Research Statement

Katherine Johnson, University of Nebraska–Lincoln

My main research interests are Extremal Combinatorics, Poset Theory, and Graph Theory. I am also interested in Probabilistic Methods and Entropy, and enjoy a study of algorithms and theoretical computer science. My thesis work has had two foci; most of my research has been computing the weak discrepancy of various posets (see section 2), but I also have results calculating the linear extension diameter of grids (discussed in section 3). I am excited to continue work in these areas, and other similar problems, which I will describe in section 4, Future Work.

1 INTRODUCTION

A **partially ordered set**, or poset, is a collection of elements along with an ordering relation between some of the pairs of elements that is transitive and antisymmetric. Posets are a valuable abstraction of many real-life situations. For example, in a company, there is often a hierarchical system of bosses and workers from various departments, but not all workers are comparable. In an Emergency Room, the patients form a poset where one patient is comparable to another if it is clear that one of their conditions is more severe, e.g. a broken foot is worse than a broken toe but may or may not be worse than a case of the flu. [18]

A **linear extension** L of a poset P is a total ordering of the elements that respects the initial partial order. So $x <_P y$ in the poset implies $x <_L y$ in the linear extension. This is a natural idea to study; for instance, consider the ER example above. We would like to totally order the patients so we know in what order to call them to see the doctor. The ones in the most critical condition should be seen first, and the ones with the minor problems last, but how ought we order the people in the middle?

It is reasonable that patients with incomparable problems shouldn't be terribly far apart in the triage schedule. Therefore, a linear extension might be considered "good" if incomparable elements appear near to one another. We call the maximum distance between two incomparable elements the **discrepancy** of the extension. Then we can measure the **linear discrepancy** of a *poset* by finding the minimum discrepancy over all its linear extensions, which is an intuitive way of measuring just how good the best linear extension of that poset can be. This idea of linear discrepancy was first introduced by Tanenbaum, Trenk, and Fishburn in 2001 in [18], and has become a popular topic to study in the decade since (see also [11, 12, 13, 14, 15]).

In certain situations, such as the first example, it makes sense to weaken the definition of a linear extension by allowing elements of the poset to be given the same label (a total ordering can be thought of as a labeling with natural numbers), while still requiring consistency with the original partial order. This is known as a **weak labeling**. It is more logical to weakly label the employees of a company with pay grades than to linearly order them, because employees can have the same salary. Similar to linear discrepancy, the **weak discrepancy** measures how nicely we can weakly label the elements of the poset. This was actually worked on a bit before linear discrepancy, in a 1998 paper by Gimbel and Trenk [10]. More recently, there was a paper in 2010 by Choi and West [5], and two more papers in 2011 by Howard and Young [14], and Shuchat, Shull, and Trenk [17]. I calculated the weak discrepancy of several posets that were previously unknown, including grids, the permutohedron, and the partition lattice, and am currently working with Young's lattice, all of which I will discuss in more detail in section 2.

I also studied how different two linear extensions could be, i.e. what is the most number of pairs of elements that could be different in two linear extensions? This is known as the **linear extension diameter**, because it is the diameter of the **linear extension graph** of the poset, which represents how similar linear extensions are to each other. In this context, we think of a linear extension as a listing of the elements of P from smallest to largest. Two linear extensions are adjacent in the graph iff they differ in two adjacent elements, e.g. when P is an antichain, any permutation is a linear extension and similarity is measured by an inversion count. This graph was introduced in a 1991 paper by Pruesse and Ruskey [16], with its diameter first studied in 1999 by Felsner and Reuter [9]. The linear extension diameter of the Boolean lattice was not computed until 2011 though, when Felsner and Massow developed a new method of looking at sub-cubes of the lattice in [8]. I was able to generalize their method to extend their results to grids, a more complicated family of posets that contains the Boolean lattice within it.

2 WEAK DISCREPANCY

Recall from the introduction that the **weak discrepancy** of a poset is the minimum over all weak labelings of the maximum distance between two incomparable elements. So if we were to consider pay grades in a company, this would be the largest pay difference between two employees, where neither employee is a supervisor of the other, and the employees have been pay graded as fairly as possible. For a labeling L, we say a and b are a **discrepant pair** of L if a is incomparable to b and their distance is the maximum distance between any two incomparable elements, e.g. the two employees whose salary differences are the greatest.

