UNIVERSITY OF CALIFORNIA SAN DIEGO

Quasi-random Boolean Functions

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Mathematics

by

Nicholas Sieger

Committee in charge:

Professor Fan Chung, Chair
Professor Samuel Buss, Co-Chair
Professor Ramamohan Paturi
Professor Lutz Warnke

2024

The Dissertation of Nicholas Sieger is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

There are many people to whom I owe thanks for their help in my Ph.D. studies. First and foremost, my advisors Fan Chung and Sam Buss, for their guidance, advice, for so many excellent research problems. I thank in particular for Math 155 and fun of teaching programming. I also thank Sam Spiro, Jason O'Neill, Gwen McKinley, Ruth Luo, and all the members of ABACUS. I learned a lot of combinatorics from you all. I also thank Mareike Dressler and Gwen McKinley for all their guidance in teaching; it was a joy serving as your TA. I also need to thank Eva Loeser, Vaki Nikitopoulos, Itai Maimon, and Gongping Niu for so many good conversations throughout my time in San Diego.

There are many people outside of campus who made my life much brighter. First and foremost, the CAC crew: Zell & Jadey, Jason, Joe (the thrower of cabbages), Joe (prince charmin), Cristian, Amanda, Michael, Will & Carly, Chris & Citta, Dana, Adam, Adriana, Lorena, Lara, Aaron, Alex, Alden, Anna, Bri, Kelly, Leo, Matt, Oscar, Rushi, and Ramon. I'm going to miss you all. And of course, the Mary Star Book Club: Joline, Michael & Mandy, Caylie, John, Scott & Caroline, Brittany & Danny, Chase & Candice, Valerie, and Mark. With them are all the folks at Mary Star: Fr. Pat, Fr. Mark, Fr. Clement, Deacon Jimmy, Fran, Maria, Alejandro, Shannon, Lupe, and Kevin. Each of you made my Sundays the highlight of each week. Last but certainly not least, I thank Michael and Maddy for sticking with me through the many twists and turns of my time in San Diego.

Finally, I need to thank Dr. Lisa and Fr. Paul for listening to all of my crazy questions and giving them serious answers. I thank my parents for their continual support and willingness to drive approximately 2000 miles so many times.

Chapters 2 and 3, in full, are a reprint of the material as it appears in Quasi-random Boolean Functions, Fan Chung and Nicholas Sieger, which is in review at the Electronic Journal of Combinatorics. The dissertation author was the primary investigator and author of this paper.

As all things should be, may this work be *ad maiorem Dei gloriam.*

ABSTRACT OF THE DISSERTATION

Quasi-random Boolean Functions

by

Nicholas Sieger

Doctor of Philosophy in Mathematics

University of California San Diego, 2024

Professor Fan Chung, Chair
Professor Samuel Buss, Co-Chair

We examine a hierarchy of equivalence classes of local quasi-random properties of Boolean Functions. In particular, we prove an equivalence between a number of properties including balanced influences, spectral discrepancy, local strong regularity, subgraph counts in a Cayley graph associated to a Boolean function, and equidistribution of additive derivatives among many others. In addition, we construct families of quasi-random Boolean functions which exhibit the properties of our equivalence theorem and separate the levels of our hierarchy. Furthermore, we relate our properties to several extant notions of pseudo-randomness for Boolean functions.

Chapter 1

Introduction

> You cannot ask a government agency to build a new 100 billion dollar supercomputer and then say 'just wire it up randomly, it should work.'
>
> (Steven J. Young)

Randomness is simultaneously powerful, infuriating, and mysterious. Random constructions can produce elusive Ramsey graphs [32], graphs of high girth and high chromatic number [5], and efficient codes [5]. Yet in all these examples, randomness conceals as much as it reveals. None of the construction mentioned here are *explicit*; they merely prove that such an object exists without providing a single clear example. Indeed, many of the problems mentioned above only received explicit solutions after decades of work (see for instance [4]), and some still await constructions which will clarify their nature.

The random constructions themselves further exhibit often miraculous properties. The same random graphs which Erdős used to prove lower bounds in Ramsey theory also possess the following properties [42] amongst many more:

- contain every small graph as a subgraph,
- expand, in the sense that the neighborhood of a set of vertices is larger than that set,
- contain Hamiltonian cycles in abundance,
- contain every possible tree on n vertices,

- cannot be disconnected without removing a large number of edges.

Indeed, there are decades of study into the properties of a random graph, and the expansion properties of random graphs are of signal importance in computation. One natural question arises: how do these properties of random graphs relate to one another? Can we say that certain graph properties must appear together?

Out of the the desire to find explicit constructions meeting random bounds and the desire to explore the properties of random graphs arose a new idea: quasi-randomness. The story begins with quasi-random graphs.

1.1 Quasi-random Graphs

Quasi-randomness first appeared in embryo in the works of Erdős, S os, Graham, Thomassen, and Wilson [31, 33, 82, 86] which explored several random graph properties which could also be applied to deterministic graphs. The form of quasi-random theorems was fully fleshed out in the seminal paper of Chung, Graham, and Wilson [20] which also introduced the name "quasi-random." As quasi-random graphs are the prototype for other theories of quasi-randomness, it is worthwhile to present their result in more detail. Consider the following graph properties for a fixed $\varepsilon > 0$:

- Every fixed graph H appears as a subgraph $(1 \pm \varepsilon)2^{-e(H)}n^{v}(H)$ times.
- The the largest eigenvalue of the adjacency matrix of G is at least $\frac{n}{2}$ and the second-largest eigenvalue of the adjacency matrix of G is at most εn.
- Between every two sets of vertices $S \subseteq V(G)$ and $T \subseteq V(G)$, there are $(1 \pm \varepsilon)\frac{|S||T|}{2}$ edges.
- The 4-cycle appears as an induced subgraph $(1 \pm \varepsilon)2^{-6}n^{4}$ times.

