Question 1 a

Different algorithms exist for this part which are mostly equivalent. Following outlines the most common solution using submodular function minimization:

First, observe that $x_e \geq 0$ constraints can be naively checked in polynomial time. However, other constraint is defined over $S \subset V$ and thus there are exponentially ($2^{|V|}$) many to check. For these, define:

$$f(S) = \sum_{e \in \delta(S)} x_e - 2$$

Here, f is submodular. This can be shown either from its definition or by noticing that it is simply the cut-size function and a constant. We know that unconstrained submodular function minimization can be done in polnomial time. Let $u, v \in V$ and define $g(S) = f(S \cup \{u\})$. Then, find the minimum of:

$$\min_{S \subseteq V \setminus \{v\}} g(S)$$

for each (u, v) pair. This is achieved in polynomial time.

Another (equivalent) solution involves finding the min-weighted-cut of the graph induced by the candidate solution x, which can also be done in polynomial time.

Question 1 b

For finding the dual of the problem, one can define variables as all possible non-empty and strict subsets of V ($x_e \leftrightarrow y_S$). Hence, the dual LP becomes:

$$\begin{aligned} \max \quad & \sum_{S \subset V, S \neq \emptyset} 2y_S & \\ \text{s.t.} \quad & \sum_{S \subset V, S \neq \emptyset, e \in \delta(S)} y_S \leq c_e & \forall e \in E \\ & y_S \geq 0 & \forall S \subset V, S \neq \emptyset \end{aligned}$$

Question 2

First, we show the statement in the hint. In the question, the bias is given as "at least ε", but let us set the bias exactly ε since this is the worst case and we can still work it out.

Without loss of generality, consider a fixed coin. Let X_i be indicator where $X_i = 1$ if i-th flip of the coin is heads and $X_i = 0$ otherwise. Observe that X_i's are independent Bernoulli random variables. For a fair coin, $E[X_i^{fair}] = \frac{1}{2}$ and for a biased coin, $E[X_i^{biased}] = \frac{1}{2} + \varepsilon$. Assume that we flip this coin t times and define the sum of outcomes as: $Y = \sum_i^t X_i$.

Then, compute the expected value of the sum r.v.s:

$$E[Y^{fair}] = \sum_{i=1}^{t} E[X_i^{fair}] = \frac{t}{2}$$

$$E[Y^{biased}] = \sum_{i=1}^{t} E[X_i^{biased}] = \frac{t}{2} + \varepsilon t$$

With these expressions, we can try to estimate the bias of a coin, e.g. how much the sum of outcomes differ from the expected value for a fair coin. Now the question is: how likely will we fail? Here, a failure can be one of the following:

- Mistaking a fair coin for a biased coin: $Y^{fair} \geq \frac{t}{2} + \varepsilon t$
- Mistaking a biased coin for a fair coin: $Y^{biased} \leq \frac{t}{2}$

Using Chernoff Bounds (*) found in the lecture notes or equivalent forms, let us compute these probabilities.

- Probability of mistaking a fair coin:

$$P[Y^{fair} \geq (1 + \delta)E[Y^{fair}]] \leq \exp\left(-\frac{\delta^2}{2 + \delta} E[Y^{fair}]\right)$$

Let $\delta = 2\varepsilon$ as $(1+2\varepsilon)\frac{t}{2} = \frac{t}{2} + \varepsilon t$:

$$\begin{aligned} P[Y^{fair} \geq (1+2\varepsilon)(\frac{t}{2})] &\leq \exp\left(-\frac{4\varepsilon^2}{2+2\varepsilon}\cdot\frac{t}{2}\right) \\ &\leq \exp\left(-\frac{2}{3}\varepsilon^2 t\right) \qquad (\text{using } 0<\varepsilon<\frac{1}{2}) \end{aligned}$$

- Probability of mistaking a biased coin:

$$P[Y^{biased} \leq (1-\delta)E[Y^{biased}]] \leq \exp\left(-\frac{\delta^2}{2}E[Y^{biased}]\right)$$

Let $\delta = 2\varepsilon$ as $(1-2\varepsilon)(\frac{t}{2}+\varepsilon t) < \frac{t}{2}$:

$$\begin{aligned} P[Y^{biased} \leq (1-2\varepsilon)(\frac{t}{2}+\varepsilon t)] &\leq \exp\left(-\frac{4\varepsilon^2}{2}\cdot\left(\frac{t}{2}+\varepsilon t\right)\right) \\ &\leq \exp\left(-\varepsilon^2 t\right) \qquad (\text{using } 0<\varepsilon<\frac{1}{2}) \end{aligned}$$

Guided by the hint and the expressions above, set $t = \frac{C}{\varepsilon^2}\log n$ and substitute:

$$\begin{aligned} P[\text{mistaking fair coin}] &\leq \exp\left(-\frac{2}{3}\varepsilon^2\cdot\frac{C}{\varepsilon^2}\log n\right) = n^{-\frac{2}{3}C} \leq \frac{1}{10n} \\ P[\text{mistaking biased coin}] &\leq \exp\left(-\varepsilon^2\cdot\frac{C}{\varepsilon^2}\log n\right) = n^{-C} \leq \frac{1}{10n} \end{aligned}$$

Note that the inequalities above can be satisfied with sufficiently large selection of the constant C given that $n > 1$ (there is more than one coin).

