NEGATIVE DEPENDENCE IN DISCRETE PROBABILITY : NOTIONS, MODELS AND CONSEQUENCES

by

Recep Altar Çiçeksiz

B.S., Mathematics, Boğaziçi University, 2018

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Mathematics

Boğaziçi University

2021

ACKNOWLEDGEMENTS

Firstly, I would like to thank Mohan Ravichandran and Ümit Işlak who has been a mentor and a friend over these last few years. For the support they have given me both in terms of mathematics and life vision, I feel indebted to both of them. I would also like to thank Zafeirakis Zafeirakopoulos for being in my jury.

I would like to thank my friends Turan Karakurt, Doğa Ulaş Budak, Ege Baran Yalçın, Kerem Kubilay, Arda Başarıcı, Hacım Şengül, Jülide Miray Özkan, Ali Mehran, Şilan Tank who have supported me throughout my life and especially the process of writing this thesis. Especially, I would like to give my gratitude and appreciation to the one person, Ayşe Başak Savrun, who has supported me no matter what. Without them, I would not have been the person I am now.

Lastly but most importantly, I would like to thank my sister, Nur Şeyma Çiçeksiz; my father, Abdulkadir Çiçeksiz and my mother, Zuhal Çiçeksiz for always being there for me.

ABSTRACT

NEGATIVE DEPENDENCE IN DISCRETE PROBABILITY : NOTIONS, MODELS AND CONSEQUENCES

In this thesis we study the negative dependence properties of random cluster measure. First, we introduce the notions of negative correlation and how they relate to each other. We observe that this important notion is difficult to confirm. The uniform spanning tree measure and the uniform spanning forest measure is the measures that we mainly focus as the crucial examples that shows negative dependence properties. Later, we focus on the properties of the random cluster measure which generalizes the uniform spanning tree and uniform spanning forest measures. We prove negative edge dependence of the random cluster measure on the complete graph, giving a partial solution to a well known conjecture of Grimmett, Winkler and Wagner. We than consider a natural problem concerning correlations between collection of connectivity events in graphs with respect to the random cluster measure and relate this natural problem to the Grimmett, Winkler and Wagner conjecture.

ÖZET

AYRIK OLASILIKTA NEGATİF BAĞIMLILIK: KAVRAMLAR, MODELLER VE SONUÇLAR

Bu tezde rasgele yığıntı ölçülerinin negatif bağımlılık özellikleri incelenmektedir. Öncelikle negatif korelasyon kavramlarını tanıtıp aralarındaki ilişkilere bakacağız. Bu kavramların modellerde kanıtlanmasının ne kadar zor olacağını gözlemleyeceğiz. Düzgün ağaç ve orman ölçüleri bu konudaki en önemli örneklerimiz olacak. Sonrasında rasgele yığıntı ölçülerinin düzgün ağaç ve orman ölçülerini nasıl genelleştirdiğini göreceğiz. Tam çizgelerdeki rasgele yığıntı ölçülerinin negatif bağımlı olduğunu kanıtlayacağız. Bu kanıt, Grimmett, Winkler ve Wagner'in bir savına kısmi bir çözüm vermektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS . . . **iii**

ABSTRACT . . . **iv**

ÖZET . . . **v**

LIST OF FIGURES . . . **ix**

1. INTRODUCTION . . . **1**

2. PRELIMINARIES . . . **7**

- 2.1. DISCRETE MEASURES AND SETS . . . 7
- 2.2. GRAPH THEORY . . . 9
  - 2.2.1. Operations on Graphs . . . 10
  - 2.2.2. Trees and Forests . . . 11
  - 2.2.3. Examples of Graphs . . . 12
    - 2.2.3.1. Complete Graphs . . . 12
    - 2.2.3.2. Complete Bipartite Graphs . . . 12
    - 2.2.3.3. Hypercube . . . 13
  - 2.2.4. Electrical Networks . . . 13
    - 2.2.4.1. Harmonic Functions . . . 14
- 2.3. MATROID THEORY . . . 15
  - 2.3.1. Examples of Matroids . . . 17
    - 2.3.1.1. Representable Matroids . . . 17
    - 2.3.1.2. Graphic Matroids . . . 18
    - 2.3.1.3. Uniform Matroids . . . 18
    - 2.3.1.4. Some Interesting Matroid Examples . . . 18
- 2.4. MARKOV CHAINS AND RANDOM WALKS . . . 19

3. POSITIVE ASSOCIATION . . . **23**

3.1. Monotonic Measures and Inequalities 23
3.2. Positive Association 27

4. NOTIONS OF NEGATIVE DEPENDENCE 30

4.1. Operations on Measures 30
4.1.0.1. Polynomial Conversion 31
4.2. Notions of Negative Dependence 33
4.2.1. NLC 33
4.2.1.1. h-NLC 34
4.2.1.2. h-NLC+ 34
4.2.2. p-NC 34
4.2.3. NA 35
4.2.3.1. CNA 35
4.2.3.2. CNA+ 35
4.2.4. RAYLEIGH MEASURES 36
4.2.4.1. PHR 37
4.2.4.2. STRONGLY RAYLEIGH 37
4.2.5. LORENTZIANITY 38
4.2.6. LOG CONCAVITY 39
4.2.7. STOCHASTIC COVERING PROPERTY 39
4.3. RELATIONS AMONG THE NOTIONS OF NEGATIVE DEPENDENCE 40
4.3.1. Implications 40
4.3.2. Counter Examples 47

5. MODELS 48

5.1. URN MODELS 48
5.2. UST, USF AND UCS MEASURES 50
5.2.1. UST 51
5.2.1.1. Random Walks on the Graph 51
5.2.2. USF 56
5.2.3. UCS 58

5.3. RANDOM CLUSTER MODEL 58
5.3.1. Basic Properties of Random Cluster Measure 60
5.3.1.1. Ising/Potts Model 60
5.3.1.2. UST,USF and UCS as limits of RCM 65
5.3.2. Positive Association 67

6. NEGATIVE EDGE DEPENDENCE IN RCM **69**

6.1. Connected Subgraphs of the Complete Graph 70
6.1.1. Asymptotics 72
6.2. The General Random cluster model for the Complete Graph at 1 . . . 73
6.3. Approaches 74
6.4. Adjacent Edges 75
6.5. Using bridges 80
6.5.1. Inequality for Non-adjacent Edges in K_n 83
6.5.2. Inequality for Adjacent Edges in K_n 86
6.6. Connection probabilities 87

7. CONCLUSION **91**

REFERENCES **93**

LIST OF FIGURES

Figure 2.1 A graph and one of its subgraphs. 10

Figure 2.2 G/e, $G+f$, and $G.e$. 11

Figure 2.3 A spanning tree and a spanning forest of the graph in Figure 2.1. 11

Figure 2.4 Complete Graphs K_3, K_4 and K_5. 12

Figure 2.5 Complete Bipartite Graphs $K_{2,2}$, $K_{3,3}$ and $K_{2,3}$. 13

Figure 2.6 The Hypercube Q^2 and Q^3. 13

Figure 2.7 An Electrical Network with a Battery on the Edge $e = xy$. . . . 14

Figure 2.8 Configuration of the Column Matroid A. 18

Figure 2.9 The Fano Configuration. 18

Figure 3.1 The lattice of Example 3.2.1. 29

Figure 4.1 Relations among notions of negative dependence. 46

Figure 5.1 K_4 and its Spanning Trees. 52

Figure 5.2 A graph. 53

Figure 5.3 SRW_A on the graph in Figure 5.2. 54

Figure 5.4 Random Cluster Model with $q = 2$ on a 2048×2048 box. 59

Figure 5.5 Ising Model at $T = 0.000611$. 61

Figure 5.6 Ising Model at $T = 2.333625$. 62

Figure 5.7 Conditional Measures of μ. 65

Figure 6.1 Both edges are bridges without a common component. 83

Figure 6.2 Both edges are bridges with a common component. 84

Figure 6.3 Connected graph with two bridges. 84

Figure 6.4 Joint bridges. 85

1. INTRODUCTION

A large part of classical probability theory deals with the properties of collections of independent random variables. Examples include central limit theorems, empirical process theory, stochastic processes with independent increments and point processes with trivial covariance matrices, the so called Poisson Point Processes and what will be most relevant for us, percolation theory.

There are of course several natural models which incorporate interaction between random variables that arise in mathematics, physics and computer science. Models such as ferromagnetic Ising and Potts models, Cox processes and several natural Markov Chains leads to ensembles of random variables that are positively correlated : If one of the random variables exceeds its mean, then the others exceed their mean as well. Positive correlation can be formalized in a number of different ways and significantly, these different formalisms often turn out to be closely related. One important and influential result in this area is the famous Fortuin-Kasteleyn-Ginibre (FKG) inequality, a local to global transfer theorem, that is of fundamental importance in the study of positive correlated (or in statistical physics language, ferromagnetic systems). Coupled with the FKG and related inequalities and intuitions from physics, the study of positively correlated ensembles of variables has been developed into a systematic and coherent theory.

Negatively correlated ensembles are widely encountered as well in a range of disciplines. Notable examples include Exclusion processes, Determinantal point processes and a large number of urn models. Another important and fundamental example arises from the study of electrical networks on graphs, an example that is closely related to Spanning Tree measures on graphs. In this regard, a famous theorem of Rayleigh from the 1850's says that conditioning on an edge being in a random spanning tree, any other edge is less likely to occur in the tree. This result has been extremely influential and has led to important results in combinatorics, probability, statistical physics and theoretical computer science. In probability, this is the foundation of the important theory of Loop Erased random walk, that was developed by Gregory Lawler in the 1980's following a suggestion of Peter Nelson. In Statistical physics, this leads to the celebrated result of Robin Pemantle that establishes a unique Gibbs measure for spanning trees in lattices. In theoretical computer science, this leads to the two famous algorithms for sampling random spanning trees, the first due to David Aldous and Andrei Broder that uses non backtracking walks and the second due to David Bruce Wilson that uses Loop Erased random walk. Further, Rayleigh's result on negative correlation was fundamental to recent advances on the Travelling Salesman Problem due to Shayan Oveis Gharan and co-authors.

However, when one seeks to generalize this result of Rayleigh, one encounters a major problem : Analogues of the FKG inequality fail and there are no useful local to global theorems. Indeed, for a number of reasons we will touch upon in this thesis, it is expected that there are no such theorems. Owing to this unfortunate fact, there is no overarching theory of negative correlation or dependence and instead, there are a large number of competing definitions, many of which are difficult to relate to others. It is intuitively clear that for any notion of negative correlation, a sum of a large number of negatively correlated random variables should exhibit central limit behaviour, but the correct level of generality where this can be proven is still open. Even more frustatingly, while showing that a system is positively correlated is usually routine, doing the same for negative correlation is non-trivial and indeed, there are a number of models where negative correlation is expected to hold, but for which these remain conjectures. Around 2000, Robin Pemantle laid out this problem in an influential article and called for a theory of negative dependence. A major advance on this was made in work by Julius Borcea, Petter Branden and Thomas Liggett in 2008, who introduced the class of Strongly Rayleigh measures. However, as we will see that there are several natural models that do not fit inside this framework and for which a number of basic questions remain open. This thesis deals with one of these models, the famous Random Cluster model of Cees Fortuin and Peter Kasteleyn.

Negative dependence is a property of a probability measure that arises naturally in many models, physical and mathematical, but that does not have a good theoretical background. The notion of negative dependence arises naturally as a counterpart of positive dependence. A probability measure is positively associated, if for any increasing functions, the expected value of the product of the functions is greater than or equal to the product of the individual expected values, and if they are always equal we say that the measure is independent, or a product measure. Intuitively, this may be thought as the functions supporting each other positively. The product under the probability measure is bigger than the individual ones. For instance, an increasing random variable is always positively associated with itself, since once we know one value the other goes according to the first one. If one increases so does the other and so forth. The beautiful part of the theory of positive association is that it is relatively straight-forward to prove that a probability measure is positively associated. This is due to Fortuin-Kasteleyn-Ginibre theorem, which states that upon proving an inequality, the FKG inequality, on the space which ties the levels of the lattice space, the probability measure is positively associated. This fact is really surprising and easy, for it discards any computation on functions.

On the other hand, the concept of negative association is not so straight-forward to show. First, we see that a random variable can not be negatively associated with itself. Hence, a need arises to define these concepts on disjoint subsets of coordinates which complicates things. Most importantly, the reverse inequality of FKG does not imply negative association. Thus, if a probability measure is negatively associated, the proof must be intricate involving functions and their expectations.

In this thesis, our main motivation is to study the concepts of negative dependence and their relations and the random cluster measure which is defined on the subgraphs of a graph and prove a negative dependence result for some classes of graphs. Throughout the years, different concepts of negative dependence has been developed, but there is no big framework like the positive counterpart. The notions arise from the models for which some kind of negative dependence is expected.

In Chapter Two, we define preliminary concepts: Probability measures and concepts, graph theoretic concepts and matroid theory. Graph theoretic concepts are used as a general framework for negative dependence notions. For instance, a natural framework is the spanning trees and spanning forests of a graph. It was proved that the edges e and f in a uniformly chosen spanning tree are negatively dependent. Since the spanning trees do not contain any cycles, it is reasonable to expect that the inclusion of an edge would drop the probability of the inclusion of another edge. This has been rigourously proven, which we follow, yet the same concept is still open for uniform spanning forests, which we mention as well. Matroids are introduced as a generalization of spanning trees and graphs, and due to the interesting examples and counter-examples they bring forth to the table.

In Chapter Three, we study basic properties of positive association in the spirit of showing the theoretical differences between two concepts that are very similar at the beginning, yet differ very much in their properties.

In Chapter Four, we get into different notions of negative dependence introduced in [1] and [2]. There are many generalizations of negative edge dependence and negative association concepts which arises both from definitions and from models. A natural approach here is to work with the properties of polynomials, in our case the polynomials will be the partition functions and generating polynomials of a discrete probability measure. We introduce in this section Rayleigh and Strongly Rayleigh polynomials and Lorentzian polynomials. Rayleigh and Strongly Rayleigh polynomials arises from electrical networks and their properties and are one of the strongest notions which we will mention. Lorentzian polynomials on the other hand have arised from algebraic geometry and have nice ties to discrete objects and what we study. After this relations among distinct notions of negative dependence, we explore some consequences: what kind of concentration inequalities exists when the random variables are negatively associated and so forth.

In Chapter Five, we delve into the world of models. We introduce two models in actuality in three sections. The first model is one where the theory flourished from almost, called Urn models. We follow the 1994 paper of Dubhashi and Ranjan [3], and prove negative association when the random variables involved are Bernoulli and have sum 1 all the time. As we will see, this is one of the most intuitive notion of negative dependence. The sum being constant, we expect that the probability of a urn being occupied falls lower whenever another urn is occupied. This is established using the existing theory and actually the FKG inequality which is intriguing. Later, we describe the models on graphs and their subgraphs. As mentioned above, spanning trees are an important class of objects in the theory of negative dependence. We prove that they satisfy negative edge dependence and also consider natural generalizations of this which are spanning forests and connected subgraphs. Grimmett asked the following conjecture in his book [4]:

Conjecture 1.0.1. For any finite graph $G = (V, E)$, the uniform spanning forest measure USF and the uniform connected subgraph measure UCS are negative-edge dependent.

We follow the spirit of this conjecture and actually work with another conjecture of Grimmett:

Conjecture 1.0.2. *Prove some notion of negative dependence in the Random cluster measure when $q < 1$.*

The foremost model we introduce and examine is Random cluster measure on the graphs. The exciting attribution of the random cluster measure is its generality and concreteness. Even though, it is relatively easy to define, it generalizes many models that are prominent in Statistical physics and at the same time, it is not easy so easy to explore due to the nature of its partition function and complexity. It is a natural extension of Ising/Potts models that are extensively studied, and the only difference between the random cluster measure and the well understood Erdős-Renyi measure is the introduction of a new parameter q which adds the connected components of a graph into consideration and changes everything.

It is proved, for instance in [4], that when $q > 1$, the random cluster measure is positively associated. We will see that the proof is relatively straight forward once we establish FKG inequality and the characterization of positive association with this simple inequality. However, there is almost nothing known when $q < 1$. It is expected as stated above in Grimmett's conjecture that the random cluster model will be negatively associated, yet there is no proof or partial proof of it.

In Chapter Six, we work towards proving negative edge dependence in random cluster measure when $q < 1$ on some classes of graphs. We define a parametric version of random cluster measure, and prove that it satisfies p-NC when $\lambda > 1$ and $q < 1$ when the underlying graph is a complete graph K_n or a complete bipartite graph $K_{r,s}$ and multipartite graph. We discuss the scope of our technique which is to study the partition function and translate the question into a counting problem.

2. PRELIMINARIES

2.1. DISCRETE MEASURES AND SETS

First, let us start by defining our sets. Let E be a finite set, and denote by 2^E, the power set of E and by $\{0,1\}^E$, we mean the $0-1$ sequences with the elements labeled from the set E. We write $[n]$ to denote the set $\{1,\ldots,n\}$

We will use $\{0,1\}^E$ and 2^E interchangeably, so let us mention the correspondence between those sets. First, we observe that any element $a \in \{0,1\}^E$ defines an element of $S \in 2^E$ via the natural mapping $e \in S$ if and only if $a(e) = 1$.

Example 2.1.1. *Let $E = \{1,2,3\}$, then the elements of the set $\{0,1\}^E$ and the corresponding sets are the following:*

$$\begin{aligned}
(0,0,0) &\longleftrightarrow \emptyset \\
(0,0,1) &\longleftrightarrow \{3\} \\
(0,1,0) &\longleftrightarrow \{2\} \\
(0,1,1) &\longleftrightarrow \{2,3\} \\
(1,0,0) &\longleftrightarrow \{1\} \\
(1,0,1) &\longleftrightarrow \{1,3\} \\
(1,1,0) &\longleftrightarrow \{1,2\} \\
(1,1,1) &\longleftrightarrow \{1,2,3\}.
\end{aligned}$$

The first entry corresponds to the set element 1, and so forth.

We observe that the set $\{0,1\}^E$ or 2^E is a partially ordered set, with entry-wise comparison. That is, we say $x \leq y$ if $x(e) \leq y(e)$ for all $e \in E$. Let $(A, \leq)$ be a partially ordered set and $S \subset A$. An element x is said to be an **upper bound** of S, if we have $s \leq x$ for all $s \in S$. $x \in A$ is said to be the **least upper bound** if it is an upper bound and for all y which is an upper bound of S we have $x \leq y$. We denote the least upper bound of two elements x, y by $x \vee y$, i.e. $x \vee y(e) = \max\{x(e), y(e)\}$. An element x is said to be an **lower bound** of S, if we have $s \geq x$ for all $s \in S$. $x \in A$ is said to be the **greatest lower bound** if it is a lower bound and for all y which is a lower bound of S we have $x \geq y$. We denote the greatest lower bound of two elements x, y by $x \wedge y$, i.e. $x \wedge y(e) = \min\{x(e), y(e)\}$. A partially ordered set A where each pair of elements (x, y) has a greatest lower bound and a least upper bound is called a **lattice**. With the above order, $\{0,1\}^E$ is a lattice.

The operations of **switching on** and **switching off** for an element $x \in \{0,1\}^E$ and $e \in E$ is respectively defined as follows:

$$x^e(f) = \begin{cases} x(f), & \text{if} f \neq e \\ 1, & \text{if} f = e. \end{cases} \tag{2.1}$$

$$x_e(f) = \begin{cases} x(f), & \text{if} f \neq e \\ 0, & \text{if} f = e. \end{cases} \tag{2.2}$$

Remark 2.1.1. *The term "switching on e" denoted by x^e and "switching off e" comes from graphs. We think of the configuration and adding or deleting, respectively, the edge e to the configuration.*

We define a metric on $\{0,1\}^E$ which is called the **Hamming distance** by $d_H(x, y) := |\{i : x_i \neq y_i\}|$. We say that μ is a **probability measure** if it is a measure on the finite set E and $\mu(E) = 1$. Note that for any finite measure μ' on E, we can define a probability measure by setting $\mu(A) = \frac{\mu'(A)}{\mu'(E)}$, for any $A \subset E$. We call a probability measure μ on Σ **strictly positive** if $\mu(x) > 0$ for all $x \in \Sigma$.

An event $\mathcal{A}$ is said to be an **increasing event** if it is upward-closed. That is, if $x \in \mathcal{A}$, then $y \in \mathcal{A}$ if $x \leq y$. In terms of subsets, this becomes, if $S \in \mathcal{A}$, then for any $S \subset T$, $T \in \mathcal{A}$.

We define the total variation distance between two measures μ, ν on $(\Sigma, \mathcal{A})$.

Definition 2.1.1. Let μ, ν be two discrete measures on $(\Sigma, \mathcal{A})$. Then the **total variation distance**, denoted by $d_{\mathrm{TV}}(\mu, \nu)$ is defined as

$$d_{\mathrm{TV}}(\mu, \nu) = \sup_{A \in \mathcal{A}} |\mu(A) - \nu(A)|. \tag{2.3}$$

Remark 2.1.2. *The total variation distance on a discrete set Σ is equal to*

$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{x \in \Sigma} |\mu(x) - \nu(x)|. \tag{2.4}$$

2.2. GRAPH THEORY

A **graph** is a pair of sets (V, E), where $E \subset V^2$. We call V the **vertex set** of the graph, and E the **edge set** of the graph. Let $x, y \in V$ be vertices, then $e = \langle x, y \rangle$ is the edge that connects x and y. If there is no confusion, we will denote the edge e by xy as well. Vertices x, y are called **adjacent** if $\langle x, y \rangle \in E$. This defines a relation which we denote by $x \sim y$. If x, y are not adjacent, then we write $x \not\sim y$. Since we are not considering ordered pairs, $\langle x, y \rangle$ mean the same edge. We can draw a graph $G = (V, E)$ by putting points in place of the vertices and draw the corresponding lines between the vertices for the edges of the graph.

A graph $G' = (V', E')$ with $V' \subset V$, and $E' \subset E$ is called a **subgraph** of the graph $G = (V, E)$.

If G' contains all the edges in E among the vertices of V', G' is called the **induced subgraph** of G by V', and is denoted by $G[V']$. If $V' = V$, then the subgraph is called **spanning**.

A **path** is graph of the form

$$V = \{v_0, ..., v_n\} \quad E = \{v_0v_1, v_1v_2, ..., v_{n-1}v_n\}.$$

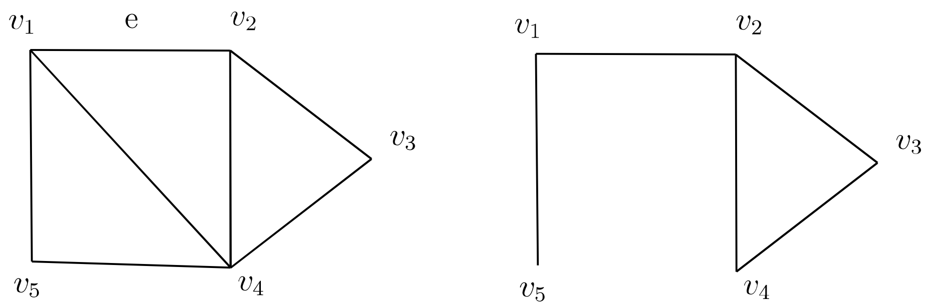


Figure 2.1. A graph and one of its subgraphs.

