The simple random walk and max-degree walk on a directed graph

Ravi Montenegro *

Abstract

We bound total variation and L^∞ mixing times, spectral gap and magnitudes of the complex valued eigenvalues of general (non-reversible non-lazy) Markov chains with a minor expansion property. The resulting bounds for the (non-lazy) simple and max-degree walks on a (directed) graph are of the optimal order. It follows that, within a factor of two or four, the worst case of each of these mixing time and eigenvalue quantities is a walk on a cycle with clockwise drift.

Keywords : Markov chain, evolving sets, Eulerian graph, spectral gap, eigenvalues.

1 Introduction

Markov chains are a key tool in approximation algorithms for combinatorial counting problems and for sampling from discrete spaces. Surprisingly, little is known about the convergence rate of a Markov chain with no holding probability. Even for the simple random walk (i.e. nearest neighbor walk) on an undirected graph the order of magnitude for the slowest converging walk seems to be unknown.

More specifically, consider a connected undirected graph with m edges, n vertices and maximum degree d. The lazy (i.e. strongly aperiodic) simple random walk is known to converge in $O(m^2 \log(m/\epsilon))$ steps, because for instance the conductance is at least $1/2m$, and so the lazy max-degree walk mixes in time $O(n^2 d^2 \log(nd/\epsilon))$ as well. However, comparable (or better) bounds appear to be unknown in the non-lazy case. We remedy this by giving new bounds for simple and max-degree walks which are better then these, apply to directed graphs, require no holding probability, and are nearly sharp.

To state our results, define an Eulerian graph to be a strongly connected directed graph such that each vertex v has the same in and out-degrees $deg(v)$. This is the natural directed analog of an undirected graph, as any undirected graph can be made into an Eulerian graph by replacing each undirected edge with two directed edges. Two natural walks on a graph will be considered. For the simple random walk choose a neighbor uniformly at random and go there, while in the max-degree walk choose a neighbor with probability $1/d$ each and otherwise do nothing.

Which (non-lazy) directed walks mix rapidly? Certainly it is necessary that the walk not get stuck drifting between sets of equal sizes, such as from one bipartition to another (e.g. simple walk on a cycle with an even number of vertices). To avoid this it is enough that if the walk starts in a set of size $\pi(A) \leq 1/2$, then the set of adjacent vertices has size $> \pi(A)$. For instance, a

*Department of Mathematical Sciences, University of Massachusetts Lowell, Lowell, MA 01854, ravi_montenegro@uml.edu.

max-degree walk on a strongly connected graph with a self-loop at each vertex. It will be found that this expansion condition is also sufficient.

We now give our main results. Note that $\tau(\epsilon)$ is total variation mixing time (time to converge at an "average" vertex), $\tau_\infty(\epsilon)$ is L^∞ mixing time (time to converge at every vertex), $\lambda_i \neq 1$ is any non-trivial (complex-valued) eigenvalue of the transition matrix, λ is the spectral gap, and $N(A) = \{y \in V : \exists x \in A, \mathsf{P}(x,y) > 0\}$ is the neighborhood of set A, and $\pi(v)$ is the stationary distribution.

Corollary 1.1. *The simple random walk on an Eulerian graph with m edges satisfies*

$$\lambda \geq 1 - \cos\frac{2\pi}{m} \approx \frac{2\pi^2}{m^2}\,.$$

If it satisfies the expansion condition that

$$\forall A \subset V,\ \pi(A) \leq 1/2,\ \forall v \in N(A):\ \pi\left(N(A) \setminus v\right) \geq \pi(A)$$

then also

$$\begin{aligned}
1 - |\lambda_i| &\geq 1 - \cos\frac{2\pi}{m} \approx \frac{2\pi^2}{m^2} \\
\tau(\epsilon) &\leq \frac{1}{-\log\cos\frac{2\pi}{m}} \log\frac{1-2/m}{\epsilon} \approx \frac{m^2}{2\pi^2}\log\frac{1}{\epsilon} \\
\tau_\infty(\epsilon) &\leq \min\left\{\frac{\log\frac{m-2}{2} + \log\frac{1}{\epsilon}}{-\log\cos\frac{2\pi}{m}},\ \frac{m^2}{6} + \frac{m^2}{8}\log\frac{1}{\epsilon}\right\}\delta_{\epsilon\leq 1} + m^2\frac{1+3\epsilon}{3(1+\epsilon)^3}\,\delta_{\epsilon>1} \\
&\approx \frac{m^2}{2\pi^2}\log\frac{m-2}{2\epsilon}\,\delta_{\epsilon<1/m} + \frac{m^2}{8}\log\frac{4}{\epsilon}\,\delta_{\epsilon\in[1/m,1]} + \frac{m^2}{(1+\epsilon)^2}\,\delta_{\epsilon>1}
\end{aligned}$$

For the lazy simple random walk the bound on λ is a factor two smaller, the expansion condition is replaced by strong connectivity, and in the remaining bounds replace m by $2m$.

It follows that every lazy simple Eulerian walk converges in the same $\tau(\epsilon) = O(m^2\log(1/\epsilon))$ steps required for a cycle walk, improving on and generalizing the classical result $\tau(\epsilon) = O(m^2\log(m/\epsilon))$ for a lazy simple undirected walk. This can be further improved on by an order of magnitude in the special case of a walk on a regular graph, or equivalently of a max-degree walk.

Corollary 1.2. *The max-degree walk on an Eulerian graph with n vertices and max-degree d satisfies*

$$\lambda \geq \frac{2}{d}\left(1 - \cos\frac{\pi}{n}\right) \approx \frac{\pi^2}{n^2 d}\,.$$

If it satisfies the expansion condition that

$$\forall A \subset V,\ |A| \leq |V|/2:\ |N(A)| > |A|$$

then also

$$
\begin{aligned}
1-|\lambda_i| &\geq \frac{2}{d}\left(1-\cos\frac{\pi}{n}\right) \approx \frac{\pi^2}{n^2 d} \\
\tau(\epsilon) &\leq \frac{1}{-\log(1-\frac{2}{d}(1-\cos\frac{\pi}{n}))}\log\frac{1-1/n}{\epsilon} \approx \frac{n^2 d}{\pi^2}\log\frac{1}{\epsilon} \\
\tau_\infty(\epsilon) &\leq \min\left\{\frac{\log(n-1)+\log\frac{1}{\epsilon}}{-\log(1-\frac{2}{d}(1-\cos\frac{\pi}{n}))},\ \frac{n^2 d}{3}+\frac{n^2 d}{4}\log\frac{1}{\epsilon}\right\}\delta_{\epsilon\leq 1} + n^2 d\,\frac{2}{3}\,\frac{1+3\epsilon}{(1+\epsilon)^3}\,\delta_{\epsilon>1} \\
&\approx \frac{n^2 d}{\pi^2}\log\frac{n-1}{\epsilon}\,\delta_{\epsilon<1/n} + \frac{n^2 d}{4}\log\frac{4}{\epsilon}\,\delta_{\epsilon\in[1/n,1]} + \frac{2\,n^2 d}{(1+\epsilon)^2}\,\delta_{\epsilon>1}
\end{aligned}
$$

For the lazy max-degree walk the bounds on λ is a factor two smaller, the expansion condition is replaced by strong connectivity, and in the remaining bounds replace d by $2d$.

