Maximum entropy mixing time of circulant Markov processes

Konstantin Avrachenkov[a], Laura Cottatellucci[b], Lorenzo Maggi[c], Yong-Hua Mao[d]

[a] *INRIA, BP95, 06902 Sophia Antipolis Cedex, France, e-mail: k.avrachenkov@sophia.inria.fr*
[b] *EURECOM, Department of Mobile Communications, 06410 Biot, France, e-mail: laura.cottatellucci@eurecom.fr*
[c] *Saarland University, Chair of Economic Theory, 66123 Saarbrücken, Germany, e-mail: lorenzo.maggi@uni-saarland.de*
[d] *Beijing Normal University, School of Mathematical Sciences & Laboratory of Mathematics and Complex Systems, Ministry of Education, 100875 Beijing, China, e-mail: maoyh@bnu.edu.cn*

Abstract

We consider both discrete-time irreducible Markov chains with circulant transition probability matrix $\mathbf{P}$ and continuous-time irreducible Markov processes with circulant transition rate matrix $\mathbf{Q}$. In both cases we provide an expression of all the moments of the mixing time. In the discrete case, we prove that all the moments of the mixing time associated with the transition probability matrix $\alpha\mathbf{P} + [1-\alpha]\mathbf{P}^*$ are maximum in the interval $0 \leq \alpha \leq 1$ when $\alpha = 1/2$, where $\mathbf{P}^*$ is the transition probability matrix of the time-reversed chain. Similarly, in the continuous case, we show that all the moments of the mixing time associated with the transition rate matrix $\alpha\mathbf{Q} + [1-\alpha]\mathbf{Q}^*$ are also maximum in the interval $0 \leq \alpha \leq 1$ when $\alpha = 1/2$, where $\mathbf{Q}^*$ is the time-reversed transition rate matrix.

Keywords: circulant Markov process, maximum mixing time, moments mixing time

1. Introduction

In this paper we consider both discrete irreducible Markov chains with circulant transition probability matrix $\mathbf{P}$ and continuous-time irreducible Markov processes with circulant transition rate matrix $\mathbf{Q}$. Therefore, the adjacency matrix[1] of their associated graph is circulant as well. Hence, our model can be viewed either as a general case of random walk[2] on unweighted circulant graphs or as a particular case of random walk on weighted circulant graphs, in which the matrix of weights is circulant as well. Examples of circulant graphs are complete graph, crown graph $2n-1$, Paley graph of prime order, Möbius ladder, cocktail party graph, Andrásfai graph, antiprism graph, complete bipartite graph, cycle graph, octahedral graph, pentatope graph, prism graphs, square graph, tetrahedral graph, triangle graph, and utility graph (Weisstein). Moreover, connected circulant graphs are Cayley graphs (Pegg et al.).

Circulant graphs have been extensively studied in literature over the years from both a combinatorics and an algebraic point of view (Morris, 2007; Elspas and Turner, 1970; Codenotti et al., 1998; So, 2006). They play a fundamental role in telecommunication networks, VLSI design, and distributed computation (Bermond et al., 1995; Leighton, 1992; Cai et al., 1999; Mans,

[1]The adjacency matrix of a continuous Markov process is associated with its subordinated chain (Brémaud, 1999) with transition probability matrix $\mathbf{K} = \mathbf{Q}/\lambda + \mathbf{I}$, for some $\lambda > 0$.

[2]In the continuous case, the number of steps of the random walk over time is distributed as a Poisson variable with parameter λ.

1997), since they are a natural extension of rings, with increased connectivity.

Let us give an overview of the content of this paper. In Section 2 we recapitulate some useful definitions and results on Markov processes. We provide an expression for all the moments of the mixing time of both discrete and continuous circulant Markov processes in Sections 3.1 and 3.2, respectively. In Section 4 we prove a conjecture by Aldous and Fill (2002), restricted to the case of circulant discrete Markov chains, claiming that the mixing time of the matrix $\alpha\mathbf{P} + [1-\alpha]\mathbf{P}^*$, where $\mathbf{P}^*$ is the time-reversed of $\mathbf{P}$, is maximum in the interval $0 \leq \alpha \leq 1$ at $\alpha = 1/2$. We refer to such mixing time as maximum entropy mixing time. We extend the validity of Aldous and Fill's conjecture to all the moments of the maximum entropy mixing time, and we provide analogous results in the continuous case. We dub these results as maximum entropy theorems of mixing time.

