1. **Design Markov Chain**

(a) *Proof.* We prove that graph G is strongly connected. Consider $\sigma, \sigma' \in V$, if $\sigma = \sigma'$, then path p is the self-loop. If $\sigma \neq \sigma'$, then we can construct a path p from σ to σ'. In each step of the path, let i be the first index such that $\sigma_i \neq \sigma'_i$, we add the edge that swaps σ_i with σ_j for some index j where $\sigma_j = \sigma'_i$. We claim that such index $j \geq i$, and after the swap, the first i integers of the destination vertex equals $\sigma'_1 \cdots \sigma'_i$. For $i = 1$, every other $j > 1$, and we get σ'_1 at index 1 after one swap, so claim is true. Now we prove claim is true for $i > 1$. By induction hypothesis, the first $i-1$ integers of the vertex is $\sigma'_1 \cdots \sigma'_{i-1}$ and that $\sigma'_i \neq \sigma'_k = \sigma_k$ for $k = 1, \cdots, i-1$ since σ' is a permutation. So there is some $j \neq k$ for $k = 1, \cdots, i-1$ such that $\sigma_j = \sigma'_i$ and we swap σ_j with σ_i such that the first i integers in the permutation is equal to $\sigma'_1 \cdots \sigma'_i$, hence the claim is true in the inductive case. By $i = n$, we have $\sigma'_1 \cdots \sigma'_n$ as the resulting vertex. Since choice of σ, σ' arbitrary, we can always find such path, so graph is strongly connected. $\square$

(b) To find the transition probabilities P on Markov chain on G such that stationary distribution p has $p_\sigma \propto 2^{-inv(\sigma)}$, we first note that the maximum number of degrees is same for all vertices, specifically $r = \binom{n}{2}$, since we have edge connecting two vertices by arbitrarily swapping 2 numbers. Therefore,

$$P_{ij} = \frac{1}{\binom{n}{2}} \min\{1, 2^{inv(\sigma_1) - inv(\sigma_2)}\} \qquad P_{ii} = 1 - \sum_{j \neq i} P_{ij}$$

2. **Streaming**

(a) We express $\mathsf{E}\{c\}$ in terms of $f_1 = |\{t : \sigma_t = i\}|$. Let c_j be arbitrary where $j \in [k]$, then we can express c_j as follows

$$c_j = f_1 h_j(1) + \cdots + f_n h_j(n) = \sum_{i=1}^{n} f_i h_j(i)$$

Also note that $(h_j(i))^2 : [n] \to \{1\}$ outputs 1 with probability 1, so then $\mathsf{E}\left\{(h_j(i))^2\right\} = 1$. Also we have that $\mathsf{E}\{h_j(i)\} = 0$ for arbitrary hash function h_j and arbitrary input $i = 1, \cdots, n$.

$$\begin{aligned}
\mathsf{E}\left\{c_j^2\right\} &= \mathsf{E}\left\{(f_1 h_j(1) + \cdots + f_n h_j(n))^2\right\} \\
&= \mathsf{E}\left\{\sum_{i=1}^{n} f_i^2 (h_j(i))^2 + \sum_{i' \neq i}^{n} f_{i'} f_i h_j(i') h_j(i)\right\} \\
&= \sum_{i=1}^{n} f_i^2 \mathsf{E}\left\{(h_j(i))^2\right\} + \mathsf{E}\left\{\sum_{i' \neq i}^{n} f_{i'} f_i h_j(i') h_j(i)\right\} \\
&= \sum_{i=1}^{n} f_i^2 + \sum_{i' \neq i}^{n} f_{i'} f_i \mathsf{E}\left\{h_j(i') h_j(i)\right\} \\
&= \sum_{i=1}^{n} f_i^2 + \sum_{i' \neq i}^{n} f_{i'} f_i \mathsf{E}\left\{h_j(i')\right\} \mathsf{E}\left\{h_j(i)\right\} \qquad (h_j(i), h_j(i') \text{ independent}) \\
&= \sum_{i=1}^{n} f_i^2
\end{aligned}$$

So then, $\mathsf{E}\{c\} = \mathsf{E}\left\{\frac{1}{k}\left(c_1^2 + \cdots + c_k^2\right)\right\} = \frac{1}{k}\sum_{j=1}^{k} \mathsf{E}\left\{c_j^2\right\} = \frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{n} f_i^2 = \sum_{i=1}^{n} f_i^2$

(b) Note

$$(1-\epsilon)\mathsf{E}\{c\} \le c \le (1+\epsilon)\mathsf{E}\{c\} \iff |\,\mathsf{E}\{c\} - c\,| \le \epsilon \mathsf{E}\{c\}$$