A path which starts and ends at the same vertex is called a **cycle**. We will say that a graph a path exists between x and y in a graph G, if there is a path graph P so that P is a subgraph of G. A graph is called **connected** if there is a path between any of its vertices. A **connected component** or simply component of a graph is one of its maximally connected subgraphs.

2.2.1. Operations on Graphs

We define three basic operations on the edges of a graph: deletion, addition and contraction.

Let $G = (V, E)$ be a finite graph with $e = xy \in E$. We define the graph $G/\{e\}$ (G/e when there is no confusion) to be the graph obtained when we delete the edge $e = xy$.

Let $G = (V, E)$ be a finite graph with $e = xy \notin E$. We define the graph $G + e$ or $G + xy$ to be the graph obtained when we add the edge e to the graph G.

Let $G = (V, E)$ be a finite graph with $e = xy \in E$. We define the graph $G.e$ as the multi-graph obtained when we identify the incident vertices x and y of the edge e as one vertex.

Example 2.2.1. *Let $G = (V, E)$ and $e = v_1v_2$ be the graph and the edge which is in Figure 1. Also, let $f = v_2v_5$ be another edge. Then the following figure shows G/e, $G + f$ and $G.e$ in order.*

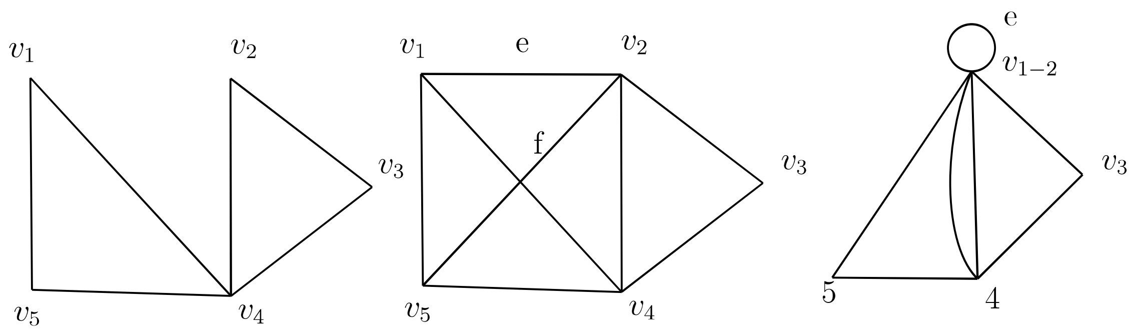


Figure 2.2. G/e, $G+f$, and $G.e$.

2.2.2. Trees and Forests

A graph which contains no cycles is called a **forest**. A connected forest is called a **tree**. We will generally denote a forest by $\mathcal{F}$, and a tree by $\mathcal{T}$. Let us see some properties of a tree.

Proposition 2.2.1. *The following are equivalent.*

(i)$\mathcal{T}$ *is a tree.*

(ii)Any two vertices x, y *of the graph* $\mathcal{T}$ *is connected by a single path* P.

(iii)$\mathcal{T}$ *is connected and has* $n-1$ *edges.*

(iv)$\mathcal{T}$ *doesn't contain a cycle and has* $n-1$ *edges.*

(v)$\mathcal{T}$ *is connected and for any edge* $e \in E(\mathcal{T})$, $\mathcal{T}/\{e\}$ *is disconnected.*

(vi)$\mathcal{T}$ *doesn't contain any cycle, and for any* $x \not\sim y$, *the graph* $\mathcal{T}+\langle x, y\rangle$ *contains a cycle.*

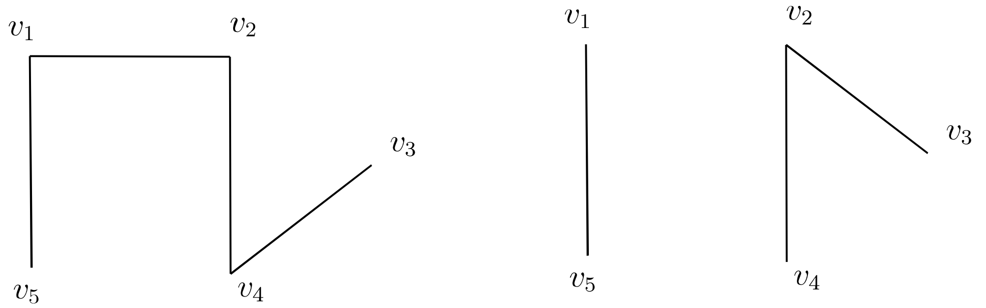


Figure 2.3. A spanning tree and a spanning forest of the graph in Figure 2.1.

These give some characterizations for the trees which we will use. We say that a subgraph $\mathcal{T}$ of $G = (V, E)$ is called a **spanning tree** if it is a tree and $V(\mathcal{T}) = V$. We will see that spanning trees will be an important example of a model for negative dependence.

2.2.3. Examples of Graphs

We give some classes of graphs that we will be working with.

<u>2.2.3.1. Complete Graphs.</u> Let n be the number of vertices that we will construct our graph on. The complete graph on n vertices is the one where every possible edge is present. Therefore, we have $\binom{n}{2}$ edges. We denote the complete graph on n vertices by K_n.

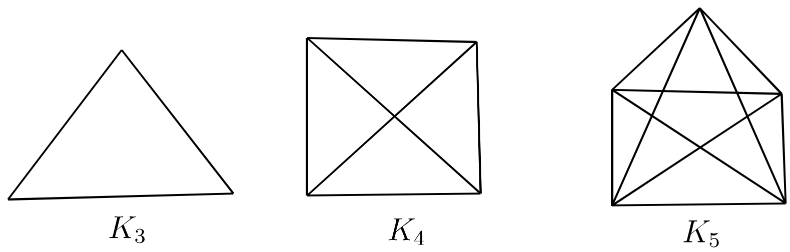


Figure 2.4. Complete Graphs K_3, K_4 and K_5.

Theorem 2.2.1 (Cayley's Theorem). *The complete graph K_n has n^{n-2} spanning trees.*

<u>2.2.3.2. Complete Bipartite Graphs.</u> Another class of graphs is the bipartite graphs. A bipartite graph $G = (V, E)$ is a graph whose vertices can be partitioned into two sets V_1 and V_2 with $V_1 \dot{\cup} V_2 = V$ and the edge set $E \subset V_1 \times V_2$. The complete bipartite graph is the one where each edge exists between the vertex set V_1 and V_2. We denote the complete bipartite graphs with $|V_1| = r$ and $|V_2| = s$ by $K_{r,s}$.

Remark 2.2.1. *If the vertex set is partitioned into m parts $(V_i)_{i=1}^m$, with $\dot{\cup}_{i=1}^m V_i = V$, and the edges are between the distinct sets V_i, V_j for $i \neq j$, we call it a multi-partite graph. The complete multi-partite graph where all possible edges exist with $|V_i| = r_i$ is denoted by $K_{r_1,\ldots,r_m}$.*

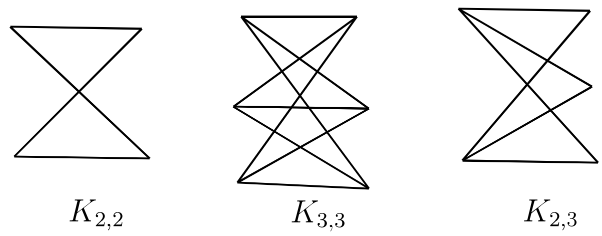


Figure 2.5. Complete Bipartite Graphs $K_{2,2}$, $K_{3,3}$ and $K_{2,3}$.

2.2.3.3. Hypercube. Let $V = \{0, 1\}^n$. Also let the edge set E be defined according to the hamming distance being one or not, i.e. if $d_H(x, y) = 1$, $x \sim y$, otherwise $x \not\sim y$. We call the graph $G = (V, E)$ n-hypercube (or n Boolean cube), and denote it as Q^n (or $\mathcal{B}^n$).

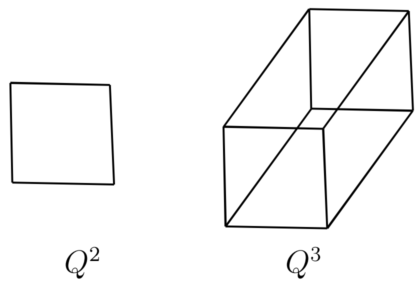


Figure 2.6. The Hypercube Q^2 and Q^3.

2.2.4. Electrical Networks

Electrical networks in general are physical objects that is formed by a voltage source and resistors. We will think of it as a **weighted graph**. Let $w : E \to \mathbb{R}^+$ be a function, and $R : E \to \mathbb{R}^+$ be the resistance.

An **electrical network** is graph with a function attached to its edges. The reciprocal of the resistance is called **conductance**. We think of the weight function w as the conductance that is attached to the edge.

For a specific edge $e = xy$, we consider a battery that is connected to its terminals, which makes the current flow.

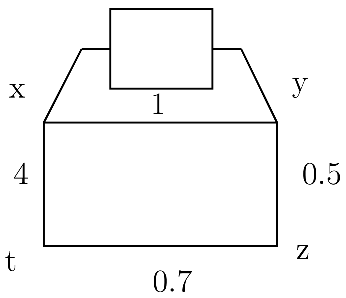


Figure 2.7. An Electrical Network with a Battery on the Edge $e = xy$.

The electrical properties of such a graph is governed by Kirchoff's laws which we state. The first Kirchoff law is that to each vertex, there is a real number attached which is its voltage. We will think of it as a function $v : V \to \mathbb{R}$. The second Kirchoff law is that to each oriented edge, there is a real number attached which is its current. We think of it as a function $i : E \to \mathbb{R}^+$. Since we are working on oriented edges, we have anti-symmetry for the function i, that is $i(xy) = -i(yx)$. The third Kirchoff law is the one concerning the current flows. For any vertex, the sum of the currents incoming should equal to the sum of the current outgoing from the vertex, i.e.

$$\sum_{v \sim x} i(vx) = 0. \tag{2.5}$$

Let $u, v \in V$, and $e = uv$. Then the second Kirchoff law states the following

$$[V(u) - V(v)] = i(e)R(e). \tag{2.6}$$

Combining the second law and the third law, we have that

$$0 = \sum_{v \sim x} i(vx) = \sum_{v \sim x} [V(v) - V(x)]w(xy) = V(v)d(v) - \sum_{v \sim x} V(x)w(xy). \tag{2.7}$$

The above equation leads us to define harmonic functions.

<u>2.2.4.1. Harmonic Functions.</u> Let f be a function on the vertex set V of a weighted graph $G = (V, E, w)$.

Definition 2.2.1. The **excess** of a function f at a vertex $v \in V$is the defined as

$$0 = \sum_{v \sim x} i(vx) = \sum_{v \sim x} [V(v) - V(x)]w(xy) = V(v)d(v) - \sum_{v \sim x} V(x)w(xy). \tag{2.8}$$

Definition 2.2.2. A function $f : V \to \mathbb{R}$ is said to be **harmonic** at the vertex v if the excess of f at v is 0.

We state some theorems apropos to harmonic functions, which we do not prove.

Theorem 2.2.2 (Maximum Principle)**.** *Let $f : V \to \mathbb{R}$ be a function on a connected finite graph $G = (V, E)$ that is harmonic except possibly on a finite set of vertices $S = \{v_1, ..., v_k\}$. Then, f attains its maximum and minimum on the set S. If f is harmonic on all of V, then it is constant.*

Theorem 2.2.3. *Let $f, g : V \to \mathbb{R}$ be two functions on a finite, connected, weighted graph $G = (V, E, w)$. Suppose $f(v_i) = g(v_i) = c_i$ on a set of vertices $S = \{v_1, ..., v_k\}$, and that f and g are harmonic on the set V/S. Then $f = g$.*

We will use harmonic functions to tie electrical networks and spanning trees which gives us a good relation among distinct subjects.

Remark 2.2.2. *Note that the equation*

$$V(v)d(v) = \sum_{v \sim x} V(x)w(xy), \tag{2.9}$$

from the previous section actually shows that the Voltage function V is harmonic.

2.3. MATROID THEORY

Matroids are objects that generalize finite matrices and finite graphs in a way. We define what a matroid is and some basic properties and basic examples of matroids. Since, they naturally generalize graphs and spanning trees, matroids and their bases will provide us with a good background for negative dependence properties and notions. Some counter-examples will be given in the next sections for some conjectural implications among negative dependence notions.

Definition 2.3.1. A **Matroid** is $\mathcal{M} = (E, \mathcal{I})$ consisting of a finite set E, called the **ground** set, and a collection $\mathcal{I}$ of subsets of E called the **independent sets** of the matroids satisfying the following:

(i)$\emptyset \in \mathcal{I}$,

(ii)If $I \in \mathcal{I}$, and $I' \subset I$, then $I' \in \mathcal{I}$,

(iii)If $I_1, I_2 \in \mathcal{I}$, and $|I_1| < |I_2|$, then there is an element $e \in I_2 \setminus I_1$ so that $I_1 \cup \{e\} \in \mathcal{I}$.

A subset that is not in $\mathcal{I}$ is called **dependent**. We will call a minimal dependent subset, that is a dependent subset of E whose subsets are all independent, a **circuit**.

Remark 2.3.1. *Circuits determine the independent sets of a matroid and independent sets determine the circuits of a matroid. Thus, they are somewhat exchangeable.*

Definition 2.3.2. A maximal independent subset of E is called a **basis** of the matroid $\mathcal{M} = (E, \mathcal{I})$. The set of basis of a matroid is called the **base set** and is denoted by $\mathcal{B}$.

Theorem 2.3.1. *(i)Let $B_1, B_2 \in \mathcal{B}$ for a matroid $\mathcal{M}$. Suppose $e \in B_1 \setminus B_2$. Then there exists an element $f \in B_2 \setminus B_1$ such that $(B_1 \setminus e) \cup f \in \mathcal{B}$.*

(ii)Let $B_1, B_2 \in \mathcal{B}$ for a matroid $\mathcal{M}$. Then $|B_1| = |B_2|$. That is all the members of the base set have the same cardinality.

Remark 2.3.2. *The first part of the theorem can be thought as the basis exchange lemma. We can remove an element from one basis, and add another one from another basis to get a new basis. This all become clear when we think of the basis elements of a vector space.*

We now define two important notions for matroids: the restriction and the rank of a matroid.

Definition 2.3.3. Let $A \subset E$. Consider the set $\mathcal{I}|A := \{I \subset A : I \in \mathcal{I}\}$. Together with A, the pair $(A, \mathcal{I}|A)$ becomes a matroid denoted by $\mathcal{M}|A$, and is called the **restriction of $\mathcal{M}$ to** A. We define the **rank of a subset** A, denoted by $\mathrm{rk}(A)$, to be the cardinality of the basis of the matroid $\mathcal{M}|A$.

Remark 2.3.3. *Note that rank can be thought of as a function* $\mathrm{rk} : 2^E \to \mathbb{Z}^+$. *Thus, we can attach the rank for an element in* $\{0,1\}^E$ *naturally. It will be seen that rank is almost the same as the number of connected components of a subgraph.*

2.3.1. Examples of Matroids

2.3.1.1. Representable Matroids. A natural class to consider is vectors. The definitions of independent sets and basis is derived from the theory of linear matroids. A **representable matroid** is a matroid $\mathcal{M} = (E, \mathcal{I})$ where the independent sets are the independent sets of a vector space over a field $\mathbb{F}$, and the ground set is the vector space. We denote the matroid by writing a matrix and considering its column vectors. Let us see an example.

Example 2.3.1. *Consider the matrix*

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

We think of its column vector over a field. Say $\mathbb{R}$. *We can represent this matroid as a configuration in the following way:*

- *Each column of the matrix is a point in our Matroid, i.e. a column vector* $x \in E$,
- *If three vectors* x, y *and* z *are linearly dependent, then they will be collinear in our configuration.*

Considering this we have the following configuration for the matroid A:

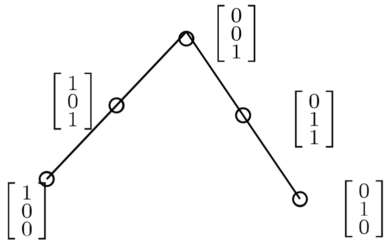


Figure 2.8. Configuration of the Column Matroid A.

2.3.1.2. Graphic Matroids. Another natural class to consider is graphic matroids. A **graphic matroid** is a matroid $\mathcal{M} = (E, \mathcal{I})$, with the ground set being the edges of a graph and the independent sets are the spanning forests of the given graph. We state a known theorem:

Theorem 2.3.2. *Every graphic matroid is representable.*

2.3.1.3. Uniform Matroids. Let E be a finite set with $|E| = n$, and let $\mathcal{I} = \{I \subset E : |I| \leq m\}$. Note that $m \leq n$. Then we call the matroid $\mathcal{U}_{m,n} = (E, \mathcal{I})$, the uniform matroid of size m over a set of size n.

2.3.1.4. Some Interesting Matroid Examples. 1)Fano Matroid. The Fano configuration of projective geometry gives us a matroid. It is denoted by F_7.

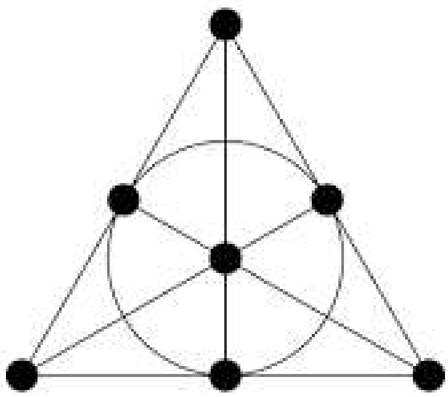


Figure 2.9. The Fano Configuration.

The Fano configuration can be represented as the following matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

2) $\mathcal{S}_8$. This matroid will be an important one for some of the counter-examples. We give its matrix form:

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

We state some of its properties:

Remark 2.3.4.

- *It is $\mathbb{F}$ representable, if and only if the characteristic of the field $\mathbb{F}$ is two.*
- *It is not graphic.*
- *It has a unique element x and a unique element $y \neq x$ such that $\mathcal{S}_8 \setminus x \cong F_7*$ and $\mathcal{S}_8 \setminus y \cong F_7$.*

2.4. MARKOV CHAINS AND RANDOM WALKS

We define the basics of Markov chains and random walks on a graph $G = (V, E)$. A **markov chain** is a process that satisfies a memoryless property called the **Markov property** on a finite set $\mathcal{S}$ with transition probabilities $p(x, y)$. We call $\mathcal{S}$ as the **state space**, and the elements $s \in \mathcal{S}$ as **states**. More precisely we have

Definition 2.4.1. Let P be a $k \times k$ matrix with elements $\{p_{i,j} : i, j \in [k]\}$. A random process $X = (X_0, X_1, \ldots)$ with the finite state space $S = \{s_1, \ldots, s_k\}$ is said to be a **(homogeneous) Markov chain with transition matrix** P, if for all n and $i, j \in [k]$, and all $i_0, \ldots, i_{n-1} \in [k]$, we have the **Markov property**,

$$\begin{aligned}\mathbb{P}(X_{n+1} = s_j | X_0 = s_{i_0}, X_1 =& s_{i_1} \ldots X_{n-1} = s_{i_{n-1}}, X_n = s_i) \\ =& \mathbb{P}(X_{n+1} = s_j | X_n = s_i) \\ =& p_{i,j}.\end{aligned}$$

A **random walk** on a graph $G = (V, E)$ is a Markov chain with the vertex set as state space and transition probabilities defining the walk of a particle on the vertices of the graph.

Theorem 2.4.1. *For a Markov chain $X = (X_0, \ldots)$ with state space $S = \{s_1, \ldots, s_k\}$, transition matrix P, and a given initial distribution μ_0, we have for any time n, the distribution μ_n of X_n as follows: $\mu_n = \mu_o P^n$.*

Now, we define some basic properties of a Markov chain.

Definition 2.4.2. A Markov chain X with transition matrix P is called **irreducible** if for any two state $s_i, s_j \in S$, there is an integer $m \in \mathbb{Z}$ such that $P^m_{i,j} > 0$.

Definition 2.4.3. Let $\mathcal{T}(s_i) = \{t \geq 1 : P^t_{i,i} > 0\}$. The **period** of the state s_i of the Markov chain is defined as the greatest common divisor of the elements in $\mathcal{T}(s_i)$, denoted as $\gcd \mathcal{T}(s_i)$.

Proposition 2.4.1. *If the Markov chain X is irreducible, then $\gcd \mathcal{T}(s_i) = \gcd \mathcal{T}(s_j)$ for any $s_i, s_j \in S$.*

Definition 2.4.4. A Markov chain X is said to **aperiodic** if the period of any state is 1.

Example 2.4.1. *We define a Markov chain on the graph which is in Figure 1. A natural way to define a simple random walk on a graph is to give the following probabilities:*

$$p_{i,j} = \begin{cases} \dfrac{1}{deg(v_i)}, & \text{if } v_i \sim v_j \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

Then for our example, the transition matrix becomes:

$$P = \begin{pmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

We observe that since our graph is connected, the corresponding simple random walk is irreducible.

Definition 2.4.5. We call a distribution η, a **stationary distribution** for the Markov chain X with transition matrix P, if η is a probability measure on the state space S, and it satisfies $\eta P = \eta$.

We state one of the main theorems of finite state Markov chains.

Theorem 2.4.2. *For an irreducible, aperiodic Markov chain X, there exists at least one stationary distribution.*

Another important theorem in finite state Markov chain is the convergence to the stationary distribution under suitable conditions.

Theorem 2.4.3. *Let X be an irreducible, aperiodic Markov chain with transition matrix P and an initial distribution μ_0. Then, for any distribution η which is stationary we have as $n \to \infty$, $\mu_n \to_{TV} \eta$.*

Remark 2.4.1. *This also shows the uniqueness of the stationary distribution under suitable assumptions for the Markov chain, i.e. aperiodicity and irreducibility.*

Definition 2.4.6. *Let a Markov chain X with transition matrix P be given. A distribution η on the state space of the Markov chain is said to be* ***reversible*** *if it satisfies the* ***detailed balance equations****: $\eta_i P_{i,j} = \eta_j P_{j,i}$, for all $i, j \in [k]$.*

Remark 2.4.2. *We call a Markov chain reversible, if such a distribution exists.*

Theorem 2.4.4. *If a Markov chain X is reversible, say with the distribution η, then η is a stationary distribution for X.*

3. POSITIVE ASSOCIATION

3.1. Monotonic Measures and Inequalities

We prove some inequalities concerning the ordering of measures and some correlation inequalities which will be used in the following sections.

Theorem 3.1.1 (Holley Inequality). *Let μ_1 and μ_2 be strictly positive measures on the finite space $(\Sigma, \mathcal{A})$ such that*

$$\mu_2(x_1 \vee x_2)\mu_1(x_1 \wedge x_2) \geq \mu_1(x_1)\mu_2(x_2), \quad x_1, x_2 \in \Sigma. \tag{3.1}$$

Then,

$$\mathbb{E}_{\mu_1}[X] \leq \mathbb{E}_{\mu_2}[X], \tag{3.2}$$

for increasing functions $X : \Sigma \to \mathbb{R}$.

Proof. We will prove the inequality using Markov chains. The main idea is to construct a coupling of μ_1 and μ_2 which is supported on the sub-diagonal regime, that is on the set $\{(x, y) \in \Sigma \times \Sigma : x \leq y\}$.

