Problem statement. Given a network with its graph $G=(V,E)$; nodes $V=\{1,\ldots,n\}$ and $E$ be edge-set such that $G$ is connected.

Let nodes have load $x_i > 0$ be load of node $i$. We wish to balance load among these $n$ nodes evenly. That is, we wish to assign load

$$X_{\text{ave}} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

to all nodes.

**Goal.** Assign load $X_{\text{ave}}$ at all nodes through only local operations.

We will be interested in linear iterative algorithm. Specifically, let $P = [p_{ij}] \in \mathbb{R}_{+}^{n \times n}$ be matrix s.t. $p_{ij} > 0$ only if $(i,j) \in E$.

Initially:

$$X(0) = [x_i^0] \in \mathbb{R}_+^n$$

Iteration $k$:

$$X(k) = X^T(k-1) P$$

that is, $x_i(k-1)$ sends $p_{ij} \cdot x_i(k-1)$ amount to node $j$; node $i$ sets its new value to sum of values received from its neighbors including itself.
Theorem. Let $P$ satisfy the following conditions:
(a) $P$ be a doubly stochastic matrix
\[ \sum_k P_{ik} = 1 = \sum_k P_{kj} \] for all $i,j$.
(b) $P$ be irreducible and aperiodic.

Then, \( \lim_{k \to \infty} x(k) = x_{\text{ave}} \).

Proof. We wish to show that \( 1^T P = 1^T \) and then we want to show the convergence to desired limit.

\[
1^T P = \left[ \left( \sum_k P_{kj} \cdot 1 \right) \right]^T = \left[ \sum_k P_{kj} \right]^T = \left[ 1 \right]^T.
\]

Next, convergence. For this we will use the coupling argument. Let, \( Y_t \) be a discrete time, time-homogeneous Markov chain (HMC) on space \( V = \{1, \ldots, n\} \). Let its transition matrix be $P$, let $Z_t$ be another such Markov chain with the same transition matrix $P$. 
We create joint distribution of \((Y_t, Z_t)\) as follows:

\[Z_0\] has uniform distribution \(\frac{1}{n}\).

\[Y_0\] has an arbitrary distribution, say \(\mu\).

\(Y_{t+1}\) is generated from \(Y_t\) as per transition matrix \(P\).

\(Z_{t+1}\) is generated from \(Z_t\) as follows:

- till \(Z_t\) becomes same as \(Y_t\), it evolves independently as per transition matrix \(P\).
- else \(Z_{t+1} = Y_{t+1}\).

Let \(\tau = \inf \{ t \geq 1 : Z_t = Y_t \}\). Then, \(\tau\) is a stopping time. By definition, \(Y_t\) is HMC on \(\mathcal{E}_1, \ldots, \mathcal{E}_n\) with transition matrix \(P\). The \(Z_t\) is an HMC too with the same transition matrix \(P\). Since \(Z_0\sim\text{uniform}\) and \(\tau\) is stationary distribution of \(P\), the distribution of \(Z_t\) remains \(\frac{1}{n}\) for all \(t \geq 1\).

We claim that

\[d_{TV}(Y_t, Z_t) \leq P(\tau > t)\]
\[ P(Y_t \in A) - P(Z_t \in A) = P(Y_t \in A; z \geq T) + P(Y_t \in A; z < T) - P(Z_t \in A; z \geq T) - P(Z_t \in A; z < T) \]

Now \( P(Y_t \in A; z \leq T) = P(Z_t \in A; z \leq T) \).

Therefore, \( P(Y_t \in A) - P(Z_t \in A) \leq P(Y_t \in A; z \geq T) \leq P(Z_t). \)

Similarly, \( P(Y_t \in A) - P(Z_t \in A) \geq -P(Z_t). \)

Thus, \( \left| P(Y_t \in A) - P(Z_t \in A) \right| \leq P(Z_t). \)

This is true for all \( A \).