2.1 Grids

Formally, a **grid** $m_1 \times m_2 \times \cdots \times m_n$ is a poset with elements $x_1 x_2 \ldots x_n$ having $0 \leq x_i \leq m_i - 1$, where $x_1 \ldots x_n \leq y_1 \ldots y_n$ if $x_i \leq y_i$ for all i. Grids are important and interesting to study because they're a fundamental example, simply the product of chains. Chains are already totally ordered, meaning they have a unique linear extension, and hence the grid is basically a study of how linear extensions interact with each other. For example, Figure 1 is the product of a chain of length 2 and one of length 4.

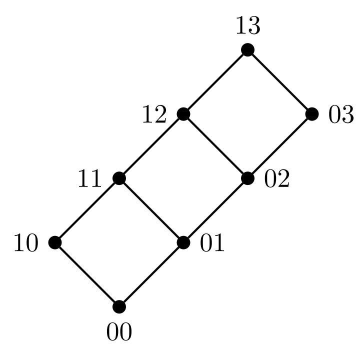

Figure 1: The grid 2×4.

For the grid above, a typical first labeling would be summing the indices at each vertex. This labeling has discrepancy 2 because 10 and 03 have labels 1 and 3, respectively, and these are a discrepant pair. However, we can do better. If we label each vertex by twice the first index, plus the second, then we get a discrepancy of 1, because we've moved these two elements closer together without pushing any others too far apart.

This is rather unexpected, because for the Boolean lattice, the optimal weak labeling is summing the indices, i.e. labeling by level. What is more surprising is that for any grid, the best we can do is to stretch the smallest dimension to the next smallest dimension, so to multiply the smallest dimension index by the ratio of the next smallest to it. In the above example, this would be multiplying the first index by $\lfloor \frac{4}{2} \rfloor = 2$ which is precisely what I showed. Therefore, our freedom for an optimal labeling lies only in the smallest two dimensions. Precisely, my result is that the weak discrepancy of the grid $m_1 \times \cdots \times m_n$ with the m_i nondecreasing is $\sum_{k=2}^{n}(m_k - 1) - \left\lfloor \frac{m_2}{m_1} \right\rfloor$.

2.2 Other Posets

I have also calculated the weak discrepancy of the permutohedron, which is the poset of permutations under the weak Bruhat order, and the partition lattice, which is the poset of partitions where two are comparable if one is the other broken into smaller pieces. For examples of these, see Figures 2 and 3, respectively. The grid $1 \times 2 \times \cdots \times n$ is an extension of the permutohedron. However, this does not directly give either a lower bound or an upper bound, because the permutohedron is not an *induced* subposet. (Induced subposets do have a bounded discrepancy, as shown in [18].) Regardless, I was able to calculate the weak discrepancy of the permutohedron using similar methods that I had created for the proof of the grid discrepancy. The partition lattice is an entirely different poset, unrelated to the grid, but has the very nice structure of long incomparable chains, so I was able to show the best labeling is simply by level.

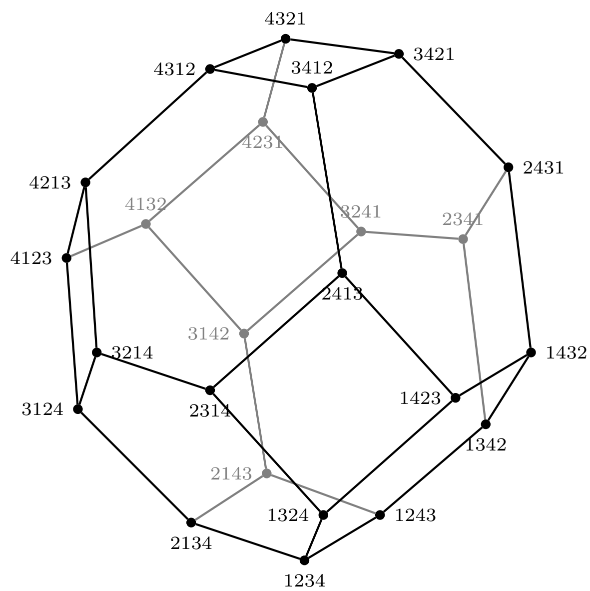

Figure 2: The permutohedron of [4].

Currently, I am working on computing the weak discrepancy of Young's lattice, which is the poset of Ferrer's diagrams ordered by inclusion. See Figure 4 for an example. This is an induced subposet of the grid, so we have immediately an upper bound for the weak discrepancy, but it is not always tight. For this poset, it is usually in our best interest to shift all of the dimensions by some amount, and it's not obvious what that amount is, because the pieces of the grid that are missing vary with each dimension. However, I was still able to compute the weak discrepancy for all 2-dimensional Young's lattices, and it seems likely that my methods will generalize to higher dimensions.

