Lewis University
Dr. James Girard Summer Undergraduate Research Program 2022
Faculty Mentor - Project Application
Mathematical Modeling of Self-Assembling DNA

Dr. Amanda Harsy
Department of Engineering, Computing, and Mathematical Sciences
CoAST

January 9, 2024

Mathematical Modeling of Self-Assembling DNA

Abstract

Self-assembly is a term used to describe the process of a collection of components combining to form an organized structure without external direction. The unique properties of double-stranded DNA molecules make DNA a valuable structural material with which to form nanostructures, and the field of DNA nanotechnology is largely based on this premise. By modeling nanostructures with discrete graphs, efficient DNA self-assembly becomes a mathematical puzzle. These nanostructures have wide-ranging applications, such as containers for the transport and release of nano-cargos, templates for the controlled growth of nano-objects, and in drug-delivery methods. This research project centers around the exploration of the graph theoretical and combinatorial properties of DNA self-assembly, as well as development of computational tools to aid in answering fundamental questions that arise.

Introduction and Background

Within biological systems, DNA is used as a means for storing information that guides the processes that build the elements of living systems out of building blocks based on amino acids. The field of DNA nanotechnology uses DNA as building blocks to construct molecular structures. The information from the DNA, embedded in their nucleotide sequences and following basic base-paring rules, can be used to guide their self-assembly into a desired nanostructure. DNA self-assembly, and self-assembly in general, is a rapidly advancing field, with [24, 27] providing good overviews. Synthetic DNA molecules have been designed that self-assemble into given nanostructures, starting with branched DNA molecules [14, 31], nanoscale arrays [32, 33], numerous polyhedra [3, 9, 10, 28, 39], arbitrary graphs [13, 25, 34], a variety of DNA and RNA knots [20, 21, 30], and the first macroscopic self-assembled 3D DNA crystals [40]. This has led to molecular scaffoldings made of DNA which have wide-ranging potential, such as containers for the transport and release of nano-cargos, templates for the controlled growth of nano-objects, biomolecular computing, biosensors and in drug-delivery methods (see [1, 6, 7, 15, 18, 22, 23, 38, 26, 36]).

Furthermore, new experiments demonstrating the feasibility of engineering synthetic DNA nanoscale constructs offer great promise for emergent applications in nanoelectronics, biosensors, biomolecular computing, drug delivery systems, and directed organic synthesis, all of which lead to more effective diagnosis and treatment of illness. [37], [19], [17], [35], and [18] give excellent overviews of self-assembly methods and emerging applications. In particular, given recent advancements in nanotechnology and the discovery of new laboratory techniques using the Watson-Crick complementary properties of DNA strands, graph theory has become useful in the study of self-assembling DNA complexes. DNA origami is also advancing very rapidly toward its potential as a drug delivery mechanism. For example, [35] and [41] discuss how DNA assembly can improve the treatment of cancer, and [16] outlines how self-assembled DNA can be used to help overcome drug resistance in breast cancer cells. Already various labs have produced closed containers self-assembled out of DNA strands which can be made to encapsulate particles, and then their opening deliberately triggered. See [2] and [8] and the HarvarDNAno video [29] which well illustrates both the process and promise of this new technique.

Since modeling this self-assembling process requires designing the component molecular building

blocks, which often are modeled through surface meshes, lattice subsets, and other graph-like structures, construction methods developed with concepts from graph theory have resulted in signi - cantly increased e ciency. For example, one recent focus in DNA nanotechnology is the formation of nanotubes which can be modeled using a lattice graph. The rules governing the structure of these nanostructures are not yet well understood, and this naturally o ers open problems in the realm of applied graph theory.

Methodology

This research focuses on a construction method for self-assembling DNA structures which involves branched junction molecules whose exible *k*-arms are double strands of DNA. The introduction of a graph-theoretical formalization for exploring the combinatorial properties within the self-assembly of branched junction molecules was rst introduced by Jonoska et al. in [12]. The centers of these branched molecules form the vertices of the constructed complex. One arm of the molecule extends longer than the other and its end forms a cohesive end-type which can bond with other arms with the corresponding complementary nucleotide base pairing (called a bond-edge). See Figures 1 and 2.