Therefore, \( d_{TV}(Y_t, Z_t) = \sup_{A \subseteq V} \left| P(Y_t \in A) - P(Z_t \in A) \right| \leq P(Z_t). \)

So if we show that \( P(Z < \infty) = 1 \), then \( \lim_{t \to \infty} d_{TV}(Y_t, Z_t) = 0. \)

That is, the distribution of \( Y_t \) converges to \( \frac{1}{n} \). Next, we prove remaining claim of \( P(Z < \infty) = 1 \) to complete the proof.
Note that the way we have define the probability distribution \( p_t \) \((Y_t,Z_t)\); it is sufficient to prove the following: let \((i,j) \in \mathbb{E} \times \mathbb{E} \times [1,2,\ldots,n]^2\) be any starting point. Every time, each coordinate transits independently according to transition matrix \( P \). That is,

\[
P((i,j) \rightarrow (i',j')) = P_{ii'} P_{jj'}
\]

Now, we show that such product MC is irreducible since \( P \) is irreducible, aperiodic: for large enough \( n \), there is positive probability of \( i \rightarrow k, j \rightarrow l \) for all \((k,l)\). Thus we have an irreducible Markov chain on \( \mathbb{E} \times \mathbb{E} \times [1,2,\ldots,n]^2 \). Since it is finite state it is positive recurrent. That is, starting from any \((i,j)\), with probability 1 the MC reaches \((k,k)\), \( k = 1,\ldots,n \). That is

\[
P(\tau < \infty) = 1. \quad \text{(QED)}
\]

Next, a more precise version of the above theorem.
Theorem [Perron-Frobenius]

Let \( P \) be \( n \times n \) non-negative valued matrix. Let it be irreducible. Then, there is an eigenvalue \( \lambda \geq 0 \) with vector \( v \) s.t. \( Pv = \lambda v \). Further all other e.v. \( \lambda \) be s.t. \( |\lambda| < \lambda \).

Proof. We have seen the proof this (almost) through the uniqueness of invariant distribution for positive recurrent MC and convergence to it as above. \( \square \)
**FACT:** Let $P$ be symmetric, doubly stochastic matrix. Then it has $n$ real valued eigenvalues with $\lambda_{\text{max}} = 1$; $\lambda_2, \ldots, \lambda_n$ s.t.

$$ |\lambda_i| < 1 \text{ for } i = 2, \ldots, n. $$

Let $v_1, \ldots, v_n$ be corresponding eigenvectors. Then $v_1, \ldots, v_n$ are orthogonal. That is

$$ \langle v_i, v_j \rangle = 0 $$

More generally, let $v_i$ be s.t. $\|v_i\| = 1$. Then for any vector $x \in \mathbb{R}^n$,

$$ x = \sum_{i=1}^{n} \langle x, v_i \rangle \cdot v_i $$

$$ P^k x = \sum_{i=1}^{n} \langle x, v_i \rangle \cdot \lambda_i^k \cdot v_i $$

Since $|\lambda_i| < 1$; $\lambda_i^k \to 0$ for $i = 2, \ldots, n$, while $\lambda_1 = 1$. Therefore,

$$ \lim_{k \to \infty} P^k x = \langle x, v_1 \rangle \cdot v_1 $$

For example, if $P$ had $1$ as e.v. then,

$$ \lim_{k \to \infty} P^k x \to \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right) \cdot 1 $$
Mixing Time. Let $\Pi$ be stationary distribution of a Markov chain with matrix $P$. Let $X_t$ be Markov chain state at time $t$. Let $\mu_i$ be distribution of $X_t$ with $X_0 = i$. Let 

$$T_\epsilon = \inf \{t : \max_i |\mu_i - \frac{\Pi}{\|\Pi\|} < \epsilon\}.$$ 

Then $T_\epsilon$ is called “mixing time” of Markov Chain.