How sharp are these bounds? For the simple random walk on the cycle with an odd number of vertices n (so $m = 2n$ and $d = 2$) the spectral gap bound is off by a factor of 4, the eigenvalue bounds are exact, and the upper bounds on $\tau(\epsilon)$ become lower bounds if $\log\frac{1-\pi_*}{\epsilon}$ is replaced by $\log\frac{1}{2\epsilon}$ (where $\pi_* = 2/m$ and $\pi_* = 1/n$ respectively). More generally, we define a precise notion of rate of expansion, and show that a cycle walk with clockwise drift will be within a factor two of being the slowest mixing, not only among simple or max-degree walks, but among all Markov chains with this rate of expansion.

An interesting aspect of our argument is that it uses the Evolving set methodology of Morris and Peres [4], in an improved form given by Montenegro and Tetali [3] which bounds total variation distance directly, without going through L^2 distance. Related bounds also show that with relative entropy and L^2 mixing times the cycle walk is again nearly the slowest walk.

The paper proceeds as follows. In Section 2 we review the Evolving set methodology. This is followed in Section 3 with a proof of our main mixing result, a generalization of the simple and max-degree Eulerian walks considered above. In Section 4 this is extended to a bound on convergence rates in distances other than total variation. The Appendix contains proofs of inequalities used in showing our results.

2 Review of Mixing and Evolving Sets

We begin by reviewing mixing time theory, and particularly Evolving Set ideas.

Let P be a finite irreducible Markov kernel on state space V with stationary distribution π, that is, P is a $|V|\times|V|$ matrix with entries in $[0,1]$, row sums are one, V is connected under P ($\forall x, y \in V\, \exists t : \mathsf{P}^t(x,y) > 0$), and π is a distribution on V with $\pi\mathsf{P} = \pi$. The time-reversal P^* is given by $\mathsf{P}^*(x,y) = \frac{\pi(y)\mathsf{P}(y,x)}{\pi(x)}$ and has stationary distribution π as well. If $A, B \subset V$ the ergodic flow from A to B is given by $\mathsf{Q}(A,B) = \sum_{x\in A, y\in B}\pi(x)\mathsf{P}(x,y)$. Given initial distribution σ, the t-step discrete time distribution is given by $\sigma\mathsf{P}^t$.

If the walk is strongly connected and aperiodic then $\sigma\mathsf{P}^t \xrightarrow{t\to\infty} \pi$. The rate of convergence can be measured by the variation distance

$$\|\sigma-\pi\|_{TV} = \frac{1}{2}\sum_{x\in V}|\sigma(x)-\pi(x)|\,.$$

The mixing time $\tau(\epsilon)$ denotes the worst-case number of steps required for the total variation distance $\|\mathsf{P}^t(x,\cdot) - \pi\|_{TV}$ to drop to ϵ.

Many bounds on mixing time are shown by working with the spectral gap, which is just the gap between the two largest eigenvalues of the walk $\frac{\mathsf{P}+\mathsf{P}^*}{2}$, that is,

$$\lambda = \min_{i\neq 0} 1 - \lambda_i\left(\frac{\mathsf{P}+\mathsf{P}^*}{2}\right) = \inf_{\mathrm{Var}(f)\neq 0} \frac{\frac{1}{2}\sum_{x,y\in V}(f(x)-f(y))^2\pi(x)\mathsf{P}(x,y)}{\frac{1}{2}\sum_{x,y\in V}(f(x)-f(y))^2\pi(x)\pi(y)}$$

where $\{\lambda_i(\mathsf{K})\}$ denotes the eigenvalues of Markov chain K, and $\lambda_0(\mathsf{K}) = 1$.

Our results use the Evolving set methodology of Morris and Peres [4].

Definition 2.1. Given set $A \subset V$, a step of the *evolving set process* is given by choosing $u \in [0,1]$ uniformly at random, and transitioning to the set

$$A_u = \{y \in V : \mathsf{Q}(A,y) \geq u\,\pi(y)\} = \{y \in V : \mathsf{P}^*(y,A) \geq u\}\,.$$

The evolving set process is thus a random walk on sets. By Theorem 4.6 and Corollary 4.9 of [3] it relates to the original Markov chain by

$$\|\mathsf{P}^t(x,\cdot) - \pi\|_{TV} \leq \frac{1}{\pi(x)}\mathsf{E}\pi(S_t)(1-\pi(S_t)) \tag{2.1}$$

where S_t denotes the t-th state of an evolving set walk starting at $S_0 = \{x\}$. The set size $\pi(S_t) \xrightarrow{t\to\infty} \{0,1\}$, and the rate of convergence in a single step can be measured by the root profile of [4] or the following generalization of [3]:

Definition 2.2. Given $f : [0,1] \to \mathbb{R}_+$ let

$$\mathcal{C}_f(A) = \frac{\int_0^1 f(\pi(A_u))\,du}{f(\pi(A))}\,.$$

The f-congestion is $\mathcal{C}_f = \max_{\pi(A)\leq 1/2} \mathcal{C}_f(A)$, and the f-congestion profile is any function satisfying $\mathcal{C}_f(r) \geq \max_{\pi(A)\leq r} \mathcal{C}_f(A)$ for $r \in (0,1)$.

For instance, the root profile of [4] is $\psi(r) = 1 - \mathcal{C}_{\sqrt{a}}(r)$. The f-congestion is related to mixing time and spectral gap by Theorem 2.3 below; the proof is identical to that of Theorem 3.2 of [2], but with (2.1) instead of a bound of [4].