Let us provide some remarks about notation. The imaginary unit is referred to as j. Let $c \in \mathbb{C}$. Its complex conjugate is $\overline{c}$; $\Re(c)$ and $\Im(c)$ indicate its real and its imaginary part, respectively. The convolution between two functions a and b is denoted by $a \star b$. $\mathbb{N}_0$ is equivalent to $\mathbb{N} \cup \{0\}$. $\mathbb{R}_0^+$ is the set of nonnegative real numbers. We use bold fonts for matrices and vectors. $\mathbf{B}_{n,m}$ is the (n, m) component of matrix $\mathbf{B}$. The acronym ROC stands for Region Of Convergence.

2. Background on Markov processes

2.1. Discrete-time Markov chains

Let $\mathbf{P}$ be the stochastic transition probability matrix of a discrete-time homogeneous Markov chain (DT-HMC) on the finite state space $\mathcal{S}$, having cardinality N. The matrix $\mathbf{P}$ is irreducible when each state is reachable from any state with positive probability. The stationary distribution of $\mathbf{P}$ is the column vector $\boldsymbol{\pi} \in \mathbb{R}^N : \boldsymbol{\pi}^T\mathbf{P} = \boldsymbol{\pi}^T$, $\sum_{i=1}^N \boldsymbol{\pi}_i = 1$. If $\mathbf{P}$ is irreducible, then $\boldsymbol{\pi}$ is unique and $\boldsymbol{\pi}_i > 0$ for all i (Brémaud, 1999). Hence, in such a case we can define the transition probability matrix $\mathbf{P}^*$ of the time-reversed chain, such that $\mathbf{P}^*_{n,m} = \mathbf{P}_{m,n}\boldsymbol{\pi}_m/\boldsymbol{\pi}_n$, for all n, m.

Theorem 2.1 (Brémaud (1999))**.** *Let $\mathbf{P}$ be an irreducible transition probability matrix matrix. Let $\{\lambda_i\}_i$ be its eigenvalues. Then, 1 is a simple eigenvalue of $\mathbf{P}$ (i.e., it has multiplicity 1) and $|\lambda_i| \leq 1$, for all i's.*

The k-th moment of the hitting time from state s_n to state s_m, where $k \in \mathbb{N}$, is defined as

$$\mathrm{E}_n\left(T_m^k\right) \equiv \mathrm{E}\left(\inf\left\{i^k \geq 0, i \in \mathbb{N}_0 : S_i = s_m\right\} \middle| S_0 = s_n\right),$$

where S_i is the state of the Markov chain at time step i. We denote by $\Psi^{(k)}(\mathbf{P})$ the k-th moment of the mixing time associated with $\mathbf{P}$, i.e.

$$\Psi^{(k)}(\mathbf{P}) = \sum_{n=1}^{N}\sum_{m=1}^{N} \boldsymbol{\pi}_n\boldsymbol{\pi}_m \,\mathrm{E}_n\left(T_m^k\right). \tag{1}$$

Let the matrix $\mathbf{P}$ be circulant, i.e. there exist $\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{N-1}$ such that $\mathbf{P}_{n,m} = \mathbf{c}_{m-n \bmod (N)}$. Hence, its eigenvalues $\{\lambda_i\}_{0\leq i\leq N-1}$ of $\mathbf{P}$ can be computed as (Gray, 2006)

$$\lambda_i = \sum_{n=0}^{N-1} \mathbf{c}_n \exp(-j2\pi ni/N), \qquad 0 \le i \le N-1. \tag{2}$$

If $\mathbf{P}$ is also irreducible, then the column vector $\boldsymbol{\pi} = \mathbf{1}_N/N$ is its unique stationary distribution.

