Spectral gap of Markov chains on a cycle

Balázs Gerencsér[1] and Julien Hendrickx[1] and Paul Van Dooren[1]

***Abstract*— We search for the Markov chain with the optimal mixing rate where transitions are restricted to happen along a cycle of the states. We show that homogeneous, reversible chains are locally optimal for perturbations that make them inhomogeneous and non-reversible. Moreover, we show the optimality holds globally if only a single type of perturbation (either inhomogeneous or non-reversible) is applied. However, we conjecture global optimality holds for mixed perturbations as well, which is backed by simulation results. This paper complements previous results on bounds for mixing times of general Markov chains on the cycle [1].**

I. Introduction

Markov chains appear as the essential building block for several applications, mostly targeting a randomized approach. Complicated simulations often benefit from the Markov chain Monte Carlo framework, see Metropolis et al. [2], Hastings [3] and Jerrum [4]. Distributed systems use average consensus for collaborative computation, Olfati-Saber et al. [5] or Blondel et al. [6]. Similarly for networked control systems, like flocking in Reynolds [7], these are all also often based on Markov chain techniques. The performance of such applications heavily rely on utilizing a proper Markov chain with good mixing properties, see Olshevsky, Tsitsiklis [8] and Boyd et al. [9] for details. In this paper we study the challenge of optimizing this performance for the instructional case when the connection graph forms a cycle.

Although sometimes we can only analyze a given Markov chain, often we have the freedom to choose it according to our taste, given some constraints on the allowed transitions. It is a very natural approach to search for the one with the best performance. However, as we will see, finding the optimal chain is quite hard even in the simplest setting. A Markov chain is efficient if it quickly approaches its stationary distribution regardless of the initial condition. We want to capture this speed via the asymptotic rate given as

$$\tilde{\gamma} = -\max_x \limsup_{k\to\infty} \frac{1}{k} \log \|xP^k - \pi\|, \qquad (1)$$

where x ranges through possible starting distributions, P is the transition matrix, and π is the stationary distribution. The larger this is, the faster the Markov chain mixes. This is very straightforward to characterize given the eigenvalues $1 = \lambda_1, \lambda_2, \dots, \lambda_n$. By defining the *spectral gap* as

$$\gamma = 1 - \max_{\lambda_i \neq 1} |\lambda_i|,$$

we get a quantity that is essentially the same as the rate $\tilde{\gamma}$ defined above in (1) (the ratio $\gamma/\tilde{\gamma}$ is asymptotically 1 as they approach 0, see [10] for a general introduction on Markov chains). Maximizing γ for a given connectivity graph is not an easy task in general. Computational schemes exists to formulate the problem as an SDP, see Boyd et al. [11],[12]. However, for these methods we need to additionally impose that the Markov chain must be reversible. By *reversibility* we mean

$$\pi_i P_{ij} = \pi_j P_{ji}, \qquad \forall i, j.$$

This signifies a certain symmetry: on every edge, transitions occur with the same frequency in the two directions. The final goal would be to get an analytic solution, without assuming the technical requirement of reversibility.

In this paper we target this ambitious goal for the simplest case. Assume the connectivity graph is a cycle on the n nodes. The simplest Markov chain to consider is the symmetric random walk, where every transition happens with a predefined probability d and the chain stays put with probability $1 - 2d$. It is well known that the spectral gap of the symmetric random walk is $\Theta(1/n^2)$. Previous work [1] shows that the mixing time can be at most constant factor better for any other Markov chain. Now we want to show that the symmetric random walk is actually the best possible choice. To do this, we start with benchmark Markov chains. We choose those which are reversible and which are *homogeneous*, meaning that the transition matrix P is invariant under cyclic permutation of the nodes. We then investigate the effect of perturbations which cause to lose reversibility, homogeneity, or both. We will show that reversible homogeneous Markov chains are globally optimal if only one type of perturbation is allowed and they are at least locally optimal when both types of perturbations are present. Moreover, numerical experiments suggests that global optimality also holds in this general case.

In the end, we demonstrate that there is no hope to speed up a symmetric random walk on the cycle to improve the mixing rate, we need to enrich the set of possible transitions in order to achieve this.

