HIROSHIMA MATH. J.
30 (2000), 271–299

An axiomatic approach to the cut-off phenomenon for random walks on large distance-regular graphs

To the Memory of Dr. Hitoshi Mizumachi

Akihito HORA

(Received June 17, 1999)

ABSTRACT. The cut-off phenomenon is a sort of critical phenomenon which one often observes in the process of convergence to equilibrium for various Markov chains including card shuffling and diffusion of sparse gases. This article aims at developing an axiomatic approach to this phenomenon on a nice class of distance-regular graphs. Following the formulation through large volume limits, we present a rigorous criterion for the cut-off phenomenon in terms of spectral data of the adjacency matrix of the graph.

1. Introduction

The *cut-off phenomenon* (abbreviated to COP) is widely observed in the process of convergence to equilibrium for Markov chains. It is a critical phenomenon owing to the huge cardinality of the state space of the chain and is well understood through a large volume limit of the system. Initiated by P. Diaconis, the study of this phenomenon has now grown to enjoy considerable literature. Let us consider a Markov chain on finite state space X with transition probability matrix P and invariant probability π. We assume the convergence to equilibrium of the chain:

$$(P^k)_{x,y} \to \pi(y) \qquad \text{as } k \to \infty \qquad \text{for } \forall x, y \in X,$$

which is in fact assured under mild conditions. The total variation distance

$$\|(P^k)_{x,\cdot} - \pi\| = \frac{1}{2}\sum_{y \in X} |(P^k)_{x,y} - \pi(y)| \tag{1}$$

will describe the convergence more quantitatively. In this article, we treat Markov chains enjoying some spatial symmetry, which then implies that the

2000 *Mathematics Subject Classification*: 60J10, 60C05, 82C27, 05E30.
Key words and phrases: random walk, cut-off phenomenon, distance-regular graph, critical time to reach equilibrium

invariant probability is uniform and (1) is independent of the choice of x. Hence we set

$$D(k) = \frac{1}{2|X|}\sum_{x,y\in X}\left|(P^k)_{x,y} - \frac{1}{|X|}\right| \qquad (= (1)). \tag{2}$$

Roughly speaking, $D(k)$ often takes a sudden transition from 1 to 0 at specific critical time k_c. A naive description of the COP is as follows.

DEFINITION 1. Consider a directed family of Markov chains parametrized by $\lambda \in \Lambda$ (directed set) and set $D^{(\lambda)}(k)$ as (2) for each chain. If one can take $0 < k_c^{(\lambda)} \to \infty$ (as $\lambda \to \infty$) such that the scaled graph of function $D^{(\lambda)}(k)$, $\{(k/k_c^{(\lambda)}, D^{(\lambda)}(k)) \mid k \in \mathbf{N}\}$, converges to the graph of step function

$$S(x) = \begin{cases} 1 & \text{for } 0 < x \le 1 \\ 0 & \text{for } 1 \le x \end{cases},$$

namely to $(0,1] \times \{1\} \cup \{1\} \times [0,1] \cup [1,\infty) \times \{0\}$ as subsets of $\mathbf{R}^2$ as $\lambda \to \infty$, one says that the COP occurs for this family of chains and call $k_c^{(\lambda)}$ (determined up to $(1+o(1))$ multiple) the critical time to reach equilibrium.

A more detailed formulation of the COP is presented in §3.

Comprehensive expositions of the COP are due to Diaconis ([7], [8]). The best studied models in which the COP occurs are random walks on some finite groups and their homogeneous spaces. They include, for example, shuffling cards ([13], [1]), the Ehrenfests urn model and related ones ([9], [27], [20]), the Bernoulli-Laplace diffusion model ([14], [15]) and its q-analogue ([6]), and some matrix groups over a finite field ([16]). Furthermore, several models on compact groups (like classical groups) and compact homogeneous spaces were treated in [23], [24], [21], [22], [26] and [28]. See the bibliographies in [7], [8] and [25] for other works not cited here. In [8] Diaconis proposed an essential understanding of what causes the COP. He pointed out the decisive role of high multiplicity of the second largest eigenvalue of a transition matrix, which results from high symmetry of the system, and presented an intuitive explanation based on the upper bound lemma. On the other hand, in order to prove the COP precisely in an actual model, however, one needs still additional information on the model discussed.

The results for random walks with high symmetry on finite homogeneous spaces would have extension in several directions. One is to weaken (or to break) the symmetry of a transition matrix. In particular, the invariant probability may be no more uniform. This extension is quite important from the viewpoint of applications, especially to statistical mechanics. Interesting results in such a direction are presented in [10], [11], [12], [8] and [25]. Another is a generalization of the algebraic structure which describes symmetry of the system. In [7] Diaconis already mentioned two candidates beyond groups: hypergroups and association schemes. For example, if $X = G/K$ is a homogeneous space of finite group G, $G\backslash X \times X$ and $K\backslash G/K$ have canonically the structure of an association scheme and a hypergroup respectively. These structures naturally come into the present context if one recalls that, in the case of a homogeneous space, an essential role is played not by the group itself but by the Hecke algebra associated with it at least from a methodological viewpoint. In [26], [27] and [28], Voit dealt with the COP for some models related to hypergroups (in particular, polynomial hypergroups). As for the COP on association schemes, we mention [4] discussed below (and the present article also).

The purpose of this article is to propose rigorous and practical criteria for the COP which say more beyond verification in individual models or intuitive understanding based on degeneration of the second eigenvalue and at the same time are applicable to the models associated with a nice class of distance-regular graphs. The *distance-regular graphs* (abbreviated to DRG) are an important subclass of the association schemes. We include a concise review on DRGs in §2. This article contains the full proofs of our results announced in [18] and [19] as well as several refinements. A significant change of the situation during preparation of this article was appearance of [4]. Belsley investigated there the COP for all the known families of q-DRGs and showed that the critical time coincides with the diameter of the graph. (Here "q-" suggests symbolically the objects concerning e.g. matrices or vector spaces over a finite field.) Although our new concrete examples in [18] and [19] are now included in [4], we think that developing an axiomatic approach to the COP enjoys meanings because the classification of the DRGs is still far beyond the scope. We will present the criteria for the COP which are expressed fully in terms of the *spectral data* of the DRGs. Our method is a prolongation of the harmonic-analytic one developed first by Diaconis and Shahshahani in [13] and [14]. We also note that the method works on general commutative association schemes under minor trivial modifications. (See also [20].)

We organize the subsequent sections as follows.

§2 begins with a brief review on DRGs. Then we introduce random walks on them and state some basic properties such as the upper bound lemma. §3 is the core of the article. We present a precise formulation of the COP, making much of its statistical-mechanical aspect. Comparison of the scales between the critical time and small deviation around it is important. Estimating the distance $D(k)$ in (2) through harmonic-analytic tools, we reach the main results on criteria for the COP. Until §3, we keep the article self-contained. §4 is devoted to illustration in concrete models. We refer to [2] for explicit spectral data of concrete DRGs.

I would like to dedicate this article to Dr. Hitoshi Mizumachi who shared a great interest in algebraic probability with us.

2. Random walk on distance-regular graph

2.1 Distance-regular graph (DRG)

Let $\Gamma = (X, E)$ be a finite connected graph where X is a vertex set and E is an edge set. Two vertices x, y are said to be adjacent (denoted by $x \sim y$) if they are joined by an edge in E. The canonical distance between $x, y \in X$ is denoted by $\partial(x, y)$. $d = \max_{x,y \in X} \partial(x, y)$ is called the diameter of Γ. Then $\partial(x, y) \in \{0, 1, \dots, d\}$. Γ is called a DRG if for $\forall h, i, j \in \{0, 1, \dots, d\}$ the following quantity does not depend on the choice of x, y whenever $\partial(x, y) = h$:

$$|\{z \in X \mid \partial(x, z) = i, \partial(z, y) = j\}| = p_{ij}^h. \tag{3}$$

In particular, the degree of Γ is $\kappa = p_{11}^0$. The ith adjacency matrix A_i of Γ is the $|X| \times |X|$ matrix whose (x, y)-entry is

$$(A_i)_{x,y} = \begin{cases} 1 & \text{if } \partial(x, y) = i \\ 0 & \text{if } \partial(x, y) \neq i \end{cases} \qquad (x, y \in X; i = 0, 1, \dots, d).$$

In particular, $A_0 = I$ (identity). A_1 is simply denoted by A and called the adjacency matrix of Γ. Condition (3) of distance-regularity is translated to the linearizing formula for adjacency matrices:

$$A_i A_j = \sum_{h=0}^{d} p_{ij}^h A_h, \tag{4}$$

which one checks immediately by comparing the entries of both sides. The following two are the best known DRGs.

Hamming graph For $d, n \in \mathbf{N}$, let F be an n-set (i.e. $|F| = n$) and $X = F^d$ (d direct product). Joining two vertices $x = (x_j)_{j=1}^d$, $y = (y_j)_{j=1}^d \in X$ by an edge if $|\{j | x_j \neq y_j\}| = 1$, one gets Hamming graph $H(d,n)$. The distance is given by $\partial(x,y) = |\{j | x_j \neq y_j\}|$. The diameter is d.

Johnson graph For $v, d \in \mathbf{N}$, let S be a v-set and $X = \{x \subset S \,|\, |x| = d\}$. One can assume $2d \leq v$ without loss of generality. Joining $x, y \in X$ by an edge if $|x \cap y| = d - 1$, one gets Johnson graph $J(v,d)$. The distance is given by $\partial(x,y) = d - |x \cap y|$. The diameter is d.

It is easy to check that $H(d,n)$ and $J(v,d)$ are DRGs from the definitions.

In the rest of this subsection, let $\Gamma = (X, E)$ be a DRG with diameter d.

