CS5330: Assignment for Week 10

Due: Tuesday, 7th Apr 2020.

Here are solution sketches to the Week 10 problems. If anything is unclear, please talk to me or your TA.

1. Consider the graph on $\mathbb{Z}_n^d = \{0, 1, \ldots, n-1\}^d$ where vertices $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ are adjacent if for some $j \in [d]$, $x_j - y_j = \pm 1 \mod n$ but for all $i \neq j$, $x_i = y_i$. The lazy random walk on $\mathbb{Z}_n^d$ behaves as follows: if the walk is currently at node x, then with probability 1/2, it stays at x; otherwise, it picks $i \in [d]$ uniformly at random and conditioned on that choice of i, it moves to[1] $(x + e^i) \mod n$ with probability 1/2 and $(x - e^i) \mod n$ with probability 1/2.

   Generalize the analysis shown in class for the lazy random walk on the cycle ($\mathbb{Z}_n$) to show that for the lazy random walk on $\mathbb{Z}_n^d$, $t_{\text{mix}}(\epsilon) = O(d^2 n^2/\epsilon)$.

   Fix any two vertices x and y. Consider the following coupling (X_t, Y_t) between two lazy random walks on $\mathbb{Z}_n^d$, one starting at x and the other starting at y. Choose a coordinate $j \in [d]$ uniformly at random. If the two walks agree on the j'th coordinate, then with probability 1/2, set $X_{t+1} = X_t$ and $Y_{t+1} = Y_t$, with probability 1/4, set $X_{t+1} = X_t + e^j \pmod n$ and $Y_{t+1} = Y_t + e^j \pmod n$, and with probability 1/4, set $X_{t+1} = X_t - e^j \pmod n$ and $Y_{t+1} = Y_t - e^j \pmod n$. If the two walks disagree on the j'th coordinate, then randomly choose one of the two walks to move ($+e^j \pmod n$ with probability 1/2 and $-e^j \pmod n$ with probability 1/2), and leave the other one fixed. It is easy to see that this is a valid coupling.

   If we restrict attention to the j'th coordinate and only to the steps when j is selected, then this is exactly the coupling of the n-cycle studied in class. Since j is selected with probability $1/d$, the j'th coordinate couples in $O(dn^2)$ steps in expectation. Therefore by Markov, we get that the probability that j'th coordinate is not coupled after $O(d^2n^2/\varepsilon)$ steps is ε/d. So, by the union bound, with probability at most ε, all the coordinates are not the same after $O(d^2n^2/\varepsilon)$ steps.

2. Read Theorem 12.6 (restated below in the notation used in class) and its proof in MU.

[1] e^i is the d-dimensional vector that is 0 everywhere but 1 at the i'th coordinate

Theorem 1. *Let $\boldsymbol{P}$ be the transition matrix for a finite, irreducible, aperiodic Markov chain M_t with $t_{\mathrm{mix}}(c) \le T$ for some $c < 1/2$. Then, for this Markov chain, $t_{\mathrm{mix}}(\epsilon) \le \lceil \ln \epsilon / \ln(2c) \rceil T$.*

Use it to improve the dependence of the mixing time on ϵ for the lazy walk on the cycle and the lazy walk on $\mathbb{Z}_n^d$ in Problem 1 above.

By the above theorem, $t_{\mathrm{mix}}(\varepsilon) = t_{\mathrm{mix}}(1/3) \cdot O(\log 1/\varepsilon)$. Hence, for $\mathbb{Z}_n^d$, the improved bound on the mixing time is $O(d^2 n^2 \log(1/\varepsilon))$.

3. (Exercise 11.6 in MU) The problem of counting the number of solutions to a knapsack instance can be defined as follows: Given items with sizes $a_1, a_2, \ldots, a_n > 0$ and an integer $b > 0$, find the number of vectors $(x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ such that $\sum_{i=1}^n a_i x_i \le b$. The number b can be thought of as the size of a knapsack, and the x_i denote whether or not each item is put into the knapsack. Counting solutions corresponds to counting the number of different sets of items that can be placed in the knapsack without exceeding its capacity.