Each of the these properties makes sense both for a fixed graph and for a random graph. Furthermore, random graphs satisfy all of these properties in expectation. *A priori*, there is little reason to believe that these properties have anything to do with each other. After all, the first property cosniders subgraph counts of *every finite graph* whereas the last property considers a single graph, the 4-cycle. Furthermore, the second property can be determined in polynomial time, whereas there is no direct means of determining the third property.

The surprising result of Chung, Graham, and Wilson is that these four properties (and several others) are *equivalent* to each other for a fixed graph. The notion of equivalence here is subtle and important. Given such different conditions as in the above list, it is unreasonable to demand that the properties have on-the-nose identical error bounds ε, but the error should not grow too rapidly either. Chung, Graham, and Wilson balance these concerns as follows. Two graph properties $P_1(\varepsilon)$ and $P_2(\varepsilon)$ are *equivalent* if for every $\varepsilon > 0$, there is a $\delta > 0$ such that $P_1(\delta) \implies P_2(\varepsilon)$ and vice-versa, and the δ does not depend on the size of the graph in either direction. The theorem's goal is to organize the many properties of the list into equivalence classes and thereby give a precise definition of a "random-like" graph. As a consequence, Chung, Graham, and Wilson term a graph satisfying any one of the above properties *quasi-random*.

Of course, a beautiful equivalence of various graph properties is much less interesting in the absence of concrete examples of graphs possessing the properties of the theorem. So Chung, Graham, and Wilson proceed to give two primary examples. The first is the example of Paley graphs, whose vertices are the integers mod p for some prime p and whose edges are defined by quadratic residues. The second is graphs defined by parity conditions; the vertex set is the set of all subsets of $[n]$ of odd size and there is an edge between two sets if their intersection also has odd size. Both graphs can be readily shown to possess one of the properties of the quasi-random equivalence theorem, and therefore they possess all of the properties.

In summary, a theory of quasi-randomness has three main components:

- A list of properties, each of which is possessed by a random object of the relevant type.
- An equivalence theorem proving that the properties in the list are equivalent to each other up to small losses in the parameters.

- A construction of a object which possesses one (and therefore all) of the properties on in the list.

The chief value of the theorem arises from the diversity of the list of properties in the first step. The list ought to contain several properties which are not obviously equivalent (like the 4-cycle count property and the finite graph counting property for quasi-random graphs), and some properties which are easily computable and some which are not (such as the eigenvalue property and the expansion property for quasi-random graphs).

1.1.1 Extensions of quasi-randomness

There are also several related fields which are deeply intertwined with quasi-randomness. The first is that of regularity lemmas. One of the properties in the Chung-Graham-Wilson Theorem is ε-regularity, which states that any two subsets X and Y of size at least $\varepsilon|G|$ have nearly $\frac{1}{2}|X||Y|$ edges between them. Szemeredi's regularity lemma [79] asserts that every sufficiently large dense graph has an ε-regular partition, i.e. a partition of the vertex set such that the induced bipartite graph between almost all pairs of parts is ε-regular. Put simply, Szemeredi's regularity lemmas that large dense graphs are quasi-random. The power of the regularity lemma is hard to overstate, as it allows for a deterministic dense graph to be treated with probabilistic techniques.

The theory of graph limits expands upon the ideas found in the regularity lemma, and has a similar relationship to quasi-randomness. The theory of graph limits models graph via symmetric functions on the unit square, i.e., maps $\rho : [0,1]^2 \to [0,1]$ such that $\rho(x,y) = \rho(y,x)$. These maps are termed *graphons* to distinguish them from discrete graphs. There are several different ways to define the topology on graphons, in fact, each of the quasi-random properties in the Chung-Graham-Wilson Theorem provides a notion of topology. In this context of graphons, the Chung-Graham-Wilson Theorem states that all of the topologies induced by these graph properties are equivalent. Such topological equivalences expedite many proofs, for instance the construction of graph limits is easily defined using subgraph counts of arbitrary graphs, while proving convergence of a sequence of graphs to a particular graph limit is more readily handled by the C_4-norm induced by the count of 4-cycles. We refer the reader to the text of Lovasz [66] for additional details.

Quasi-randomness, regularity lemmas, and graph limits continued a fruitful dialogue in the years following the Chung-Graham-Wilson Theorem. Chung and Graham studied quasi-random tournaments in [16], and a limit notion of tournaments received further study in [10, 51, 83]. Chung and Graham studied subsets of the integers mod n in [19], which tied into the study of arithmetic progressions in works of Roth [73] and Szemeredi [79]. The limit objects for subsets of the integers were fully fleshed out in the works of Green and Tao [81]. As the Chung-Graham-Wilson Theorem and Szemeredi's regularity lemma applied only to dense graphs, there was much interest in finding analogues of these theories for sparse graphs. Chung and Graham gave a theory of quasi-randomness for sparse graphs in [21], Lovasz gave two theories of graph limits for sparse graphs in [66], and a regularity lemma for sparse graphs appeared in [60]. Later on, Cooper studied quasi-random permutations in [28], followed by the work of Chan, Kral, Noel, Pehvoa, Sharifzadeh and Volec [13] and Kral and Pikhurko [63] who expanded the list of quasi-random properties. The limit theory of permutations appears in several works, notably Hoppen, Kohayakawa, Moreia, Rath, and Menezes [54]. The work of Garbe, Hancock, Hladky, and Sharifzadeh [30, 43] built a quasi-random theorem for Latin squares and the corresponding limit theory. Griffiths extended the Chung-Graham-Wilson Theorem to oriented graphs in [49], mirroring the directed graph limits in [66]. Gowers considered quasi-random groups [46].