The rest of the paper is organized as follows. In Section II we introduce the tools required, also identifying the inhomogeneity and non-reversibility in the transition matrix. In Section III we analyze reversible homogeneous Markov chains and find the fastest among them. Then in Section IV we present the optimality results for different perturbations.

[1]Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM), Université catholique de Louvain, Louvain-la-Neuve, 1348 Belgium firstname.lastname@uclouvain.be Their work is supported by the DYSCO Network (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian Federal Science Policy Office, and by the Concerted Research Action (ARC) of the French Community of Belgium.

We complement our results with numerical simulations in Section V, finally we conclude in Section VI.

II. Preliminaries

We focus our attention to Markov chains on a cycle. Therefore we consider the set $\mathcal{S}$ of transition matrices of the type

$$S = \begin{bmatrix} a_1 & b_1 & & & c_n \\ c_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & b_{n-1} \\ b_n & & & c_{n-1} & a_n \end{bmatrix} \quad (2)$$

where moreover S is non-negative and doubly stochastic, i.e.

$$e^T S = e^T, \quad e^T S^T = e^T, \quad e^T := [1, 1, \dots, 1].$$

We want to capture separately the inhomogeneity and the non-reversibility of these Markov chains. First, among the matrices of $\mathcal{S}$ we select the subset $\mathcal{S}_0$ of reversible, homogeneous transition matrices. These have the form

$$S_0(d) = \begin{bmatrix} 1-2d & d & & & d \\ d & 1-2d & d & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & d \\ d & & & d & 1-2d \end{bmatrix}, \quad (3)$$
$$0 \le d \le \frac{1}{2}.$$

As we will see soon, the two types of perturbations correspond to adding some element from two sets of matrices. First, inhomogeneity will be captured by adding an element from the set $\mathcal{H}$ of the following (symmetric) matrices:

$$\begin{bmatrix} -\delta_n - \delta_1 & \delta_1 & & & \delta_n \\ \delta_1 & -\delta_1 - \delta_2 & \delta_2 & & \\ & \ddots & \ddots & & \\ & & \ddots & \delta_{n-1} & \\ \delta_n & & \delta_{n-1} & -\delta_{n-1} - \delta_n & \end{bmatrix}, \quad (4)$$
$$\sum_i \delta_i = 0.$$

The non-reversibility effect will be represented by a single matrix R as follows:

$$R = \begin{bmatrix} 0 & 1 & & & -1 \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 1 & & & -1 & 0 \end{bmatrix}. \quad (5)$$

We show that it is indeed possible to separate the inhomogeneous and non-reversible effects from the homogeneous reversible part. The following lemma is an extension of Lemma 1 of [1].

Lemma 1: Given $S \in \mathcal{S}$ there exist unique $S_0 \in \mathcal{S}_0$, $H \in \mathcal{H}, r \in \mathbb{R}$ such that

$$S = S_0 + H + rR. \quad (6)$$

Moreover, this decomposition is orthogonal w.r.t. the Frobenius norm $\|.\|_F$.

Proof: Let us take the decomposition

$$S = \frac{S + S^T}{2} + \frac{S - S^T}{2} =: \tilde{S} + Q.$$

We claim the part Q is of the form rR. Observe that

$$Q = -Q^T,$$
$$e^T Q = e^T Q^T = 0^T.$$

Let us denote Q_{12} by r, then by the skew-symmetry we get $Q_{21} = -r$, and based on the row sums we further arrive at $Q_{23} = r$. By repeating this argument we finally get $Q = rR$.

The remaining part $\tilde{S}$ is symmetric. Choose d such that $1 - 2d = \frac{1}{n}\operatorname{tr}(\tilde{S})$. It is easy to verify that $0 \le d \le 1/2$. Take $S_0 = S_0(d)$ using this d and define $H = \tilde{S} - S_0$. Clearly $S_0 \in \mathcal{S}_0$ by explicitly defining it this way. We need to show $H \in \mathcal{H}$. Both $\tilde{S}$ and S_0 are doubly stochastic and symmetric thus we have

$$e^T H = e^T H^T = 0^T,$$
$$H = H^T.$$

Moreover, the definition of d asserts that the sum of the diagonal of H is 0. These imply $H \in \mathcal{H}$.

