Ad hoc lower bounds for mixing times

If one can find $x \in E, A \subset E$ and $t_0 \geq 0$ such that

$$\left| P^T(x, A) - \pi(A) \right| > \varepsilon,$$

then $||P^T(x, \cdot) - \pi||_{\mathrm{TV}} > \varepsilon$, and, from the definition of mixing times, this implies that $t_{\mathrm{mix}}(\varepsilon) t_0$.
In the case $\varepsilon = 1/4$, note that it is enough, for example, to choose A such that $\pi(A^c) \geq 1/2$, $P^t(x, A) > 3/4$.

Ad hoc lower bounds for mixing times

To prove such lower bound, one needs a good enough understanding of the stationary probability π, and of the law of the chain started at x at time T, $P^T(x, \cdot)$.
It is often the case that when T is taken small enough, the chain will typically remain in a well-choosen A neighborhood of x, while the stationary distribution is more evenly spread on the whole of E.
In the case of a reversible chain / conductance model on a graph $\mathcal{G}$, such a neighborhood can be often taken as a ball centered at x for the graph distance with appropriate radius.
The same idea still applies in the general case, only the underlying graph becomes oriented (it is that of the diagramm of the chain).

Example : lazy SSRW on $\frac{\mathbb{Z}}{n\mathbb{Z}}$

Let X be the lazy SSRW on $\frac{\mathbb{Z}}{n\mathbb{Z}}$. By symmetry considerations we may as well consider the walk started at $x = 0$.
Consider (the open ball) $A = B(0, n/4)$. Since the invariant probability is uniform $\pi(A) \leq 1/2$.
Note that if Y is the lazy SSRW on $\mathbb{Z}$ started at 0 then $X = Y \bmod n$. The asymptotic behaviour of Y_k is well known : $\mathbb{E}[Y_1] = 0, \mathrm{Var}[Y_1] = 1/2$, so by central limit theorem $\frac{Y_k}{\sqrt{k}}$ converges in law to $\frac{Z}{\sqrt{2}}$ with $Z \sim \mathcal{N}(0, 1/2)$. This ensures that the probability that $Y_{\delta n^2}$ is at a distance less than $n/4$ from the origin converges to $\mathbb{P}(|Z| < \frac{\sqrt{2}}{4\sqrt{\delta}})$. Glancing at a table for the normal distribution, we see that this probability is greater that $3/4$ already for $\frac{\sqrt{2}}{4\sqrt{\delta}} > 1.16$, so, e.g., $\delta = \frac{1}{8}$ will work for us.
It must be obvious by now that we are not trying to optimize the constant factors in this reasoning.

Of course if $Y_{\delta n^2}$ remains close to the origin, it must be true as well of $X_{\delta n^2}$, and we deduce that for large enough values of n,

$$\mathbb{P}_0(|X_{\frac{n^2}{16}}| \leq \frac{n}{4}) > 3/4.$$

This shows that $P^t(0, A) > 3/4$ and we conclude that $t_{\mathrm{mix}} \geq \frac{n^2}{16}$ for large enough values of n.
As we have seen in the previous paragraph, $t_{\mathrm{mix}} \leq n^2$, so we have found the order of magnitude of t_{mix}.
It is not so interesting to go beyond : more precise computations along the same analysis would provide $t_{\mathrm{mix}}(\varepsilon) \sim C(\varepsilon)n^2$, however $C(\varepsilon) \to 0$ when $\varepsilon \to 1$, and $C(\varepsilon) \to +\infty$ when $\varepsilon \to 0$.

Example : lazy SRW on hypercube

Let X be the lazy SRW on $H_d := \left(\frac{\mathbb{Z}}{2\mathbb{Z}}\right)^d$. By symmetry considerations we may as well consider the walk started at $(0, ..., 0)$. We also fix $\varepsilon \in (0,1)$.
Consider $A_K = \{x \in H_d : \sum_{i=1}^d x_i \leq n/2 - K\sqrt{n}\}$. Obviously, since the invariant probability is uniform and more than half of the points in H_d have coordinates which sum up at or above $n/2$, $\pi(A_K) \leq \frac{1}{2}$ as long as $K \geq 0$. Better yet, choosing a random point Y in H_d according to π amounts to choosing each of its coordinates independently at random according to $\mathrm{Ber}(1/2)$ distribution, so one has for such point $\sum_{i=1}^d Y_i \sim \mathrm{Bin}(n, 1/2)$. By CLT,

$$\pi(A_K) = \mathbb{P}(\sum_{i=1}^d Y_i \leq \frac{n/2}{-} K\sqrt{n}) \to \mathbb{P}(Z \leq -\frac{K}{2})$$

where $Z \sim \mathcal{N}(0,1)$.
In the end, we may choose $\pi(A_K)$ smaller than $\varepsilon/2$, say, as long as $K = K(\varepsilon)$ is choosen large enough.