There are, however, two objects whose theory of quasi-randomness requires special attention: hypergraphs and Boolean functions.

1.1.2 The Trouble of Quasi-random Hypergraphs

Chung and Graham considered quasi-random hypergraphs throughout their work on quasi-randomness for other objects, and found that hypergraphs were a different beast entirely. There are several key steps in the Chung-Graham-Wilson Theorem which fail for hypergraphs.

The first issue is *expander mixing lemma*, which appears as one of the main steps in the proof of the Chung-Graham-Wilson Theorem. The expander mixing lemma requires a notion of eigenvalues of a graph, and oo even make such a definition, a graph needs to be represented by a matrix. For ordinary graphs, the adjacency matrix (and several related matrices) provide several combinatorially meaningful eigenvalues. For hypergraphs, where the edges contain 3 or more elements, the natural analogue of the adjacency matrix is a k-dimensional array of numbers called a hypermatrix or tensor. There are two equivalent definitions of matrix eigenvalues needed to prove the expander mixing lemma, the Rayleigh quotient and the defintion via eigenvectors, and these definitions are no longer equivalent for hypermatrices. Indeed, a quasi-random theorem involving eigenvalues of hypermatrices remains an open problem [27].

In light of these difficulties, the simpler case of linear hypergraphs attracted some attention. A hypergraph is *linear* if each pair of phyeredge intersects in no more than one vertex. As Friedman and Widgerson [40] showed, there is a sensible analogue of the eigenvalue properties for linear k-uniform hypergraphs. Lenz and Mubayi [64, 65] extended this theme and gave a hierarchy of different eigenvalue properties of linear k-uniform hypergraphs. The works of Rödl, Schacht, and Kohawakaya [61] added a corresponding regularity lemma.

Quasi-randomness for linear hypergraphs contains properties closely related to quasi-randomness for graphs, which indicates that the true difficulty of hypergraphs arises from edges intersecting in more than one vertex. Indeed, such intersections lead to the second and more profound difficulty in defining quasi-randomness for hypergraphs. As Chung and Graham discovered in [15], a hypergraph may be ε-regular (in the sense of graphs) yet fail to be quasi-random. Their construction works as follows. They begin with a randomly chosen $k-1$-uniform hypergraph G, and then form a k-uniform hypergraph H by adding a k-edge $\{v_1, \dots, v_k\}$ if the induced subgraph $G[\{v_1, \dots, v_k\}]$ contains an even number of hyperedges in G. As shown in [15], H fails to contain any copies of the specific hypergraph consisting of all but one k-subset of $k+1$ vertices. As containing every small hypergraph is a signal property of quasi-randomness, H cannot be quasi-random. Nonetheless, H is ε-regular in the sense of graphs [15].

The true analogue of ε-regularity for hypergraphs requires the consideration of the *l-shadow* of the hypergraph, i.e., the l-uniform hypergraph consisting of all l-sets which are contained in at least one edge in the original hypergraph [45]. For a 3-uniform hypergraph to be quasi-random, it must not only have the "correct" number of hyperedges between any two sets, but nearly half of all triangles in its 2-shadow must be contained in a 3-edge. As one considers k-uniform hypergraphs, the situation only becomes worse, as a k-uniform hypergraph must be quasi-random with respect to every l-shadow for $l \leq k$.

Despite the absence of eigenvalue properties and the difficulties of ε-regularity, Chung and Graham's study produced a theory of quasi-randomness for k-uniform hypergraphs over a series of works [14, 15, 17, 18, 22, 24]. One of the more surprising equivalences from quasi-random graphs remained even in the setting of hypergraphs. For graphs, the count of copies of C_4 determine the counts of arbitrary graphs. For k-graph, Chung and Graham show in [15] that if the number of copies of each k-graph on $2k$ vertices is close to its value on a random k-graph, then the same holds for any k-graph on t vertices for $t \geq 2k$. When specialized to $k = 2$, this definition is nearly identical to the count of 4-cycles in the Chung-Graham-Wilson Theorem.

The corresponding regularity lemmas and several new quasi-random properties for hypergraphs appeared in works of Chung, Frankl, Rödl, Schacht, Kohawakaya, and Nagle [36, 62, 69, 76]. Additionally, Towsner [84] gathered many of the quasi-random properties together into a larger theory and began a hypergraph limit theory; this was followed by [26] who gave more combinatorial proofs of the same result. More relevant for our work, Castro-Silva explored the connections between hypergraphs and additive combinatorics in [11, 12].

1.2 Boolean Functions

Perhaps the most subtle of the objects thus far is that Boolean functions. At the highest level, a *Boolean Functions* is a map from binary strings of length n to $\{True, False\}$. Boolean functions can encode a wide variety of mathematical and computational objects, such as decision problems, error-correcting codes, communication and cryptographic protocols, among others. Given so many applications, Boolean functions are extremely well-studied in coding theory, cryptography, and computational complexity among many other areas of computer science and data science. For each application, many researchers have developed tools and perspectives unique to each area to study these Boolean functions and have isolated key properties of Boolean functions, for instance the sensitivity of the function to changes in each coordinate, the size of its Fourier coefficients, or the distance of its support viewed as a binary code.