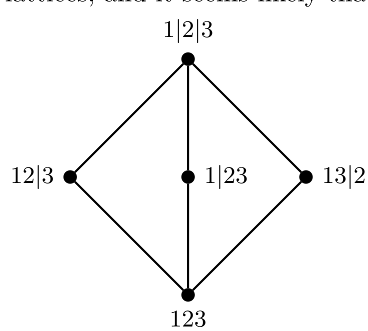

Figure 3: The partition lattice for $n = 3$.

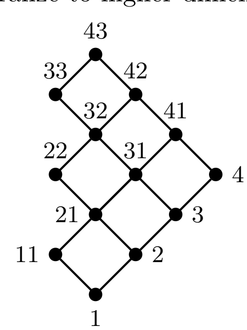

Figure 4: The Young's lattice with $\lambda = 43$.

3 LINEAR EXTENSION DIAMETER

As previously mentioned, the linear extension diameter is a measure of how different two linear extensions can be. It is also an important property of the **linear extension graph** of a poset, which is constructed in the following manner. The vertices are simply all of the linear extensions of the poset P, and two vertices are adjacent if they only differ in a flip, i.e. we have ordered them almost identically except that the order of two adjacent incomparable elements has been changed. This is a natural and interesting graph to study. While this graph contains a Hamiltonian path [19] and its minimum and maximum degrees are known [7] for grids, there are still some mysteries to solve. For example, we do not know the number of vertices; the theory of probabilistic entropy was further developed specifically to solve this problem and provides upper bounds that are tight in some cases, but lower bounds are still unsatisfying [3].

My main result was to compute the **linear extension diameter** of grids, which is the diameter of the linear extension graph, and to characterize the **diametral pairs**, which are the pairs of vertices that are the farthest apart.

In [8], Felsner and Massow compute the linear extension diameter of the Boolean lattice, i.e. 2^n, which had been an open question for 12 years. They consider the binary lexicographic ordering of the elements, with respect to an ordering σ of the indices, and then show that the linear extension L_σ defined by this is a diametral pair with the linear extension created from $\bar{\sigma}$. In order to do this, they consider the sets of ordered pairs of elements of the poset that have a certain difference and intersection, which are the sub-cubes of the poset, and count the number of reversals in these sets.

In order to generalize this to arbitrary grids, I needed to create a similar ordering and discover which sub-posets to consider to count reversals. In this way, I was able to calculate the exact diameter of the linear extension graph of grids, which is $\frac{1}{4}\left(\prod_k m_k\right)\left(\prod_k m_k - \sum_k m_k + n - 1\right)$. However, to show that all diametral pairs look like L_σ and $L_{\bar{\sigma}}$, the generalization was more difficult. The proof required a strong induction showing that the diametral pairs of a grid will produce the diametral pairs of its subgrids when restricted.

The last part of this work was to consider critical pairs and show that grids are diametrally reversing. A **critical pair** of a poset is a pair of elements x and y such that everything that is less than x is also less than y and everything greater than y is also greater than x, but x and y are incomparable. So if we were to decide that $y < x$ in an extension, there could be no elements between them; it feels more natural that we should have $x < y$. In a sense, these are pairs of elements that aren't comparable, but for all practical purposes, should be.

Now, if (x, y) is a critical pair of our poset, then we say that a linear extension reverses (x, y) if $y < x$ in the extension, which is what we don't expect to happen. A poset is **diametrally reversing** if every linear extension contained in a diametral pair reverses a critical pair. Because I had characterized the linear extensions of grids that appeared in diametral pairs, I was able to show that each of these reverses a critical pair, so grids are diametrally reversing.

4 FUTURE WORK

After I compute the weak discrepancy of Young's lattice, I am interested in studying further the relationships between the weak discrepancy of a poset, its fractional weak discrepancy, t-weak discrepancy, total discrepancy, and linear discrepancy, which are all variations of the same idea. In particular, the fractional weak discrepancy, which allows fractional labelings while requiring that comparable elements have labels at least 1 apart, would make an excellent undergraduate research project, because similar methods to my work will provide good upper bounds with which to start, and possibly even the correct labeling and therefore a tight lower bound too. Other interesting posets to study would be the composition lattice [1], the restricted integer partition poset, and products of combinations of these posets.

The t-weak discrepancy allows the multiple usage of a label, like weak discrepancy, but only up to t times. This is an interesting idea because it has practical applications too. For instance, in our original ER example, we may be able to have multiple patients being treated simultaneously, but our limit is that there are only t doctors. The methods I have developed would be an excellent starting place for this problem; however the shifts would need to be altered to match the restriction, so this would be a more complicated research direction.

