Lower Bounds for the Spectral Gap of Markov chains

Christian Meise
Universität Bielefeld

Abstract. Well-known inequalities for the spectral gap of a Markov chain, like Poincaré's and Cheeger's Inequality, do not perform well if the transition graph of the Markov chain is strongly connected. Using the coupling methods we derive a lower bound for the spectral gap of a discrete-time Markov chain. Another lower bound is derived via hitting times. Finally the bounds are discussed using nearest-neighbor random walks on the circle, star and d-dimensional cube, for which the inequalities mentioned above can be considerably improved.

Keywords: Coupling, Markov chain, Spectral Gap

1 Introduction and Results

Let $(M, \mathcal{M})$ be a measurable space. We will denote by $\mathcal{M}_b$ the set of $\mathcal{M}$-measurable and bounded functions $f : M \to \mathbb{R}$. Any transition kernel Q from $(M, \mathcal{M})$ to a measurable space $(N, \mathcal{N})$ determines a linear operator from $\mathcal{M}_b$ to $\mathcal{N}_b$ by the definition

$$[Qf](x) = \int_N f(y) Q(x, dy), \; x \in M. \tag{1}$$

Moreover we have $\|Qf\| \leq \|f\|$, where $\|f\| = \sup\{\, f(x) \mid x \in M \,\}$ denotes the supremum norm.
In our applications we choose $M = N$ and M to be a finite set. If Q is time-reversible with respect to the probability measure π the operator $[Q\cdot]$ defined in (1) is self-adjoint, so that all the eigenvalues of Q are real. By the assumption of irreducibility the constant eigenfunctions are the only eigenfunctions for the eigenvalue 1 (compare [SE]). Therefore we can enumerate the eigenvalues in the form

$$1 = \beta_0 > \beta_1 \geq \beta_2 \geq \ldots \geq \beta_{|M|-1} \geq -1.$$

The last inequality is strict if and only if the chain Q is aperiodic.

Let $\Omega = S^{\mathbb{N}}$, where $(S, \mathcal{S})$ is a measurable space and denote by $Y_n(\omega) = \omega_n$ the coordinate process. We assume that there exists a stationary measure $\mu \neq 0$ for the transition kernel Q. Recall that a stationary measure for countable infinite S in general does not exist any more, (compare page 54 in [JA86]). Denote by $T_y = \inf\{\, n \geq 0 \,|\, Y_n = y \,\}$ the first hitting time of the point y. As a lower bound for the spectral gap we get

Proposition 1.1 *(Hitting Bound) Let $f \in \mathcal{S}_b$ be an eigenvector for $\beta \neq 1$ of the transition kernel Q on $(S, \mathcal{S})$ and suppose that $E_x(T_y) < \infty$ for all $x, y \in S$. Then*

$$|1 - \beta| \geq \frac{1}{\sup_{x,y \in S} E_x(T_y)}. \tag{2}$$

As we will see in the applications the hitting bound performs well for the nearest-neighbor random walk on the circle, while it is not precise for nearest-neighbor random walk on the star and the cube. The proof of Proposition 1.1 is based on a martingale argument, that is similar to that used in the proof the coupling bound.

Recall the definition of a coupling:

Definition 1.2 *(Coupling of probability measures).*
Let $(\Omega_i, \mathcal{A}_i, P_i)$ for $i \in \{1, 2\}$ be probability spaces. A probability measure $\widetilde{P}$ on

$$(\Omega, \mathcal{A}) := (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$$

is called a coupling of (P_1, P_2), if the following marginal conditions are satisfied:
i) $\widetilde{P}(A_1 \times \Omega_2) = P_1(A_1)$ *for all* $A_1 \in \mathcal{A}_1$
ii) $\widetilde{P}(\Omega_1 \times A_2) = P_2(A_2)$ *for all* $A_2 \in \mathcal{A}_2$.

Let $\Omega_1 = \Omega_2 = S^{\mathbb{N}}$ for $(S, \mathcal{S})$ be a measurable space and $Q : S \times \mathcal{S} \to [0, 1]$ be a transition kernel on S. The coupling measure is defined on $(\Omega, \mathcal{A}) := ((S^2)^{\mathbb{N}_0}, \otimes_{k=0}^{\infty} \mathcal{S}^2)$ and the coordinate process is denoted by $Y_n(\omega) = (Y_n^1(\omega), Y_n^2(\omega)) = (\omega_n^1, \omega_n^2)$, $\omega \in \Omega_1 \times \Omega_2$. Let T be the hitting time for the diagonal in $S \times S$, e.g.

$$T = \inf\{\ n \geq 0 \ \mid\ Y_n^1 = Y_n^2 \ \}.$$

In the sequel we will be interested in measures $\widetilde{P}$ on the sequence space $(\Omega, \mathcal{A})$, such that a given sequence $(\widetilde{Z}_n)_n$ of random variables is a martingale. For special coupling measures this martingale property holds. In order to be able to specify these special coupling measures we introduce

Definition 1.3 *(Coupling of transition kernels). Let*

$$Q_i : \Omega_i \times \mathcal{A}_i \to [0, 1] \quad i \in \{1, 2\}$$

be transition kernels. A transition kernel

$$\tilde{Q} : (\Omega_1 \times \Omega_2) \times (\mathcal{A}_1 \otimes \mathcal{A}_2) \to [0,1]$$

is called a coupling of (Q_1, Q_2)*, if*

i) $\tilde{Q}((\omega_1,\omega_2), A_1 \times \Omega_2) = Q_1(\omega_1, A_1)$ *for all* $(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2,\ A_1 \in \mathcal{A}_1$

ii) $\tilde{Q}((\omega_1,\omega_2), \Omega_1 \times A_2) = Q_2(\omega_2, A_2)$ *for all* $(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2,\ A_2 \in \mathcal{A}_2$.

For any eigenvector f to the eigenvalue $\beta \neq 1$ of the transition kernel Q we define $g : S \times S \to \mathbb{R}$ by $g(y_1, y_2) := f(y_1) - f(y_2),\ y_1, y_2 \in S$.
Set

$$\tilde{Z}_n = g(Y_n) + (1-\beta)\sum_{k=0}^{n-1} g(Y_k) \tag{3}$$

and introduce the following spaces of probability measures

$$\tilde{\mathcal{P}} = \tilde{\mathcal{P}}(s_0) := \left\{ \tilde{P} \mid \tilde{Z}_0 = g(s_0)\ \tilde{P} - a.s.,\ (\tilde{Z}_n)_n \text{ is a martingale under } \tilde{P} \right\}$$

and

$$\tilde{\mathcal{P}}_{coupling}(s_0) := \left\{ \tilde{P} \mid \tilde{P} \text{ is coupling of } (P_{s_0^1}, P_{s_0^2}) \text{ that is induced by coupling kernels} \right\}.$$

The latter space is the space of *special couplings* mentioned above, e.g. we will show that

$$\emptyset \neq \tilde{\mathcal{P}}_{coupling}(s_0) \subset \tilde{\mathcal{P}}(s_0).$$

Example 3.6 illustrates that $(\tilde{Z}_n)_n$ is generally not a martingale for arbitrary coupling measures of $(P_{s_0^1}, P_{s_0^2})$. Besides there exist non-coupling measures such that $(\tilde{Z}_n)_n$ is a martingale. Compare example 3.7.
Using the martingale property of $(\tilde{Z}_n)_n$ for measures $\tilde{P} \in \tilde{\mathcal{P}}$, we can show

Theorem 1.4 *Let Q be a transition kernel on $(S, \mathcal{S})$ and $f \in \mathcal{S}_b$ be an eigenvector for the eigenvalue $\beta \neq 1$ of Q.
Assume that $s_0 = (s_0^1, s_0^2) \in S \times S$ is chosen in such a way, that*

$$f(s_0^1) - f(s_0^2) = \sup\{ f(x) - f(y) \mid x, y \in S \}, \tag{4}$$

then we have

$$|1-\beta| \geq \frac{1}{\inf_{\tilde{P} \in \tilde{\mathcal{P}}(s_0)} E_{\tilde{P}}(T)}. \tag{5}$$

Theorem 1.4 implies the following Theorem, but Theorem 1.5 is also true if the assumption (4) does not hold.

Theorem 1.5 *(Coupling Bound) Let Q be a transition kernel on $(S,\mathcal{S})$ and $f \in \mathcal{S}_b$ be an eigenvector to the eigenvalue $\beta \neq 1$. Then*

$$|1-\beta| \geq \inf_{\bar{s}\in S\times S} \frac{1}{\inf_{\widetilde{P}\in\widetilde{\mathcal{P}}(\bar{s})} E_{\widetilde{P}}(T)}. \tag{6}$$

In [CHWA92] (compare also [CH94]) Chen and Wang give a lower bound for the spectral gap of a reversible time-continuous Markov process $(X_t)_t$. More precisely they show that under certain assumptions $gap(\Omega) \geq \left(\max_{x_1\neq x_2} E_{x_1,x_2}(T)\right)^{-1}$. Here Ω is the operator on $\mathcal{S}_b$ given by $\Omega f(x) = \int_S (f(y)-f(x))q(x,dy)$ and q is the Q-matrix of $(X_t)_t$. Therefore we can say that Theorem 1.5 is the time-discrete version of the result from Chen and Wang.