Through a variety of transformations, the theories of quasi-randomness for graphs and hypergraphs apply to a theory of quasi-random Boolean functions. Indeed, in their on quasi-random subsets of $\mathbb{Z}/n\mathbb{Z}$ [19], Chung and Graham include a property that a graph associated to a subset with quasi-random in the sense of the Chung-Graham-Wilson Theorem, and a similar graph construction produces new properties for Boolean functions. In Chung and Tetali's work [23], the theory of quasi-random hypergraphs was directly extended to Boolean functions. This would be extended further in the works of Gowers on Szemeredi's Theorem [48].

Beyond works directly associated with quasi-randomness, the idea of random-like Boolean functions appears in both cryptography and theoretical computer science. In these fields, a random-like Boolean function has been formalized in several ways, and there is a terminological conflict that needs to be addressed.

1.2.1 Pseudo-random vs Quasi-random

A close cousin of quasi-randomness the study of pseudo-randomness, which frequently arises in theoretical computer science [85], additive combinatorics [47], and number theory [80]. While the names are quite similar, there are some major conceptual differences between the two.

The typical context for pseudo-randomness is a "structure vs pseudo-randomness" argument, of which Roth's theorem [73] is an excellent example. Roth studied the size of subsets S of $[n]$ which contained no three-term arithmetic progression, i.e., a sequence of three elements of the form $a, a+d, a+2d$ for fixed $a, d \in \mathbb{N}$. One readily verify that a uniformly random subset of $[n]$

contains many three term arithmetic progressions; but the desired result considers arbitrary sets. Roth noticed that the number of three term arithmetic progressions can be captured by a Fourier analytic formula, and this observation gave rise to a clever proof. Given a subset S of the set $[n]$ with no three-term arithmetic progressions, Roth divides the argument into two cases. First, he considerd the “pseudo-random” case, where the Fourier coefficients of the (indicator function of the) set S are small. As a set with small Fourier coefficients must have many three-term arithmetic progressions, Roth concluded that S must have a large Fourier coefficient. This is the “structured” case. Here, he used the large Fourier coefficient to construct a large arithmetic progression on which the set S has increased density. By reducing to this long arithmetic progression, the argument can be repeated until a three-term progression is found. As a consequence any set with constant density must contain a three-term progression, and so any set with no such progression must have density which tends to 0 as n becomes large.

Comparing Roth’s theorem with quasi-randomness, one can see several key differences between the two concepts. pseudo-random properties serve as a dividing line, separating arbitrary objects into “structured” and “random-like” sets. Thus pseudo-random properties come paired with some sort of structure. In Roth’s Theorem, the structure was a large arithmetic progression, and the proof used different techniques based on whether the pseudo-random property of small Fourier coefficients held. Indeed, a pseudo-random property is useless without a structured counterpart which must appear whenever the property fails. pseudo-random properties stand apart from each other, and are rarely placed into equivalence theorems as in quasi-randomness. quasi-random properties, by contrast, need not have a structured counterpart, and are introduced with a view towards an equivalence theorem involving many other properties.

It is certainly possible that a quasi-random property may have a structured object which appears whenever the property fails, for instance a graph which is not ε-regular has a pair of sets with very few or very many edges between them. Furthermore, there are many quasi-random properties referenced as pseudo-random and vice-versa; indeed, the pseudo-random property of small Fourier coefficients appearing in Roth’s Theorem is one of the main quasi-random properties in Chung and Graham's work on quasi-random subsets of $\mathbb{Z}/n\mathbb{Z}$. Nonetheless, the typical focus of pseudo-random properties is frequently limited to the context of a particular argument. It is of great interest to try to extend pseudo-random properties to the broader context of a quasi-random theorem, and we small consider this problem in Chapter 3.

1.2.2 Influences of Boolean Functions

One notion of pseudorandom property for Boolean functions, the notion of influences, continually appears in a variety of applications. Given a Boolean function $f : \mathbb{F}_2^n \to \{1, -1\}$, the *ith-influence* of a coordinate i if the fraction of inputs $x \in \mathbb{F}_2^n$ where flipping the value of x_i changes the output of the function. If the function does not depend at all on coordinate i, then the ith influence is 0, and if $f(x) = x_i$, then the ith influence is 1. For other functions, the influence lies between these two extremes.

Influences first appeared in the context of genetics and social choice theory in the work of Penrose [72] on majority votes, Banzhaf on weighted voting rules [6], and Coleman [25] on coalitions. Here a Boolean function is seen as a voting rule and influence tracks the power of an individual voter of the outcome of the election. In this context, the maximum influence is of great interest, and we note the Tribes function of Ben-or and Linial [7] which is gives a voting rule where all voters have small influence and the KKL Theorem [34, 57] which shows that the tribes example is best possible.

In a totally different direction, Freidgut [38] showed that influences provide a characterization of the presence of *sharp thresholds* in the study of properties of random graphs. Ever since, Friedgut's Theorem has been an essential tool in the study of graph thresholds, see [1, 37, 41].

However, a random Boolean function does not have small influences, i.e., small influences do not give a truly pseudorandom or quasirandom property. For the restricted class of montone Boolean functions, small influences do give a speudorandom property, and monotone Boolean functions are a primary focus in the context of social choice and thresholds. For general functions, however, a somewhat different notion of influence arose in the context of property testing.

The basic question of property testing is simple: given a Boolean function, does it possess a given property P? Property testing first appeared in the work of Rubinfeld and Sudan [75], followed by works of Goldrecih, Goldwasser, and Ron [44] and Friedl and Sudan [39]. There are variety of properties for which efficient tests are known, and many of these properties appear in quasirandom theorems. For instance, the gowers norms which arise in quasi-random hypergraphs provide a test for $\mathbb{F}_2$-degree [77].

