phenomenon is clearly captured through a certain infinite volume limit of a family of Markov chains.

Let us consider a family of nearest neighbor random walks on hypercube \mathbb{Z}_2^d indexed by $d \in \mathbb{N}$, where each chain either moves to one of the nearest vertices or stays at the present vertex with equal probability $1/(d+1)$ in one unit time. This is the Ehrenfests urn model with 2 urns and d balls. Diaconis, Graham and Morrison [3] showed the following exact asymptotic result. We set $D^{(d)}(k)$ as in (2).

Theorem 0. (*Diaconis-Graham-Morrison*) *In a nearest neighbor random walk on \mathbb{Z}_2^d ,*

$$D^{(d)}\left(\left[\frac{1}{4}d \log d + \theta d\right]\right) \longrightarrow \text{Erf}\left(\frac{e^{-2\theta}}{2\sqrt{2}}\right) \quad (d \rightarrow \infty) \quad \text{for } \theta \in \mathbb{R}.$$

Here, *Erf* is the error function:

$$\text{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

In [7], Voit derived Theorem 0 by using a different strategy from [3].

The purpose of this article is to establish an exact asymptotic result for random walks on Hamming graph $H(d, n)$, which is an extended Ehrenfests

urn model with n urns and d balls. The hypercube Z_2^d appears as $H(d, 2)$. The point here is that not only d but also n can grow. The following are our results. We set $D^{(d,n)}(k)$ as in (2) for a simple random walk on $H(d, n)$.

Theorem 1. *Let τ be a positive constant. For $\theta \in \mathbf{R}$, as $d \rightarrow \infty$ and $n/d \rightarrow \tau$,*

$$D^{(d,n)}\left(\left[\frac{1}{2}\left(1-\frac{1}{n}\right)d(\log(n-1)d+\theta)\right]\right) \rightarrow \|p_{1/\tau} - P_{(1/\tau) + (e^{-\theta/2}/\sqrt{\tau})}\|_{\text{var}}$$

compact-uniformly in θ , where p_α denotes the Poisson distribution with intensity α .

Theorem 2. *Let $n \geq 3$ (not necessarily fixed nor bounded). For $\theta \in \mathbf{R}$, as $d \rightarrow \infty$ and $n/d \rightarrow 0$,*

$$D^{(d,n)}\left(\left[\frac{1}{2}\left(1-\frac{1}{n}\right)d(\log(n-1)d+\theta)\right]\right) \rightarrow \text{Erf}\left(\frac{e^{-\theta/2}}{2\sqrt{2}}\right)$$

compact-uniformly in θ .

Remark. Restriction $n \geq 3$ in Theorem 2 is to avoid periodicity of the simple random walk. For the simple random walk on $H(d, 2)$ ($=Z_2^d$), (2) does not converge to zero as $k \rightarrow \infty$ since $H(d, 2)$ is a bipartite graph. Hence, one needs a modification: e.g. to consider nearest neighbor random walks as in [3] (above Theorem 0) or to treat even and odd times separately as in [7], Theorem 1.5.

Now we supplement explanation of why the above theorems show us a critical phenomenon and how they describe fine behavior of the Markov chain near the critical time. Let us consider the situation of Theorem 1 or Theorem 2. Note that the limit function, say denoted by $c(\theta)$, satisfies $0 \leq c(\theta) \leq 1$, $c(-\infty) = 1$ and $c(\infty) = 0$. For arbitrarily given (very small) $\varepsilon > 0$, take $\theta_\varepsilon > 0$ such that

$$c(\theta) \begin{cases} \geq 1 - \varepsilon & \text{if } \theta \leq -\theta_\varepsilon \\ \leq \varepsilon & \text{if } \theta \geq \theta_\varepsilon. \end{cases}$$

We set $k_c^{(d,n)} = (1 - (1/n))(d/2)\log(n-1)d$ and $h^{(d,n)} = (1 - (1/n))d/2$. Theorem 1 and Theorem 2 assure that $D^{(d,n)}([k_c^{(d,n)} + \theta h^{(d,n)}])$ is almost 1 for $\theta \leq -\theta_\varepsilon$ and is almost 0 for $\theta \geq \theta_\varepsilon$. Note that $h^{(d,n)} \ll k_c^{(d,n)} \ll |X^{(d,n)}|$ (=the cardinality of the

state space $=n^d$). In our interpretation, $k_c^{(d,n)}$ is regarded as a macroscopic time. Then, $|\mathcal{X}^{(d,n)}|$ is almost infinite and deviation $h^{(d,n)}$ is macroscopically negligible in the infinite volume limit $(d,n) \rightarrow \infty$. Here θ plays the role of a microscopic time parameter. One thus sees that the value of $D^{(d,n)}(k)$ drops suddenly from 1 to 0 quite near the critical time $k_c^{(d,n)}$ for large (d,n) and that limit function $c(\theta)$ describes the manner in which the chain approaches equilibrium around the critical time in a microscopic time scale.