First, let us see how to construct a Markov chain with a given measure μ as its stationary distribution. Let $\Sigma = \{0, 1\}^E$, and μ be a strictly positive probability measure on $(\Sigma, \mathcal{A})$. Define a Markov chain $X : \Sigma^2 \to \mathbb{R}$ on the state space Σ, as follows:

$$\begin{aligned} X(x_e, x^e) &= 1, \\ X(x^e, x_e) &= \frac{\mu(x_e)}{\mu(x^e)}, \\ X(x, y) &= 0, \quad \text{for any other}(x, y) \in \Sigma^2. \end{aligned}$$

The diagonal entries $X(x, x)$ is chosen so that $\sum_{y \in \Sigma} X(x, y) = 0$ for all $x \in \Sigma$.

Remark 3.1.1. *Let us observe how the Markov chain operates between states. Suppose we are on a state where the edge e is closed. Then we switch on the edge with rate 1. Suppose we are on a state where the edge e is open, then we switch off the edge with rate* $\frac{\mu(x_e)}{\mu(x^e)}$.

We first observe that the Markov chain defined is irreducible and aperiodic. Since, we are working with switching on and off operators, we can always reach $\mathbf{1}$, all ones state by switching on the appropriate elements of x with positive probability. Then, again with positive probability, we can switch off the appropriate edges of $\mathbf{1}$ to reach another desired state y. Hence, there exists a t so that $P^t_{x,y} > 0$. Actually, with a little deliberation, we can see that t can be taken as $t = d_H(x, y)$. Since, $P_{x,x}$ is positive, we see that X is aperiodic as well. Hence, there exists a stationary distribution.

To prove that μ is the stationary distribution, we show that μ satisfies the detailed balance equations. That is, we need to show that

$$\mu(x)X(x, y) = \mu(y)X(y, x). \tag{3.3}$$

We only need to focus on terms of the form (x_e, x^e). Thus,

$$\begin{aligned}\mu(x_e)X(x_e, x^e) &= \mu(x_e),\\ \mu(x^e)X(x^e, x_e) &= \mu(x^e)\frac{\mu(x_e)}{\mu(x^e)} = \mu(x_e).\end{aligned}$$

Thus μ is the stationary distribution of the given Markov chain.

Now, we will perform this procedure for two measures μ_1 and μ_2 that satisfy the hypotheses of the theorem. Let S be the set of all ordered pairs (x, y) of configurations in $\Sigma = \{0,1\}^E$ satisfying $x \leq y$. Define the function $H : \Sigma \times \Sigma \to \mathbb{R}$ by

$$\begin{aligned} H(x_e, y; x^e, y^e) &= 1, \\ H(x, y^e; x_e, y_e) &= \frac{\mu_2(y_e)}{\mu_2(y^e)}, \\ H(x^e, y^e; x_e, y^e) &= \frac{\mu_1(x_e)}{\mu_1(x^e)} - \frac{\mu_2(y_e)}{\mu_2(y^e)}, \end{aligned}$$

for all $(x, y) \in S$ and $e \in E$. Also, set $H = 0$ for all other off-diagonal entries.

The diagonal terms $H(x, y; x, y)$ are chosen so that we have

$$\sum_{(t,s)\in S} H(x, y; t, s) = 0, \quad (x, y) \in S. \tag{3.4}$$

Remark 3.1.2. Let us observe how H operates. We see from the first equation that, if e is not in x, we add the edge e, switch on e with rate 1, and add it to y as well if e is not present in y. The second equation gives the rate of switching off the edge e from x and y where e is present in y. The third equation gives the rate for an edge e which is present in both x and y and we switch off e in x but not from y. Note that

$$\mu_2(x_1 \vee x_2)\mu_1(x_1 \wedge x_2) \geq \mu_1(x_1)\mu_2(x_2), \quad x_1, x_2 \in \Sigma. \tag{3.5}$$

by the inequality given in the theorem, this rate is positive as well.

Let $(Y_t, Z_t)_{t\geq 0}$ be a Markov chain on S with generator H,, and set $(Y_0, Z_0) = (\mathbf{0}, \mathbf{1})$. We write $\mathbb{P}$ for the appropriate probability measure. The transitions preserve the ordering of the components of the states, hence we may assume that the chain satisfies

$$\mathbb{P}(Y_t \leq Z_t \text{ for all } t) = 1. \tag{3.6}$$

Observe that by the definition of the generator H, $(Y_t : t \geq 0)$ is a Markov chain with generator μ_1 and $(Z_t : t \geq 0)$ is a Markov chain with generator μ_2.

Let η be a stationary distribution for the paired chain $(Y_t, Z_t)_{t\geq 0}$. Since Y and Z have unique stationary distributions μ_1 and μ_2 respectively by the first part of the proof, we see that η has marginal measures μ_1 and μ_2. Since $\mathbb{P}(Y_t \leq Z_t \text{ for all} t) = 1$, we have

$$\eta(S) = \eta\left(\{(x, y) : x \leq y\}\right) = 1, \tag{3.7}$$

and η is the coupling of the measures μ_1 and μ_2 that is desired.

Now, let $(x, y) \in S$ be chosen according to the measure η. Then we have

$$\mathbb{E}_{\mu_1}[f] = \eta(f(x)) \leq \eta(f(y)) = \mathbb{E}_{\mu_2}[f], \tag{3.8}$$

for any increasing function f on Σ. Therefore, we have the desired result, i.e. $\mu_2 \succeq \mu_1$.

■

Theorem 3.1.2. *A pair μ_1, μ_2 of strictly positive probability measures on $(\Sigma, \mathcal{A})$ satisfies 3.1 if and only if it satisfies the following two inequalities:*

$$\mu_2(x^e)\mu_1(x_e) \geq \mu_1(x^e)\mu_2(x_e), \quad x \in \Sigma, e \in E,$$
$$\mu_2(x^{ef})\mu_1(x_{ef}) \geq \mu_1(x^e_f)\mu_2(x^f_e), \quad x \in \Sigma, e, f \in E.$$

Remark 3.1.3. *The Holley inequality 3.1 is equivalent to a monotonicity condition in one point conditional distributions. We state it as a theorem below.*

Theorem 3.1.3. *Let μ_1 and μ_2 be two strictly positive probability measures on $(\Sigma, \mathcal{A})$. They satisfy the Holley inequality 3.1 if and only if the one point conditional distributions satisfies*

$$\begin{aligned}&\mu_2\left(x(e) = 1 | x(f) = \xi(f) \textit{ for all } f \in E \setminus e\right)\\ \geq &\mu_1\left(x(e) = 1 | x(f) = \psi(f) \textit{ for all } f \in E \setminus e\right),\end{aligned}$$

for all $e \in E$, and all pairs $\xi, \psi \in \Sigma$ satisying $\psi \leq \xi$.

3.2. Positive Association

We start with defining the notion of Positive association. Let E be a finite set, and $\Sigma = \{0,1\}^E$. We say that a measure μ is positively associated if for all increasing functions f, g, we have

$$\mathbb{E}_\mu[fg] \geq \mathbb{E}_\mu[f]\mathbb{E}_\mu[g]. \tag{3.9}$$

We can write this as

$$\int fg d\mu \geq \int f d\mu \int g d\mu. \tag{3.10}$$

The beauty and completeness of positive association comes from the fact that one can check if a measure is positively associated by only checking a lattice condition on the space Σ, which is called the positive lattice condition.

Definition 3.2.1. A measure μ is said to satisfy the Positive Lattice Condition if we have

$$\mu(x \vee y)\mu(x \wedge y) \geq \mu(x)\mu(y) \quad \forall x, y \in \Sigma. \tag{3.11}$$

Theorem 3.2.1 (FKG)**.** *If a strictly positive measure μ on Σ satisfies the positive lattice condition, it is positively associated.*

Proof. We will prove the theorem using Holley's inequality. Now, assume that μ satisfies the positive lattice condition, and let f and g be increasing functions. Define a new function $\tilde{g} = g + a$ where $a > 0$. We observe that

$$\mathbb{E}_\mu[f\tilde{g}] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[\tilde{g}] = \mathbb{E}_\mu[fg] - \mathbb{E}_\mu[f]\mathbb{E}_\mu[g]. \tag{3.12}$$

Hence, by choosing $a > 0$ sufficiently large, we can assume that g is a strictly positive function on Σ. To use Holley's inequality, we need to have two measures on the space

$(\Sigma, \mathcal{F})$. We define $\mu_1 := \mu$, and

$$\mu_2(\sigma) := \frac{g(\sigma)\mu(\sigma)}{\sum_{\sigma\in\Sigma} g(\sigma)\mu(\sigma)}, \tag{3.13}$$

for all $\sigma \in \Sigma$. By Holley's inequality, we have $\mathbb{E}_{\mu_1}[f] \leq \mathbb{E}_{\mu_2}[f]$ for all increasing functions f. Thus,

$$\sum_{\sigma\in\Sigma} f(\sigma)\mu(\sigma) \leq \frac{\sum_{\sigma\in\Sigma} f(\sigma)g(\sigma)\mu(\sigma)}{\sum_{\sigma\in\Sigma} g(\sigma)\mu(\sigma)}, \tag{3.14}$$

which is exactly

$$\mathbb{E}_\mu[fg] \geq \mathbb{E}_\mu[f]\mathbb{E}_\mu[g]. \tag{3.15}$$

■

Remark 3.2.1.
- *For a strictly positive measure μ, it is sufficient that μ satisfies the positive lattice condition, but it is not necessary.*
- *We will see that negative lattice condition which will be defined analogously is not sufficient. This is what makes negative dependence more tricky.*

An example is proposed by J. Steif is the following:

Example 3.2.1. Let $a, b \in (0,1)$, and let μ_0, μ_1 be the probability measure on $\{0,1\}^3$ given by

$$\mu_0(010) = \mu_0(001) = a, \quad \mu_0(000) = 1 - 2a;$$
$$\mu_1(111) = \mu_1(100) = \frac{1}{2}.$$

Set $\mu = b\mu_0 + (1-b)\mu_1$. We have

$$\mu(010) = \mu(001) = ab, \;\; \mu(000) = (1-2a)b$$
$$\mu(100) = \mu(111) = \frac{1-b}{2}, \;\; \mu(011) = \mu(101) = \mu(110) = 0.$$

μ does not satisfy the positive lattice condition, but is positively associated.

We can visualize the lattice as follows:

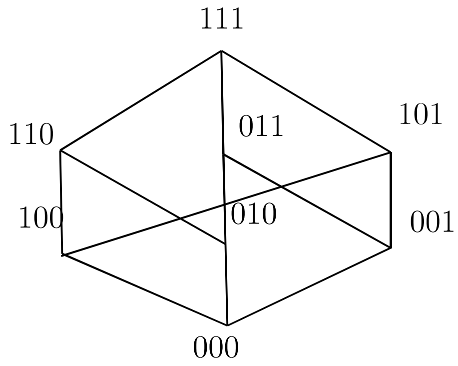


Figure 3.1. The lattice of Example 3.2.1.

The measure μ in the example 3.0.1 is not strictly positive, but taking a convex combination of μ, and a new measure, we can make it strictly positive as well. Lastly, we state a theorem which makes it even easier to check the FKG inequalities.

Theorem 3.2.2. *Let μ be a strictly positive measure on $(\Sigma, \mathcal{A})$. μ satisfies the positive lattice condition if and only if the inequality in the positive lattice condition holds for all incomparable pairs $x_1, x_2 \in \Sigma$ with $d_H(x_1, x_2) = 2$.*

4. NOTIONS OF NEGATIVE DEPENDENCE

We now start defining several notions of negative dependence. We mainly follow [1] and [2] in this chapter. The reason there are several notions of negative dependence is that we do not have the corresponding FKG theorem in the negative dependence case. First, we define some operations that gives us new measures which are natural in the way that we want the negative dependence notions we define to satisfy them.

4.1. Operations on Measures

1) Projection. Given a subset E' of our original set E, let μ' be the projection of μ onto $\{0,1\}^{E'}$. This corresponds to forgetting the variables in E/E'. It is reasonable to expect the new measure to satisfy negative dependence property the original measure satisfies as well. An example would be to define a measure on the subgraph of a graph. In this case E/E' would be the edges that we deleted to obtain the subgraph.

2) Conditioning. Given a subset A of our original subset E, we may fix the states of the elements in A. That is for a given $\psi \in \{0,1\}^A$, we consider the conditional distribution $\mu' = (\mu | X_e = \psi(e) \text{ for all } e \in A)$. It is reasonable for the conditioned measure to satisfy the negative dependence properties the original measure satisfies as well. An example would be to fix some edge set and assign them values before hand, and condition on the states given to get the new measure. For instance, we may want to construct a measure on the set of subgraphs where the edge e is present always and f is not present. Here $A = \{e, f\}$ and $\psi \in \{0,1\}^A = (1,0)$. Several examples such as spanning trees, random cluster measure and the Ising model which we will consider, are all closed under conditioning.

3) Products. If the measure μ_1 and the measure μ_2 are negatively dependent, then it is reasonable to expect that the measure $\mu_1 \times \mu_2$ to be negatively dependent as well.

4) Relabeling. Let $\pi \in S_n$ be a permutation. Define $\mu'(x) = \mu(\pi(x))$. It is reasonable to expect that the new measure obtained by relabeling satisfies the negative dependence properties that the original measure satisfies. An example would be to just renaming the edges of a graph.

5) External Fields. Let $w : E \to \mathbb{R}^+$ be a weight function. Then we consider the weighted measure μ' obtained in the following way:

$$\mu'(X_e = \psi(e) : e \in E) \propto \prod_{e \in E} w(e)^{\psi(e)} \mu(X_e = \psi(e) : e \in E). \tag{4.1}$$

We make this a probability measure by multiplying each value with the normalizing constant, which is

$$\sum_{\psi \in \{0,1\}^E} \mu'(X_e = \psi(e)). \tag{4.2}$$

The name external field is taken from Ising model.

<u>4.1.0.1. Polynomial Conversion.</u> We will convert all of the above into a language of polynomials. Suppose we have a measure $\mu \in \{0,1\}^E$, where E is a finite set. We write the generating polynomial for the given measure μ by the values it takes on each subset of E. To a subset $S \subset E$, assign the monomial $\chi_S = \prod_{e \in S} z_e$. Then the generating polynomial that corresponds to μ denoted by $g_\mu(z_1, ...z_n)$ is

$$g_\mu(z_1, ..., z_n) = \sum_{S \subset E} \mu(S)\chi_S. \tag{4.3}$$

Observe that there is a one-to-one correspondence between the measures on $\{0,1\}^E$, and the polynomials $g(z_1, ..., z_{|E|} = \sum_{S \subset E} a_S \chi_S$ with the property $||g||_{L^1} = 1$. We have $a_S = \mu(S)$.

Let us look at the operations under this perspective.

1) Projection. Let μ be a measure on the set E, and $E' \subset E$. Suppose the generating polynomial of μ is g_μ, then the projection of μ onto the subset E' gives a new measure μ' whose generating polynomial is related to the generating polynomial of μ as below:

$$g_{\mu'}(z_e; e \in E') = g_\mu(z_e; e \in E)|_{z_e=1 \text{ for } e\in E/E'}. \tag{4.4}$$

2) Conditioning. Let μ be the measure on the set E with the generating polynomial g_μ, and $A \subset E$. Given $\psi \in \{0,1\}^A$, if we condition on μ with ψ fixed, we get a new measure $\mu' = \mu(X_e|X_e = \psi(e) \text{ for } e \in A)$ with the generating polynomial

$$g_{\mu'}(z_1, ..., z_n) = \frac{g_\mu(z_1, ..., z_n)|_{z_e=0 \text{ for } e\in A}}{g_\mu(z_1, ..., z_n)|_{z_e=0 \text{ for } e\in A; z_e=1 \text{ for } e\in E/A}}. \tag{4.5}$$

3) Products. Suppose the measure μ_1 on $\{0,1\}^{E_1}$ has generating polynomial $g_{\mu_1}(z_1, ..., z_{n_1})$ and the measure μ_2 on $\{0,1\}^{E_2}$ has generating polynomial $g_{\mu_2}(z_1, ..., z_{n_2})$, where $|E_1| = n_1$ and $|E_2| = n_2$. Then the product $\mu_1 \times \mu_2$ on $\{0,1\}^{E_1 \cup E_2}$ is the measure with generating polynomial

$$g_{\mu_1\times\mu_2}(z_1, ..., z_{n_1+n_2}) = g_{\mu_1}(z_1, ..., z_{n_1}) g_{\mu_2}(z_{n_1+1}, ..., z_{n_1+n_2}). \tag{4.6}$$

4) Relabeling. Suppose the measure μ has the generating polynomial $g_\mu(z_1, ..., z_n)$. Then for a permutation $\pi \in S_n$, the relabeled measure μ' is the measure with generating polynomial

$$g_{\mu'}(z_1, ..., z_n) = g_\mu(z_{\pi(1)}, ..., z_{\pi(n)}). \tag{4.7}$$

5) External Fields. Let $w : E \to \mathbb{R}^+$ be a weight function. If μ on $\{0,1\}^E$ is a measure with generating polynomial $g_\mu(z_1, ..., z_n)$, then the measure μ' with the external field $(w(e))_{e\in E}$ is the measure with the generating polynomial

$$g_{\mu'}(z_1, ..., z_n) = \frac{g_\mu(w(1)z_1, ..., w(n)z_n)}{g_\mu(w(1), ..., w(n))}, \tag{4.8}$$

which is well-defined if $g_\mu(w(1), ..., w(n)) \neq 0$.

4.2. Notions of Negative Dependence

4.2.1. NLC

We define negative lattice condition naturally as the reverse of positive lattice condition.

Definition 4.2.1. *A measure μ on Σ is said to satisfy the **Negative Lattice Condition** if we have*

$$\mu(x \vee y)\mu(x \wedge y) \leq \mu(x)\mu(y) \ \ \forall x, y \in \Sigma. \tag{4.9}$$

Considering this condition in terms of polynomials, we see that a measure μ satisfies the negative lattice condition if and only if the corresponding generating polynomial g_μ satisfies

$$\partial^S g_\mu(0, ..., 0)\partial^T g_\mu(0, ..., 0) \geq \partial^{S\cup T} g_\mu(0, ..., 0)\partial^{S\cap T} g_\mu(0, ..., 0), \tag{4.10}$$

where $S, T \subset E$ and ∂^S means $\dfrac{\partial^{|S|}}{\partial_{|e\in S} z_e}$.

Remark 4.2.1. *We see that NLC is closed under the operations defined above except projection.*

We will give an example to show that negative lattice condition does not imply negative association in the following sections.

4.2.1.1. h-NLC.

Definition 4.2.2. *A measure μ is said to satisfy the hereditary negative lattice condition, h-NLC, if every projection of μ satisfies NLC.*

This is a stronger property than NLC and is needed since we saw that negative lattice condition is not closed under projection.

4.2.1.2. h-NLC+.

Definition 4.2.3. *A measure μ is said to satisfy the strong hereditary negative lattice condition if all measures obtained by imposing an external field on μ satisfies h-NLC.*

4.2.2. p-NC

The notion of pairwise negative correlation is the weakest among all the notions we will define. We will prove that the other notions all imply p-NC.

Definition 4.2.4. A measure $\mu \in \{0,1\}^E$ is said to be **pairwise negatively correlated**, p-NC, if we have

$$\mathbb{E}_\mu[X_e]\mathbb{E}_\mu[X_f] \geq \mathbb{E}_\mu[X_e X_f], \tag{4.11}$$

for all $e \neq f \in E$.

In terms of the polynomials this corresponds to the inequality

$$\partial_e g_\mu(\mathbf{1})\partial_f g_\mu(\mathbf{1}) \geq \partial_e\partial_f g_\mu(\mathbf{1}) \text{ for all } e \neq f, \tag{4.12}$$

where g_μ is the generating polynomial for the measure μ, and $\mathbf{1}$ denotes the all ones vector $(1, ..., 1)$ in the corresponding space.

4.2.3. NA

The notion of negative association is the most basic one, and in a way is the corresponding notion of positive association. We define it in the same way, with just the inequality reversed. However, there is a distinction since a random variable X is always positively associated with itself, i.e. $\mathbb{E}[X^2] - \mathbb{E}[X]^2 = Var(X) \geq 0$. Therefore, we need to make an adjustment, which we make with adjusting the coordinates the functions depend on.

Definition 4.2.5. A measure μ on $\{0,1\}^E$ is called **negatively associated**, NA, if

$$\mathbb{E}_\mu[f]\mathbb{E}_\mu[g] \geq \mathbb{E}_\mu[fg], \tag{4.13}$$

for any increasing functions f, g on $\{0,1\}^E$ that depend on disjoint sets of coordinates.

We state a proposition for negative association that yields calculation free proofs in some cases.

Proposition 4.2.1. *(i) If X, Y satisfy NA and are mutually independent, then the vector (X, Y) satisfies NA.*

(ii) Suppose $X = (X_1, \ldots, X_n)$ satisfy NA. Let $(A_i)_{i\in[k]}$ be disjoint index sets, and, for $i \in [k]$, let $\psi_i : \mathbb{R}^{|A_i|} \to \mathbb{R}$ be either all non-increasing or all non-decreasing functions. Set $Y_i = \psi_i(X_i, i \in A_i)$. Then $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ satisfies NA.

4.2.3.1. CNA.

Definition 4.2.6. A measure μ is said to be **conditionally negatively associated**, CNA, if each measure μ' obtained from μ by conditioning is negatively associated.

4.2.3.2. CNA+.

Definition 4.2.7. A measure μ is said to be **strongly conditionally negatively associated**, CNA+ if each measure μ' obtained from μ by imposing an external field and projection is CNA.

Remark 4.2.2. *It can be seen from the definitions that the measures satisfying h-NLC,h-NLC+,NA,CNA,CNA+ are all p-NC.*

4.2.4. RAYLEIGH MEASURES

Let g be a polynomial with $g(\mathbf{1}) = 1$, where $\mathbf{1} = (1, ..., 1)$ is the all ones vector. We say that the polynomial g is **Rayleigh** if

$$\frac{\partial g}{\partial z_e}(x)\frac{\partial g}{\partial z_f}(x) \geq \frac{\partial^2 g}{\partial z_e \partial z_f}(x)f(x), \tag{4.14}$$

for all $x \in \mathbb{R}^n_+$, and for all $e, f \in [n]$. We say that a multi-affine polynomial with non-negative coefficients in $\mathbb{R}[z_1, ..., z_n]$ is Rayleigh if it satisfies the above condition.

Remark 4.2.3. *We call a polynomial* ***c-Rayleigh*** *if the same inequality holds with a constant c, i.e.*

$$c\frac{\partial g}{\partial z_e}(x)\frac{\partial g}{\partial z_f}(x) \geq \frac{\partial^2 g}{\partial z_e \partial z_f}(x)f(x). \tag{4.15}$$

Definition 4.2.8. *A measure μ is called* ***Rayleigh (c-Rayleigh)*** *if the corresponding generating polynomial g_μ is Rayleigh* ***(c-Rayleigh*** *respectively).*

Remark 4.2.4. *The name Rayleigh comes from the similarity to the Rayleigh monotonicity property of the effective conductance in electrical networks.*

We state some basic results concerning Rayleigh polynomials.

