Report 99-060

Right Order Spectral Gap Estimates For Generating Sets of $\mathbb{Z}_4$

Matthias Löwe
Christian Meise

ISSN: 1389-2355

RIGHT ORDER SPECTRAL GAP ESTIMATES FOR GENERATING SETS OF $\mathbb{Z}_4$

MATTHIAS LÖWE AND CHRISTIAN MEISE

Abstract. Using coupling arguments, a distance method and Zeifman's method we give sharp estimates on the spectral gap for a special case of the class of Markov chains on generating n-tuples of Abelian groups. In our case the group is $\mathbb{Z}_4^n$.

Keywords: Markov chains, spectral gap, coupling

AMS 1991 subject classifications. 60J10, 60K25

1. Introduction

The problem considered in this paper is settled in the area of speed of convergence of Markov chains. In its most general form it has been posed by D. Aldous in a survey talk on the ICM 1994. It basically treats the problem of the speed of convergence of some algorithms used in computer algebra packages in order to produce random elements from a group G.

Generally speaking the setting is as follows: We are given a finite, directed, labelled graph $\mathcal{G} = (V, E)$ that is reversible (i.e. if $i \to j \in E$ then also $j \to i \in E$) and a finite group G. Moreover let Γ be a set that generates G. Classically, one would proceed in the following way to produce a random element from G. Start with the identity and successively choose random elements from Γ and multiply by them from the right.

A new algorithm proposed by Celler et. al. [CGMNO95] suggests to run a Markov chain on V^G with $|V| > |\Gamma|$, instead. More precisely, they propound to start with labelling $|\Gamma|$ vertices of V with the different elements of Γ and the rest by the identity. The random walk then proceeds by selecting randomly an edge from E (i.e. with the uniform distribution on all edges) and if this edge is $i \to j$ by replacing x_j (the element associated with vertex j) by $x_i x_j$ with probability 1/2 and by $x_i x_j^{-1}$ with probability 1/2.

In particular, if $\mathcal{G} = K_n$, the complete directed graph on n vertices, this walk has the uniform probability on all admissible n-tuples from G (that is to say, all $x \in G^n$ with $\langle x \rangle = G$) as its invariant measure. This does not quite imply that also the distribution of any coordinate of the walk converges to the uniform distribution on G. For example if $n = 2$, $\mathcal{G} = K_2$, and $G = \mathbb{Z}_2$ the fact that the first coordinate of the walk is 0 automatically implies that the other is 1. Hence the probability of having a 1 in the first coordinate is 2/3 rather than 1/2. But as n gets larger the difference between the uniform distribution on G (in each coordinate) and the one induced by the uniform distribution on all admissible configurations becomes more and more negligible.

Aldous' question now asks for the speed of convergence of this walk. In the two papers [DS96] and [DS97] Diaconis and Saloff-Coste analyse the above walk for different choices of $\mathcal{G}$ and G. In particular, for $\mathcal{G} = K_n$ and $G = \mathbb{Z}_2$ they show in [DS96] that about $\mathcal{O}(n^2 \log n)$ steps of the walk suffice to come close to stationarity.

This paper deals with with the case $\mathcal{G} = K_n$ and $G = \mathbb{Z}_4$ (the case $\mathcal{G} = K_n$ and $G = \mathbb{Z}_3$ has been studied in a previous paper by one of the authors (see [Me98]).
By applying a new technique – the so-called distance method, which has its roots in the theory of diffusions on manifolds – we are able to show that the spectral gap of the above chain is of order $\mathcal{O}(\frac{1}{n})$ (implying that after $\mathcal{O}(n\log(n))$ steps the walk is already close to the invariant measure such that the algorithm is very effective). This also suggests universality in the speed of convergence of the algorithm proposed by Celler et al. independent of the group G (since bounds on the spectral gap of the same order already were obtained for the cases $G = \mathbb{Z}_2$ by Chung and Graham [CG97] and $\mathbb{Z}_3$ by one of the authors [Me98]).

The so-called distance method employed in this paper is an estimate which allows to bound the difference between an arbitrary eigenvalue of an irreducible Markov chain and one. In particular, the spectral gap of a Markov chain is essentially given by the difference of its second largest eigenvalue and one. On the other hand, this spectral gap mainly governs the speed of convergence of the corresponding Markov chain. Hence the distance method also allows to estimate how quickly a Markov chain approaches its equilibrium distribution. Its basic idea is to compare a given unknown eigenfunction of the second-largest eigenvalue with a positive (distance) function. If the transition matrix satisfies some contraction property with respect to the distance function, we obtain an upper estimate of the second-largest eigenvalue. A more explicit explanation will be given in the third paragraph. For the origins of this method applied to time-continuous Markov processes the reader is referred to [CW94], [Ch96]. For first applications for discrete-time Markov-chains see [Me98].
Another method used in our proofs is due to Zeifman, cf. [Ze91]. It allows estimates on the rate of convergence for birth and death Markov chains. Similar results were recently obtained by Miclo using Hardy's inequality [Mi99].

The rest of the paper is organised as follows. Section 2 contains the basic setup, notations, and a statement of the results. In Section 3 we will explain our main tools, namely the distance method mentioned above and its connection to path coupling of Markov chains. Section 4 finally contains the proof.

2. Notations and the Result

Before we describe our main result let us agree on some notation.

So in the following let Q be the transition matrix of an irreducible, aperiodic Markov Y_n chain on a finite state space X, and let $L := L(Q) := Id - Q$ bet the associated (discrete) infinitesimal generator. Here, slightly overloading notation, we denote by Id the identity operator, no matter on which space it is defined. By irreducibility there exists a uniquely determined stationary distribution π of Q (that is $\pi Q = \pi$) that charges every point in X. Assume that Q is reversible with respect to π, that is the detailed balance equations

$$\pi(y)Q(y,x) = \pi(x)Q(x,y)$$

are fulfilled for every pair $x, y \in X$.
This especially implies that the eigenvalues of Q and hence the eigenvalues of L are real. We write the latter in the form

$$0 = \lambda_0 < \lambda_1 \leq \ldots \leq \lambda_{|X|-1} \leq 2.$$

From the theory of Markov chains it is well known that the distribution of Y_n converges to π independent of the starting point Y_0 and that the speed of convergence is mainly governed by the so-called spectral gap of Q. This is defined by

$$\operatorname{gap}(L) := \lambda_1 \vee (2 - \lambda_{|X|-1}).$$

By adding some holding to the Markov chain Y_n, for example by considering $\tilde{Q} := \frac{1}{2}(Q+I)$ instead of Q we can always achieve that (for the new Markov chain) $2 - \lambda_{|X|-1} \geq 1$ such that the spectral gap of the new chain agrees with the second largest eigenvalue λ_1 of $L(\tilde{Q})$.

This second largest eigenvalue (actually any eigenvalue) has a representation involving Dirichlet forms and L^2 norms of functions. So, for complex-valued functions f, g on X the scalar product on $L^2(\pi)$ is defined by

$$\langle f, g \rangle_{L^2(\pi)} = \sum_{x \in X} f(x)\overline{g(x)}\pi(x)$$

and the corresponding norm on $L^2(\pi)$ is denoted by $\|f\|_{L^2(\pi)}$.

Then the minimax characterisation of eigenvalues, see [HJ85] p. 176 states that

$$\lambda_1 = \inf \left\{ \mathcal{E}(f,f) \;:\; \|f\|_{L^2(\pi)} = 1,\, E_\pi(f) = 0 \right\}. \tag{1}$$

As already mentioned above, lower bounds on λ_1 yield upper bounds for the rate of convergence of the underlying Markov chain.

In this paper such bounds will be obtained mainly by using the so-called distance method to be explained in Section 3.