2.2. Continuous-time Markov processes

In the continuous case, we still consider the state space S to have cardinality N. Let $\{S(t)\}_{t\in\mathbb{R}}$ be a continuous-time homogeneous Markov process (CT-HMP) on S. Let $\widetilde{\mathbf{P}}(t)$ be, for any $t \in \mathbb{R}_0^+$, the transition probability matrix such that prob $\left(S(t_2) = s_{i_2} \,\middle|\, S(t_1) = s_{i_1}\right) = \widetilde{\mathbf{P}}_{i_1,i_2}(t_2 - t_1)$, for all $t_2 \ge t_1$. The transition rate matrix $\mathbf{Q}$ is defined as the component-wise limit of $[\widetilde{\mathbf{P}}(h) - \widetilde{\mathbf{P}}(0)]/h$ when $h \downarrow 0$. The following result is well known.

Theorem 2.2. *Let* $\mathbf{Q}$ *be the transition rate matrix of an irreducible CT-HMP. Then, the null eigenvalue has multiplicity 1.*

Note that, if we suppose $\mathbf{Q}$ to be circulant, then $\widetilde{\mathbf{P}}(t) = \exp(\mathbf{Q}t) = \sum_{i=0}^{\infty}[\mathbf{Q}t]^i/i!$ is circulant as well, for all $t \ge 0$. The k-th moment of the hitting time from state s_n to s_m is here defined as

$$\mathrm{E}_n\left(\widetilde{T}_m^k\right) \equiv \mathrm{E}\left(\inf\left\{t^k,\ t \in \mathbb{R}_0^+ : S(t) = s_m\right\} \,\middle|\, S_0 = s_n\right).$$

The expression of the k-th moment of the mixing time is analogous to the one in the discrete case (1), and we call it $\widetilde{\Psi}^{(k)}(\mathbf{Q})$. The time-reversed CT-HMP is characterized by the transition rate matrix $\mathbf{Q}^*$, such that $\mathbf{Q}^*_{n,m} = \mathbf{Q}_{m,n}\,\widetilde{\boldsymbol{\pi}}_m/\widetilde{\boldsymbol{\pi}}_n$, where $\widetilde{\boldsymbol{\pi}}$ is the stationary distribution of $\mathbf{Q}$.

We assume that the reader is familiar with the fundamentals of zeta and Laplace transform. A suitable reference is (Brown and Churchill, 1996).

3. Moments of mixing time

3.1. Discrete-time Markov chains

In this section we consider an irreducible DT-HMC with circulant transition probability matrix and we will provide a formula for all the moments of its mixing time.

Theorem 3.1. *Let* $\mathbf{P}$ *be the circulant transition probability matrix of an irreducible DT-HMC. Let* $\lambda_1, \ldots, \lambda_{N-1}$ *be the eigenvalues of* $\mathbf{P}$ *which are different from 1. Then,*

$$\Psi^{(k)}(\mathbf{P}) = (-1)^k \frac{d^k}{dz^k} \left[1 + [z-1] \sum_{i=1}^{N-1} [z - \lambda_i]^{-1} \right]^{-1} \Bigg|_{z=1}, \quad z \in \mathbb{C}. \qquad (3)$$

Proof. Let the discrete-argument function $f^{(n,m)}$ be the probability mass function associated with the hitting time T_m when the initial state is s_n, i.e.

$$f^{(n,m)}(i) \equiv \text{prob}\left(T_m = i \,\middle|\, S_0 = s_n\right), \qquad \forall\, i \in \mathbb{N}_0,$$

and $f^{(n,m)}(i) \equiv 0$ for $i < 0$. Since $\mathbf{P}$ is circulant, then for all $i \in \mathbb{Z}$, $k \in \mathbb{N}$, $1 \le n, m \le N$, $E_n\left[T_m^k\right] = E_{n+i-1 \bmod (N)+1}\left[T_{m+i-1 \bmod (N)+1}^k\right]$. Thus, the expression $\sum_{m=1}^N \mathrm{E}_n(T_m^k)$, for all $k \in \mathbb{N}$, does not depend on the initial state n. Hence, it is straightforward to see that the k-th moment of the mixing time $\Psi^{(k)}(\mathbf{P})$ can be expressed, for any $1 \le n \le N$, as

$$\Psi^{(k)}(\mathbf{P}) = \frac{1}{N} \sum_{m=1}^{N} \sum_{i \in \mathbb{N}_0} i^k\, f^{(n,m)}(i), \quad \forall\, k \in \mathbb{N}.$$