Theorem 2.3. *Suppose $f : [0,1] \to \mathbb{R}_+$ is such that $\forall a \in (0,1/2] :\ 0 < f(a) \leq f(1-a)$. For every finite irreducible Markov chain, if $x \in V$ then the t-step walk satisfies*

$$\|\mathsf{P}^t(x,\cdot) - \pi\|_{TV} \leq \left(\max_{\pi(A)\leq 1/2} \frac{\pi(A)\pi(A^c)}{f(\pi(A))}\right)\frac{f(\pi(x))}{\pi(x)}\,\mathcal{C}_f^t\,.$$

Also, every (complex valued) eigenvalue $\lambda_i \neq 1$ of P satisfies

$$1 - |\lambda_i| \geq 1 - \mathcal{C}_f\,.$$

If only $\forall a \in (0,1) :\ f(a) > 0$ then both results still hold, but with $\mathcal{C}_f = \max_{A\subset V} \mathcal{C}_f(A)$.

The f-congestion can often be written in terms of isoperimetric (i.e. geometric) quantities via a straightforward optimization, by use of whatever quantities are fixed by the isoperimetric term, plus Jensen's Inequality that $\int f(g(x))\,d\mu(x) \leq f(\int g(x)\,d\mu(x))$ when f is concave and μ is a probability measure, and the identity $\int_0^1 \pi(A_u)\,du = \pi(A)$. A particularly useful quantity to consider is the *modified ergodic flow* from set A:

$$\Psi(A) = \frac{1}{2}\int_0^1 |\pi(A_u) - \pi(A)|\,du$$

Because $\int_0^1 \pi(A_u)\,du = \pi(A)$ then $\Psi(A)$ is the area below $\pi(A_u)$ and above $\pi(A)$, while also the area below $\pi(A)$ and above $\pi(A_u)$. By Lemma 4.17 of [3] this is as the smallest ergodic flow from set A into a set of size $\pi(A^c)$, that is

$$\Psi(A) = \min_{\substack{B\subset V,\, v\in V, \\ \pi(B)\leq\pi(A^c)<\pi(B\cup v)}} \mathsf{Q}(A,B) + \frac{\pi(A^c)-\pi(B)}{\pi(v)}\,\mathsf{Q}(A,v)\,.$$

When π is uniform this simplifies to $\Psi(A) = \min_{\pi(B)=\pi(A^c)} \mathsf{Q}(A,B)$, while if a walk is *lazy* (i.e. $\mathsf{P}(x,x) \geq 1/2\ \forall x \in V$) then $\Psi(A) = \mathsf{Q}(A,A^c)$ with worst case $B = A^c$.

What is a good choice of function f for the f-congestion? In Example 4.4 of [2] it was suggested that if the modified ergodic flow $\Psi(A) \geq C$ for every $A \subset V$ and constant C not depending on set size, then it is best to work with $f(a) = \sin(\pi a)$. For instance, the simple random walk on an odd length cycle with $\Psi(A) \geq 1/m$. Then, by Theorem 2.3,

$$\|\mathsf{P}^t(x,\cdot) - \pi\|_{TV} \leq (1-\pi_*)\mathcal{C}^t_{\sin(\pi a)} \quad \text{and} \quad 1 - |\lambda_i| \geq 1 - \mathcal{C}_{\sin(\pi a)}\,,$$

where $\pi_* = \min_{v\in V}\pi(v)$.

3 General random walks

We now set out to show our main result, eigenvalue and total variation mixing bounds for general random walks (the L^∞ case will be dealt with in the next section). Two corollaries of this will be the specific walks on Eulerian graphs discussed in the introduction. In particular, we will find that even when general Markov chains are considered, a walk with clockwise drift on a cycle is still within a factor two of being the slowest mixing Markov chain.

The modified ergodic flow $\Psi(A)$ will play a key role in our proof, but our main theorem will involve a slightly weaker quantity. In practice these two will usually be the same. Given $A \subset V$, let $\hat{\mathsf{Q}}(A,x) = \min\{\mathsf{Q}(A,x), \pi(x)/2\}$ and define

$$\hat{\Psi}(A) \;=\; \min_{\substack{B\subset V,\, v\in V, \\ \pi(B)\leq\pi(A^c)<\pi(B\cup v)}} \sum_{x\in B} \hat{\mathsf{Q}}(A,x) + \frac{\pi(A^c)-\pi(B)}{\pi(v)}\hat{\mathsf{Q}}(A,v)\,.$$

As with $\Psi(A)$, for a uniform distribution $\hat{\Psi}(A) = \min_{\pi(B)=\pi(A^c)} \hat{\mathsf{Q}}(A,B)$. For a lazy walk $\hat{\Psi}(A) = \hat{\mathsf{Q}}(A,A^c) = \mathsf{Q}(A,A^c) = \Psi(A)$, or if $\Psi(A) \leq \Delta_{min}/2$ (defined below) then again $\hat{\Psi}(A) = \Psi(A)$.

To motivate the form of our main result, we note that in their work on Blocking conductance Kannan, Lovász and Montenegro [1] show that the square of conductance can often be replaced by a product of a measure of vertex boundary and a measure of edge expansion. Likewise, our general bound will involve a product of edge expansion $\hat{\Psi}_{min}$ with a measure of vertex boundary $\hat{A}_{max}$, rather than just the square of edge expansion which is found in most isoperimetric results.

Theorem 3.1. *Given a finite Markov chain, let*

$$\begin{array}{rclrcl} \hat{\Psi}_{min} & = & \min_{\pi(A)\leq 1/2} \hat{\Psi}(A) & \hat{A}_{min} & = & \min\{\hat{\Psi}_{min},\, \Delta_{min}/2\} \\ \Delta_{min} & = & \min_{\substack{A,B\subset V, \\ \pi(A)\neq\pi(B)}} |\pi(A)-\pi(B)| & \hat{A}_{max} & = & \max\{\hat{\Psi}_{min},\, \Delta_{min}/2\} \\ \mathsf{Q}_{min} & = & \min_{A\subset V} \mathsf{Q}(A,A^c) & \pi_* & = & \min_{v\in V} \pi(v)\,. \end{array}$$