Proposition 4.2.2. If $g(z_1, ..., z_n)$ is Rayleigh then the following polynomials are also Rayleigh:

(1) $\partial^S g(z_1, ..., z_n)$ for any $S \subset [n]$,

(2) $g(z_1 + a_1, ..., z_n + a_n)$ for any $a_i \geq 0$,

(3) $g(z_1, ..., z_n)|_{z_i = a_i}$ for any $i \in [n]$ and $a_i \geq 0$,

(4) $g(a_1 z_1, ..., a_n z_n)$ for any $a_i \geq 0$,

(5) The polynomial $\hat{g}$, called the inversion of g, $g(\frac{1}{z_1}, ..., \frac{1}{z_n}) \prod_{i=1}^{n} z_i$,

(6) The polynomial in nk variables $g\left(\frac{1}{k}\sum_{i=1}^{k} z_{1i}, ..., \frac{1}{k}\sum_{i=1}^{k} z_{ni}\right)$.

4.2.4.1. PHR.

Definition 4.2.9. A measure whose generating polynomial is homogeneous and Rayleigh is called a **homogeneous Rayleigh measure**. We denote by PHR, the set of all measures that are projections of homogeneous Rayleigh measures.

4.2.4.2. STRONGLY RAYLEIGH.

Definition 4.2.10. A polynomial $g \in \mathbb{C}[z_1, ..., z_n]$ is called **stable** if $g(z_1, ..., z_n) \neq 0$ whenever $Im(z_j) > 0$ for $1 \leq j \leq n$. If the polynomial has real coefficients, we call it **real stable**.

Remark 4.2.5. *Multivariable stable polynomials are related to many other important notions. We do not dwell on them too much.*

Now, we define Strongly Rayleigh measures which will have another equivalent definition in the way Rayleigh measures was defined.

Definition 4.2.11. A measure μ is called **strongly Rayleigh** if the corresponding generating polynomial g_μ is real stable.

Branden proved in his paper Polynomials with the half-plane property and matroid theory, that the real stability is equivalent to the following property.

Proposition 4.2.3. *A polynomial $g \in \mathbb{C}[z_1, ..., z_n]$ is real stable if and only if it has non-negative coefficients and it satisfies*

$$\frac{\partial g}{\partial z_i}(x)\frac{\partial g}{\partial z_j}(x) \geq \frac{\partial^2 g}{\partial z_i \partial z_j}(x)g(x), \tag{4.16}$$

for all $x \in \mathbb{R}^n$ and $1 \leq i \neq j \leq n$.

This is the same as the definition of Rayleigh measure with the only difference being is that we want the inequality to be satisfied not only on the positive reals, but on all of them. Hence, it is easily seen that being strongly Rayleigh implies being Rayleigh.

4.2.5. LORENTZIANITY

Definition 4.2.12. We define H_n^d to be the set of degree d homogeneous polynomials of in $\mathbb{R}[z_1, ..., z_n]$, and P_n^d to be the set of polynomials with positive coefficients in H_n^d.

Recall that the hessian of a function is defined as the matrix

$$H_f = (\partial_i \partial_j f)_{i,j=1}^n, \tag{4.17}$$

where ∂_i is the partial derivative with respect to the i^{th} variable z_i.

Definition 4.2.13 (Lorentzian Polynomials)**.** We define $L_n^2 \subset H_n^2$ to be the open subset of quadratic forms with Lorentzian signature $(+, -, \dots, -)$, i.e. the Hessian has only one positive eigenvalue. For $d > 2$, we define the open subset $L_n^d \subset H_n^d$ by setting

$$L_n^d = \{f \in H_n^d : \partial_i f \in L_n^{d-1} \text{ for all } i\}.$$

The polynomials in L_n^d are called **strictly Lorentzian**. The limits of the strictly Lorentzian polynomials are called **Lorentzian**, since this is a closure, we can denote the Lorentzian polynomials by $\overline{L_n^d}$.

Remark 4.2.6. *This polynomials not only show up in the context of negative dependence but actually they are the volume polynomials of convex bodies and projective varieties. They connect the discrete and continuous notions of convexity.*

4.2.6. LOG CONCAVITY

Now, we define log concavity for sequences and we will see how it relates to negative dependence notions in the following sections. Suppose we have a finite sequence of non-negative real numbers, $\{a_i\}_{i=1}^n$.

Definition 4.2.14. We say that a sequence of non-negative real numbers $\{a_i\}_{i=1}^n$ is **log concave** if

$$a_i^2 \geq a_{i-1}a_{i+1}, \tag{4.18}$$

for all $1 \leq i \leq n$.

Definition 4.2.15. We say that the sequence $\{a_i\}_{i=1}^n$ has no internal zeroes if the indices of the non-zero terms forms an interval.

Example 4.2.1. *A natural example of a log concave sequence is the binomial numbers. Fixing n, we have that the sequence $\{\binom{n}{k}\}_{k=0}^n$ is log concave.*

Definition 4.2.16. We say that a sequence $\{a_i\}_{i=0}^n$ is **strongly log concave** if $\{i!a_i\}_{i=0}^n$ is log concave and it has no internal zeroes. We say that the sequence is **ultra log concave** if $\{\frac{a_i}{\binom{n}{i}}\}_{i=0}^n$ is log concave and it has no internal zeroes.

Remark 4.2.7. *We have that $ULC \Rightarrow SLC \Rightarrow LC$.*

4.2.7. STOCHASTIC COVERING PROPERTY

We first define the notion of stochastic domination and covering.

Definition 4.2.17. Let μ, ν be measures on $\{0,1\}^E$. We say that the measure μ **stochastically dominates** the measure ν if we have

$$\mu(A) \geq \nu(A), \tag{4.19}$$

for every upwardly closed event A. We denote stochastic domination by $\mu \succeq \nu$.

Proposition 4.2.4. *Let μ, ν be measures on $\{0,1\}^n$. $\mu \succeq \nu$ if and only if there is a coupling, a measure ρ on $\{0,1\}^n \times \{0,1\}^n$ with marginals μ and ν which is supported on the set $\{(x,y) : x \geq y\}$.*

Definition 4.2.18. We say that x covers y in a partially ordered set if $x > y$ and there is no z such that $x > z > y$. We denote this relation by $x \dot{>} y$.

Definition 4.2.19. Let μ, ν be measures on $\{0,1\}^n$. The measure μ **stochastically covers** the measure μ, if there is a coupling, that is a measure ρ on $\{0,1\}^n \times \{0,1\}^n$ with marginals μ and ν, so that ρ is supported on the set $\{x, y\}$ where either $x = y$, or x covers y. We denote this relation by $\mu \triangleright \nu$.

Remark 4.2.8. *Stochasting covering is the same as stochastic domination except we want the coupling to be on the elements that differ by one in hamming distance.*

Definition 4.2.20. A measure μ on $\{0,1\}^E$ is said to have the **stochastic covering property** if for every $S \subset [n]$ and $x, y \in \{0,1\}^E$ with $x \dot{>} y$, the conditional measure $(\mu | X_j = x_j$ for $j \in S)$ stochastically covers the conditional measure $(\mu | X_j = y_j$ for $j \in S)$.

Remark 4.2.9. *Suppose $x \geq y$, and we compare the measures $\mu_x = (\mu | X_j = x_j$ for $j \in S)$, and $\mu_y = (\mu | X_j = y_j$ for $j \in S)$ on $\{0,1\}^{S^c}$. If the measure μ and its conditionalizations are negatively associated, then it follows that $\mu_x \succeq \mu_y$.*

4.3. RELATIONS AMONG THE NOTIONS OF NEGATIVE DEPENDENCE

4.3.1. Implications

Now, we prove some relations among the notions we defined above. There are many open conjectures that needs proving in this part and we state them in order as well.

First of all, let us observe the basic relations.

Proposition 4.3.1. *We have the following implications: CNA+* $\implies$ *CNA* $\implies$ *NA.*

Proof. These implications follow from the definitions. ■

Proposition 4.3.2. *We have the following implication: h-NLC* $\implies$ *NLC.*

Proof. Since h-NLC is just NLC with an extra hereditary condition, the implication follows. ■

Proposition 4.3.3. *We have the following implication: NA* $\implies$ *p-NC.*

Proof. By definition, negative association gives us for any increasing functions that depend on a disjoint set of coordinates f, g

$$\mathbb{E}[f(X)g(X)] \leq \mathbb{E}[f(X)]\mathbb{E}[g(X)]. \tag{4.20}$$

Note that $f(X) = X_e$ and $g(X) = X_f$ are increasing functions and depend on disjoint set of variables, therefore we have

$$\mathbb{E}[X_e X_f] \leq \mathbb{E}[X_e]\mathbb{E}[X_f], \tag{4.21}$$

which is pairwise negative dependence. ■

Proposition 4.3.4. *We have the following implication: h-NLC* $\implies$ *p-NC.*

Proposition 4.3.5. *We see that if a measure is strongly Rayleigh, then it is also Rayleigh.*

Proof. This follows from the definitions of the respective notions. ■

The above implications are trivial via the definitions. We state some non-trivial implications:

Proposition 4.3.6. *h-NLC* $\Leftrightarrow$ *Rayleigh.*

Proof. Wagner proved the left implication in his paper [11], Theorem 4.4. For the right implication, we follow the proof in [1]. Let $i \neq j \in [n]$ and set

$$g(z_i, z_j) = f(z_1, \ldots, z_n)|_{z_k=1, k\in[n]\setminus\{i,j\}}. \tag{4.22}$$

The quadratic polynomial we get is real stable by NLC properties. Now, using external fields and properties of the Rayleigh polynomials together with the fact that

$$\partial_i(\mathbf{1})\partial_j(\mathbf{1}) \geq \partial_i\partial_j(\mathbf{1}),$$

we see that the inequality also holds for $f(x_1z_1, \ldots, x_nz_n)$ whenever $x_i \geq 0$ for $i \in [n]$, which means that f is Rayleigh. ∎

Remark 4.3.1. *We also have a similar condition for h-NlC and Rayleigh polynomials. A polynomial f is h-NLC if and only if it satisfies the Rayleigh inequality for all $z = (z_1, \ldots, z_n)$ with $z_i \in \{0, 1, \infty\}$ for all $i \in [n]$.*

We give the celebrated result of Feder and Mihail:

Theorem 4.3.1 (Feder-Mihail)**.** *Let $\mathcal{A}$ be a set of measures closed under conditioning and projections. Suppose all of the measures in A is p-NC, i.e. satisfy pairwise negative correlation for X_e, X_f. If for each $\mu \in A$ and an increasing event Q, there is an edge e so that we have the following correlation*

$$\mu(X_e\mathbb{I}_Q) \geq \mu(X_e)\mu(Q),$$

then each measure in $\mathcal{A}$ is negatively associated.

Proof. We follow the proof given in [2] Theorem 1.3, which is the same as the original proof of Feder and Mihail. We will use induction on the rank of the lattice measures are supported on. Suppose the negative association property holds for all measures that are supported on the lattice of rank n and fewer. Note that when $n = 1$, this is vacuously true. We show that the two events X_e and an arbitrary increasing event Q not depending on X_e satisfies negative dependence. The other cases are similar.

If $\mathbb{P}(X_e = X_f = 1) = 0$, with $f \neq 0$ we do not have anything to do in the induction step. Assume this is not the case. By the assumption given in the theorem, for $(\mu|e)$ there is some $f \neq e$ for which we have

$$\mu(Q|X_e = X_f = 1) \geq \mu(Q|X_e = 1). \tag{4.23}$$

Now we write the following

$$\begin{aligned}
\mu(Q|X_e = 1) &= \mu(Q|X_e = X_f = 1)\mu(X_f = 1|X_e = 1) \\
&\quad + \mu(Q|X_e = 1, X_f = 0)\mu(X_f = 0|X_e = 1) \\
\mu(Q|X_e = 0) &= \mu(Q|X_e = 0, X_f = 1)\mu(X_f = 1|X_e = 0) \\
&\quad + \mu(Q|X_e = X_f = 0)\mu(X_f = 0|X_e = 0).
\end{aligned}$$

Now, we can compare the terms on the right-hand sides of the two equations:

- $\mu(X_f = 1|X_e = 1) \leq \mu(X_f = 1|X_e = 0)$ by the assumption that the measures in $\mathcal{A}$ are p-NC,
- $\mu(Q|X_e = X_f = 1) \leq \mu(Q|X_e = 0, X_f = 1)$ and $\mu(Q|X_e = 1, X_f = 0) \leq \mu(Q|X_e = X_f = 0)$due to the fact that $(\mu|X_f = 1)$ and $(\mu|X_f = 0)$ respectively is assumed to be negatively associated by induction hypothesis,
- $\mu(Q|X_e = X_f = 1) \geq \mu(Q|X_e = 1, X_f = 0)$ by the choice of f.

With this four inequalities, we observe that $\mu(Q|X_e = 1) \leq \mu(Q|X_e = 0)$ which proves the NA in this special case. The general increasing events follows in a similar manner.

■

Remark 4.3.2. *Feder and Mihail proved this in the case of a class of matroids which is called Balanced Matroids in their paper [12].*

Proposition 4.3.7. *A Strongly Rayleigh measure (and also a PHR measure) is CNA+. This implications are strict.*

Proof. Let $\mathcal{A}$ be the class of probability measures satisfying the following conditions:

- each $\mu \in \mathcal{A}$ is a measure on $\{0,1\}^E$ for some finite set $E \subset \{1,2,\ldots\}$, and
- μ has a stable homogeneous generating polynomial.

Stable homogeneous polynomials are p-NC and are closed under conditioning. Hence they satisfy the hypothesis of Feder and Mihail's theorem 4.3.1. Thus, all the measures in $\mathcal{A}$ are in NA. Now, we let $\overline{\mathcal{A}}$ be the class of measures satisfying the following conditions:

- each $\mu \in \mathcal{A}$ is a measure on $\{0,1\}^E$ for some finite set $E \subset \{1,2,\ldots\}$, and
- μ has a stable generating polynomial.

We have that $\overline{\mu} \in \overline{\mathcal{A}}$ is a projection of a measure $\mu \in \mathcal{A}$, and since NA is preserved under projection is itself in NA. Since the Strongly Rayleigh measures are preserved under projections and external fields, this class contains Strongly Rayleigh measures and we are done. ∎

Remark 4.3.3. *Since homogeneous Rayleigh polynomials need not be stable, we see that the strongly Rayleigh implying PHR is also strict. We will see that PHR$\Rightarrow$CNA+ is also strict. Hence, strongly Rayleigh $\Rightarrow$ CNA+ is also strict.*

Proposition 4.3.8. *A strongly Rayleigh measure is Lorentzian.*

Proof. We prove the proposition by induction on d. Strongly Rayleigh polynomials are stable polynomials S_n^d and note that the derivative as an operator is an open map sending $S_n^d \to S_n^{d-1}$, i.e. sending the interior of S_n^d into the interior of S_n^{d-1}. Hence, we prove that S_n^2 is in L_n^2. For quadratic forms, we prove that the Hessian has Lorentzian signature $(+,-,\ldots,-)$ if and only if for any non-zero $u \in \mathbb{R}^n_{\geq 0}$, the univariate polynomial $f(xu-v)$ in x has two distinct real zeroes not parallel to u for any $v \in \mathbb{R}^n$. We note that the polynomial $\frac{1}{2}f(xu-v)$ has discriminant

$$(u^t H_f v)^2 - (u^t H_f u)(v^t H_f v), \tag{4.24}$$

where H_f is the Hessian of f. Supposing $f \in L_n^2$, we know that the entries of H_f are all positive and $u^t H_f u > 0$ for all non-zero u. Now, by Cauchy's interlacing theorem, we have that for any $v \in \mathbb{R}^n$, the restriction of the Hessian to the plane spanned by u, v has signature $(+,-)$, hence

$$\det \begin{pmatrix} u^t H_f u & u^t H_f v \\ u^t H_f v & v^t H_f v \end{pmatrix} = -[(u^t H_f v)^2 - (u^t H_f u)(v^t H_f v)] < 0, \tag{4.25}$$

hence we have the desired result. ∎

Proposition 4.3.9. *Lorentzian polynomials of degree d are $2\left(1-\frac{1}{d}\right)$-Rayleigh.*

Proof. Euler's formula for homogeneous polynomials gives us the following:

$$u^t H_f(u) u = d(d-1)f(u), \quad \text{and} \quad u^t H_f(u) e_i = (d-1)\partial_i f(u), \tag{4.26}$$

for any fixed $u \in \mathbb{R}^n_{\geq 0}$.

We consider the restriction of H_f to the plane spanned by u and $v_t = e_i + te_j$ for some real parameter t.By Cauchy's interlacing theorem, this new restricted Hessian also has one positive eigenvalue, i.e. we have

$$(u^t H_f v_t)^2 - (u^t H_f u)(v_t^t H_f v_t) \geq 0, \tag{4.27}$$

for all $t \in \mathbb{R}$. Now applying Euler's formula for homogeneous polynomials we get:

$$(d-1)^2(\partial_i f(u) + t\partial_j f(u))^2 - d(d-1)f(\partial_i^2 f + 2t\partial_i\partial_j f + t^2\partial_j^2 f) \geq 0. \tag{4.28}$$

Hence, for all $t \in \mathbb{R}$, we have

$$(d-1)^2(\partial_i f + t\partial_j f)^2 - 2td(d-1) f\partial_i\partial_j f \geq 0, \tag{4.29}$$

which means that the discriminant should be non-positive, which turns out as

$$f\partial_i\partial_j f - 2\left(1 - \frac{1}{d}\right)\partial_i\partial_j \leq 0. \tag{4.30}$$

This inequality shows that f is $2\left(1-\frac{1}{d}\right)$-Rayleigh. ∎

We have the following figure for the implications:

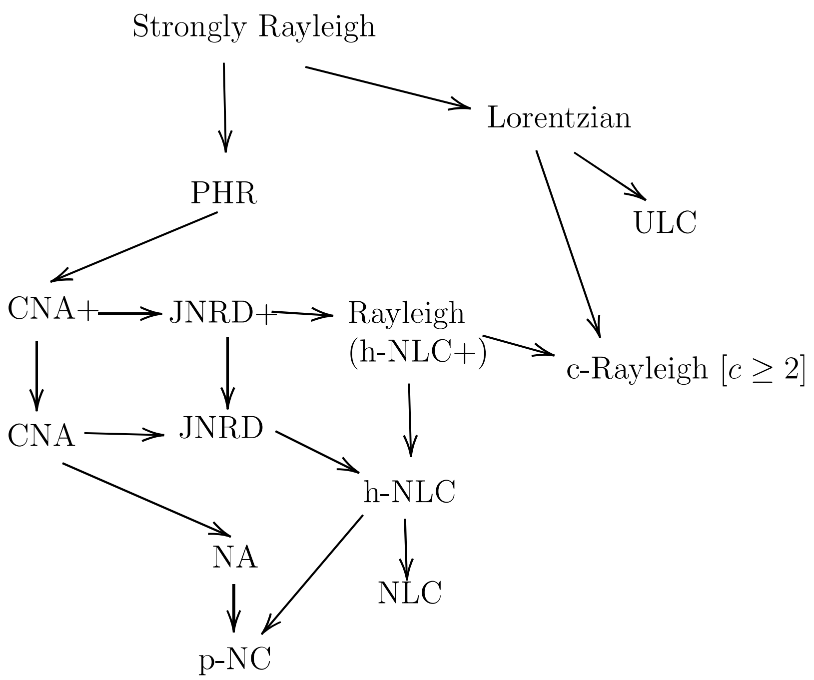


Figure 4.1. Relations among notions of negative dependence.

4.3.2. Counter Examples

We start with stating a simple counter example to the inverse FKG theorem, i.e. NLC $\not\Rightarrow$ NA.

Example 4.3.1 ([1]). *Consider the measure on $\{0,1\}^3$ whose generating polynomial is given as below:*

$$g_\mu(z) = \frac{1}{15}(3 + 2z_1 + 2z_2 + 2z_3 + 2z_1z_2 + 2z_1z_3 + 2z_2z_3), \tag{4.31}$$

i.e. the measure is concentrated on the 1-element and 2-element subsets. Recall the NLC inequality $\mu(x \vee y)\mu(x \wedge y) \leq \mu(x)\mu(y)$. Let $x = (1,0,0)$ and $y = (0,1,0)$, then $x \vee y = (1,1,0)$ and $x \wedge y = (0,0,0)$. Thus we have $\frac{3}{15}\frac{2}{15} \geq \frac{2}{15}\frac{2}{15}$ which shows that the measure μ does not satisfy the NLC inequality. However, the measure is negatively associated.

Remark 4.3.4. *Another beautiful example is due to T. M. Liggett in his paper [13] Theorem 3.5.*

Example 4.3.2 ([2]). *Consider the measure on $\{0,1\}^3$ whose generating polynomial is given as below:*

$$g_\mu(z) = \frac{1}{4+21\epsilon}(z_1 + z_2 + z_3 + 10\epsilon z_1z_2 + z_1z_3 + 10\epsilon z_2z_3 + \epsilon z_1z_2z_3). \tag{4.32}$$

When $0 \leq \epsilon \leq 8$, this measure satisfies the CNA conditions and all its implications. However, when $\epsilon > 0$, with the external field $(\lambda, 1, 1)$ for any $\lambda < \frac{\epsilon}{1-\epsilon}$ gives us a measure where z_2 and z_3 are positively correlated. Hence the external fields violate the negative association. This shows that the implications from CNA+, JNRD+ and h-NLC+ to CNA, JNRD, h-NLC are strict respectively.

5. MODELS

5.1. URN MODELS

One model where negative dependence is prominent is Urn models. Suppose we have n balls and m bins to distribute these balls into. Let O_j denote the number of balls inside bin j for $1 \leq j \leq m$.

Intuitively, we see that if some balls are in some bins, then since the number of balls decreases, the probability of a bin being filled decreases. Somehow, there should be a negative dependence amongst the distribution laws of O_j's. We will call these numbers **occupancy numbers**.

Suppose the balls and the bins are identical and the probability that the ball will be put into bin j is p_j, with $\sum_{j=1}^{m} p_j = 1$ then the distribution will be the multinomial distribution. That is, if we have the vector $(x_1, ..., x_m)$ where x_i is the number of balls in the bin i, we have

$$\mathbb{P}[\mathbf{O} = \mathbf{x}] = \frac{m!}{x_1! \dots x_m!} p_1^{x_1} \dots p_m^{x_m}, \tag{5.1}$$

if $\sum_{i=1}^{m} x_i = n$. Otherwise, the probability is zero.

We will prove a proposition first.

Proposition 5.1.1 (0-1 Law). *If the 0-1 random variables, $(X_i)_{1 \leq i \leq n}$ satisfy*

$$\sum_{i=1}^{n} X_i = 1, \tag{5.2}$$

then they satisfy the negative association property.

Proof. Suppose $X_i = 1$, and the others are 0. Suppose I and J are the disjoint subsets that the functions in NA depend on. That is, $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are actually functions of the form $f : \mathbb{R}^|I| \to \mathbb{R}$ and $g : \mathbb{R}^|J| \to \mathbb{R}$. We can write this as $f(a_i.i \in I)$ and $g(a_j, j \in J)$. Suppose also that f and g are non-decreasing. We can arrange the coordinates so that I is the first I of coordinates and J is last $|J|$ many coordinates. Now, we have

$$
\begin{aligned}
f(0,\ldots,0) \leq f(0,\ldots,1) \leq \ldots \leq f(1,0,\ldots,0),\\
g(0,\ldots,0) \leq g(1,\ldots,0) \leq \ldots \leq g(0,0,\ldots,1).
\end{aligned}
$$

This inequalities follows from the choice of the set of coordinates that the functions depend on.

Define a new random variable X taking values in $[n]$ as follows: $X_i = 1 \implies X = i$.