Thanks to a well known property of zeta transform (see e.g. Brown and Churchill (1996)) we can write

$$\Psi^{(k)}(\mathbf{P}) = \frac{[-1]^k}{N} \frac{d^k}{dz^k} \sum_{m=1}^{N} F^{(n,m)}(z) \Big|_{z=1} \tag{4}$$

where $F^{(n,m)}$ is the zeta transforms of $f^{(n,m)}$. Let us define $p^{(n,m)}$ as the discrete-time function associated with the probability of transition from state s_n to state s_m, i.e. $p^{(n,m)}(i) \equiv (\mathbf{P}^i)_{n,m}$, for all $i \in \mathbb{N}_0$, and $p^{(n,m)}(i) \equiv 0$, for all $i < 0$. The following recursive property holds for all $1 \leq n, m \leq N$, $i \in \mathbb{Z}$ (see Chung (1967), Theorem 2, p. 21)

$$f^{(n,m)}(i) = p^{(n,m)}(i) - \sum_{\tau=0}^{i-1} p^{(m,m)}(i-\tau)\, f^{(n,m)}(\tau),$$

which can be rewritten as $p^{(n,m)} = f^{(n,m)} \star p^{(m,m)}$. Let $P^{(n,m)}$ be the zeta transforms of $p^{(n,m)}$. Then, thanks to the convolution theorem of zeta transform (see e.g. Brown and Churchill (1996)),

$$F^{(n,m)}(z) = \frac{P^{(n,m)}(z)}{P^{(m,m)}(z)}, \quad \forall\, n, m,\ z \in \mathrm{ROC}(F^{(n,m)}) \cap \mathrm{ROC}(P^{(m,m)}). \tag{5}$$

Let u be the discrete step function, i.e. $u(i) = 1$ for $i \in \mathbb{N}_0$ and $u(i) = 0$ for $i < 0$. Let $\mathcal{U}(z) = [1 - z^{-1}]^{-1}$ be its zeta transform. Then, thanks to the linearity of zeta transform, we can write

$$\sum_{m=1}^{N} P^{(n,m)}(z) = \mathcal{U}(z), \qquad |z| > 1.$$

Since any power of a circulant matrix is still circulant, then we can define the discrete-time function $\omega \equiv p^{(m,m)}$, for all $m \in [1; N]$. Let Ω be the zeta transform of ω. Let $\lambda_0^{(i)}, \ldots, \lambda_{N-1}^{(i)}$ be the eigenvalues of $\mathbf{P}^i$, for $i \geq 2$. Then, by elementary properties of eigenvalues,

$$\boldsymbol{\omega}(i) = \frac{1}{N}\,\mathrm{trace}(\mathbf{P}^i) = \frac{1}{N}\sum_{n=0}^{N-1}\lambda_n^{(i)} = \frac{1}{N}\sum_{n=0}^{N-1}\lambda_n^i, \quad \forall\, i \in \mathbb{N}_0$$

and $\boldsymbol{\omega}(i) = 0$ for all $i < 0$. By Theorem 2.1, $\mathbf{P}$ has only one eigenvalue equal to 1, i.e. $\lambda_0 = 1 > |\lambda_i|$, for all i. Then,

$$\Omega(z) = \frac{1}{N}\,\mathcal{U}(z) + \frac{1}{N}\sum_{i=1}^{N-1}\frac{1}{1-\lambda_i\, z^{-1}}, \qquad |z| > 1.$$

Hence, we can say that, for any $n \in [1; N]$,

$$\sum_{m=1}^{N} F^{(n,m)}(z) = \frac{\mathcal{U}(z)}{\Omega(z)}, \qquad |z| > 1. \tag{6}$$