There the situation will be such that the Markov chain Y_n lives on the state space $\mathbb{Z}_4^n$ described in the introduction. More precisely, the state space X of the walk is given by

$$X = \{ x \in \mathbb{Z}_4^n \; : < \{x_1, \dots, x_n\} >= \mathbb{Z}_4 \}.$$

Here, for $A \subset \mathbb{Z}_4^n$, we denote by $\langle A \rangle$ the group generated by the set A. Given that the Markov chain is in state $g = (g_1, \dots, g_n)$, we choose an edge (i,j) in the edge set of K_n uniformly at random. Then with probability $1/2$ the element g_j is replaced by $g_i g_j$ or $g_i g_j^{-1}$ respectively.

This implies that the transition probabilities of the walk Y_n described by this procedure are given by

$$P(Y_{n+1} = (g_1, \dots g_j', \dots, g_n) | Y_n = (g_1, \dots g_j, \dots, g_n)) =$$
$$\left(\frac{1}{2n(n-1)} \sum_{\substack{i \neq j \\ g_i g_j = g_j'}} 1 \right) + \left(\frac{1}{2n(n-1)} \sum_{\substack{i \neq j \\ g_i g_j^{-1} = g_j'}} 1 \right).$$

Before we state our result let us finally introduce two more notations. For $x \in X$ we define

$$n_1(x) = |\{ 1 \leq i \leq n \;:\; x_i \in \{1,3\}\}|,$$

and

$$n_2(x) = |\{ 1 \leq i \leq n \;:\; x_i = 2\}|.$$

Moreover, we let $n(x) := (n_1(x), n_2(x))$ and call $n(x)$ the index of x.

For this random walk we have the following results that show that a bound on the spectral gap of $\mathcal{O}(n^{-1})$ not only can be achieved but also is optimal.

Theorem 2.1. *For the random walk Y_n described above the following bounds on the spectral gap $\lambda_1(Q)$ hold true. Let $c_1 = 4$ and $c_2 = \frac{1}{12}$. Then*

$$\frac{c_1}{n} + \geq \lambda_1(Q) \geq \frac{c_2}{n} + \mathcal{O}(n^{-2}). \tag{2}$$

Remark 2.2. *Note that with some additional work in the proofs presented below the bounds on c_2 might be slightly improved. We consider this of minor interest as long as c_1 and c_2 do not match. Theorem 2.1 above already shows that the spectral gap is of order $\frac{1}{n}$.*

3. The distance method

In this section we will have a closer look at our basic tool, the so-called distance-method. To this end we will use the notation introduced in Section 2 and let $\beta \neq 1$ be an eigenvalue of Q with corresponding eigenfunction f. Assume that there exists a function $d : X \to \mathbb{R}^+$ such that $[Qd](x) \leq \alpha d(x)$ for all $x \in X$ and some $\alpha \leq 1$. Here $[Q\cdot]$ denotes the L^2 operator associated with Q, that is

$$[Qf](x) = \sum_{y \in X} Q(x,y) f(y)$$

for all $f \in L^2(\pi)$.
Let

$$m := \max \left\{ \frac{|f(x)|}{d(x)} \; : \; x \in X \right\}.$$

Choose $x \in X$ such that $|f(x)| / d(x) = m$. Then we immediately obtain

$$\begin{aligned} |\beta f(x)| &= |[Qf](x)| = \left| \sum_{y \in X} Q(x,y) f(y) \right| \\ &\leq \sum_{y \in X} Q(x,y) \, |f(y)| = \sum_{y \in X} Q(x,y) \frac{|f(y)|}{d(y)} d(y) \\ &\leq \frac{|f(x)|}{d(x)} [Qd](x) \leq \alpha \, |f(x)| \, . \end{aligned}$$

Dividing by $|f(x)|$ we conclude $|\beta| \leq \alpha$. Thus we have proved

Proposition 3.1. *Let $d : X \to \mathbb{R}^+$ be a function, such that $[Qd](x) \leq \alpha d(x)$ for all $x \in X$ and a fixed $\alpha \leq 1$. Then for all eigenvalues $\beta \neq 1$ of Q, $|\beta| \leq \alpha$.*

It is well known that another method to proof estimates on the second largest eigenvalue of Q to couple two chains driven by Q such that the average time until they meet becomes small. The argument below combines this coupling approach with the distance method introduced above.

Proposition 3.2. *Let $d : X \times X \to \mathbb{R}^+$ be a function, such that for some coupling $\widetilde{Q}$ of (Q,Q) the inequality $[\widetilde{Q}d](x,y) \leq \alpha d(x,y)$ for all $x \neq y$ and some $\alpha \leq 1$ holds true. Then any eigenvalue $\beta \neq 1$ of Q satisfies $|\beta| \leq \alpha$.*

Proof. Similarly to Proposition 3.1 we define

$$m := \max \left\{ \frac{|f(y) - f(x)|}{d(x,y)} \; : \; x \neq y \in X \right\}.$$

Let (x,y) be such that $|g(x,y)| / d(x,y) = m$, where $g(x,y) = f(y) - f(x)$. We conclude

$$|\beta g(x,y)| = \left|\left[\widetilde{Q}g\right](x,y)\right| \leq m\left[\widetilde{Q}d\right](x,y) \leq m\alpha d(x,y).$$

Hence $|\beta| \leq \alpha$. □

Remarks 3.3. 1. *One might wonder why such easy calculations may result in a powerful tool for estimating the spectral gap of a Markov chain. The point and the most difficult part in the application of the distance method is that we are still free to choose two major ingredients of this method: the coupling and the distance function.*

The basic trick then is to find a distance function that models "the landscape" given by $|f(y) - f(x)|$ as closely as possible and on the other hand still allows for a nice coupling.

2. *Note that the technique of Proposition 3.2 also works if we let*

$$m := \max\left\{ \frac{|f(y) - f(x)|}{d(x,y)} \; : \; (x,y) \in S \subset X \times X \right\}$$

for some set $S \subset X \times X$ which is invariant under the coupling $\widetilde{Q}$, that is for some set $S \subset X \times X$ with the property that

$$\sum_{(x',y') \in S} \widetilde{Q}((x,y),(x',y')) = 1$$

whenever $(x,y) \in S$.

3. *In the applications we will use the distance method in the following way. Suppose that there is a set $M \subset X \times X$ of the form $M = \dot{\bigcup}_{i=1}^{k} M_i \times M_i$, such that there is a coupling $\widetilde{Q}$ keeping the "levels" $M_j \times M_j$ fixed. By that we mean that there is a coupling for which $(x,y) \in M_i \times M_i$ and $\widetilde{Q}((x,y),(x',y')) > 0$ implies $(x',y') \in M_j \times M_j$ for some j. Thus, the bivariate coupled chain never escapes from M, and when starting in the same M_i the two particles at the same time are always in the same M_j. Assume that there is a coupling $\widetilde{Q}$ of Q on M that contracts a function $d : X \times X \to \mathbb{R}_0^+$, that is, $[\widetilde{Q}d](x,y) \leq \alpha d(x,y)$ for some $\alpha < 1$ for all $(x,y) \in M_i \times M_i$, $i \in \{1,2,\ldots,k\}$ (indeed d may even depend on j). If there is an eigenfunction f to an eigenvalue β that is non-constant on some M_i, we obtain the estimate $|\beta| \leq \alpha$.*

Otherwise all eigenfunctions f for β are constant on each M_i, and we may identify all points in each M_i and analyse an induced Markov chain. If, for example, $Q(x, M_j) = Q(y, M_j)$ for all $x, y \in M_i$ and $i, j \in \{1,2,\ldots,k\}$ we can consider the induced Markov chain given by $K(i,j) := Q(x, M_j)$, where $x \in M_i$ is arbitrary. It is known that the spectrum of K is contained in the spectrum of Q. However, here we are interested in the second-largest or the smallest eigenvalues only. Hence the fact that all the eigenfunctions on the sets M_i are constant yields that the second-largest or the smallest eigenvalue of K must coincide with those of Q.