To determine the fair coin, we flip each coin $t = \frac{C}{\varepsilon^2}\log n$ times and pick the one with smallest number of heads. Here, we may fail if we misclassify any of the n coins. Notice that this is a conservative approach since the coins are not independent - there are exactly 1 fair and $n-1$ biased coins. So, tighter bounds can be found but this suffices for our problem.

$$\begin{aligned} P[\text{success}] &= 1 - P[\text{fail}] \\ &= 1 - P[(\text{mistaking coin } 1)\vee\ldots\vee(\text{mistaking coin } n)] \\ &\geq 1 - \sum_{k=1}^{n} P[(\text{mistaking coin } k)] \qquad (\text{by union bound}) \\ &\geq 1 - n\cdot\frac{1}{10n} \\ &\geq \frac{9}{10} \end{aligned}$$

Recap: Alice flips each coin $\frac{C}{\varepsilon^2}\log n$, a total of $O(\frac{1}{\varepsilon^2}n\log n)$ flips, and selects the one which resulted in the smallest number of heads. With $> 90\%$ probability she will find the fair coin.

(*) Chernoff Bounds used here for reference

- Upper Tail: $P[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2}{2+\delta}\mu} \qquad \forall\delta > 0$
- Lower Tail: $P[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2}{2}\mu} \qquad \forall\delta > 0$

Question 3

First we show that $\gamma_{MAX-CUT}(G) \geq m/2$. Consider the following process: initialize two empty sets S_1, S_2. Iterate over all vertices in v, and let $E_{v,1}$ be the set of edges connecting v to S_1, and $E_{v,2}$ to S_2. If $|E_{v,1}| \geq |E_{v,2}|$, add it to S_2. Otherwise, add it to S_1. In the end, (S_1, S_2) is a cut. Observe that its size is at least $m/2$, since the sets $\{E_{v,1}, E_{v,2}\}_{v\in V}$ form a partition of E, and for each vertex v the bigger of the sets $E_{v,1}, E_{v,2}$ is a subset of $E(S_1, S_2)$.

Recall that in Homework 2 problem 1 we constructed a graph sparsifier G' with the following guarantee by sampling edges with probability $p = \min 1, Cn/(\epsilon^2 m)$:

$$E(S, V\setminus S) - \epsilon m \leq (1/p)E'(S, V\setminus S) \leq E(S, V\setminus S) - \epsilon m$$

for any cut S. Hence for any cut of size at least $m/4$

$$E(S, V \setminus S)(1 - 4\epsilon) \leq (1/p)E'(S, V \setminus S) \leq E(S, V \setminus S)(1 + 4\epsilon)$$

Suppose we set ϵ in this sparsifier to be equal to $\epsilon/4$. Then if we were to find a max cut S' in G', it's value in G would still be at least $m/4$, which would mean that for it the relative guarantee also holds. Therefore, the size of S' is $1 \pm \epsilon$ close to the size of the max cut. Using Chernoff bound, you can show that with high probability there is only $O(n/\epsilon^2)$ edges in G'.

The algorithm goes as follows:

1. Sample each edge with probability p from the stream.
2. Find max cut S in the sampled graph.
3. Return S

Question 4

We consider the offline problem, in which we have an erroneous predictor Previous work: -in the case when the error is bounded Gamlath et al and Ergun et al (Learning-Augmented k-means Clustering) perform outlier detection and show that the centroids of the remaining predicted clustering they compute have good quality -In the case of perfect predictions Gamlath et al (Semi-Supervised Algorithms for Approximately Optimal and Accurate Clustering) begin with an approximate solution and split the area around each centroid into rings. improve the solution by querying sufficiently in each ring

We propose to begin with an approximate solution, then gather information about the quality of the prediction allowing us to either improve the solution or lower bound the error.

Question 5 a

Note that for a fixed choice of S, the $|S_{ALG}|$ is the number of vertices in S, less the vertices in S such that there is an edge in the graph with both endpoints in S. For each $v \in V$ let X_v be the indicator random variable that takes on value 1 if v was included in S (before removal) and 0 otherwise. Then for each edge uv let Z_{uv} be the indicator random variable that takes on value 1 if both u and v were included in S (before removal) and 0 otherwise. Then by the union bound the probability that a single vertex $v \in V$ is included after removal is at least

$$X_v - \sum_{u:uv \in E} Z_{uv}.$$

By design of our algorithm $\mathbb{E}[X_v] = \frac{x_v^*}{\sqrt{m}}$ and $\mathbb{E}[Z_{uv}] = \frac{x_v^* x_u^*}{m}$. Then by linearity of expectation

$$\mathbb{E}[|S_{alg}|] \geq \sum_{v \in V} \frac{x_v^*}{\sqrt{m}} - \sum_{u,v \in V: uv \in E} \frac{x_v^* x_u^*}{m}.$$

Question 5 b

Note that by feasibility of the linear program we have that for each edge $e = uv$ that $x_u^* + x_v^* \leq 1$. As a result $x_u^* \cdot x_v^*$ is maximized when $x_u^* = x_v^* = 1/2$. Hence ,

$$\sum_{u,v \in V: uv \in E} \frac{x_v^* x_u^*}{m} \leq \frac{1}{4m} \leq 1/2.$$

Question 5 c

We apply parts (b) and (c) to obtain that

$$\mathbb{E}[|S_{alg}|] \geq \sum_{v \in V} \frac{x_v^*}{\sqrt{m}} - 1.$$

Since we assumed that $\sum_{v \in V} x_v^* \geq 2\sqrt{m}$ we get that

$$\mathbb{E}[|S_{alg}|] \geq \sum_{v \in V} \frac{x_v^*}{\sqrt{m}} - \frac{2\sqrt{m}}{2\sqrt{m}} \geq \sum_{v \in V} \frac{x_v^*}{2\sqrt{m}}.$$

and hence since the claim holds.