Since $F^{(n,m)}$ is the zeta transform of a probability distribution, then the point $z = 1$ belongs to the interior of its region of convergence, hence we are allowed to cancel out the discontinuity of (6) in $z = 1$. Therefore, $\sum_{m=1}^{N} F^{(n,m)}(z)$ is analytic in $z = 1$, and we can rewrite (4) as

$$\Psi^{(k)}(\mathbf{P}) = \lim_{z\to 1}\frac{[-1]^k}{N}\,\frac{d^k}{dz^k}\,\frac{\mathcal{U}(z)}{\Omega(z)}. \tag{7}$$

It is straightforward to see that (7) coincides with (3), *q.e.d.*. □

For $k = 1$, equation (3) reduces to the classic expression for mixing time (see e.g. Aldous and Fill (2002); Hunter (2006)), i.e. $\Psi^{(1)}(\mathbf{P}) = \sum_{i=1}^{N-1}[1-\lambda_i]^{-1}$. For $k = 2$, we obtain the second moment of the mixing time:

$$\Psi^{(2)}(\mathbf{P}) = 2\,\left[\Psi^{(1)}(\mathbf{P})\right]^2 + \Psi^{(1)}(\mathbf{P}) + 2\sum_{i=1}^{N-1}\lambda_i\,[1-\lambda_i]^{-2}\,. \tag{8}$$

3.2. Continuous-time Markov processes

In this section we will provide an expression for the moments of the mixing time of an irreducible CT-HMP with circulant transition rate matrix. We will omit the proof, which is similar to the one of Theorem 3.1.

Theorem 3.2. *Let* $\mathbf{Q}$ *be the circulant transition rate matrix of an irreducible CT-HMP. Let* $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_{N-1}$ *be the nonnull eigenvalues of* $\mathbf{Q}$*. Then,*

$$\widetilde{\Psi}^{(k)}(\mathbf{Q}) = [-1]^k \frac{d^k}{ds^k} \left[1 + s \sum_{i=1}^{N-1} [s - \widetilde{\lambda}_i]^{-1} \right]^{-1} \Bigg|_{s=0}, \quad s \in \mathbb{C}. \tag{9}$$

Substituting $k = 1$ into (9), we obtain the classic mixing time expression for continuous-time HMC, i.e. $\widetilde{\Psi}^{(1)}(\mathbf{Q}) = -\sum_{i=1}^{N-1} \widetilde{\lambda}_i^{-1}$. For $k = 2$, we obtain the second moment of the mixing time:

$$\widetilde{\Psi}^{(2)}(\mathbf{Q}) = 2 \left[\widetilde{\Psi}^{(1)}(\mathbf{Q}) \right]^2 + 2 \sum_{i=1}^{N-1} \widetilde{\lambda}_i^{-2}.$$

4. Maximum entropy mixing time theorems

In this section we prove Conjecture 24 in Chapter 9 of (Aldous and Fill, 2002), in the case of a DT-HMC with circulant transition probability matrix. We also extend its validity to all the moments of the mixing time. Moreover, we provide analogous results for CT-HMP with circulant transition rate matrix.

Theorem 4.1. *Let* $\mathbf{P}$ *be the circulant transition probability matrix of an irreducible DT-HMC. Let* $\mathbf{P}^*$ *be the transition probability matrix of the time-reversed chain. Then, all the moments of the mixing time* $\Psi^{(k)}(\mathbf{P}^{(\alpha)})$*,* $k \in \mathbb{N}$*, associated with the transition probability matrix* $\mathbf{P}^{(\alpha)} = \alpha\mathbf{P} + [1-\alpha]\mathbf{P}^*$ *attain their maximum in the interval* $0 \leq \alpha \leq 1$ *at* $\alpha = 1/2$*.*

Proof. Let $\{\lambda_i\}_i$ be the eigenvalues of $\mathbf{P}$. Let us define $\mathbf{P}^{(\alpha)} = \alpha\mathbf{P} + [1-\alpha]\mathbf{P}^*$. Hence $\mathbf{P}^{(\alpha)}$ is circulant, too. Let $\{\phi_i^{(\alpha)}\}_i$ be the eigenvalues of $\mathbf{P}^{(\alpha)}$. Firstly we observe that

$$\frac{d^k}{dz^k}\frac{z-1}{z-\phi_i^{(\alpha)}} = \frac{[-1]^{k-1}\,k!\,\left[1-\phi_i^{(\alpha)}\right]}{\left[z-\phi_i^{(\alpha)}\right]^{k+1}}. \tag{10}$$