2 Applications

In the applications we choose coupling kernels such as

1) Classical Coupling. Let X be a discrete set and Q a transition kernel on X. Then the classical coupling $\widetilde{Q}_c$ is given by

$$\begin{aligned} \widetilde{Q}_c((x_1,x_2),(y_1,y_2)) &:= Q(x_1,y_1)\cdot Q(x_2,y_2) \quad \text{for } y_1\neq y_2 \\ \widetilde{Q}_c((x,x),(y,y)) &:= Q(x,y). \end{aligned}$$

The kernel $\widetilde{Q}_c$ describes a random walk of two particles each of them marginally behaving according to Q. After the first meeting of two the particles they only move together through the state space. Doeblin used this coupling in a 1938 paper [D38] for proving the Ergodic Theorem for regular Markov chains with finite state space.

2) Wasserstein's Coupling.

$$\begin{aligned} \widetilde{Q}((x_1,x_2),(y,y)) &:= Q(x_1,y)\wedge Q(x_2,y) \\ \widetilde{Q}((x_1,x_2),(y_1,y_2)) &:= \frac{(Q(x_1,y_1)-Q(x_2,y_1))^+\,(Q(x_2,y_2)-Q(x_1,y_2))^+}{1-\sum_{z\in X} Q(x_1,z)\wedge Q(x_2,z)} \\ &\quad\ \text{for } x_1\neq x_2,\ y_1\neq y_2. \end{aligned}$$

In the case of nearest-neighbor random walk on the star the classical coupling and Wasserstein's coupling coincide.
Before discussing applications, we show that any sequence of coupling kernels kernels induce a coupling of probability measures on the corresponding sequence space. The converse is false as is illustrated by example 3.6.

Proposition 2.1 *Let $(\Omega_n^1, \mathcal{A}_n^1)$ and $(\Omega_n^2, \mathcal{A}_n^2)$ for $n \geq 0$ be measurable spaces. Furthermore we set*

$$\Omega \; := \; \prod_{n=0}^{\infty} \left(\Omega_n^1 \times \Omega_n^2\right).$$

and we assume that

$$\widetilde{Q}_n : (\Omega_0 \times \ldots \times \Omega_n) \times \left(\mathcal{A}_n^1 \otimes \mathcal{A}_n^2\right) \to [0,1]$$

is a coupling of the kernels

$$Q_n : \left(\Omega_0^1 \times \ldots \times \Omega_n^1\right) \times \mathcal{A}_{n+1} \to [0,1]$$
$$Q_n' : \left(\Omega_0^2 \times \ldots \times \Omega_n^2\right) \times \mathcal{A}_{n+1}' \to [0,1].$$

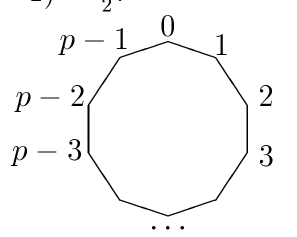
Let $\widetilde{\mu}$ be any coupling of the initial distributions μ_1 resp. μ_2 given on Ω_0^1 resp. Ω_0^2.
Under these conditions $\widetilde{P}_{\widetilde{\mu}}$, which is uniquely determined on

$$\left(\Omega, \otimes_{n=0}^{\infty}(\mathcal{A}_n^1 \otimes \mathcal{A}_n^2)\right)$$

by the initial distribution $\widetilde{\mu}$ and the kernels $(\widetilde{Q}_n)_n$, is a coupling (as probability measure) of (P_{μ_1}, P_{μ_2}).

2.1 The Circle $\mathbf{Z}_p$

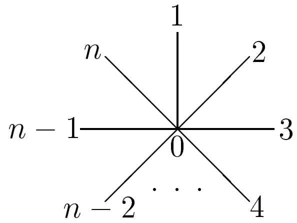
Let $X := \mathbf{Z}_p = \{0, \ldots, p-1\}$ denote the points of the circle where p is an odd number. The Markov chain is given by nearest-neighbor random walk with equal probability on the set of neighbor, e.g. for $a \in \mathbf{Z}_p$ the transition matrix is given by $Q(a, a+1) = \frac{1}{2}, \quad Q(a, a-1) = \frac{1}{2}$.



The hitting bound gives $1 - \beta_1 \geq \frac{4}{(p-1)(p+1)}$. Compared with the exact value $1 - \beta_1 = \frac{2\pi^2}{p^2} + O(p^{-4})$ the bound gives the right order. Using the classical coupling Theorem 1.5 yields $1 - \beta_1 \geq \frac{2}{p-1} = \frac{p-3}{p-1}$, which contains no information for large p. The Poincaré Inequality gives $1 - \beta_1 \geq \frac{8p}{(p-1)^2(p+1)}$, which is of the right order, compare [DS91]. The Cheeger Inequality gives also the right order $1 - \beta_1 \geq \frac{2p^2}{(p^2-1)^2}$, compare [DS91]. The hitting bound is slightly better than the Cheeger Inequality and slightly worse than the Poincaré Inequality.
In the next example we will see that the coupling bound can be better than the Cheeger Inequality.

2.2 The Star

For this purpose we take the star with $n+1$ points, e.g. the state space is given by $X := \{0, \ldots, n\}$ and every point $k \geq 1$ is only connected with the interior point 0.



The transition probabilities are given by the matrix Q:

$$Q = \begin{pmatrix} \theta & \frac{1-\theta}{n} & \cdots & \frac{1-\theta}{n} \\ 1 & 0 & & \\ \vdots & & \ddots & \\ 1 & & & 0 \end{pmatrix}$$

where $\theta \in]0, 1[$ denotes the one-step probability of staying at 0.

The matrix Q has the eigenvalues

$$\beta_0 := 1 \quad \beta_1 := 0 \quad \beta_2 := \theta - 1.$$

The eigenvalue β_1 is of multiplicity $(n-1)$.
This can be verified in the following way:
First recall that only the constant vectors are eigenvectors for the eigenvalue 1. The (non-constant) vector

$$\begin{pmatrix} 1 \\ \frac{1}{\theta-1} \\ \vdots \\ \frac{1}{\theta-1} \end{pmatrix}$$

is an eigenvector for the eigenvalue $\beta_2 = \theta - 1$. Therefore the set of $(n-1)$ linear independent vectors $v^i \in \mathbb{R}^{n+1}, \quad i \in \{1, \ldots, n-1\}$ given by

$$v^i(1) := 1 \quad v^i(i) := -1 \quad v^i(j) := 0 \quad \text{for } j \notin \{1, i\}$$

is a basis of the eigenspace corresponding to the eigenvalue 0. $\Box$
The hitting bound yields $1 - \beta_1 \geq \frac{1-\theta}{(n+1)(1-\theta)+n-1}$, which is uninformative for small θ and large n. Again using the classical coupling $\widetilde{Q}_c$ application of the coupling bound gives $1 - \beta_1 \geq \theta$.

Taking the vector v^1 as eigenvector for β_1, we see that the function $g(x,y) = f(x) - f(y) = v^1(x) - v^1(y)$ attains its supremum for $x = 1$ and $y = 2$. Therefore we can choose the starting point $s_0 := (1,2)$. Applying Theorem 1.4 we get:

$$1 - \beta_1 \geq \frac{1}{E_{s_0,\widetilde{Q}_c}(T)} = 1.$$

Therefore the estimate of Theorem 1.4 is sharp. The Poincaré Inequality gives

$$1 - \beta_1 \geq \frac{2-\theta}{3 - 2\theta - 2\frac{1-\theta}{n}}.$$

Since the right side is always greater than θ the coupling bound is never better than the Poincaré Inequality. The Cheeger Inequality gives the bound $1 - \beta_1 \geq \frac{1}{2}$. Therefore for $\theta > \frac{1}{2}$ the Cheeger bound is worse than the coupling bound.

2.3 The Cube $\mathbf{Z}_2^d$

Calculating the mean values E_k of the hitting time to $(1,\ldots,1)$ starting in any point with k zeroes shows that

$$E_k = \sum_{j=0}^{k-1} \frac{j!}{(d-1)(d-2)\ldots(d-j)} \left(\binom{d}{j+1} + \binom{d}{j+2} + \ldots + \binom{d}{d} \right).$$

The hitting bound yields a non-interesting estimate $1 - \beta \geq \frac{1}{E_d}$ since $E_d = o(2^d)$. In order to apply the coupling bound we use a coupling which is presented by Aldous and Diaconis in [AD87]. The advantage of this coupling is, that it does not allow the two particles, randomly moving through the state space, to increase their distance to each other. For $\bar{y} = (y_1, y_2)$ let

$$V(\bar{y}) := \{ \, i \mid y_1(i) \neq y_2(i) \, \}$$

the set of components, at which y_1 and y_2 differ. Furthermore let $k(\bar{y}) := |V(\bar{y})|$ the Hamming distance of y_1 and y_2.
Introduce $I := \{1,\ldots,d\}$, the transitions are given by

$$\begin{aligned} &\widetilde{Q}(Y_{n+1} = (y_1 + e_i, y_2 + e_i) \mid Y_n = \bar{y}) = \tfrac{1}{d+1} && \text{for } i \in I \setminus V(y_1,y_2) \\ &\widetilde{Q}(Y_{n+1} = (y_1 + e_j, y_2 + e_i) \mid Y_n = \bar{y}) = \tfrac{1}{(d+1)k(\bar{y})} && \text{for } i,j \in \{0\} \cup V(\bar{y}) \text{ with } i \neq j. \end{aligned} \tag{7}$$

Using this coupling yields the bound

$$1 - \beta_1 \geq \frac{2}{(d+1)\,(\ln(d) + 1)}.$$

Comparing this estimate with the exact value $1 - \beta_1 = \frac{2}{d+1}$, we state a deviation of order $\ln(d)$.