4. Proofs

Before we step into the details let us quickly exhibit the principal structure of the proof. It is basically split into two parts depending on whether the eigenfunctions to β_1 are all constant on the "level sets" of $n(\cdot)$, or not. This shows how Remark 3.3.3 is realized on our setting. So we distinguish between

A) There is an eigenfunction f to the eigenvalue β_1 of Q such that $f(x) = f(y)$ for all $x, y \in X$ with $n(x) = n(y)$. In this case, we can "glue all states with the same index $n(\cdot)$ together" and the analysis of the Markov chain Q reduces to that of an associated Markov chain on a new state space. This state space is given by a subset of the two-dimensional lattice, namely by

$$Y = \{(i,j) \, : \, i \geq 1, \, j \geq 0, \, i+j \leq n\}.$$

The corresponding transition probabilities are defined in (3) while Figure 1 illustrates the state space for the case $n = 7$.

and

B) For all eigenfunctions f to β_1 of Q there exist $x, y \in X$ such that $n(x) = n(y)$ and $f(x) \neq f(y)$.

The analysis of this case is split into different subcases. If x and y not only have the same index $n(\cdot)$ but also the same number of 1's, then x can be obtained by a permutation of y. As a permutation is nothing but the product of several transpositions, there also have to be x and y with $f(x) \neq f(y)$ such that x can be obtained from y by a simple transposition, which means they have Hamming distance two. According to in which arguments x and y differ we obtain four different subcases (as 1 and 3 can be treated identically). On the other hand, if e.g. x contains more 1's than y (and thus y contains more 3's) by the same sort of reduction argument we can assume that x contains exactly one more 1 than y and that they differ in precisely one position (say the first), where x is 1 while y is 3.

These cases will be treated at the end of this section.

Let us now turn to the proof of the main theorem.

Proof of Proposition 2.1

As indicated above the proof is divided into several parts

Case A) There exists an eigenfunction f to β_1 of Q such that for all $x, y \in X$ with $n(x) = n(y)$ we have $f(x) = f(y)$. Let us now identify all points $x \in X$ which have the same index $n(x)$. This results in a new space

$$Y = \{(i,j) \, : \, i \geq 1, \, j \geq 0, \, i+j \leq n\}.$$

Here i stands for $n_1(x)$, while j abbreviates $n_2(x)$. Under the above assumption f induces a function $\tilde{f}$ via $\tilde{f}(i,j) = f(x)$, for an $x \in X$ with $n(x) = (i,j)$. The above assumption guarantees that $\tilde{f}$ is well defined.

Moreover the analysis of the Markov chain reduces to the analysis of the Markov chain K on Y, where the transition matrix K is given by

$$K((i,j),(i',j')) := P\left(n(X_t) = (i',j') \mid n(X_{t-1}) = (i,j)\right).$$

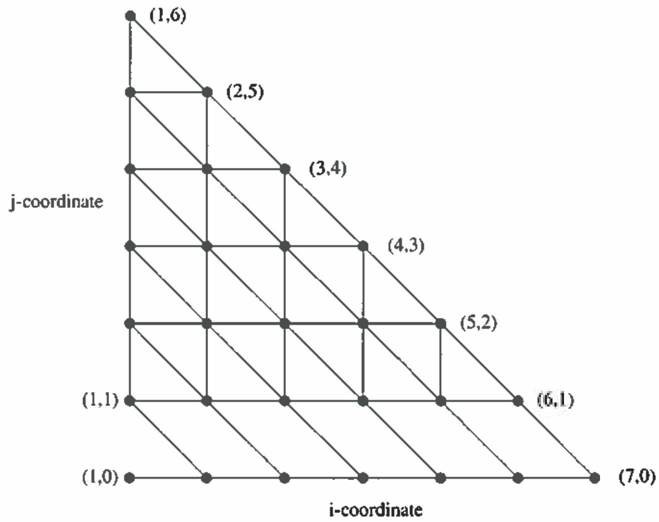

FIGURE 1. State space Y in the case $n = 7$.

More precisely the transition probabilities are

$$
\begin{array}{rlr}
K\left((i,j),(i+1,j)\right) = & \frac{(n-i-j)i}{n(n-1)} & \rightarrow \\
K\left((i,j),(i-1,j)\right) = & \frac{i(i-1)}{2n(n-1)} & \leftarrow \\
K\left((i,j),(i,j+1)\right) = & \frac{(n-i-j)j}{n(n-1)} & \uparrow \\
K\left((i,j),(i,j-1)\right) = & \frac{j(j-1)}{n(n-1)} & \downarrow \\
K\left((i,j),(i+1,j-1)\right) = & \frac{ij}{n(n-1)} & \searrow \\
K\left((i,j),(i-1,j+1)\right) = & \frac{i(i-1)}{2n(n-1)} & \nwarrow \\
K\left((i,j),(i,j)\right) = & \frac{ij}{n(n-1)} + \frac{n-i-j}{n} & \bullet
\end{array}
\tag{3}
$$

Here the arrows to the right side illustrate to which of the neighboring points in Y the particle is moving, e.g. '$\rightarrow$' means a move to the right (increment of the first coordinate) and '$\bullet$' indicates that the particle stays at its current position.

Using the minimax characterisation of the second largest eigenvalue quoted above in Section 2 (see 1), also see e.g. [HJ85]) together with the fact that f induces $\tilde{f}$, it is straightforward to check that the second-largest eigenvalue of Q coincides with that of K.
Therefore we estimate the spectral gap of the Markov chain K on the grid defined in (3).

Proposition 4.1. *The first positive eigenvalue λ_1 of the discrete generator $L = I - K$, where K is given by (3) satisfies*

$$\lambda_1 \geq \frac{1}{6}\frac{1}{n-1} + \mathcal{O}(n^{-3/2}).$$

Proof. Again, we differentiate between two cases.

First assume that there is an eigenfunction f of K such that $f((i,j)) = f((i,j'))$ for all i, j, j' with $(i,j) \in Y$ and $(i,j') \in Y$. Similar to the above reduction, in this case we can reduce the problem for estimating the second-largest eigenvalue of the chain K to that of the one-dimensional chain M with state space $Z = \{1, 2, \dots, n\}$ and with transition probabilities

$$M(i,i+1) = \frac{(n-i)i}{n(n-1)} \tag{4}$$

$$M(i,i-1) = \frac{i(i-1)}{2n(n-1)}. \tag{5}$$

This chain can and has been studied by Zeifman's method (cf. [Ze91], [Ze98]). In [Me98] it has been shown that indeed the Assumptions of Theorem 1 in [Ze91] are fulfilled implying that for any $\varepsilon > 0$ there is N such that $\lambda_1(M) \geq (1-\varepsilon)/(n-1)$ for all $n \geq N$, cf. also the corresponding estimates [Me98]. Similar results were recently obtained by Miclo [Mi99] by using Hardy's inequality. He is able to show that the spectral gap of the above chain is of order $1/n$ but not that the constant is actually one (which is enough for the applications we have in mind).