It remains to show the decomposition is orthogonal. The inner product associated to the Frobenius norm is $\langle A, B \rangle_F = \sum_{i,j} \bar{A}_{i,j} B_{i,j}$. Clearly $\langle S_0, R \rangle_F = \langle H, R \rangle_F = 0$ because of the symmetry and skew-symmetry of the matrices. We also have $\langle S_0, H \rangle_F = (1 + 4d) \sum_i \delta_i = 0$. ∎

III. Reversible homogeneous Markov chains

It is a simple exercise to compute the spectral gap for the reversible homogeneous matrices in $\mathcal{S}_0$ and to find the optimal one among them as shown below.

Proposition 1: The spectral gap of a matrix $S_0(d) \in \mathcal{S}_0$ is

$$\gamma = \min\left(2d\left(1 - \cos\frac{2\pi}{n}\right), \right.$$
$$\left. 2 - 2d\left(1 - \cos\frac{2\pi\lceil\frac{n+1}{2}\rceil}{n}\right)\right).$$

Among the matrices in $\mathcal{S}_0$ we get the largest spectral gap for some $d = \frac{1}{2} + O\left(\frac{1}{n^2}\right)$. This optimal spectral gap is

$$\gamma = \frac{2\pi^2}{n^2} + O\left(\frac{1}{n^4}\right).$$

Proof: Since all the matrices of $\mathcal{S}_0$ are circulant matrices, the eigenvectors are the columns of

$$W := \begin{bmatrix} 1 & 1 & 1 & & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2n-2} & \dots & \omega^{(n-1)^2} \end{bmatrix}, \quad (7)$$
$$\omega := \exp(2\pi i/n).$$

For general properties of circulant matrices, see [13]. Also, the eigenvalues can be obtained by multiplying the first row of our matrix $S_0(d)$ with W:

$$[\lambda_1, \lambda_2, \dots, \lambda_n] = [1-2d, d, 0, \dots, 0, d]\, W.$$

In our case we get

$$\lambda_k = 1 - 2d\left(1 - \cos\frac{2\pi(k-1)}{n}\right).$$

All these eigenvalues are real, therefore the spectral gap γ is determined by the eigenvalue $\lambda_2, \dots, \lambda_n$ which is closest to 1 or -1. It is thus easy to check that

$$\gamma = \min\left(2d\left(1-\cos\frac{2\pi}{n}\right), 2 - 2d\left(1 - \cos\frac{2\pi\lceil\frac{n+1}{2}\rceil}{n}\right)\right).$$

We get the maximal γ by the proper choice of d when the two terms in the minimization become equal. This happens when

$$d = \frac{1}{1-\cos\frac{2\pi}{n} + 1 - \cos\frac{2\pi\lceil\frac{n+1}{2}\rceil}{n}} = \frac{1}{2} + O\left(\frac{1}{n^2}\right).$$

Plugging this back to the expression of γ we arrive at

$$\begin{aligned}\gamma &= \left(1 + O\left(\frac{1}{n^2}\right)\right)\left(1 - \cos\frac{2\pi}{n}\right)\\ &= \frac{2\pi^2}{n^2} + O\left(\frac{1}{n^4}\right).\end{aligned}$$

■

IV. Local behavior

In this section we investigate the effect of different types of perturbations applied to reversible, homogeneous Markov chains.

Proposition 2: For any $S_0 \in \mathcal{S}_0$ homogeneous reversible transition matrix and $r \in \mathbb{R}$ we have

$$\gamma(S_0 + rR) \le \gamma(S_0),$$

as long as $S_0 + rR \in \mathcal{S}$, in other words as long as it stays in the domain of doubly stochastic matrices. The matrix R is the one corresponding to non-reversible perturbations as defined in (5).