Let $\mu$ be any starting distribution. As before,

$$\mu = \sum_{i=1}^{\infty} \langle \mu, v_i \rangle \cdot v_i$$

Then,

$$P^k \mu = \sum_{i=1}^{\infty} \langle \mu, v_i \rangle \cdot \lambda_i^k \cdot v_i$$

Therefore

$$\|P^k \mu - \langle \mu, u_1 \rangle u_1\|_2 = \|\sum_{i=2}^{\infty} \langle \mu, v_i \rangle \lambda_i^k v_i\|_2$$

(a) 

$$\leq \|\lambda_i^k\|_2 \left[\sum_{i=2}^{\infty} \langle \mu, v_i \rangle^2\right]$$

(b) 

$$\leq |\lambda_{\max}|^k.$$ 

Here, (a) follows due to definition $|\lambda_{\max}| = \max_{i=2,\ldots,n} |\lambda_i|$; (b) follows due to fact that

$$1 = \|\mu\|_1 = \|\mu\|_2 = \sum_{i=1}^{\infty} \langle \mu, v_i \rangle^2.$$
Clearly, $\langle u, v_1 \rangle v_1 = \pi$. Therefore,

$$T_\varepsilon \leq \inf \{ k : |\lambda_{\max}|^k < \varepsilon^\delta \}.$$

$$= \frac{\log \varepsilon^\delta}{\log |\lambda_{\max}|^k}$$

Usually $|\lambda_{\max}|$ is close to 1. Then,

$$\log |\lambda_{\max}| = \log (1 - (1 - |\lambda_{\max}|))$$

$$= -(1 - |\lambda_{\max}|).$$

Thus,

$$T_\varepsilon \approx \frac{\log \varepsilon^\delta}{(1 - |\lambda_{\max}|)}$$
Question: How to generate $P$ so that $P$ has uniform stationary distribution.

A simple algorithm: Let $d^* = \max_{i \in G} \deg_i$ of $G$.
Then, set

$$P_{ij} = \frac{1}{2d^*} \text{ for all } (ij) \in E;$$

set

$$P_{ii} = 1 - \left(\sum_{j \neq i} P_{ij}\right) = 1 - \frac{d_i}{2d^*}.$$

Clearly, $P$ is symmetric, doubly stochastic matrix. Therefore, it works.

**General Method:** Metropolis-Hastings Method.

Suppose we want $(\pi_i)$ as stationary distribution.

We only know $(\frac{\pi_i}{\pi_j})$ for $(ij) \in E$.

Suppose we have another Markov chain, say $Q = [Q_{ij}]$, we will generate $P$ with $\pi$ as stationary distribution.

$$R_{ij} = \frac{\pi_j}{\pi_i} \frac{Q_{ji}}{Q_{ij}}$$
At state \( i \), generate \( j \) according to \( \pi_{ij} \).
Draw \( u \) according to \( U[0,1] \). If \( u < \frac{1}{R_{ij}} \), accept \( j \). Else, remain at \( i \).

Note that \( R_{ij} = \frac{1}{R_{ji}} \). Suppose \( R_{ij} > 1 \).

Therefore,

\[
\pi_i \cdot P_{ij} = \pi_i \cdot \pi_{ij} = \frac{\pi_i}{\pi_j} \cdot \frac{\pi_{ij}}{\pi_{ji}} \cdot \pi_{ji} = \pi_j \cdot R_{ji} \cdot \pi_{ji} = \pi_j \cdot P_{ji}.
\]

Thus, we have got reversible \( P \) with \( \pi \) as stationary distribution.
Some examples.

1. Complete graph.

- \( P = \left[ \frac{1}{n} \right] \).
- For any \( M = (\mu_i) ; \ \Sigma \mu_i = 1 ; \ \mu_i \geq 0 \)

then \( M^T P = \left[ \sum_{j=1}^{n} \mu_j \cdot P_{ji} \right]^T \)

\[ = \left[ \frac{1}{n} \left( \sum_1^N \mu_j \right) \right]^T \]

\[ = \left[ \frac{1}{n} \right]^T \]

Thus, \( T_{\epsilon} = 1 \) for all \( \epsilon \geq 0 \).