LEMMA 1. *$B = [p_{1j}^h]_{j,h}$ is a $(d+1) \times (d+1)$ tridiagonal matrix,*

which immediately follows from (3).

LEMMA 2. *For $\forall i = 0, 1, \ldots, d$, there exists polynomial $v_i(x)$ of degree i such that $A_i = v_i(A)$. Here $v_0(x) = 1$ and $v_0(A) = I$.*

PROOF. It suffices to show that, for $i = 0, 1, \ldots, d$, A^i is expressed as a linear combination of $A_0, A_1, \ldots, A_i$ with a positive coefficient of A_i. Assume that $A^i = \sum_{h=0}^{i} c_h A_h$ holds with $c_i > 0$ ($i = 0, 1, \ldots, d-1$). Using (4) and Lemma 1, we have

$$\begin{aligned} A^{i+1} &= \sum_{h=0}^{i} c_h A A_h = \sum_{h=0}^{i} \sum_{j=0}^{d} c_h p_{1h}^j A_j \\ &= c_0 A_1 + \sum_{h=1}^{i} c_h (p_{1h}^{h-1} A_{h-1} + p_{1h}^h A_h + p_{1h}^{h+1} A_{h+1}), \end{aligned}$$

where the coefficient of A_{i+1} is $c_i p_{1i}^{i+1} > 0$. The proof is completed by induction on i. $\square$

The $\mathbf{C}$-algebra generated by A and I is denoted by $\mathscr{A}(\Gamma)$ and called the adjacency algebra of Γ. Since $A_0, A_1, \ldots, A_d$ are obviously linearly independent, (4) and Lemma 2 imply that $\{A_0, A_1, \ldots, A_d\}$ is a basis of $\mathscr{A}(\Gamma)$.

$\dim \mathscr{A}(\Gamma)$ $(= d+1)$ is equal to the degree of the minimal polynomial of A. Hence symmetric matrix A has $d+1$ distinct eigenvalues and enjoys spectral decomposition

$$A = \sum_{j=0}^{d} \theta_j E_j, \qquad \theta_0 > \theta_1 > \cdots > \theta_d. \tag{5}$$

Here E_j's satisfy $E_j^2 = E_j = E_j^*$, $E_i E_j = \delta_{ij} E_i$ and $E_0 + \cdots + E_d = I$. In view of Lemma 2 and (5), $\{E_0, E_1, \dots, E_d\}$ is also a basis of $\mathscr{A}(\Gamma)$.

LEMMA 3. *In* (5), $\theta_0 = \kappa$ *and* $|\theta_j| \le \kappa$. θ_0 *is a simple eigenvalue.*

PROOF. Let $Au = \alpha u$, $\alpha \in \mathbf{R}$, $u \in \mathbf{R}^{|X|}$, $u \neq 0$. Take an entry of u, say u_a $(a \in X)$, which has maximal absolute value. Since there exist κ vertices b adjacent to a in $\sum_{b \sim a} u_b = \alpha u_a$, we have $|\alpha| \le \kappa$. Applying the argument again to $\alpha = \kappa$, we have $u_b = u_a$ for $b \sim a$. Since Γ is connected, the eigenvector belonging to κ must have identical entries. □

Throughout this article, J denotes the $|X| \times |X|$ matrix having identical entries 1. Lemma 3 shows $E_0 = |X|^{-1} J$ in (5). Each A_i has spectral decomposition

$$A_i = \sum_{j=0}^{d} p_i(j) E_j \qquad (i = 0, 1, \dots, d) \tag{6}$$

where $p_1(j) = \theta_j$ and $p_i(j) = v_i(\theta_j)$. It is obvious from the definition that A_i's are closed under the Hadamard product $\circ$ (i.e. entry-wise product) of matrices and hence so is $\mathscr{A}(\Gamma)$. Then projectors E_j's are linearized with respect to the Hadamard product as

$$(|X|E_i) \circ (|X|E_j) = \sum_{h=0}^{d} q_{ij}^h (|X|E_h). \tag{7}$$

q_{ij}^h is called a Krein parameter. A dual property of Lemma 1 is that $B_1^* = [q_{1j}^h]_{j,h}$ is a tridiagonal matrix. A DRG satisfying this property is said to be Q-polynomial. We summarize a few remarks which are important in general but will not be used later. See [2].

REMARK 1. All the Krein parameters are nonnegative.

REMARK 2. The Q-polynomial property of a DRG holds if and only if, for $\forall i = 0, 1, \dots, d$, there exists polynomial $v_i^*(x)$ of degree i such that $E_i = v_i^*(E_1)$. Here the multiplication of matrices in the expression $v_i^*(E_1)$ is taken under the Hadamard product. (Compare this with Lemma 2.)

REMARK 3. The notions and properties above of DRGs are extended to commutative association schemes with minor modifications, including in particular the linearizing formula and the spectral decompositions for adjacency matrices. The DRGs are characterized by the property of Lemma 1, said to be P-polynomial, in the commutative association schemes.

2.2 Random walk

Random walks are distinguished from other Markov chains by spatial homogeneity. Let P be the transition matrix of a Markov chain on homogeneous space X of finite group G. The action of G induces the orbital decomposition $X \times X = \Lambda_0 \cup \Lambda_1 \cup \cdots \cup \Lambda_d$. One sees

$$\begin{aligned} \text{spatial homogeneity} \quad &\Leftrightarrow \quad P_{gx,gy} = P_{x,y} \qquad \text{for } \forall x, y \in X, \forall g \in G \\ &\Leftrightarrow \quad P_{x,y} \text{ is constant on each } \Lambda_i \\ &\qquad\ \ \text{as a fucntion on } X \times X. \end{aligned} \tag{8}$$

Let us now consider DRG Γ with vertex set X and a Markov chain on X. Setting $R_i = \{(x, y) \in X \times X \mid \partial(x, y) = i\}$ $(i = 0, 1, \ldots, d)$ where ∂ is the distance and d is the diameter, one has the decomposition $X \times X = R_0 \cup R_1 \cup \cdots \cup R_d$. Transition matrix P of this chain satisfies (8) if and only if P is a linear combination of A_i's, namely P belongs to the adjacency algebra $\mathscr{A}(\Gamma)$. Keeping this discussion in mind, we introduce a random walk on a DRG as follows.

DEFINITION 2. A Markov chain on DRG Γ with transition matrix P [resp. transition semigroup $e^{t(P-I)}$] is called a discrete [resp. continuous] time random walk if $P \in \mathscr{A}(\Gamma)$.

Since $\|H_{x,\cdot}\| = \frac{1}{2}\sum_{y \in X} |H_{x,y}|$ does not depend on $x \in X$ for $H \in \mathscr{A}(\Gamma)$, (1) with uniform π is independent of the choice of x if $P \in \mathscr{A}(\Gamma)$. Hence, as in (2), we consider

$$D(k) = \frac{1}{2|X|} \sum_{x,y \in X} |(P^k - E_0)_{x,y}| \qquad (k \in \mathbf{N}) \text{ in discrete time} \tag{9}$$

$$C(t) = \frac{1}{2|X|} \sum_{x,y \in X} |(e^{t(P-I)} - E_0)_{x,y}| \qquad (t \geq 0) \text{ in continuous time} \tag{10}$$

for a random walk on Γ (where $E_0 = |X|^{-1}J$). Set $\kappa_i = p_{ii}^0 =$ the number of i-neighbors of each vertex. In particular, $\kappa_0 = 1$ and $\kappa_1 = \kappa$ (degree). We note

$$\kappa_i = p_i(0) \qquad (i = 0, 1, \ldots, d) \tag{11}$$

in (6). In fact, multiplying E_0 in (6), we have

$$\kappa_i E_0 = A_i E_0 = \sum_{j=0}^{d} p_i(j) E_j E_0 = p_i(0) E_0$$

since any entry of $A_i J$ is κ_i. Transition matrix $P \in \mathscr{A}(\Gamma)$ is expressed as a convex combination of stochastic matrices A_i/κ_i's:

$$P = \sum_{i=0}^{d} \frac{w_i}{\kappa_i} A_i \qquad \left(w_i \geq 0, \quad \sum_{i=0}^{d} w_i = 1 \right). \tag{12}$$

What was called the upper bound lemma due to Diaconis and Shahshahani [13] now takes the following form. Set $m_j = \operatorname{rank} E_j$ $(j = 0, 1, \ldots, d)$.

Proposition 1. *A random walk having transition matrix P as in* (12) *yields*

$$D(k)^2 \leq \frac{1}{4} \sum_{j=1}^{d} m_j \left| \sum_{i=0}^{d} w_i \frac{p_i(j)}{\kappa_i} \right|^{2k} \tag{13}$$

$$C(t)^2 \leq \frac{1}{4} \sum_{j=1}^{d} m_j \exp\left\{ -2t \sum_{i=0}^{d} w_i \left(1 - \frac{p_i(j)}{\kappa_i} \right) \right\}. \tag{14}$$

Proof. Using the spectral decomposition (6) of A_i's, we have

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

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

and by noting (11)

$$P^k - E_0 = \sum_{j=1}^{d} \left(\sum_{i=0}^{d} w_i \frac{p_i(j)}{\kappa_i} \right)^k E_j.$$

Since $\{E_j\}_{j=0}^{d}$ is a complete orthogonal system of projectors, we get

$$\|P^k - E_0\|_{HS}^2 = \sum_{j=1}^{d} m_j \left| \sum_{i=0}^{d} w_i \frac{p_i(j)}{\kappa_i} \right|^{2k}$$

where $\| \cdot \|_{HS}$ denotes the Hilbert-Schmidt norm. Now that the Schwarz inequality yields

$$D(k) \leq \frac{1}{2|X|}\|P^k - E_0\|_{HS}\sqrt{|X|^2} = \frac{1}{2}\|P^k - E_0\|_{HS},$$

(13) holds. Similarly, using

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

$$\|e^{t(P-I)} - E_0\|_{HS}^2 = \sum_{j=1}^{d} m_j \exp\left\{-2t\sum_{i=0}^{d} w_i\left(1 - \frac{p_i(j)}{\kappa_i}\right)\right\},$$

we get (14). □

If $D(k)$ or $C(t) \to 0$ as k or $t \to \infty$, the random walk is often said to be asymptotically equidistributed. In [17] we characterized asymptotic equidistribution of a random walk on a finite group in terms of the support of the one-step transition of the walk and 1-dimensional characters of the group. It is possible to write a similar characterization for a random walk on a DRG by taking up the relation between the support of w_i's and $p_i(j)$'s.