Then,

$$\begin{array}{rcl} \tau(\epsilon) & \leq & \dfrac{1}{-\log\left(1-2\frac{\hat{A}_{min}}{\Delta_{min}}\,(1-\cos(2\pi\,\hat{A}_{max}))\right)} \log\dfrac{1-\pi_*}{\epsilon} \\ & \approx & \dfrac{1}{2\pi^2\hat{\Psi}_{min}\hat{A}_{max}} \log\dfrac{1-\pi_*}{\epsilon} \\ 1-|\lambda_i| & \geq & 2\dfrac{\hat{A}_{min}}{\Delta_{min}}\,(1-\cos(2\pi\hat{A}_{max})) \approx 2\pi^2\,\hat{\Psi}_{min}\,\hat{A}_{max} \\ \lambda & \geq & \dfrac{2\mathsf{Q}_{min}}{\pi_*}(1-\cos(\pi\pi_*)) \approx \pi^2\,\pi_*\mathsf{Q}_{min} \end{array}$$

Proof of Corollaries 1.1 and 1.2 (see Section 4 for the L^∞-bounds). First to Corollary 1.1. Suppose that $\pi(A)\leq 1/2$, and B,v are such that $\Psi(A)=\mathsf{Q}(A,B)+\frac{\pi(A^c)-\pi(B)}{\pi(v)}\mathsf{Q}(A,v)$. If $N(A)\subseteq B^c$ then $\pi(N(A)\setminus v)\leq\pi(B^c\setminus v)=1-\pi(B\cup v)<\pi(A)$, contradicting the expansion condition. Hence, $N(A)\cap B\neq\emptyset$ and so $\exists x\in A,\ y\in B$ with $\mathsf{P}(x,y)>0$, and so $\Psi(A)\geq\mathsf{Q}(A,B)\geq\pi(x)\mathsf{P}(x,y)\geq 1/m$. Likewise, for some $B\subset V$, $\hat{\Psi}(A)\geq\hat{\mathsf{Q}}(A,B)\geq\min\{\mathsf{Q}(x,y),\pi(y)/2\}$ and so $\hat{\Psi}(A)\geq 1/m$ if $\pi(y)=deg(y)/m\geq 2/m$. If $deg(y)=1$ then $N(\{y\})$ has only a single vertex v, and so $\pi(N(\{y\})\setminus v)=0$ contradicting the expansion condition. It follows that $\pi(y)\geq\pi_*\geq 2/m$. Corollary 1.1 then follows from Theorem 3.1 and the bound $\Delta_{min}\geq 1/m$. Corollary 1.2 follows similarly, but with $\hat{\Psi}(A)\geq 1/nd$ and $\Delta_{min}=1/n$. □

Note that the max-degree walk is actually the same as the simple random walk when each vertex x has $d-deg(x)$ self-loops added, and yet Corollary 1.2 is much better than that induced by Corollary 1.1. To understand this, recall that, in keeping with the intuition of Blocking Conductance, Theorem 3.1 will greatly improve on a bound involving edge-expansion alone (i.e. $\Psi(A)$ or $\hat{\Psi}(A)$) if $\Delta_{min}\gg\Psi_{min}$. In fact, the max-degree walk had $\Delta_{min}=1/n\gg 1/nd=\Psi_{min}$.

The theorem gets us very close to answering the question of what is the worst of all random walks, as shown by the following examples.

Example 3.2. Consider the simple random walk on a cycle (with $m=2n$ edges). Since $m=2n$ and $d=2$ then Corollaries 1.1 and 1.2 are the same. In Example 3.3 we find that Corollary 1.2 is exact for the eigenvalue gap and essentially sharp for mixing times, and hence Corollary 1.1 is equally good.

Example 3.3. Consider a max-degree walk on a cycle with an odd number of vertices n, such that at each vertex there are $d-1$ edges pointing in the clockwise direction, and 1 edge pointing in the counterclockwise direction.

This walk has an eigenvalue $\lambda_k = \frac{d-1}{d} e^{\pi i(n-1)/n} + \frac{1}{d} e^{-\pi i(n-1)/n}$ with eigenvector $f(x) = e^{\pi i x(n-1)/n}$ where the vertices are labeled clockwise as $x \in \{0, 1, \ldots, n-1\}$. Then

$$1 - |\lambda_k| = 1 - \sqrt{1 - \frac{4}{d}(1-1/d)\sin^2\frac{\pi}{n}} \approx \frac{2}{d}(1-1/d)\left(\frac{\pi}{n}\right)^2 \approx \frac{2\pi^2}{n^2 d}\,.$$

Corollary 1.2 gives a fairly similar bound of

$$\min 1 - |\lambda_k| \geq \frac{2}{d}(1 - \cos(\pi/n)) \approx \frac{\pi^2}{n^2 d}\,.$$

The upper and lower bounds are equal at $d = 2$, and within a factor two of equality when $d > 2$.

For spectral gap, note that $\frac{\mathsf{P}+\mathsf{P}^*}{2}$ is just the simple random walk on a cycle, and the largest eigenvalue of this is $\cos(2\pi/n)$. Consequently $\lambda = 1 - \cos(2\pi/n) \approx \frac{2\pi^2}{n^2}$. By Theorem 3.1 every walk with $\mathsf{Q}_{min} = 1/n$ and $\pi_* = 1/n$ satisfy$\lambda \geq 2(1 - \cos(\pi/n)) \approx \frac{\pi^2}{n^2}$, and so our drifting walk is within a factor two of having the worse spectral gap among all walks with $\mathsf{Q}_{min} = 1/n$ and $\pi_* = 1/n$. Although Corollary 1.2 is quite poor for this example, it is only off by a factor of four when considering instead the simple random walk on a cycle with $d-2$ self-loops (and $\lambda = 1 - \frac{2}{d}(1 - \cos(2\pi/n))$.

Likewise, the upper and lower bounds on mixing time are quite similar, with

$$\tau(\epsilon) \geq \frac{1}{-\log|\lambda_k|}\log\frac{1}{2\epsilon} \geq \frac{1}{-\log\sqrt{1 - \frac{4}{d}(1-1/d)\sin^2\frac{\pi(n-1)}{n}}}\log\frac{1}{2\epsilon} \approx \frac{n^2 d}{2\pi^2}\log\frac{1}{2\epsilon}$$

while the upper bound is

$$\tau(\epsilon) \leq \frac{1}{-\log(1 - \frac{2}{d}(1 - \cos(\pi/n))}\log\frac{1-1/n}{\epsilon} \approx \frac{n^2 d}{\pi^2}\log\frac{1}{\epsilon}\,.$$

The bounds are nearly equivalent at $d = 2$, and within a factor two of equality when $d > 2$. When $n = 3$ and $d = 2$ then the lower bound can be sharpened slightly to be exactly equal to the upper bound.