In [2], Section 4, Diaconis proposed a general understanding of what causes the cut-off phenomenon. The essential role is played by high multiplicity of the second eigenvalue of a transition matrix, which is due to some symmetry (like Assumption in §2). In order to perform a precise proof of the cut-off phenomenon, however, one needs more detailed information of the model discussed. (In [5] and [6], we proposed a definition of the cut-off phenomenon and gave some general criteria in terms of the spectrum and what is called a Krein parameter.) It would be worth while to work out a careful evaluation in concrete models. We also mention [8] dealing with a continuous state space.

Our plan for the proof of Theorem 1 and Theorem 2 is a prolongation of the approach in [7]. By virtue of the very nice structure of the spectrum of $H(d,n)$ (see §2), continuous time simple random walks on $H(d,n)$ are easy to handle. Thus we first discuss the relation between a continuous time case and a discrete time case (Theorem 3 in §3). Next the results in a continuous time case are derived (Theorem 4 in §4 and Theorem 5 in §5). Finally we combine them to obtain Theorem 1 and Theorem 2. §2 is devoted to preliminary computations and a bit general discussions on transition probabilities.

Acknowledgement. The study in this article was motivated by discussions with Prof. M. Voit, to whom I would like to express deep appreciation. I am grateful also to the referee for valuable comments based on careful reading.

§2. Random Walk on a Hamming Graph $H(d,n)$

Explicit functional calculus for a transition operator via spectral decomposition enables us to investigate time evolution of a Markov chain in great detail. If the transition probability enjoys a certain symmetry, one can make use of algebraic structure associated with the state space for that purpose. In this section we perform functional calculus for transition matrices by using algebraic machinery because we feel that such a general treatment would make the process of computation more transparent. After that, we

specialize our formulas onto Hamming graphs. However, if a reader has little interest in algebraic argument, he can skip to formula (11) which is a closed expression of transition probabilities at an arbitrary time of the *continuous time* simple random walk on a Hamming graph. Alternatively, one gets (11) via probabilistic argument by using independence of the coordinate variables as is done in [3], Section 2 for the continuous time Ehrenfests urn model. We refer also to [7], Section 3 for some analytical discussions related to the Fourier-Krawtchouk transformation.

We begin with a quick review of Hamming graphs and some related notions. See [1], Chap. III for details. Let F be an n -set and $X = F^d$ where $d, n \geq 2$. Set $\partial(x, y) = |\{i | x_i \neq y_i\}|$ for $x = (x_i), y = (y_i) \in X$. Distance ∂ induces a relation on $X \times X$ by $(x, y) \in R_i \Leftrightarrow \partial(x, y) = i$ ($i = 0, 1, \dots, d$). Undirected graph (X, R_1) with vertices X and edges R_1 is called Hamming graph $H(d, n)$. Adjacency matrix A_i is defined by

$$(A_i)_{x,y} = \begin{cases} 1 & \text{if } (x, y) \in R_i \\ 0 & \text{if } (x, y) \notin R_i. \end{cases}$$

The complex linear combinations of A_0 (=identity), A_1, \dots, A_d form a commutative algebra \mathcal{A} , which is called the Bose-Mesner algebra of $H(d, n)$. Actually, A_2, \dots, A_d are expressed as polynomials in A_1 since $H(d, n)$ is distance-regular. Set $\kappa_i = |\{y \in X | \partial(x, y) = i\}|$ (the right-hand side being independent of $x \in X$). κ_1 is the degree of each vertex. A_0, A_1, \dots, A_d are simultaneously diagonalized by primitive idempotents E_0, E_1, \dots, E_d in \mathcal{A} . Here E_0 denotes the matrix whose entries are all $1/|X|$. Coefficients $p_i(j)$ and $q_i(j)$ are determined by

$$(3) \quad A_i = \sum_{j=0}^d p_i(j)E_j, \quad |X|E_i = \sum_{j=0}^d q_i(j)A_j.$$

Set $m_i = \text{rank } E_i$. We have $p_i(0) = \kappa_i$ and $q_i(0) = m_i$.

Let us consider a Markov chain on X with transition matrix P . We assume the spatial symmetry of P that it is constant on each orbit R_i :

Assumption. $\partial(x, y) = \partial(x', y') \Rightarrow (P)_{x,y} = (P)_{x',y'}$
 or equivalently that P belongs to Bose-Mesner algebra \mathcal{A} . Then the transition probability takes the form of

$$(4) \quad P = \sum_{i=0}^d \frac{w_i}{\kappa_i} A_i \quad \text{where} \quad w_i \geq 0, \quad \sum_{i=0}^d w_i = 1.$$

In particular, the chain is called a *simple random walk* when $w_1 = 1$. P^k ($k \in \mathbb{N}$) and $e^{t(P-I)}$ ($t \geq 0$) are also constant on each orbit. Let $P_k(h)$ [resp. $P(t, h)$] denote $(P^k)_{x,y}$ [resp. $(e^{t(P-I)})_{x,y}$] for $(x, y) \in R_h$.