A closer look to the eigenvectors allows us to take neighboring points such as $((0,\dots,0),(1,0,\dots,0)) =: s_0$ as starting points. Application of Proposition 1.4 yields

$$1-\beta_1 \geq \frac{2}{E_{s_0,\widetilde{Q}}(T)} = \frac{2}{E_1} = \frac{2}{d+1},$$

which is sharp.
The Poincaré Inequality gives

$$1-\beta_1 \geq \frac{2d2^{d-1}}{(d+1)^2 d2^{d-1}} = \frac{2}{(d+1)^2},$$

which is off by a factor of d.
The Cheeger Inequality gives $1-\beta_1 \geq \frac{1}{2(d+1)^2}$, so that this bound is also off by a factor of d, compare [DS91].

3 Proofs

Let $(S,\mathcal{S})$ be a measurable space and for $n \geq 0$ let $Q_n : S^{n+1} \times \mathcal{S} \to [0,1]$ be transition kernels, where $S^n := \underbrace{S \times \ldots \times S}_{n\,times}$ and $\mathcal{S}^n := \underbrace{\mathcal{S} \times \ldots \times \mathcal{S}}_{n\,times}$. Introduce the *discrete generator* $L_n := I_n - Q_n$. Here I_n denotes the kernel defined by

$$I_n(x_0,\dots,x_n,A) = I(x_n,A) = \begin{cases} 1 & : \quad x_n \in A \\ 0 & : \quad x_n \notin A. \end{cases}$$

Furthermore we denote by Y_n the coordinate process. Recall that for any initial distribution μ on $(S,\mathcal{S})$ there exists exactly one measure P_μ on $(\Omega, \otimes_{n=0}^{\infty}\mathcal{S})$, such that for any bounded $\mathcal{S}^{n+1}$-measurable function $f : S^{n+1} \to \mathbb{R}$

$$\int_\Omega f \circ (Y_0,\dots,Y_n)\, dP_\mu = \int_{S^{n+1}} f\, dP_\mu^{(n)},$$

where $P_\mu^{(n)}$ is defined recursively:
For $n = 0$ set $P_\mu^{(0)} = \mu$. Given $P_\mu^{(n-1)}$ define $P_\mu^{(n)}$ on $(S^{n+1},\mathcal{S}^{n+1})$ by

$$P_\mu^{(n+1)}(A) = P_\mu^{(n)} \otimes Q_n(A) = \int_{S^n}\int_S 1_A(x,y) Q_n(x,dy) P_\mu^{(n)}(dx).$$

By the following Lemma any probability measure induced by an initial distribution and (not necessarily Markovian) transition kernels solves the *discrete martingale problem*.

Lemma 3.1 *For any $f \in \mathcal{S}_b$ and any initial distribution μ on $(S,\mathcal{S})$ the sequence $(Z_n)_{n\geq 0}$ defined by*

$$Z_n := f(Y_n) + \sum_{k=0}^{n-1} [L_k f](Y_0,\dots,Y_k) \tag{8}$$

is a martingale under P_μ with respect to the natural filtration $(\mathcal{F}_n)_{n\geq 0}$, e.g. $\mathcal{F}_n := \sigma(\, Y_k \mid 0 \leq k \leq n)$.

Proof: Since the function $(Y_0, \ldots, Y_k) : \Omega \to S^{k+1}$ is $\mathcal{F}_k$-measurable and $[L_k f] : S^{k+1} \to \mathbb{R}$ is $\mathcal{S}^{k+1}$-measurable the function $[L_k f](Y_0, \ldots, Y_k) : \Omega \to \mathbb{R}$ is (as composition) $\mathcal{F}_k \subset \mathcal{F}_n$-measurable. Therefore $(Z_n)_n$ is $(\mathcal{F}_n)_n$-adapted.
For the martingale property the main observation is

$$E(\, f(Y_{n+1}) \mid \mathcal{F}_n) \;=\; [Q_n f](Y_0, \ldots, Y_n). \tag{9}$$

In order to prove (9) let $A \in \mathcal{F}_n$. Since A is of the form $A = A_n \times S \times S \times \ldots$ with $A_n \in \otimes_{i=0}^n \mathcal{S}$, we conclude

$$\begin{aligned}
\int_A f(Y_0, \ldots, Y_{n+1}) dP_\mu &= \int_{A_n} \int_S f(y) Q_n(x, dy) P_\mu^{(n)}(dx) \\
&= \int_{A_n} [Q_n f](x) P_\mu^{(n)}(dx) \\
&= \int_A [Q_n f] dP_\mu.
\end{aligned}$$

Using $[L_n f](Y_0, \ldots, Y_n) = f(Y_n) - [P_n f](Y_0, \ldots, Y_n)$, we conclude

$$\begin{aligned}
E(\, Z_{n+1} - Z_n \mid \mathcal{F}_n) &= E(\, f(Y_{n+1}) - f(Y_n) + [L_n f](Y_0, \ldots, Y_n) \mid \mathcal{F}_n) \\
&= 0. \quad \square
\end{aligned}$$

3.1 Proof of the Hitting Bound

Taking $Q_k = Q$ for all k Lemma 3.1 shows that $(Z_n)_n$ given by

$$Z_n = f(Y_n) + \sum_{k=0}^{n-1} [(I - Q) f](Y_k)$$

is a martingale with respect to the natural filtration for any $f \in \mathcal{S}_b$.
Using the martingale property of $(Z_n)_n$ the following Lemma is derived:

Lemma 3.2 *Let $f \in \mathcal{S}_b$ be an eigenvector for $\beta \neq 1$ of Q and suppose that $E_x(T_y) < \infty$. Under these assumptions we have the integral representation*

$$f(y) = f(x) + E_x \left(\sum_{k=0}^{T_y - 1} [Lf](Y_k) \right). \tag{10}$$

Proof: By the martingale property of $(Z_n)_n$ under P_x we have

$$f(x) \quad = \quad E_x(f(Y_{\min(T_y,n)-1})) + (1-\beta)E_x\left(\sum_{k=0}^{\min(T_y,n)-1} f(Y_k)\right).$$

Let n tend to infinity and apply Lebesgue's convergence Theorem twice. This yields the assertion. $\square$

Assume that μ is a stationary measure of the kernel $Q : S \times \mathcal{S} \to [0,1]$. Let $f \in \mathcal{S}_b$ be an eigenvector for the eigenvalue $\beta \neq 1$ of Q.

By Lemma 3.1 $(Z_n)_{n\geq 0}$ is a martingale under P_μ. Notice that $E_{P_\mu}(f(Y_k)) = E_{P_\mu}(f(Y_0)) = E_\mu(f)$ is independent of k. Hence

$$E_\mu(f) = E_{P_\mu}(f(Z_0)) = E_{P_\mu}(Z_1) = (2-\beta)E_\mu(f).$$

Therefore $E_\mu(f) = 0$ for $\beta \neq 1$. Since $f \neq 0$ there exists $x \in S$ with $f(x) \neq 0$. Suppose that $f(x) > 0$. If there does not exist $y \in S$ with $f(y) \leq 0$, it necessarily follows that $\mu(\{f > 0\}) = 0$, which yields the contradiction $1 = \mu(\{f \leq 0\}) = \mu(\emptyset) = 0$.