I also plan to work on computing, or at least providing better bounds for, the linear extension diameter of the other posets that I have studied, improving the results of [9]. I think it would be interesting to consider a weak labeling graph that connects weak labelings that are similar, and to study how different two weak labelings could be (of course, one would need to restrict the labels to come from a fixed set).

My final research idea related to posets is investigating what a random labeling of a fixed poset "looks like." In [7], Cooper investigates the average jump number of a linear extension. I would like to develop a similar theory for a random weak labeling, and see what that produces.

Another combinatorial structure that fascinates me is Universal Cycles. The basic idea is to enumerate all of the objects in a set (e.g. $0-1$ vectors or permutations) by moving from one to the next in such a way that only one aspect of the object has changed. Universal Cycles were introduced in a 1992 paper by Chung, Diaconis, and Graham [6], but I recently became interested in them because of a 2010 paper by Brockman, Kay, and Snively [4] that extended the idea to Universal Cycles of graphs. In this case, the aspect that is changing is a single vertex. I worked with Jamie Radcliffe and Stephen Hartke to correct a mistake in this paper and show that Universal Cycles do exist for labeled trees. I would like to determine whether they exist for unlabeled trees as well, an open question that was included in the Brockman, Kay, and Snively paper.

5 REFERENCES

[1] Anders Björner and Richard Stanley. An analogue of young's lattice for compositions, 2005. `http://arxiv.org/abs/math/0508043`.

[2] Graham Brightwell and Mareike Massow. Diametral pairs of linear extensions, 2008. `http://arxiv.org/abs/math/0809.1828`.

[3] Graham R. Brightwell and Prasad Tetali. The number of linear extensions of the Boolean lattice. *Order*, 20(4):333–345 (2004), 2003.

[4] Greg Brockman, Bill Kay, and Emma E. Snively. On universal cycles of labeled graphs. *Electron. J. Combin.*, 17(1):Research Paper 4, 9, 2010.

[5] Jeong-Ok Choi and Douglas B. West. Forbidden subposets for fractional weak discrepancy at most k. *European J. Combin.*, 31(8):1957–1963, 2010.

[6] Fan Chung, Persi Diaconis, and Ron Graham. Universal cycles for combinatorial structures. *Discrete Math.*, 110(1-3):43–59, 1992.

[7] Joshua Cooper. Random linear extensions of grids, 2006. `http://arxiv.org/abs/math/0602509`.

[8] Stefan Felsner and Mareike Massow. Linear extension diameter of downset lattices of two-dimensional posets. *SIAM J. Discrete Math.*, 25(1):112–129, 2011.

[9] Stefan Felsner and Klaus Reuter. The linear extension diameter of a poset. *SIAM J. Discrete Math.*, 12(3):360–373, 1999.

[10] John G. Gimbel and Ann N. Trenk. On the weakness of an ordered set. *SIAM J. Discrete Math.*, 11(4):655–663, 1998.

[11] Sungpyo Hong, Jong-Yoon Hyun, Hyun Kwang Kim, and Sang-Mok Kim. Linear discrepancy of the product of two chains. *Order*, 22(1):63–72, 2005.

[12] David Howard, Randy Shull, Noah Streib, and Ann N. Trenk. The total linear discrepancy of an ordered set. *Discrete Math.*, 310(5):1022–1025, 2010.

[13] David M. Howard, Mitchel T. Keller, and Stephen J. Young. A characterization of partially ordered sets with linear discrepancy equal to 2. *Order*, 24(3):139–153, 2007.

[14] David M. Howard and Stephen J. Young. When linear and weak discrepancy are equal. *Discrete Math.*, 311(4):252–257, 2011.

[15] Mitchel T. Keller and Stephen J. Young. Degree bounds for linear discrepancy of interval orders and disconnected posets. *Discrete Math.*, 310(15-16):2198–2203, 2010.

[16] Gara Pruesse and Frank Ruskey. Generating the linear extensions of certain posets by transpositions. *SIAM J. Discrete Math.*, 4(3):413–422, 1991.

[17] Alan Shuchat, Randy Shull, and Ann N. Trenk. The total weak discrepancy of a partially ordered set. *Ars Math. Contemp.*, 4(1):95–109, 2011.

[18] Paul J. Tanenbaum, Ann N. Trenk, and Peter C. Fishburn. Linear discrepancy and weak discrepancy of partially ordered sets. *Order*, 18(3):201–225, 2001.

[19] Douglas B. West. Generating linear extensions by adjacent transpositions. *J. Combin. Theory Ser. B*, 58(1):58–64, 1993.