If, on the other hand, for any eigenfunction f of K there exists $(i,j) \in Y$ and $(i,j') \in Y$ such that $f((i,j)) \neq f((i,j'))$ we may apply a path coupling argument. To this end let us first introduce the coupling $\widetilde{K}$ of K that we will use. For i,j such that $n-i-2j+1 \geq 0$ we let

$$\begin{array}{rcll}
\widetilde{K}\left([(i,j),(i,j-1)],[(i+1,j),(i+1,j-1)]\right) &=& \frac{(n-i-j)i}{n(n-1)} & \rightarrow \ \rightarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i-1,j),(i-1,j-1)]\right) &=& \frac{i(i-1)}{2n(n-1)} & \leftarrow \ \leftarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j+1),(i,j)]\right) &=& \frac{(n-i-(j-1))(j-1)}{n(n-1)} & \uparrow \ \uparrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j-1),(i,j-2)]\right) &=& \frac{(j-1)(j-2)}{n(n-1)} & \downarrow \ \downarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i+1,j-1),(i+1,j-1)]\right) &=& \frac{i}{n(n-1)} & \searrow \ \rightarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i+1,j-1),(i+1,j-2)]\right) &=& \frac{i(j-1)}{n(n-1)} & \searrow \ \searrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i-1,j+1),(i-1,j)]\right) &=& \frac{i(i-1)}{2n(n-1)} & \nwarrow \ \nwarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j),(i,j-1)]\right) &=& \frac{(n-i-j)(n-1)+ij}{n(n-1)} & \bullet \ \bullet \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j+1),(i,j-1)]\right) &=& \frac{n-i-2j+1}{n(n-1)} & \uparrow \ \bullet \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j-1),(i,j-1)]\right) &=& \frac{2(j-1)}{n(n-1)} & \downarrow \ \bullet
\end{array}$$

For i,j such that $n-i-2j+1 < 0$ we let

$$\begin{array}{rcll}
\widetilde{K}\left([(i,j),(i,j-1)],[(i+1,j),(i+1,j-1)]\right) &=& \frac{(n-i-j)i}{n(n-1)} & \rightarrow \ \rightarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i-1,j),(i-1,j-1)]\right) &=& \frac{i(i-1)}{2n(n-1)} & \leftarrow \ \leftarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j+1),(i,j)]\right) &=& \frac{(n-i-j)j}{n(n-1)} & \uparrow \ \uparrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j-1),(i,j-2)]\right) &=& \frac{(j-1)(j-2)}{n(n-1)} & \downarrow \ \downarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i+1,j-1),(i+1,j-1)]\right) &=& \frac{i}{n(n-1)} & \searrow \ \rightarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i+1,j-1),(i+1,j-2)]\right) &=& \frac{i(j-1)}{n(n-1)} & \searrow \ \searrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i-1,j+1),(i-1,j)]\right) &=& \frac{i(i-1)}{2n(n-1)} & \nwarrow \ \nwarrow \\
\widetilde{K}\left([(i,j),(i,j-1)],[(i,j),(i,j-1)]\right) &=& \frac{(n-i-j+1)(n-1)+(j-1)(i-2)}{n(n-1)} & \bullet \ \bullet
\end{array}$$

$$\widetilde{K}\left([(i,j),(i,j-1)],[(i,j),(i,j)]\right) = \tfrac{2j+i-n-1}{n(n-1)} \quad \bullet \ \uparrow$$
$$\widetilde{K}\left([(i,j),(i,j-1)],[(i,j-1),(i,j-1)]\right) = \tfrac{2(j-1)}{n(n-1)} \quad \downarrow \ \bullet$$

Some tedious, but elementary calculations show that indeed $\widetilde{K}$ couples (K,K). Moreover, consider the set

$$S = \{\, [(i,j),(i,j')] \ : \ (i,j)\in Y,\ (i,j')\in Y \,\}.$$

and observe that the coupling $\widetilde{K}$ presented above guarantee that the coupled process will never leave S given that it starts in S. For this reason and in view of Remark 3.3.3 it is clearly not necessary to define transition probabilities of $\widetilde{K}$ for elements not in S. Next, we need to provide a distance function for neighboring elements in Y. For $(i,j),(i,j-1)\in Y$ we define,

$$d((i,j),(i,j-1)) = \begin{cases} \frac{1}{i(j+1)} & : \quad (i,j)\neq(1,1) \\ \frac{5}{4} & : \quad (i,j)=(1,1). \end{cases}$$

For elements $(i,j_1),(i,j_2)\in Y$ with $j_1<j_2$ we define

$$d((i,j_1),(i,j_2)) = \sum_{l=j_1+1}^{j_2} d((i,l),(i,l-1)). \tag{6}$$

Now we will prove an estimate of the type

$$\left[\widetilde{K}d\right]((i,j),(i,j-1)) \le \alpha d((i,j),(i,j-1)), \tag{7}$$

for all $(i,j),(i,j-1)\in Y$ for some fixed $\alpha<1$. Using (6) and the linearity of $\widetilde{K}$ for $(i,j_2),(i,j_1)\in Y$ this clearly yields

$$\left[\widetilde{K}d\right]((i,j_2),(i,j_1)) \le \alpha d((i,j_2),(i,j_1)),$$

for all $(i,j_2),(i,j_1)\in Y$ with $j_1<j_2$. Thus the distance method presented in Section 3 yields that $\lambda_1\le\alpha$ (and actually that $\lambda_i\le\alpha$ for every $i=1,\dots,|X|-1$).
In the following considerations we will use the abbreviation

$$d(i,j-1) = d((i,j),(i,j-1)).$$

Let us distinguish the following cases:
Case I: $n-i-2j<0$.
This is again divided into several subcases.
Case Ia: $n-i-2j<0$ **and** $i\ge2$, $j\ge2$.
According to the distance method sketched in Section 3 and our goal to bound the spectral gap of K by c/n we need to calculate $[\widetilde{K}d]$ and bound it appropriately. Let us start with computing $[\widetilde{K}]$.

$$\begin{aligned}
&n(n-1)[\widetilde{K}d]((i,j),(i,j-1)) \\
&= (n-i-j)id(i+1,j-1) + \frac{i(i-1)}{2}d(i-1,j-1) \\
&+(n-i-j)jd(i,j) + (j-1)(j-2)d(i,j-2) + i(j-1)d(i+1,j-2) \\
&+\frac{i(i-1)}{2}d(i-1,j) + (n-i-j+1)(n-1)d(i,j-1) + (i-2)(j-1)d(i,j-1).
\end{aligned}$$

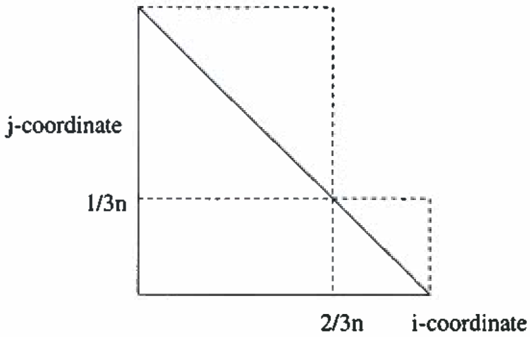

FIGURE 2. The rectangles cover the triangular state space.

Introducing the function

$$f : (i,j) \mapsto 2n(n-1)\left[\frac{[\widetilde{K}d]((i,j),(i,j-1))}{d(i,j-1)} - 1\right]$$

we we thus obtain

$$f(i,j) = -\frac{1}{2}\frac{4nij - 3i^2 + i^3 + 2ni + 2nj - 6i^2j}{(i+1)(j+1)} + \mathcal{O}(1).$$

As both nominator and denominator of f are linear functions in j one quickly checks that the function $j \mapsto f(i,j)$ is either monotonely decreasing or increasing (depending on the fixed value of i). We give estimates for $f(i,j)$ in each of the two rectangles indicated in Figure 2.

Case Ia1: $n - i - 2j < 0, i \geq 2$, $j \geq 2$ **and** $i \leq 2/3n$.
Using the monotonicity of the function $j \mapsto f(i,j)$we have

$$f(i,j) \leq f(i,2) \vee f(i,n).$$

Note that

$$f(i,2) = -\frac{1}{6}\frac{4n + 10ni - 15i^2 + i^3}{i} \leq -n + \mathcal{O}(1),$$

and

$$f(i,n) = -\frac{1}{2}\frac{2n^2 - 6i^2n + 2in + i^3 + 4n^2i}{(n+1)(i+1)} \leq -\frac{2}{9}n + \mathcal{O}(1).$$

for all $2 \leq i \leq 2/3n$.