This shows

Lemma 3.3 *Under the above assumptions: $E_\mu(f) = 0$ for $\beta \neq 1$. Moreover there exists $x, y \in S$ such that one of the inequalities $f(x) \geq 0$ and $f(y) \leq 0$ is strict.*

Now we are able to prove the Proposition 1.1

Proof of Proposition 1.1: Without loss of generality we may assume that f satisfies

$$\sup\{\, f(z) \mid z \in S\,\} = \sup\{\, |f(z)| \mid z \in S\,\} = 1.$$

Choose $x \in S$ such that $f(x) \geq 1 - \varepsilon$, $(1 > \varepsilon > 0$ arbitrary). By Lemma 3.3 we can choose $y \in S$ with $f(y) \leq 0$. By assumption we have $E_x(T_y) < \infty$, so that the integral representation (3.2) yields

$$f(x) = f(y) + (1-\beta)E_x\left(\sum_{k=0}^{T_y-1} f(Y_k)\right).$$

Since $f(Y_k) \leq 1$, we conclude

$$\begin{aligned} |1-\beta| \quad &\geq \quad \frac{f(x)-f(y)}{E_x(T_y)} \geq \frac{1-\varepsilon}{E_x(T_y)} \\ &\geq \quad \frac{1-\varepsilon}{\sup_{x,y\in S} E_x(T_y)}. \quad \square \end{aligned}$$

3.2 Proof of the Coupling Bound

Let $\widetilde{Q}$ be a coupling of (Q_1, Q_2) where Q_1 resp. Q_2 is a transition kernel from $(R_1, \mathcal{R}_1)$ to $(S_1, \mathcal{S}_1)$ resp. from $(R_2, \mathcal{R}_2)$ to $(S_2, \mathcal{S}_2)$. We denote by $\mathcal{S}_{i,b}$ ($i \in \{1, 2\}$) the set of real-valued, bounded and $\mathcal{S}_i$-measurable functions defined on S_i. We can regard $f \in \mathcal{S}_{i,b}$ as a function on $S_1 \times S_2$ by defining $f(s_1, s_2) := s_i$.

Lemma 3.4 *In the situation described above we have*

$$\begin{aligned} [\widetilde{Q}f](r_1, r_2) &= [Q_1 f](r_1) \quad \textit{for } r_1 \in R_1, r_2 \in R_2,\ f \in \mathcal{S}_{1,b} \\ [\widetilde{Q}f](r_1, r_2) &= [Q_2 f](r_2) \quad \textit{for } r_1 \in R_1, r_2 \in R_2,\ f \in \mathcal{S}_{2,b}. \end{aligned} \tag{11}$$

In particular if $f \in \mathcal{S}_{i,b}$ is an eigenfunction of Q_i then it can be regarded as an eigenfunction of $\widetilde{Q}$.

Proof: Without loss of generality just prove the first of the equations (11). At first this is done for indicator functions $f = 1_A$, $A \in \mathcal{S}_1$:

$$\begin{aligned} [\widetilde{Q}1_A](x_1, x_2) &= \int_{S_1 \times S_2} 1_A(y_1, y_2) \widetilde{Q}((x_1, x_2), d(y_1, y_2)) \\ &= \widetilde{Q}(x_1, x_2, A \times S_2) \\ &= Q_1(x_1, A) \\ &= [Q_1 1_A](x_1). \end{aligned}$$

Using linearity of the operators $[\widetilde{Q}\cdot]$ and $[Q_1\cdot]$, (11) is obtained for elementary functions. Applying the monotone convergence Theorem, we get (11) for non-negative or bounded functions. $\Box$

Let $\widetilde{Q}$ be a coupling kernel of Q and let $f \in \mathcal{S}_b$ denote an eigenvector for the eigenvalue β of Q. By Lemma 3.4 the function $g : S \times S \to \mathbb{R}$ given by $g(y_1, y_2) := f(y_1) - f(y_2),\ y_1, y_2 \in S$ is an eigenvector for the eigenvalue β of $\widetilde{Q}$. Note that $[(\widetilde{I} - \widetilde{Q})g] = (1 - \beta)g$. Using Lemma 3.1 yields that $(\widetilde{Z}_n)_n$ defined in (3) is a martingale under the induced measure $\widetilde{P}_{\widetilde{\mu},\widetilde{Q}}$, where $\widetilde{\mu}$ is arbitrary. Therefore any coupling kernel of Q yields a coupling measure contained in $\widetilde{\mathcal{P}} = \widetilde{\mathcal{P}}(s_0)$.

Proof of Theorem 1.4: Compare [CHWA92] or [CH94]. Let $\widetilde{P} \in \widetilde{\mathcal{P}}(s_0)$ be arbitrary. Without loss of generality we can assume, that $f(s_0^1) - f(s_0^2) = 1$. Since $(\widetilde{Z}_n)_n$ is a martingale, this also holds for the stopped sequence $(\widetilde{Z}_{\min(n,T)})_n$. Since $\widetilde{Z}_0 = g(s_0^1, s_0^2) = 1$ $\widetilde{P}$-almost surely we get

$$1 = E_{\widetilde{P}}(\widetilde{Z}_0) = E_{\widetilde{P}}\left(g(Y^1_{\min(n,T)}, Y^2_{\min(n,T)})\right) + (1 - \beta) E_{\widetilde{P}}\left(\sum_{m=0}^{\min(n,T)-1} g(Y^1_m, Y^2_m)\right)$$

Using absolute values on the right side, we get

$$\begin{aligned}1 \quad &\leq \quad E_{\widetilde{P}}\left(|g\left(Y^1_{\min(n,T)}, Y^2_{\min(n,T)}\right)|\right) + |1-\beta| E_{\widetilde{P}}\left(\sum_{m=0}^{\min(n,T)-1} \underbrace{|g(Y^1_m, Y^2_m)|}_{\leq 1}\right) \\ &\leq \quad E_{\widetilde{P}}\left(|f\left(Y^1_{\min(n,T)}\right) - f\left(Y^2_{\min(n,T)}\right)|\right) + |1-\beta| E_{\widetilde{P}}(\min(n,T))\end{aligned}$$

As n tends to infinity, we get from the monotone convergence theorem

$$E_{\widetilde{P}}\left(\min(n,T)\right) \uparrow E_{\widetilde{P}}(T)$$

and from Lebesgue's convergence Theorem

$$E_{\widetilde{P}}\left(|f\left(Y^1_{\min(n,T)}\right) - f\left(Y^2_{\min(n,T)}\right)|\right) \downarrow 0.$$

We may assume that $T < \infty$ $\widetilde{P}$-almost surely, because otherwise the inequality (5) is trivially true by using the convention $\frac{1}{\infty} = 0$.
Finally we have

$$1 \leq |1-\beta| E_{\widetilde{P}}\left(T\right)$$

and so

$$|1-\beta| \geq \frac{1}{E_{\widetilde{P}}\left(T\right)}.$$

Since $\widetilde{P} \in \widetilde{\mathcal{P}}(s_0)$ has been arbitrary the proof is complete. □

In order to profit from the estimate given in the last theorem one has to know for which points s^1_0, s^2_0 the supremum over the differences of eigenvector-components is achieved. Since we are interested in knowledge concerning the eigenvalue we are in general far from having this knowledge concerning the eigenvector. So for practical use we give the following rougher estimate

$$\frac{1}{\inf_{\widetilde{P} \in \widetilde{\mathcal{P}}(s_0)} E_{\widetilde{P}}(T)} \geq \inf_{\bar{x} \in X \times X} \frac{1}{\inf_{\widetilde{P} \in \widetilde{\mathcal{P}}(\bar{x})} E_{\widetilde{P}}\left(T\right)},$$

which yields Theorem 1.5.

Remark 3.5 *Notice that in Theorem 1.5 the assumption (4) is not necessary any more:*
For any $\varepsilon > 0$ choose $s_0 \in S \times S$ such that $f(s^1_0) - f(s^2_0) \geq 1 - \varepsilon$. Using the same conclusions as in the proof of Theorem 1.4, we derive

$$|1-\beta| \geq \frac{1-\varepsilon}{E_{\widetilde{P}}(T)} \geq \inf_{\bar{s} \in S \times S} \frac{1-\varepsilon}{\inf_{\widetilde{P} \in \widetilde{\mathcal{P}}(\bar{s})}(T)}.$$

This yields (6) since ε has been arbitrary.

The next example shows that the sequence $(\widetilde{Z}_n)$ is unfortunately not always a martingale.

Example 3.6 *Take $S = \{x, y\}$, and for $\alpha, \alpha' \in]0, 1[$ let P_x resp. P_y be the induced measure on $S^{\mathbb{N}_0}$ by the initial distribution ε_x resp. ε_y and the transition kernel*

$$Q = \begin{pmatrix} \alpha & 1-\alpha \\ 1-\alpha' & \alpha' \end{pmatrix}. \tag{12}$$

Let $\widetilde{Q}_1$ be the classical coupling of Q and for $c \in [0, 1[$ define $\widetilde{Q}_2$ such that

$$\begin{array}{ll} \widetilde{Q}_2((x,x),\{y\}\times S) = c & \widetilde{Q}_2((x,x),\{x\}\times S) = 1-c \\ \widetilde{Q}_2((x,y),\{y\}\times S) = 1 & \widetilde{Q}_2((x,y),\{x\}\times S) = 0 \\ \widetilde{Q}_2((y,x),\{y\}\times S) = 0 & \widetilde{Q}_2((y,x),\{x\}\times S) = 1 \\ \widetilde{Q}_2((y,y),\{y\}\times S) = 1 & \widetilde{Q}_2((y,y),\{x\}\times S) = 0 \end{array}$$

$$\begin{array}{ll} \widetilde{Q}_2((x,x),S\times\{x\}) = 1 & \widetilde{Q}_2((x,x),S\times\{y\}) = 0 \\ \widetilde{Q}_2((x,y),S\times\{x\}) = 0 & \widetilde{Q}_2((x,y),S\times\{y\}) = 1 \\ \widetilde{Q}_2((y,x),S\times\{x\}) = 0 & \widetilde{Q}_2((y,x),S\times\{y\}) = 1 \\ \widetilde{Q}_2((y,y),S\times\{x\}) = \frac{1-\alpha'}{1-\alpha} & \widetilde{Q}_2((y,y),S\times\{y\}) = 1-\frac{1-\alpha'}{1-\alpha}. \end{array}$$