Example 3.4. Consider a general Markov chain R with uniform stationary distribution $\pi(\cdot) = 1/n$. Note that $\Psi_{\min,\mathsf{R}} \leq \Psi(\{v\}) \leq \pi(v)(1-\pi(v))$ for every $v \in V$, and so $\Psi_{min,\mathsf{R}} \leq \pi_*(1-\pi_*)$. Hence, in Theorem 3.1 we have $\hat{\Psi}_{\min,\mathsf{R}} \geq \frac{1}{2}\Psi_{\min,\mathsf{R}}$ and $A_{max} \geq 1/2n$ and so the upper bound on mixing time in Theorem 3.1 is roughly $\frac{2n}{\pi^2\Psi_{\min,\mathsf{R}}}\log\frac{1-\pi_*}{\epsilon}$.

On the other hand, the Markov chain of Example 3.3 can be generalized to the walk $\mathsf{P}(x, x+1) = \alpha \in [1/2, 1]$ and $\mathsf{P}(x, x-1) = 1-\alpha$, with $\Psi_{\min,\mathsf{P}} = (1-\alpha)\pi_*$ an eigenvalue $\lambda_k = \alpha e^{\pi i(n-1)/n} + (1-\alpha)e^{-\pi i(n-1)/n}$, and hence $\tau(\epsilon)$ lower bounded by roughly $\frac{n^2}{2\pi^2(1-\alpha)}\log\frac{1}{2\epsilon}$. If $\Psi_{min,\mathsf{R}} \leq \frac{1}{2}\pi_*$ then when $\alpha = 1 - \frac{\Psi_{min,\mathsf{R}}}{\pi_*} \geq 1/2$ then $\Psi_{\min,\mathsf{P}} = \Psi_{\min,\mathsf{R}}$ and the lower bound becomes $\frac{n}{2\pi^2\Psi_{\min,\mathsf{R}}}\log\frac{1}{2\epsilon}$, which shows that the mixing time of this cycle is at most four times as fast as that of the walk R.

Proof of Theorem 3.1. As suggested in the preliminaries, we will study the f-congestion $\mathcal{C}_{\sin(\pi a)}$. This will be done in two steps. First, we show a result appropriate for max-degree random walks. Then we consider a case relevant to the simple random walk.

Fix some set $A \subset V$.

First consider the case that $\Psi(A) < \Delta_{min}/2$.

Notice that if $\pi(A_u) > \pi(A)$ then $\pi(A_u) - \pi(A) \geq \Delta_{min}$, while if $\pi(A_u) < \pi(A)$ then $\pi(A) - \pi(A_u) \geq \Delta_{min}$. Since $\pi(A_u)$ is decreasing with $\Psi(A) = \int_{\pi(A_u)>\pi(A)} (\pi(A_u) - \pi(A))\, du < \Delta_{min}/2$ by assumption, it follows that $\pi(A_u) \leq \pi(A)$ when $u \geq \frac{\Psi(A)}{\Delta_{min}}$. Hence

$$\Psi(A) = \int_{\pi(A_u)>\pi(A)} (\pi(A_u) - \pi(A))\, du = \int_0^{\Psi(A)/\Delta_{min}} (\pi(A_u) - \pi(A))\, du\,.$$

Likewise $\pi(A_u) \geq \pi(A)$ when $u \leq 1 - \frac{\Psi(A)}{\Delta_{min}}$, and so $\Psi(A) = \int_{1-\Psi(A)/\Delta_{min}}^1 (\pi(A) - \pi(A_u))\, du$. Combining these, $\pi(A_u) = \pi(A)$ when $u \in \left[\frac{\Psi(A)}{\Delta_{min}}, 1 - \frac{\Psi(A)}{\Delta_{min}}\right]$.

Hence, if we let $f(a) = \sin(\pi a)$ and $M = \frac{\Psi(A)}{\Delta_{min}}$ then by Jensen's Inequality

$$\begin{aligned}
\mathcal{C}_f(A) &= \frac{M \int_0^M f(\pi(A_u))\, \frac{du}{M} + (1-2M) f(\pi(A)) + M \int_{1-M}^1 f(\pi(A_u))\, \frac{du}{M}}{f(\pi(A))} \\
&\leq \frac{M\, f\left(\pi(A) + \frac{\Psi(A)}{M}\right) + (1-2M) f(\pi(A)) + M\, f\left(\pi(A) - \frac{\Psi(A)}{M}\right)}{f(\pi(A))} \\
&= 1 - 2\frac{\Psi(A)}{\Delta_{min}} \left(1 - \cos(\pi \Delta_{min})\right) \\
&= 1 - 2\frac{\hat{\Psi}(A)}{\Delta_{min}} \left(1 - \cos(\pi \Delta_{min})\right)\,.
\end{aligned} \tag{3.2}$$

Now, consider the case that $\Psi(A) \geq \Delta_{min}/2$.

Choose $\wp \in [0,1]$ such that $\pi(A_u) \geq \pi(A)$ if $u < \wp$ and $\pi(A_u) \leq \pi(A)$ if $u > \wp$, so that $\Psi(A) = \int_0^{\wp} (\pi(A_u) - \pi(A))\, du = \int_{\wp}^1 (\pi(A) - \pi(A_u))\, du$. Then, if $\hat{\wp} = \min\{\wp, 1/2\}$ it follows that $\hat{\Psi}(A) = \int_0^{\hat{\wp}} (\pi(A_u) - \pi(A))\, du = \int_{\hat{\wp}}^1 (\pi(A) - \pi(A_u))\, du$. Hence, if $x = \pi(A)$ then

$$\begin{aligned}
\int_0^1 f(\pi(A_u))\, du &= \hat{\wp} \int_0^{\hat{\wp}} f(\pi(A_u))\, \frac{du}{\hat{\wp}} + (1-\hat{\wp}) \int_{\hat{\wp}}^1 f(\pi(A_u))\, \frac{du}{1-\hat{\wp}} \\
&\leq \hat{\wp} \sin\left(\pi\left(x + \frac{\hat{\Psi}(A)}{\hat{\wp}}\right)\right) + (1-\hat{\wp}) \sin\left(\pi\left(x - \frac{\hat{\Psi}(A)}{1-\hat{\wp}}\right)\right) \\
&\leq \sin(\pi x)\, \cos(2\pi \hat{\Psi}(A))
\end{aligned}$$

by Lemma 4.4 in the Appendix. Hence, $\mathcal{C}_{\sin(\pi a)}(A) \leq \cos(2\pi \hat{\Psi}(A))$.