2. Ring graph. Total \( n \) nodes.

- Each node has two neighbors.

\[ P = \left[ P_{ij} \right] ; \]

\[ P_{ij} = \begin{cases} 
\frac{1}{4} & j \equiv i+1 \mod n \text{ or } j \equiv i-1 \mod n \\
\frac{1}{2} & i=j \\
0 & \text{o.w.}
\end{cases} \]
Then, \( T_\epsilon = C \cdot n^2 \cdot \log \epsilon^{-1} \).

Thus: Graph topology affects performance.

Next, randomized Gossip algorithm.
Randomized Gossip Algorithm

* The above describe algorithm allows for all pairs to communicate simultaneously. This means that number of operation come be equal to number of edges at each time. For example, in complete graph it is $\Theta(n^2)$. In the ring graph it is $\Theta(n) \times \Theta(n^2 \log \varepsilon^{-1}) = \Theta(n^3 \log \varepsilon^{-1})$. This may be a lot more (redundant).

Further, under certain networks these operations may not be possible to perform simultaneously.

We consider, wireless sensor network model. Here each node can communicate to almost one other node simultaneously. That is, the simultaneously communicating edges form matching in the graph. Now the question is:

What is the penalty of matching constraint on time to compute avg or load balance?
As before, we have Graph $G = (V, E)$ and matrix $P$ s.t. it is irreducible, aperiodic, and $\frac{1}{n}$ as stationary distribution. That is $P$ is a doubly stochastic matrix.

**Property [Birkhoff–Von Neumann]**

Let, $\Lambda = \{ Q \in R_{+}^{n \times n} : \sum_{j} Q_{ij} = 1; \sum_{i} Q_{ji} = 1; \forall i \}$

be set of all doubly stochastic matrices. It is convex and its extreme points are 0-1 matrices, known as permutation matrices.

Due to above property, $P$ can be written as

$$P = \sum_{k=1}^{n^2} \alpha_k \Pi_k; \sum \alpha_k = 1; \alpha_k \geq 0.$$

$$\Pi_k \in \Lambda_{0-1} = \{ Q \in \Lambda : Q_{ij} \in \{0,1\} \}.$$

Clearly, $\Pi_k$ represents permutation of $n$ numbers.
Algorithm.

Initially: $x(0) = \begin{bmatrix} x^0 \end{bmatrix}$

Iterate: $k^{th}$ iteration, pick one of $(\Pi_i)_{i=1}^n$ with probabilities $(\alpha_i)_{i=1}^n$. Let $\Pi(k)$ be randomly chosen permutation. Then,

$$x(k) = \frac{1}{2} (I + \Pi(k)) x(k-1).$$

Analysis:

First moment:

$$E[x(k)] = E\left[\frac{1}{2} (I + \Pi(k)) x(k-1)\right]$$

$$= E \left[ E \left[ \frac{1}{2} (I + \Pi(k)) x(k-1) | x(k-1) \right] \right]$$

$$= E \left[ \frac{1}{2} \sum_{i=1}^n E \left[ I + \Pi(k) \right] | x(k-1) \right] x(k-1)$$

$$= \frac{1}{2} \sum_{i=1}^n E \left[ (I + E[\Pi(k)]) x(k-1) \right]$$

But $E[\Pi(k)] = \sum_{i=1}^n \alpha_i \Pi_i = P$
Therefore, \[ E \left[ x(k) \right] = \frac{1}{2} (I + P) E \left[ x(k-1) \right]. \]

Let's study \( W = \frac{1}{2} (I + P) \). Clearly, \( W \) is doubly stochastic since \( P \) is. Therefore, \( W \) has \( \frac{1}{n} \mathbf{1} \) as it's eigenvector (stability and convergence).