To represent these *k*-armed molecules and sticky end types, we will use a vertex of degree *k* like the ones in Figure 1. We call these *k*-armed molecules \tiles." We represent the complementary sticky-ends or *bond-edges* of these tiles using letter labels as in Figures 1 and 2. The letters a, b, c, $\hat{a}$, $\hat{b}$, $\hat{c}$, etc. represent the un-adjoined arms sticking o of the *k*-armed molecules. For example, bond end types a and $\hat{a}$ represent complementary sequences of bases as in Figure 2. Thus a and $\hat{a}$ could self-assemble in order to form a bond-edge. Thus we can combinatorially represent a *k*-armed branched-junction molecule with bond-end types $a_1; :::; a_k$ using a tile $T = (a_1; :::; a_k)$.

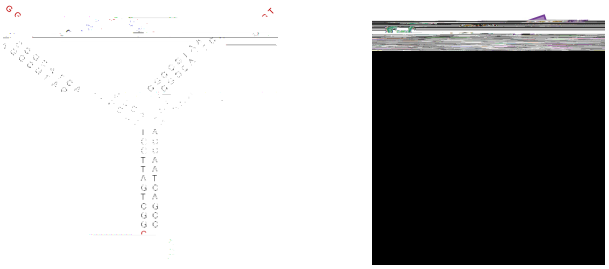

Figure 1: *3*-armed branched junction molecule (left) with example tile representation (right)

Figure 2: Representing the complementary cohesive end types

We assume these tiles havcan3(armsmed)Tl0(the)]eme theseTl0(th)]emek

Figure 3: A pot of 4 tile types.

Theoretically, we assume that once a tile type is made, a pot with that tile type can have \in nitely many" of those tile types. Once we have a pot, those tiles can self-assemble into complexes either complete or incomplete as seen in Figure 4. Labs generally want to create complete complexes.

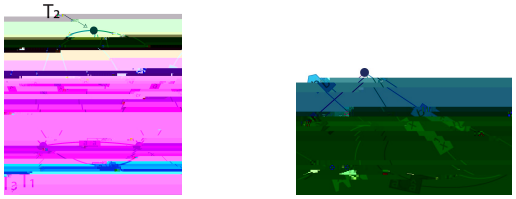

Figure 4: Complete vs Incomplete Complex

Research Aims and Goals

In this research, we explore the underlying *graph-theoretical* structure of nanostructure construction and related design strategy problems for self-assembling DNA structures which involve branched junction molecules whose exible *k*-arms are double strands of DNA. Recall, we call the *k*-armed molecule a \tile" and represent it as a vertex of degree *k* in a graph. We say a collection of tiles (a pot) realizes a graph, *G*, if the collection of tiles constructs the same structure as *G*. Thus, we can investigate design process questions regarding which types of nal structures can be constructed from a given pot of tiles. Inversely, we can also nd a pot of tiles that will realize a given target graph. Our group primarily focuses on the latter problem. Along with graph theory and combinatorics, we also use linear algebra and programming to help answer these questions. The central focus of our research is to nd the minimum number of tile types and bond-edge types necessary to construct a target graph *G* by modular assembly under a variety of laboratory conditions, and furthermore nd explicit pots realizing these minima. This includes creating general design theory and nding accurate bounds for the number of tile types and bond-edge types for a variety of new graphs. Thus, the fundamental questions for thistheer(is)-4701s includ:x

a, target a is the minimum number of tile types od create the

graph? We denote thisth(minim)28(um)-334(n)28(um)28(b)-28(er)-(th(b)28(y)]TJ

on target graph, that as

Scenario 1: A graph with a fewer number of vertices than the target graph may be realized from a pot of tiles.

Scenario 2: A graph on the same number of vertices, but not isomorphic to the target graph may be realized from a given pot of tiles. However, no graph on fewer vertices may be realized.

Scenario 3: No graph on fewer vertices nor non-isomorphic graphs on the same number of vertices as the target graph may be realized from a given pot of tiles.