Combine these two cases, maximize over sets $A \subset V$, and apply Theorem 2.3 to obtain the mixing time and eigenvalue bounds.

For the spectral gap, note that

$$\lambda = 2 \min 1 - \lambda_i \left(\frac{\mathsf{I}}{2} + \frac{\mathsf{P} + \mathsf{P}^*}{4}\right) = 2 \min 1 - \left|\lambda_i \left(\frac{\mathsf{I}}{2} + \frac{\mathsf{P} + \mathsf{P}^*}{4}\right)\right|\,.$$

Hence it suffices to study eigenvalues of $\mathsf{P}' = \frac{\mathsf{I}}{2} + \frac{\mathsf{P}+\mathsf{P}^*}{4}$. However, P' is a lazy walk and so $\hat{\Psi}(A) = \Psi(A) = \mathsf{Q}_{\mathsf{P}'}(A, A^c)$. This is in turn half the ergodic flow $\mathsf{Q}_{\frac{\mathsf{P}+\mathsf{P}^*}{2}}(A, A^c)$, and so $\hat{\Psi}(A) = \frac{1}{2}\mathsf{Q}_{\frac{\mathsf{P}+\mathsf{P}^*}{2}}(A, A^c) = \frac{1}{2}\mathsf{Q}_{\mathsf{P}}(A, A^c)$ (since $\mathsf{Q}_{\mathsf{P}}(A, A^c) = \mathsf{Q}_{\mathsf{P}^*}(A, A^c) = \mathsf{Q}_{\frac{\mathsf{P}+\mathsf{P}^*}{2}}(A, A^c)$). In short,

$$\hat{\Psi}_{min}\left(\frac{\mathsf{I}}{2} + \frac{\mathsf{P}+\mathsf{P}^*}{4}\right) = \frac{1}{2}\min_{A\subset V}\mathsf{Q}_{\mathsf{P}}(A, A^c) = \frac{1}{2}\mathsf{Q}_{min}(\mathsf{P})\,.$$

Before applying the eigenvalue bounds proven earlier, note for a lazy walk that $\Psi(A) = \mathsf{Q}(A, A^c)$, with $\frac{\mathsf{Q}(A,x)}{\pi(x)} < \frac{1}{2}$ only if $x \in A^c$, and $\frac{\mathsf{Q}(A,x)}{\pi(x)} > \frac{1}{2}$ only if $x \in A$. It follows that if $\pi(A_u) > \pi(A)$ then $A \subsetneq A_u$ and so $\pi(A_u) \geq \pi(A) + \pi_*$. Likewise, if $\pi(A_u) < \pi(A)$ then $A \supsetneq A_u$ and so $\pi(A_u) \leq \pi(A) - \pi_*$. Hence, when studying a lazy walk (such as $\frac{\mathsf{I}}{2} + \frac{\mathsf{P}+\mathsf{P}^*}{4}$), Δ_{min} may be replaced by π_* in our earlier analysis. But $\pi_*\left(\frac{\mathsf{I}}{2} + \frac{\mathsf{P}+\mathsf{P}^*}{4}\right) = \pi_*(\mathsf{P})$, and so the spectral bound follows from the earlier eigenvalue bounds. □

4 Other distances

Different applications may require different measures of the convergence rate. The total variation distance is the most widely used and measures distance from stationary at an average vertex. The much stronger L^∞ or relative pointwise distance measures distance from stationary at the worst vertex. In this section we show mixing bounds on L^2 distance, which are again within a small constant factor of those for the walk on a cycle with clockwise drift, and infer from this a bound on L^∞ distance as well. The interested reader can easily use the approach of this section to construct similar bounds for other distances, such as relative entropy.

Given distributions σ and π the L^2 distance is defined by

$$\|\sigma/\pi - 1\|_{2,\pi} = \sqrt{\sum_{x\in V}\pi(x)\left(\frac{\sigma(x)}{\pi(x)} - 1\right)^2}$$

The mixing time $\tau_2(\epsilon)$ is the worst case number of steps required for a walk to reach distance ϵ in L^2 distance. This can be bounded by applying the technique used to prove Theorem 2.3 to evolving set bounds on L^2 convergence [3] (Theorem 4.6 and Corollary 4.9) to show

$$\tau_2(\epsilon) \leq \frac{1}{1 - \mathcal{C}_{\sin(\pi a)}}\left(\frac{1}{2}\log\frac{1-\pi_*}{\pi_*} + \log\frac{1}{\epsilon}\right)\,. \tag{4.3}$$

When ϵ is large then this can be improved further via equation (4.5) of [3]:

Theorem 4.1. *When $r\left(1 - \mathcal{C}_{\sqrt{a(1-a)}}\left(\frac{1}{1+r^2}\right)\right)$ is convex then*

$$\tau_2(\epsilon) \leq \left\lceil \int_{\pi_*}^{1/(1+\epsilon^2)} \frac{dr}{2r(1-r)(1-\mathcal{C}_{\sqrt{a(1-a)}}(r))} \right\rceil\,.$$

It remains to bound the profile $\mathcal{C}_{\sqrt{a(1-a)}}(r)$:

Lemma 4.2. *Let* $\Psi_{min} = \min_{A\subset V} \Psi(A)$ *and* $A_{max} = \max\{\Psi_{min}, \frac{\Delta_{min}}{2}\}$*. Then we may take*

$$1 - \mathcal{C}_{\sqrt{a(1-a)}}(r) = \frac{\Psi_{min} A_{max}}{2r^2(1-r)^2}\, \delta_{r\le 1/2} + 8\Psi_{min} A_{max}\, \delta_{r>1/2}\,.$$

Proof. In Theorem 4.16 of [3] the authors show that

$$1 - \mathcal{C}_{\sqrt{a(1-a)}}(A) \ge \frac{\Psi(A)^2}{2\pi(A)^2\pi(A^c)^2}\,.$$