Suppose \( x \) is e.v. of \( P \). Then,

\[
Wx = \frac{1}{2} (I + P)x = \frac{1}{2} x + \frac{1}{2} \lambda x
\]

\[
= \frac{1}{2} (1 + \lambda)x.
\]

Thus, \( x \) is e.v. of \( W \) as well with eigenvalue \( \frac{1}{2} (1 + \lambda) \). Since \( P \) is symmetric all of its e.v. are real and due to it being doubly stochastic they have norm \( < 1 \) except the one. Therefore, e.v. of \( W \) are beta \( [0, 1) \) except one equal \( 1 \) with eigenvector \( \frac{1}{n} \mathbf{1} \).

Therefore,

\[
E \left[ x(k) \right] = W^k x(0) \rightarrow x = \mathbf{1}
\text{ as } k \rightarrow \infty.
\]

Question: does \( x(k) \rightarrow x = \mathbf{1} \) ?
For this, we need bound on “deviation.”

Define, \( y(k) = x(k) - x_{\text{ave}} \).

Then, \( y(k+1) = x(k+1) - x_{\text{ave}} \).

\[
= \frac{1}{2} (I + \Pi(k)) x(k) - x_{\text{ave}} \Pi
\]

\[
= \frac{1}{2} (I + \Pi(k)) x(k) - \frac{1}{2} (I + \Pi(k)) x_{\text{ave}} \Pi
\]

\[
= \frac{1}{2} (I + \Pi(k)) \left[ x(k) - x_{\text{ave}} \Pi \right]
\]

\[
= \frac{1}{2} (I + \Pi(k)) y(k).
\]

Thus, \( y(\cdot) \) obeys the same “linear dynamics” as \( x(\cdot) \). We wish to study 2nd moment of \( y(\cdot) \).

\[
E\left[ y(k)^T y(k) \right] = \frac{1}{4} E\left[ y(k-1)^T (I + \Pi(k)) (I + \Pi(k))^T y(k-1) \right]
\]

\[
= \frac{1}{4} E\left[ y(k-1)^T \right] E\left[ (I + \Pi(k)^T + \Pi(k) + \Pi(k)^T \Pi(k)) y(k-1) \right]
\]
For permutation matrices \( \Pi^T(k) \Pi(k) = I \). That is,
\[
E \left[ I + \Pi^T(k) + \Pi(k) + \Pi^T(k) \Pi(k) \right] = 2I + E[\Pi^T(k)] + E[\Pi(k)]
\]
But \( E[\Pi(k)] + E[\Pi^T(k)] = P + P^T \)

Therefore:
\[
E\left[ Y(k)^T Y(k) \right] = \frac{1}{2} E\left[ Y(k-1)^T \left( I + \frac{P + P^T}{2} \right) Y(k-1) \right]
\]

Now, both \( P \) and \( P^T \) have \( \frac{1}{n} \mathbf{1} \) as eigenvector. More generally, \( Q = \frac{1}{2} \left( I + \frac{P + P^T}{2} \right) \) has \( \frac{1}{n} \mathbf{1} \) as eigenvector.

Further, \( Q \) is symmetric. Therefore, it has eigenvectors that form an orthonormal basis.

Now, \( Y(k) = X(k) - \text{Xave} \mathbf{1} \) is s.t.
\[
\langle Y(k), \mathbf{1} \rangle = \langle X(k), \mathbf{1} \rangle - \text{Xave} \langle \mathbf{1}, \mathbf{1} \rangle = n \text{Xave} - n \text{Xave} = 0.
\]

Therefore \( Y(k) \perp \mathbf{1} \). So, we have
\[
Y(k)^T Q Y(k) \leq \lambda_{\text{max}}(Q) \cdot Y(k)^T Y(k).
\]

where \( \lambda_{\text{max}}(Q) = \max_{i=1} \lambda_i(Q) \)
Therefore, \( E[ y(x)^{T} y(x) ] \leq \lambda_{\text{max}}(Q) \cdot E[y(x)^{T} y(x)] \) \\
\leq \lambda_{\text{max}}(Q) \cdot y(0)^{T} y(0) \\
\leq \lambda_{\text{max}}(Q) \cdot x(0)^{T} x(0) \\

