Random walks and algorithms - handout exercises

1. (M) Prove that there exist constants $d \in \mathbb{N}$ and $c > 0$ such that the following holds: let $G(n,d)$ be a uniform random d-regular graph on n vertices. Let $\lambda_1(G) \geq \lambda_2(G) \geq ... \geq \lambda_n(G)$ be the eigenvalues of its adjacency matrix. Then

$$\mathbb{P}\Big(\lambda_1(G(n,d)) - \lambda_2(G(n,d)) > cd\Big) \to 1, \quad \text{as } n \to \infty$$

   Remark: Picking a uniform graph out of all possible d-regular graphs is not a very tractable construction. It turns out that there is a constructive way to do it, which is much easier to work with. The idea is to consider the n vertices, each connected to d half-edges and take a uniform perfect matching on these nd half-edges to determine how to connect them. If we condition on the resulting graph to be simple (hence, to have no loops and no double-edges), the graph we get turns out to be a uniform sample of a $G(n,d)$ graph (why?).

2. Let K be a convex body in $\mathbb{R}^d$, suppose that it is encapsulated between a ball of radius 1 and a ball of radius n^{10}. Define $b(K) = \frac{\int_K x dx}{\int_K dx}$ to be the barycenter of K. Suppose you're given access to K via a membership oracle (= a black box which answers the question "is $x \in K$).

   (a) (M) Find a polynomial time algorithm to approximate $b(K)$ up to a small error. Hence prove that for every $\varepsilon > 0$ the algorithm generates a point Y such that $\mathbb{P}(|Y - b(K)| < \varepsilon) > 1 - o(1)$ using $poly(n, \varepsilon)$ steps.

   (b) (M) Do the same thing for estimating the covariance matrix of K, hence, if X is a random vector uniformly distributed in K, find a polynomial time algorithm that approximates its covariance matrix

$$Cov(X) := \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]$$

   in the sense that the algorithm output a matrix M such that $2Cov(X) \succeq M \succeq Cov(X)$ with high probability.

3. (M) Let $f(x)$ be a density in $\mathbb{R}^n$ with bounded support. We define the "hit-and-run" Markov chain (whose state space is $\mathbb{R}^n$) as follows: suppose that we are at some position X_t. Then X_{t+1} is generated by: (i) Generating a uniform random direction $\theta \in \mathbb{S}^{n-1}$. (ii) Considering the one-dimensional affine subspace $X_t + \text{span}(\theta)$, and taking X_t to be a random point in this subspace picked according to the density proportional to $f(x)$.

   Prove that the stationary distribution of this Markov chain is $f(x)$. Bonus: suppose that $f(x)$ is the indicator of a convex body. Can you come up with an idea how to prove that the chain mixes in polynomial time?

4. (M) Define the $1-3$-tree as the following infinite tree. Fix a root with two children. Now suppose that all nodes except for the root has either one or three children, so that the k-th level has 2^k nodes (hence half the nodes on each level have 3 descendants and the other half have 1 descendant) and so that if the tree is drawn on a plane, the degree of the nodes is monotone from left to right (thus, a node of degree 4 always has an ancestor of degree 4, unless it's the root). Is the random walk on this tree transient or recurrent? How many distinct infinite rays does this tree have?

5. (R) Consider the following self-interacting random walk on $\mathbb{Z}^2$: whenever a vertex is reached by the walker for the first time, the walker takes a step either right or left with probabilities $(1/2, 1/2)$ and whenever the walker reaches a vertex which has already been visited, the walker goes either up or down with probabilities $(1/2, 1/2)$. Is this random walk transient or recurrent? Is there limiting shape for the range? (hence, is it true that after rescaling the space properly the set of vertices visited by the random walk before time n looks like an interval?).

6. (E) Let $Z_1, ..., Z_T$ be random variables. Let $X_1, ..., X_T$ be real valued random variables such that X_i is a (deterministic) function of Z_i and such that for every $1 \leq t < T$ and $z_1, .., z_t$, one has that the increment $X_{t+1} - X_t$, conditioned on the event $\{Z_1 = z_1, ..., Z_t = z_t\}$, is a Gaussian with mean zero and variance at most 1. Prove that

$$\mathbb{P}(|X_T| > \lambda\sqrt{T}) < 2\exp(-\lambda^2/2), \ \forall \lambda > 0.$$

7. Let $X_0, X_1, X_2, ...$ be an irreducible and aperiodic Markov chain with stationary measure π and let $S_0, S_1, S_2, ...$ be the associated evolving set process with $S_0 = \{X_0\} = \{x_0\}$ being some deterministic point. Denote by $v(t)$ the total-variation distance between X_t and the stationary measure.

   (a) (E) Prove that $\mathbb{P}(X_t = y) = \frac{\pi(y)}{\pi(x_0)}\mathbb{P}(y \in S_t)$ for all y, t.

   (b) (M) Let T be the first time that S_t is either the empty-set or the whole state-space. Prove that
   $$v(t) \leq \frac{1}{\pi(x_0)}\mathbb{P}(T > t).$$

   (c) (H) Prove that
   $$v(t) \leq \frac{C}{\pi(x_0)}\mathbb{E}\sqrt{\min(\pi(S_t), 1 - \pi(S_t))}.$$

   (d) (H) Can you think of a way to use this idea to show that a Cheeger inequality implies mixing bounds?