The lemma follows unless $\Psi_{min}(A) < \Delta_{min}/2$ for some $\pi(A) \le r$. But then equation (3.2) applied to $f(a) = \sqrt{a(1-a)}$, with $X = \frac{1}{2} + \frac{\Delta_{min}}{2\pi(A)}$ and $Y = \frac{1}{2} - \frac{\Delta_{min}}{2\pi(A^c)}$, implies

$$\begin{aligned} 1 - \mathcal{C}_f(A) &\ge 2\frac{\Psi(A)}{\Delta_{min}}\left(1 - \sqrt{XY} - \sqrt{(1-X)(1-Y)}\right) \\ &\ge 2\frac{\Psi(A)}{\Delta_{min}}\left(1 - \sqrt{1-(X-Y)^2}\right) \\ &\ge \frac{\Psi(A)\Delta_{min}}{4\pi(A)^2\pi(A^c)^2} \end{aligned}$$

where the second inequality was Lemma 4.3 of [2], i.e. $\sqrt{XY} + \sqrt{(1-X)(1-Y)} \le \sqrt{1-(X-Y)^2}$, and the final inequality is $\sqrt{1-x} \le 1 - x/2$. □

By combining this with equation (4.3) when $\epsilon \le 1$ and Theorem 4.1 when $\epsilon > 1$, we obtain the bound:

Corollary 4.3.

$$\tau_2(\epsilon) \le \left\lceil \frac{\frac{2}{3} + \log\frac{1}{\epsilon}}{8\Psi_{min} A_{max}}\, \delta_{\epsilon\le 1} + \frac{1+3\epsilon^2}{6\Psi_{min} A_{max}(1+\epsilon^2)^3}\, \delta_{\epsilon>1} \right\rceil$$

Proof of L^∞ cases in Corollaries 1.1 and 1.2. We use the relation $\tau_\infty(\epsilon) \le \tau_{2,\mathsf{P}}(\sqrt{\epsilon}) + \tau_{2,\mathsf{P}^*}(\sqrt{\epsilon})$ (see e.g. Appendix of [3]). Also, suppose $A, B \subset V$.

First Corollary 1.2. Since π is uniform then $\Psi(A) = \min_{\pi(B)=\pi(A^c)} \mathsf{Q}(A, B)$. If $\pi(B) = \pi(A^c)$ then

$$\begin{aligned} &\mathsf{Q}_{\mathsf{P}}(A, B) = \pi(B) - \mathsf{Q}_{\mathsf{P}}(A^c, B) \\ &\quad = \quad \pi(B) - \pi(A^c) + \mathsf{Q}_{\mathsf{P}}(A^c, B^c) = \mathsf{Q}_{\mathsf{P}^*}(B^c, A^c)\,, \end{aligned}$$

and so Ψ_{min} is the same for P and P^*. Finish with Corollary 4.3 using the conditions $\Delta_{min} \ge 1/n$, $\Psi_{min} \ge 1/nd$ and $A_{max} \ge 1/2n$.

Now to Corollary 1.1. Note that $\Psi_{min,\mathsf{P}} \ge 1/m$ was shown in the proof of the total variation case. It remains to bound Ψ_{min,P^*}. For some $v \in V$ and $\pi(B) \le \pi(A^c) < \pi(B \cup v)$ then, arguing as above,

$$\begin{aligned} \Psi_{min,\mathsf{P}^*} &= \mathsf{Q}_{\mathsf{P}^*}(A, B) + \frac{\pi(A^c) - \pi(B)}{\pi(v)}\, \mathsf{Q}_{\mathsf{P}^*}(A, v) \\ &= \mathsf{Q}_{\mathsf{P}}(B^c \setminus v, A^c) + \left(1 - \frac{\pi(A^c) - \pi(B)}{\pi(v)}\right) \mathsf{Q}_{\mathsf{P}}(v, A^c)\,. \end{aligned}$$

Arguing as in the proof of Corollaries 1.1 and 1.2 appearing after Theorem 3.1, if $\pi(C) \leq 1/2$ and $\pi(D) \leq \pi(C^c) < \pi(D \cup v)$ then $\mathsf{Q}(C, D) \geq 1/m$. Then, if $C = B^c \setminus v$ and $D = A^c$ it follows that $\Psi_{min,\mathsf{P}^*} \geq \mathsf{Q}_{\mathsf{P}}(B^c \setminus v, A^c) \geq 1/m$. Hence $\Delta_{min} \geq 1/m$, $\Psi_{min} \geq 1/m$ and $A_{max} \geq 1/2m$ for both P and P^*. Finish with Corollary 4.3. □

References

[1] R. Kannan, L. Lovász and R. Montenegro, "Blocking conductance and mixing in random walks," *Combinatorics, Probability and Computing*, vol. 15, no. 4, pp. 541–570, 2006.

[2] R. Montenegro, "Sharp edge, vertex, and mixed Cheeger inequalities for finite Markov kernels," *Electronic Communications in Probability*, vol. 12, pp. 377–389, 2007.

[3] R. Montenegro and P. Tetali, *Mathematical aspects of mixing times in Markov chains*, volume 1:3 of *Foundations and Trends in Theoretical Computer Science*, now publishers, Boston-Delft, June 2006.

[4] B. Morris and Y. Peres, "Evolving sets, mixing and heat kernel bounds," *Probability Theory and Related Fields*, vol. 133, no. 2, pp. 245–266, 2005.

Appendix

We have left for the Appendix the proof of an inequality key to our main theorem.

Lemma 4.4. *Given $a, b \in [0, 1/2]$ and $c \in [0, 1/4]$, if $c \leq b(1 - \max\{a, b\})$ then*

$$b \sin\left(\pi\left(a + \frac{c}{b}\right)\right) + (1-b)\sin\left(\pi\left(a - \frac{c}{1-b}\right)\right) \leq \sin(\pi a)\cos(2\pi c)\,.$$

Proof. Suppose $a = 0$. Observe that if $x \in [c, \infty)$ then

$$\frac{d}{dx} x \sin\frac{\pi c}{x} = \sin\frac{\pi c}{x} - \frac{\pi c}{x}\cos\frac{\pi c}{x} \geq 0$$

because $\sin y \geq y \cos y$ when $y \in [0, \pi]$. But $1 - b \geq b \geq c$ and so

$$b \sin\frac{\pi c}{b} - (1-b)\sin\frac{\pi c}{1-b} \leq 0\,. \tag{4.4}$$

The lemma then follows.

