Solution for Exam

1. (a) The algorithm is as follows: first take arbitrary client \( c_1 \in C \), and take \( p_1 \) the supplier closest to \( c_1 \) into \( P \). Now, for \( i = 2, \ldots, k \), take a client maximizing \( d(c_i, P) \), and add to \( P \) the supplier \( p_i \in S \) closest to \( c_i \). (Note that we may have chosen a supplier or a client more than once).

Consider the optimal solution \( o_1, \ldots, o_k \in P \), and the \( k \) clients \( c_1, \ldots, c_k \). Let \( \sigma : [k] \to [k] \) be a map matching the \( i \)-th client to its closest supplier from \( o_1, \ldots, o_k \). If \( \sigma \) is a bijection, meaning every client is serviced by a different supplier in the optimal solution, then we claim that for every \( 1 \leq i \leq k \) there is a supplier from \( P \) in \( B(o_i, 2\text{OPT}) \). This is because the algorithm chose for client \( j \) with \( \sigma(j) = i \) the closest supplier \( p_j \) to be in \( P \), in particular \( d(p_j, o_i) \leq d(p_j, c_j) + d(c_j, o_i) \leq 2\text{OPT} \). This suggests that all clients are within 3OPT from some supplier \( p_j \).

Otherwise, it must be the case that \( \sigma \) is not a bijection, by the pigeonhole principle there is a supplier \( o_i \) servicing two clients \( c_j, c_h \) (that is, \( \sigma(j) = \sigma(h) = i \)). Assume w.l.o.g that \( h < j \), then \( d(p_h, c_j) \leq d(p_h, c_h) + d(c_h, o_i) + d(o_i, c_j) \leq 3\text{OPT} \) (because \( d(p_h, c_h) \leq d(o_i, c_h) \)). So we chose client \( c_j \) even though it has been covered by \( p_h \) with distance of at most 3OPT, as we chose the client furthest away from \( P \), all clients are covered by 3OPT.

(b) Given an input graph \( G = (V, E) \) for the dominating set problem (for convenient assume that \( (x, x) \in E \) for all \( x \in V \)), take two disjoint copies of \( V \) call them \( C \) and \( S \), and define the following metric:

\[
    d(x, y) = \begin{cases} 
    0 & x = y \\
    1 & |S \cap \{x, y\}| = 1 \land (x, y) \in E \\
    3 & |S \cap \{x, y\}| = 1 \land (x, y) \notin E \\
    2 & \text{otherwise}
    \end{cases}
\]

It can be verified that this is indeed a metric, we have to show that if there is a dominating set of size \( \leq k \) then OPT = 1 and if there is no dominating set of size \( \leq k \) then OPT = 3.

For the first case, if \( D \) is a dominating set of size \( k \), then taking the copy of \( D \) in \( S \) is a set of suppliers of distance 1 from all clients. For the second case, every solution of size \( k \) for the suppliers problem cannot be a dominating set, so there must exist a client of distance 3 from the \( k \) suppliers. This suggests that \( 3 - \epsilon \) approximation for any \( \epsilon > 0 \) will give us a polynomial time algorithm that can decide the existence of a dominating set of size \( k \).

2. (a) Fix some element \( a \in E \), and the \( f \) sets \( S_1, \ldots, S_f \) that contain \( a \). For every \( 1 \leq i \leq f \), let \( A_i \) be the event that the algorithm picks \( S_i \) and none of the other sets, note that \( \text{Pr}[A_i] = p \cdot (1 - p)^{f-1} \). Since the events \( A_i \) are disjoint, \( \text{Pr}[\exists i, A_i] = \sum_{i=1}^{f} \text{Pr}[A_i] = fp \cdot (1 - p)^{f-1} \). The expected number of covered elements, by linearity of expectation, is \( nf \cdot p(1 - p)^{f-1} \). Taking derivative with respect to \( p \) suggests that this is maximized for \( p = 1/f \), which yields expectation of \( n(1 - 1/f)^{f-1} \approx n/e \).

(b) If \( f = 1 \) the problem is trivial (simply choose all sets), so assume \( f > 1 \). Sampling every set with probability \( 1/f \), if an element is covered by \( k \) sets, then the probability that it
is uniquely covered is
\[ k \cdot \frac{1}{f} \left( 1 - \frac{1}{f} \right)^{k-1} \geq \frac{1}{2} \cdot \left( 1 - \frac{1}{f} \right)^f \geq \frac{1}{8}, \]
where we used that \( k \geq f/2 \) in the first inequality and that \( k \leq f \) and \( (1 - 1/f)^f \geq 1/4 \) for all \( f \geq 2 \). Thus the expected number of uniquely covered elements is at least \( n/8 \). Since \( \text{OPT} \leq n \), our algorithm is \( 8 \) approximation (in expectation).