8. (M) Suppose that you're given N directions $\theta_i \in \mathbb{S}^{n-1}$, for $1 \leq i \leq N$. Suggest an algorithm to approximate the spherical surface area of the set

$$S = \mathbb{S}^{n-1} \cap \{x; \langle x, \theta_i \rangle > 0, \ \forall 1 \leq i \leq N\}$$

up to a small multiplicative error, with complexity $\mathrm{poly}(nN)$.

9. (H) The goal of this question is to find a variant of the constructive hypergraph discrepancy algorithm for the case where, instead of ± 1, one assigns to each point a value from some given subset of $\mathbb{R}^d$.

Let $\mathcal{L} = \{s_1, .., s_k\} \subset \mathbb{R}^d$ such that $\|s_i\|_2 \le 1$ for all i and such that $\mathcal{L}$ contains the origin in its convex hull. Let $V = [n]$ be a vertex set, and let $\mathcal{S} \subset 2^V$ be a family of m subsets. The goal is to find a "coloring" $\xi : V \to \mathcal{L}$ such that for all $S \in \mathcal{S}$,

$$\left\| \sum_{v \in S} \xi(v) \right\|_2 < D$$

for a number D as small as possible. Find a way to tweak the algorithm we saw in class so that this can be done for

$$D = C(d)\sqrt{n \log((m+n)/n)}$$

where $C(d)$ depends only on d. What is the best dependence on d that can be attained?

10. (E) Let $S_1, S_2, ...$ be the evolving set process associated with a Markov chain with state space $\mathcal{S}$. Denote by $K(S, S_1)$ its transition kernel. Consider the Doob transform associated to the process (what we called in class "modified process") with transition kernel

$$\tilde{K}(S, S_1) = K(S, S_1)\frac{\mu(S_1)}{\mu(S)}$$

where μ is the stationary measure of the Markov chain. Prove that the modified process is the same as the original process conditioned on the event $\{S_i \neq \emptyset \text{ for all } i\}$.

11. (M) In [Andersen-Peres, Proposition 5] (the paper on the evolving sets process), it is shown that for most points in a set A whose conductance is ϕ, the escape time from A (hence the typical time it takes a random walk to exit the set A) is bounded by $O(1/\phi)$. Note that this bound is *not* tight, for example, if A is an interval in $\mathbb{Z}$ (in this case, the escape time is of the order $1/\phi^2$). Can you find a family of bounded-degree graphs and respective subsets A where this bound is actually tight? Bonus: (H?) can you think of a natural sufficient condition on the graph and on the set under which this bound can be improved to $1/\phi^2$?

12. Consider the random walk $X_0, X_1, ..$ on $K = \{-1, 1\}^d$ where $X_0 = (1, ..., 1)$ and at each step, a coordinate is selected at random, and is then toggled with probability $\frac{1}{2}$. Let u be the uniform measure on K and define $v_d(t)$ to be the total variation distance between the law of X_t and u.

    (a) (E) Prove that for a universal constant $C > 0$ large enough one has that

    $$\lim_{d \to \infty} v_d(Cd \log d) = 0$$

    (hence, the random walk mixes in less than $Cd \log d$ steps).

    (b) (H) Prove that for every $\delta > 0$ one has

    $$\lim_{d \to \infty} v_d\big((1-\delta)\tfrac{1}{2} d \log d\big) = 1$$

    and on the other hand

    $$\lim_{d \to \infty} v_d\big((1+\delta)\tfrac{1}{2} d \log d\big) = 0$$

    This phenomenon is called "cutoff", which intuitively means that the mixing of the random walk occurs in a small window around the time $\frac{1}{2}\log_2 d$.

13. (M) Let $K \subset \mathbb{R}^d$ be a convex body. Assume that $D \subset K \subset RD$, where D is the unit Euclidean ball. Let $f : K \to [-M, M]$ be a function with Lipschitz constant L. A. Provided access to a membership oracle to K and a function which evaluates f in time $O(1)$ for a point $x \in K$, find a way to approximate $\int_K f(x)dx$ up to an additive error of ε (with high probability), with complexity $poly(d, \varepsilon, L, M, R)$. B. What happens if one omits the assumption that f is L-Lipschitz? How should the algorithm be modified to overcome this?

14. (M) The Binary Branching Random walk is defined as follows: a single particle starts from a prescribed vertex (say, the origin). At every time step, every particle duplicates itself, dividing into two particles, which will now each do an independent random walk, so that after t iterations one has 2^t particles, which all do an independent step. Prove that this random walk is recurrent in $\mathbb{Z}^3$ in the sense that there are infinite times in which the origin will be occupied by some particle. Is it true that from some point in time and on the origin will *always* be occupied?

15. Let G be a a connected infinite graph with degrees bounded by d, rooted at some vertex v_0. Let $v_0 = X_0, X_1, X_2, ...$ be a random walk starting from v_0. Use evolving sets to prove the following claims:

    (a) (M) For every time $t > 0$ one has

    $$\mathbb{P}(X_t = v_0) < \frac{c(d)}{\sqrt{t}}$$

    for a constant $c(d)$ depending only on d.

    (b) (M) Suppose now that G is an expander in the sense that for every $A \subset G$ one has $\frac{|\partial A|}{|A|} > \phi > 0$. Prove that for all $t > 0$,

    $$\mathbb{P}(X_t = v_0) < C(d, \phi) \exp(-tc(d, \phi)).$$

    with $C(d, \phi), c(d, \phi) > 0$.

    (c) (H) Suppose that you don't know what G is, you only know that it has a bounded degree. You're given the sequence $t_1, t_2, ...$ of return times of X_t to v_0. Can you determine whether or not the graph G is finite or infinite only by looking at this sequence?