Case Ia2: $n - i - 2j < 0, i \geq 2$, $j \geq 2$ **and** $2/3n < i \leq n$.
Again using the monotonicity of $j \mapsto f(i,j)$, we obtain that

$$f(i,j) \leq f(i,2) \vee f(i,n/3).$$

Estimating $f(i,2)$ and $f(i,n/3)$ we conclude that

$$f(i,j) \leq -\frac{1}{2}n + O(1),$$

for all $2/3n \leq i \leq n$.

Combining the estimates of the subcases shows that

$$f(i,j) \leq -\frac{2}{9}n + \mathcal{O}(1),$$

for all $i \geq 2$, $j \geq 2$ with $i + j \leq n$.

Case Ib: $n-i-2j<0$ and $i=1$, $j\geq 2$.
Here we obtain

$$\frac{[\widetilde{K}d]((1,j),(1,j-1))}{d(1,j-1)}-1=-\frac{1}{4n(n-1)}\frac{3jn+n-j-1}{j+1}\leq-\frac{1}{4(n-1)},$$

for all $2\leq j\leq n$.

Case Ic: $n-i-2j<0$ and $i\geq 2$, $j=1$.
The condition $n-i-2j<0$ implies that $i\geq n-1$ and hence $i=n-1$, since $i+j\leq n$. Therefore, we only need to compute

$$2n(n-1)\left[\frac{[\widetilde{K}d]((n-1,1),(n-1,0))}{d(n-1,0)}-1\right]=-\frac{1}{4}n^2+O(n).$$

So altogether we obtain

$$f(i,j)\leq-\frac{2}{9}n+\mathcal{O}(1),$$

in Case I. This by the definition of $f(i,j)$ yields

$$[\widetilde{K}d](i,j-1)\leq\alpha d(i,j-1)$$

with $\alpha\leq 1-\frac{2}{9n}+\mathcal{O}(\frac{1}{n^2})$.

Case II: $n-i-2j\geq 0$.

Case IIa: $n-i-2j\geq 0$ and $i\geq 2$, $j\geq 2$.
Similar to the above we obtain

$$\begin{aligned}n(n-1)[\widetilde{K}d]((i,j),(i,j-1)) &= (n-i-j)id(i+1,j-1)+\frac{i(i-1)}{2}d(i-1,j-1)\\ &\quad+(n-i-j+1)(j-1)d(i,j)+(j-1)(j-2)d(i,j-2)\\ &\quad+i(j-1)d(i+1,j-2)+i(i-1)/2d(i-1,j)\\ &\quad+(n-i-j)(n-1)d(i,j-1)+ijd(i,j-1)\\ &\quad+(n-i-2j+1)[d(i,j-1)+d(i,j)].\end{aligned}$$

Defining the function

$$f:(i,j)\mapsto 2n(n-1)\left[\frac{[\widetilde{K}d]((i,j),(i,j-1))}{d(i,j-1)}-1\right],$$

we obtain

$$f(i,j) = -\frac{-2n+4j^2-i^2-4i^2j+4j^2i+i^3+2nij}{(i+1)(j+1)}+\mathcal{O}(1).$$

Our aim is to give an upper estimate for the expression to the right side under the condition $n-i-2j\geq 0$, $i\geq 2$ and $j\geq 2$. For the sake of convenience we will consider the function $h(i,j):=f(i-1,j-1)$ rather than f. Considering h as a function of j for fixed i, we obtain that the first derivative has zeros $j=1/2\sqrt{i^2-2n}$ and $j=-1/2\sqrt{i^2-2n}$. Clearly, the negative value for j is meaningless for our problem.

Case IIa1: $n-i-2j \geq 0$, $i \geq 2$, $j \geq 2$ and $i < \sqrt{2n}$.
In this case $h'(i,\cdot)$ has no real zeros at all, so that the function $h(i,\cdot)$ is monotone. Thus

$$\max_{3 \leq j \leq n/2+1} h(i,j) \leq h(i,3) \vee h(i,n/2+1).$$

Note that

$$h(i,3) = -\frac{1}{3}\frac{36i+4ni-12i^2-6n+i^3}{i} \leq -\frac{2}{3}n$$

for all $3 \leq i \leq n$. Moreover, we have

$$h(i,n/2+1) = -2\frac{2n^2i-2ni^2-n^2+i^3}{i(n+2)} + \mathcal{O}(1) \leq -3n + \mathcal{O}(1)$$

for all $3 \leq i \leq \sqrt{2n}$.

Case IIa2: $n-i-2j \geq 0$, $i \geq 2$, $j \geq 2$ and $i \geq \sqrt{2n}$.
As $h(i,\cdot)$ has two local extrema, one with a positive and one with a negative j one readily sees that in this case the function $j \mapsto h(i,j)$, $j \geq 1$, is concave in our domain of interest. Therefore $h(i,j)$ is maximal for $j_0 = 1/2\sqrt{i^2-2n}$. Writing i in the form $i = \alpha(i)\sqrt{n}$ we have

$$h(i,j_0) = -2\frac{2i^2-4n}{\sqrt{i^2-2n}} + 4i + 2\frac{n}{i} - 2n = -2g(\alpha(i)) - 2n,$$

where g denotes the concave function

$$g(x) = \sqrt{n}\left[\frac{(2x^2-4)}{\sqrt{x^2-2}} - 2x - \frac{1}{x}\right].$$

Note that $\lim_{x\to\sqrt{2}} g(x) = -2^{3/2}\sqrt{n} - 2^{-1/2}\sqrt{n}$ and $g(\sqrt{n}) = O(1)$. This yields that

$$\max_{\sqrt{2n}\leq i \leq n} h(i,j_0) \leq \sup_{\sqrt{2}\leq\alpha\leq\sqrt{n}} [-2g(\alpha) - 2n] \leq -2n + O(\sqrt{n}).$$

Hence in Case IIa we obtain that

$$\max h(i,j) \leq -\frac{2}{3}n + \mathcal{O}(\sqrt{n})$$

(where the max is taken over all pairs (i,j) that are admissible for Case IIa) implying that

$$[\widetilde{K}d](i,j-1) \leq \alpha d(i,j-1)$$

with $\alpha \leq 1 - \frac{2}{3n} + \mathcal{O}(\frac{1}{n^{3/2}})$.

Case IIb: $n-i-2j \geq 0$ and $i \geq 3$, $j = 1$.
Here the following estimate holds

$$\frac{[\widetilde{K}d]((i,1),(i,0))}{d(i,0)} - 1 \quad = \quad -\frac{1}{4n(n-1)}\frac{2(i-1)n+2-i^2+i^3}{i+1} \leq -\frac{n+5}{4n(n-1)},$$

for all $3 \leq i \leq n$.

Cases IIc: $n-i-2j \geq 0$ and $i = 2$, $j = 1$ or $n-i-2j \geq 0$ and $i = 1$, $j = 1$.
Here we compute

$$\frac{[\widetilde{K}d]((2,1),(2,0))}{d(2,0)} - 1 = -\frac{1}{3(n-1)}$$

and

$$\frac{[\widetilde{K}d]((1,1),(1,0))}{d(1,0)} - 1 = -\frac{1}{5(n-1)} - \frac{3}{5n(n-1)}.$$

Finally we have to treat
Case IId: $n-i-2j \geq 0$ and $i=1$, $j \geq 2$.

$$\begin{aligned}\frac{[\widetilde{K}d]((1,j),(1,j-1))}{d(1,j-1)} - 1 &= -\frac{1}{2n(n-1)}\frac{(j-1)n+4j^2+3j-1}{j+1} \\ &\leq -\frac{1}{6(n-1)} - \frac{7}{2n(n-1)}.\end{aligned}$$

Concluding we see that in all the cases considered above

$$[\widetilde{K}d](i,j-1) \leq \alpha d(i,j-1)$$

with $\alpha \leq 1 - \frac{1}{6(n-1)} + \mathcal{O}(\frac{1}{n^{3/2}})$. which in turn implies that

$$\lambda_1 \geq \frac{1}{6(n-1)} + \mathcal{O}(\frac{1}{n^{3/2}}).$$

This finishes the analysis of the spectral gap in Part A. □

Let us now turn to Part B, that is the case where for every eigenfunction f to the second largest eigenvalue of Q there are $x, y \in X$ with $n(x) = n(y)$ but $f(x) \neq f(y)$. Here we will use the following strategy: First we show that the spectral gap is of the order $\mathcal{O}(n^{-1})$ in a subcase of the case where $n(x) = n(y)$ but x and y contain a different number of 1's. Then we treat the case where $n(x) = n(y)$, x and y contain the same number of 1's, but $f(x) \neq f(y)$. In a final step we see that these two cases already suffice to cover the general case.