For example we max assume that $\widetilde{Q}_2$ satisfies

$$\widetilde{Q}_2((u_1,u_2),(v_1,v_2)) = \widetilde{Q}_2((u_1,u_2),(v_1,S)) \cdot \widetilde{Q}_2((u_1,u_2),(S,v_2)),$$

where the marginal values are given from above.
We may suppose that $\alpha' = 1 - \frac{\alpha}{1-c}$. Consequently the right marginal condition is satisfied:

$$\begin{aligned} \widetilde{\mu}\otimes\widetilde{Q}_1\otimes\widetilde{Q}_2(\{x\}\times S\times\{x\}\times S\times\{x\}\times S) &= \alpha^2 \\ &= \varepsilon_x\otimes Q\otimes Q(x,x,x) \\ \widetilde{\mu}\otimes\widetilde{Q}_1\otimes\widetilde{Q}_2(\{x\}\times S\times\{x\}\times S\times\{y\}\times S) &= \alpha(1-\alpha) \\ &= \varepsilon_x\otimes Q\otimes Q(x,x,y) \\ \widetilde{\mu}\otimes\widetilde{Q}_1\otimes\widetilde{Q}_2(\{x\}\times S\times\{y\}\times S\times\{x\}\times S) &= (1-\alpha)(1-\alpha') \\ &= \varepsilon_x\otimes Q\otimes Q(x,y,x) \\ \widetilde{\mu}\otimes\widetilde{Q}_1\otimes\widetilde{Q}_2(\{x\}\times S\times\{y\}\times S\times\{y\}\times S) &= (1-\alpha)\alpha' \\ &= \varepsilon_x\otimes Q\otimes Q(x,y,y). \end{aligned}$$

In the same way the left marginal condition can be verified.
We maintain that $(\widetilde{Z}_n)_n$ is no martingale under $\widetilde{P}_{\widetilde{\mu}}$ for $\alpha' \neq 1-2\alpha$:

First note that the matrix Q has the eigenvalues $\beta_0 = 1$ and $\beta_1 = \alpha + \alpha' - 1$ with eigenvectors

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \frac{\alpha-1}{1-\alpha'} \\ 1 \end{pmatrix}. \tag{13}$$

Letting $g(x_1, x_2) = f(x_1) - f(x_2)$ and $A = \{(x,y),(x,x)\} \times S^2 \times \ldots \in \mathcal{F}_1$, we show that for $\alpha' \neq 1 - 2\alpha$

$$\int_A g(Y_2) d\widetilde{P}_{\widetilde{\mu}} \neq \int_A \beta_1 g(Y_1) d\widetilde{P}_{\widetilde{\mu}}. \tag{14}$$

The right side of (14) is zero, while the left side is equal to

$$\begin{aligned}
& g(x,y)\widetilde{P}^{(2)}_{\widetilde{\mu}}((x,y),(x,x),(x,y)) + g(y,x)\widetilde{P}^{(2)}_{\widetilde{\mu}}((x,y),(x,x),(y,x)) \\
= \; & g(x,y)\alpha(1-\alpha')(1-c) + g(y,x)\alpha(1-\alpha')c \\
= \; & \alpha \frac{2-(\alpha+\alpha')}{1-\beta}(1-2\alpha-\alpha') \neq 0.
\end{aligned}$$

Therefore $E_{\widetilde{P}_{\widetilde{\mu}}}(\, g(Y_2) \,|\, \mathcal{F}_1\,) \neq \beta_1 g(Y_1)$, the latter being equivalent to $E_{\widetilde{P}_{\widetilde{\mu}}}(\, \widetilde{Z}_2 \,|\, \mathcal{F}_1\,) \neq \widetilde{Z}_1$.

We want to give another example in order to illustrate that there are non-coupling measures $\widetilde{P}$ such that $(\widetilde{Z}_n)_n$, defined in (3), is a martingale.

Example 3.7 *For Q as in (12) we choose a kernel $\widetilde{Q}$ on $S \times S$ such that*

$$[\widetilde{Q}g](x,y) = \beta_1 g(x,y), \; x, y \in S, \tag{15}$$

where $g(x,y) = f(x) - f(y)$ and f is the eigenvector in (13) for β_1.
The induced measure for an arbitrary initial distribution $\widetilde{\mu}$ and any kernel $\widetilde{Q}$ satisfying (15) is a martingale measure for $(\widetilde{Z}_n)_n$.
Let $S = \{x, y\}$. Then the condition (15) is equivalent to the following four conditions

$$\begin{aligned}
\widetilde{Q}((x,x),(x,y)) &= \widetilde{Q}((x,x),(y,x)) \\
\widetilde{Q}((y,y),(x,y)) &= \widetilde{Q}((y,y),(y,x)) \\
\widetilde{Q}((x,y),(x,y)) &= \beta_1 + \widetilde{Q}((x,y),(y,x)) \\
\widetilde{Q}((y,x),(x,y)) &= \widetilde{Q}((y,x),(y,x)) - \beta_1.
\end{aligned}$$

Let us choose $\widetilde{Q}$ such that $\widetilde{Q}((x,y),(x,y)) = \frac{1+\beta_1}{2}$ and $\widetilde{Q}((x,y),(y,x)) = \frac{1-\beta_1}{2}$. For $\alpha \neq \alpha'$ and $\widetilde{\mu} = \varepsilon_x \otimes \varepsilon_y$, we conclude

$$\begin{aligned}
\widetilde{\mu} \otimes \widetilde{Q}((x,S),(x,S)) &= \widetilde{Q}((x,y),(x,x)) + \widetilde{Q}((x,y),(x,y)) \\
&= \frac{1+\beta}{2} = \frac{\alpha+\alpha'}{2} \\
&\neq \alpha = \varepsilon_x \otimes Q(x,x).
\end{aligned}$$

Therefore the measure $\widetilde{P}_{\widetilde{\mu},\widetilde{Q}}$ is not a coupling of $(P_{x,Q}, P_{y,Q})$.

Proof of Proposition 2.1: Let $\mathcal{A}^1 = \otimes_{n=0}^{\infty}\mathcal{A}_n^1$ resp. $\mathcal{A}^2 = \otimes_{n=0}^{\infty}\mathcal{A}_n^2$ denote the product-σ-algebra on $\Omega^1 := \prod_{n=0}^{\infty}\Omega_n^1$ resp. $\Omega^2 := \prod_{n=0}^{\infty}\Omega_n^2$.
Since we can identify Ω with the set $\Omega^1 \times \Omega^2$ via $(\omega_0, \ldots) \mapsto ((\omega_0^1, \ldots), (\omega_0^2, \ldots))$ we can write

$$\left(\Omega, \otimes_{n=0}^{\infty}(\mathcal{A}_n^1 \otimes \mathcal{A}_n^2)\right) = \left(\Omega^1 \times \Omega^2, \mathcal{A}^1 \otimes \mathcal{A}^2\right).$$

By symmetry it suffices to show, that for all $A \in \mathcal{A}^1$ the equality

$$\widetilde{P}_{\widetilde{\mu}}(A \times \Omega^2) = P_\mu(A) \tag{16}$$

holds.
Since $\mathcal{A}^1$ is generated by the $\cap$-stable system

$$\left\{ A_0 \times A_1 \times \ldots \times A_n \;\mid n \in \mathbb{N}_0,\; A_i \in \mathcal{A}_i^1,\; i \in \{0, \ldots, n\}\right\},$$

it it sufficient to show (16) for sets $A = A_0 \times A_1 \times \ldots \times A_n \times \Omega_{n+1}^1 \times \ldots$.
Instead of (16) we will show that for any function $f \in \mathcal{F}_{n,b}^1$ the identity

$$\int_\Omega f d\widetilde{P}_{\widetilde{\mu}} = \int_{\Omega^1} f dP_{\mu_1}, \tag{17}$$

holds, where $\mathcal{F}_n^1 = \sigma(\,Y_k^1 \;\mid\; 0 \le k \le n\,)$ and f is regarded as a function on Ω via $f(\omega) := f(\omega_0^1, \ldots, \omega_n^1, \ldots)$. Having proved (17) yields (16) using $f = 1_{A_0 \times \ldots A_n \times \Omega_{n+1}^1 \times \ldots}$.
The proof of (17) is obtained by induction over n.
For $n = 0$ we obtain (17) by

$$\int_\Omega f d\widetilde{P}_{\widetilde{\mu}} \;=\; \int_{\Omega_0} f(Y_0) d\widetilde{P}_{\widetilde{\mu}}^{(0)} = \int_{\Omega_0} f(Y_0) d\widetilde{\mu} = \int_{\Omega_0^1} f(Y_0^1) d\mu_1 = \int_{\Omega^1} f dP_{\mu_1}.$$