Proposition 1. *We have*

$$(5) \quad P_k(h) = \frac{1}{|X|} \sum_{j=0}^d \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) \right)^k q_j(h),$$

$$(6) \quad P(t, h) = \frac{1}{|X|} \sum_{j=0}^d \exp \left\{ t \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) - 1 \right) \right\} q_j(h).$$

Proof. Let $A \circ B$ denote the Hadamard product (viz. entry-wise product) of matrices A, B and $\tau(A)$ the sum of the entries of A . From (3) and (4),

$$(7) \quad P^k = \sum_{j=0}^d \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) \right)^k E_j,$$

$$P_k(h) = \frac{1}{\kappa_h |X|} \tau(P^k \circ A_h) = \frac{1}{\kappa_h |X|} \sum_{j=0}^d \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) \right)^k \tau(E_j \circ A_h).$$

Since $\tau(E_j \circ A_h) = \kappa_h q_j(h)$ holds by (3), we have (5). Using

$$(8) \quad e^{t(P-I)} = \sum_{j=0}^d \exp t \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) - 1 \right) E_j,$$

we obtain (6) similarly. ■

Remark 1. Proposition 1 and its proof remain valid without any modifications for every commutative association scheme.

Let us turn to Hamming graph $H(d, n)$. In terms of Krawtchouk polynomial

$$K_i(u) = \sum_{l=0}^i (-n)^l (n-1)^{i-l} \binom{d-l}{i-l} \binom{u}{l},$$

we have $p_i(j) = q_i(j) = K_i(j)$. See [1], III.2. In particular,

$$(9) \quad p_1(j) = (n-1)d - nj, \quad \kappa_j = m_j = (n-1)^j \binom{d}{j} \quad (j=0, 1, \dots, d).$$

Now we specialize (5) and (6) in a simple random walk on $H(d, n)$:

$$(10) \quad P_k(h) = \frac{1}{|X|} \sum_{j=0}^d \left(\frac{p_1(j)}{\kappa_1} \right)^k q_j(h) = \frac{1}{n^d} \sum_{j=0}^d \left(1 - \frac{nj}{(n-1)d} \right)^k K_j(h).$$

Since the eigenvalues of $H(d, n)$ form an arithmetic progression and $K_i(u)$ has the following generating function (see [1], III.2):

$$\sum_{i=0}^{\infty} K_i(u)t^i = (1 + (n-1)t)^d - u(1-t)^d,$$

we obtain

$$(11) \quad \begin{aligned} P(t, h) &= \frac{1}{|X|} \sum_{j=0}^d \exp t \left(\frac{p_1(j)}{\kappa_1} - 1 \right) q_j(h) = \frac{1}{n^d} \sum_{j=0}^d \exp \left(-t \frac{nj}{(n-1)d} \right) K_j(h) \\ &= \frac{1}{n^d} \left\{ 1 + (n-1) \exp \left(-\frac{nt}{(n-1)d} \right) \right\}^{d-h} \left\{ 1 - \exp \left(-\frac{nt}{(n-1)d} \right) \right\}^h. \end{aligned}$$

Remark 2. (9) and (11) show that $\kappa_h P(t, h)$ ($h=0, 1, \dots, d$) is a binomial distribution.

The cut-off phenomenon occurs in a simple random walk on $H(d, n)$ if $n \geq 3$ and n/d is bounded above. The critical time is given by $(d/2)(1 - (1/n)) \log(n-1)d$.

§3. Continuous vs Discrete Time

From (7) and (8), we have, for $k \in \mathbb{N}$,

$$(12) \quad \begin{aligned} P^k - e^{k(P-I)} &= \sum_{j=1}^d \left\{ \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) \right)^k - \exp k \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) - 1 \right) \right\} E_j, \\ \|P^k - e^{k(P-I)}\|_{HS}^2 &= \sum_{j=1}^d m_j \left| \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) \right)^k - \exp k \left(\sum_{i=0}^d \frac{w_i}{\kappa_i} p_i(j) - 1 \right) \right|^2 \end{aligned}$$

where $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm of a matrix.

Theorem 3. *Let us consider a simple random walk on $H(d, n)$ ($n \geq 3$) and assume $\log n / \log d$ is bounded. For $\forall \alpha \in (0, 1/2)$ and $\forall \theta_0 > 0$, there exists a positive constant $A = A(\alpha, \theta_0)$ depending only on α and θ_0 , and locally bounded, such that, for*

$$(13) \quad k = \left\lceil \frac{1}{2} \left(1 - \frac{1}{n} \right) d (\log(n-1)d + \theta) \right\rceil \quad \text{where } |\theta| \leq \theta_0,$$

$$\|P^k - e^{k(P-I)}\|_{HS} \leq A \left(\frac{\log d}{d} \right)^{2\alpha}.$$

Here d in (13) is large enough to satisfy $\log(n-1)d - \theta_0 \geq 0$ for given θ_0 .