So, let us first consider a special case of this situation, where these x and y, although having the same signature $n(\cdot)$ contain a different number of 1's, say x has one more 1 than y or vice versa. Moreover we assume x and y differ in just one coordinate, that is to say, that $d_H(x,y) = 1$, where $d_H(x,y) = \sum_{i=1}^n \delta_{x_i \neq y_i}$ is the Hamming distance. Let us call the set of all such pairs (x,y), $\mathcal{S}_1$, so

(8)

$$\mathcal{S}_1 := \{(x,y) \in X \times X : n(x) = n(y), |\sum_{i=1}^n \delta_{x_i=1} - \sum_{i=1}^n \delta_{y_i=1}| = 1, \text{ and } d_H(x,y) = 1\}.$$

So we will now in a first step see, that if $f(x) \neq f(y)$ for a pair $(x,y) \in \mathcal{S}_1$ we obtain a spectral gap for the Markov chain Y_n of the order $\mathcal{O}(n^{-1})$.
To this end we define a rather obvious coupling for such pairs $(x,y) \in \mathcal{S}_1$. Without loss of generality we assume that $x = (1, x_2, \ldots, x_n)$, and $y = (3, y_2 \ldots y_n)$ with $x_j = y_j$ for $j \geq 2$. We define the coupling $\widetilde{Q}$ by

$$\begin{aligned}\widetilde{Q}((x,y),[(2,x_2\ldots x_n),(2,y_2\ldots y_n)]) &:= \frac{i-1}{2n(n-1)} \\ \widetilde{Q}((x,y),[(0,x_2\ldots x_n),(0,y_2\ldots y_n)]) &:= \frac{i-1}{2n(n-1)} \\ \widetilde{Q}((x,y),(x',y')) &:= Q(x,x') = Q(y,y'),\end{aligned}$$

where again $(x', y') \in \mathcal{S}_1$.
For fixed $c \in]0,1[$ we define a distance d on pairs $(x,y) \in \mathcal{S}_1$ (observe that this in particular implies $n_1(x) = n_1(y)$) as follows

$$d(x,y) = \prod_{j=1}^{cn-n_1(x)} (1 + j/n) .$$

Since d actually depends on the number of 1's and 3's only we may write $d(x,y) = d(n_1(x))$.

Let us now see that the coupling $\widetilde{Q}$ contracts the function d. Observe that for any value of $i := n_1(x) \geq 1$

$$\begin{aligned} \left[\widetilde{Q}d\right](x,y) &= \frac{(i-1)^2}{n(n-1)} d(i-1) + \frac{i(n-i)}{n(n-1)} d(i+1) \\ &\quad + \left[1 - \frac{i(i-1)}{n(n-1)} - \frac{i(n-i)}{n(n-1)}\right] d(i). \end{aligned}$$

Note that we have used that with probability $(i-1)/[n(n-1)]$ the pair x, y hits the diagonal and d is zero in this case. Multiplying by $n(n-1)/d(i)$ and subtracting 1 gives

$$\begin{aligned} \frac{n(n-1)}{d(i)} \left(\left[\widetilde{Q}d\right](x,y) - 1 \right) &= (i-1)^2 \frac{d(i-1)}{d(i)} + i(n-i) \frac{d(i+1)}{d(i)} \\ &\quad + \left[-i(i-1) - i(n-i)\right] . \end{aligned}$$

Now we first consider the case $2 \leq n_1(x) \leq cn - 1$. Note that in this case

$$\frac{d(i-1)}{d(i)} = 1 + (cn - i + 1)/n$$

and

$$\frac{d(i+1)}{d(i)} = 1/[1 + (cn-i)/n].$$

Setting $c = 1/10$ we arrive at

$$\frac{n(n-1)}{d(i)} \left(\left[\widetilde{Q}d\right](x,y) - 1 \right) =$$

$$\frac{1}{10n(11n - 10i)} \left(121 i^2 n^2 - 220 i^3 n + 100 i^4 - 10 i n^3 + 450 i^2 n - 300 i^3 - 132 i n^2\right) + \mathcal{O}(1)$$

Consider the function

$$f := i \mapsto 121 i^2 n^2 - 220 i^3 n + 100 i^4 - 10 i n^3 + 450 i^2 n - 300 i^3 - 132 i n^2.$$

Calculation of f'' and obvious estimates show that f is a convex function on the interval $[2, n/10]$. Thus f is maximal at one of the endpoints of this interval. Using that $f(n/10 - 1) = -17n^3 + \mathcal{O}(n^2)$, $f(2) = -20n^3 + \mathcal{O}(n^2)$ and $11n - 10i \geq 10n$ for $2 \leq i \leq n/10$, this yields

$$\frac{n(n-1)}{d(i)} \left(\left[\widetilde{Q}d\right](x,y) - 1 \right) \leq -\frac{17}{100} n + \mathcal{O}(1)$$

implying that $\lambda_1 \geq \frac{17}{100}\frac{1}{n} + \mathcal{O}(n^{-2})$.

Let us now turn to the cases when $n_1(x)$ is either 1 or $n/10$: If $n_1(x) = 1$, a straightforward calculation following the above lines shows that

$$\frac{n(n-1)}{d(1)}\left(\left[\tilde{Q}d\right](x,y) - 1\right) \quad \leq \quad -\frac{1}{11}n + \mathcal{O}(1).$$

For $n_1(x) = n/10$ we check in the same way that

$$\frac{n(n-1)}{d(n/10)}\left(\left[\tilde{Q}d\right](x,y) - 1\right) \quad \leq \quad -\frac{9}{100}n + \mathcal{O}(1).$$

If finally $n_1(x) \geq n/10 + 1$, d is constant and therefore the estimates reduces to

$$\frac{n(n-1)}{d(i)}\left(\left[\tilde{Q}d\right](x,y) - 1\right) \quad = \quad 1 - \frac{i-1}{n(n-1)} \leq 1 - \frac{1}{10n} + \mathcal{O}(n^{-2}).$$

The considerations show that in the situation described above the spectral gap $\lambda_1(Q)$ is bounded from below by $\frac{1}{11n}$.

Next we will treat the case where for each eigenfunction f to β_1 there exist $x, y \in X$ with $n(x) = n(y)$, $f(x) \neq n(y)$, and x and y contain the same number of 1's. Obviously, then y is nothing but a permutation of x. As we decompose any permutation into a number of transpositions, there also have to be x and y with the above properties such that x differs from y by a transposition of two coordinates (without loss of generality the first two), only. Otherwise f would be constant on all transpositions and hence also on all permutations in contrast to what we assumed above.

We distinguish four different cases

Case I: $x = (03*)$, $y = (30*)$
Case II: $x = (31*)$, $y = (13*)$
Case III: $x = (02*)$, $y = (20*)$
Case IV: $x = (12*)$, $y = (21*)$.