Mentorship Plan

To help keep the students engaged in the research, the students will meet 2-3 times a week with Dr. Harsy in addition to attending the SURE seminars and the weekly Math Research Seminars. During the Math Research Seminars, students will present their current progress on their research to the other students and faculty working on math research. This allows them to practice presenting and gives them a chance to get feedback from other mathematicians. Most students present bi-weekly. Students will also be expected to keep a research journal and complete a final write-up and research report/paper by the end of the summer. The schedules below provide details and timelines for the students and will help keep them accountable and organized in completing their work. Many of the students who have worked on this research have presented their results at local, regional, or national mathematics conferences and I would expect the students in this project to present at similar conferences over the following year.

Regardless of the project the student chooses, I have a particular way which I prepare students for this research. Below is a detailed plan for the first two weeks (6 meetings) of the research which provides background and allows students to learn the typical techniques, notation, and logic used in this research.

Day 1: Introduction to research topic and notation: Students will read *Using DNA self-assembly design strategies to motivate graph theory concepts* pages 96-101 along with the first three chapters of Alan Tucker's *Applied Combinatorics*.

Day 2: Scenario 1 Techniques: Students will begin reading *Minimal tile and bond-edge types for self-assembling DNA graphs* pages 239-245, and the Scenario 1 section (page 247-249). Create and justify optimal tiling for known graphs like $C_n; K_n; S_n; W_n;$ and P_n in Scenario 1.

Day 3: Scenario 2 Techniques and Introduction to Using the Construction Matrix: Students will continue reading *Minimal tile and bond-edge types for self-assembling DNA graphs*, specially pages 245-247 and the Scenario 2 section on pages 250-253. Students will start creating and justifying optimal tiling for known graphs like $C_n; K_n; S_n; W_n;$ and P_n in Scenario 2. The student will be introduced to the main linear algebra construct used in this research.

Day 4: Scenario 2 Techniques Continued: Students will finish reading *Minimal tile and bond-edge types for self-assembling DNA graphs* pages 254-256, 260-264. Students should have finished creating and justifying optimal tiling for known graphs like $C_n; K_n; S_n; W_n;$ and P_n in Scenario 2.

Day 5: Scenario 3 Techniques: Students should have finished reading *Minimal tile and bond-edge types for self-assembling DNA graphs* and should start creating and justifying optimal tiling for known graphs like $C_n; K_n; S_n; W_n;$ and P_n in Scenario 3.

Day 6: Scenario 3 Techniques Continued: Students should be done creating and justifying optimal tiling for known graphs like $C_n; K_n; S_n; W_n;$ and P_n in all 3 scenarios and will decide on which project/graphs they would like to explore on their own.

Proposed General Timeline with Project Goals

Weeks 1-2: Read the first three chapters of Alan Tucker's *Applied Combinatorics* along with some basic papers including [4] and [5] which introduce the project. Students should work on exercises and confirm

results from these papers.

Weeks 3-7: Develop design strategies to create minimum pots for di erent bipartite, tripartite, or lattice-based graphs in Scenarios 1,2, and 3. Formalize the results by constructing proofs or counter examples.

Weeks 8-9: Finalize and summarize results, attend and present at MAA MathFest.

Week 10: Finish main draft of paper and prepare for SURE presentation

Timeline for Student Write-Up/Paper Task List:

Week 1: Do a literary search for the current research in Modeling DNA Self-Assembly. Read [4] and [5] which introduce the project. Complete a typed 1-3 paragraph synopsis of what might be useful in the articles.

Week 2: Work with Dr. Harsy to decide on the focus of their research and the point of their paper. At

[22] Kyle Lund, Anthony J Manzo, Nadine Dabby, Nicole Michelotti, Alexander Johnson-Buck, Jeanette Nangreave, Steven Taylor, Renjun Pei, Milan N Stojanovic, Nils G Walter, et al. Molecular robots guided by prescriptive landscapes. *Nature*, 465(7295):206, 2010.

