Mathematical foundations of sampling connected balanced graph partitions

organized by
Sarah Cannon and Daryl DeFord

Workshop Summary

This workshop brought together researchers from a variety of institutions and research areas. The main focus was on the problem of sampling partitions of graph into connected pieces of equal sizes. While motivation for this problem comes from the computational study of political districting, the workshop focused on the underlying mathematical questions. This involves the intersection of graph theory, probability, combinatorics, computational geometry, and more. The workshop resulted in both new theoretical results about existing sampling methods (such as proving bounds on the mixing time of the Recombination Markov chain on the complete graph K_n) as well as developing new sampling methods (such as a new Markov chain we later named the Balanced Up-Down (BUD) walk). Overall, the organizers are extremely pleased with the outcomes of this workshop and feel it accomplished it goals.

Each morning featured two talks, by an individual or a small group, on a variety of topics. After an open problem session on Monday, afternoons were spent divided into subgroups working on specific questions. The open problem session was particularly productive, generating approximately thirty questions for future research. This was an excellent outcome for the workshop, since as a relatively new research field there have not been many opportunities for experts in the field to come together to identify promising research directions. Several subgroups have continued working on these problems throughout Summer 2025 over Zoom, after forming initial connections at AIM, and we anticipate multiple research papers to arise from these efforts. The work done by the various subgroups is summarized below.

Irreducibility. This group considered several Markov chains over the space of balanced graph partitions, and asked whether they were irreducible, that is, whether the moves of the Markov chain suffice to move between every pair of partitions. Along the way, the group came up with a new Markov chain that walks over the set of "balanced" spanning trees of a graph — trees which can be split into k equal-sized parts. We called this Markov chain the Balanced Up-Down (BUD) walk. This chain has the potential to be more powerful than those from the ReCom family (the current standard approach), since it can change more than two districts in a single step. While we determined that it is not as powerful as ReCom in general, the group is currently exploring whether there are particular settings (i.e., exact district balance, on grid graphs) in which it is irreducible. The group has met several times during Summer 2025, and is in the process of proving a variety of irreducibility results about the BUD walk and implementing this new random walk to see how it performs in practice.

k-splittability/tree metrics. This group focused on graph-theoretic metrics for determining whether a given tree or families of trees can be partitioned into k-equally sized subtrees, based on interpolating stars (not splittable) and paths (splittable). A series of computational experiments found strong correlations between the proportion of splittable trees of a fixed order and selected metrics. This group also studied the corresponding question of splittability of spanning trees of grids, motivated by the other research happening at the workshop, observing the surprising result that even in this setting diameter $n-2$ is not enough to guarantee splittability, through a series of pathological examples. Finally, the group considered relationships between this problem and the up-down walk, using the Markov chain to determine a metric on the space of trees to address the question of how many peturbations are necessary to turn a non-splittable tree into a splittable one.

Mixing of Recombination on K_n. The group worked to prove fast mixing time results for ReCom on the complete graph, as a starting point for proving fast mixing results for ReCom on more general families of graphs. With some effort, we first proved upper and lower bounds on the diameter of the configuration space, which matched up to a constant. Later in the week, a coupling was proposed to show fast mixing for this process, and we showed that it indeed gave a mixing time upper bound of order $k^2 \log n$. We determined that the dependence on n in the mixing time was necessary, and after the workshop, we were able to turn our argument into a lower bound of order $\log n$. These are some of the first results saying anything about the mixing time of Recombination Markov chains, and even though the complete graph is far from the graphs that are most relevant for applications, this is an important first step. Correspondence has continued during Summer 2025, with an upcoming Zoom meeting scheduled to discuss follow-up work.

Generating Dual Graphs The central question considered by this group was, "How can we develop models of graphs arising in real geographic data?" These are the graphs for which we're most interested in sampling connected balanced partitions, and a better understanding of these graphs could provide useful insights that help design and analyze sampling methods. The approach used to better understand these graphs was to attempt to build a random graph model that produces graphs resembling the dual graphs used in practice. We discussed previous work, generated a list of possible alternative methods for generating random dual graphs, and discussed methods for evaluating random models for dual graphs - what criteria make the most sense for evaluating whether a particular random graph model is a good one? This group has held one zoom meeting and has designed experiments to evaluate competing models, including applying deep learning techniques.

MST vs. UST. The following question arose during the Monday workshop problem session: "Is it true that on any graph, a random minimum spanning tree (MST) is weakly less likely to be partitionable into two equal-size pieces than a uniformly random spanning tree (UST)?" On the first day, our group obtained a negative answer by providing a counterexample with 6 vertices and 7 edges. On the second day, we worked towards obtaining a positive answer on grids and complete graphs. For complete graphs, a hopeful tool was the "path rotation" moves from Models of Random Spanning Trees (https://arxiv.org/abs/2407.20226), although we obtained only small partial results. After the second day, the group dispersed and focused their energy on other working groups.