Technique 5: Randomized Rounding

One round of randomized rounding:
Pick set S_j uniformly at random with probability $1 - x_j$ (for all j).

Version A: Repeat rounds until you have a cover.

Version B: Repeat for s rounds. If you have a cover STOP.
Otherwise, repeat the whole algorithm.

Probability that $u \in U$ is not covered (in one round):

$$\begin{aligned}\Pr[u \text{ not covered in one round}] \\ &= \prod_{j:u\in S_j}(1-x_j) \le \prod_{j:u\in S_j} e^{-x_j} \\ &= e^{-\sum_{j:u\in S_j} x_j} \le e^{-1} \ .\end{aligned}$$

Probability that $u \in U$ is not covered (after ℓ rounds):

$$\Pr[u \text{ not covered after } \ell \text{ round}] \le \frac{1}{e^{\ell}} \ .$$

$$
\begin{aligned}
&\Pr[\exists u \in U \text{ not covered after } \ell \text{ round}] \\
&\quad = \Pr[u_1 \text{ not covered} \vee u_2 \text{ not covered} \vee \ldots \vee u_n \text{ not covered}] \\
&\quad \le \sum_i \Pr[u_i \text{ not covered after } \ell \text{ rounds}] \le ne^{-\ell} \enspace .
\end{aligned}
$$

Lemma 5

With high probability $\mathcal{O}(\log n)$ *rounds suffice.*

With high probability:

For any constant α the number of rounds is at most $\mathcal{O}(\log n)$ with probability at least $1 - n^{-\alpha}$.

Proof: We have

$$\Pr[\#\text{rounds} \geq (\alpha + 1)\ln n] \leq ne^{-(\alpha+1)\ln n} = n^{-\alpha} \ .$$

Expected Cost

- Version A.
  Repeat for $s = (\alpha + 1)\ln n$ rounds. If you don't have a cover simply take for each element u the cheapest set that contains u.

  $$E[\text{cost}] \le (\alpha+1)\ln n \cdot \text{cost}(LP) + (n \cdot \text{OPT})n^{-\alpha} = \mathcal{O}(\ln n) \cdot \text{OPT}$$

Expected Cost

- Version B.
  Repeat for $s = (\alpha + 1)\ln n$ rounds. If you don't have a cover simply repeat the whole process.

$$E[\text{cost}] = \Pr[\text{success}] \cdot E[\text{cost} \mid \text{success}] + \Pr[\text{no success}] \cdot E[\text{cost} \mid \text{no success}]$$

This means

$$\begin{aligned} &E[\text{cost} \mid \text{success}] \\ &= \frac{1}{\Pr[\text{succ.}]}\Big(E[\text{cost}] - \Pr[\text{no success}] \cdot E[\text{cost} \mid \text{no success}]\Big) \\ &\le \frac{1}{\Pr[\text{succ.}]} E[\text{cost}] \le \frac{1}{1 - n^{-\alpha}} (\alpha + 1)\ln n \cdot \text{cost}(\text{LP}) \\ &\le 2(\alpha + 1)\ln n \cdot \text{OPT} \end{aligned}$$

for $n \ge 2$ and $\alpha \ge 1$.

Randomized rounding gives an $\mathcal{O}(\log n)$ approximation. The running time is polynomial with high probability.

Theorem 6 (without proof)

There is no approximation algorithm for set cover with approximation guarantee better than $\frac{1}{2}\log n$ unless NP has quasi-polynomial time algorithms (algorithms with running time $2^{\mathrm{poly}(\log n)}$).

Integrality Gap

The integrality gap of the SetCover LP is $\Omega(\log n)$.

- $n = 2^k - 1$
- Elements are all vectors $\boldsymbol{i}$ over $GF[2]$ of length k (excluding zero vector).
- Every vector $\boldsymbol{j}$ defines a set as follows

$$S_{\boldsymbol{j}} := \{\boldsymbol{i} \mid \boldsymbol{i} \cdot \boldsymbol{j} = 1\}$$

- each set contains 2^{k-1} vectors; each vector is contained in 2^{k-1} sets
- $x_i = \frac{1}{2^{k-1}} = \frac{2}{n+1}$ is fractional solution.

Integrality Gap

Every collection of $p < k$ sets does not cover all elements.

Hence, we get a gap of $\Omega(\log n)$.

Techniques:

- Deterministic Rounding
- Rounding of the Dual
- Primal Dual
- Greedy
- Randomized Rounding
- Local Search
- Rounding Data + Dynamic Programming