Example : lazy SRW on hypercube

Now recall that the walk X can be performed by, at each step, choosing a coordinate at random and replacing it with a Bernoulli(1/2) variable. The number of coordinates which have been updated at time $t \in \mathbb{N}$ is exactly the number of coupons, say C_t, collected by the collector at time t, and, writing S_t for the sum of the coordinates of X_t, $S_t = \sum_{i=1}^{C_t} \xi_i$, where $(\xi_i)_{i \geq 1}$ is a sequence of i.i.d. Bernoulli(1/2). Thus

$$\mathbb{P}(S_t \leq \frac{n}{2} - K\sqrt{n}) \geq \mathbb{P}(C_t \leq n - L\sqrt{n})\mathbb{P}\left(\sum_{i=1}^{n-L\sqrt{n}} \xi_i \leq \frac{n}{2} - K\sqrt{n}\right).$$

Again by CLT, for $K = K(\varepsilon)$ as above, we may choose $L = L(\varepsilon)$ large enough so that the second term in the above product remains above $1 - \varepsilon/4$, say. As for the first term, we proceed as in the analysis of the coupon collector problem to show that it also remains above $1 - \varepsilon/4$ as long as $t \leq n\log(n) - C(\varepsilon)n$.

Example : lazy SRW on hypercube

In the end we have proven that for some constant $C(\varepsilon)$ depending on ε,

$$P^{n\log n - C(\varepsilon)n}((0, ..., 0), A) \leq 1 - \varepsilon/2.$$

while $\pi(A) < \varepsilon/2$, hence $t_{\mathrm{mix}}(\varepsilon) \leq n\log(n) - C(\varepsilon)n$.
Thanks to the upper bound of paragraph 3.3 we conclude that $t_{\mathrm{mix}}(\varepsilon) \sim n\log(n)$ as $n \to \infty$ whatever the value of $\varepsilon \in (0, 1)$.
The fact that the asymptotics of $t_{\mathrm{mix}}(\varepsilon)$ does not depend on ε is known as the *cutoff* phenomenon.

Diameter bound

Diameter bound is based exactly on the same idea, namely that $t_{\mathrm{mix}} \geq \mathrm{diam}(\mathcal{G})/2$, where $\mathrm{diam}(\mathcal{G}) = \max_{x,y \in E} d(x,y)$, and d is the graph distance.

It suffices to choose x and y such that $d(x,y) = \mathrm{diam}(\mathcal{G})$, it must be either $B(x, d(x,y)/2)$ or $B(y, d(x,y)/2)$ must have mass less than $1/2$ for invariant probability, so we can choose A to be this set.

Lower bound through Cheeger constant : heuristics

One can also prove a lower bound for the mixing time involving the inverse of the Cheeger constant. The main idea is to cut the graph at its bottleneck, that is, choosing the set A of invariant probability less than $1/2$ by minimizing, at stationarity, the proportion of $\pi(A)$ which escapes A. The Cheeger constant Φ^* precisely is this minimal proportion. Now starting the chain from the restriction π_A of its stationary distribution on A, it is possible to show that the Cheeger constant appears as the total variation distance between the distribution of the chain at times 0 and 1. Because $||\pi_A P^t - \pi_A P^{t+1}||_{TV}$ can only decrease with t, it must be that after $1/4\Phi^*$ steps, at most half of this initial distribution may have traveled to A^c, yielding $t_{\mathrm{mix}} \geq \frac{1}{4\Phi^*}$.
Of course, bounds involving the Cheeger constant are well-suited for graphs exhibiting a clear bottleneck.

Further results

When the chain is reversible, aperiodic, we have seen in 3.4 that one may essentially bound t_{mix} with t_{relax}.
There are deeper and fruitful spectral techniques which we will not investigate in this course, such as Wilson's method (altough it requires being able to produce an eigenfunction, and ideally, the one realizing the absolute spectral gap). More often than not, an explicit spectral decomposition is not available. It is possible however, to bound above the spectral gap by twice the bottleneck ratio Φ^*.
See chapter XIII of Levin,Peres,Wilmer.