(c) For an element \( a \in E \) let \( f(a) \) be the number of sets that contain \( a \). For all integers \( 1 \leq i \leq \log m + 1 \) let \( F_i = \{ a \in E : 2^{i-1} \leq f(a) < 2^i \} \), these are the elements who are covered by at most \( 2^i \) and at least \( 2^{i-1} \) sets. Since every element is included in exactly one \( F_i \), and there are \( \log m + 1 \) values of \( i \), it follows that there exists \( i \) such that \( |F_i| \geq \frac{n}{\log m + 1} \) (clearly this \( F_i \) can be found efficiently). Apply the algorithm of (b) on the elements \( F_i \) and the given set system induced on \( F_i \) with parameter \( f = 2^i \). This gives in expectation at least \( |F_i|/8 \geq \frac{n}{16 \log m} \) uniquely covered elements. Again \( \text{OPT} \leq n \), so this is \( \Omega(1/\log m) \) approximation.

(d) To obtain a factor of \( \Omega(1/\log n) \), every element \( e_i \) chooses some arbitrary set that contains it \( S_j \), and we only consider these \( n \) sets when running the algorithm from (c). Since every element can still be covered, and we get that there are at least \( \frac{n}{\log n + 1} \) elements with similar frequencies, so this is \( \Omega(1/\log n) \) approximation.

3. (a) The algorithm is to put a vertex \( i \) with probability \( 1/2 \) in \( L \), independently of all other choices. If \( X_{ij} \) is an indicator random variable for the event that the edge \( (i,j) \in A \) has \( i \in L \) and \( j \in R \), then clearly \( \text{Pr}[X_{ij}] = 1/4 \). By linearity of expectation, the expected size of the cut is
\[
\mathbb{E} \left[ \sum_{(i,j) \in A} w_{ij} X_{ij} \right] = \sum_{(i,j) \in A} w_{ij} \mathbb{E}[X_{ij}] = W/4,
\]
where \( W \) is the total weight of all edges. Since \( \text{OPT} \leq W \), we have a 1/4-approximation.

(b) Consider the complete directed graph with all weights equal 1, then any cut with \( |L| = \ell \) will have \( \ell(n - \ell) \) edges crossing from \( L \) to \( R \). This expression is maximized for \( \ell = n/2 \) and gives \( n^2/4 \) crossing edges. However, the total number of edges is \( m = n(n-1) \), so the cut size is \( m/4 + O(n) = m/4 + o(m) \).

(c) The variable \( x_i \) indicate whether \( i \in L \), and the variable \( z_{ij} \) indicates whether \( i \in L \) and also \( j \in R \) (so that this edge contributes to the cut). A feasible solution to the integer program can be transformed to a solution of Directed-Max-Cut by putting vertex \( i \) in \( L \) iff \( x_i = 1 \). The value of this solution is \( \sum_{(i,j) \in A} w_{ij} x_i(1 - x_j) = \sum_{(i,j) \in A} w_{ij} z_{ij} \), because if \( x_i = 1 \) and \( x_j = 0 \) then the integer program will put \( z_{ij} = 1 \), otherwise it is constrained to have \( z_{ij} = 0 \). Equivalently, any solution to Directed-Max-Cut can be transformed to a feasible integer program solution by letting \( x_i = 1 \) iff \( i \in L \), and \( z_{ij} = 1 \) iff \( x_i = 1 \) and \( x_j = 0 \).

(d) Let \( x^*, z^* \) be an optimal solution to the linear program, in which we relax the integrality constraint on \( x_i \). Consider some edge \( (i,j) \in A \), the probability that \( i \in L \) and \( j \in R \), that is, the probability that it contributes to the directed max cut, is
\[
\left( \frac{1}{4} + \frac{x_i^*}{2} \right) \cdot \left( \frac{3}{4} - \frac{x_j^*}{2} \right) = \left( \frac{1}{4} + \frac{x_i^*}{2} \right) \cdot \left( \frac{1}{4} + \frac{1 - x_j^*}{2} \right) \geq \left( \frac{1}{4} + \frac{z_{ij}^*}{2} \right)^2,
\]
where we used the LP constraints: that \( x^*_i \geq z^*_{ij} \) and \( (1 - x^*_j) \geq z^*_{ij} \). We can bound this as follows

\[
\left( \frac{1}{4} + \frac{z^*_{ij}}{2} \right)^2 = \left( \frac{1}{4} - \frac{z^*_{ij}}{2} \right)^2 + \frac{z^*_{ij}}{2} \geq \frac{z^*_{ij}}{2}.
\]