By induction hypothesis we know that (17) is satisfied for $f \in \mathcal{F}_{n-1,b}^1$.
Using the notation $\Omega_n = \Omega_n^1 \times \Omega_n^2$ we have

$$\begin{aligned}
&\int_\Omega f d\widetilde{P}_{\widetilde{\mu}} \\
=\;& \int_{\Omega_0 \times \ldots \times \Omega_n} f\left(Y_0, \ldots, Y_n\right) d\widetilde{P}_{\widetilde{\mu}}^{(n)} \\
=\;& \int_{\Omega_0 \times \ldots \times \Omega_{n-1}} \int_{\Omega_n} f(\omega_0, \ldots, \omega_n) \widetilde{Q}_{n-1}(\omega_0, \ldots, \omega_{n-1}, d\omega_n) \widetilde{P}_{\widetilde{\mu}}^{(n-1)}(d(\omega_0, \ldots, \omega_{n-1}).
\end{aligned}$$

Applying Lemma 3.4 to the inner integral with $R_i = \Omega_0^i \times \ldots \times \Omega_{n-1}^i$, $S_i = \Omega_n^i$ for $i \in \{1, 2\}$ and $f \in \mathcal{R}_{1,b}$, we conclude

$$\begin{aligned}
&\int_{\Omega_n} f(\omega_0, \ldots, \omega_n) \widetilde{Q}_{n-1}(\omega_0, \ldots, \omega_{n-1}, d\omega_n) \\
=\;& \int_{\Omega_n^1} f(\omega_0^1, \ldots, \omega_n^1) Q_{n-1}(\omega_0^1, \ldots, \omega_{n-1}^1, d\omega_n^1).
\end{aligned}$$

The function

$$(\omega_0^1, \ldots, \omega_{n-1}^1) \mapsto [Q_{n-1}f](\omega_0^1, \ldots, \omega_{n-1}^1)$$

is bounded and $\mathcal{F}_{n-1}^1$-measurable, so that we have by induction hypothesis

$$\int_\Omega [Q_{n-1}f] d\widetilde{P}_{\widetilde{\mu}} = \int_{\Omega^1} [Q_{n-1}f] dP_{\mu_1}.$$

This yields

$$\begin{aligned}
& \int_\Omega f d\widetilde{P}_{\widetilde{\mu}} \\
= & \int_{\Omega_0 \times \ldots \times \Omega_{n-1}} [Q_{n-1}f](Y_0, \ldots, Y_{n-1}) \widetilde{P}_{\widetilde{\mu}}^{(n-1)} \\
= & \int_\Omega [Q_{n-1}f] d\widetilde{P}_{\widetilde{\mu}} \\
= & \int_{\Omega^1} [Q_{n-1}f] dP_{\mu_1} \\
= & \int_{\Omega_0^1 \times \ldots \times \Omega_{n-1}^1} [Q_{n-1}f](Y_0^1, \ldots, Y_{n-1}^1) dP_{\mu_1}^{(n-1)} \\
= & \int_{\Omega_0^1 \times \ldots \times \Omega_{n-1}^1} \int_{\Omega_n^1} f(\omega_0^1, \ldots, \omega_n^1) Q_{n-1}(\omega_0^1, \ldots, \omega_{n-1}^1, d\omega_n^1) P_{\mu_1}^{(n-1)}(d(\omega_0^1, \ldots, \omega_{n-1}^1)) \\
= & \int_{\Omega_0^1 \times \ldots \times \Omega_n^1} f(Y_0^1, \ldots, Y_n^1) dP_{\mu_1}^{(n)} \\
= & \int_{\Omega_1} f dP_{\mu_1}. \quad \square
\end{aligned}$$

The converse of Proposition 2.1 is false in the following sense. Let $\widetilde{P}$ be a coupling as probability measure on $\Omega = (S \times S)^{\mathbb{N}_1}$ of (P_1, P_2). If S is a Polish space there exists corresponding transition kernels $\widetilde{Q}_n : (S \times S)^n \times \mathcal{S} \otimes \mathcal{S} \to [0,1]$ and an initial distribution $\widetilde{\mu}$ such that $\widetilde{P}$ is the measure that is induced by $\widetilde{\mu}$ and $(\widetilde{Q}_n)_{n \geq 1}$. Let μ_1 resp. μ_2 and $(Q_n^1)_n$ resp. (Q_n^2) denote the corresponding initial distribution and transition kernels to P_1 resp. P_2. Then $\widetilde{Q}_n$ is in general not a coupling of (Q_n^1, Q_n^2). This is also shown by example 3.6.

3.3 Proofs for bounds in the applications

For purposes of comparison we cite the Poincaré Inequality from [DS91]. In [DS91] the Markov chain Q on a finite set X is assumed throughout to be time-reversible with respect to the probability distribution π, e.g. $\pi(x)Q(x,y) = \pi(y)Q(y,x) =: \alpha(x,y) \; \forall x, y \in X$.

Proposition 3.8 *(Poincaré Inequality) For any irreducible Markov chain Q the second largest eigenvalue satisfies*

$$\beta_1 \leq 1 - \frac{1}{\kappa},$$

with κ defined by

$$\kappa := \max_e \left\{ \sum_{\gamma_{xy} \ni e} |\gamma_{xy}|_\alpha \pi(x)\pi(y) \right\}.$$

Here for any path γ_{xy} the path length is given by

$$|\gamma_{xy}|_\alpha := \sum_{e \in \gamma_{xy}} \frac{1}{\alpha(e)}.$$

Proof: [DS91].
For nearest-neighbor random walk on a connected graph (X, E) the proposition specializes to

Corollary 3.9 *Let (X, E) be a connected graph. Then for any choice of paths $\Gamma = \{ \gamma_{xy} \mid x, y \in X, \; x \neq y\}$*

$$\beta_1 \leq 1 - \frac{2|E|}{d_*^2 \gamma_* b}.$$

Here $d_* := \max\{ d(x) \}$, $\gamma_* := \max\{ l(\gamma) \mid \gamma \in \Gamma\}$ and

$$b := \max |\{ \gamma \in \Gamma \mid e \in \gamma\}|.$$

Proof: [DS91].

3.4 Proofs for bounds of the circle $\mathbf{Z}_p$

The Hitting Bound.
Using the hitting-bound we first calculate the mean-values E_k, where E_k for $k \in \{0, \ldots, \frac{p-1}{2}\}$ denotes the mean value of the hitting time until absorption in 0 starting in k. Solving the (uniquely solvable) equations

$$E_k = \frac{1}{2}(E_{k-1} + E_{k+1}) + 1, \; k \in \{1, \ldots, \frac{p-1}{2}\} \tag{18}$$

yields $E_k = k(p-k)$.
We conclude that the supremum $\sup_{x,y \in X} E_x(T_y)$ is achieved for $k = \frac{p-1}{2}$, so that the hitting bound gives

$$1 - \beta_1 \geq \frac{4}{(p-1)(p+1)}.$$

The Coupling Bound.
For any two points $x, y \in X$ denote $l(x, y)$ the length of a shortest path from x to y. Let $(x, y) \in X^2$ be the starting pair of the coupling process. By symmetry the mean value $E_{(x,y),\widetilde{Q}_c}(T)$ is the same if the starting pair (x, y) is replaced by (x', y') with $l(x, y) = l(x', y')$.

Therefore we will shortly write E_k instead of $E_{(x,y),Q}(T)$ for $l(x,y) = k \in \{0, \ldots, \frac{p-1}{2}\}$.
The mean values E_k have to satisfy the following equations

$$E_k = \frac{1}{2}(E_{k-2} + E_{k+2}) + 2,\ k \in \{1, \ldots, \frac{p-1}{2}\} \tag{19}$$

with the marginal conditions

$$E_{-1} = E_1, \quad E_0 = 0, \quad E_{\frac{p+1}{2}} = E_{\frac{p-1}{2}} \quad \text{and} \quad E_{\frac{p+3}{2}} = E_{\frac{p-3}{2}}. \tag{20}$$

It is only a matter of verification to see that (19) and (20) is satisfied by

$$\begin{aligned} & E_k = (p-k)\left(\tfrac{p+k}{2}\right) && \text{for } k \text{ odd} \\ \text{and} \quad & E_k = \left(p - \tfrac{k}{2}\right) k && \text{for } k \text{ even.} \end{aligned}$$

In order to show uniqueness let $(x_i), (y_i)$, $i \in \{-1, \ldots, \frac{p+3}{2}\}$ be two solutions of (19) and (20).
Set

$$M := \max\left\{ |x_i - y_i| \ : \ i \in \{1, \ldots, \frac{p-1}{2}\} \right\}$$

and

$$i_0 := \min\left\{ i \in \{1, \ldots, \frac{p-1}{2}\} \ : \ |x_i - y_i| = M \right\}.$$

