Project Proposal: Hereditary and Vector Discrepancy of Set Systems

Alantha Newman
CNRS, G-SCOP, Grenoble, France.
Email: alantha.newman@grenoble-inp.fr.

Given a set system $\mathcal{S} = \{S_1, S_2, \dots S_m\}$ on n elements, the *discrepancy* of the set system $\mathcal{S}$ is the following quantity:

$$disc(\mathcal{S}) \quad := \quad \min_{\chi \in \{\pm 1\}} \; \max_{S_j \in \mathcal{S}} \; |\sum_{i \in S_j} \chi(i)|. \tag{1}$$

In other words, suppose we wish to color each element in the set system red or blue. Our goal is to find a coloring so that the maximum difference in the number of red and blue elements over all the sets is minimized. Recently, there have been breakthrough results due to Bansal (2010) and Lovett and Meka (2012), who gave efficient algorithms for computing approximate bounds on this quantity.

Other notions that are less well-understood are *hereditary discrepancy*, *vector discrepancy* and *hereditary vector discrepancy*. The hereditary discrepancy is the maximum value obtained for the discrepancy of the set system when the set system is restricted to some subset of elements.

$$herdisc(\mathcal{S}) \quad := \quad \max_{W \subseteq [n]} \; \min_{\chi \in \{\pm 1\}} \; \max_{S_j \in \mathcal{S}} \; |\sum_{i \in S_j \cap W} \chi(i)|. \tag{2}$$

This quantity is considered to be more robust and can be approximated to within a better factor than the discrepancy. However, the current algorithms are rely on more sophisticated geometric arguments.

The goal of this project is obtain a good understanding (and, hopefully, simplifications) of the current algorithms for efficiently computing bounds on the discrepancy and the hereditary discrepancy of a set system, and to explore the relationships and bounds between the various notions of discrepancy. Besides this general goal, there are a number of specific open problems related to computing the discrepancy of certain set systems that can be addressed. The project can also include implementation of algorithms if the student desires.