We also define two new functions

$$
f'(i) = \begin{cases} f(0,\ldots,1,\ldots,0), & \text{if } i \in I \\ f(0,\ldots,0,\ldots,0), & \text{if } i \notin I \end{cases}
$$

$$
g'(i) = \begin{cases} g(0,\ldots,1,\ldots,0), & \text{if } i \in J \\ g(0,\ldots,0,\ldots,0), & \text{if } i \notin J \end{cases},
$$

where the 1 is in position i. We observe that $\mathbb{E}[f'(X)] = \mathbb{E}[f(X_i, i \in I)]$, $\mathbb{E}[g'(X)] = \mathbb{E}[g(X_i, i \in J)]$, and $\mathbb{E}[f'(X)g'(X)] = \mathbb{E}[f(X_i, i \in I)g(X_i, i \in J)]$. We have by FKG inequality $\mathbb{E}[f'(X)g'(X)] \leq \mathbb{E}[f'(X)]\mathbb{E}[g'(X)]$, since f' and g' are non-increasing functions. This proves the desired inequality. ■

Remark 5.1.1. *We defined the NA for probability measures. Here the probability measure is the measure that comes naturally with the given random variables* X_i.

Now, let us work with Urn models. Instead of working with identical balls, let us name the balls as well. We define the following random variables:

$$O_{i,k} = \begin{cases} 1, & \text{if ball } k \text{ falls into urn } i \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

Proposition 5.1.2. *For any fixed* k*, we have that* $(O_{i,k}, i \in [m])$*, satisfies NA.*

Proof. Note that the ball k can only fall into 1 urn. Hence, $\sum_{i=1}^{m} O_{i,k} = 1$. The proposition follows from the 0-1 law 5.1.1. ■

Remark 5.1.2. *Since the balls are independent, we can take in the above proposition 5.1.2* $(O_{i,k}, i \in [m], k \in [n])$.

Proposition 5.1.3. *The vector* $\boldsymbol{O} = (O_1, \ldots, O_m)$ *satisfies NA.*

Proof. Observe that $O_i = \sum_{k=1}^{n} O_{i,k}$, and the sum is a non-decreasing function for all $i \in [m]$. Now, using the above proposition 5.1.2 with the remark 5.1.2 and the second part of proposition 4.2.1, we get the desired result. ■

5.2. UST, USF AND UCS MEASURES

We define two measures on the set of subgraphs of a graph. The first one is the measure called the uniform spanning tree measure,UST, which is supported on the spanning trees of a graph, and the second one is the measure called the uniform spanning forest measure, USF, which is supported on the spanning forests of a graph.

5.2.1. UST

We prove in this section that the uniform spanning tree measure is negative edge dependent. Let $G = (V, E)$ be a finite, connected graph. The **uniform spanning tree measure** denoted by μ^{st} is the measure on the subgraphs of G concentrated on the spanning trees. Let $Z^{G,\mathrm{st}}$ denote the number of spanning trees of the graph G. Note that if there are no confusions, we suppress the graph G. Then we have for a subgraph $\sigma \in \{0,1\}^E$,

$$\mu^{\mathrm{st}}(\sigma) = \begin{cases} \dfrac{1}{Z^{G,\mathrm{st}}}, & \text{if } \sigma \text{ is a spanning tree} \\ 0, & \text{otherwise.} \end{cases} \tag{5.4}$$

Let us look at an example. First, let us restate p-NC in terms of probabilities. Observe the following $\mathbb{E}[X_e] = \mathbb{P}(e \in \mathcal{T})$, where $\mathcal{T}$ is chosen according to UST measure. Hence, p-NC becomes $\mathbb{P}(e, f \in \mathcal{T}) \leq \mathbb{P}(e \in \mathcal{T})\mathbb{P}(f \in \mathcal{T})$.

Example 5.2.1. *Suppose our graph $G = (V, E)$ is K_4. Let us look at its spanning trees which are given in 5.1. We count the number of spanning trees that contain e, that contain f, and that contain e, f. Let us denote the spanning trees that contain a set of edges $e_1, \ldots, e_k$ by $\mathcal{T}(e_1, \ldots, e_k)$. Now, we have $|\mathcal{T}^{K_4}(e)| = 8, \quad |\mathcal{T}^{K_4}(f)| = 8, \quad$ and $|\mathcal{T}^{K_4}(e, f) = 3|$. Therefore, we have*

$$\mathbb{P}(e, f \in \mathcal{T}) = \frac{3}{16} \leq \frac{4}{16} = \frac{8}{16}\frac{8}{16} = \mathbb{P}(e \in \mathcal{T})\mathbb{P}(f \in \mathcal{T}). \tag{5.5}$$

<u>5.2.1.1. Random Walks on the Graph.</u> We state a Markov chain that generates UST measure as its stationary distribution. We will generate a simple random walk on a finite, connected, d-regular graph G. The assumption of d-regularity can be bypassed with some work.

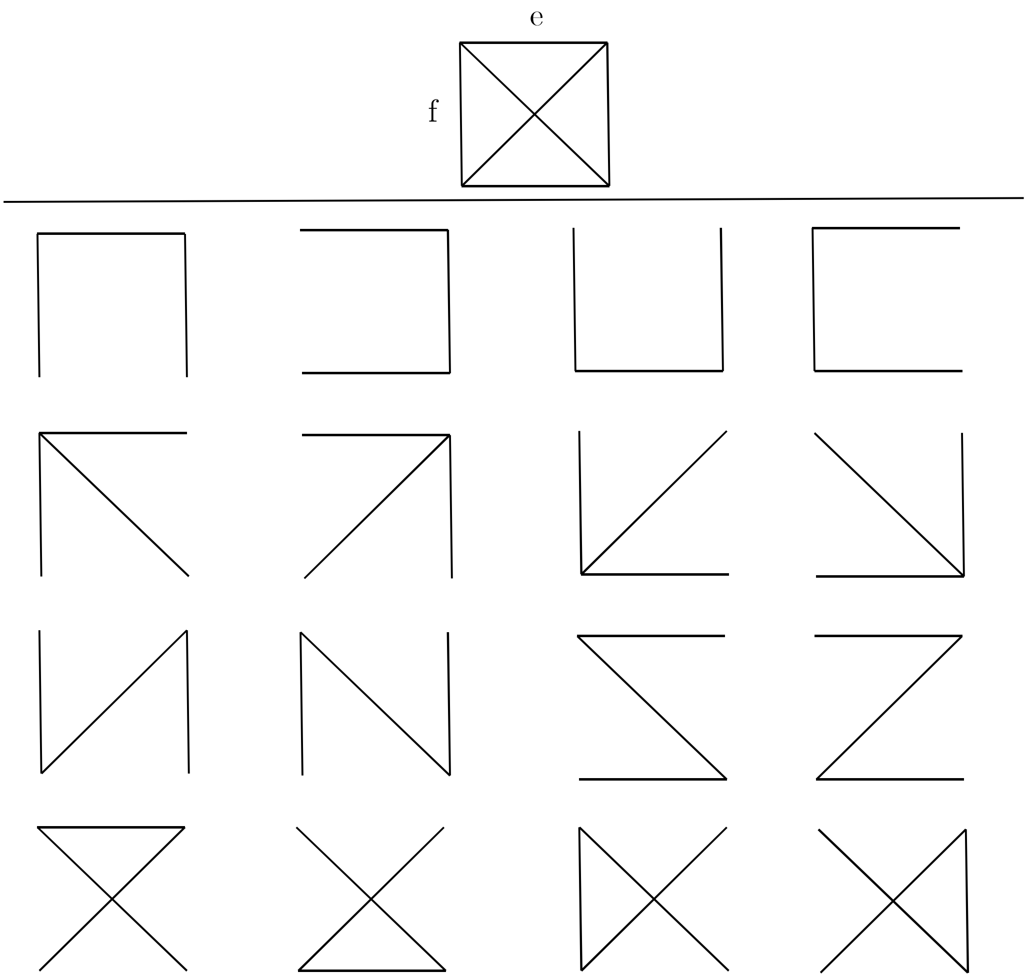


Figure 5.1. K_4 and its Spanning Trees.

The process starts by picking a vertex, v_0. Consider this as time $t = 0$. Then, at time $t = 1$, the walk will move along an edge to another vertex v_i which is adjacent to v_0, denoted $v_i \sim v_0$, with probability $\frac{1}{d}$.

Remark 5.2.1. *We can define a similar random walk on a finite, connected graph G, by assigning the probabilities as $\dfrac{1}{deg(v)}$ for each vertex.*

The given random walk is a Markov chain with state space $V(G)$, and probabilities

$$p(x, y) = \begin{cases} \dfrac{1}{d} & \text{if } x \sim y, \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

Observe that since the graph is connected, the Markov chain which we denote by SRW is irreducible. We can arrive from any vertex x to any other vertex y via moving on a path. The existence of such a path is guaranteed by the connectedness of the graph. Since the Markov chain is irreducible, we know that there is a stationary distribution.

We denote the simple random walk started at the vertex x by SRW_x^G. We begin by a construction that will give us a directed spanning tree with a root. However, we can simply forget the root and the orientations of the edges and the directed spanning tree becomes a spanning tree.

Suppose we perform a simple random walk on the graph G, but only take the edge that arrives at the vertex x if the edge is the first one that touches x, and we put the direction as $x \to y$, where y is the vertex from which we arrived at x. In this way, we obviously get a directed graph, and since we put the directions backwards, the root becomes the first vertex we started the process. Since we do not include the edges that overlap an already reached vertex, the process ends with a tree.

Remark 5.2.2. *It is important to note that the process will terminate with probability 1.*

Let us give an example. Suppose we perform the simple random walk on the following graph starting from A.

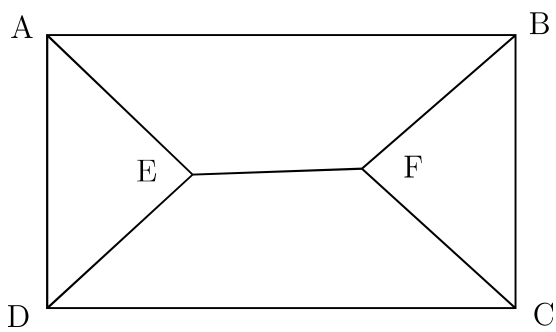


Figure 5.2. A graph.

Assume that the simple random walk SRW_A generate $ABFBCDAE$. Since we only take the edges that hit the vertices first, this gives the directed tree rooted at A:

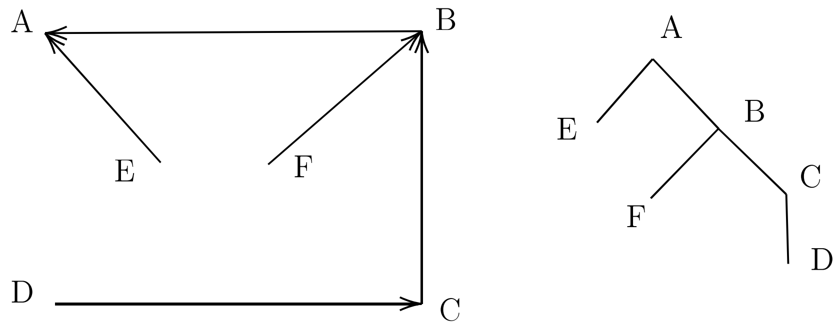


Figure 5.3. SRW_A on the graph in Figure 5.2.

Theorem 5.2.1. *The resulting measure is the uniform spanning tree measure.*

Let us return to the question of finding the probability that a given edge $e \in V$ is in a Spanning tree. We will formulate this problem in the setting of random walks. Let $e = \langle x, y \rangle$, that is e is the edge with vertices x and y. Consider the simple random walk that has been discussed above. It generates a uniform spanning tree. The question now becomes, what is the probability that the given edge e is in the uniform spanning tree generated by SRW_x^G. [Any vertex is fine, but for a given edge e we pick one of its vertices.] Considering the algorithm, we see that the only way, the edge will be in the resulting configuration is that the first time SRW_x^G hits the vertex y, is from the vertex x. Since otherwise, we would not add the edge $\langle x, y \rangle$. Hence, we have $\mathbb{P}(e \in T) = \mathbb{P}(SRW_x^G$ hits the vertex y first from the vertex x). We'll see that this is a much more tractable probability. We define the following function

$$h_{ab}(x) = \mathbb{P}(SRW_a^G \text{ from the state } x \text{ hits the vertex b before a}). \tag{5.7}$$

Remark 5.2.3.

- Note that $h_{ab}(a) = 0$, and $h_{ab}(b) = 1$.
- Observe that due to the nature of the simple random walk SRW, from a vertex, we can only pass to its neighbours, hence we can write a recursive formula of $h_{ab}(x)$, which involves the neighbours of x.

Suppose the time the random walk is at the vertex x is $\kappa - 1$, then at time κ, it will be on one of the neighbours of x. Hence we have

$$h_{ab}(x) = \sum_{y \sim x} \mathbb{P}(t_\kappa = y | t_{\kappa-1} = x) h_{ab}(y). \tag{5.8}$$

If we let the weights of the edges to be the transition probabilities, we see that

$$h_{ab}(x) = \sum_{y \sim x} h_{ab}(y) w(\langle x, y \rangle). \tag{5.9}$$

Hence, besides from the terminal vertices $T = \{a, b\}$, the function h_{ab} is a harmonic function and the terminal vertices satisfy $h_{ab}(a) = 0, h_{ab}(b) = 1$

Remark 5.2.4. *Observe that the same properties hold for the Voltage function* $V : V \to \mathbb{R}^+$ *as well.*

The theorems we stated in the Harmonic function section 2.2.4.1 together with **??** yield that there is only one function that satisfies the required properties. Thus, we get the following:

Theorem 5.2.2. *Let G be a finite, connected weighted graph. Let a, b be vertices of G. For any vertex x, the probability of SRW_x^G reaching a before b is equal to the voltage at x in G when the voltages at a and b are fixed to 0 and 1 respectively.*

Now, let us look at the currents in terms of the random walks and spanning trees.

Theorem 5.2.3 (Main Theorem of UST)**.** *The uniform spanning tree measure UST is negative edge dependent, or satisfies p-NC.*

Before proving the theorem, we state a physical principle:

Proposition 5.2.1 (Rayleigh Monotonicity Principle)**.** *The effective resistance of a circuit cannot increase when a new resistor is added.*

Proof of 5.2.3. Let us look at $\frac{\mathcal{T}(s,xy,t)}{|\mathcal{T}|}$, where $\mathcal{T}(s,xy,t)$ is the number of spanning trees whose unique path from s to t passes through the edge $e = xy$ in the direction $x \to y$, and $|\mathcal{T}|$ is the number of all spanning trees of G. From what we have done in this section we observe that this quantity is the same as the current flowing through the edge $e = xy$ in the direction $x \to y$ when a unit current flows from s to t.

Now, let $e = xy$, and μ^{st} be the UST measure on G. $\mu^{\text{st}}(X_e)$ equals the current flowing along e when a unit current flows through G from source x to sink y. By Ohm's Law, this equals the potential difference between x and y, which in turn equals the effective resistance $R(e)$ of the network between x and y. Let $f \neq e \in E$, and look at $G.f$, the graph obtained from G via the contraction of f. There is a one–one correspondence between spanning trees of $G.f$ and spanning trees of G containing f. Therefore, $\mu^{\text{st}}(X_e|X_f)$ equals the effective resistance $R^{G.f}(e)$ of the electrical network $G.f$ between x and y.

The Rayleigh principle states that the effective resistance of a network is a non-decreasing function of the individual edge-resistances. It follows that $R^{G.f}(e) \leq R^G(e)$, and thus $\mu^{\text{st}}(X_e|X_f) \leq \mu^{\text{st}}(X_e)$ which is $\mathbb{P}(e, f \in \mathcal{T}) \leq \mathbb{P}(e \in \mathcal{T})\mathbb{P}(f \in \mathcal{T})$. ∎

5.2.2. USF

We define the **uniform spanning forest measure** denoted as μ^{sf} to be the measure defined on the subgraphs $\sigma \in \{0,1\}^E$ of a graph $G = (V, E)$ as follows.:

$$\mu^{\text{sf}}(\sigma) = \begin{cases} \dfrac{1}{Z^{G,\text{sf}}}, & \text{if } \sigma \text{ is a forest,} \\ 0, & \text{otherwise.} \end{cases} \tag{5.10}$$

Here $Z^{G,\text{sf}}$ denotes the number of spanning forests of the graph G. In general, almost nothing is known about this model. We prove it is p-NC for some classes of graphs in the last section. Grimmett and Winkler proved the following using an algorithmic approach:

Theorem 5.2.4. *If a graph $G = (V, E)$ has eight or fewer vertices, or has nine vertices with eighteen and fewer edges, then the USF measure on G is negative edge dependent.*

Stark proved the following for USF in 2008:

Theorem 5.2.5. *Let $\mathcal{F}$ be a spanning forest of K_n chosen accordingly from USF. Then for sufficiently large n we have*

$$\mu^{K_n,sf}(e, f \in \mathcal{F}) \leq \mu^{K_n,sf}(e \in \mathcal{F})\mu^{K_n,sf}(f \in \mathcal{F}), \tag{5.11}$$

or for lesser confusion

$$\mathbb{P}(e, f \in \mathcal{F}) \leq \mathbb{P}(e \in \mathcal{F})\mathbb{P}(f \in \mathcal{F}), \tag{5.12}$$

i.e the measure $\mu^{K_n,sf}$ is p-NC for large enough n.

Remark 5.2.5. Stark proves some asymptotics for the above results which is as follows:

$$\mathbb{P}(e \in \mathcal{F})\mathbb{P}(f \in \mathcal{F}) = \frac{4}{n^2} - \frac{4}{n^3} - \frac{23}{n^4} + O\left(\frac{1}{n^5}\right), \tag{5.13}$$

and for adjacent edges e, f,

$$\mathbb{P}(e, f \in \mathcal{F}) = \frac{3}{n^2} + O\left(\frac{1}{n^3}\right), \tag{5.14}$$

and for non-adjacent edges e, f,

$$\mathbb{P}(e, f \in \mathcal{F}) = \frac{4}{n^2} - \frac{4}{n^3} - \frac{27}{n^4} + O\left(\frac{1}{n^5}\right). \tag{5.15}$$

Note that we need the analysis upto order $O(n^{-5})$ due to non-adjacent edges.

A natural question to ask at this point is conditioning. What happens when we condition on the connected component size of the forest?

5.2.3. UCS

We define the **uniform connected subgraph measure** denoted as μ^{cs} to be the measure defined on the subgraphs $\sigma \in \{0,1\}^E$ of a graph $G = (V, E)$ as follows:

$$\mu^{\mathrm{sf}}(\sigma) = \begin{cases} \dfrac{1}{Z^{G,\mathrm{cs}}}, & \text{if } \kappa(\sigma) = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{5.16}$$

Here $Z^{G,\mathrm{cs}}$ denotes the number of spanning forests of the graph G.

5.3. RANDOM CLUSTER MODEL

Let $G = (V, E)$ be a graph with $|V| = m$, and $|E| = n$. A random cluster measure on the graph G is a measure on the subgraphs of G. By a subgraph of G, we mean a graph which has the same vertex set V, and an edge set $E' \subset E$. As the state space, we take the set $\Sigma = \{0,1\}^E$, where the vector $\sigma = (\sigma(e))_{e\in E}$ will be defined according to the edge being present or not. That is $\sigma(e) = 1$ if the edge is present in the subgraph and $\sigma(e) = 0$ if the edge is not present in the subgraph. Define $o(\sigma) = |e : \sigma(e) = 1|$. We will mainly work on finite graphs, hence we can order them. We define the parameter κ to the element σ so that $\kappa(\sigma)$ is the number of connected components of the subgraph. We assign the probability to a subgraph σ with the parameters $p \in [0,1]$ and $q \in (0,\infty)$ proportional to

$$\mu^G_{p,q}(\sigma) \propto \prod_{e\in E} p^{o(\sigma)}(1-p)^{n-o(\sigma)}q^{\kappa(\sigma)}. \tag{5.17}$$

Define the partition function

$$Z^G_{p,q} = \sum_{\sigma\in\Sigma} \mu^G_{p,q}(\sigma). \tag{5.18}$$

The partition function is the normalizing constant for the probability measure, i.e.

$$\mu^G_{p,q}(\sigma) = \frac{1}{Z^G_{p,q}} \prod_{e \in E} p^{o(\sigma)}(1-p)^{n-o(\sigma)} q^{\kappa(\sigma)}. \tag{5.19}$$

We observe that the insertion of the term q^κ differentiates the measure from the usual product. When $q = 1$, the measure becomes the product measure.

Note that the behaviour of the measure is very different in the cases $q < 1$ and $q > 1$. When $q > 1$ the measure concentrates on the subgraphs with more connected components, whereas when $q < 1$ the measure concentrates on the subgraphs with fewer connected components. The strength of the random cluster model is that it generalizes many other models.

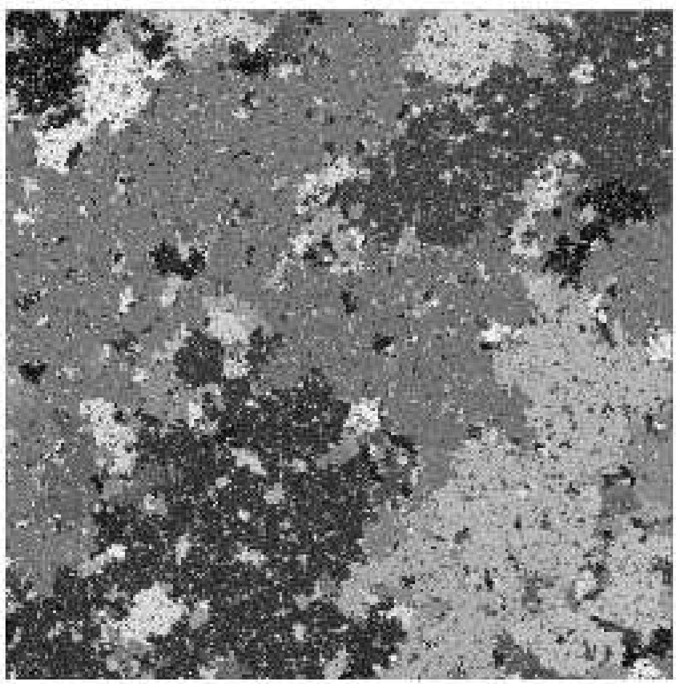


Figure 5.4. Random Cluster Model with $q = 2$ on a 2048×2048 box.

5.3.1. Basic Properties of Random Cluster Measure

Many properties of the random cluster measure is determined and can be observed through the partition function. We state some properties of the random cluster measure. We focus on comparison inequalities and how UST, USF, UCS measures arises as a limit of the random cluster model.

5.3.1.1. Ising/Potts Model. Ising/Potts models first arises in statistical physics. It is used to model particle interactions. We suppose that the particles (particles of a magnetized iron or electrons interacting with each other) are positioned on the vertices of a lattice such as $\mathbb{Z}^d$. In general, they may be positioned on the vertices of any graph. The lattice $\mathbb{Z}^d$ is of importance from a physical perspective.

Let $G = (V, E)$ be a finite graph. We think of the vertices as being occupied by a particle with a random spin. In the Ising model, the spin space has 2 elements which we denote by $\{-1, 1\}$, and in the Potts model which is a slight generalization of the Ising model, the spin space has $q \in \mathbb{Z}^+$ elements which we denote by $\{1, \ldots, q\}$. Hence, our sample space is $\Sigma = \{-1, 1\}^V$ in the Ising model and $\{1, \ldots, q\}^V$ in the Potts model.