If $a > 0$ then let $h(a, b, c)$ denote the left side of the inequality in the lemma. By the identity $\sin(x + y) = \sin(x)\cos(y) + \sin(y)\cos(x)$ then

$$\frac{h(a,b,c)}{\sin(\pi a)} = b\cos\frac{\pi c}{b} + (1-b)\cos\frac{\pi c}{1-b} + \cot(\pi a)\left(b\sin\frac{\pi c}{b} - (1-b)\sin\frac{\pi c}{1-b}\right)$$

Since $a \in [0, 1/2]$ then $\cot(\pi a)$ is decreasing, and by (4.4) then $\frac{h(a,b,c)}{\sin(\pi a)}$ is increasing in a. Hence, if $b \geq 2c$ then $\frac{h(a,b,c)}{\sin(\pi a)} \leq \frac{h(1/2,b,c)}{\sin(\pi/2)} = h(1/2, b, c)$, otherwise $\frac{h(a,b,c)}{\sin(\pi a)} \leq \frac{h(1-c/b,b,c)}{\sin(\pi(1-c/b))}$.

When $b \geq 2c$ then $\frac{\pi c}{b}, \frac{\pi c}{1-b} \in [0, \pi/2]$ since $1 - b \geq \frac{1}{2} \geq 2c$. By concavity of $\cos x$ when $x \in [0, \pi/2]$, it follows that

$$\lambda \cos \frac{\pi c}{b} + (1-\lambda) \cos \frac{\pi c}{1-b} \leq \cos\left(\lambda \frac{\pi c}{b} + (1-\lambda) \frac{\pi c}{1-b}\right)$$

for every $\lambda \in [0, 1]$. When $\lambda = b$ it follows that $h(1/2, b, c) \leq \cos(2\pi c)$.

When $b < 2c$ then, since $\cot(\pi(1-x)) = -\cot(\pi x)$,

$$\begin{aligned} \frac{h(1-c/b, b, c)}{\sin(\pi(1-c/b))} &= \frac{(1-b)\sin \frac{\pi c}{b(1-b)}}{\sin \frac{\pi c}{b}} \\ &\leq \quad (1-2c)\cos \frac{\pi c}{1-2c} = h(1/2, 2c, c) \leq \cos(2\pi c) \end{aligned}$$

where the first inequality is Lemma 4.5 with $x = \frac{\pi c}{b(1-b)}$ and $y = \frac{\pi c}{b}$, and the second inequality follows from the case of $b \geq 2c$. $\square$

The following Lemma was required in the preceding proof.

Lemma 4.5. *If $x, y \in [\pi/2, \pi]$, $x > y$, and $\mathrm{sinc}(z) := \frac{\sin z}{z}$ then*

$$\frac{\mathrm{sinc}(x)}{\mathrm{sinc}(y)} \leq \left(1 - \frac{2}{\pi} y \left(1 - \frac{y}{x}\right)\right) \cos \frac{y\left(1-\frac{y}{x}\right)}{1 - \frac{2}{\pi} y \left(1 - \frac{y}{x}\right)}\,.$$

Since $\mathrm{sinc}(y)$ is decreasing when $y \in [\pi/2, \pi]$ then this measures how much sinc drops between y and x. A slightly weaker result that is perhaps a bit easier to grasp is

$$\forall x, y \in [\pi/2, \pi],\ x > y : \ \frac{\mathrm{sinc}(x)}{\mathrm{sinc}(y)} \leq \cos\left(2y\left(1 - \frac{y}{x}\right)\right)\,.$$

Proof. Rewrite the problem as one of showing that

$$f(x, y) := \left(1 - \frac{2}{\pi} y \left(1 - \frac{y}{x}\right)\right) \cos \frac{y\left(1-\frac{y}{x}\right)}{1 - \frac{2}{\pi} y \left(1 - \frac{y}{x}\right)} - \frac{\sin x}{x} \frac{y}{\sin y} \geq 0\,.$$

The third partial is

$$\begin{aligned} \frac{\partial^3 f}{\partial y^3} \quad &= \quad -\frac{\pi^5 x^2 (2y-x)^3}{(\pi x - 2xy + 2y^2)^5} \sin \frac{y\left(1-\frac{y}{x}\right)}{1 - \frac{2}{\pi} y \left(1 - \frac{y}{x}\right)} \\ &\quad -\frac{6\pi^3 x (2y-x)(x(\pi - x) + 2y(x-y))}{(\pi x - 2xy + 2y^2)^4} \cos \frac{y\left(1-\frac{y}{x}\right)}{1 - \frac{2}{\pi} y \left(1 - \frac{y}{x}\right)} \\ &\quad +\frac{6y}{x} \frac{\sin x}{\sin y} \cot^3 y - \frac{6}{x} \frac{\sin x}{\sin y} \cot^2 y + \frac{5y}{x} \frac{\sin x}{\sin y} \cot y - \frac{3}{x} \frac{\sin x}{\sin y} \\ &\leq \quad 0 \end{aligned}$$

The inequality is because every term is negative (note that $\frac{y(1-y/x)}{1-\frac{2}{\pi}y(1-y/x)} \in [0, \pi/2]$ and $\pi x - 2xy + 2y^2 > 0$).

It follows that the second partial is decreasing, and so for each $x \in [\pi/2, \pi]$ there are three possible cases: strictly convex in y, convex then concave in y, or strictly concave in y. If $\frac{\partial f}{\partial y}(x, \pi/2) \geq 0$ then in each case the minimum is at an endpoint, i.e. $f(x, y) \geq \min\{f(x, \pi/2), f(x, x)\} = 0$, and we are done. It remains to consider the first partial:

$$\begin{aligned}
\frac{\partial f}{\partial y}(x, \pi/2) &= \frac{2x(\pi - x)\cos(\pi - x) - (1 - 2(\pi - x))\sin(\pi - x)}{\pi x} \\
&\geq \min_{x \in [\pi/2, \pi]} \left\{0, \frac{2x(\pi - x)(1 - \frac{2}{\pi}(\pi - x)) - (1 - 2(\pi - x))(\pi - x)}{\pi x}\right\} \\
&= \min_{x \in [\pi/2, \pi]} \left\{0, \frac{\pi - x}{\pi x}\left(2\pi - 1 - \frac{4x(\pi - x)}{\pi}\right)\right\} = 0
\end{aligned}$$

The inequality is because the expression is trivially positive if $\pi - x \in [1/2, \pi/2]$, whereas if $a := \pi - x \in [0, 1/2]$ then use the relations $\cos a \geq 1 - \frac{2}{\pi}a$ and $\sin a \leq a$. $\square$