Here $*$ indicates that the configurations x and y agree in the third to n'th coordinate. The other two cases $x = (23*)$, $y = (32*)$, and $x = (01*)$, $y = (10*)$ do not need special consideration here, since the rôles of 1 and 3 are completely identical.

First observe that Case II is already covered by the considerations at the beginning of the treatment of Part B. By the arguments given there we might assume that any eigenfunction f to β_1 is constant on $\mathcal{S}_1$ (as defined in (8)).
But then, if $x = (31*)$ and $y = (13*)$, we obtain

$$f(x) - f(y) = (f(x) - f(\tilde{x})) + (f(\tilde{x}) - f(y))$$

where $\tilde{x} = (11*)$. Now $(x, \tilde{x}) \in \mathcal{S}_1$, as well as $(\tilde{x}, y) \in \mathcal{S}_1$ implying that

$$f(x) - f(y) = 0.$$

Hence f must be constant in Case II.

Let us thus turn to the remaining Cases I, III, and IV and put $x = (x_1x_2x_3\ldots x_n)$, and $y = (y_1y_2y_3\ldots y_n)$ where x_1, x_2, y_1, and y_2 are chosen according to the case we wish to consider and in each of these cases $x_i = y_i$ for $i \geq 3$.
The distance for x, y such that $n(x) = (i,j)$, $n(y) = (i,j)$ is defined by

$$d(x,y) = \begin{cases} \frac{1}{i} & : \quad \text{for } x, y \text{ as in Cases I and IV} \\ \frac{1}{i} + \frac{2}{3n} & : \quad \text{for } x, y \text{ as in Case III.} \end{cases}$$

The coupling that we use is very natural. If the walk associated with x proposes to add x_i^s, $s = \pm 1$ to x_j, and $i, j \geq 3$, the walk associated with y will add y_i^s, $s = \pm 1$ to y_j. If the walk associated with x proposes to add x_1^s, $s = \pm 1$ to x_j, $j \geq 2$, the walk associated with y will y_2^s, $s = \pm 1$ to y_j and, correspondingly, if the walk associated with x proposes to add x_2^s, $s = \pm 1$ to x_j, $j \geq 2$, the walk associated with y will y_1^s, $s = \pm 1$ to y_j. Finally, when the walk associated with x proposes to add x_i^s, $s = \pm 1$ to x_1, $i \geq 2$, the walk associated with y will y_i^s, $s = \pm 1$ to y_2 and vice versa, when the walk associated with x proposes to add x_i^s, $s = \pm 1$ to x_2, $i \geq 2$, the walk associated with y will y_i^s, $s = \pm 1$ to y_1.
Let us call the kernel associated with this coupling $\widetilde{Q}$.
Note that under $\widetilde{Q}$ a configuration starting in one of the Cases I-IV might enter another one of these cases. However also notice that $\widetilde{Q}$ keeps the Hamming distance of x and y at the value 2 until the two particles meet.

For all three cases we need to show that $\widetilde{Q}$ contracts d, that is

$$[\widetilde{Q}d](x,y) \leq (1-\alpha)d(x,y),$$

for all x, y as in Cases I, III, or IV and some $\alpha > 0$. This, obviously, is equivalent to

$$\left[[\widetilde{Q}d](x,y) - d(x,y)\right] / d(x,y) \leq -\alpha.$$

We define $d_1(i,j) = 1/i$ and $d_2(i,j) = 1/i + 2/(3n)$. Thus for x and y such that $n(x) = n(y) = (i,j)$ we have

$$d(x,y) = \begin{cases} d_1(i,j) & : \text{ for } x,y \text{ as in Cases I and IV} \\ d_2(i,j) & : \text{ for } x,y \text{ as in Case III.} \end{cases}$$

Note that in either case $d(x,y)$ only depends on $i := n_1(x)$ and $j := n_2(x)$. We set

$$f(i,j) := 2n(n-1)\left[[\widetilde{Q}d](x,y) - d(x,y)\right] / d(x,y).$$

Let us start with considering
Case IV) $x = (12...)$, $y = (21...)$.
The term of interest is given by

$$\begin{aligned} f(i,j) \quad := \quad & \Big[(n-i-j)id_1(i+1,j) + (n-i-j)jd_1(i,j+1) + (i-1)^2/2d_1(i-1,j) \\ & +(i-1)/2d_2(i-1,j) + j(j-1)d_1(i,j-1) + i(j-1)d_1(i+1,j-1) \\ & (i-1)^2/2d_1(i-1,j+1) + ijd_1(i,j) - (i+j)(n-1)d_1(i,j)\Big] / d_1(i,j) \qquad (9) \\ = \quad & \frac{1}{6}\frac{(2i^2+3in-6n^2-2+3n)i}{n(i+1)} \\ \leq \quad & \frac{-n}{6}\frac{i}{i+1} \leq -\frac{n}{12}. \end{aligned}$$

Let us explain the summands in (9) in greater detail. The first summand arises, when 1 or 3 is added to a zero component. This happens with probability $i(n-i-j)/(2n(n-1))$. The second summand stems from situations where 2 is added to a zero component. For the third summand 1 or 3 is added to a 1 or 3 (except the component x_1 resp. y_2) such that the new component is zero.
For the fourth summand a 1 or 3 has been added to the first component x_1 resp. y_2

such that we get to Case III.
The fifth summand results form adding a 2 to another 2.
The sixth summand occurs when 1 is added to a component 2 except the component $x_2 = 2$, $y_1 = 2$ (if we add 1 to the component $x_2 = 2$, $y_1 = 2$ we are done by the considerations at the beginning of Part B).
The seventh summand arises, when 1 or 3 is added to a 1 or 3 (except the component $x_1 = 2$, $y_2 = 1$) to give a 2 component.
The eighth summand shows up when 2 is added to a 1 or 3.
And finally the ninth summand results from the transformation shown above.

Case I) $x = (01...)$, $y = (10...)$.
By the same arguments as above the term of interest is given by

$$\begin{aligned} f(i,j) &:= \Big[(n-i-j-1)id_1(i+1,j) + (n-i-j)jd_1(i,j+1) + (i-1)^2/2d_1(i-1,j) \\ &\quad +j(j-1)d(i,j-1) + ijd(i+1,j-1) \\ &\quad (i-1)^2/2d_1(i-1,j+1) + (i-1)/2d_2(i-1,j+1) \\ &\quad +ijd_1(i,j) - (i+j)(n-1)d_1(i,j)\Big]/d_1(i,j) \\ &= \frac{1}{6}\frac{(2i^2+3in-6n^2-2+3n)i}{n(i+1)} \\ (10) &\leq \frac{-n}{6}\frac{i}{i+1} \leq -\frac{n}{12}. \end{aligned}$$

In Case I we get exactly the same term as in Case IV.

Case III) $x = (02...)$, $y = (20...)$.
Here the term of interest is given by

$$\begin{aligned} f(i,j) &:= \Big[(n-i-j-1)id_2(i+1,j) + id_1(i+1,j) \\ &\quad +(n-i-j-1)jd_2(i,j+1) + i(i-1)/2d_2(i-1,j) \\ &\quad +(j-1)^2d_2(i,j-1) + i(j-1)d_2(i+1,j-1) \\ &\quad +id_1(i+1,j-1) + +i(i-1)/2d_2(i-1,j+1) + ijd_2(i,j) \\ &\quad -(i+j)(n-1)d_2(i,j)\Big]/d_2(i,j) \\ &= -\frac{4i^3-6i^2n+4i^2j+2i^2+3in^2+6ijn-2i-6in+4ij-3n+6jn}{(i+1)(3n+2i)}. \end{aligned}$$

We show that the function

$$h(i,j) := \frac{4i^3-6i^2n+4i^2j+3in^2+6ijn}{(i+1)(3n+2i)}$$

is positive of order n for $1 \leq i \leq n$, $0 \leq j \leq n$. More precisely assume that we would know that $h(i,j) \geq cn$ where c is a positive constant. Then we obtain immediately

$$f(i,j) \leq -cn + \mathcal{O}(1).$$

Now, the derivate of the function $j \mapsto h(i,j)$, i fixed, is given by $2i/(i+1)$. Thus $j \mapsto h(i,j)$ is increasing in j for each i. Thus for each i, $h(i,j)$ is smallest for $j = 1$.