One of the central problems in property testing is the testing of *dictators*, i.e., testing if a Boolean function depends on a single input. These tests are at the heart of the famed PCP theorem [8, 9, 29]. As shown by [55], testing whether a Boolean function is a dictator can be done via a form of noise stability, known as "small stable influences." Random Boolean functions do possess small stable influences and these works implicitly define a pseudorandom property which is opposed to functions which depend only on a few coordinates. Efficient tests for dictators were developed in a series of works due to Bellare, Goldreich, Sudan, Parnas, Ron, and Samorodnitsky [71] culminating in the works of Hastad [55, 56].

1.3 Local Quasi-randomness

The goal of this thesis is to organize a range of properties of Boolean functions into a hierarchy of equivalence classes in the same style as the quasi-random graphs and hypergraphs in [14, 20, 24]. Our properties are local in nature, forming a hierarchy depending on a local parameter d. For instance, one of our main properties, the Balanced Influences Property, concerns the influences of all vectors of Hamming weight at most d. Another property considers subgraph counts of 4-cycles in the associated Cayley graph with location restrictions depending on d. There is a second parameter in the descriptions of our properties, an error bound ε which controls our notion of equivalence between properties. For two properties $P_1(d,\varepsilon)$ and $P_2(d,\varepsilon)$, we say that P_1 implies P_2 if for every $\varepsilon > 0$ there is a $\delta > 0$ such that $P_1(d,\delta)$ implies $P_2(d,\varepsilon)$ where δ only depends on d and ε. If P_1 and P_2 imply each other, then we say that P_1 and P_2 are equivalent.

In our main theorem, we show how a number of known analytic properties of Boolean functions, such as the *k*-th order strict avalanche criterion, restrictions of the function having small Fourier coefficients, and discrepancy of the Fourier coefficients, can be either strengthened or weakened so as to become equivalent to one another. Motivated by the enumeration of "sub-patterns" within a larger object, we further show that several combinatorial properties of graphs built from our Boolean function are equivalent with these analytic properties. These combinatorial properties include local 4-cycle counts, a local sameness property, counts of rainbow embeddings of graphs and a co-degree condition on a Cayley graph defined from the Boolean function. We summarize the main theorem and its proof in Figure 1.1. Finally, we give an explicit construction of a family of Boolean functions which exhibits the properties in our main theorem. As it turns out, our construction depends crucially on the existence of good binary codes. As will be indicated throughout the thesis, the properties that we discuss here are satisfied by a random Boolean function, and therefore are called *quasi-random* in the spirit of [20] .

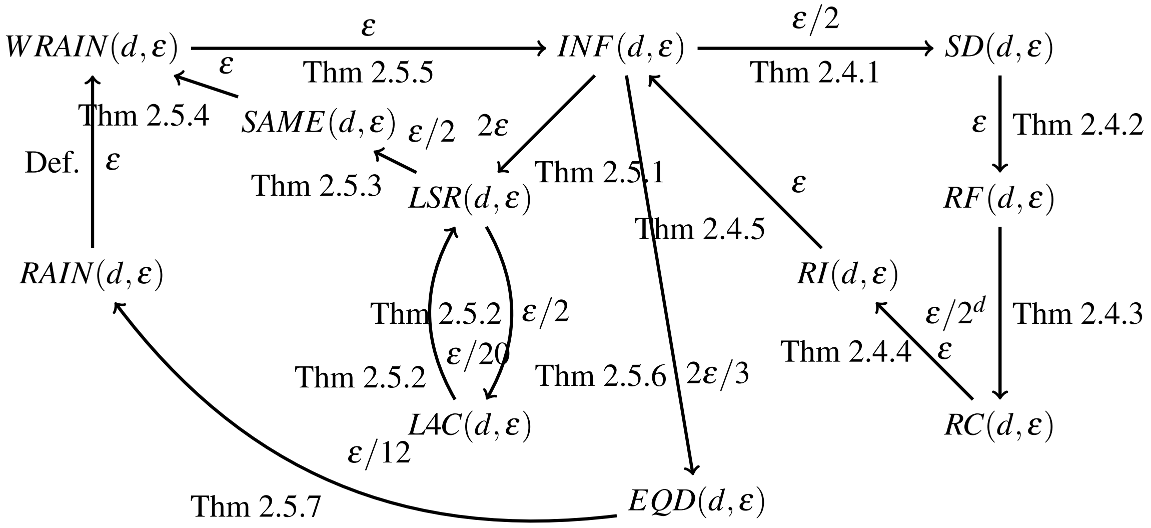

Figure 1.1. The implications in Theorem 2.3.2. Each edge gives the loss in ε and the reference to the theorem in which the implication is shown.