Next we claim that $\Psi^{(k)}(\mathbf{P}^{(\alpha)})$ can be written as a finite sum of terms of this form:

$$C_1\left[1+\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{-n_0}\prod_{l=1}^{L}\left[\frac{d^{k_l}}{dz^{k_l}}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{n_l}\Bigg|_{z=1} \tag{11}$$

$$= C_2\prod_{l=1}^{L}\left[\sum_{i=1}^{N-1}\left[1-\phi_i^{(\alpha)}\right]^{-k_l}\right]^{n_l}, \tag{12}$$

where $L, n_l \in \mathbb{N}$, such that

$$C_2 = C_1\prod_{l=1}^{L} k_l![-1]^{[k_l-1]n_l} \geq 0. \tag{13}$$

This can be proved inductively. For $k=1$ this is true, since

$$-\frac{d}{dz}\left[1+\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{-1}\Bigg|_{z=1} = \left[1+\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{-2}\frac{d}{dz}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\Bigg|_{z=1}.$$

Now suppose the claim is true for $k \geq 1$. Then, if we compute the derivative with respect to z of the expression in z in (11) and we multiply it by -1, then we obtain

$$C_1 n_0\left[1+\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{-n_0-1}\left[\frac{d}{dz}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]\prod_{l=1}^{L}\left[\frac{d^{k_l}}{dz^{k_l}}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{n_l} +$$
$$- C_1\left[1+\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{-n_0}\sum_{h=1}^{L} n_h\left[\frac{d^{k_h+1}}{dz^{k_h+1}}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]\left[\frac{d^{k_h}}{dz^{k_h}}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{n_h-1}\times$$
$$\times\prod_{l=1,l\neq h}^{L}\left[\frac{d^{k_l}}{dz^{k_l}}\sum_{i=1}^{N-1}\frac{z-1}{z-\phi_i^{(\alpha)}}\right]^{n_l} \tag{14}$$

It is easy to check that the coefficients of each addend in (14) still satisfy the rule in (13). Now we prove that the expression in (12) attains its maximum at $\alpha = 1/2$. First, we note that the eigenvalues of $\mathbf{P}$ make conjugate pairs and they are equal to the eigenvalues of $\mathbf{P}^*$. Then, exploiting formula (2) we derive $\phi_i^{(\alpha)} = \Re(\lambda_i) + j[2\alpha - 1]\Im(\lambda_i)$, for all i's. Next we claim that, since $\{\phi_i^{(\alpha)}\}_i$ make conjugate pairs too, then we can write (12) as

$$C_2 \prod_{l=1}^{L} \left[\sum_{i=1}^{N-1} \left[1 - \phi_i^{(\alpha)}\right]^{-k_l} \right]^{n_l} = C_2 \prod_{l=1}^{L} \left[\sum_{i=1}^{N-1} \Re\left(\left[1 - \phi_i^{(\alpha)}\right]^{-k_l} \right) \right]^{n_l}. \qquad (15)$$

For all $1 \le i \le N-1$, $1 \le l \le L$,

$$\begin{aligned} \Re\left(\left[1 - \phi_i^{(1/2)}\right]^{-k_l} \right) &= \left| \left[1 - \phi_i^{(1/2)}\right]^{-k_l} \right| \\ &\ge \left[[1 - \Re(\phi_i^{(\alpha)})]^2 + [1 - 2\alpha]^2 [\Im(\phi_i^{(\alpha)})]^2 \right]^{-k_l/2}, \quad \forall \alpha \in [0;1] \\ &\ge \left| \Re\left(\left[1 - \phi_i^{(\alpha)}\right]^{-k_l} \right) \right|, \quad \forall \alpha \in [0;1]. \end{aligned}$$

Therefore, for all $l = 1, \dots, L$,

$$\left[\sum_{i=1}^{N-1} \Re\left(\left[1 - \phi_i^{(1/2)}\right]^{-k_l} \right) \right]^{n_l} \ge \left| \left[\sum_{i=1}^{N-1} \Re\left(\left[1 - \phi_i^{(\alpha)}\right]^{-k_l} \right) \right]^{n_l} \right|, \quad \forall\, \alpha \in [0;1].$$

Then, we can conclude that the expression in (15) attains its maximum at $\alpha = 1/2$ and the thesis is proved. □

The following Corollary follows straightforward from the proof of Theorem 4.1.