(14) immediately gives the following rough estimate in terms of the "spectral gap". Setting

$$\gamma = \min_{j=1,2,\ldots,d} 1 - \sum_{i=0}^{d} w_i \frac{p_i(j)}{\kappa_i},$$

we have

$$C(t)^2 \leq \frac{1}{4}\sum_{j=1}^{d} m_j e^{-2\gamma t} = \frac{|X|-1}{4}e^{-2\gamma t} = \frac{1}{4}\exp\{\log(|X|-1) - 2\gamma t\}. \tag{15}$$

Assume $\gamma > 0$ (which corresponds to irreducibility of the random walk). One should not deduce from (15) that the time to reach equilibrium is $t = (2\gamma)^{-1}$ or $t = (2\gamma)^{-1}\log(|X|-1)$. We will illustrate that the correct critical time is actually of an intermediate order of the two (as the size of the system grows).

A random walk on a DRG is said to be simple if $P = A/\kappa$. For a simple random walk, the upper bound lemma (Proposition 1) takes the form of

$$D(k)^2 \leq \frac{1}{4}\sum_{j=1}^{d} m_j \left(\frac{\theta_j}{\kappa}\right)^{2k} \tag{16}$$

$$C(t)^2 \leq \frac{1}{4}\sum_{j=1}^{d} m_j \exp\left\{-2t\left(1 - \frac{\theta_j}{\kappa}\right)\right\}. \tag{17}$$

PROPOSITION 2. *A simple random walk on a DRG satisfies that*
(i) *if it is of continuous time, the distribution tends to the uniform probability as* $t \to \infty$,
(ii) *if it is of discrete time and the DRG is not bipartite, so does the distribution as* $k \to \infty$.

PROOF is immediate from (16) and (17) if one notes Lemma 3 and Lemma 4 below.

LEMMA 4. *A DRG is bipartite if* $\theta_d = -\kappa$.

PROOF. Let $u \in \mathbf{R}^{|X|}$ be an eigenvector of the adjacency matrix belonging to $-\kappa$ and u_a $(a \in X)$ an entry of u having maximal absolute value. We can assume $u_a = 1$. Since $\sum_{b \sim a} u_b = -\kappa u_a$ holds, we see $u_b = -1$ for $b \sim a$. Thus the vertices x are divided into two classes according as $u_x = \pm 1$ so that adjacent vertices are in different classes. □

3. The cut-off phenomenon

3.1 Formulation of the cut-off phenomenon (COP)

In this subsection, we present a more quantitaitve description of the COP than Definition 1 in Introduction.

Let us consider a family of random walks on DRGs parametrized by $\lambda \in \Lambda$, Λ being a directed set. λ is usually a (multi-)parameter concerning the size of the system. Let each walk on DRG $\Gamma^{(\lambda)} = (X^{(\lambda)}, E^{(\lambda)})$ have transition matrix $P^{(\lambda)}$ and start at an initial vertex. $D^{(\lambda)}(k)$ [resp. $C^{(\lambda)}(t)$] denotes the total variation distance between the distribution at time $k \in \mathbf{N}$ [resp. $t \geq 0$] of the discrete [resp. continuous] time random walk and the uniform probability on $X^{(\lambda)}$ which is defined as in (9) [resp. (10)].

DEFINITION 3. Assume that one can take $k_c^{(\lambda)} > 0$ for each $\lambda \in \Lambda$ satisfying the following conditions:
(i) $k_c^{(\lambda)} \to \infty$ and $k_c^{(\lambda)}/|X^{(\lambda)}| \to 0$ as $\lambda \to \infty$
(ii) $\forall \varepsilon > 0$, $\exists \lambda_\varepsilon \in \Lambda$ and $\exists h_\varepsilon^{(\lambda)} > 0$ such that $h_\varepsilon^{(\lambda)}/k_c^{(\lambda)} \to 0$ holds as $\lambda \to \infty$ and, if $\lambda > \lambda_\varepsilon$,

$$0 \leq k \leq k_c^{(\lambda)} - h_\varepsilon^{(\lambda)} \quad \Rightarrow \quad D^{(\lambda)}(k) \geq 1 - \varepsilon$$

$$k \geq k_c^{(\lambda)} + h_\varepsilon^{(\lambda)} \quad \Rightarrow \quad D^{(\lambda)}(k) \leq \varepsilon.$$

Then we say that the COP occurs for this family of discrete time random walks and call $k_c^{(\lambda)}$ the critical time to reach equilibrium. We give the same definition to continuous time random walks, replacing $D^{(\lambda)}(k)$ and $k_c^{(\lambda)}$ by $C^{(\lambda)}(t)$ and $t_c^{(\lambda)}$ respectively.

Remark 1. Clearly Definition 3 is a refinement of Definition 1. Since we treat an aperiodic Markov chain with the uniform invariant probability, the mean recurrence time of the chain is $|X^{(\lambda)}|$. We thus see three different scales of time

$$h_\varepsilon^{(\lambda)} \ll k_c^{(\lambda)} \ll |X^{(\lambda)}| \qquad \text{as } \lambda \to \infty.$$

It is appropriate that $k_c^{(\lambda)}$ is regarded as a macroscopic time. Then the system discussed is far before it returns to the initial situation.

Remark 2. Saloff-Coste proposed in [25] a weaker version of the COP ("weak ℓ^p-cutoff").

Remark 3. Diaconis raised in [8] the definition of the COP as follows. Here we use some different notations from the original ones. Under condition (i) of Definition 3, let us assume that one finds $h^{(\lambda)} > 0$ satisfying $h^{(\lambda)}/k_c^{(\lambda)} \to 0$ and

$$D^{(\lambda)}(k_c^{(\lambda)} + \theta h^{(\lambda)}) \to c(\theta) \qquad \text{compact-uniformly in } \theta \in \mathbf{R} \tag{18}$$

as $\lambda \to \infty$ where $c(\theta) : \mathbf{R} \to [0,1]$ is a function such that $c(-\infty) = 1$ and $c(\infty) = 0$. While Definition 3 provides a macrosopic understanding of the COP, (18) describes in more detail the deviation from the equilibrium around critical time $k_c^{(\lambda)}$ with respect to a smaller scale (may be microscopic) of time ($\asymp h^{(\lambda)}$). Concrete models in which the COP is verified to the extent of (18) are given in [9], [3], [27], [28] and [20].

3.2 Upper estimate

In the present and the next sections, we discuss the upper and the lower estimates respectively of total variation distance (9) and (10) for simple random walks on DRGs. Our aim is to establish those estimates done by the spectral datum of the adjacency matrix of a DRG. We consider a directed family $\{\Gamma^{(\lambda)}\}_{\lambda \in \Lambda}$ of DRGs (each DRG having diameter $d^{(\lambda)}$) and a *simple random walk* on $\Gamma^{(\lambda)}$ starting at some vertex. As a notational remark, let superscript $^{(\lambda)}$ denote quantities on $\Gamma^{(\lambda)}$ such as $\kappa^{(\lambda)}$ (degree), $\theta_j^{(\lambda)}$ (eigenvalue of adjacency matrix $A^{(\lambda)}$), $m_j^{(\lambda)}$ (its multiplicity), $C^{(\lambda)}(t)$, $D^{(\lambda)}(k)$ etc. In particular, $m_1^{(\lambda)}$—the multiplicity of the second eigenvalue—is simply denoted by $m^{(\lambda)}$ because it frequently appears and plays an essential role in the sequel.

Theorem 1. *Assume* $\exists \lambda_0 \in \Lambda$ *such that*

$$M = \sup_{\lambda > \lambda_0} \sum_{\substack{\theta \le \theta_1^{(\lambda)} \\ \theta:\, eigenvalue\ of\ A^{(\lambda)}}} \frac{1}{{m^{(\lambda)}}^{(\kappa^{(\lambda)}-\theta)/(\kappa^{(\lambda)}-\theta_1^{(\lambda)})}} < \infty. \tag{19}$$

Here the sum over θ's is taken multiply according to the multiplicity of θ. Then, at time

$$t = \frac{\kappa^{(\lambda)}}{2(\kappa^{(\lambda)} - \theta_1^{(\lambda)})}(\log m^{(\lambda)} + c) \qquad (c > 0), \tag{20}$$

we have

$$C^{(\lambda)}(t) \le \frac{\sqrt{M}}{2} e^{-c/2}. \tag{21}$$

PROOF. We omit superscript $^{(\lambda)}$ for simplicity of notations. Putting t of (20) into (17), we have

$$\begin{aligned} C(t)^2 &\le \frac{1}{4}\sum_{j=1}^{d} m_j \exp\left\{ -\frac{\kappa - \theta_j}{\kappa - \theta_1}(\log m + c)\right\} \\ &= \frac{1}{4}\sum_{j=1}^{d} \exp\left(-\frac{\kappa - \theta_j}{\kappa - \theta_1}c\right) \frac{m_j}{m^{(\kappa-\theta_j)/(\kappa-\theta_1)}} \le \frac{1}{4}e^{-c}M \end{aligned}$$

namely (21). □

REMARK. Let us refer to the first nontrivial eigenvalue $1 - (\theta_1/\kappa)$ of $I - A/\kappa$ as the spectral gap of DRG Γ. Theorem 1 *cannot* be applied if the following two conditions hold:

(i) the spectral gap does not vanish for $\{\Gamma^{(\lambda)}\}_{\lambda \in \Lambda}$ i.e. $\liminf_{\lambda \to \infty} \cdot (1 - (\theta_1^{(\lambda)}/\kappa^{(\lambda)})) > 0$

(ii) $|X^{(\lambda)}|$ is of larger order than any polynomial in $m^{(\lambda)}$.