We therefore arrive at

$$h(i,j) \geq h(i,1) = \frac{i(4i^2 - 6in + 4i + 6n + 3n^2)}{(i+1)(3n+2i)}.$$

Now for $i = 1$

$$h(1,1) = \frac{3n^2}{2(3n+2)} \geq \frac{3}{10}n.$$

For $i \geq 2$ we may bound $i/(i+1)$ by $2/3 \leq i/(i+1) \leq 1$ in our further considerations. The function

$$i \mapsto \frac{4i^2 - 6in + 4i + 6n + 3n^2}{3n + 2i}$$

is minimal for $i_0 = -3/2n + 1/2\sqrt{21}n$. Hence we obtain

$$\begin{aligned} h(i_0, 1) &\geq -\frac{2}{3}\frac{n(15\sqrt{21}n - 69n - 2\sqrt{21} + 6)}{-3n + \sqrt{21}n + 2} \\ &\approx -n\frac{2}{3}\frac{-0,2614n - 2\sqrt{21} + 6}{1,5826n + 2}\frac{1}{10}n + \mathcal{O}(1). \end{aligned}$$

Thus also in Case III

$$f(i,j) \leq -cn$$

with $c = \frac{1}{10} + \mathcal{O}(n^{-1})$.

Therefore, in all Cases I-IV the second smallest eigenvalue λ_1 of our kernel $L = I - Q$ is bounded by

$$\lambda_1 \geq \frac{1}{12n} + \mathcal{O}(n^{-2}).$$

Let us see, what we have obtained by now. In Part A of this proof we have shown that the lower bound of our main Theorem holds true in case that every eigenfunction to the second largest eigenvalue β_1 is constant on the level sets of the signature $n(\cdot)$. In Part B we thus only need to consider the case where there is an eigenfunction f to β_1 and there are x and y with $n(x) = n(y)$ but $f(x) \neq f(y)$. At beginning of Part B we saw that we may assume that $f(x) = f(y)$ if $(x,y) \in \mathcal{S}_1$ (as defined in 8), otherwise we would be done. Moreover, in the four cases above we saw that we also may assume that $f(x) = f(\pi(x))$, for every permutation $\pi \in S_n$ (the symmetric group on n elements).

Let us now eventually turn to the most general case of Part B, that is there are x and y with $n(x) = n(y)$ but $f(x) \neq f(y)$. By the cases considered above we may assume that x and y contain a different number of 1's (otherwise x could be obtained form y by a permutation), say x contains more 1's than y. But then there also exists a pair $(x, \tilde{y})$ with $n(x) = n(\tilde{y})$ and $f(x) \neq f(\tilde{y})$ such that x contains exactly one more 1 than $\tilde{y}$. This is true, since we may assume that f is constant on $\mathcal{S}_1$ and we can reach such a $\tilde{y}$ from y via "a path" in $\mathcal{S}_1$ by successively flipping 3's into 1's. But then there also exists a pair $(x, \hat{y}) \in \mathcal{S}_1$ with $f(x) \neq f(\hat{y})$ because we can reach such a $\hat{y}$ from $\tilde{y}$ by permutation. On the other hand such a pair $(x, \hat{y}) \in \mathcal{S}_1$ cannot exist as we have shown that we may assume that f is constant on $\mathcal{S}_1$.
This shows that the cases treated above already cover the full regime of Part B and finishes the proof of the lower bound in Theorem 2.1.

For the upper bound in Theorem 2.1 note that the spectrum of an induced Markov chain always is contained in the spectrum of the original Markov chain. In particular, the spectrum of our chain Q is contained in the spectrum of the chain M as defined in (4) and (5) with

$$\begin{aligned} M(i,i+1) &= \frac{(n-i)i}{n(n-1)} \\ M(i,i-1) &= \frac{i(i-1)}{2n(n-1)}. \end{aligned}$$

Let π' denote the invariant measure of M and take any function

$$f:\{1,\dots,n\}\to\mathbb{R}$$

with $f(i)=0$ for all $i\in\{3,\dots,n\}$ and $E_{\pi'}(f)=0$. Using $(a-b)^2\le 2(a^2+b^2)$ we conclude for the Dirichlet form $\mathcal{E}(f,f)$ with respect to M and π'

$$\mathcal{E}(f,f)=\frac{1}{2}\sum_{1\le i,j\le n}(f(i)-f(j))^2\pi'(i)M(i,j)\le\frac{4(n-2)}{n(n-1)}||f||_2^2$$

where $||\cdot||_2$ denotes the L^2-norm with respect to π'. Since the second smallest eigenvalue λ_1' of $Id-M$ has a representation as

$$\lambda_1'=\inf\left\{\frac{\mathcal{E}(f,f)}{||f||_2^2},E_{\pi'}(f)=0,f\not\equiv 0\right\}$$

we arrive at

$$\lambda_1'\le\frac{4(n-2)}{n(n-1)}.$$

This proves the upper bound in Theorem 2.1 and therefore finishes the proof. □

References

[CGMNO95] F. Celler, C. Leedham-Green, S. Murray, A. Niemeyer, E. O'Brien, *Generating random elements of a finite group*, Communications in Algebra No. 23, 4931–4948 (1995).

[CW94] M.F. Chen, F.Y. Wang, *Applications of coupling method to the first eigenvalue on manifold*, Science in China, Series A **37** No. 1, 1–14.

[Ch96] M.F. Chen, *Estimation of Spectral Gap for Markov Chains*, Acta Mathematica Sinica, New Series **12** No. 4, 337–360 (1996).

[CG97] F.R.K. Chung, R.L. Graham, *Stratified Random Walks on the n-Cube*, Random Struct. Alg., **11**, 199–222 (1997).

[DS91] P. Diaconis, D. Stroock, *Geometric bounds for eigenvalues of Markov chains*, The Annals of Appl. Prob., No. 1, 36–61 (1991).

[DS97] P. Diaconis, L. Saloff-Coste, *Walks on generating sets of groups*, Preprint (1997).

[DS96] P. Diaconis, L. Saloff-Coste, *Walks on generating sets of Abelian groups*, Probab. Theory Rel. Fields 105, 393–421 (1996).

[HJ85] R. Horn, C. Johnson, *Matrix Analysis*, Cambridge University Press (1985).

[Me98] C. Meise, *On Spectral Gap estimates via distance arguments*, submitted.

[Me99] C. Meise, *On Spectral Gap estimates of a Markov chain via hitting times and coupling*, to appear: J. of Appl. Prob. (1999).

[Mi99] L. Miclo, *An example Of Applicaton of Discrete Hardy's Type Inequalities*, Markov Proc. Related Fields **5**, 319-330 (1999).

[Ze91] A.I. Zeifman, *Some estimates of the rate of convergence for birth and death processes*, J. Appl. Prob. **28**, 268–277 (1991).

[Ze98] B.L. Granovsky, A.I. Zeifman, *The N-limit of spectral gap of random walks associated with complete graphs*, Preprint (1998).

(Matthias Löwe) EURANDOM, P.O. Box 513, 5600 MB Eindhoven, Netherlands
E-mail address, Matthias Löwe: `lowe@eurandom.tue.nl`

(Christian Meise) Fakultät für Mathematik, Universität Bielefeld, Postfach 100 131, D-33501 Bielefeld, Germany
E-mail address, Christian Meise: `meise@mathematik.uni-bielefeld.de`