Corollary 4.2. *If* $\mathbf{P}$ *has at least one pair of non-real eigenvalues then* $\alpha = 1/2$ *is the unique point of maximum of* $\Psi^{(k)}(\mathbf{P}^{(\alpha)})$*, for all* $k \in \mathbb{N}$.

The name of Theorem 4.1 follows from the fact that the mixing time $\Psi^{(k)}(\mathbf{P}^{(\alpha)})$ is maximum when the entropy of the distribution $(\alpha, 1-\alpha)$ is maximum. Next we prove the analogue of Theorem 4.1 for continuous-time HMCs.

Theorem 4.3. *Let* $\mathbf{Q}$ *be the circulant transition rate matrix of an irreducible CT-HMP. Let* $\mathbf{Q}^*$ *be the transition rate matrix of the time-reversed process. Then, all the moments of the mixing time* $\widetilde{\Psi}^{(k)}(\mathbf{Q}^{(\alpha)})$, $k \in \mathbb{N}$, *associated with the transition rate matrix* $\mathbf{Q}^{(\alpha)} = \alpha\mathbf{Q} + [1-\alpha]\mathbf{Q}^*$ *attain their maximum in the interval* $0 \leq \alpha \leq 1$ *at* $\alpha = 1/2$.

Proof. The proof follows the same lines as the one of Theorem 4.1. Let $\{\widetilde{\lambda}_i\}_i$ be the eigenvalues of $\mathbf{Q}$. Let us define $\mathbf{Q}^{(\alpha)} = \alpha\mathbf{Q} + [1-\alpha]\mathbf{Q}^*$. Let $\widetilde{\phi}_i^{(\alpha)} = \Re(\widetilde{\lambda}_i) + j[2\alpha - 1]\Im(\widetilde{\lambda}_i)$ be the i-th eigenvalue of $\mathbf{Q}^{(\alpha)}$. Firstly we observe that

$$\frac{d^k}{ds^k}\frac{s}{s-\widetilde{\phi}_i^{(\alpha)}} = -\frac{\widetilde{\phi}_i^{(\alpha)}\, k!}{[\widetilde{\phi}_i^{(\alpha)} - s]^{k+1}} \tag{16}$$

By induction on k, similarly as in Theorem 4.1, we can show that $\widetilde{\Psi}^{(k)}(\mathbf{Q}^{(\alpha)})$ can be written as a finite sum of terms of the form:

$$\begin{aligned} &C\left[1+\sum_{i=1}^{N-1}\frac{s}{s-\widetilde{\phi}_i^{(\alpha)}}\right]^{-n_0}\prod_{l=1}^{L}\left[[-1]^{k_l-1}\frac{d^{k_l}}{dz^{k_l}}\sum_{i=1}^{N-1}\frac{s}{s-\widetilde{\phi}_i^{(\alpha)}}\right]^{n_l}\Bigg|_{s=0} \\ &= C\prod_{l=1}^{L}k_l!\prod_{l=1}^{L}\left[\sum_{i=1}^{N-1}\left[-\widetilde{\phi}_i^{(\alpha)}\right]^{-k_l}\right]^{n_l} \end{aligned} \tag{17}$$

with $L, n_l \in \mathbb{N}$, for $l = 0, \ldots, L$, and $C \geq 0$. Next we claim that, since $\{\phi_i^{(\alpha)}\}_i$ make conjugate pairs, then $\sum_{i=1}^{N-1}[-\widetilde{\phi}_i^{(\alpha)}]^{-k_l} = \sum_{i=1}^{N-1}\Re\big([-\widetilde{\phi}_i^{(\alpha)}]^{-k_l}\big)$. Recall that $\Re(\widetilde{\lambda}_i) < 0$ for all $i = 1, \ldots, N-1$. Then,

$$\Re\left(\left[-\widetilde{\phi}_i^{(1/2)}\right]^{-k_l}\right) \geq \left[\left[\Re(\widetilde{\lambda}_i)\right]^2 + [2\alpha - 1]^2 \left[\Im(\widetilde{\lambda}_i)\right]^2\right]^{-k_l/2}$$
$$\geq \left|\Re\left(\left[-\widetilde{\phi}_i^{(\alpha)}\right]^{-k_l}\right)\right|, \quad 1 \leq i \leq N-1.$$