The probability measure $\mu_{\beta,J,h}$ on Σ has three parameters with $\beta, J \in [0, \infty)$ and $h \in \mathbb{R}$ is given by

$$\mu_{\beta,J,h}(\sigma) = \frac{1}{Z_I} e^{-\beta H(\sigma)}, \quad \sigma \in \Sigma, \tag{5.20}$$

where the hamiltonian $H : \Sigma \to \mathbb{R}$ is the following energy function

$$H(\sigma) = -J \sum_{e=xy\in E} \sigma(x)\sigma(y) - h \sum_{x\in V} \sigma_x. \tag{5.21}$$

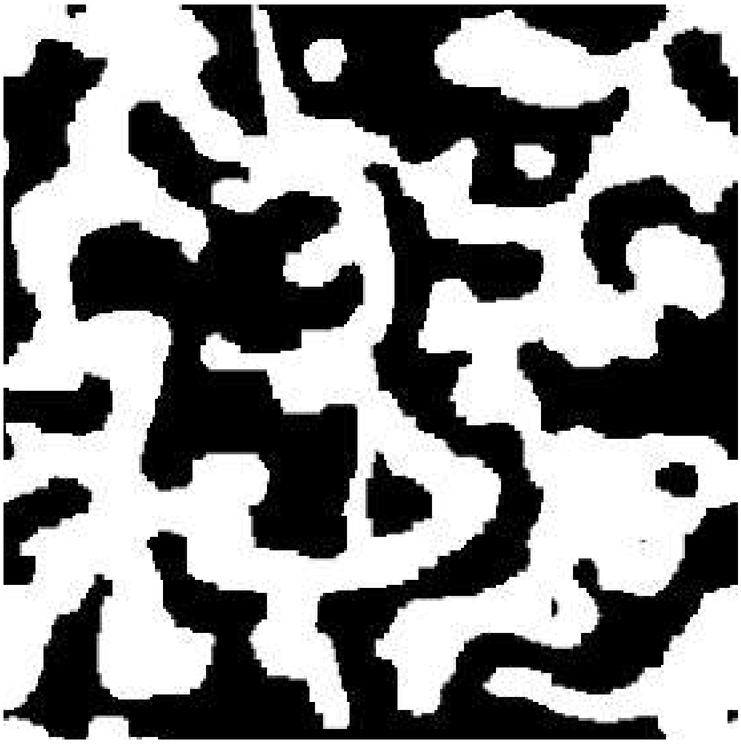


Figure 5.5. Ising Model at $T = 0.000611$.

Here, Z_I is the partition function for the Ising model, which is

$$Z_I = \sum_{\sigma \in \Sigma} e^{-\beta H(\sigma)}. \tag{5.22}$$

In physical terms β denotes the inverse temperature of the system $\frac{1}{T}$, and J is the bonding energy of the particles.

Remark 5.3.1.
- If we look at the Hamiltonian, we observe that there are two parts to it. The first one is the interaction part, that is for any edge the vertices incident to the edge interacts with each other which is seen in $\sigma(x)\sigma(y)$ for $e = xy$. The second one is the part that shows the energy of the vertex in and of itself in a manner of putting.
- Since the interaction J in this basic model does not depend on edges, we can think of β and J as a single parameter. They occur only as βJ.

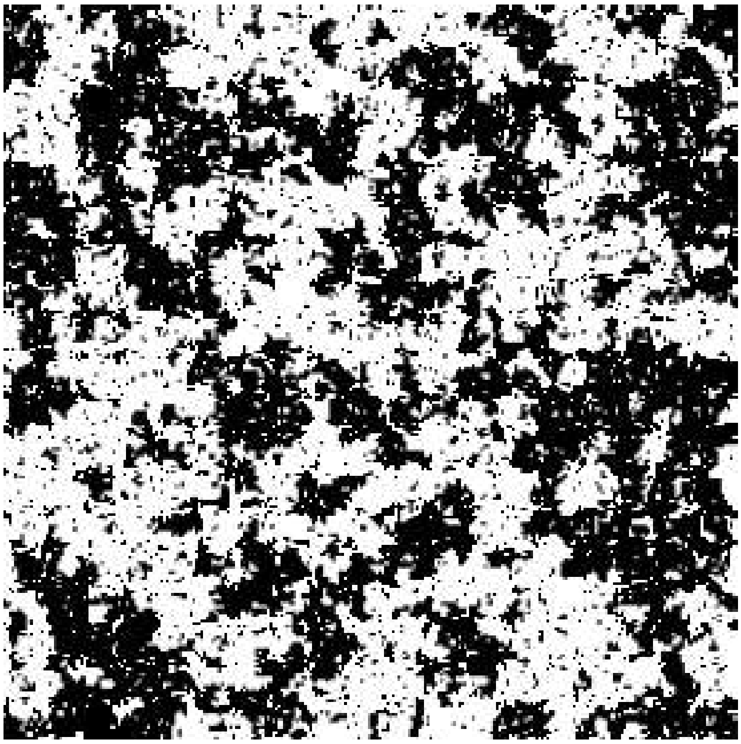


Figure 5.6. Ising Model at $T = 2.333625$.

For the Potts model, we have q states $\{1, \ldots q\}$. Let δ_{xy} denote the Kronecker delta symbol, i.e.

$$\delta_{xy} = \begin{cases} 1, & x = y \\ 0, & x \neq y. \end{cases} \tag{5.23}$$

The Potts measure is

$$\mu_{\beta,q}(\sigma) = \frac{1}{Z_P} e^{-\beta H^P(\sigma)}, \quad \sigma \in \Sigma, \tag{5.24}$$

where the hamiltonian $H^P : \Sigma \to \mathbb{R}$ is given by

$$H^P(\sigma) = -\sum_{e=xy \in E} \delta_e(\sigma), \tag{5.25}$$

with $\delta_e(\sigma) = \delta_{\sigma(x)\sigma(y)}$. The partition function for the Potts model then becomes

$$Z_P = \sum_{\sigma \in \Sigma} e^{-\beta H^P(\sigma)}. \tag{5.26}$$

We will see that Ising and Potts models are special cases of the random cluster model when we fix q. To see this, we state a coupling in the edge and vertex space of a graph where the marginal on the edge space is the random cluster model and the marginal on the vertex space is the Ising/Potts model.

Let $q \in \{2, 3, \ldots\}$ and $p \in [0, 1]$ and let $G = (V, E)$ be a finite graph. Consider the product space $\Sigma \times \Omega$ where $\Sigma = \{1, 2, \ldots q\}^V$ and $\Omega = \{0, 1\}^E$. We define a probability measure proportional to the following quantity,

$$\mu(\sigma, \omega) \propto \prod_{e \in E} [(1-p)\delta_{\omega(e)0} + p\delta_{\omega(e)1}\delta_e(\sigma)], \quad (\sigma, \omega) \in \Sigma \times \Omega. \tag{5.27}$$

The exact probability is again by multiplying with the normalizing constant which is $\dfrac{1}{\sum_{(\sigma,\omega)\in\Sigma\times\Omega} \mu(\sigma, \omega)}$.

Now, we look at the marginal probability measures.

Theorem 5.3.1 (Marginals of μ.)**.** *Let $q \in \{2, 3, ...\}$ and $p = 1 - e^{-\beta} \in [0, 1]$.*

(a)Marginal on Σ. The marginal measure $\mu_\Sigma(\sigma) = \sum_{\omega\in\Omega} \mu(\sigma, \omega)$ on Σ is the Potts measure

$$\mu_\Sigma(\sigma) = \frac{1}{Z_P} \exp\left(\beta \sum_{e \in E} \delta_e(\sigma)\right), \quad \sigma \in \Sigma. \tag{5.28}$$

(b)Marginal on Ω. The marginal measure $\mu_\Omega(\omega) = \sum_{\sigma\in\Sigma} \mu(\sigma, \omega)$ on Ω is the random cluster measure

$$\mu_\Omega(\omega) = \frac{1}{Z_{p,q}^G} \left(\prod_{e \in E} p^{o(e)}(1-p)^{1-o(e)}\right) q^{\kappa(\omega)}, \quad \omega \in \Omega. \tag{5.29}$$

(c)Partition Functions. We have that

$$\sum_{\omega\in\Omega}\left(\prod_{e\in E} p^{o(e)}(1-p)^{1-o(e)}\right) q^{\kappa(\omega)} = \sum_{\sigma\in\Sigma}\prod_{e\in E}\exp\left(\beta(\delta_e(\sigma)-1)\right), \tag{5.30}$$

which is the same as

$$Z^G_{p,q} = e^{-\beta|E|} Z_P(\beta,q). \tag{5.31}$$

We also look at the conditional measures of μ.

Theorem 5.3.2 (Conditional Measures of μ.). *Let μ and p be as in the above Theorem 5.3.1.*

(a)For $\omega\in\Omega$, the conditional measure $\mu(\cdot|\omega)$ on Σ is the measure we get when we put random spins on the clusters determined by ω of which there are $\kappa(\omega)$ many. Each cluster gets one of the spins $\{1,\dots,q\}$ at uniformly random and independently.

(b)For $\sigma\in\Sigma$, the conditional measure $\mu(\cdot|\sigma)$ on Σ is the measure we get as follows. If $e = xy$ is such that the spins are not the same, i.e. $\sigma(x)\neq\sigma(y)$, do not include the edge, i.e. $o(e) = 0$. If the spins are the same, i.e. $\sigma(x) = \sigma(y)$, we will include the edge with probability p, i.e.

$$o(e) = \begin{cases} 1, & \textit{with probability } p \\ 0, & \textit{otherwise,} \end{cases} \tag{5.32}$$

with the values of distinct e being independent (conditionally) random variables.

Remark 5.3.2. The conditioning in 5.3.2 is called the Edward-Sokal coupling. We give below a visual explanation.

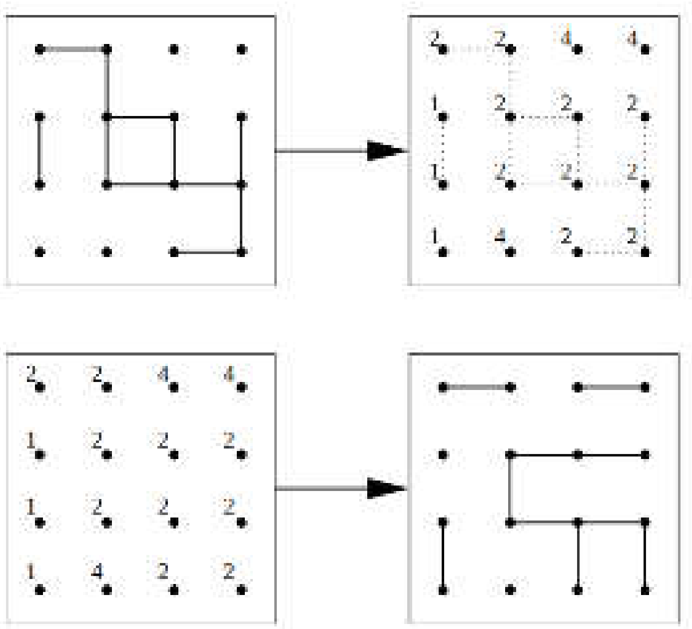


Figure 5.7. Conditional Measures of μ.

5.3.1.2. UST,USF and UCS as limits of RCM. We look at the limits of the random cluster model as $q \to 0$. The behaviour of $\frac{q}{p}$ will change the limiting measure as we will see. We will look at the weak limits and focus on the partition function $Z^G_{p,q}$. Let $G = (V, E)$ with $|E| = n$.

First, consider the weak limit as $q \to 0$ when $p \in (0, 1)$ is fixed. We have the following for the partition function

$$Z^G_{p,q} = \sum_{\sigma \in \{0,1\}^E} p^{o(\sigma)}(1-p)^{n-o(\sigma)} q^{\kappa(\sigma)}. \tag{5.33}$$

Since we take the limit $q \to 0$, observe that the dominant terms of the $Z^G_{p,q}$ are those with minimal $\kappa(\sigma)$. Therefore, the dominant terms are those with $\kappa(\sigma) = 1$, which are the connected subgraphs of G. Hence $\lim_{q \downarrow 0} \mu_{p,q}$ is the product measure $\mu_{p,1}$ conditioned on the resulting graph being connected, i.e. $\mu_{p,q} \Rightarrow \mu_r^{\text{cs}}$, where $r = \frac{p}{1-p}$, and

$$\mu_r^{\text{cs}}(\sigma) = \begin{cases} \dfrac{1}{Z_{\text{cs}}} r^{o(\sigma)}, & \text{if } \kappa(\sigma) = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{5.34}$$

Here, Z_{cs} is the appropriate normalizing constant. Now, if we take $p = \frac{1}{2}$, we see that $\mu_{p,q} \Rightarrow \text{UCS}$.

Now, we look at the weak limit when we send both p and q to 0. First, let us rearrange the partition function as

$$Z_{p,q}^G = (1-p)^n \sum_{\sigma \in \{0,1\}^E} \left(\frac{p}{1-p}\right)^{o(\sigma)+\kappa(\sigma)} \left(\frac{q(1-p)}{p}\right)^{\kappa(\sigma)}. \tag{5.35}$$

Note that $\frac{p}{1-p} \to 0$ since $p \to 0$. If we also let $\frac{q}{p} \to 0$, that is if p tends to 0 slower than q, we see that $\frac{q(1-p)}{p} \to 0$. Thus, the dominant terms will be the ones when both $o(\sigma) + \kappa(\sigma)$ and $\kappa(\sigma)$ is minimized. We know that $\kappa(\sigma) \geq 1$ and $o(\sigma) + \kappa(\sigma) \geq |V|$. These two are satisfied as equality when the subgraph is a spanning tree. Therefore, the limiting measure will be concentrated around the spanning trees. We have $\mu_{p,q} \Rightarrow \text{UST}$.

Now, suppose that $\frac{p}{q} = a$, that is q and p tends to 0 at the same rate. We can plug in $p = aq$ into the partition function and upon rearranging we have

$$Z_{p,q}^G = (1-aq)^n \sum_{\sigma \in \{0,1\}^E} \left(\frac{a}{1-aq}\right)^{o(\sigma)} q^{o(\sigma)+\kappa(\sigma)}. \tag{5.36}$$

Note that in this case we only need to minimize $o(\sigma) + \kappa(\sigma) \geq |V|$ which happens if and only if σ is a forest, i.e. $o(\sigma) + \kappa(\sigma) = |V|$ if σ is a forest. Therefore, we have that $\mu_{p,q} \Rightarrow \mu_a^{\text{sf}}$. where μ_a^{sf} is the measure on forest with edge weight a, i.e.

$$\mu_a^{\text{sf}}(\sigma) = \begin{cases} \dfrac{1}{Z_{\text{sf}}} a^{o(\sigma)}, & \text{if } \sigma \text{ is a forest} \\ 0, & \text{otherwise.} \end{cases} \tag{5.37}$$

Here Z_{sf} is the appropriate normalizing constant. Now, when $a = 1$, we get that $\mu_{p,q} \Rightarrow$ USF.

Now, we summarize what we have done as follows:

Theorem 5.3.3. *As $q \downarrow 0$, we have that*

$$\mu_{p,q} \Rightarrow \begin{cases} UCS, & \textit{if } p = \frac{1}{2} \\ UST, & \textit{if } p \to 0 \textit{ and } \frac{q}{p} \to 0 \\ USF, & \textit{if } p = q. \end{cases} \tag{5.38}$$

5.3.2. Positive Association

When $q = 1$, we know that the random cluster measure is just the Erdos-Renyi measure with edge probability p. The edges are independent in this case.

In this section, we prove that when $q > 1$ the random cluster measure $\mu_{p,q}^G$ is positively associated. To prove positive association, we show that the random cluster measure satisfies the positive lattice condition.

Theorem 5.3.4. *Let $p \in (0, 1)$, and $q \in [1, \infty)$.*

(i)The random cluster measure $\mu_{p,q}$ is strictly positive and satisfies the positive lattice condition.

(ii)The random cluster measure $\mu_{p,q}$ is strongly positively associated, and in particular we have that

$$\mathbb{E}_{\mu_{p,q}}[XY] \geq \mathbb{E}_{\mu_{p,q}}[X]\mathbb{E}_{\mu_{p,q}}[Y] \;\; for\ increasing\ X, Y : \Sigma \to \mathbb{R},$$
$$\mu_{p,q}(A \cap B) \geq \mu_{p,q}(A)\mu_{p,q}(B) \;\; for\ increasing\ A, B \in \mathcal{F}.$$

Proof. We only need to prove part (i), since the second one follows. Now, we check the positive lattice condition. We observe that $o(x \vee y) + o(x \wedge y) = o(x) + o(y)$. Taking logarithms we see that it is okay to prove $\kappa(x \vee y) + \kappa(x \wedge y) \geq \kappa(x) + \kappa(y)$. We can focus on the states that differ by two edges exactly, which means that there is a configuration z and two edges e, f such that $x = z_f^e$ and $y = z_e^f$. Now, we can omit the other edges and write $x = 10$ and $y = 01$ indicating the edges e, f being open or not. Let C_f be the indicator function of connection of the endvertices of f being connected by no open path of $E \setminus f$. We observe that C_f is a decreasing random variable, and hence $C_f(10) \leq C_f(00)$. Hence, we have $\kappa(10) - \kappa(11) = C_f(10) \leq C_f(00) = \kappa(00) - \kappa(01)$, which is the desired inequality. ∎

6. NEGATIVE EDGE DEPENDENCE IN RCM

The following are our work joint with Mohan Ravichandran is in preparation. We can redefine the random-cluster model by suppressing the probability p of an edge being open and adding a new complex variable z_e to each edge. One can think of it as putting the probability p into the variable z_e. In this way, the multivariate partition function of the random cluster model becomes

$$Z_G(\{z_e\}, q) = \sum_{S \subset E} q^{\kappa(S)} \prod_{e \in S} z_e. \tag{6.1}$$

We fix two edges e, f. We will prove the negative edge dependence, p-NC when $q < 1$ and the underlying graph is a graph of our choice which will be mentioned later on. Since, we work with the fixed edges, we are interested in the polynomial that we get by setting $z_e = x$ and $z_f = y$ and the other variables to z. We may write this polynomial as $p(x, y, z) = p_{00}(z) + p_{10}(z)x + p_{01}(z)y + p_{11}(z)xy$, where $p_{00}(z)$ is the sum over the configurations not containing the edges e, f and $p_{11}(z)$ is the sum over all configurations containing both e and f and $p_{10}(z)$ and $p_{01}(z)$ is the sum over all configurations containing e and not f, and containing f but not e respectively.

With the polynomials we wrote above, the negative edge dependence is equivalent to $p_{10}\, p_{01} \geq p_{00}\, p_{11}$.

Remark 6.0.1. Note that this expression is an equality when $q = 1$ since

$$Z_G(\{z_e\}, 1) = \prod_{g \in E} (1 + z_g), \tag{6.2}$$

and the Rayleigh polynomial is zero. This is reasonable and expected since when $q = 1$, we have independence among each edge.

The inequality also holds for q sufficiently close to 1 : Letting $q = 1 - \epsilon$,

$$Z_G(\{z_e\}, 1-\epsilon) = \prod_{g \in E}(1+z_g) - \epsilon \sum_{S \subset E} \kappa(S) z^S + O(\epsilon^2). \tag{6.3}$$

The Rayleigh expression can be written as follows.

$$\begin{aligned}
\partial_e Z_G(\{z_e\}, q) =& \frac{1}{1+z_e} \prod_{g \in E}(1+z_g) - \frac{\epsilon}{z_e} \sum_{S \ni e} \kappa(S) z^S, \\
\partial_f Z_G(\{z_e\}, q) =& \frac{1}{1+z_f} \prod_{g \in E}(1+z_g) - \frac{\epsilon}{z_f} \sum_{S \ni f} \kappa(S) z^S, \\
\partial_{e,f} Z_G(\{z_e\}, q) =& \frac{1}{(1+z_e)(1+z_f)} \prod_{g \in E}(1+z_g) - \frac{\epsilon}{z_e\, z_f} \sum_{S \ni e,f} \kappa(S) z^S.
\end{aligned}$$

The desired inequality reduces to showing

$$\sum_T \left[\kappa(T \cup \{e\}) + \kappa(T \cup \{f\}) - \kappa(T \cup \{e,f\}) - \kappa(T)\right] z^T \geq 0, \tag{6.4}$$

where the sum is over all subsets of $E \setminus \{e, f\}$. The expression is term by term non-negative by the submodularity of the rank function. Note that this works just as well for matroids.

6.1. Connected Subgraphs of the Complete Graph

Let $G(n) = 2^{\binom{n}{2}}$ be the number of graphs on n labeled vertices and let $C(n)$ be the number of connected subgraphs on n vertices. Letting A and B be the exponential generating functions of the sequences $G(n)$ and $C(n)$, as follows

$$A(x) = \sum_{n=0}^{\infty} 2^{\binom{n}{2}} \frac{x^n}{n!}, \quad B(x) = x + \sum_{n=1}^{\infty} C(n) \frac{x^n}{n!}. \tag{6.5}$$

The exponential formula says that $e^{B(x)} = A(x)$. Consequently, $B(x) = \log\left(\sum_{n=0}^{\infty} 2^{\binom{n}{2}} \frac{x^n}{n!}\right)$. The coefficients $C(n)$ have been well studied and the first few numbers are as follows : $1, 1, 1, 4, 38, 728, 26704, 1866256, 251548592, 66296291072, 34496488594816$. This is the sequence $A001187$ in the OEIS.

Lemma 6.1.1. *Let $U(n)$ be the number of connected subgraphs of K_n containing the edge $(1,2)$. Then*

$$U(n) = \frac{C(n)}{2} + \frac{1}{2}\sum_{k=1}^{n-1} \frac{C(k)}{(k-1)!}\frac{C(n-k)}{(n-k-1)!}. \tag{6.6}$$

Consequently, we have that

$$\mathbf{U}'' = \frac{\mathbf{C}'' + (\mathbf{C}')^2}{2}. \tag{6.7}$$

Lemma 6.1.2. *Let $V(n)$ be the number of connected subgraphs of K_n containing the edges $(1,2)$ and $(2,3)$. Then*

$$V(n) = \frac{U(n)}{2} + \frac{1}{2}\sum_{k=2}^{n-1} U(k)C(n-k)\binom{n-3}{k-2}. \tag{6.8}$$

Consequently,

$$\mathbf{V}^{(3)} = \frac{\mathbf{C}^{(3)} + 3\mathbf{C}''\mathbf{C}' + (\mathbf{C}')^3}{4}. \tag{6.9}$$

Lemma 6.1.3. *Let $W(n)$ be the number of connected subgraphs of K_n containing the edges $(1,2)$ and $(3,4)$. Then*

$$W(n) = \frac{U(n)}{2} + \sum_{k=3}^{n-1} U(k)C(n-k)\binom{n-4}{k-3}.$$
$$\mathbf{W}^{(4)} = \frac{\mathbf{C}^{(4)} + 4\mathbf{C}^{(3)}\mathbf{C}' + 2(\mathbf{C}'')^2 + 4\mathbf{C}''(\mathbf{C}')^2}{4}.$$

We may rewrite the results as follows.