In fact, taking $0 < \delta < \liminf_{\lambda \to \infty}(1 - (\theta_1^{(\lambda)}/\kappa^{(\lambda)}))$, we have

$$\frac{\kappa - \theta_j}{\kappa - \theta_1} \le \frac{2\kappa}{\kappa - \theta_1} = \frac{2}{1 - (\theta_1/\kappa)} \le \frac{2}{\delta} \qquad \text{for large } \lambda$$

and hence

$$\sum_{j=1}^{d} \frac{m_j}{m^{(\kappa-\theta_j)/(\kappa-\theta_1)}} \ge \sum_{j=1}^{d} \frac{m_j}{m^{2/\delta}} = \frac{|X| - 1}{m^{2/\delta}} \to \infty$$

as $\lambda \to \infty$. We will recall this fact when we discuss the COP on DRGs of q-analogue type in §4.

THEOREM 2. *Assume* $\exists \lambda_0 \in \Lambda$ *such that*

$$N = \sup_{\lambda > \lambda_0} \sum_{\substack{\theta \le \theta_1^{(\lambda)} \\ \theta : \text{eigenvalue of } A^{(\lambda)}}} \frac{1}{{m^{(\lambda)}}^{\log|\theta/\kappa^{(\lambda)}|/\log(\theta_1^{(\lambda)}/\kappa^{(\lambda)})}} < \infty \tag{22}$$

and

$$|\theta_{d^{(\lambda)}}^{(\lambda)}| \le \theta_1^{(\lambda)} \qquad (\lambda > \lambda_0) \tag{23}$$

holds. Then, at time

$$k = \left\lceil \frac{\log m^{(\lambda)} + c}{2\log(\kappa^{(\lambda)}/\theta_1^{(\lambda)})} \right\rceil \qquad (c > 0), \tag{24}$$

we have

$$D^{(\lambda)}(k) \le \frac{\sqrt{N}}{2} e^{-c/2}.$$

PROOF. The proof goes ahead in parallel with that of Theorem 1. Note that (23) implies $|\theta_j| \le \theta_1$ for $j = 1, \ldots, d$. If a θ_j is equal to 0, we have only to adopt $\log 0 = -\infty$. □

REMARK 1. Theorem 1 and Theorem 2 show the importance of the degeneration of the second eigenvalue θ_1 in the upper estimate. However, the assumptions (19) and (22) are not stated in the way of having structural stability. This will be overcome if we replace multiplicity m of θ_1 by the cardinality of the (appropriately formulated) "second cluster of eigenvalues".

REMARK 2. In [18] and [19] we presented further statements by decomposing the assumptions (19) and (22) into smaller pieces of conditions. We omit them here because the decomposition might be a bit artificial.

3.3 Lower estimate

We begin with rewriting $D(k)$ and $C(t)$ given by (9) and (10) respectively. Let $\circ$ denote the Hadamard product of matrices and τ the summation of all the entries of a matrix. If S is an $|X| \times |X|$ stochastic matrix, $|X|^{-1}\tau(S \circ \cdot)$ defines the probability measure on $X \times X$ which assigns $|X|^{-1}\tau(S \circ A)$ to matrix A regarded as a function on $X \times X$. Then (9) and (10) yield

$$D(k) = \| |X|^{-1}\tau(P^k \circ \cdot) - |X|^{-1}\tau(E_0 \circ \cdot) \| \tag{25}$$

$$C(t) = \| |X|^{-1}\tau(e^{t(P-I)} \circ \cdot) - |X|^{-1}\tau(E_0 \circ \cdot) \|. \tag{26}$$

The following inequality plays a basic role in our lower estimate.

PROPOSITION 3. *Let Q_1 and Q_2 be probabilities on a measurable space $(\Omega, \mathscr{B})$. Assume that we find an* **R**-*valued measurable function f on Ω which has mean $\mu > 0$ [resp. 0] and variance $\sigma^2 > 0$ [resp. 1] with respect to Q_1 [resp. Q_2]. Then, for any r such that $0 < r < \mu$, we have*

$$\|Q_1 - Q_2\| = \max_{B \in \mathscr{B}} |Q_1(B) - Q_2(B)| \geq 1 - \frac{1}{r^2} - \frac{\sigma^2}{(\mu - r)^2}. \tag{27}$$

PROOF. Using the Chebychev inequality, we have

$$Q_2(|f| \leq r) \geq 1 - r^{-2},$$

$$Q_1(|f| \leq r) \leq Q_1(|f - \mu| \geq \mu - r) \leq \sigma^2/(\mu - r)^2$$

and hence the desired inequality. □

In [14], Diaconis and Shahshahani developed a method to obtain good lower bounds by combining the above inequality with elementary (or zonal) spherical functions in the case of a Gel'fand pair of a symmetric group and its subgroup. See also [7] Chapter 3, [23] §5, and [24] §2. In this subsection we modify their method to be adapted to our aim.

Let us consider a directed family $\{\Gamma^{(\lambda)}\}_{\lambda \in \Lambda}$ of *Q-polynomial* DRGs (each DRG having diameter $d^{(\lambda)}$) and a simple random walk on $\Gamma^{(\lambda)}$ starting at some vertex. See §§2.1 for the definition of a Q-polynomial DRG.

THEOREM 3. *Assume that*

$$m^{(\lambda)} \to \infty \qquad \textit{as } \lambda \to \infty \tag{28}$$

$$\frac{\kappa^{(\lambda)} - \theta_2^{(\lambda)}}{2(\kappa^{(\lambda)} - \theta_1^{(\lambda)})} = 1 + \frac{o(1)}{\log m^{(\lambda)}} \qquad \textit{as } \lambda \to \infty \tag{29}$$

$$\exists \beta > 0 \textit{ and } \exists \lambda_1 \in \Lambda \textit{ such that } \forall \lambda > \lambda_1, \qquad (q_{11}^{1(\lambda)})^2 \leq \beta^2 m^{(\lambda)}. \tag{30}$$

(q_{11}^1 is a Krein parameter which appeared in (7).*) Then, at time*

$$t = \frac{\kappa^{(\lambda)}}{2(\kappa^{(\lambda)} - \theta_1^{(\lambda)})}(\log m^{(\lambda)} - c) \qquad (0 < c < \log m^{(\lambda)}), \tag{31}$$

we have $\forall \varepsilon > 0$, $\exists c_\varepsilon > 0$ and $\lambda_\varepsilon \in \Lambda$ such that

$$\lambda > \lambda_\varepsilon \textit{ and } \log m^{(\lambda)} > c > c_\varepsilon \qquad \Rightarrow \qquad C^{(\lambda)}(t) > 1 - \varepsilon. \tag{32}$$

PROOF. For simplicity of notations, we omit superscript $^{(\lambda)}$. The terminology (including the spherical functions) and some technicalities used here are summarized in Appendix 2.

(Step 1) We apply Proposition 3 to (26), putting $\Omega = X \times X$, $Q_1 = |X|^{-1}\tau(e^{t(P-I)} \circ \cdot)$, $Q_2 = |X|^{-1}\tau(E_0 \circ \cdot)$ and taking as f the *1st elementary spherical function on a DRG* (see §§A2.1) under necessary normalization. The point-wise product of functions on Ω agrees with the Hadamard product of matrices. From (62) and (63) we see

$$\frac{1}{|X|}\tau\left(E_0 \circ \sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}A_i\right) = \frac{1}{|X|}\sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}p_i(0) = 0,$$

$$\frac{1}{|X|}\tau\left(E_0 \circ \left(\sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}A_i\right)^{\circ 2}\right) = \frac{1}{|X|}\sum_{i=0}^{d}\frac{p_i(1)^2}{\kappa_i^2}\kappa_i = \frac{1}{m},$$

and hence that $f = \sqrt{m}\sum_{i=0}^{d}(p_i(1)/\kappa_i)A_i$ has mean 0 and variance 1 with respect to Q_2. Let μ and σ^2 denote the mean and the variance of f with respect to Q_1 respectively. (62)–(65) yield

$$\begin{aligned}\mu &= \frac{1}{|X|}\tau\left(e^{(t(P-I)} \circ \sqrt{m}\sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}A_i\right) = \frac{\sqrt{m}}{|X|}\sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}\sum_{j=0}^{d}\exp t\left(\frac{\theta_j}{\kappa}-1\right)\tau(E_j \circ A_i)\\ &= \frac{\sqrt{m}}{|X|}\sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}\sum_{j=0}^{d}m_j p_i(j)\exp t\left(\frac{\theta_j}{\kappa}-1\right)\\ &= \frac{\sqrt{m}}{|X|}\sum_{j=0}^{d}m_j\exp t\left(\frac{\theta_j}{\kappa}-1\right)\frac{|X|}{m}\delta_{1j} = \sqrt{m}\exp t\left(\frac{\theta_1}{\kappa}-1\right)\end{aligned}$$

and

$$\begin{aligned}&\frac{1}{|X|}\tau\left(e^{(t(P-I)} \circ \left(\sqrt{m}\sum_{i=0}^{d}\frac{p_i(1)}{\kappa_i}A_i\right)^{\circ 2}\right)\\ &\quad = \frac{m}{|X|}\sum_{i=0}^{d}\frac{p_i(1)^2}{\kappa_i^2}\sum_{j=0}^{d}m_j p_i(j)\exp t\left(\frac{\theta_j}{\kappa}-1\right)\\ &\quad = \frac{m}{|X|}\sum_{j=0}^{d}m_j\exp t\left(\frac{\theta_j}{\kappa}-1\right)q_{1j}^1\frac{|X|}{mm_j}\\ &\quad = 1 + q_{11}^1\exp t\left(\frac{\theta_1}{\kappa}-1\right) + (m-1-q_{11}^1)\exp t\left(\frac{\theta_2}{\kappa}-1\right)\end{aligned}$$

where tridiagonality of $[q_{1i}^j]_{i,j}$ and $q_{10}^1 = 1$ (directly seen from (7)) are used in the last equality. Combining these with (31), we have

$$\mu = e^{c/2} \tag{33}$$

$$\sigma^2 = 1 + \frac{q_{11}^1}{\sqrt{m}} e^{c/2} + (m - 1 - q_{11}^1) \exp\left\{ -\frac{\kappa - \theta_2}{2(\kappa - \theta_1)} (\log m - c) \right\} - e^c. \tag{34}$$