Because: \( y(0)^{T} y(0) = \langle y(0), y(0) \rangle \) \\
= \langle x(0) - x_{\text{ave}} 1, x(0) - x_{\text{ave}} 1 \rangle \\
= \langle x(0), x(0) \rangle - 2 x_{\text{ave}} \langle x(0), y \rangle + x_{\text{ave}}^2 \langle y, y \rangle \\
= \langle x(0), x(0) \rangle - 2 x_{\text{ave}} \cdot y + 2 x_{\text{ave}} \langle x, y \rangle \\
= \langle x(0), x(0) \rangle - x_{\text{ave}}^2 - n + n x_{\text{ave}}^2 \\
= \langle x(0), x(0) \rangle - n x_{\text{ave}}^2 \leq \langle x(0), x(0) \rangle \\

By chebychev’s inequality: \\
\text{Pr}( y(x)^{T} y(x) \geq \epsilon^2 \| x(0) \|_{2}^2 ) \leq \frac{E[y(x)^{T} y(x)]}{\epsilon^2 \| x(0) \|_{2}^2} \leq \frac{\lambda_{\text{max}}(Q) \cdot \| x(0) \|_{2}^2}{\epsilon^2 \| x(0) \|_{2}^2} = \frac{1}{\epsilon^2} \lambda_{\text{max}}(Q)
Thus: \( k(\varepsilon) \approx \frac{3 \log \frac{\varepsilon}{\Delta}}{\log \lambda_{\max}(Q)} \); we have
\[
P \left( \|X(k) - X_{\text{ave}} \|_2^2 \geq \varepsilon^2 \|x(0)\|_2^2 \right) \leq \varepsilon
\]
for \( k \geq k(\varepsilon) \).

Now: \( \lambda_{\max}(Q) = \lambda_{\max} \left( \frac{1}{2} \left( I + \frac{p + p^T}{\varepsilon^2} \right) \right) \).

Since \( \frac{p + p^T}{\varepsilon^2} \) is symmetric matrix, their e.v.'s are real \( \in \{-1, 1\} \). Let they be
\( \lambda_1 \geq \lambda_2 \ldots \geq \lambda_n \geq -1 \). Then, from above clearly:
\[
\lambda_{\max}(Q) = \frac{1}{2} \left( 1 + \lambda_2 \left( \frac{p + p^T}{\varepsilon^2} \right) \right).
\]

If \( p \) were symmetric, then \( \lambda_2 \left( \frac{p + p^T}{\varepsilon^2} \right) = \lambda_2(p) \).

Thus: \( \log \lambda_{\max}(Q) = \log \frac{1}{2} \left( 1 + \lambda_2(p) \right) \).

That is,
\[
k(\varepsilon) = \frac{3 \log \frac{\varepsilon}{\Delta}}{\log \frac{1}{2} \left( 1 + \lambda_2(p) \right)}
\]
**Lower Bound.**

Let \( Q \) have eigenvalues \( 1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0 \) as discussed above. The corresponding eigenvectors are \( \frac{1}{\sqrt{n}} \, V_1, V_2, \ldots, V_n \). Select,

\[
x(0) = \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{n}} \, V_1 + V_2 \right) \Rightarrow y(0) = \frac{V_2}{\sqrt{2}}.
\]

Here \( ||x(0)|| = 1 \). From above calculations,

\[
E[ y(k) ] = \left[ \frac{1}{2} (I + P) \right]^k \cdot y(0).
\]

\[
= Q^k \cdot y(0), \quad \text{for } P \text{ symmetric}.
\]

\[
= \lambda_{\max}^k \cdot \frac{V_2}{\sqrt{2}}
\]