   Argue that if we have an FPUAS for the knapsack problem, then we can derive an FPRAS for the problem. To set the problem up properly, assume without loss of generality that $a_1 \le a_2 \le \cdots \le a_n$. Let $b_0 = 0$ and $b_i = \sum_{j=1}^i a_j$. Let $\Omega(b_i)$ be the set of vectors $(x_1, \ldots, x_n) \in \{0,1\}^n$ that satisfy $\sum_{j=1}^n a_j \le b_i$. Let k be the smallest integer such that $b_k \ge b$. Consider the equation:

$$|\Omega(b)| = \frac{|\Omega(b)|}{|\Omega(b_{k-1})|} \times \frac{|\Omega(b_{k-1})|}{|\Omega(b_{k-2})|} \times \cdots \times \frac{|\Omega(b_1)|}{|\Omega(b_0)|} \times |\Omega(b_0)|$$

   You will need to argue that $|\Omega(b_{i-1})|/|\Omega(b_i)|$ is not too small. Specifically, argue that $|\Omega(b_i)| \le (n+1) \cdot |\Omega(b_{i-1})|$.

   Clearly, $\Omega(b_0) = \varnothing$ and $\Omega(b_0) \subseteq \Omega(b_1) \subseteq \cdots \subseteq \Omega(b_k)$.

   Let's argue that $|\Omega(b_i) \setminus \Omega(b_{i-1})| \le n \cdot |\Omega(b_{i-1})|$. For any vector $x \in \Omega(b_i) \setminus \Omega(b_{i-1})$, let i^* be the smallest number $\ge i$ such that $x_{i^*} = 1$. Note that i^* must exist because otherwise $\sum_j x_j a_j \le b_{i-1}$, contradicting our assumption $x \notin \Omega(b_{i-1})$. Define $f_i : \Omega(b_i) \setminus \Omega(b_{i-1}) \to \Omega(b_{i-1})$ to be the vector that is the same as x on all coordinates except i^* where it is zero. $f_i(x) \in \Omega(b_{i-1})$ because its size is at most $b_i - a_{i^*} \le b_{i-1}$. Additionally, since x and $f_i(x)$ differ at exactly one out of the n coordinates, $|f_i^{-1}(y)| \le n$ for any $y \in \Omega(b_{i-1})$. So, $|\Omega(b_i) \setminus \Omega(b_{i-1})| \le n \cdot |\Omega(b_{i-1})|$, and therefore, $|\Omega(b_i)| \le (n+1) \cdot |\Omega(b_{i-1})|$.

   Suppose we use the knapsack FPAUS to sample from $\Omega(b_i)$ with the guarantee that its output distribution is $\frac{\varepsilon}{4k(n+1)}$-close to the uniform distribution in TV distance. Suppose you draw m samples from this sampler, and let Z be the expected number of the samples that lie in $\Omega(b_{i-1})$. Using the definition of TV distance, $\mathbb{E}[Z] \ge m\left(\frac{|\Omega(b_{i-1})|}{|\Omega(b_i)|} - \frac{\varepsilon}{4k(n+1)}\right) \ge m\frac{|\Omega(b_{i-1})|}{|\Omega(b_i)|}\left(1 - \frac{\varepsilon}{4k}\right)$, where we used that $\frac{|\Omega(b_{i-1})|}{|\Omega(b_i)|} \ge \frac{1}{n+1}$. Similarly, $\mathbb{E}[Z] \le m\frac{|\Omega(b_{i-1})|}{|\Omega(b_i)|}\left(1 + \frac{\varepsilon}{4k}\right)$, in particular smaller than $5m/4$. Finally, using the Chernoff bound, we get $\Pr[|Z - \mathbb{E}[Z]| > \frac{\varepsilon}{4k}\mathbb{E}[Z]] \le 2\exp\left(-\frac{\varepsilon^2}{48k^2}\frac{5m}{4}\right)$.

   If $m = O(k^2 \varepsilon^{-2} \log(k/\delta))$, then $\left(1 - \frac{\varepsilon}{4k}\right)\mathbb{E}[Z] \le Z \le \left(1 - \frac{\varepsilon}{4k}\right)\mathbb{E}[Z]$ with probability at least $1 - \delta/k$. Moreover, using the above bounds on $\mathbb{E}[Z]$, we get that with probability at least $1 - \delta/k$, Z is within a multiplicative factor $\left(1 \pm \frac{\varepsilon}{4k}\right)^2 = \left(1 \pm \frac{\varepsilon}{2k}\right)$ of $|\Omega(b_{i-1})|/|\Omega(b_i)|$.

Applying this approach for all $i \in [k]$, we get that with probability $1 - \delta$, we get estimates Z_i such that each is within a $(1 \pm \varepsilon/2k)$ multiplicative factor of $|\Omega(b_{i-1})|/|\Omega(b_i)|$. Our final estimate $\frac{1}{Z_k} \times \frac{1}{Z_{k-1}} \times \cdots \times \frac{1}{Z_1} \times 1$ is a $(1 \pm \varepsilon)$ multiplicative approximation of $|\Omega(b)|$ with probability $1 - \delta$.