Proof. In a simple random walk on $H(d, n)$, (12) with (9) yields

$$(14) \quad \|P^k - e^{k(P-I)}\|_{HS}^2 = \sum_{j=1}^d (n-1)^j \binom{d}{j} \left| \left(1 - \frac{nj}{(n-1)d} \right)^k - \exp \left(-k \frac{nj}{(n-1)d} \right) \right|^2.$$

(A reader who skipped §2 will be able to get (14) directly by using diagonalization of symmetric matrices.)

We divide the right-hand side into $\Sigma_{j=1}^J$ and $\Sigma_{j=J+1}^d$ to estimate it. J is specified later in Step 3.

(Step 1) The first sum $\Sigma_{j=1}^J$ in (14) is

$$(15) \quad \sum_{j=1}^J (n-1)^j \binom{d}{j} \exp \left(-2k \frac{nj}{(n-1)d} \right) \left| \exp k \left\{ \log \left(1 - \frac{nj}{(n-1)d} \right) + \frac{nj}{(n-1)d} \right\} - 1 \right|^2.$$

Let $J \leq (d/2)(1 - (1/n))$. Using $-x^2 \leq \log(1-x) + x \leq 0$ for $x \leq 1/2$, we get, for $j=1, \dots, J$ and for k in (13),

$$\begin{aligned} & \left| \exp k \left\{ \log \left(1 - \frac{nj}{(n-1)d} \right) + \frac{nj}{(n-1)d} \right\} - 1 \right|^2 \\ & \leq k^2 \left| \log \left(1 - \frac{nj}{(n-1)d} \right) + \frac{nj}{(n-1)d} \right|^2 \leq k^2 \left(\frac{nj}{(n-1)d} \right)^4 \leq \frac{n^2 J^4}{4(n-1)^2 d^2} (\log(n-1)d + \theta)^2. \end{aligned}$$

Then, since $k > (d/2)(1 - (1/n))(\log(n-1)d + \theta) - 1$,

$$\begin{aligned} (15) & \leq \frac{n^2 J^4}{4(n-1)^2 d^2} (\log(n-1)d + \theta)^2 \sum_{j=1}^J (n-1)^j \binom{d}{j} e^4 \frac{1}{(n-1)^j d^j} e^{-j\theta} \\ & \leq \frac{e^4}{2} \frac{n^2 J^4}{(n-1)^2 d^2} \left\{ (\log(n-1)d)^2 + \theta^2 \right\} \sum_{j=1}^{\infty} \frac{1}{j!} e^{-j\theta} \\ & \leq (\exp e^{-\theta} - 1) \frac{J^4}{d^2} \left\{ (\log(n-1)d)^2 + \theta^2 \right\}. \end{aligned}$$

(Step 2) The second sum $\Sigma_{j=J+1}^d$ in (14) is majorized by

$$2 \sum_{j=J+1}^d (n-1)^j \binom{d}{j} \left(1 - \frac{nj}{(n-1)d}\right)^{2k} + 2 \sum_{j=J+1}^d (n-1)^j \binom{d}{j} \exp\left(-2k \frac{nj}{(n-1)d}\right).$$

Putting k in (13) and using $j! \geq \sqrt{2\pi} j^{j+(1/2)} e^{-j}$, we get

$$\begin{aligned} \sum_{j=J+1}^d (n-1)^j \binom{d}{j} \exp\left(-2k \frac{nj}{(n-1)d}\right) &\leq e^4 \sum_{j=J+1}^d \frac{1}{j!} e^{-j\theta} \\ &\leq \frac{e^4}{\sqrt{2\pi}} \sum_{j=J+1}^d \exp\left\{-j \log j \left(1 + \frac{1}{2j} + \frac{\theta-1}{\log j}\right)\right\} \leq K_1 J^{-\beta J} \end{aligned}$$

where $K_1 > 0, \beta > 0$ are absolute constants, if $\log J \geq 2(|\theta| + 1)$. In fact,

$$1 + \frac{1}{2J} + \frac{\theta-1}{\log J} \geq 1 - \frac{|\theta|+1}{\log J} \geq \frac{1}{2}.$$

Next, using $\log(1-x) \leq -x$ ($x < 1$) and dividing the sum into two, we get

$$\begin{aligned} &\sum_{j=J+1}^d (n-1)^j \binom{d}{j} \left(1 - \frac{nj}{(n-1)d}\right)^{2k} \\ &\leq \sum_{j=J+1}^{\lfloor (1-(1/n)d) \rfloor} (n-1)^j \binom{d}{j} \exp\left(-2k \frac{nj}{(n-1)d}\right) \\ &\quad + \sum_{j=\lfloor (1-(1/n)d) \rfloor + 1}^d (n-1)^j \binom{d}{j} \left(\frac{nj}{(n-1)d} - 1\right)^{2k}. \end{aligned}$$

Estimation of the first sum is already done above.