Since the equations (19) have to be fulfilled, we get

$$\begin{aligned} M = |x_{i_0} - y_{i_0}| &= \frac{1}{2}|x_{i_0-2} - y_{i_0-2} + x_{i_0+2} - y_{i_0+2}| \\ &\leq \frac{1}{2}\left(|x_{i_0-2} - y_{i_0-2}| + |x_{i_0+2} - y_{i_0+2}|\right) \end{aligned}$$

If $i_0 > 1$, it follows $|x_{i_0-2} - y_{i_0-2}| < M$. Because of $|x_{i_0+2} - y_{i_0+2}| \leq M$ this yields the contradiction $M < M$, so that we are done.
In the case $i_0 = 1$ assume first, that $\frac{p-1}{2}$ is odd.
It follows from (19) for $k \in \{1, \ldots, \frac{p-1}{2}\}$

$$x_{k+2} - y_{k+2} = 2(x_k - y_k) - (x_{k-2} - y_{k-2}).$$

At first for $k = 1$ we conclude from this that $x_3 - y_3 = x_1 - y_1$ since $x_{-1} - y_{-1} = x_1 - y_1$.
Next it follows for $k = 3$ also $x_5 - y_5 = x_3 - y_3$.
By continuing this yields

$$M = x_1 - y_1 = x_3 - y_3 = x_5 - y_5 = \ldots = x_{\frac{p-1}{2}} - y_{\frac{p-1}{2}} = y_{\frac{p+3}{2}} - y_{\frac{p+3}{2}}.$$

Now we conclude backwards from the equations

$$x_{k-2} - y_{k-2} = 2(x_k - y_k) - (x_{k+2} - y_{k+2}),\ k \in \{1, \ldots, \frac{p-1}{2}\}$$

by using $x_{\frac{p+3}{2}} - y_{\frac{p+3}{2}} = x_{\frac{p-3}{2}} - y_{\frac{p-3}{2}}$ the identity

$$M = x_{\frac{p-7}{2}} - y_{\frac{p-7}{2}} = \ldots = x_2 - y_2 = x_0 - y_0 = 0.$$

This means $M = 0$ which shows that any two solutions $(x_i), (y_i)$ have to be equal in the case $\frac{p-1}{2}$ odd.
If $\frac{p-1}{2}$ is even, we conclude in the same manner

$$M = x_1 - y_1 = x_3 - y_3 = \ldots = x_{\frac{p-3}{2}} - y_{\frac{p-3}{2}} = \underbrace{x_{\frac{p+1}{2}} - y_{\frac{p+1}{2}}}_{=x_{\frac{p-1}{2}} - y_{\frac{p-1}{2}}}.$$

Then it follows backwards

$$M = x_{\frac{p-5}{2}} - y_{\frac{p-5}{2}} = \ldots = x_2 - y_2 = x_0 - y_0 = 0$$

and again we get $x_i = y_i$, $i \in \{-1, \ldots, \frac{p+3}{2}\}$.
This shows that (3.4) is the uniquely determined solution of (19).
Theorem 1.5 yields

$$(1 - \beta_1) \geq \frac{1}{E_1} = \frac{2}{p-1} = \frac{p-3}{p-1}.$$

3.5 Proof of the bounds for the star

The Coupling Bound.
In the sequel we calculate the mean-values for the case of the classical coupling. Again the mean-value of $T := \inf\{n \geq 0 \mid Y_n \in \Delta\}$ only depends on the distance of the two starting points $x, y \in X$.

The following equations have to be satisfied by the mean-values:

$$\begin{aligned} E_0 &= 0 \\ E_1 &= E_1\,(1-\theta) + 1 \\ E_2 &= 1 \end{aligned}$$

We compute

$$\begin{aligned} E_1 &= \frac{1}{\theta} \\ E_2 &= 1 \end{aligned}$$

The Poincaré Inequality.

Fixing the edge $(1,0)$, we see that there are $(n-1)$ paths of length 2 leading from point 1 to any of the points $2,\dots,n$ and that there is one path from 1 to 0 of length 1. By using that the stationary measure π of Q is given by

$$\begin{aligned}\pi(0) &:= \frac{1}{2-\theta}\\ \pi(i) &:= \frac{1-\theta}{2-\theta}\frac{1}{n} \quad \text{for } i\in\{1,\dots,n\},\end{aligned}$$

we conclude

$$\kappa = (n-1)\left(2n\frac{2-\theta}{1-\theta}\right)\pi(1)^2 + \frac{2-\theta}{1-\theta}n\pi(1)\pi(0) = \frac{3n-2n\theta-2(1-\theta)}{n(2-\theta)}.$$

Therefore

$$1-\beta_1 \geq \frac{2-\theta}{3-2\theta-2\frac{1-\theta}{n}}.$$

3.6 Proof for the bounds of the Cube $\mathbf{Z}_2^d$

In the sequel we calculate first the eigenvalues of nearest-neighbor random walk on the cube with holding probability θ. For further details on group representational arguments the reader is referred to [Dia88].
The set $\mathbf{Z}_2^d = \{0,1\}^d$ is together with adding in components modulo 2 an Abelian group of order 2^d. Recall that, given a vector space V, any group homomorphism $\rho : \mathbf{Z}_2^d \to GL(V)$ is called a representation of $\mathbf{Z}_2^d$ of dimension $d_\rho := dim(V)$. A subspace $W \subset V$ is called stable, if

$$\rho(g)(W) \subset W \quad \text{for all } g \in \mathbf{Z}_2^d.$$

A representation ρ is called irreducible, if there exist only the trivial stable subspaces $\{0\}$ and V.
Let $e_0 = (0,\dots,0)$ and for $i \in \{0,\dots,d\}$ define $e_i \in \mathbf{Z}_2^d$ by $e_i(j) = \delta_{ij}$ $(j \in \{1,\dots,d\})$.
Of course for $V = \mathbb{R}$ any representation must be irreducible. A representation $\rho : \mathbf{Z}_2^d \to \mathbb{R}$ is uniquely determined by its values on $\{e_1,\dots,e_d\}$. Moreover any element $x \in \mathbf{Z}_2^d$ induces an irreducible representation ρ_x by

$$\rho_x(e_i) := \begin{cases} -1 & : \quad x_i = 1 \\ 1 & : \quad x_i = 0. \end{cases}$$

Let Q be the transition kernel given by

$$\begin{aligned} Q(f,f) &:= \theta \in [0,1]\\ Q(f,f+e_i) &:= \frac{1-\theta}{d},\ i\in\{1,\dots,d\}\\ Q(f,f+g) &:= 0 \quad \text{for } g \notin \{e_0,\dots,e_d\}.\end{aligned}$$

Then we get for the eigenvalues

Proposition 3.10 *The eigenvalues of Q are given by*

$$\beta_j = 1 - \frac{2(1-\theta)}{d} j$$

with multiplicity $\binom{d}{j}$ *for* $j \in \{0, \ldots, d-1\}$.

Proof: For any $x \in \mathbf{Z}_2^d$ the vector $\rho_x = (\rho_x(g))_{g \in \mathbf{Z}_2^d}$ is an eigenvector for the eigenvalue $\beta_{w(x)}$, where $w(x) := \sum_{i=1}^d x_i$. The last assertion follows immediately from

$$\begin{aligned}
[Q\rho_x](f) &= \sum_{g \in \mathbf{Z}_2^d} Q(f,g)\rho_x(g) \\
&= \sum_{i=0}^{d} Q(f, f+e_i)\rho_x(f+e_i) \\
&= \left(\theta + \frac{(1-\theta)}{d} \sum_{i=1}^{d} \rho_x(e_i)\right) \rho_x(f) \\
&= \left(\theta + \frac{(1-\theta)}{d}(d - w(x) - w(x))\right) \rho_x(f) \\
&= \left(1 - \frac{2(1-\theta)}{d} w(x)\right) \rho_x(f) \\
&= \beta_{w(x)}\rho_x(f), \quad f \in \mathbf{Z}_2^d
\end{aligned}$$

This means $[Q\rho_x] = \beta_{w(x)}\rho_x$.
The linear independence is concluded by the fact (cm. **Theorem 3.**, [Dia88] Chapter 2B), that the irreducible representations of $\mathbf{Z}_2^d$ are orthonormal with respect to the positive definite bilinear form

$$(\phi, \psi) := \frac{1}{2^d} \sum_{g \in \mathbf{Z}_2^d} \phi(g)\psi(g^{-1}). \quad \Box$$

The Coupling Bound.
First checking the left marginal condition of $\widetilde{Q}$ defined in (7) yields for $i \in I \backslash V(\bar{y})$

$$\begin{aligned}
\widetilde{Q}(Y_{n+1} = (y_1 + e_i, \mathbf{Z}_2^d) \mid Y_n = \bar{y}) &= \sum_{z \in \mathbf{Z}_2^d} \underbrace{\widetilde{Q}(Y_{n+1} = (y_1 + e_i, z) \mid Y_n = \bar{y})}_{=0, \text{ for } z \neq y_1 + e_i} \\
&= \frac{1}{d+1}.
\end{aligned}$$