[23] Tosan Omabegho, Ruojie Sha, and Nadrian C Seeman. A bipedal dna brownian motor with coordinated legs. *Science*, 324(5923):67{71, 2009.

[24] John A Pelesko. *Self assembly: the science of things that put themselves together*. Chapman and Hall/CRC, 2007.

[25] Phiset Sa-Ardyen, Natasa Jonoska, and Nadrian C Seeman. Self-assembling dna graphs. In *International Workshop on DNA-Based Computers*, pages 1{9. Springer, 2002.

[26] Nadrian C Seeman. An overview of structural dna nanotechnology. *Molecular biotechnology*, 37(3):246, 2007.

[27] Nadrian C Seeman. *Structural DNA nanotechnology*. Cambridge University Press, 2016.

[28] Wei Sun, Etienne Boulais, Yera Hakobyan, Wei Li Wang, Amy Guan, Mark Bathe, and Peng Yin. Casting inorganic structures with dna molds. *Science*, page 1258361, 2014.

[29] swagasaurus. The harvardnanos present: Dna origami containers, Nov 2011.

[30] Hui Wang, Russell J Di Gate, and Nadrian C Seeman. An rna topoisomerase. *Proceedings of the National Academy of Sciences*, 93(18):9477{9482, 1996.

[31] Yinli Wang, John E Mueller, Berries Kemper, and Nadrian C Seeman. Assembly and characterization of five-arm and six-arm dna branched junctions. *Biochemistry*, 30(23):5667{5674, 1991.

[32] Erik Winfree. *Algorithmic self-assembly of DNA*. PhD thesis, California Institute of Technology, 1998.

[33] Erik Winfree, Furong Liu, Lisa A Wenzler, and Nadrian C Seeman. Design and self-assembly of two-

Budget

I am requesting no funds for supplies or software.

Description of any additional Funding you will be using for your proposed research and how they will be used in this project.

I have no additional funding for this research.

Criteria for Student Applicants

This research can be introduced to students with little mathematical background and is open to all majors. The basic graph theory and techniques for this project will be taught to the student as part of the research mentoring program. Depending on the student's background and interests, the research may also involve programming, computer graphics, art, biology, and geometry.

SURE Seminars

I am interested in being a Mock presentation supervisor which I could do on the 5th of August. I could also help prepare and present the following seminars remotely: Presentation Skills, Preparing for Graduate School, Literature Search and Library Resources, Resume Writing and Marketing YOU, and Interview Skills. I would also be willing to do a LaTeX seminar as well.

SURE Agreement

The James Girard Summer Undergraduate Research Program is designed to support the execution of this proposed project by the faculty mentor and a single undergraduate student. After review of faculty proposals, selected projects will be advertised to Lewis University students, and all interested undergraduates will then be required to apply into the program, denoting the project for which they would like to be considered. Student applications will be reviewed for completeness by the program director, and then forwarded to the appropriate faculty mentor for final selection of a candidate. Faculty may submit up to 2 projects for funding through the program. Although faculty mentors may also mentor additional students in the summer not funded through the program, the weekly program events and presentations will be exclusive for students in the program.

By submitting this application, you are agreeing to the following responsibilities of a S.U.R.E. Faculty Mentor:

- Working closely with your student to ensure a worthwhile educational experience. Regular interactions with your student (a minimum of once a week, but more frequently is encouraged) are an expectation. Interaction with other mentors and students is strongly encouraged.
- Participating in the welcome and orientation day.
- Leading at least one of the weekly workshops for the entire group of S.U.R.E. participants.
- Writing at least one blog related to your area of expertise for the program website.
- Participating in the Summer Research Symposium.

This application will be reviewed by a faculty panel for acceptance into the program { determination of selected projects will be communicated after review. Project descriptions will then be made available to Lewis University undergraduate students, who can apply to the program and speci c projects online via our website. Student applicants will be matched with mentors using a selection process where mentors rank interested students based on their applications and students rank projects based on their interests.

Any questions and all completed applications should be sent to Brittany Stephenson (S.U.R.E. Director) at bstephenson@lewisu.edu