The second sum = $\sum_{l=0}^{\lfloor d/n \rfloor} (n-1)^{d-l} \binom{d}{l} \left(\frac{n(d-l)}{(n-1)d} - 1\right)^{2k}$

$$\begin{aligned} &\leq \sum_{l=0}^{\lfloor d/n \rfloor} \frac{1}{l!} \exp\left\{(d-l)\log(n-1) + l\log d + 2k\left(\log \frac{1}{n-1} + \log\left(1 - \frac{nl}{d}\right)\right)\right\} \\ &\leq \sum_{l=0}^{\lfloor d/n \rfloor} \frac{1}{l!} \exp\left\{d\log(n-1) + l\log d - \left(1 - \frac{1}{n}\right)d\log(n-1)(\log(n-1)d + \theta) \right. \\ &\quad \left. + 2\log(n-1) + \left(1 - \frac{1}{n}\right)d\log\left(1 - \frac{nl}{d}\right)(\log(n-1)d + \theta) - 2\log\left(1 - \frac{nl}{d}\right)\right\} \\ (16) \quad &= \sum_{l=0}^{\lfloor d/n \rfloor} \frac{1}{l!} \exp\left\{d\log(n-1)\log(n-1)d\left(\frac{1}{\log(n-1)d} - \left(1 - \frac{1}{n}\right)\left(1 + \frac{\theta}{\log(n-1)d}\right)\right) \right. \\ &\quad \left. + \frac{2}{d\log(n-1)d}\right) + l\log d + \left(1 - \frac{1}{n}\right)d\log\left(1 - \frac{nl}{d}\right)(\log(n-1)d + \theta) \end{aligned}$$

$$-2\log\left(1-\frac{nl}{d}\right)\}.$$

Here, for $l=0, 1, \dots, [d/n]$,

$$\begin{aligned} l\log d + \left\{ \left(1-\frac{1}{n}\right)d(\log(n-1)d + \theta) - 2 \right\} \log\left(1-\frac{nl}{d}\right) \\ \leq l\log d - l(n-1)(\log d - |\theta|) + \frac{2nl}{d} \leq l(n-2)\left(2|\theta| - \log d + \frac{2n}{(n-2)d}\right) \leq 0 \end{aligned}$$

if $\log d > 2|\theta| + 3$. (Note $n \geq 3$.) We can then deduce

$$(16) \leq \sum_{l=0}^{[d/n]} \frac{1}{l!} \exp\left\{-\gamma d \log(n-1) \log(n-1)d\right\} \leq K_2 d^{-\gamma d}$$

where $K_2 > 0, \gamma > 0$ are absolute constants.

(Step 3) We have obtained, for $\exp 2(|\theta| + 1) < J \leq (d/2)(1 - (1/n))$ and $d > \exp(2|\theta| + 3)$,

$$(14) \leq (\exp e^{-\theta} - 1) \frac{J^4}{d^2} \{(\log(n-1)d)^2 + \theta^2\} + 4K_1 J^{-\beta J} + 2K_2 d^{-\gamma d}.$$

Now set $J = [(d/\log d)^{(1/2) - \alpha}]$ for given $\alpha \in (0, 1/2)$. Then, if $d/\log d > \exp\{2(|\theta| + 1)/((1/2) - \alpha)\}$,

$$\begin{aligned} (14) \leq \left\{ (\exp e^{-\theta} - 1) \frac{(\log(n-1)d)^2 + \theta^2}{(\log d)^2} + 4K_1 J^{-\beta J} \left(\frac{d}{\log d}\right)^{4\alpha} \right. \\ \left. + 2K_2 d^{-\gamma d} \left(\frac{d}{\log d}\right)^{4\alpha} \right\} \left(\frac{\log d}{d}\right)^{4\alpha}. \end{aligned}$$

Since

$$\begin{aligned} \left(\frac{d}{\log d}\right)^{4\alpha} J^{-\beta J} &\leq \left(\frac{d}{\log d}\right)^{4\alpha} 2^{\beta J} \left(\frac{d}{\log d}\right)^{-\beta J((1/2) - \alpha)} \\ &= \exp\left\{\left(4\alpha - \beta J\left(\frac{1}{2} - \alpha\right) + \beta J \log 2 \frac{1}{\log(d/\log d)}\right) \log\left(\frac{d}{\log d}\right)\right\} \rightarrow 0 \quad \text{as } d \rightarrow \infty, \end{aligned}$$

we can take $K(\alpha) > 0$ depending only on α such that

$$(14) \leq \left\{ (\exp e^{-\theta} - 1) \frac{(\log(n-1)d)^2 + \theta^2}{(\log d)^2} + K(\alpha) \right\} \left(\frac{\log d}{d} \right)^{4\alpha}$$

$$\text{if } \frac{d}{\log d} > \exp \frac{2(|\theta| + 1)}{(1/2) - \alpha}.$$

This yields the desired inequality since $\log n / \log d$ is assumed to be bounded. ■

§4. The Case Where $n/d \rightarrow \tau$ ($0 < \tau < \infty$)

In a continuous time Markov chain (with spatial symmetry), closeness to the equilibrium at time t is measured by

$$(17) \quad C(t) = \frac{1}{2|X|} \sum_{x,y \in X} |(e^{t(P-J)})_{x,y} - (E_0)_{x,y}|$$

similarly to (2). Let $C^{(d,n)}(t)$ denote this quantity in a simple random walk on $H(d,n)$.