(Step 2) Proposition 3 and (26) yield, for $0 < 2r < \mu$,

$$C(t) = \|Q_1 - Q_2\| \geq 1 - \frac{1}{r^2} - \frac{\sigma^2}{(\mu - r)^2}$$

$$= 1 - \frac{1}{r^2} - \frac{\sigma^2}{\mu^2}\left(1 - \frac{r}{\mu}\right)^{-2} \geq 1 - \frac{1}{r^2} - \frac{4\sigma^2}{\mu^2}. \tag{35}$$

Note that μ and σ^2 depend on both λ and c. Applying (29) and (30) to (33) and (34), we have

$$0 < \frac{\sigma^2}{\mu^2} \leq e^{-c} + \frac{q_{11}^1}{\sqrt{m}} e^{-c/2} + m \exp\left\{ \left(1 + \frac{o(1)}{\log m}\right)(c - \log m) \right\} e^{-c} - 1$$

$$\leq e^{-c} + \beta e^{-c/2} + \exp\left\{ \left(\frac{c}{\log m} - 1\right) o(1) \right\} - 1 \qquad (\text{if } \lambda > \lambda_1)$$

$$= e^{-c} + \beta e^{-c/2} + o(1).$$

(28) enables us to take $c(< \log m) \to \infty$ under $\lambda \to \infty$. Hence we get $\sigma^2/\mu^2 \to 0$ as $\lambda \to \infty$ and $c \to \infty$ in (35). Moreover, $r(< \mu/2 = e^{c/2}/2)$ can be arbitrarily large whenever $c \to \infty$ in (35). This completes the proof of (32). □

THEOREM 4. *Assume that*

$$\left\{ \frac{\log(\kappa^{(\lambda)}/\theta_1^{(\lambda)})}{\log m^{(\lambda)}} \,\middle|\, \lambda \in \Lambda \right\} \text{ is bounded above and } \theta_2^{(\lambda)} > 0 \qquad (\forall \lambda \in \Lambda) \tag{36}$$

$$\frac{\log(\theta_2^{(\lambda)}/\kappa^{(\lambda)})}{2\log(\theta_1^{(\lambda)}/\kappa^{(\lambda)})} = 1 + \frac{o(1)}{\log m^{(\lambda)}} \qquad \text{as } \lambda \to \infty \tag{37}$$

as well as (28) *and* (30). *Then, at time*

$$k = \left\lfloor \frac{\log m^{(\lambda)} - c}{2\log(\kappa^{(\lambda)}/\theta_1^{(\lambda)})} \right\rfloor \qquad (0 < c < \log m^{(\lambda)}), \tag{38}$$

we have $\forall \varepsilon > 0$, $\exists c_\varepsilon > 0$ *and* $\exists \lambda_\varepsilon \in \Lambda$ *such that*

$$\lambda > \lambda_\varepsilon \quad \text{and} \quad \log m^{(\lambda)} > c > c_\varepsilon \quad \Rightarrow \quad D^{(\lambda)}(k) > 1 - \varepsilon.$$

PROOF. The proof of Theorem 3 almost works well under the replacement of $e^{t(P-I)}$ by P^k. Instead of (33) and (34) we now see

$$\mu = \sqrt{m}\left(\frac{\theta_1}{\kappa}\right)^k \geq e^{c/2} \tag{39}$$

$$\sigma^2 = 1 + q_{11}^1\left(\frac{\theta_1}{\kappa}\right)^k + (m - 1 - q_{11}^1)\left(\frac{\theta_2}{\kappa}\right)^k - m\left(\frac{\theta_1}{\kappa}\right)^{2k}. \tag{40}$$

Applying (30) to (39) and (40), we have

$$0 < \frac{\sigma^2}{\mu^2} \leq e^{-c} + \beta e^{-c/2} + \frac{(\theta_2/\kappa)^k}{(\theta_1/\kappa)^{2k}} - 1. \tag{41}$$

Here (37) and (38) yield

$$\frac{(\theta_2/\kappa)^k}{(\theta_1/\kappa)^{2k}} = \exp\left\{k\frac{2\log(\theta_1/\kappa)}{\log m}o(1)\right\},$$

$$\frac{c}{\log m} - 1 \leq k\frac{2\log(\theta_1/\kappa)}{\log m} < \frac{c}{\log m} - 1 + \frac{2\log(\kappa/\theta_1)}{\log m}.$$

Putting these into (41) and combining (36) with it, we get

$$0 < \frac{\sigma^2}{\mu^2} \leq e^{-c} + \beta e^{-c/2} + o(1).$$

This inequality completes the proof under (28) exactly in the same way as Step 2 in the proof of Theorem 3. □

3.4 Criterion for the COP

Combining Theorems 1–4 with Definition 3, we are led to the conditions under which the COP occurs for *simple random walks on Q-polynomial DRGs.* In Definition 3, we required that the critical time should be of smaller order than the cardinality of the state space. This is motivated by a physical point of view (Remark 1 after Definition 3) and, however, is not logically indispensable for deducing the COP. Hence we will not include here "$k_c^{(\lambda)}/|X^{(\lambda)}| \to 0$ as $\lambda \to \infty$" in the criterion. It is directly verified in concrete models.

- The case of continuous time

Under (19), (28), (29) *and* (30), *one observes the COP. Seen from* (20) *and* (31), *the critical time to reach equilibrium is given by*

$$t_c^{(\lambda)} = \frac{\kappa^{(\lambda)}}{2(\kappa^{(\lambda)} - \theta_1^{(\lambda)})}\log m^{(\lambda)}, \tag{42}$$

and furthermore the fluctuation around the critical time by

$$h_\varepsilon^{(\lambda)} \asymp \frac{\kappa^{(\lambda)}}{2(\kappa^{(\lambda)} - \theta_1^{(\lambda)})}.$$

- The case of discrete time

Under (22), (23), (28), (30), (36) *and* (37), *one observes the COP. The critical time to reach equilibrium is given by*

$$k_c^{(\lambda)} = \frac{\log m^{(\lambda)}}{2\log(\kappa^{(\lambda)}/\theta_1^{(\lambda)})}, \tag{43}$$

and the fluctuation around the critical time by

$$h_\varepsilon^{(\lambda)} \asymp \frac{1}{2\log(\kappa^{(\lambda)}/\theta_1^{(\lambda)})}. \tag{44}$$

Remark 1. If the spectral gap $1-(\theta_1^{(\lambda)}/\kappa^{(\lambda)}) \to 0$ as $\lambda \to \infty$, critical times (42) and (43) asymptotically coincide. If the spectral gap does not vanish, some difference can occur between the continuous and the discrete time cases. See Remark after Theorem 1.

Remark 2. The conditions summarized above are written in terms of the spectrum of adjacency matrix A

$$\begin{pmatrix} \theta_0 & \theta_1 & \dots & \theta_d \\ m_0 & m_1 & \dots & m_d \end{pmatrix} \quad \text{where} \quad \begin{array}{l} \theta_0 = \kappa \\ m_0 = 1, \quad m_1 = m \end{array}$$

and Krein parameter q_{11}^1 of the DRG considered. Here we note that q_{11}^1 can be computed from projector E_1 in the spectral decomposition (5) of A. Indeed, multiplying E_1 in (7) for $i = j = 1$, we have

$$(|X|E_1)^{\circ 2}E_1 = \sum_h q_{11}^h |X|E_h E_1 = q_{11}^1 |X|E_1.$$

Then taking the trace yields

$$q_{11}^1 = \frac{|X|}{m}\,\mathrm{tr}(E_1^{\circ 2}E_1). \tag{45}$$

Consequently, it suffices to observe the spectral data of the adjacency matrix in order to apply our criterion for the COP.

4. Application to concrete models

The simple random walks on Hamming graph $H(d,n)$ and Johnson graph $J(v,d)$ are classical models imitating diffusion of sparse gases, which are called the Ehrenfests urn model and the Bernoulli-Laplace diffusion model respectively. As is cited in Introduction, the COP for them has been well studied. Hence we do not present the details of our criteria applied to $H(d,n)$ and $J(v,d)$ but summarize the results in Appendix 1 for convenience.

In this section, we illustrate the effect of our criteria for the COP with

- DRG associated with quadratic forms over a finite field
- bipartite half of Hamming graph $H(r,2)$.

We refer to [2] Chapter III for Q-polynomiality and spectral data of these DRGs.