Proposition 6.1.1. *Let* $\mathbf{G}$, $\mathbf{C}$, $\mathbf{U}$, $\mathbf{V}$ *and* $\mathbf{W}$ *be the exponential generating functions of all graphs, connected graphs, connected graphs containing $(1,2)$, connected graphs* *containing* $(1,2)\cup(2,3)$ *and connected graphs containing* $(1,2)\cup(3,4)$. *Then*

$$\begin{aligned}\mathbf{C}' =& \frac{\mathbf{G}'}{\mathbf{G}},\\ \mathbf{U}'' =& \frac{\mathbf{G}''}{2\,\mathbf{G}},\\ \mathbf{V}^{(\mathbf{3})} =& \frac{\mathbf{G}^{(\mathbf{3})}}{4\,\mathbf{G}},\\ \mathbf{W}^{(\mathbf{4})} =& \frac{\mathbf{G}^{(\mathbf{4})}}{4\,\mathbf{G}} - \left(\frac{\mathbf{G}^{(\mathbf{2})}}{2\,\mathbf{G}}\right)^2.\end{aligned}$$

6.1.1. Asymptotics

Proposition 6.1.2. *The exponential generating function for connected subgraphs of* K_n *satisfies*

$$\mathbf{C}(x) = \sum_{n=1}^{\infty} \frac{x^n}{n!}\left[\sum_{\substack{n_1+\ldots+n_k=n,\\ n_1,\ldots,n_k>0}} \frac{(-1)^{k-1}}{k}\binom{n}{n_1,\ldots,n_k} 2^{\binom{n_1}{2}+\ldots+\binom{n}{n_k}}\right]. \tag{6.10}$$

We have the asymptotics

$$C(n) = 2^{\binom{n}{2}}\left[1 - \frac{2n}{2^n} - \frac{128}{3}\frac{n(n-1)(n-2)}{2^{7n/2}} + \ldots\right]. \tag{6.11}$$

Lemma 6.1.4. *The coefficients satisfy the asymptotics*

$$C(n) = 2^{\binom{n}{2}}\left[1 - \frac{2n}{2^n} + O\left(\frac{1}{2^{7n/2}}\right)\right]. \tag{6.12}$$

For connected subgraphs containing a particular edge

$$U(n) = 2^{\binom{n}{2}}\left[\frac{1}{2} - \frac{n-2}{2^n} - \frac{4}{2^{2n}} + O\left(\frac{1}{2^{7n/2}}\right)\right]. \tag{6.13}$$

For connected subgraphs containing two disjoint edges, we have

$$W(n) = 2^{\binom{n}{2}}\left[\frac{1}{4} - \frac{n-4}{2^{n+1}} - \frac{6}{2^{2n}} + O\left(\frac{1}{2^{7n/2}}\right)\right]. \tag{6.14}$$

Consequently,

$$U(n)^2 - C(n)W(n) = 4^{\binom{n}{2}}\left[\frac{2}{4^n} + O\left(\frac{1}{8^n}\right)\right] \tag{6.15}$$

$$C(n) = 2^{\binom{n}{2}}\left[1 - \frac{2n}{2^n} - \frac{n(n-1)}{2^{2n-2}} + \frac{n^3\theta}{2^{3n}}\right],$$

$$U(n) = 2^{\binom{n}{2}}\left[\frac{1}{2} - \frac{n-2}{2^n} - \frac{4}{2^{2n}} - \frac{5n(n-1)}{2^{2n-1}} + \frac{n^3}{2^{3n}}\theta'\right],$$

where $\theta, \theta' < 0.1$ *whenever* $n > 10$.

6.2. The General Random cluster model for the Complete Graph at 1

Let C_k be the exponential generating function for subgraphs with k connected components and with . Then $C_k = k!\, C^k$. If U_k is the exponential generating function for subgraphs with connected components and containing a fixed edge e, we have that

$$U_k(n) = \sum_{k=2}^{n-1} U(k)C(n-k)\binom{n-2}{k-2}, \tag{6.16}$$

yielding that

$$U_k'' = \frac{U''(C^{k-1})}{(k-1)!}. \tag{6.17}$$

Moreover, if W_k is the exponential generating function for subgraphs with connected components and containing two fixed non-adjacent edges e and f, we have that

$$W_k(n) = \sum_{m=4}^{n-k+2} W_2(m)C_{k-2}(n-m)\binom{n-4}{m-4}$$
$$+ \sum_{m=4}^{n-k+1} W(m)C_{k-1}(n-m)\binom{n-4}{m-4},$$

yielding

$$W_k^{(4)} = \frac{W^{(4)}C^{k-1}}{(k-1)!} + \frac{(U'')^2C^{k-2})}{(k-2)!}. \tag{6.18}$$

6.3. Approaches

Let $G = (V, E)$ be a graph and let $e, f \in E$ and let $0 < q < 1$. Let Z_G be the partition function of the random cluster model,

$$Z_G(\{z_e\}, q) = \sum_{S \subset E} q^{\kappa(S)} \prod_{e \in S} z_e, \tag{6.19}$$

and let $\Delta_{e,f}$ be the Rayleigh difference

$$\Delta_{e,f}(\{z_e\}, q) = (\partial_e Z_G)(\partial_f Z_G) - Z_G(\partial_e \partial_f Z_G). \tag{6.20}$$

Let us consider this as a polynomial in q, i.e. let us treat the $\{z_e\}$ as fixed but arbitrary. Note that this polynomial vanishes when $q = 1$ by the independence of the edges.

Conjecture 6.3.1. *The polynomial f defined as*

$$f(q) = \frac{\Delta_{e,f}(\{z_e\}, q)}{1 - q}, \tag{6.21}$$

has positive coefficients.

Remark 6.3.1. *Observe that this will be more than showing negative edge dependence actually. Since, in the end we only want the positivity of the whole function not coefficient by coefficient positivity.*

The following will prove that the random cluster model on K_n is negatively correlated at 1.

Conjecture 6.3.2. *Let*

$$A = \{(S, T), S \ni e, \not\ni f, T \ni f, \not\ni e \mid \kappa(S) + \kappa(T) = k\} \tag{6.22}$$

$$B = \{(S, T), S \ni e, f, T \not\ni e, f \mid \kappa(S) + \kappa(T) = k\}. \tag{6.23}$$

Then for any $k > 2$, $|A| \geq |B|$.

6.4. Adjacent Edges

Let us fix some notation.

Definition 6.4.1. Given $n, k \in \mathbb{N}$ and collections of edges $e_1, \ldots, e_k$ and $f_1, \ldots, f_l$ and collections of vertices $\mathbf{V} = (V_1, \ldots, V_j)$, we will use $C_k(e_1, \ldots, e_k, \bar{f}_1, \ldots, \bar{f}_l, \mathbf{V})$, to represent the number of all subgraphs S on n labeled vertices such that

- The graphs have exactly k connected components.
- The vertices in V_i for each $i \in [j]$ lie in the same connected component.
- The vertices in V_i and $V_{i'}$ for $i \neq i'$ lie in different connected components.
- The graphs contain the edges $e_1, \ldots, e_k$.
- The graphs do not contains the edges $f_1, \ldots, f_l$.

We will use the usual economies, writing $C_1(\cdots)$ as $C(\cdots)$ and $C_i()$ as C_i. Note that in this notation,

- The number of connected graphs may be written as C.
- The number of connected graphs containing a given edge e is $C(e)$.
- Suppose the edge $e = (1, 2)$. Then the number of connected graphs containing e which become disconnected when we remove this edge is $C_2(\bar{e}, [\{1\}, \{2\}])$.

Conjecture 6.4.1. *Consider the complete graph K_n. Then for any $k \geq 2$,*

$$\sum_{i+j \leq k} [C_i(e, \bar{f})C_j(f, \bar{e}) - C_i(e, f)C_j(\bar{e}, \bar{f})] > 0. \tag{6.24}$$

We further conjecture that the analogous statement holds for all graphs.

Let $e = (1, 2)$ and $f = (2, 3)$. Let us consider the general expression

$$I_k = \sum_{s+t \leq k} C_s(e, \bar{f})C_t(f, \bar{e}) - C_s(e, f)C_t(\bar{e}, \bar{f}). \tag{6.25}$$

We have that

$$\begin{aligned} C_s(e,\bar f) &= C_s(\bar e,\bar f,[\{1,2\}]) + C_{s+1}(\bar f,[\{1\},\{2\}]) \\ &= C_s(\bar e,\bar f) - C_s(\bar f,[\{1\},\{2\}]) + C_{s+1}(\bar f,[\{1\},\{2\}]). \end{aligned}$$

We also have that

$$\begin{aligned} C_s(e,f) &= C_s(f,\bar e,[\{1,2\}]) + C_{s+1}(f,[\{1\},\{2,3\}]) \\ &= C_s(f,\bar e) - C_s(f,1,2) + C_{s+1}(f,[\{1\},\{2,3\}]). \end{aligned}$$

We now compute

$$\begin{aligned} I_k &= \sum_{s+t\le k} C_s(e,\bar f)\,C_t(\bar e,f) - C_s(e,f)\,C_t(\bar e,\bar f) \\ &= \sum_{s+t\le k} \left[C_s(\bar e,\bar f) - C_s(\bar f,[\{1\},\{2\}]) + C_{s+1}(\bar f,[\{1\},\{2\}])\right]\ C_t(\bar e f) \\ &- \sum_{s+t\le k} \left[C_s(f,\bar e) - C_s(f,1,2) + C_{s+1}(f,[\{1\},\{2,3\}])\right]\ C_t(\bar e,\bar f) \\ &= \sum_{s+t\le k} \left[C_{s+1}(\bar f,[\{1\},\{2\}]) - C_s(\bar f,[\{1\},\{2\}])\right] C_t(\bar e,f) \\ &- \sum_{s+t\le k} \left[C_{s+1}(f,[\{1\},\{2\}]) - C_s(f,[\{1\},\{2\}])\right] C_t(\bar e,\bar f) \\ &= \sum_{s+t=k+1} C_s(\bar f,1,2)C_t(\bar e,f) - \sum_{s+t=k+1} C_s(f,1,2)C_t(\bar e,\bar f) \\ &= \sum_{s+t=k+1} C_s(\bar f,1,2)\left[C_t(\bar e,\bar f,23) + C_{t+1}(\bar e,2,3)\right] - \sum_{s+t=k+1} C_s(f,1,2)C_t(\bar e,\bar f). \end{aligned}$$

This means that I_k equals

$$\sum_{s+t=k+1} C_s(\bar f,1,2)C_t(\bar e,f,12) - \sum_{s+t=k+1} C_s(f,1,2)C_t(\bar e,\bar f,12). \tag{6.26}$$

Similarly we have,

$$\begin{aligned} C_t(f,\bar{e}) &= C_t(\bar{e},\bar{f},[\{2,3\}]) + C_{t+1}(\bar{e},[\{2\},\{3\}]) \\ &= C_t(\bar{e},\bar{f},[\{2,3\}]) + C_{t+1}(\bar{e},[\{1,2\},\{3\}]) + \\ &\quad C_{t+1}([\{1\},\{2\},\{3\}]) + C_{t+1}([\{1,3\},\{2\}]). \end{aligned}$$

Note that

$$C_s(f,1,2) = C_s(\bar{f},1,23) + C_{s+1}(1,2,3).$$
$$C_s(\bar{f},1,2) = C_s(1,2,3) + C_s(\bar{f},1,23) + C_s(13,2).$$

Remark 6.4.1. We see that

$$\begin{aligned} I_k &= \sum_{s+t\leq k} C_s(e,\bar{f})C_t(f,\bar{e}) - C_s(e,f)C_t(\bar{e},\bar{f}) \\ &\geq \sum_{s+t\leq k} \left[C_s(\bar{e},\bar{f}) + C_{s+1}([\{1,3\},\{2\}])\right] C_t(f,\bar{e}) - \\ &\quad [C_s(f,\bar{e}) + C_{s+1}([\{1\},\{2,3\}])]\, C_t(\bar{e},\bar{f}) \\ &= \sum_{s+t\leq k} C_{s+1}([\{1,3\},\{2\}])\, C_t(f,\bar{e}) - C_{s+1}([\{1\},\{2,3\}])\, C_t(\bar{e},\bar{f}) \\ &= \sum_{s=1}^{k-1} C_{s+1}([\{1,3\},\{2\}]) \sum_{t=1}^{k-s} \left[C_t(f,\bar{e}) - C_t(\bar{e},\bar{f})\right], \\ &\geq 0. \end{aligned}$$

Remark 6.4.2. Given two connected subgraphs S, T such that S contains e but not f and T contains f but not e, we could consider the map $(S,T) \to (S \cup \{f\}, T \setminus \{f\})$. This is a bijection between

- Pairs (S,T) of connected subgraphs such that S contains e but not f and T contains f but not e and additionally such that f is a not a cut edge in T. Consider the complements of these sets. There are two configurations of T that we will analyze separately.
  - $T \setminus \{f\}$ is a union of two connected graphs such that $\{1,2\}$ lies in one

component and $\{3\}$ in another.

– $T \setminus \{f\}$ is a union of two connected graphs such that $\{1,3\}$ lies in one component and $\{2\}$ in another.

- Pairs (S,T) of connected subgraphs such that S contains e and f and T contains neither e nor f and such that e is not a cut edge in S.

Remark 6.4.3. The set of connected graphs on $[n]$ contains as a subset graphs formed from

- Taking a connected graph on $[n-1]$ and connecting the vertex n to one or more of the vertices in $[n-1]$: Note that these are graphs G such that $G \setminus \{n\}$ is connected.
- Taking a connected graph on $[n-1] \setminus \{i\}$, for some $i \in [n-1]$ and connecting the vertex n to one or more vertices in $[n-1] \setminus \{i\}$ as well as to $\{i\}$. These graphs are disconnected when we remove the vertex n and are thus distinct from the ones above.

Thus, $C_n \geq (2^{n-1}-1)\, C_{n-1} + (n-1)C_{n-2}(2^{n-2}-1)$. We rewrite this as

$$\frac{C_n}{C_{n-1}} \geq 2^{n-1} - 1 + (n-1)(2^{n-2}+1)\frac{C_{n-2}}{C_{n-1}}. \tag{6.27}$$

Using the facts that $C_{n-1} \leq G_{n-1}, C_{n-2} \geq \dfrac{G_{n-2}}{2}$, the first by simple inclusion and the second from our induction, we see that

$$\frac{C_n}{C_{n-1}} \geq 2^{n-1} - 1 + (n-1)(2^{n-2}+1)\frac{G_{n-2}}{2G_{n-1}} = 2^{n-1} - 1 + (n-1)\frac{2^{n-2}+1}{2^{n-2}}. \tag{6.28}$$

This last expression is at least 2^{n-1} for $n \geq 3$.

Another Proof. We know that there is an injection from $G_n/C_n \to C_n$ via complement map. Any disconnected graph has connected complement. Moreover, $|G_n/C_n| + |C_n| = |G_n|$. We also know that there exist graphs $R \in C_n$ such that $R^c \in C_n$, hence $|C_n| \geq |G_n/C_n|$. Hence, $2C_n \geq G_n$ which proves that $C_n \geq 2^{\binom{n}{2}-1}$. ∎

Lemma 6.4.1. *For any $n > 2$, we have that*

$$\#\{S \in \mathcal{C}(G) \mid (1,2) \in S \text{ and is a cut edge in } S\} \leq \frac{C_n}{2}. \tag{6.29}$$

Proof. Denoting the quantity on the left by I, we have that

$$\begin{aligned}
I &= \sum_{k=1}^{n-1} C_k C_{n-k} \binom{n-2}{k-1} \\
&= 2C_{n-1} + 2(n-2)C_2 C_{n-2} + \sum_{k=3}^{n-3} C_k C_{n-k} \binom{n-2}{k-1} \\
&\leq 2C_{n-1} + \sum_{k=3}^{n-3} 2^{\binom{k}{2}+\binom{n-k}{2}+n-2} \\
&\leq 2^{\binom{n-1}{2}} + 2(n-2)2^{\binom{n-2}{2}} + (n-5)2^{\binom{n-1}{2}} \\
&\leq 3\,n\,2^{\binom{n-1}{2}}
\end{aligned}$$

When $n \geq 6$, this is less than $2^{\binom{n}{2}} - 1 \leq C_n$ and this proves the assertion in this case. For smaller n, we can verify this directly. ∎

Lemma 6.4.2. *We have that*

$$C_2([\{1,2\},\{3,4\}]) \geq 2\,C_{n-2} + 8(n-4)C_{n-3}. \tag{6.30}$$

Proof. The first quantity on the right counts graphs with 2 connected components, one being either the single edge $(1,2)$ or $(3,4)$ and the other being an arbitrary connected graph on the remaining $n-2$ vertices. The second quantity counts a graph which is the union of a connected graph on 3 vertices, with these three vertices being one of $\{1,2,x\}$, or $\{3,4,x\}$, where x is any vertex in $[5:n]$ and the other being a graph on the remaining $n-3$ vertices. There are 4 connected graphs on 3 vertices and there are $n-4$ choices for x, giving us the coefficient $8(n-4)$ for C_{n-3}. ∎

6.5. Using bridges

Recall that we need to show the following.

$$\sum_{s+t\leq k} C_s(\bar{e},f)C_t(e,\bar{f}) \geq \sum_{s+t\leq k} C_s(e,f)C_t(\bar{e},\bar{f}). \tag{6.31}$$

Let us use the notation

$$A_k = \{(S,T) \mid S \in \mathcal{C}_s(\bar{e},f),\ T \in \mathcal{C}_t(e,\bar{f}),\ s+t \leq k\}, \tag{6.32}$$

$$B_k = \{(S,T) \mid S \in \mathcal{C}_s(e,f),\ T \in \mathcal{C}_t(\bar{e},\bar{f}),\ s+t \leq k\}. \tag{6.33}$$

Given two configurations S,T such that $S \in \mathcal{C}_s(\bar{e},f)$ and $T \in \mathcal{C}_t(e,\bar{f})$, we consider the new pair, $\bar{S} = S \cup \{e\}, \quad \bar{T} = T \setminus \{e\}$. This is a bijection between the following two sets,

- $A_k \setminus C_k$, where $C_k = \{(S,T) \mid S \in \mathcal{C}_s((e),f),\ T \in \mathcal{C}_t(\langle e\rangle,\bar{f}),\ s+t = k\}$.
- $B_k \setminus D_k$, where $D_k = \{(S,T) \mid S \in \mathcal{C}_s(\langle e\rangle,f),\ T \in \mathcal{C}_t((e),\bar{f}),\ s+t = k\}$.

We thus need to show that

$$\sum_{s+t=k} C_s(\langle e\rangle,f)C_t((e),\bar{f}) \leq \sum_{s+t=k} C_s((e),f)C_t(\langle e\rangle,\bar{f}). \tag{6.34}$$

We observe that the following will yield our desired negative correlation.

$$\sum_{s+t\leq k} C_s(\langle e\rangle,f)C_t((e),\bar{f}) \leq \sum_{s+t\leq k} C_s((e),f)C_t(\langle e\rangle,\bar{f}). \tag{6.35}$$

Let us now switch the f from left to right. We will then need to show that

$$\begin{aligned}
&\sum_{s+t=k} C_s(\langle e\rangle, \langle f\rangle)\,[C_t((e),(f)) - C_t(\langle e, f\rangle)] \leq \\
&\sum_{s+t=k} [C_s((e),\langle f\rangle) + C_s(\langle e, f\rangle)]\;[C_t(\langle e\rangle,(f)) + C_t(\langle e, f\rangle] \\
+&\sum_{s+t=k} C_s((e),(f))\,C_t(\langle e, f\rangle).
\end{aligned}$$

This can be written as

$$\sum_{s+t=k} C_s(\langle e\rangle, \langle f\rangle)\,C_t((e),(f)) \leq \sum_{s+t=k} C_s((e),\langle f\rangle)\,C_t(\langle e\rangle,(f)) + \sum_{s+t=k} C_s(e,f)\,C_t(\langle e, f\rangle).$$

Once again, we see that it is actually enough to show that

$$\sum_{s+t\leq k} C_s(\langle e\rangle, \langle f\rangle)\,C_t((e),(f)) \leq \sum_{s+t\leq k} C_s((e),\langle f\rangle)\,C_t(\langle e\rangle,(f)) + \sum_{s+t\leq k} C_s(e,f)\,C_t(\langle e, f\rangle).$$

We will show the above by showing that for n large enough ($n = 10$ will do) and for every $k \leq n$,

$$\sum_{t\leq k} C_t(\langle e\rangle, \langle f\rangle) \leq \sum_{t\leq k} C_t(\langle e, f\rangle). \tag{6.36}$$

Definition 6.5.1. *The expression $C_s(\langle e\rangle, (f))$ will denote the number of sugraphs with s components containing both e and f and such that removing e (without removing f) increases κ by 1 but removing f (without removing e) does not affect κ.*

Similarly, The expression $C_s(\langle e\rangle, (\bar{f}))$ will denote the number of sugraphs with s components containing e but not f and such that removing e (without adding f) increases κ by 1 but adding f (without removing e) does not affect κ.

Recall that we need to show the following.

$$\sum_{s+t\leq k} C_s(\langle e\rangle, f)C_t((e),\bar{f}) \leq \sum_{s+t\leq k} C_s((e),f)C_t(\langle e\rangle,\bar{f}). \tag{6.37}$$

Let us denote the left and right hand sides by L and R respectively. We have that

$$
\begin{aligned}
R &= \sum_{s+t\leq k} C_s((e),f)C_t(\langle e\rangle,\bar{f}) \\
&= \sum_{s+t\leq k} \left[C_s((e),\langle f\rangle) + C_s((e),(f))\right] \left[C_t(\langle e\rangle,\langle\bar{f}\rangle) + C_t(\langle e\rangle,(\bar{f}))\right].
\end{aligned}
$$

For the left hand side we have that

$$
\begin{aligned}
L &= \sum_{s+t\leq k} C_s(\langle e\rangle,f)\, C_t((e),\bar{f}) \\
&= \sum_{s+t\leq k} \left[C_s(\langle e\rangle,\langle f\rangle) + C_s(\langle e\rangle,(f))\right] \left[C_t((e),\langle\bar{f}\rangle) + C_t((e),\left(\bar{f}\right))\right] \\
&= \sum_{s+t\leq k} \left[C_{s+1}(\langle e\rangle,\langle\bar{f}\rangle) + C_s(\langle e\rangle,\left(\bar{f}\right)) - C_s(\langle e,f\rangle)\right] \\
&\quad \left[C_{t-1}((e),\langle f\rangle) + C_t((e),(f)) - C_t(\langle e,f\rangle)\right] \\
&= \sum_{s+t\leq k} \left[C_s(\langle e\rangle,\langle\bar{f}\rangle) + C_s(\langle e\rangle,\left(\bar{f}\right))\right] \left[C_t((e),\langle f\rangle) + C_t((e),(f))\right] \\
&+ \sum_{s+t=k+1} C_s(\langle e\rangle,\langle\bar{f}\rangle)\, C_t((e),(f)) - \sum_{s+t=k} C_s(\langle e\rangle,(\bar{f}))\, C_t((e),\langle f\rangle) \\
&- \sum_{s+t\leq k} C_s(\langle e,f\rangle) \left[C_t((e),\bar{f}) + C_t(\langle e\rangle,f) - C_t(\langle e,f\rangle\right] \\
&= R + \sum_{s+t=k} C_s(\langle e\rangle,\langle f\rangle)\, C_t((e),(f)) - \sum_{s+t=k} C_s(\langle e\rangle,(f))\, C_t((e),\langle f\rangle) \\
&- \sum_{s+t\leq k} C_s(\langle e,f\rangle) \left[C_t((e),\bar{f}) + C_t(\langle e\rangle,f) + C_t(\langle e,f\rangle\right] \\
&= R + \sum_{s+t=k} C_s(\langle e\rangle,\langle f\rangle)\, C_t((e),(f)) - \sum_{s+t=k} C_s(\langle e\rangle,(f))\, C_t((e),\langle f\rangle) \\
&- \sum_{s+t\leq k} C_s(\langle e,f\rangle) \left[C_t((e),(f)) + C_{t-1}((e),\langle f\rangle) + C_t(\langle e\rangle,f)\right] \\
&\leq R + \sum_{s+t=k} \left[C_s(\langle e\rangle,\langle f\rangle) - \sum_{j=1}^{s} C_j(\langle e,f\rangle)\right] C_t((e),(f)) \\
&\leq R.
\end{aligned}
$$

We will be done if we can prove that for all s,

$$C_s(\langle e\rangle, \langle f\rangle) \leq \sum_{j=1}^{s} C_j(\langle e, f\rangle). \tag{6.38}$$

6.5.1. Inequality for Non-adjacent Edges in K_n

Proposition 6.5.1. Proving the inequality

$$C_s(\langle e\rangle, \langle f\rangle) \leq \sum_{j=1}^{s} C_j(\langle e, f\rangle). \tag{6.39}$$

for $s = 1$ is enough.