Theorem 4. *Let τ be a positive constant. For $\theta \in \mathbf{R}$, as $d \rightarrow \infty$ and $n/d \rightarrow \tau$,*

$$C^{(d,n)} \left(\frac{1}{2} \left(1 - \frac{1}{n} \right) d (\log(n-1)d + \theta) \right) \longrightarrow \| p_{1/\tau} - p_{(1/\tau) + (e^{-\theta/2}/\sqrt{\tau})} \|_{\text{var}}$$

(compact-uniformly in θ), where p_α denotes the Poisson distribution with intensity α .

Proof. (Step 1) Following [3], Proposition 1, we derive an expression of $C^{(d,n)}(t)$. (17) and (11) yield

$$(18) \quad C^{(d,n)}(t) = \frac{1}{2} \sum_{h=0}^d \kappa_h \left| P(t, h) - \frac{1}{|X|} \right|$$

$$= \frac{1}{2n^d} \sum_{h=0}^d (n-1)^h \binom{d}{h}$$

$$\left| \left(1 + (n-1) \exp \left(-\frac{nt}{(n-1)d} \right) \right)^{d-h} \left(1 - \exp \left(-\frac{nt}{(n-1)d} \right) \right)^h - 1 \right|.$$

Note that the inside of $|\cdot|$ in (18) is decreasing in h . Putting $t = (d/2)(1 - (1/n))(\log(n-1)d + \theta)$ into (18) and setting

$$(19) \quad \phi(d, n, \theta) = \frac{d \log(1 + e^{-\theta/2} \sqrt{(n-1)/d})}{\log(1 + e^{-\theta/2} \sqrt{(n-1)/d}) - \log(1 - e^{-\theta/2} / \sqrt{(n-1)d})},$$

we get

$$\begin{aligned}
 C^{(d,n)}(t) &= \sum_{\phi \leq h \leq d} \frac{(n-1)^h}{n^d} \binom{d}{h} \left\{ 1 - \left(1 + e^{-\theta/2} \sqrt{\frac{n-1}{d}} \right)^{d-h} \left(1 - \frac{e^{-\theta/2}}{\sqrt{(n-1)d}} \right)^h \right\} \\
 (20) \quad &= \sum_{0 \leq j \leq d-\phi} \binom{1}{n}^j \left(1 - \frac{1}{n} \right)^{d-j} \binom{d}{j} \\
 &\quad \left\{ 1 - \left(1 + e^{-\theta/2} \sqrt{\frac{n-1}{d}} \right)^j \left(1 - \frac{e^{-\theta/2}}{\sqrt{(n-1)d}} \right)^{d-j} \right\}.
 \end{aligned}$$

(Step 2) From (19), as $d \rightarrow \infty$ and $n/d \rightarrow \tau$,

$$d - \phi(d, n, \theta) \rightarrow \theta^* := \frac{1}{e^{\theta/2} \sqrt{\tau} \log(1 + e^{-\theta/2} \sqrt{\tau})}.$$

θ^* is characterized by

$$j \leq \theta^* \Leftrightarrow e^{-1/\tau} \frac{1}{j!} \left(\frac{1}{\tau} \right)^j \geq e^{-((1/\tau) + (e^{-\theta/2}/\sqrt{\tau}))} \frac{1}{j!} \left(\frac{1}{\tau} + \frac{e^{-\theta/2}}{\sqrt{\tau}} \right)^j.$$

Noting that

$$1 - \left(1 + e^{-\theta/2} \sqrt{\frac{n-1}{d}} \right)^{\theta^*} \left(1 - \frac{e^{-\theta/2}}{\sqrt{(n-1)d}} \right)^{d-\theta^*} \rightarrow 0 \quad \text{as } \begin{cases} d \rightarrow \infty \\ n/d \rightarrow \tau \end{cases}$$