Proof: Both S_0 and rR are circulant matrices so they share eigenvectors. Consequently, the eigenvalues of $S_0 + rR$ are the sums of the corresponding eigenvalues of S_0 and rR. Note that rR has only imaginary eigenvalues (and 0). The modulus of the real eigenvalues of S_0 can only increase by adding imaginary numbers, so the spectral gap can only decrease. ■

Proposition 3: For any $S_0 \in \mathcal{S}_0$ and $H \in \mathcal{H}$ perturbation matrix for inhomogeneity we have

$$\gamma(S_0 + H) \le \gamma(S_0),$$

as long as $S_0 + H \in \mathcal{S}$.

Proof: Notice that when performing unitary similarity transformations on matrices, their orthogonality is preserved (when using the inner product for complex matrices). In particular, the Discrete Fourier Transform performed via W (as defined in (7)) maps pairs of orthogonal matrices to other pairs of orthogonal matrices. Let us use

$$\frac{1}{n}W^*(S_0 + H)W = \frac{1}{n}W^* S_0 W + \frac{1}{n}W^* H W =: \hat{S}_0 + \hat{H}.$$

We know that the set $\hat{\mathcal{S}}_0 = \{\hat{S}_0 : S_0 \in \mathcal{S}_0\}$ is a subset of diagonal matrices (since S_0 is circulant). Observe that $\mathcal{H}$ is orthogonal to not only $\mathcal{S}_0$, but to all circulant matrices, as the sum of elements is 0 along the diagonal and all its circularly shifted variants. Therefore $\hat{\mathcal{H}} = \{\hat{H}\}$ is orthogonal to all diagonal matrices. This means that $\hat{H}$ is zero on the diagonal. Moreover, it follows from (4) that the first column and row of $\hat{H}$ is zero, and also that it is Hermitian since H is symmetric. We thus have that

$$\hat{S}_0 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix}, \tag{8}$$

$$\hat{H} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \times & \dots & \times \\ 0 & \times & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \times \\ 0 & \times & \dots & \times & 0 \end{bmatrix}. \tag{9}$$

To bound the spectral gap we need to evaluate the spectral radius of the $(n-1)\times(n-1)$ trailing submatrix of $\hat{S}_0 + \hat{H}$, which we denote by

$$\tilde{S}_0 := \begin{bmatrix} \lambda_2 & 0 & \dots & 0 \\ 0 & \lambda_3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}, \tag{10}$$

$$\tilde{H} := \begin{bmatrix} 0 & \times & \dots & \times \\ \times & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ \times & \dots & \times & 0 \end{bmatrix}. \tag{11}$$

We claim that the spectrum of $\tilde{S}_0$ lies in the interior of the spectral radius of $\tilde{S}_0 + \tilde{H}$. Indeed, for $e_i = (0, 0, \dots, 1, \dots, 0)^T \in \mathbb{R}^{n-1}$,

$$(\tilde{S}_0 + \tilde{H})e_i = \lambda_{i+1}e_i + v_i.$$

Here we see $\langle e_i, v_i\rangle = 0$ just by checking where the non-zero entries are in both vectors. This implies

$$\|(\tilde{S}_0 + \tilde{H})e_i\|_2 = \|\lambda_{i+1}e_i + v_i\|_2 \ge |\lambda_{i+1}|.$$

For Hermitian matrices the largest eigenvalue modulus and the norm coincide, so we conclude that the largest eigenvalue modulus of $\tilde{S}_0 + \tilde{H}$ is at least $\max_{i=2}^n |\lambda_i|$. The claim of the proposition follows. ■

Theorem 4: For any $S_0 \in \mathcal{S}_0$ the spectral gap $\gamma(.)$ has a local optimum at S_0 in the domain $(S_0 + \mathcal{H} + \mathbb{R}R) \cap \mathcal{S}$.

Moreover it is a strong local optimum with a minor exception: if n is even and $\operatorname{tr}(S_0) = O\left(\frac{1}{n}\right)$, $\gamma(.)$ is locally constant for rR perturbations.