4.1 Quadratic forms over a finite field

Fix a finite field $GF(p^f)$ where $p \neq 2$ is a prime number. Let X be the set of the $(n-1)\times(n-1)$ symmetric matrices over $GF(p^f)$. Joining two vertices $x, y \in X$ by an edge if $\operatorname{rank}(x-y) = 1$ or 2, one gets a Q-polynomial DRG $\Gamma^{(n)}$. The diameter of $\Gamma^{(n)}$ is $d = \lfloor n/2 \rfloor$. Set $R_i = \{(x,y) \in X \times X \mid \partial(x,y) = i\}$ $(i = 0,1,\dots,d)$ where ∂ denotes the canonical distance on X. One sees $(x,y) \in R_i \Leftrightarrow \operatorname{rank}(x-y) = 2i-1$ or $2i$. It is known that $\Gamma^{(n)}$ is *not* a distance-transitive graph. (A finite connected graph is said to be distance-transitive if the automorphism group of the graph acts transitively on each R_i.) See [2] and [5]. Hence the present model really goes beyond the framework of homogeneous spaces of groups. We have the spectral data of $\Gamma^{(n)}$:

$$\theta_j = (q-1)^{-1}(1 + q^{n-j-1/2} - q^{(n-1)/2} - q^{n/2}) \qquad (j = 0,1,\dots,d) \tag{46}$$

$$m_j = q^{j(n-j-1/2)} \prod_{i=1}^{j} \frac{(1-q^{i-1-n/2})(1-q^{i-(n+1)/2})}{1-q^{-i}} \qquad (j = 1,\dots,d) \tag{47}$$

$$q_{11}^1 = q^{n/2}(q-1)^{-1}(q + q^{1/2} - q^{-1/2} - 1 + q^{-n/2} - q^{-(n-3)/2}) - 1 \tag{48}$$

where we set $q = p^{2f}$. Note that the conditions in Remark after Theorem 1 hold, namely (i) the spectral gap does not vanish:

$$1 - \frac{\theta_1}{\kappa} = \frac{q^{n-1/2} - q^{n-3/2}}{1 + q^{n-1/2} - q^{(n-1)/2} - q^{n/2}} \to 1 - \frac{1}{q} > 0 \qquad \text{as } n \to \infty$$

and (ii) $|X| \asymp q^{n^2 \times \text{constant}} \gg$ any polynomial of $m \asymp q^{n \times \text{constant}}$ as $n \to \infty$. Since Theorem 1 cannot be applied in this situation, we consider only the discrete time simple random walk on $\Gamma^{(n)}$.

Let us check the conditions summarized in §§3.4 under the assumption that n is sufficiently large. In the following, the terms indicated by the Landau notation $O(\cdot)$ depend only on n and are independent of j. Note $\theta_d < 0$ and $\theta_{d-1} > 0$. (23) follows from

$$\theta_1 - |\theta_d| = \theta_1 + \theta_d > (q-1)^{-1}(q^{n-3/2} - 4q^{n/2}) > 0.$$

(46) yields

$$\begin{aligned}
&\log(\theta_1/\kappa) = -\log q + O(q^{-n/2}) \\
&\log(\theta_j/\kappa) = -j\log q + \log(1 + q^{j-n+1/2} - q^{j-n/2} - q^{j-(n-1)/2}) + O(q^{-n/2}) \\
&\qquad (j = 1, \dots, d-1) \\
&\log|\theta_d/\kappa| = -d\log q \pm (\log q)/2 + O(q^{-n/2}) \quad \text{when } n \text{ is even/odd respectively}
\end{aligned} \tag{49}$$

and hence, for $j = 1, \dots, d-1$,

$$\frac{\log(\theta_j/\kappa)}{\log(\theta_1/\kappa)} \geq \frac{j + O(q^{-n/2})}{1 + O(q^{-n/2})}, \qquad \frac{\log|\theta_d/\kappa|}{\log(\theta_1/\kappa)} \geq \frac{d - 1/2 + O(q^{-n/2})}{1 + O(q^{-n/2})}. \tag{50}$$

On the other hand, (47) yields

$$m_j \leq \gamma q^{j(n-j-1/2)} \qquad \text{where } 1 < \gamma = 1\Big/\prod_{i=1}^{\infty}(1 - q^{-i}) < \infty. \tag{51}$$

Combining (50) and (51) with $\log m = n\log q + O(1)$, we see

$$\begin{aligned}
\sum_{\theta \leq \theta_1} \frac{1}{m^{\log|\theta/\kappa|/\log(\theta_1/\kappa)}} &= \sum_{j=1}^{d} \frac{m_j}{m^{\log|\theta_j/\kappa|/\log(\theta_1/\kappa)}} \\
&\leq \sum_{j=1}^{d-1} \frac{\gamma q^{j(n-j-1/2)}}{(q^n e^{O(1)})^{(j+O(q^{-n/2}))/(1+O(q^{-n/2}))}} \\
&\quad + \frac{\gamma q^{d(n-d-1/2)}}{(q^n e^{O(1)})^{(d-1/2+O(q^{-n/2}))/(1+O(q^{-n/2}))}}
\end{aligned}$$

and hence (22). (28) and (36) are obvious from (46) and (47). (49) yields

$$\frac{\log(\theta_2/\kappa)}{2\log(\theta_1/\kappa)} - 1 = \frac{O(q^{-n/2})}{-2\log q + O(q^{-n/2})},$$

which implies (37) holds. Lastly, (30) immediately follows from (48).

We thus observe the COP for the discrete time simple random walks. The critical time (43) is expressed as

$$k_c^{(n)} = \frac{n - 3/2 - \log(1 - q^{-1})/\log q + O(q^{-n/2})}{2 + O(q^{-n/2})} \sim \frac{n}{2} \sim d \qquad \text{as } n \to \infty. \tag{52}$$

$k_c^{(n)}$ is of far smaller order than $|X^{(n)}|$. $h_\varepsilon^{(n)}$ in (44) remains finite as $n \to \infty$. Hence one finds the COP quite sharp.

REMARK 1. Similar analysis proceeds on other DRGs of q-analogue type including a DRG associated with bilinear forms over a finite field (q-Hamming graph) and q-analogue of the Johnson graph (§§A1.3). The discrete time simple random walks on them admit the critical time asymptotically equal to diameter d. See [4] for more detailed information.

REMARK 2. It is obvious that the critical time to reach equilibrium for the discrete time simple random walk must be, if it exists, at least diameter d of the graph. Hence the lower estimate is conceptually easy in the case of q-analogue type DRGs. It may be surprising that time d is also sufficient as is shown in (52). To find the reason, one can check so rapid growth of $\kappa_i = |R_i|$ that R_d occupies a dominantly large part in the decomposition $X \times X = \bigcup_{i=0}^{d} R_i$. This is also the case for other q-analogue type DRGs.

4.2 Bipartite half of $H(r, 2)$

Since Hamming graph $H(r, 2)$ is bipartite, the (discrete time) simple random walk on it is periodic. Defining anew two vertices $x, y \in \{0, 1\}^r$ to be adjacent if they are at distance 2 in $H(r, 2)$, one gets a bipartite half $\Gamma^{(r)}$ of $H(r, 2)$. $\Gamma^{(r)}$ is a Q-polynomial DRG with 2^{r-1} vertices and diameter $d = \lfloor r/2 \rfloor$. The simple random walk on $\Gamma^{(r)}$ is slightly different from the restriction of that on $H(r, 2)$ in even steps in that the latter walk can come back to the initial vertex after 2 steps. We have the spectral data of $\Gamma^{(r)}$:

$$\theta_j = r(r-1)/2 - 2j(r-j) \qquad (j = 0, 1, \ldots, d) \tag{53}$$

$$m_j = \binom{r}{j} \quad (j = 0, 1, \ldots, d-1) \quad m_d = \frac{1}{2^\varepsilon}\binom{r}{d} \quad \varepsilon = \begin{cases} 0 & \text{if } r \text{ is odd} \\ 1 & \text{if } r \text{ is even} \end{cases} \tag{54}$$

$$q_{11}^1 = 0. \tag{55}$$

We consider the simple random walk on $\Gamma^{(r)}$ in continuous and discrete time. The spectral gap vanishes:

$$1 - (\theta_1/\kappa) = 4/r \to 0 \qquad \text{as } r \to \infty.$$

Let us check the conditions in §§3.4 under the assumption that r is sufficiently large. We begin with the upper estimate for the continuous time case. (53) and (54) yield

$$\frac{\kappa - \theta_j}{\kappa - \theta_1} > j\left(1 - \frac{j}{r}\right), \qquad \log m_j \le j \log r - \log j!, \qquad m = r.$$

Hence we have

$$\begin{aligned} \frac{m_j}{m^{(\kappa-\theta_j)/(\kappa-\theta_1)}} &\le \exp\{j \log r - \log j! - j(1 - (j/r)) \log r\} \\ &\le \exp\{-j(\log j - 1 - (j/r) \log r)\}, \end{aligned}$$

where we used $\log j! \geq j \log j - j$. Taking δ such that

$$\lim_{r\to\infty} \min_{3\leq j\leq r/2} (\log j - 1 - (j/r)\log r) > \delta > 0,$$

we have, if r is sufficiently large,

$$\sum_{j=1}^{d} \frac{m_j}{m^{(\kappa-\theta_j)/(\kappa-\theta_1)}} \leq 1 + \frac{1}{2} e^{(4/r)\log r} + \sum_{j=3}^{\infty} e^{-j\delta}. \tag{56}$$

This implies (19).