holds even if $\theta^* \in \mathbb{N}$, we obtain

$$\begin{aligned}
 C^{(d,n)}(t) &\rightarrow \sum_{0 \leq j \leq \theta^*} \lim \binom{1}{n}^j \left(1 - \frac{1}{n} \right)^{d-j} \binom{d}{j} \left\{ 1 - \left(1 + e^{-\theta/2} \sqrt{\frac{n-1}{d}} \right)^j \left(1 - \frac{e^{-\theta/2}}{\sqrt{(n-1)d}} \right)^{d-j} \right\} \\
 &= \sum_{0 \leq j \leq \theta^*} e^{-1/\tau} \frac{1}{j!} \left(\frac{1}{\tau} \right)^j - e^{-((1/\tau) + (e^{-\theta/2}/\sqrt{\tau}))} \frac{1}{j!} \left(\frac{1}{\tau} + \frac{e^{-\theta/2}}{\sqrt{\tau}} \right)^j \\
 &= \|P_{1/\tau} - P_{(1/\tau) + (e^{-\theta/2}/\sqrt{\tau})}\|_{\text{var}}
 \end{aligned}$$

as $d \rightarrow \infty$ and $n/d \rightarrow \tau$. ■

Proof of Theorem 1. From (2) and (17),

$$(21) \quad |D(k) - C(k)| \leq \frac{1}{2|X|} \sum_{x,y \in X} |(P^k)_{x,y} - (e^{k(P-I)})_{x,y}|$$

$$\leq \frac{1}{2} \|P^k - e^{k(P-I)}\|_{HS}.$$

Hence Theorem 1 immediately follows from Theorem 3 and Theorem 4. ■

§5. The Case Where $n/d \rightarrow 0$

In this section, let $F_{d,p}(x)$ denote the distribution function of the normalization of binomial distribution $B(d,p)$ (so as to enjoy mean 0 and variance 1) and $F(x)$ that of the normal distribution $N(0, 1)$.

Berry-Esseen Theorem. *Let d, p be unfixed in $B(d,p)$ and satisfy $0 < p \leq 1 - \delta$ for some $\delta \in (0, 1)$. Then, there exists a positive constant $C = C(\delta)$ depending only on δ such that*

$$\sup_{x \in \mathbf{R}} |F_{d,p}(x) - F(x)| \leq \frac{C}{\sqrt{pd}}.$$

Though d and p are unfixed, it can be proved similarly to the usual Berry-Esseen theorem. See e.g. [4], Chap.2.

Theorem 5. *For $\theta \in \mathbf{R}$, as $d \rightarrow \infty$ and $n/d \rightarrow 0$,*

$$C^{(d,n)} \left(\frac{1}{2} \left(1 - \frac{1}{n} \right) d (\log(n-1)d + \theta) \right) \rightarrow \text{Erf} \left(\frac{e^{-\theta/2}}{2\sqrt{2}} \right)$$

(compact-uniformly in θ).

Proof. We have

$$\begin{aligned} C^{(d,n)}(t) &= \sum_{0 \leq j \leq d-\phi} \left(\frac{1}{n}\right)^j \left(1 - \frac{1}{n}\right)^{d-j} \binom{d}{j} \\ &\quad - \sum_{0 \leq j \leq d-\phi} \left(\frac{1 + e^{-\theta/2} \sqrt{(n-1)/d}}{n}\right)^j \left(1 - \frac{1 + e^{-\theta/2} \sqrt{(n-1)/d}}{n}\right)^{d-j} \binom{d}{j} \end{aligned}$$

in (20). Let (I) and (II) denote the first and the second sum above respectively.

(Step 1) Set $p = (1 + e^{-\theta/2} \sqrt{(n-1)/d})/n$. Then,

$$(22) \quad \text{(II)} = F_{d,p} \left(\frac{d - \phi(d, n, \theta) - pd}{\sqrt{p(1-p)d}} \right).$$

As is shown in Step 2,

$$(23) \quad \frac{d - \phi(d, n, \theta) - pd}{\sqrt{p(1-p)d}} \longrightarrow -\frac{1}{2}e^{-\theta/2} \quad \text{as} \quad \begin{cases} d \rightarrow \infty \\ n/d \rightarrow 0 \end{cases}$$

holds. Hence

$$(24) \quad F\left(\frac{d - \phi(d, n, \theta) - pd}{\sqrt{p(1-p)d}}\right) \longrightarrow F\left(-\frac{1}{2}e^{-\theta/2}\right) \quad \text{as} \quad \begin{cases} d \rightarrow \infty \\ n/d \rightarrow 0. \end{cases}$$

For given $\theta \in \mathbf{R}$, if d is large enough,

$$1 - p = 1 - \frac{1}{n} - e^{-\theta/2} \sqrt{1 - \frac{1}{n} \frac{1}{\sqrt{nd}}} \geq \frac{1}{2}.$$

Then, Berry-Esseen theorem ensures

$$(25) \quad \left| F_{a,p}\left(\frac{d - \phi(d, n, \theta) - pd}{\sqrt{p(1-p)d}}\right) - F\left(\frac{d - \phi(d, n, \theta) - pd}{\sqrt{p(1-p)d}}\right) \right| \\ \leq \frac{C}{\sqrt{pd}} = \frac{C}{\sqrt{(d/n)(1 + e^{-\theta/2})\sqrt{(n-1)/d}}} \longrightarrow 0.$$