Proof. We will prove that $C_s(\langle e\rangle, \langle f\rangle)$ is monotone decreasing. Let us write a recursion on the number of vertices for $C_s(\langle e\rangle, \langle f\rangle)$. First, suppose e and f are non-adjacent, say $e = 12$ and $f = 34$. Let $s \geq 2$ be fixed. We will denote by C^i, the number of connected subgraphs on i vertices, and by C^i_s, the number of subgraphs on i vertices with s connected components. The edges e and f can both be bridges in the following two ways:

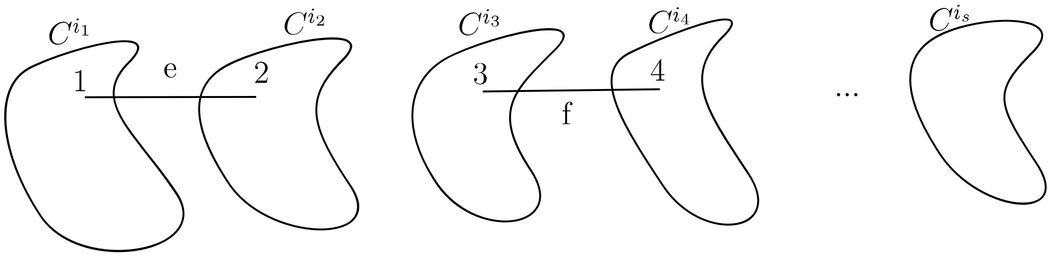


Figure 6.1. Both edges are bridges without a common component.

In the first case, only the first four component is important. We can arrange the edges among themselves in four ways. For the first one, we need to choose $i_1 - 1$ vertices out of $n - 4$. For the second one, we need to choose $i_2 - 1$ vertices out of $n - 4 - (i_1 - 1)$ vertices and so forth. The remaining $s - 4$ components have no condition on them, hence we can write them as $C^{n-i_1-i_2-i_3-i_4}_{s-4}$.

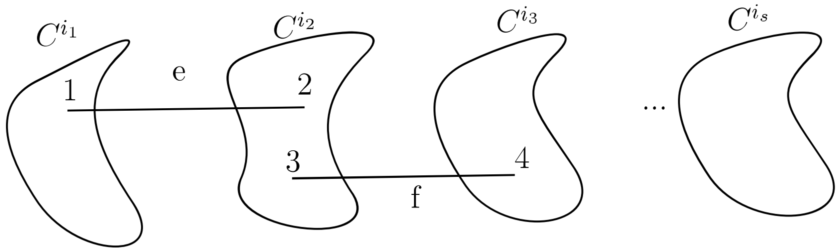


Figure 6.2. Both edges are bridges with a common component.

Hence, overall we have

$$4\sum_{i_1,i_2,i_3,i_4\geq 1}\binom{n-4}{i_1-1}\binom{n-i_1-3}{i_2-1}\binom{n-i_1-i_2-2}{i_3-1}\binom{n-i_1-i_2-i_3-1}{i_4-1}C^{i_1}C^{i_2}C^{i_3}C^{i_4}C_{s-4}^{n-i_1-i_2-i_3-i_4}.$$

The second case's analysis is similar and yields the following:

$$2\sum_{\substack{i_1,i_3\geq 1\\ i_2\geq 2}}\binom{n-4}{i_1-1}\binom{n-i_1-3}{i_2-2}\binom{n-i_1-i_2-1}{i_3-1}C^{i_1}C^{i_2}C^{i_3}C_{s-3}^{n-i_1-i_2-i_3} \tag{6.40}$$

Now, we know that C_s is itself decreasing. Hence, $C_s(\langle e\rangle,\langle f\rangle)$ is also decreasing. ■

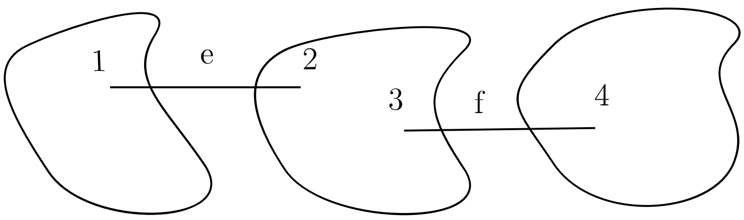


Figure 6.3. Connected graph with two bridges.

Let us look at $C(\langle e\rangle,\langle f\rangle)$. Since both $e=12$ and $f=34$ are bridges, we observe that there should be three connected components which are tied by e and f as in the figure below:

There are 4 different ways the vertices can be put into this pieces. Say, there are i vertices in the component that contains the vertex 1, j vertices that contains the vertices $2,3$ and k vertices that contain the vertex 4. For the first component, out of

$n-4$ vertices we need to choose $i-1$ vertex, for the second component, out of the remaining $n-i-3$ vertices we should choose $j-2$ vertices. Hence, in total we have

$$C(\langle e\rangle, \langle f\rangle) = 4 \sum_{\substack{i+j+k=n \\ i,k\geq 1, j\geq 2}} \binom{n-4}{i-1}\binom{n-i-3}{j-2} C_i C_j C_k. \tag{6.41}$$

Now, a similar argument shows for e and f to be joint bridge we should have two components which are tied by e and f together as in the figure below:

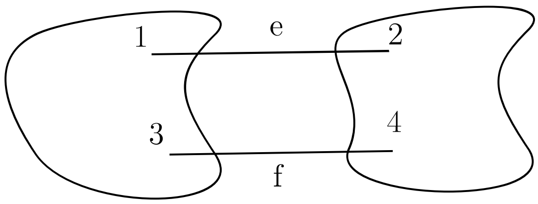


Figure 6.4. Joint bridges.

The vertices can be grouped in 2 possible ways. Say the component containing the vertices $1, 3$ has i vertices and the component containing $2, 4$ has j vertices. Then we choose $i-2$ vertices out of $n-4$ and we are done. Hence we have

$$C(\langle e, f\rangle) = 2 \sum_{\substack{i+j=n \\ i,j\geq 2}} \binom{n-4}{i-2} C_i C_j. \tag{6.42}$$

At this point, actually above, we need to distinguish our graphs since at some points we are using the properties of the complete graph K_n. Let us write a few of the dominant terms of $C(\langle e\rangle, \langle f\rangle)$. We have

$$C(\langle e\rangle, \langle f\rangle) = 4C^{n-2} + (8n-28)C^{n-3} + 20(n-3)(n-4)C^{n-4} + \dots \tag{6.43}$$

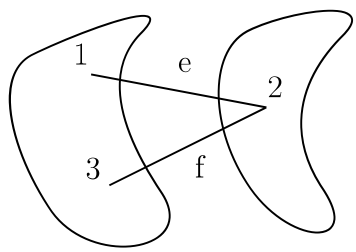
Now, let us write a few of the dominant terms of $C(\langle e, f\rangle)$. We have

$$C(\langle e, f\rangle) = 4C^{n-2} + (16n-32)C^{n-3} + 76(n-4)(n-5)C^{n-4} + \dots \tag{6.44}$$

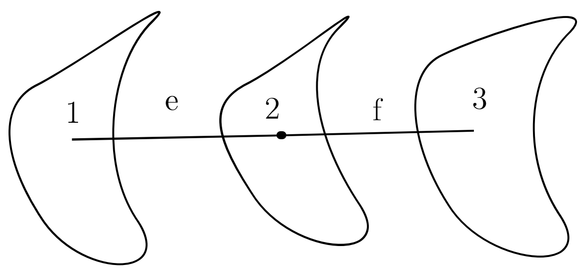
Hence, we observe that the result is correct in the case of $s=1$ for big enough n in K_n.

6.5.2. Inequality for Adjacent Edges in K_n

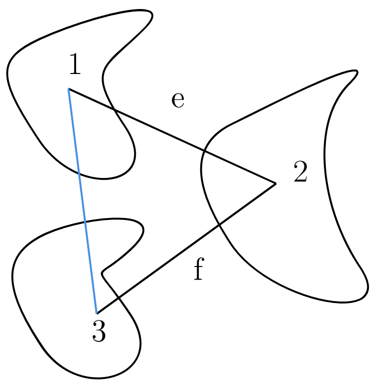
The reasoning that proving the inequality for $s = 1$ holds in this case as well. Say $e = 12$ and $f = 23$ are the adjacent edges. On one hand we have the following $C(\langle e, f\rangle)$ which can only happen in the following way:



On the other hand, we have $C(\langle e\rangle, \langle f\rangle)$ which can only happen in the following way:



In the case of the complete graph, note that we have the edge 13. Let A be a sample subgraph from $C(\langle e\rangle, \langle f\rangle)$, and note that $A \cup 13 \in C(\langle e, f\rangle)$ as shown in the following figure:



Hence we have an injection from $C(\langle e\rangle, \langle f\rangle)$ into $C(\langle e, f\rangle)$. This proves the inequality in the case of adjacent edges in K_n.

6.6. Connection probabilities

We give another perspective for the negative dependence of random cluster measures. Let $G = (V, E)$ be a graph and let u, v be vertices and let e be an edge in G. We will use the notation $\mathbb{P}_G[u \sim v]$ to denote the probability that there is an open path from u to v in a random subset drawn with respect to the random cluster measure with parameters $\{\lambda_f\}_{f\in E}$ and q.

Proposition 6.6.1. *For any graph G and vertices u, v such that the edge $f = uv$ is in G, we have that*

$$\mathbb{P}_G[f] = \frac{\dfrac{\lambda_f}{q} - \lambda_f\left(\dfrac{1-q}{q}\right)\mathbb{P}_{G\setminus\{f\}}[u \sim v]}{1 + \dfrac{\lambda_f}{q} - \lambda_f\left(\dfrac{1-q}{q}\right)\mathbb{P}_{G\setminus\{f\}}[u \sim v]}. \tag{6.45}$$

Let e be another edge in G. We have that

$$\mathbb{P}_G[f \mid e] = \frac{\dfrac{\lambda_f}{q} - \lambda_f\left(\dfrac{1-q}{q}\right)\mathbb{P}_{G\setminus\{f\}}[u \sim v \mid e]}{1 + \dfrac{\lambda_f}{q} - \lambda_f\left(\dfrac{1-q}{q}\right)\mathbb{P}_{G\setminus\{f\}}[u \sim v \mid e]}. \tag{6.46}$$

We also have that

$$\mathbb{P}_G[f] = \frac{\lambda_f}{1+\lambda_f}\left[\frac{\mathbb{P}_G[u \not\sim v]}{q} + \mathbb{P}_G[u \sim v]\right], \tag{6.47}$$

as well as

$$\mathbb{P}_G[f \mid e] = \frac{\lambda_f}{1+\lambda_f}\left[\frac{\mathbb{P}_G[u \not\sim v \mid e]}{q} + \mathbb{P}_G[u \sim v \mid e]\right]. \tag{6.48}$$

Proof. For the first part, we note that

$$\begin{aligned}
Z_G &= [u \sim v]_{G\setminus\{f\}}\,(1+\lambda_f) + [u \not\sim v]_{G\setminus\{f\}}\left(1+\frac{\lambda_f}{q}\right)\\
&= [u \sim v]_{G\setminus\{f\}}\,(1+\lambda_f) + \left(Z_{G\setminus\{f\}} - [u \sim v]_{G\setminus\{f\}}\right)\left(1+\frac{\lambda_f}{q}\right)\\
&= Z_{G\setminus\{f\}}\left(1+\frac{\lambda_f}{q}\right) - [u \sim v]_{G\setminus\{f\}}\,\lambda_f\left(\frac{1-q}{q}\right).
\end{aligned}$$

Dividing out by $Z_{G\setminus\{f\}}$, we see that

$$\frac{Z_G}{Z_{G\setminus\{f\}}} = \left(1+\frac{\lambda_f}{q}\right) - P_{G\setminus\{f\}}[u \sim v]\,\lambda_f\left(\frac{1-q}{q}\right). \tag{6.49}$$

Next, we note that

$$\begin{aligned}
[f]_G &= [u \sim v]_{G\setminus\{f\}}\,\lambda_f + [u \not\sim v]_{G\setminus\{f\}}\,\frac{\lambda_f}{q}\\
&= [u \sim v]_{G\setminus\{f\}}\,\lambda_f + \left(Z_{G\setminus\{f\}} - [u \sim v]_{G\setminus\{f\}}\right)\frac{\lambda_f}{q}\\
&= Z_{G\setminus\{f\}}\,\frac{\lambda_f}{q} - [u \sim v]_{G\setminus\{f\}}\,\lambda_f\left(\frac{1-q}{q}\right).
\end{aligned}$$

Again, dividing out by $Z_{G\setminus\{f\}}$, we see that

$$\frac{[f]_G}{Z_{G\setminus\{f\}}} = \frac{\lambda_f}{q} - P_{G\setminus\{f\}}[u \sim v]\,\lambda_f\left(\frac{1-q}{q}\right). \tag{6.50}$$

We rewrite this as

$$\begin{aligned}\frac{\lambda_f}{q} - &\mathbb{P}_{G\setminus\{f\}}[u \sim v]\,\lambda_f \left(\frac{1-q}{q}\right) = \frac{[f]_G}{Z_G}\frac{Z_G}{Z_{G\setminus\{f\}}} \\ &= \mathbb{P}_G[f]\left[\left(1+\frac{\lambda_f}{q}\right) - P_{G\setminus\{f\}}[u \sim v]\,\lambda_f\left(\frac{1-q}{q}\right)\right].\end{aligned}$$

Simplifying, this yields

$$\mathbb{P}_G[f] = \frac{\frac{\lambda_f}{q} - \lambda_f\left(\frac{1-q}{q}\right)\mathbb{P}_{G\setminus\{f\}}[u \sim v]}{1 + \frac{\lambda_f}{q} - \lambda_f\left(\frac{1-q}{q}\right)\mathbb{P}_{G\setminus\{f\}}[u \sim v]}. \tag{6.51}$$

The second part follows from applying the previous result to the graph $G/\{e\}$, the contraction of G with respect to the edge e. ■

The conjectured edge negative correlation in the random cluster model can be translated to a statement about connection probabilities.

Theorem 6.6.1. *Let G be a graph and let $f = uv$ and e be edges in G and let $0 < q < 1$ Then the following are equivalent.*

- *The vertices e, f are negatively correlated with respect to the random cluster measure $\mu_G^{q,\{\lambda_f\}_{f\in E}}$, i.e. $\mathbb{P}[f \mid e] \leq \mathbb{P}[f]$.*
- *The events $\{u \sim v\}$ and $\{e\}$ are positively correlated in $G \setminus \{f\}$, i.e. $\mathbb{P}_{G\setminus\{f\}}[u \sim v \mid e] \geq \mathbb{P}_{G\setminus\{f\}}[u \sim v]$.*
- *The events $\{u \sim v\}$ and $\{e\}$ are positively correlated in G.*

Proof. The equivalence of the first two follows from the fact that the function

$$x \longrightarrow \frac{\frac{\lambda_f}{q} - \lambda_f\left(\frac{1-q}{q}\right)x}{1 + \frac{\lambda_f}{q} - \lambda_f\left(\frac{1-q}{q}\right)x}, \tag{6.52}$$

is monotone decreasing for $q < 1$.

As for the equivalence of the first and the third, this follows from the monotonicity of

$$\frac{\lambda_f}{1+\lambda_f}\left[\frac{1-x}{q}+x\right]. \tag{6.53}$$

■

We note that for $q > 1$, we have the implication that edgewise positive correlation is equivalent to positive correlation in the second statement above and both statements are known to be true thanks to an application of the FKG inequality.

Conjecture 6.6.1 (Another Formulation)**.** *For every graph and every q, the probability with respect to the random cluster measure that there is an open path between two vertices increases when we add an edge to the graph.*

7. CONCLUSION

We end the thesis with some final remarks and open conjectures and paths for subsequent works that we plan to pursue.

The main motivation of this thesis was to investigate the notions of negative dependence and prove negative dependence for the random cluster measure. We have studied several notions of negative dependence with the strongest being strongly Rayleigh measures and the weakest being pairwise negative correlation or negative edge dependence. The implications among these classes are intricate and we have proved some of them and ended up with the general picture that is depicted in Figure 4.1. There is however still not a complete picture.

Among the problems we were unable to solve is the followings:

(i) Is there a notion of negative dependence not strong as strongly Rayleigh measures but is preserved under symmetric exclusion processes?
(ii) Is there an easy way to verify NA?

In the case of random cluster measures, much more is unknown. We have proved p-NC in the case of the complete graph, but the main conjecture is still open and we state it again.

Conjecture 7.0.1 (Main Problem)**.** *Let $G = (V, E)$ be a graph and $\mu_{p,q}$ be the random cluster measure on G. Prove that the edges are negatively correlated, i.e. $\mu_{p,q}(e, f) \leq \mu_{p,q}(e)\mu_{p,q}(f)$, whenever $q < 1$.*

We think that it would be remarkable to prove the conjecture above even in the case of lattice graphs $\mathbb{Z}^n$. This case seems to be the prominent one for physicists as well. Since our proof worked with the symmetry of the complete graphs, it does not seem to work with $\mathbb{Z}^n$, some of the inequalities involved in the proof need the extra edges for injection. Another idea to work with in the future is to use the limits of the random cluster measure to give proofs for USF and UCS measures.

We also have very basic correlation inequalities which are still open for the random cluster measure.

Conjecture 7.0.2. *Let $G = (V, E)$ be a graph and $\mu_{p,q}$ be the random cluster measure on G. The random variables X_e, where $e \in E$ and κ, which is the number of connected components is negatively correlated.*

There are many more open and important conjectures in this area. In the future, I plan on following the results of this thesis and study Random cluster measure and the general conjecture.

REFERENCES

1. Borcea J., P. Branden.,T. M. Liggett, "Negative Dependence and The Geometry of Polynomials", *Journal of American Mathematical Society*, Vol. 22, pp.521-567, July 2009.

2. Pemantle R., "Towards a Theory of Negative Dependence", *Journal of Mathematical Physics*, Vol. 41, 1371-1390, pp. 1-31, April 2000.

3. Dubhashi D., Ranjan D., "Balls and Bins: A Study in Negative Dependence", *BRICS Report Series*, RS-96-25, 0909-0878, Vol. 25, pp.1-16, July 1996.

4. Grimmett G., "The Random-Cluster Model",*Graduate Texts in Mathematics*, University of Cambridge, Wilberforce Road, Springer-Verlag, pp.1-100, July 2006.

5. Bollabas B.. "Modern Graph Theory", *Graduate Texts in Mathematics*, Springer, pp.1-20, February 2000.

6. Oxley J. G., "Matroid Theory", *Oxford Graduate Texts in Mathematics*, Vol. 21, 2nd Ed., Oxford University Press, Oxford, pp.1-30, August 2011.

7. Erickson A., "Negative Correlation Properties for Matroids", pp.1-13, November 2008.

8. Huh J., P. Branden, "Lorentzian Polynomials", **https://arxiv.org/abs/1902.03719**, accessed in July 2021.

9. Welsh D., C. Semple, "Negative Correlation in Graphs and Matroids", *Combinatorics, Probability and Computing*, pp.423-435, May 2008.

10. Preston C. J., "A Generalization of The FKG Inequalities", *Communications in Mathematical Physics*, Vol. 36, pp.1-10, December 1974.

11. Wagner D., "Negatively Correlated Matroids and Mason's Conjecture", *Annals of Combinatorics*, Vol. 12, pp.211-239, July 2008.

12. Feder T., M. Mihail, "Balanced Matroids", *STOC'92: Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing*, pp.26-38, January 1992.

13. Liggett T. M., "Negative Correlations and Particle Systems", *Markov Processes and Related Fields*, Vol. 8, Issue 4, pp.547-564, October 2002.

14. Wagner D., "Negatively Correlated Random Variables and Mason's Conjecture for Independent Sets in Matroids", *Annals of Combinatorics*, Vol. 12, No.2, pp.212-239, February 2008.

15. Pemantle R., "Hyperbolicity and Stable Polynomials in Combinatorics and Probability", *Current Developments in Mathematics*, 57-123, INT. Press., Somerville, pp.1-16, October 2011.

16. Lyons R., "Determinantal Probability Measures", *Publications in Mathematics Institute Hautes Etudes Science*, Vol. 98, 167-212, pp.1-41, June 2003.

17. D. Dubhashi, V. Priebe, D. Ranjan, "Negative Dependence Through the FKG Inequality", Technical Report RS-96-27, *BRICS Report Series*, Basic Research In Computer Science, Arhus, Denmark, pp.1-18, July 1996.

18. Choe Y., J. Oxley, A. Sokal, D. G. Wagner, "Homogeneous Multivariate Polynomials with the Half-plane Property", *Advances in Applied Mathematics*, Vol. 32, pp.88–187, December 2004.

19. Liggett T. M., "Negative Correlations and Particle Systems", *Markov Processes and Related Fields*, Vol. 8, 547–564, pp.1-15, October 2002.

20. Huh J., B. Schroter, B. Wang, "Correlation Bounds for Fields and Matroids", **https://arxiv.org/abs/1806.02675**, accessed in June 2018.

21. Sokal A., "The Multivariate Tutte Polynomial (alias Potts model) for Graphs and Matroids", *Surveys in Combinatorics*, London Mathematical Society Lecture Notes, Series 327, Cambridge University Press, Cambridge, pp.1-34, March 2005.

22. Lyons R., Y. Peres, "Probability on Networks and Trees", *Cambridge Series in Statistical and Probabilistic Mathematics*, One Liberty Plaza, New York NY 10006, USA, pp.1-100, November 2016.

23. Haggstrom O., "Random-cluster Measures and Uniform Spanning Trees", *Stochastic Probability and Applications*, Vol. 59, pp.1-37, October 1995

24. Berg J. V. D., U. Fiebig, "On a Combinatorial Conjecture Concerning Disjoint Occurrences of Events", *Annals of Probability*, Vol. 15, pp.1-17, January 1987.

25. Seymour P., D. Welsh, "Combinatorial Applications of an Inequality from Statistical Mechanics", *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 77, pp.1-27, May 1975.