In the case of the discrete time random walk, we first note that

$$\theta_j \geq 0 \quad \Leftrightarrow \quad j \leq (r - \sqrt{r})/2$$

holds from (53). (23) is immediately verified. For $j \leq (r - \sqrt{r})/2$, the concavity of $\log x$ yields

$$\frac{\log(\theta_j/\kappa)}{\log(\theta_1/\kappa)} = \frac{\log(1 - (\kappa - \theta_j)/\kappa)}{\log(1 - (\kappa - \theta_1)/\kappa)} \geq \frac{\kappa - \theta_j}{\kappa - \theta_1}.$$

Hence we have

$$\sum_{j=1}^{\lfloor (r-\sqrt{r})/2 \rfloor} \frac{m_j}{m^{\log(\theta_j/\kappa)/\log(\theta_1/\kappa)}} \leq \sum_{j=1}^{\lfloor (r-\sqrt{r})/2 \rfloor} \frac{m_j}{m^{(\kappa-\theta_j)/(\kappa-\theta_1)}} \leq \text{RHS of (56)}. \tag{57}$$

For $j > (r - \sqrt{r})/2$, we see from $0 < |\theta_j| \leq |\theta_d|$ and the concavity of $\log x$ that

$$\frac{\log|\theta_j/\kappa|}{\log(\theta_1/\kappa)} \geq \frac{\log|\theta_d/\kappa|}{\log(\theta_1/\kappa)} > \frac{\kappa - |\theta_d|}{\kappa - \theta_1} = \frac{r}{2} - \frac{d(r-d)}{r-1}.$$

Combining this with

$$m_j \leq \exp(j \log r - j \log j + j) \leq \exp\{r(1 + \log 2)/2\} \qquad \text{for } j \leq r/2,$$

we have

$$\begin{aligned}
&\sum_{j=\lceil (r-\sqrt{r})/2 \rceil}^{\lfloor r/2 \rfloor} \frac{m_j}{m^{\log|\theta_j/\kappa|/\log(\theta_1/\kappa)}} \\
&\leq \frac{\sqrt{r}}{2} \exp\left\{\frac{r}{2}(1+\log 2) - \left(\frac{r}{2} - \frac{d(r-d)}{r-1}\right)\log r\right\} \\
&\leq \frac{1}{2} \exp\left\{\frac{r}{2}(1+\log 2) + \frac{1}{2}\log r - \left(\frac{r}{2} - \frac{r^2}{4(r-1)}\right)\log r\right\} \qquad \text{since } d = \left\lfloor \frac{r}{2} \right\rfloor \\
&= \frac{1}{2} \exp r\left\{\frac{1+\log 2}{2} - \left(-\frac{1}{2r} + \frac{1}{2} - \frac{r}{4(r-1)}\right)\log r\right\} \to 0 \qquad \text{as } r \to \infty.
\end{aligned} \tag{58}$$

(57) and (58) imply (22).

It is easy to verify the conditions for the lower estimate. (28), (30) and (36) are obvious from (53)–(55). Finally, (29) and (37) follow from

$$\frac{\kappa-\theta_2}{2(\kappa-\theta_1)}-1=-\frac{1}{r-1} \qquad \text{and} \qquad \frac{\log(\theta_2/\kappa)}{2\log(\theta_1/\kappa)}-1=\frac{O(1/r^2)}{-(8/r)+O(1/r^2)}=O\left(\frac{1}{r}\right)$$

respectively.

We thus observe the COP for the simple random walks in either continuous or discrete time. The critical time (43) is given asymptotically by

$$t_c^{(r)} \sim k_c^{(r)} \sim \frac{r}{8}\log r \qquad \text{as } r \to \infty. \tag{59}$$

Needless to say, (59) is of smaller order than $|X^{(r)}|=2^{r-1}$.

Appendix 1. Related well-known models

We summarize the results on the COP for some well-known models. All the spectral data needed here are found in [2] Chapter III.

A1.1 The Ehrenfests urn model

This is the simple random walk on Hamming graph $H(d,n)$, though the original model is the case of $n=2$. To avoid periodicity for $n=2$, however, one usually takes $P=(I+A)/(\kappa+1)$ as the transition matrix instead of $P=A/\kappa$ (in other words, allows the walker to pause at each step). As spectral data of $H(d,n)$, one has

$$\theta_j=(n-1)d-nj, \qquad m_j=(n-1)^j\binom{d}{j} \qquad (j=0,1,\dots,d), \qquad q_{11}^1=n-2.$$

In particular, (30) holds if n/d is bounded. Under this assumption, one observes the COP for the simple random walk on $H(d,n)$ ($n\ge 3$) in either continuous or discrete time. The critical time is given by

$$t_c^{(d,n)} \sim k_c^{(d,n)} \sim \frac{d}{2}\left(1-\frac{1}{n}\right)(\log d+\log n).$$

Furthermore, $c(\theta)$ in (18), which describes the small fluctuation around the critical time, is explicitly computed. It takes different expressions according to the way of large volume limit $(d,n)\to\infty$ as was shown in [20].

A1.2 The Bernoulli-Laplace diffusion model

This is the simple random walk on Johnson graph $J(v,d)$, where one assumes $2d\le v$ without loss of generality. As spectral data of $J(v,d)$, one has

$$\theta_j = d(v-d) - j(v-j+1) \qquad (j = 0, 1, \ldots, d),$$

$$m_j = \binom{v}{j} - \binom{v}{j-1} \qquad (j = 1, \ldots, d), \qquad q_{11}^1 = v\left\{1 - \frac{v(v-d-1)(d-1)}{(v-2)(v-d)d}\right\} - 2.$$

In particular, (30) holds if v/d^2 is bounded. Then, one observes the COP in either continuous or discrete time with the critical time

$$t_c^{(v,d)} \sim k_c^{(v.d)} \sim \frac{d}{2}\left(1 - \frac{d}{v}\right) \log v.$$

A1.3 q-analogue of the Bernoulli-Laplace diffusion model

Replacing the symmetric group by the general linear group over a finite field, one has q-analogue of the Johnson graph. More precisely, let V be a v-dimensional vector space over $GF(q)$ and X the set of the d-dimensional subspaces of V where $2d \leq v$. Joining two vertices $x, y \in X$ if $\dim(x \cap y) = d-1$, one gets Q-polynomial DRG $J_q(v,d)$ with diameter d. Its spectral data are given by

$$\theta_j = q[d][v-d] - [j][v-j+1] \qquad (j = 0, 1, \ldots, d),$$

$$m_j = \begin{bmatrix} v \\ j \end{bmatrix} - \begin{bmatrix} v \\ j-1 \end{bmatrix} \qquad (j = 1, \ldots, d), \qquad q_{11}^1 = [v]\left(1 - \frac{[v][v-d-1][d-1]}{[v-2][v-d][d]}\right) - 2$$

where $[\cdot] = [\cdot]_q$ denotes the Gauss q-integer and $\begin{bmatrix} \cdot \\ \cdot \end{bmatrix}$ the q-binomial coefficient. Let q be fixed. The spectral gap does not vanish as $d \to \infty$. (30) holds if $v - 2d$ is bounded. Then, one observes the COP for the discrete time simple random walk on $J_q(v,d)$ with the critical time

$$k_c^{(v,d)} \sim v/2 \qquad \text{as } d \to \infty.$$

Appendix 2. Technicalities on DRG and adjacency algebra

A2.1 Elementary spherical function

We give a brief explanation of elementary spherical functions on a DRG which were used in the proof of Theorem 3 and Theorem 4 in §§3.3. To make a comparison, let G be a locally compact separable unimodular topological group and K a compact subgroup of G. Assume that (G,K) is a Gel'fand pair, namely $L_1(K\backslash G/K)$ is a commutative algebra with respect to the convolution product. An elementary spherical function ϕ on (G,K) is characterized as a normalized (i.e. $\phi(e) = 1$) K-bi-invariant simultaneous eigenfunction on G of the convolution operators $f * \cdot$ with f running over the elements in the Hecke algebra. In the present context of a DRG, the Hecke algebra should be replaced by the adjacency algebra while the convolution corresponds to the usual multiplication of matrices.

LEMMA A1. *Let $\mathscr{A}$ be the adjacency algebra of DRG Γ with vertex set X and diameter d. For each $j = 0, 1, \dots, d$, $\sum_{i=0}^{d}(p_i(j)/\kappa_i)A_i$ is a normalized simultaneous eigenfunction of $\mathscr{A}$ with respect to the multiplication in $\mathscr{A}$.*

PROOF. Normalization means that the diagonal entries are all 1. First we note that $p_{hi}^{l}\kappa_l = p_{hl}^{i}\kappa_i$ holds. (Recall (3) and $\kappa_i = p_{ii}^{0}$.) Indeed, count up the triangles xyz such that $\partial(x, y) = l$, $\partial(x, z) = h$ and $\partial(z, y) = i$ in two ways. For each fixed $j = 0, 1, \dots, d$, we have

$$A_h\left(\sum_{i=0}^{d}\frac{p_i(j)}{\kappa_i}A_i\right) = \sum_{i=0}^{d}\frac{p_i(j)}{\kappa_i}\sum_{l=0}^{d}p_{hi}^{l}A_l = \sum_{l=0}^{d}\left(\sum_{i=0}^{d}p_i(j)p_{hl}^{i}\right)\frac{A_l}{\kappa_l}. \tag{60}$$

Set $B_n = [p_{ni}^{h}]_{i,h}$ and $\Theta = [p_k(l)]_{l,k}$. Then we have

$$\Theta\,{}^{t}B_n = \mathrm{diag}(p_n(0), p_n(1), \dots, p_n(d))\Theta. \tag{61}$$

In fact, expressing a tensor product of matrices with respect to the basis of lexicographical order and denoting by I the identity matrix of degree $|X|$, we see from (6)

$$\begin{aligned}
&[E_0E_1\dots E_d](\Theta\otimes I)({}^{t}B_n\otimes I)\\
&\quad= [A_0A_1\dots A_d]({}^{t}B_n\otimes I)\\
&\quad= [\dots\sum_{h}p_{ni}^{h}A_h\dots] = [\dots A_nA_i\dots] \qquad \text{(the ith component)}\\
&\quad= A_n[A_0A_1\dots A_d] = A_n[E_0E_1\dots E_d](\Theta\otimes I)\\
&\quad= [\dots\sum_{h}p_n(h)E_hE_i\dots](\Theta\otimes I) = [\dots p_n(i)E_i\dots](\Theta\otimes I)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(the ith component)}\\
&\quad= [E_0E_1\dots E_d](\mathrm{diag}(p_n(0), p_n(1), \dots, p_n(d))\otimes I)(\Theta\otimes I).
\end{aligned}$$

(61) yields

$$\sum_{i=0}^{d}p_i(j)p_{hl}^{i} = (j, l) \text{ entry of } \Theta\,{}^{t}B_h = p_h(j)p_l(j).$$

Putting this into (60), we see that $\sum_{i=0}^{d}(p_i(j)/\kappa_i)A_i$ is an eigenfunction of A_h belonging to eigenvalue $p_h(j)$. □

Lemma A1 enables us to call $\sum_{i=0}^{d}(p_i(j)/\kappa_i)A_i$ the jth elementary spherical function on Γ.

