Set Balancing

Input: Sets $S_1, S_2, \ldots, S_m \subseteq [n]$  
Output: $\sigma: [n] \rightarrow \{-1, 1\}$ / \{red, blue\}

Objective: Find a signing which is as balanced as possible

Formally, 
$$\text{minimize } \max_{i=1}^{m} \left| \sum_{j \in S_i} \sigma(j) \right|$$

Quantity is called the Discrepancy of the Set system.

This lecture we will work with the case when any element is in at most $d$ sets

- Casting as an Integer Program:

$$\text{Min } T$$

$$\sum_{j \in S_i} \sigma_j \leq T, \quad \forall i = 1, \ldots, m$$

$$\sum_{j \in S_i} \sigma_j \geq -T, \quad \forall i = 1, \ldots, m$$

$$\sigma_j \in \{-1, 1\}$$

$$-1 \leq \sigma_j \leq 1$$
What is the value of $T$? $T = 0$ when $\sigma = 0$

**Algorithm**

- Maintain a coll of "safe" sets, $A \equiv \emptyset$
- Maintain a coll of "set" vars, $I \equiv \emptyset$
  
  // Ultimately $I = \mathbb{C}_n$

- Define a "surplus" vector $s \in \mathbb{R}^m$, a var for every set, $s \equiv_{init} all$-zeros.

- Consider $LP(A, I)$:

  $\sum x \in \prod_{c=1,i \in I}^{c \in I} x_i$

  $\sum x^i_j = s^i_i \quad \forall i \in A$

  $\sum_{j \in S_i \setminus I} x^i_j = $ un-safe sets.

- $x$ be a bfs of $LP(A, I)$

  - $\text{If: } x^i_j = -1 \text{ or } +1$, $I = I \cup j$ and $\sigma_j = x^i_j$
    
    - $\forall i \in A$, $\sum x^i_j + \sum_{j \in S_i \setminus I} x^i_j = 0$
    

  - Else If: For some set $S_i$, $|S_i \setminus I| \leq d$

    - \text{Case A: } then $A = A \cup i$
      
      We proclaim this set is safe. Why?

      \text{Note!}:
      \[ \sum x_i + \sum_{j \in S_i \setminus I} x^i_j = 0 \]
\[ \sum_{j \in S_i \cap I} + \sum_{j \in S_i \setminus I} = 0 \]

\[ \Rightarrow \left| \sum_{j \in S_i \cap I} \sigma_j \right| \leq \left| \sum_{j \in S_i \setminus I} x_j \right| < d \]

and since LHS is integer, \[ \leq d - 1 \]

Ultimately, \( j \in S_i \setminus I \) will get some \( \sigma_j \)'s,
but \( \left| \sum_{j \in S_i \setminus I} \sigma_j \right| \leq d \) \( \therefore |S_i \setminus I| \leq d \)

\[ \therefore \left| \sum_{j \in S_i} \sigma_j \right| \leq \left| \sum_{j \in S_i \cap I} \sigma_j \right| + \left| \sum_{j \in S_i \setminus I} \sigma_j \right| \]

\[ \leq 2d - 1 \]

Claim: The algorithm terminates with \( \sigma_j \in \{-1, 1\} \) for all elements.

Pf: It suffices to show one of case A or case B always occurs.

Suppose case A doesn't occur.

\[ \Rightarrow \left| \{j \mid 0 < x_j < 1/2\} \right| = \text{# of cols} \]

\[ \Rightarrow \text{rank ("const. matrix") } \wedge \text{ # of rows} \]

But each col. has \( \leq d \) ones (Problem.)
But each col. has \( \leq d \) ones \( \text{(Problem assumption)} \).

Since \( \# \text{ rows} \geq \# \text{ cols} \),

\[
\text{some row } i \text{ has } \leq d \text{ ones}
\]

i.e. \( \left| S_i \setminus I \right| \leq d \implies \text{Case 6 occurs} \)

Claim: At the end we have a \((2d-1)\)-balanced coloring/signing.

\[\text{Proof:} \quad \text{The blue stuff above proves this for safe sets.} \]

- For unsafe sets, in fact, we have perfect balance

Note: This is not a standard Approx Alg.

We didn’t compare ourselves with the “best-possible” balance. Rather, we proved that any set-system with each element in at most \( d \)-sets has a \((2d-1)\)-balanced coloring.

It turns out that the \( q \)-th of whether there is a coloring with “perfect balance”, i.e., whether \( \exists \sigma : [n] \rightarrow \{1, 2, 3\} \) s.t.

\[\forall i, \sigma(S_i) = 0 \text{ is in } P \text{ that is, it can be solved in poly-time.} \]

\[\therefore \text{If } \text{OPT} = 0, \text{ there can be an } ALG = 0 \]
opt \geq 1$, and our thm. gives a $(2d-1)$-factor alg.