We conclude that the probability that the edge \((i, j)\) contributes \( w_{ij} \) is at least \( z^*_{ij} / 2 \), so the expected contribution is

\[
\sum_{(i,j) \in A} w_{ij} \Pr[i \in L \land j \in R] \geq \sum_{(i,j) \in A} w_{ij} z^*_{ij} / 2 \geq \text{OPT} / 2.
\]

**Bonus:** For every \( i \in V \) let \( y_i \) be a \( \{-1, 1\} \) variable, and let \( y_0 \in \{-1, 1\} \) be the indicator for which value, \(-1\) or \(1\), means being in \(L\). Write the following quadratic program

\[
\begin{align*}
\text{max} & \quad \sum_{(i,j) \in A} w_{ij} \left( 1 + y_0 y_i - y_0 y_j - y_i y_j \right) / 4 \\
\text{subject to} & \quad y_i y_i = 1 \quad \forall \ i \in \{0, 1, \ldots, n\} \\
& \quad y_i \in \mathbb{R} \quad \forall \ i \in \{0, 1, \ldots, n\}.
\end{align*}
\]

It can be checked that the value of this program is equal to the value of the maximum directed cut, because every term in the summation is equal to \( w_{ij} \) if \( y_i = y_0 \) and \( y_j = -y_0 \), and is equal to 0 in the other three cases.

Relax the program to a SDP by replacing the last constraint with \( y_i \in \mathbb{R}^n \). Solve the SDP and obtain an optimal vector solution \( v_0, v_1, \ldots, v_n \). For the rounding, select uniformly at random a hyperplane \( r \in \mathbb{R}^n \), and set \( i \in L \) iff \( \text{sign}(v_0 r) = \text{sign}(v_i r) \). Let \( y_i = 1 \) if indeed \( \text{sign}(v_0 r) = \text{sign}(v_i r) \) and set \( y_i = -1 \) otherwise. Let \( X_{ij} \) be an indicator random variable for the event that \( i \in L \) and \( j \in R \), as discussed above, we have that \( X_{ij} = \frac{1 + y_0 y_i - y_0 y_j - y_i y_j}{4} \).

The expected weight of the cut returned by the algorithm is \( \mathbb{E} \left[ \sum_{(i,j) \in A} w_{ij} X_{ij} \right] \). Consider the following events:

\[
\begin{align*}
A &= \{ y_0 = y_i \neq y_j \} \\
B &= \{ y_0 = y_i = y_j \} \\
C &= \{ y_0 \neq y_i = y_j \} \\
D &= \{ y_i \neq y_0 = y_j \} \\
E_{i0} &= \{ y_0 = y_i \} \\
\bar{E}_{ij} &= \{ y_i \neq y_j \} \\
\bar{E}_{j0} &= \{ y_0 \neq y_j \}
\end{align*}
\]

It can be checked that \( A, B, C, D \) are disjoint and their probabilities sum to 1. Furthermore, \( B = E_{i0} \setminus A, C = \bar{E}_{j0} \setminus A \) and \( D = \bar{E}_{ij} \setminus A \). Thus we can write

\[
1 = \Pr[A] + \Pr[B] + \Pr[C] + \Pr[D] = \Pr[E_{i0}] + \Pr[\bar{E}_{j0}] + \Pr[\bar{E}_{ij}] - 2 \Pr[A],
\]
in other words,

\[ \Pr[A] = 1 - \frac{1}{2} \left( \Pr[E_{i0}] + \Pr[E_{j0}] + \Pr[E_{ij}] \right) \]

\[ = 1 - \frac{1}{2} (\pi - \theta_{i0} + \theta_{0j} + \theta_{ij}) \]

\[ = \frac{1 + \theta_{i0} - \theta_{j0} - \theta_{ij}}{2} \]

\[ = \frac{1 + \arccos(v_{i0}v_{0}) - \arccos(v_{0j}v_{0}) - \arccos(v_{ij})}{2}, \]

where \( \theta_{ij} \) is the angle between \( v_{i} \) and \( v_{j} \). It remains to show that this quantity is larger than \( \alpha \cdot \frac{1}{4} (1 + v_{i0}v_{0} - v_{0j}v_{0} - v_{ij}) \) for some \( \alpha > \frac{1}{2} \). This can be verified for \( \alpha = \frac{3}{4} \), say, by case analysis.