All of the theories mentioned in the previous sections center on properties of a global nature, for instance the total number of copies of a fixed subgraph as considered in the first property of Chung, Graham, and Wilson's work [20]. By contrast, our properties here are local in nature. We shall later prove that our local theory of quasi-random Boolean functions is distinct from each of these global theories, stronger than several of the global theories, and incomparable with the others. We illustrate the relationships in Figure 1.2

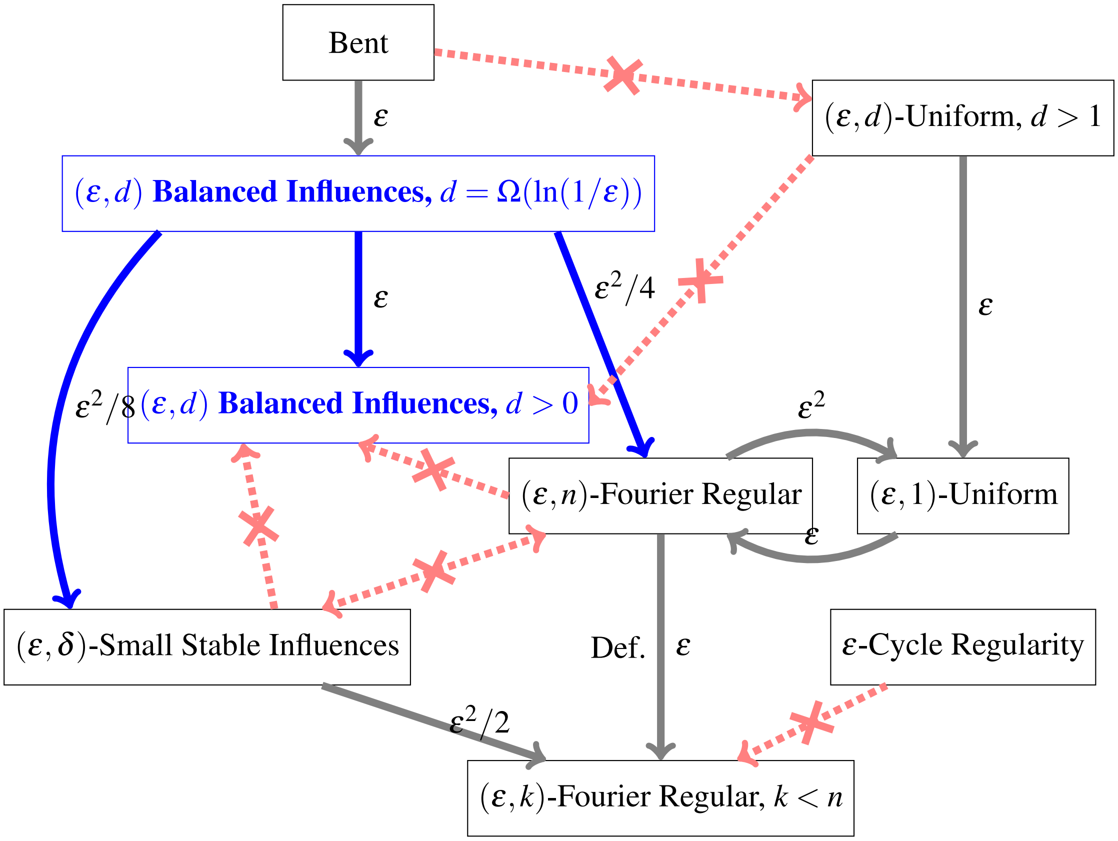
Bent
ε
(ε, d)-Uniform, $d > 1$
(ε, d) **Balanced Influences,** $d = \Omega(\ln(1/\varepsilon))$
ε
$\varepsilon^2/4$
ε
$\varepsilon^2/8$
(ε, d) **Balanced Influences,** $d > 0$
ε^2
(ε, n)-Fourier Regular
$(\varepsilon, 1)$-Uniform
ε
(ε, δ)-Small Stable Influences
Def.
ε
ε-Cycle Regularity
$\varepsilon^2/2$
(ε, k)-Fourier Regular, $k < n$

Figure 1.2. The relationships between different theories of quasi-randomness. Each box is a distinct theory of quasi-randomness. Each arrow is a strict implication. Beside each arrow we give the loss function. The results of this thesis are in bold blue text and blue arrows. Non-implications are red dotted lines with an X in the middle.

Our thesis is organized as follows. In Chapter 2, we present our local quasi-randomness theorem for Boolean functions. In Sections 2.1 and 2.2 we give the preliminaries needed to state our quasi-random properties. In Section 2.3, we state the main equivalence theorem of eleven quasi-random properties. Due to the large umber of properties and their rich connections, the proofs of the implications are divided into two sections. Section 2.4 considers influences of Boolean functions and several analytic properties. Section 2.5 considers a codegree property and

4-cycle counts amongst other combinatorial properties. We then give an explicit construction of quasi-random functions possessing the properties in our main theorem in Section 2.6. These functions also separate the levels of the hierarchy of our equivalence classes.

In Chapter 3 we discuss several extant theories of quasi-random Boolean functions. Each is presented via specific pseudo-random property, and we then give an theorem relating our quasi-random properties to the extant theory. The main results of Chapters 2 and 3 are summarized in the flowcharts found in figures Figure 1.1 and Figure 1.2.

Chapter 2

quasi-random Influences

In this chapter, we state and prove our main quasi-randomness theorem for Boolean functions. At a high level, our theorem connects a variety of analytic and combinatorial properties of Boolean functions, and these properties require an extensive list of definitions. In Sections 2.1 and 2.2 we state all of the relevant definitions. We then state our main theorem in Section 2.3

2.1 Analytic Properties of Boolean Functions

At the highest level, a Boolean function is a map from the set of binary strings of length n to True and False. There are several different mathematical interpretations of binary strings and True and False, and we will use the following:

Definition 2.1.1. A *Boolean function* to be a map $f : \mathbb{F}_2^n \to \{1,-1\}$. We will view -1 as True and 1 as False.