Proof: The eigenvalues determining $\gamma(S_0)$ can be the largest ones, $\lambda_2 = \lambda_n$, remember the indexing used in Proposition 1. It could also be the smallest one(s), $\lambda_{\frac{n}{2}+1}$ or $\lambda_{\frac{n-1}{2}}$ and $\lambda_{\frac{n+1}{2}}$ (depending on the parity of n), but for simplicity, we assume $\lambda_2 = \lambda_n$ are the ones determining $\gamma(S_0)$, a similar argument works in the other cases. In the end, we want to show that the moduli $|\lambda_2| = |\lambda_n|$ increase for any small perturbation in $\mathcal{H} + \mathbb{R}R$. Note that the matrix $S(t)$ remains a real matrix, so we still have $\lambda_2 = \bar{\lambda}_n$ through any perturbation.

The function $|\lambda_2|^2 = \lambda_2\lambda_n$ is smooth at S_0 as a function of the perturbations, see [14] for an in-depth discussion on perturbation theory. We show that it is first order constant and that the Hessian is positive definite. We may write the expression of interest as

$$|\lambda_2|^2 = \hat{\lambda}^2 + (\operatorname{Im}\lambda_2)^2, \tag{12}$$

where $\hat{\lambda}$ is tracking the eigenvalue average $\hat{\lambda} := (\lambda_2+\lambda_n)/2$.

At first let us restrict our attention to one-dimensional perturbations in an arbitrary direction. Let us fix $H \in \mathcal{H}$ and $r \in \mathbb{R}$ and define the matrix function

$$S(t) = S_0 + t(H + rR) =: S_0 + tT.$$

We apply the theory of perturbations of linear operators described in Kato ([14] section II.2). We use Kato's expansion of the "$\hat{\lambda}$ group" as a function of t:

$$\begin{aligned}
&\hat{\lambda}(t) = \hat{\lambda} + t\hat{\lambda}^{(1)} + t^2\hat{\lambda}^{(2)} + \dots,\\
&\lambda^{(1)} = \frac{1}{2}\operatorname{tr} PTP,\\
&\lambda^{(2)} = -\frac{1}{2}\operatorname{tr}(PTPTQ + PTQTP + QTPTP),
\end{aligned} \tag{13}$$

where P is the (constant) orthogonal projection to the eigenspace of the double eigenvalue $\hat{\lambda} = \lambda_2 = \lambda_n$ of S_0, T is the perturbation matrix $H + rR$ and Q is the reduced resolvent $(I-P)(S_0 - \hat{\lambda}I)^{-1}(I-P)$. Before going further we simplify the formulas by claiming

$$\operatorname{tr} PTPTQ = \operatorname{tr} TPTQP = \operatorname{tr} 0 = 0.$$

The first equation follows from $\operatorname{tr} AB = \operatorname{tr} BA$, the second one from $QP = PQ = 0$. This cancels the first term of $\hat{\lambda}^{(2)}$, and a similar argument works for the third term.

In order to evaluate $\hat{\lambda}^{(1)}, \hat{\lambda}^{(2)}$ we change bases using the W defined before, similarly as we did in the proof of Proposition 3. We get the matrices described as follows. Recall the two we identified earlier:

$$\hat{S}_0 = diag(1, \lambda_2, \lambda_3, \dots, \lambda_n),$$

$$\hat{H} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0\\ 0 & 0 & \times & \dots & \times\\ 0 & \times & 0 & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & \times\\ 0 & \times & \dots & \times & 0 \end{bmatrix}.$$

Notice that R is circulant, so $\hat{R}$ is diagonal. Moreover, it is a real skew-symmetric matrix, leading to imaginary eigenvalues in pairs (or 0).

$$\begin{aligned}
r\hat{R} &:= diag(0, ix_1, ix_2, \dots, -ix_2, -ix_1),\\
&\quad x_1, x_2, \dots \in \mathbb{R},\\
\hat{T} &:= \hat{H} + r\hat{R}.
\end{aligned}$$

The matrix P is projecting to some eigenspace which means some ones and some zeros on the diagonal once it is written in the basis of eigenvectors.