4. (a) Let \( B_{x} = B^{\circ}(x, 1/2) = \{ y : \|x - y\|_2 < 1/2 \} \). Observe that if \( x, y \in V \) are included in some independent set, then their distance is at least 1, thus \( B_{x} \cap B_{y} = \emptyset \). As \( B_{x} \) contains a square of side length \( 1/\sqrt{2} \), we have that \( 2k^{2} \) small squares cover the entire big square, so any solution can contain at most \( 2k^{2} \) points. The algorithm now is simple: iterate over every subset of size at most \( 2k^{2} \), and return the one with maximal size which is an independent set. The number of such sets is at most \( n^{2k^{2}} \), and testing independence is done in \( k^{2} \) time.

(b) If the optimal solution only uses disks that do not intersect the grid lines, then we can find it efficiently. For each non-empty square: solve optimally for the disks fully contained in that square using (a), and combine the results. This is feasible solution because no two disks from different squares intersect. To see optimality, note that our solution is at least as large as the global optimum restricted to a square, so the sum over all squares is at least as large as the optimum.

(c) The randomized algorithm is to choose \( k = 2/\epsilon \), and place a randomly shifted grid of side length \( k \). The expected number of disks from the optimal solution that are cut by a grid line is at most \( 2/k \cdot \text{OPT} \). Now use the algorithm of (b) only for disks that are fully contained in a square. The value of the optimal solution for the fully contained disks is expected to be at least \( (1 - \epsilon)\text{OPT} \), and our algorithm will find a solution at least that good.

5. (a) Consider the following probabilistic construction, we start with \( V' = V \).

- For \( i = 1, \ldots, r - 1 \): choose a vertex \( u_{i} \) uniformly at random from \( V' \), and set \( V' = V' \setminus \{u_{i}\} \);

Let \( A_{u} \) be an indicator random variable for the event that vertex \( u \) was not chosen as some \( u_{i} \), then \( \Pr[A_{u}] = 1 - (r - 1)/n \). Conditioning on some other vertex not chosen, \( \Pr[A_{u} | A_{v}] = 1 - (r - 1)/(n - 1) \), because there are \( n - 1 \) options to choose \( r - 1 \) vertices. Consider the partition of \( V \) consisting of \( r - 1 \) singleton clusters, where the \( i \)-th cluster is \( S_{i} = \{u_{i}\} \), and the remaining cluster \( S_{r} \) contains the rest of the vertices. The probability that an edge is in this cut is the probability that at least one of its end points was chosen, that is, if \( e = (u, v) \),

\[ \Pr[\neg A_{u} \lor \neg A_{v}] = 1 - \Pr[A_{u} \land A_{v}] = 1 - \left( 1 - \frac{r - 1}{n} \right) \left( 1 - \frac{r - 1}{n - 1} \right). \]
By linearity of expectation, the expected size of an $r$-way cut created by this random construction is $\left[ 1 - \left(1 - \frac{r-1}{n}\right) \left(1 - \frac{r-1}{n-1}\right) \right] m$. We conclude that there must be some cut achieving this value, and the minimum $r$-way cut can only be smaller.

(b) Fix some minimum $r$-way cut $S_1, \ldots, S_r$ of the graph of size $k$ (that is, there are $k$ edges that are not contained in a single $S_i$), we will show that it survives the contraction algorithm with probability at least $n^{-2(r-1)}$. Let $\mathcal{E}_i$ be the event that no edge of the cut is chosen in the $i$-th step, we saw in class that this implies no edge of the cut is discarded in step $i$. At the beginning of the $i$-th step we have $n - i + 1$ blobs and $m_i$ edges. Even conditioned on $\mathcal{E}_j$ for all $j < i$, we have that the minimum $r$-way cut size is bounded as in (a), the argument there considered every edge separately, so it will hold for graphs over blobs as well (which may have parallel edges). We get that

$$\Pr\left[ \bigwedge_{j=1}^{i-1} \mathcal{E}_j \right] = 1 - \frac{k}{m_i} \geq \left( 1 - \frac{r-1}{n-i+1} \right) \left( 1 - \frac{r-1}{n-i} \right) = \frac{n-i-r+2}{n-i+1} \cdot \frac{n-i-r+1}{n-i}.$$ 

Now,

$$\Pr\left[ \bigwedge_{i=1}^{n-r} \mathcal{E}_i \right] = \prod_{i=1}^{n-r} \Pr\left[ \mathcal{E}_i \mid \bigwedge_{j=1}^{i-1} \mathcal{E}_j \right] \geq \prod_{i=1}^{n-r} \frac{n-i-r+2}{n-i+1} \cdot \frac{n-i-r+1}{n-i} = \frac{r(r-1)\ldots 2}{n(n-1)\ldots(n-r+2)} \cdot \frac{(r-1)(r-2)\ldots 1}{(n-1)(n-2)\ldots(n-r+1)} \geq n^{-2(r-1)}.$$