For a proposition $P(x)$, let $[P(x)] := \begin{cases} 1 & P(x) \\ 0 & \neg P(x) \end{cases}$ denote the *indicator function* for $P(x)$. We will write 0 for the zero vector in $\mathbb{F}_2^n$ throughout, and write $\mathbf{1} \in \mathbb{F}_2^n$ for the all-ones vector. If μ is a distribution on a set Ω, and $P(x)$ is a proposition on the variable $x \in \Omega$, then $\mathbb{P}_{x\sim\mu}\left[P(x)\right]$ will denote the probability distribution that $P(x)$ holds when x is drawn from the distribution μ. Whenever we write the expectation or probability over a set, such as $\mathbb{E}_{x\in\mathbb{F}_2^n}$, the expectation or probability is taken with respect to the uniform distribution. We refer the reader to O'Donnell's book [70] for any undefined terminology.

In the following subsections, Sections 2.1.1 to 2.1.7 to §2.1.3, we state the definitions concerning various aspects of Boolean functions that will be used to define our various properties of Boolean functions.

2.1.1 Fourier Analysis

We can equip the space of all maps $g : \mathbb{F}_2^n \to \mathbb{R}$ with the following inner product:

$$\langle f, g \rangle := \mathbb{E}_{x \in \mathbb{F}_2^n} f(x)g(x) = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} f(x)g(x).$$

Definition 2.1.2. For each $\gamma \in \mathbb{F}_2^n$, the *Fourier character* $\chi_\gamma : \mathbb{F}_2^n \to \{1, -1\}$ is $\chi_\gamma(x) := (-1)^{\gamma \cdot x}$ where $\gamma \cdot x := \sum_{i=1}^n \gamma_i x_i$ is the usual dot product.

The Fourier characters form an orthonormal basis for the space of all maps $g : \mathbb{F}_2^n \to \mathbb{R}$ with the inner product as defined above.

Definition 2.1.3. For $g : \mathbb{F}_2^n \to \mathbb{R}$ the *Fourier coefficient* with respect to $\gamma \in \mathbb{F}_2^n$, denoted $\widehat{g}(\gamma)$, is $\langle g, \chi_\gamma \rangle = \mathbb{E}_{x \in \mathbb{F}_2^n} g(x)\chi_\gamma(x)$.

Notice that $\widehat{f}(0) = \langle f, 1 \rangle = \mathbb{E}_{x \in \mathbb{F}_2^n} f(x)$ is simply the average value of f.

The *density* of a Boolean function $f : \mathbb{F}_2^n \to \{1, -1\}$, denoted by $\mathrm{dens}(f)$, is $\frac{|f^{-1}(\{-1\})|}{2^n}$, which we note is precisely $\frac{1 - \widehat{f}(0)}{2}$.

Definition 2.1.4. The *convolution* of two functions g and $h : \mathbb{F}_2^n \to \mathbb{R}$ is

$$(g * h)(x) := \mathbb{E}_{y \in \mathbb{F}_2^n} g(x + y)h(y).$$

that $\widehat{g * h}(\gamma) = \widehat{g}(\gamma)\widehat{h}(\gamma)$.

2.1.2 $\mathbb{F}_2$-Degree

Every boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has an alternate representation as a map $\mathbb{F}_2^n \to \mathbb{F}_2$ defined as follows. $f(x) = \sum_{y\in\mathbb{F}_2^n} f(y)\mathbf{1}_y(x)$ where $\mathbf{1}_y(x) = \prod_{i=1}^n (1+y_i+x_i)$. If we expand out each $\mathbf{1}_y(x)$, we get a multilinear polynomial (using the fact that $x^2 = x$ over $\mathbb{F}_2$) $f(x) = \sum_{\vec{y}\in\mathbb{F}_2^n} c_{\vec{y}} x^{\vec{y}}$ where $x^{\vec{y}} := \prod_{i\in\mathrm{supp}(y)} x_i$. As this expansion is unique, we can then define the **degree** of f as $\deg(f) := \max_{\vec{y}\neq\vec{0}} |\vec{y}|$.

A Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ can be equivalently defined as a multilinear polynomial from $\mathbb{F}_2^n \to \mathbb{F}_2$, [70] where $1 \in \mathbb{F}_2$ denotes True and $0 \in \mathbb{F}_2^n$ denotes False. As the multilinear expansion of a Boolean function is unique (see [70]) each Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has a well-defined $\mathbb{F}_2$-*degree*, given by the size of the largest monomial in its multilinear expansion over $\mathbb{F}_2$.

2.1.3 Bent Functions

We consider a specific class of Boolean functions originally defined by Rothaus [74].

Definition 2.1.5. [74] For n even, a Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ is *bent* if for every $\gamma \in \mathbb{F}_2^n$ we have

$$\left|\widehat{f}(\gamma)\right| = 2^{-n/2}.$$

Note that bent functions only exist for n even.

We will use the following property of Bent functions frequently.

Proposition 2.1.6. *[74] If $g : \mathbb{F}_2^n \to \{1,-1\}$ is bent, then $(g * g)(x) = [x = 0]$.*

Definition 2.1.7. The **inner product** function $IP : \mathbb{F}_2^{2m} \to \{1,-1\}$ is defined by

$$IP(z) := (-1)^{\sum_{i=1}^m z_i z_{m+i}}.$$

The inner product function will serve as a useful example throughout this thesis. From the definition, we have the following:

Lemma 2.1.8. *IP has $\mathbb{F}_2$-degree 2.*

For the sake of completeness, we show that IP is in fact a bent function.