$$\hat{P} := diag(0, 1, 0, 0, \dots, 0, 1).$$

Similarly, Q is basically the inverse of $S_0 - \hat{\lambda}I$ but disregarding the eigenspace corresponding to $\hat{\lambda}$. This is again simple to describe in the eigenvector basis, in the following way:

$$\begin{aligned}
\hat{Q} := diag(&(1-\hat{\lambda})^{-1}, 0, (\lambda_3 - \hat{\lambda})^{-1},\\
&\dots, (\lambda_{n-1} - \hat{\lambda})^{-1}, 0).
\end{aligned}$$

With this notation, we get

$$\lambda^{(1)} = \frac{1}{2}\operatorname{tr}\hat{P}\hat{T}\hat{P}, \quad \lambda^{(2)} = -\frac{1}{2}\operatorname{tr}\hat{P}\hat{T}\hat{Q}\hat{T}\hat{P}.$$

In $\hat{P}\hat{T}\hat{P}$ only the terms $ix_1, -ix_1$ remain on the diagonal, therefore $\hat{\lambda}^{(1)} = 0$. Simple expansion of the matrix product $\hat{P}\hat{T}\hat{Q}\hat{T}\hat{P}$ shows that

$$\operatorname{tr}\hat{P}\hat{T}\hat{Q}\hat{T}\hat{P} = \sum_{k=3}^{n-1}\hat{H}_{2,k}(\lambda_k - \hat{\lambda})^{-1}\hat{H}_{k,2} \tag{14}$$

$$+ \sum_{k=3}^{n-1}\hat{H}_{n,k}(\lambda_k - \hat{\lambda})^{-1}\hat{H}_{k,n}. \tag{15}$$

By the choice of $\hat{\lambda}$ we have $\lambda_k < \hat{\lambda}$ for $k = 3, 4, \dots, n-1$. Note that $\hat{H}$ is Hermitian, so $\hat{H}_{l,k}\hat{H}_{k,l} \geq 0$, therefore all the terms of (14) are non-positive. We need to check if it can be zero. This can only happen if all $\hat{H}_{2,k}, \hat{H}_{k,2}, \hat{H}_{n,k}, \hat{H}_{k,n}$ are 0 for $k = 3, \dots, n-1$. In this case the second and last row and column of $\hat{H}$ are zero, thus all the rows of H are orthogonal to the second and last row of W. This means

$$\omega^{-1}\delta_i - (\delta_i - \delta_{i+1}) + \omega\delta_{i+1} = 0, \quad i = 1, \dots, n.$$

Knowing all δ_i are real and they sum to 0 this can happen only if they are all 0, consequently $H = 0$.

The above reasoning works for any perturbation, thus together with (13) we get

$$\hat{\lambda}^{(2)} \geq 0$$

in general and equality occurs only when the perturbation is of the form rR. This shows that the Hessian of $\hat{\lambda}$ is positive semidefinite and is only singular along the axis of rR.

In our way to understand the local behaviour of $|\lambda_2|^2$, we have so far investigated the first term of (12). Let us turn our attention to the second term, $(\operatorname{Im}\lambda_2)^2$. It is a non-negative smooth function which is 0 at S_0, so it must have zero derivative and positive semidefinite Hessian. Moreover, the Hessian is positive definite in the direction of rR as

$\operatorname{Im}\lambda_2$ changes linearly (with nonzero rate) in this direction (see the proof of Proposition IV).

To put the pieces together, we check again the expression

$$|\lambda_2|^2 = \hat{\lambda}^2 + (\operatorname{Im}\lambda_2)^2.$$

We have seen that the two functions on the right have zero derivative and positive semidefinite Hessians. Furthermore, the Hessian of the second term is strictly positive definite in the only direction for which the Hessian of the first term is singular. Consequently the sum is positive definite.

Let us comment on the other cases where we have to consider other eigenvalues as the ones with the largest moduli. Whenever these are $\lambda_{\frac{n-1}{2}}$ and $\lambda_{\frac{n+1}{2}}$, the same reasoning applies. When it happens to be $\lambda_{\frac{n}{2}+1}$, we find that $\operatorname{Im}\lambda_{\frac{n}{2}+1}$ is first order constant for rR perturbations. This is the case described as an exception in the theorem statement. ∎

V. Numerical results

Simulation output is in line with our claims, and even the global optimality of homogeneous reversible chains seem justifiable. We start with $S_0(1/3)$ for 100 nodes. This example has all positive entries at the allowed positions, thus it allows for a wide range of perturbations. We take 2 000 000 random directions for perturbations and modify the matrix in each of these directions until we exit the domain $\mathcal{S}$. In Figure 1 we plot the resulting spectral gaps against the Frobenius norm of the perturbation. We see a clear decrease in the spectral gap as me move away from S_0.