So proving $\mathsf{P}\left((1-\epsilon)\mathsf{E}\{c\} \le c \le (1+\epsilon)\mathsf{E}\{c\}\right) \ge \frac{1}{2}$ is equivalent to

$$\mathsf{P}\left(|\,\mathsf{E}\{c\} - c\,| \le \epsilon \mathsf{E}\{c\}\right) \ge \frac{1}{2} \qquad \text{or} \qquad \mathsf{P}\left(|\,\mathsf{E}\{c\} - c\,| > \epsilon \mathsf{E}\{c\}\right) < \frac{1}{2}$$

We note the previous expression can be reformulated with chebyshev's inequality

$$\mathsf{P}\left(|\,\mathsf{E}\{c\} - c\,| > \epsilon \mathsf{E}\{c\}\right) < \frac{\mathsf{V}\{c\}}{\epsilon^2 \mathsf{E}\{c\}^2}$$

Note from previous $\mathsf{E}\left\{c_j^2\right\} = \mathsf{E}\left\{c\right\} = \sum_{i=1}^{n} f_i^2$. We then compute $\mathsf{V}\left\{c\right\}$

$$\begin{aligned}
\mathsf{V}\left\{c\right\} &= \frac{1}{k^2}\sum_{j=1}^{k} \mathsf{V}\left\{c_j^2\right\} \\
&= \frac{1}{k^2}\sum_{j=1}^{k} \mathsf{E}\left\{c_j^4\right\} - \mathsf{E}\left\{c_j^2\right\}^2 \\
&\le \frac{1}{k^2}\sum_{j=1}^{k} 3\mathsf{E}\left\{c_j^2\right\}^2 - \mathsf{E}\left\{c_j^2\right\}^2 \\
&= \frac{2}{k^2}\sum_{j=1}^{k} \mathsf{E}\left\{c_j^2\right\}^2 \\
&= \frac{2}{k}\mathsf{E}\left\{c\right\}^2
\end{aligned}$$

So then

$$\mathsf{P}\left(\left|\,\mathsf{E}\left\{c\right\} - c\,\right| > \epsilon\mathsf{E}\left\{c\right\}\right) < \frac{\mathsf{V}\left\{c\right\}}{\epsilon^2\mathsf{E}\left\{c\right\}^2} \le \frac{\frac{2}{k}\mathsf{E}\left\{c\right\}^2}{\epsilon^2\mathsf{E}\left\{c\right\}^2} = \frac{2}{k\epsilon^2}$$

Let $k = \frac{4}{\epsilon^2}$ such that $\mathsf{P}\left(\left|\,\mathsf{E}\left\{c\right\} - c\,\right| > \epsilon\mathsf{E}\left\{c\right\}\right) < \frac{1}{2}$, which satisfies the probability constraint given in the problem specification.

3. **Linear Programming**

Proof. For any constraint with k terms where $k > 3$, i.e. corresponding row i of A has k nonzero entries, we can introduce $k-3$ new variable $\{y_i\}_{i=1}^{k-3}$ and convert the original constraint

$$\sum_{j=1} A_{ij}x_1 \le b_i$$

to an equivalent form.

$$\begin{aligned}
A_{il_1}x_{l_1} + A_{il_2}x_{l_2} + y_1 &\le b_i \\
y_1 &= A_{il_3}x_3 + y_2 \\
&\cdots \\
y_{k-3} &= A_{il_{k-1}}x_{k-1} + A_{il_k}x_k
\end{aligned}$$

where $\{l_1, \cdots, l_k\}$ are the indices of nonzero entries in row i of A, i.e. $A_{il_j} \neq 0$ where $l_j \in \{l_1, \cdots, l_k\}$, where the equality can be expressed as inequality constraints as follows

$$\begin{aligned} y_{j-1} - A_{il_{j+1}} x_{j+1} - y_j &\leq 0 \\ -y_{j-1} + A_{il_{j+1}} x_{j+1} + y_j &\leq 0 \end{aligned}$$

Let y' be all the extra variables introduced to all constraints. Let D' be the matrix consisting of coefficients for the converted constraints, and e' be the rhs of the converted matrices. Then

$$y = \begin{pmatrix} x \\ y' \end{pmatrix} \qquad f = \begin{pmatrix} c \\ 0 \end{pmatrix} \qquad D = \begin{pmatrix} D' \\ I \end{pmatrix} \qquad e = \begin{pmatrix} e' \\ 0 \end{pmatrix}$$

The objective is equivalent since $f^T y = c^T x$. The constrants are equivalent. Computing y, f, D, e takes polynomial times, since we are simply converting m to $m(K-2)$ constraints where $K \leq n$ is maximum number of nonzero entries in any given constraint. $\square$