A2.2 Some formulas needed

For the sake of convenience, we prove the formulas used in the proof of Theorem 3. Recall that τ denotes summation of all the entries of a matrix.

Lemma A2. *The following hold on a DRG with vertex set X and diameter d:*

$$\tau(E_j \circ A_i) = m_j p_i(j) \tag{62}$$

$$\sum_{\nu=0}^{d} \frac{1}{\kappa_\nu} p_\nu(i) p_\nu(j) = \frac{|X|}{m_i}\delta_{ij} \tag{63}$$

$$q_{ij}^h = \frac{m_i m_j}{|X|} \sum_{\nu=0}^{d} \frac{1}{\kappa_\nu^2} p_\nu(i) p_\nu(j) p_\nu(h) \tag{64}$$

$$\sum_{j=0}^{d} q_{ij}^h = m_i. \tag{65}$$

Proof. As coefficients of the base change in the adjacency algebra, $q_j(h)$'s are determined by

$$|X|E_j = \sum_{h=0}^{d} q_j(h) A_h \qquad (j = 0, 1, \ldots, d).$$

First we note

$$\begin{aligned} \operatorname{tr} A_i E_j &= \operatorname{tr} \frac{1}{|X|} \sum_{h=0}^{d} q_j(h) A_i A_h = \operatorname{tr} \frac{1}{|X|} \sum_{l=0}^{d} \sum_{h=0}^{d} q_j(h) p_{ih}^l A_l \\ &= \sum_{h=0}^{d} q_j(h) p_{ih}^0 = q_j(i)\kappa_i \end{aligned} \tag{66}$$

$$\operatorname{tr} A_i E_j = \operatorname{tr} \sum_{h=0}^{d} p_i(h) E_h E_j = \operatorname{tr} p_i(j) E_j = m_j p_i(j),$$

and hence get

$$m_j p_i(j) = q_j(i)\kappa_i. \tag{67}$$

(67) yields (62) since

$$\tau(E_j \circ A_i) = \frac{1}{|X|}\sum_{h=0}^{d} q_j(h)\tau(A_h \circ A_i) = q_j(i)\kappa_i.$$

Taking the trace of

$$|X|E_iE_j = \sum_{\nu=0}^{d} q_i(\nu)A_\nu E_j, \tag{68}$$

we have (63). In fact, (66) and (67) yield

$$\mathrm{tr}(\text{LHS of } (68)) = |X|\delta_{ij}m_j$$

$$\mathrm{tr}(\text{RHS of } (68)) = \sum_{\nu=0}^{d} q_i(\nu)\,\mathrm{tr}(A_\nu E_j) = \sum_{\nu=0}^{d} \frac{m_i p_\nu(i)}{\kappa_\nu} m_j p_\nu(j).$$

Taking the trace of

$$|X|(E_i \circ E_j)E_h = \sum_{l=0}^{d} q_{ij}^l E_l E_h = q_{ij}^h E_h, \tag{69}$$

and using (66) and (67), we have

$$q_{ij}^h = \frac{|X|}{m_h}\mathrm{tr}((E_i \circ E_j)E_h) = \frac{|X|}{m_h}\sum_{\nu=0}^{d}\sum_{\mu=0}^{d}\frac{1}{|X|^2}q_i(\nu)q_j(\mu)\,\mathrm{tr}((A_\nu \circ A_\mu)E_h)$$

$$= \frac{1}{m_h|X|}\sum_{\nu=0}^{d} q_i(\nu)q_j(\nu)m_h p_\nu(h) = \frac{1}{|X|}\sum_{\nu=0}^{d}\frac{m_i p_\nu(i)}{\kappa_\nu}\frac{m_j p_\nu(j)}{\kappa_\nu}p_\nu(h)$$

namely (64). Lastly, (65) follows from (69) since

$$\sum_{j=0}^{d} q_{ij}^h E_h = \sum_{j=0}^{d} |X|(E_i \circ E_j)E_h = |X|(E_i \circ I)E_h = m_i E_h. \qquad \square$$

References

[1] D. Aldous and P. Diaconis, Shuffling cards and stopping times, Amer. Math. Monthly **93** (1986), 333–348.

[2] E. Bannai and T. Ito, Algebraic combinatorics I: Association schemes, Benjamin/Cummings, Menlo Park, California, 1984.

[3] D. Bayer and P. Diaconis, Trailing the dovetail shuffle to its lair, Ann. Appl. Probab. **2** (1992), 294–313.

[4] E. D. Belsley, Rates of convergence of random walk on distance regular graphs, Probab. Th. Rel. Fields **112** (1998), 493–533.

[5] A. E. Brouwer, A. M. Cohen and A. Neumaier, Distance-regular graphs, Springer, Berlin Heidelberg, 1989.

[6] A. J. D'Aristotile, The nearest neighbor random walks on subspaces of a vector space and rate of convergence, J. Theoret. Probab. **8** (1995), 321–346.

[7] P. Diaconis, Group representations in probability and statistics, Inst. Math. Stat., Hayward, California, 1988.

[8] P. Diaconis, The cutoff phenomenon in finite Markov chains, Proc. Nat. Acad. Sci. USA **93**, no. 4 (1996), 1659–1664.

[9] P. Diaconis, R. L. Graham and J. A. Morrison, Asymptotic analysis of a random walk on a hypercube with many dimensions, Random Struct. Algor. **1** (1990), 51–72.

[10] P. Diaconis and P. Hanlon, Eigen analysis for some examples of the Metropolis algorithm, Contemporary Math. **138** (1992), 99–117.

[11] P. Diaconis and L. Saloff-Coste, Comparison theorems for reversible Markov chains, Ann. Appl. Probab. **3** (1993), 696–730.

[12] P. Diaconis and L. Saloff-Coste, Comparison techniques for random walks on finite groups, Ann. Probab. **21** (1993), 2131–2156.

[13] P. Diaconis and M. Shahshahani, Generating a random permutation with random transpositions, Z. Wahr. verw. Geb. **57** (1981), 159–179.

[14] P. Diaconis and M. Shahshahani, Time to reach stationarity in the Bernoulli-Laplace diffusion model, SIAM J. Math. Anal. **18** (1987), 208–218.

[15] P. Donnelly, P. Lloyd and A. Sudbury, Approach to stationarity of the Bernoulli-Laplace diffusion model, Adv. Appl. Probab. **26** (1994), 715–727.

[16] M. Hildebrand, Generating random elements in $SL_n(\mathbf{F}_q)$ by random transvections, J. Alg. Combin. **1** (1992), 133–150.

[17] A. Hora, Remarks on the shuffling problem for finite groups, Publ. RIMS Kyoto Univ. **29** (1993), 153–159.

[18] A. Hora, Towards critical phenomena for random walks on various algebraic structures, In H. Heyer and T. Hirai, (eds.), Trans. German-Japanese Symposium 1995 in Tübingen, D.+M. Gräbner, Bamberg, 1996, pp. 113–127.

[19] A. Hora, The cut-off phenomenon in random walks on association schemes, RIMS Kôkyûroku Kyoto Univ. **962** (1996), 32–41.

[20] A. Hora, The cut-off phenomenon for random walks on Hamming graphs with variable growth conditions, Publ. RIMS Kyoto Univ. **33** (1997), 695–710.

[21] U. Porod, The cut-off phenomenon for random reflections, Ann. Probab. **24** (1996), 74–96.

[22] U. Porod, The cut-off phenomenon for random reflections II: complex and quaternionic cases, Probab. Th. Rel. Fields **104** (1996), 181–210.

[23] J. S. Rosenthal, Random rotations: characters and random walks on $SO(N)$, Ann. Probab. **22** (1994), 398–423.

[24] L. Saloff-Coste, Precise estimates on the rate at which certain diffusions tend to equilibrium, Math. Z. **217** (1994), 641–677.

[25] L. Saloff-Coste, Lectures on finite Markov chains, In P. Bernard, (ed.), Lectures on probability theory and statistics, Lect. Notes Math. **1665**, Springer, Berlin Heidelberg, 1997, pp. 301–413.

[26] M. Voit, Limit theorems for compact two-point homogeneous spaces of large dimensions, J. Theoret. Probab. **9** (1996), 353–370.

[27] M. Voit, Asymptotic distributions for the Ehrenfest urn and related random walks, J. Appl. Probab. **33**, no. 3 (1996), 340–356.

[28] M. Voit, Asymptotic behavior of heat kernels on spheres of large dimensions, J. Multivariate Anal. **59** (1996), 230–248.

Department of Environmental and Mathematical Sciences
Faculty of Environmental Science and Technology
Okayama University
Okayama 700-8530, Japan
hora@math.ems.okayama-u.ac.jp