From (22), (24) and (25), we get

$$(II) \longrightarrow F\left(-\frac{1}{2}e^{-\theta/2}\right).$$

(Step 2) We show (23). Set

$$d - \phi(d, n, \theta) = d \frac{-\log(1 - (e^{-\theta/2}/\sqrt{(n-1)d}))}{\log(1 + e^{-\theta/2}\sqrt{(n-1)/d}) - \log(1 - (e^{-\theta/2}/\sqrt{(n-1)d}))} \\ = d(III)/(IV).$$

Then,

$$(III) = \frac{e^{-\theta/2}}{\sqrt{(n-1)d}} + \frac{e^{-\theta}}{2(n-1)d} + O\left(\frac{1}{(nd)^{3/2}}\right), \\ (IV) = \left(\sqrt{n-1} + \frac{1}{\sqrt{n-1}}\right) \frac{e^{-\theta/2}}{\sqrt{d}} + \frac{n(2-n)e^{-\theta}}{2(n-1)d} + O((n/d)^{3/2}),$$

$$\begin{aligned} \frac{\text{(III)}}{\text{(IV)}} &= \frac{1}{n} \frac{1 + (e^{-\theta/2}/2)\sqrt{(n-1)d} + O(1/(nd))}{1 + (e^{-\theta/2}(2-n)/2)\sqrt{(n-1)d} + O(n/d)} \\ &= \frac{1}{n} \left\{ 1 + \frac{(e^{-\theta/2}\sqrt{n-1}/2\sqrt{d}) + O(n/d)}{1 + (e^{-\theta/2}(2-n)/2)\sqrt{(n-1)d} + O(n/d)} \right\} \\ &= \frac{1}{n} + \frac{1}{\sqrt{nd}} \frac{(e^{-\theta/2}\sqrt{1-(1/n)/2}) + O(\sqrt{n/d})}{1 + O(\sqrt{n/d})}. \end{aligned}$$

Hence we get

$$(26) \quad d - \phi(d, n, \theta) = \frac{d}{n} + \sqrt{\frac{d}{n}} \frac{(e^{-\theta/2}\sqrt{1-(1/n)/2}) + O(\sqrt{n/d})}{1 + O(\sqrt{n/d})}.$$

Combining (26) with

$$pd = \frac{d}{n} + \sqrt{\frac{d}{n}} \sqrt{1 - \frac{1}{n}} e^{-\theta/2},$$

we obtain (23).

(Step 3) From (26),

$$\frac{d - \phi(d, n, \theta) - (d/n)}{\sqrt{(1 - (1/n))d/n}} = \frac{(e^{-\theta/2}/2) + O(\sqrt{n/d})}{1 + O(\sqrt{n/d})} \rightarrow \frac{1}{2} e^{-\theta/2}.$$

Hence, replacing p by $1/n$ in Step 1, we get

$$(I) \rightarrow F\left(\frac{1}{2} e^{-\theta/2}\right).$$

Consequently, we obtain

$$\begin{aligned} C^{(d,n)}(t) &= (I) - (II) \\ &\rightarrow F\left(\frac{1}{2} e^{-\theta/2}\right) - F\left(-\frac{1}{2} e^{-\theta/2}\right) = \text{Erf}\left(\frac{e^{-\theta/2}}{2\sqrt{2}}\right) \end{aligned}$$

as $d \rightarrow \infty$ and $n/d \rightarrow 0$. ■

Proof of Theorem 2. Combine Theorem 3 and (21) with Theorem 5. ■

References

- [1] Bannai, E. and Ito, T., *Algebraic combinatorics I, Association schemes*, Benjamin/Cummings, Menlo Park, California, 1984.
- [2] Diaconis, P., The cutoff phenomenon in finite Markov chains, *Proc. Nat. Acad. Sci. USA*, **93** no. 4 (1996), 1659–1664.
- [3] Diaconis, P., Graham, R. L. and Morrison, J. A., Asymptotic analysis of a random walk on a hypercube with many dimensions, *Random Struct. Algor.*, **1** no. 1 (1990), 51–72.
- [4] Durrett, R., *Probability: theory and examples*, Duxbury Press, Belmont, California, 1991.
- [5] Hora, A., Critical phenomena for random walks on P - and Q -polynomial association schemes, *Preprint*, 1995.
- [6] ———, Towards critical phenomena for random walks on various algebraic structures, In H. Heyer and T. Hirai (eds.), *Trans. German-Japanese Symposium 1995 in Tübingen*, 113–127, D. + M. Gräbner, Bamberg, 1996.
- [7] Voit, M., Asymptotic distributions for the Ehrenfest urn and related random walks, *J. Appl. Probab.*, **33** no. 3 (1996), 340–356.
- [8] ———, Asymptotic behaviour of heat kernels on spheres of large dimensions, *J. Multivariate Anal.*, **59** (1996), 230–248.