For $i \in \{0\} \cup V(\bar{y})$ we have

$$
\begin{aligned}
& \widetilde{Q}(Y_{n+1} = (y_1 + e_i, \mathbf{Z}_2^d) \mid Y_n = \bar{y}) \\
= & \sum_{z \in \mathbf{Z}_2^d} \widetilde{Q}(Y_{n+1} = (y_1 + e_i, z) \mid Y_n = \bar{y}) \\
= & \sum_{\{\, j \mid j \neq i,\ j \in \{0\} \cup V(\bar{y}) \,\}} \widetilde{Q}(Y_{n+1} = (y_1 + e_i, y_2 + e_j) \mid Y_n = \bar{y}) \\
= & \frac{1}{d+1}.
\end{aligned}
$$

Therefore the left marginal condition is fulfilled and by symmetry the right marginal condition of $\widetilde{Q}$ must also be satisfied.
We get a new Markov chain on the state space of the distances $\{0, \ldots, d\}$ with transition probabilities

$$
\begin{aligned}
R(k,k) &:= \widetilde{Q}(k(Y_{n+1}) = k \mid k(Y_n) = k) = \frac{d-k}{d+1} \\
R(k,k-1) &:= \widetilde{Q}(k(Y_{n+1}) = k-1 \mid k(Y_n) = k) = \frac{2}{d+1} \\
R(k,k-2) &:= \widetilde{Q}(k(Y_{n+1}) = k-2 \mid k(Y_n) = k) = \frac{k-1}{d+1}.
\end{aligned}
$$

Denoting by E_k for $k \in \{0, \ldots, d\}$ the mean-value $E_{(y_1,y_2),\widetilde{Q}}(T) = E_k$ for (y_1, y_2) with $k(\bar{y}) = k$, the following equations have to be satisfied

$$
\begin{aligned}
E_0 &= 0 \\
E_k &= \frac{d-k}{d+1} E_k + \frac{2}{d+1} E_{k-1} + \frac{k-1}{d+1} E_{k-2} + 1 \quad \text{for } k \in \{1, \ldots, d\}.
\end{aligned}
\tag{21}
$$

Equivalent to (21) is

$$
\begin{aligned}
E_0 &= 0 \\
E_k &= \frac{2}{k+1} E_{k-1} + \frac{k-1}{k+1} E_{k-2} + \frac{d+1}{k+1} \quad \text{for } k \in \{1, \ldots, d\}.
\end{aligned}
\tag{22}
$$

We maintain that for $k \in \{1, \ldots, d\}$ the inequalities

$$
\begin{aligned}
&i) \quad E_k \leq E_{k-1} + \tfrac{k+2}{2k(k+1)}(d+1) \\
&ii) \quad E_k \geq E_{k-1} + \tfrac{k-1}{2k(k+1)}(d+1)
\end{aligned}
\tag{23}
$$

hold.
Proof: By induction over k. For $k = 1$ we have $E_0 = 0$ and $E_1 = \frac{1}{2}(d+1)$, so

$$
\begin{aligned}
E_1 &\leq E_0 + \frac{2}{4}(d+1) \quad \text{and} \\
E_1 &\geq E_0 + \frac{0}{2 \cdot 1 \cdot 2}(d+1) = 0.
\end{aligned}
$$

By induction hypothesis $i)$ and $ii)$ of (23) is true for $l \leq k$.
ad $ii)$ Using (22) and $i)$ yields

$$\begin{aligned} E_{k+1} &= \frac{2}{k+2}E_k + \frac{k}{k+2}\underbrace{E_{k-1}}_{\geq E_k - \frac{k+2}{2k(k+1)}(d+1)} + \frac{d+1}{k+2} \\ &\geq E_k - \frac{(k+2)}{2(k+1)(k+2)}(d+1) + \frac{d+1}{k+2} \\ &= E_k + \frac{2(k+1)-(k+2)}{2(k+1)(k+2)}(d+1) \\ &= E_k + \frac{k}{2(k+1)(k+2)}(d+1) \end{aligned}$$

This shows part $ii)$ of the assertion for $k+1$.
ad $i)$ Using $ii)$ we get

$$\begin{aligned} E_{k+1} &= \frac{2}{k+2}E_k + \frac{k}{k+2}\underbrace{E_{k-1}}_{\leq E_k - \frac{(k-1)(d+1)}{2k(k+1)}} + \frac{d+1}{k+2} \\ &\leq E_k - \frac{(k-1)}{2(k+1)(k+2)}(d+1) + \frac{d+1}{k+2} \\ &= E_k + \frac{2(k+1)-(k-1)}{2(k+1)(k+2)}(d+1) \\ &= E_k + \frac{k+3}{2(k+1)(k+2)}(d+1) \end{aligned}$$

This shows (23). □
Using part $ii)$ of (23), we conclude the lower bound for E_d

$$\begin{aligned} E_d &\geq \frac{(d+1)}{2}\sum_{k=1}^{d-1}\frac{k}{(k+1)(k+2)} \\ &\geq \frac{(d+1)}{6}\sum_{k=1}^{d-1}\frac{1}{k+1} \\ &\geq \frac{(d+1)}{6}\ln\left(\frac{d+1}{2}\right). \end{aligned}$$

Using part $i)$ of (23) yields the upper bound

$$E_d \leq \frac{(d+1)}{2}\sum_{k=1}^{d}\frac{k+2}{k(k+1)}$$

$$\leq \frac{(d+1)}{2}\left(1+\sum_{k=2}^{d}\frac{1}{k}\right)$$
$$\leq \frac{(d+1)}{2}(1+\ln(d)).$$

In summary we have

$$\frac{(d+1)}{6}\ln\left(\frac{d+1}{2}\right) \leq E_d \leq \frac{(d+1)}{2}(1+\ln(d)). \tag{24}$$

Therefore $E_d = O(d\ln(d))$ and application of Proposition 1.5 yields

$$1-\beta_1 \geq \frac{1}{E_d} \geq \frac{2}{(d+1)\,(\ln(d)+1)}.$$

Taking the eigenvector ρ_{e_i} for arbitrary $i \in \{1,\ldots,d\}$, we see, that the supremum

$$\sup_{\bar{x}\in X^2} |\rho_{e_i}(x_1) - \rho_{e_i}(x_2)|$$

is attained for $s_0 = (e_0, e_i)$. Application of Proposition 1.4 with starting point s_0, yields

$$1-\beta_1 \geq \frac{2}{E_{s_0,\widetilde{Q}}(T)} = \frac{2}{E_1} = \frac{2}{d+1}.$$

Poincaré Inequality.
The stationary measure is given by $\pi(x) = \frac{1}{2^d}$, $x \in \mathbf{Z}_2^d$. For any $x \neq y \in \mathbf{Z}_2^d$ choose a path from x to y such that the different components of x and y are changed from the left to the right. There are $d2^{d-1}$ (non-oriented) edges and clearly $\gamma_* = d$. Any edge $e = (e^-, e^+)$ with $e_i^- = 1 - e_i^+$ is used by at most $2^{i-1}\cdot 2^{d-i} = 2^{d-1}$ many paths. (There are 2^{i-1} predecessors of e_i^- that can lead to 2^{d-i} many successors of e_i^+.)
Corollary 3.9 gives

$$1-\beta_1 \geq \frac{2d2^{d-1}}{(d+1)^2 d2^{d-1}} = \frac{2}{(d+1)^2},$$

which is off by a factor of d. Diaconis and Stroock show in [DS91] that the above choice of paths already yields the best value for b namely $b = 2^{d-1}$.

References

[D38] W. Doeblin, *Exposé de la théorie des chaines simples constantes de Markov a un nombre fini d'états*, Rev. Math. Union Interbalkanique **2**, 77-105 (1938).

[IM76] Dean L. Isaacson, Richard W. Madsen, *Markov Chains Theory and Applications* Wiley (1976).

[SE] E. Seneta, *Non-negative matrices and Markov chains*, Springer-Verlag, New York (1981).

[JA86] M. Jacobsen, *Co-optional Times and Invariant Measures for Transient Markov Chains*, Analytic and Geometric Stochastics, Applied Probability Trust, 49-58 (1986).

[AD87] D. Aldous, P. Diaconis, *Strong Uniform Times and Finite Random Walks*, Adavances in Applied Mathematics **8**, 69-97 (1987).

[Dia88] P. Diaconis *Group Representations in Probability and Statistics* IMS, Hayward, Calif. (1988).

[DS91] P. Diaconis, D. Stroock, *Geometric Bounds For Eigenvalues of Markov Chains* The Annals of Applied Probability, No. 1, 36-61 (1991).

[F91] J.A. Fill, *Eigenvalue bounds on convergence to stationarity for nonreversible markov chains, with an application to the exclusion process*, The Annals of Applied Probability, No. 1, 62-87 (1991).

[CHWA92] M.F. Chen, F.Y. Wang, *Applications of coupling method to the first eigenvalue on manifold*, Science in China, Technical Report, No. 215 (1994).

[CH94] M.F. Chen, *Optimal Markovian Couplings and Applications*, Acta Math. Sinica (NS) **10** (1994) (Nr. 3).