6. Let $B$ be a set of $q = \lceil 2 \ln(2/(\epsilon \cdot \delta))/\epsilon \rceil$ points chosen uniformly at random from $\binom{X}{q}$. Consider a point $x \in X$, and let $r_x > 0$ be the minimal such that $|B(x,r_x)| \geq \epsilon n/2$. Let $\mathcal{E}_x$ be an indicator random variable for the event $\{B \cap B(x,r_x) = \emptyset\}$. Say $u_1, \ldots, u_\ell$ are the points in $B(x,r_x)$ (where $\ell \geq \epsilon n/2$), and let $A_i$ be the event that $u_i \notin B$. We have that $\Pr[A_i] = 1 - q/n$, and

$$\Pr\left[ \bigwedge_{i=1}^{\ell} A_i \right] = \prod_{i=1}^{\ell} \Pr\left[ A_i \mid \bigwedge_{j<i} A_j \right].$$

Since conditioning on some $A_j$ for $j < i$ only increases the chances that $B$ contains $u_i$, we have that

$$\Pr[\mathcal{E}_x] \leq \prod_{i=1}^{\ell} \Pr[A_i] \leq \left(1 - \frac{q}{n}\right)^{\epsilon n/2} \leq e^{-\ln(2/(\epsilon \cdot \delta))} = \epsilon \cdot \delta / 2.$$
Let $A = \sum_{x \in X} E_x$, then $E[A] \leq \epsilon \cdot \delta n/2$, and by Markov inequality
\[
\Pr[A > E[A]/\delta] \leq \delta.
\]
So with probability at least $1 - \delta$ we have that at most $\epsilon n/2$ points in $X$ have no "representative" from $B$ in $B(x, r_x)$. Let $u \in X$ be such that $E_u$ does not hold, then for any $v \notin B(u, r_u)$ we claim that $\hat{d}(u, v) \leq 3d(u, v)$. To see this, consider the point $b \in B$ such that $b \in B(u, r_u)$, then
\[
\hat{d}(u, v) = \min_{b' \in B} \{d(u, b') + d(v, b')\} \leq d(u, b) + d(v, b) \leq d(u, v) + d(u, b) \leq 3d(u, v),
\]
where the third inequality is the triangle inequality, and the fourth is because $d(u, b) \leq d(u, v)$.

To see how many pairs are well preserved: each point $u$ for which $E_u$ does not hold have at least $(1 - \epsilon/2)n$ other points $v$ for which $\hat{d}(u, v)$ is a good estimate. We saw that with probability at least $1 - \delta$ there are at least $(1 - \epsilon/2)n$ such good points, we may be counting every pair twice, so the total number of good pairs is at least
\[
((1 - \epsilon/2)n)^2/2 \geq (1 - \epsilon) \left(\frac{n}{2}\right).
\]

7. Let $e_i$ be the length of the $i$-th edge of the cycle, and consider the continuous cycle $C$ of length $L = \sum_{i=1}^{n} e_i$. Place the points of the metric $X$, $a_1, \ldots, a_n$, accordingly on the cycle. For any $x \in C$ let $\bar{x}$ be its antipodal point on $C$, and define a cut $S_x = \{a \in X : a$ is on the left of the line connecting $x, \bar{x}\}$. Observe that we have an infinite set of cuts, but there are at most $n^2$ different ones, because any cut $S_x$ consists of the points $a_i, a_{i+1}, \ldots, a_j$ for some $i, j \in [n]$ (where again, indices are cyclic). Let $C(S) = \{x \in C : S_x = S\}$ be the set of points on the cycle whose induced cut is $S$. Define the weight of the cut $S$ to be $\lambda_S = \int_{C(S)} dx$, where the total length is normalized to be $\int_C dx = 2L$ (the 2 factor since we are counting every cut twice).

This gives an isometric embedding to $\ell_1$, because if $a_i, a_j$ are two points on the cycle of distance $d(a_i, a_j)$, then every point $x \in C$ which lies on the shortest path connecting $a_i$ to $a_j$ on the cycle, and its antipodal, both will contributes to a cut separating $a_i$ from $a_j$. To see this observe that the antipodal of $x$ cannot lie inside this shortest path, because otherwise this path would be of length $> L/2$, which is a contradiction. Every other cut will not separate $a_i, a_j$, so the distance is exactly preserved.