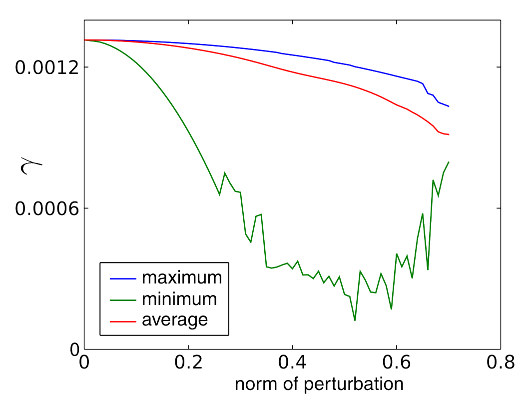

Fig. 1: Spectral gap of the perturbed matrix for 2 000 000 random perturbation directions.

There is a noisy part of the graph that seems surprising at first. In fact, as we generate random directions for the perturbations, only a very few of these directions allow a perturbation of large norm while staying in $\mathcal{S}$. As a result, when we move towards higher norms of perturbations, we see the aggregated results of less and less instances.

VI. Conclusions

In this paper we investigated the fastest Markov chain problem for the simplest non-trivial case, a cycle with n nodes.

We showed that any transition matrix can be neatly decomposed into a reversible homogeneous part and two separate perturbations, one responsible for breaking homogeneity, the other causing non-reversibility.

We presented partial results for spectral gap comparison. We have proven that reversible homogeneous Markov chains are globally optimal if we restrict perturbations to only one of the two types. Whenever both types of perturbations are present, we can still show local optimality of the reversible homogeneous Markov chains. Still, numerical simulations suggest that global optimality also holds in this more general setting.

The analytic confirmation of this general case remains for future research. Also, it would be interesting to extend the analysis for more complex connectivity graphs. There are examples where non-reversible Markov chains appear to be clearly faster than the best reversible ones [15], for example, when $\sqrt{n}$ random edges are added to the cycle. Consequently, getting rid of the reversibility condition would have a substantial impact on the performance we can achieve.

References

[1] B. Gerencsér, "Markov chain mixing time on cycles," *Stoch. Proc. Appl.*, vol. 121, no. 11, pp. 2553–2570, 2011.
[2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, 1953.
[3] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.
[4] M. Jerrum, "Mathematical foundations of the Markov chain Monte Carlo method," in *Probabilistic methods for algorithmic discrete mathematics*, vol. 16 of *Algorithms and Combinatorics*, pp. 116–165, Springer, 1998.
[5] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
[6] V. Blondel, J. M. Hendrickx, A. Olshevsky, and J. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *44th IEEE Conference on Decision and Control*, pp. 2996–3000, 2005.
[7] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM Siggraph Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987.
[8] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Review*, vol. 53, no. 4, pp. 747–772, 2011.
[9] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.
[10] P. Brémaud, *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, vol. 31. Springer, 1999.
[11] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing Markov chain on a graph," *SIAM Rev.*, vol. 46, no. 4, pp. 667–689 (electronic), 2004.
[12] S. Boyd, P. Diaconis, P. Parrilo, and L. Xiao, "Fastest mixing Markov chain on graphs with symmetries," *SIAM J. Optim.*, vol. 20, no. 2, pp. 792–819, 2009.
[13] R. M. Gray, *Toeplitz and circulant matrices: A review*. Now Publishers Inc, 2006.
[14] T. Kato, *Perturbation theory for linear operators*, vol. 132. Springer, 1995.
[15] B. Gerencsér, "Mixing times of Markov chains on a cycle with additional long range connections." arXiv:1401.1692, Submitted for Publication, 2014.