Thus, for all $1 \leq l \leq L$,

$$\left[\sum_{i=1}^{N-1}\left[-\widetilde{\phi}_i^{(1/2)}\right]^{-k_l}\right]^{n_l} \geq \left|\left[\sum_{i=1}^{N-1}\left[-\widetilde{\phi}_i^{(\alpha)}\right]^{-k_l}\right]^{n_l}\right|, \quad \forall\, \alpha \in [0;1].$$

Therefore, the expression (17) attains its maximum at $\alpha = 1/2$ and the thesis is proved. □

Corollary 4.4. *If* $\mathbf{Q}$ *has at least one pair of non-real eigenvalues then* $\alpha = 1/2$ *is the unique point of maximum of* $\widetilde{\Psi}^{(k)}(\mathbf{Q}^{(\alpha)})$*, for all* $k \in \mathbb{N}$*.*

Acknowledgement

This research was supported by “Agence Nationale de la Recherche”, with reference ANR-09-VERS-001. The fourth author (Y.H.M) is supported by 973 Project (No 2011CB808000), NSFC (No 11131003). We also would like to thank Nelly Litvak for motivating discussions.

References

Aldous, D., Fill, J., 2002. Reversible Markov Chains and Random Walks on Graphs. A draft of the book is available at ``http://www.stat.berkeley.edu/~aldous/RWG/book.html''.

Bermond, J.C., Comellas, F., Hsu, D.F., 1995. Distributed loop computer networks: a survey. Journal of Parallel and Distributed Computing 24, 2–10.

Brémaud, P., 1999. Markov chains: Gibbs fields, Monte Carlo simulation, and queues. Springer.

Brown, J.W., Churchill, R.V., 1996. Complex variables and applications, sixth edition. McGraw Hill International Editions.

Cai, J.Y., Havas, G., Mans, B., Nerurkar, A., Seifert, J.P., Shparlinski, I., 1999. On routing in circulant graphs, in: Proceedings of the 5th annual international conference on Computing and combinatorics, Springer-Verlag. pp. 360–369.

Chung, K.L., 1967. Markov Chains with Stationary Transition Probabilities. volume 104 of *Grundlehren der mathematischen Wissenschaften*. Springer. 2nd edition.

Codenotti, B., Gerace, I., Vigna, S., 1998. Hardness results and spectral techniques for combinatorial problems on circulant graphs. Linear Algebra and its Applications 285, 123–142.

Elspas, B., Turner, J., 1970. Graphs with circulant adjacency matrices. Journal of Combinatorial Theory 9, 297–307.

Gray, R.M., 2006. Toeplitz and circulant matrices: A review. Foundations and Trends in Communications and Information Theory 2, 155–239.

Hunter, J.J., 2006. Mixing times with applications to perturbed Markov chains. Linear Algebra and its Applications 417, 108–123.

Leighton, F.T., 1992. Introduction to parallel algorithms and architectures. Kauffman.

Mans, B., 1997. Optimal distributed algorithms in unlabeled tori and chordal rings. Journal of Parallel and Distributed Computing 46, 80–90.

Morris, J., 2007. Automorphism groups of circulant graphs–A survey. Graph Theory in Paris , 311–325.

Pegg, E.J., Rowland, T., Weisstein, E.W.. Cayley graph. `http://mathworld.wolfram.com/CayleyGraph.html`. From MathWorld–A Wolfram Web Resource.

So, W., 2006. Integral circulant graphs. Discrete Mathematics 306, 153–158.

Weisstein, E.W.. Circulant graph. `http://mathworld.wolfram.com/CirculantGraph.html`. From MathWorld–A Wolfram Web Resource.