Now, by Jensen's inequality:

\[
E[ y(k)^T y(k) ] \geq \left[ E[ y(k) ] \right]^T \cdot E[ y(k) ]
\]

\[
= \frac{1}{2} \cdot \lambda_{\max}^2 \cdot V_2^T \cdot V_2
\]

\[
= \frac{1}{2} \lambda_{\max}^{2k}
\]
Lemma. Let $X$ be a random variable such that $0 \leq x \leq B$. Then, for any $0 < \epsilon < B$,

$$
P(X \geq \epsilon) \geq \frac{\text{E}[X] - \epsilon}{B - \epsilon}.
$$

Proof. \[\text{E}[X] = \text{E}[X \cdot 1_{x<\epsilon}] + \text{E}[X \cdot 1_{x>\epsilon}] \leq \epsilon \text{E}[1_{x<\epsilon}] + B \text{E}[1_{x>\epsilon}] = \epsilon + (B - \epsilon) P(X \geq \epsilon)\]

Thus, \[P(X \geq \epsilon) \geq \frac{\text{E}(X) - \epsilon}{B - \epsilon}. \quad \text{Q.E.D.}\]

Let

$$
K_{\epsilon}(\epsilon) = \frac{1}{2} \frac{\log (2 \epsilon)^{-1}}{\log \lambda_{\max}^{-1}}.
$$

Now, $\|y_{(k)}\|_{2}^{2} \leq \|y_{0}\|_{2}^{2} \leq \frac{1}{2}$. Then, using above lemma we have,

$$
P(\|y_{(k)}\| \geq \epsilon) \geq \epsilon \text{ for } k < K_{\epsilon}(\epsilon).$$
Summarizing above analysis:

Theorem. Let \( P \) be a symmetric stochastic matrix. Let

\[
T^*_\epsilon = \min \{ k : \ P(\|y^k_w\| \geq \epsilon) \leq \epsilon \}.
\]

Then,

\[
\frac{1}{2} \log \frac{1}{2\epsilon} \leq T^*_\epsilon \leq \frac{2 \log \frac{1}{\epsilon}}{\log \left( \frac{1}{2} \left( 1 + \lambda^2(p) \right) \right)}.
\]

Some Implications.

- Lower bound for any \( G \). For any symmetric \( P \), \(-1 \leq \lambda_n(p) \leq \lambda_{n-1}(p) \leq \cdots \leq \lambda_1(p) = 1\).

The Trace \( (p) = \sum P_{ii} = \sum_{i=1}^{n} \lambda_i(p) \).

Since \( P \geq 0 \), Trace \( (p) \geq 0 \).

Therefore, \( 1 + \sum_{i=1}^{n} \lambda_i(p) \geq 0 \).

That is, \( (n+1) \lambda_2(p) \geq -1 \Rightarrow \lambda_2(p) \geq -\frac{1}{n+1} \).

Thus, \( \frac{1}{2} \left( 1 + \lambda_2(p) \right) \geq \frac{1}{2} \left( 1 - \frac{1}{n+1} \right) \).

Therefore, \( T^*_\epsilon \geq \frac{1}{2} \frac{\log \frac{1}{2\epsilon}}{\log \left( \frac{2(n+1)}{n-2} \right)} \approx \Omega \left( \log \frac{1}{\epsilon} \right) \).
- Thus,

\[ T^R_E = \Theta \left( \log \frac{1}{\varepsilon} + T_E \right). \]

That is, \( \Theta \left( \log \frac{1}{\varepsilon} \right) \) is the additive penalty due to "matching constraint".

- For expander graphs, \( \lambda_2(p) < -\varepsilon \);
  therefore \( T^R_E = \Theta \left( \log \frac{1}{\varepsilon} \right) \).

- For ring graph, \( \lambda_2(p) = 1 - \frac{1}{n^2} \).
  Therefore \( T^R_E = \Theta \left( n^2 \log \frac{1}{\varepsilon} \right) \).