Lemma 2.1.9. *[[74]] $IP : \mathbb{F}_2^{2m} \to \{1,-1\}$ is a Bent function.*

Proof. Fix $\gamma \in \mathbb{F}_2^{2m}$, and let $\gamma_1, \gamma_2 \in \mathbb{F}_2^m$ denote the first m bits of γ and the last m bits respectively. For $x \in \mathbb{F}_2^{2m}$, define x_1, x_2 similarly. Then,

$$\begin{aligned}
\widehat{IP}(\gamma) &= \mathbb{E}_{x \in \mathbb{F}_2^{2m}} IP(x)\chi_\gamma(x) \\
&= \mathbb{E}_{x_1 \in \mathbb{F}_2^m} \mathbb{E}_{x_2 \in \mathbb{F}_2^m} (-1)^{x_1 \cdot x_2 + \gamma_1 \cdot x_1 + \gamma_2 \cdot x_2} \\
&= \mathbb{E}_{x_1 \in \mathbb{F}_2^m} (-1)^{\gamma_1 \cdot x_1} \mathbb{E}_{x_2 \in \mathbb{F}_2^m} (-1)^{(x_1 + \gamma_2) \cdot x_2} \\
&= \mathbb{E}_{x_1 \in \mathbb{F}_2^m} (-1)^{\gamma_1 \cdot x_1} [x_1 = \gamma_2] \qquad (2.1) \\
&= (-1)^{\gamma_1 \cdot \gamma_2} 2^{-m}
\end{aligned}$$

where we use the fact that Fourier characters are orthogonal in Equation (2.1)). Thus IP is a bent function. We remark for later use that IP has $\mathbb{F}_2$-degree 2 as it is equal to the degree 2 polynomial $\sum_{i=1}^m z_i z_{m+i}$. □

2.1.4 Hamming Weight

We will also need to track the size of individual vectors in $\mathbb{F}_2^n$.

Definition 2.1.10. For a vector $x \in \mathbb{F}_2^n$, its *Hamming weight*, denoted $|x|$, is the number of nonzero entries in x.

Similarly, the *Hamming distance* between two vectors x and $y \in \mathbb{F}_2^n$ is $|x - y|$. The *Hamming ball* of radius d in $\mathbb{F}_2^n$ and centered at the vector $x \in \mathbb{F}_2^n$, denoted by $B_d(n,x)$, is $\{y \in \mathbb{F}_2^n : |x - y| \leq d\}$.

The following definition will be useful in our combinatorial properties.

Definition 2.1.11. For a subset $S \subseteq \mathbb{F}_2^n$, its *diameter* is $\mathrm{diam}(S) := \max_{x,y \in S} |x - y|$.

2.1.5 The influences of Boolean functions

The notion of “influences” is prominent in both analysis of Boolean functions and cryptography.

Definition 2.1.12. For $\gamma \in \mathbb{F}_2^n$, the γ-*Influence* of f is

$$\mathrm{I}_\gamma[f] := \mathbb{P}_{x\in\mathbb{F}_2^n}\left[f(x) \neq f(x+\gamma)\right].$$

Note that $\mathrm{I}_0[f]$ is always 0. Furthermore, for $\gamma \in \mathbb{F}_2^n$ with $\gamma_i = 1$ and $\gamma_j = 0$ for $j \neq i$, $\mathrm{I}_\gamma[f]$ is precisely the influence of coordinate i as studied extensively in O’Donnell [70]. We note the work of Keevash et al [58] which considers a related generalization of influences in the context of hypercontractivity.

The following property of the γ-influences will be quite useful later.

Lemma 2.1.13. *For any fixed $\gamma \in \mathbb{F}_2^n$, a Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ satisfies*

$$f * f(\gamma) = 1 - 2\mathrm{I}_\gamma[f].$$

Proof. By definition of γ-influence,

$$\begin{aligned}
1 - 2\mathrm{I}_\gamma[f] &= 1 - 2\mathbb{P}[f(x) \neq f(x+\gamma)] \\
&= \mathbb{E}_{x\in\mathbb{F}_2^n}\left(1 - 2[f(x) \neq f(x+\gamma)]\right) \\
&= \mathbb{E}_{x\in\mathbb{F}_2^n} f(x)f(x+\gamma) \qquad (2.2) \\
&= f * f(\gamma)
\end{aligned}$$

where we use the fact that $f(x) \in \{1,-1\}$ in Equation (2.2)). □

2.1.6 The spectral sampling of Boolean functions

Parseval's Theorem states that for $f : \mathbb{F}_2^n \to \{1,-1\}$,

$$\sum_{\gamma \in \mathbb{F}_2^n} \widehat{f}(\gamma)^2 = \mathbb{E}_{x \in \mathbb{F}_2^n}\left[f(x)^2\right] = 1.$$

Thus the Fourier coefficients of f define a probability distribution on $\mathbb{F}_2^n$ as follows:

Definition 2.1.14. For a fixed Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$, the *Spectral Sample* $\mathscr{S}_f$ is the distribution on $\mathbb{F}_2^n$ where

$$\mathbb{P}_{\gamma \sim \mathscr{S}_f}[\gamma = \delta] = \widehat{f}(\delta)^2$$

for each fixed $\delta \in \mathbb{F}_2^n$.

2.1.7 Subcubes and the counting of subcubes

Let $[n]$ denote the set $\{1,\dots,n\}$, and for $S \subseteq [n]$, let $\overline{S}$ denote $[n] \setminus S$. Given a set $S \subseteq [n]$, and two vectors $x \in \mathbb{F}_2^S$, $y \in \mathbb{F}_2^{\overline{S}}$, let $x \underset{S}{\otimes} y$ denote the vector where

$$(x \underset{S}{\otimes} y)_i = \begin{cases} x_i & i \in S \\ y_i & i \in \overline{S} \end{cases}.$$

Definition 2.1.15. The *subcube* defined by a set $S \subseteq [n]$ and a vector $z \in \mathbb{F}_2^{\overline{S}}$ is the set

$$C(S,z) := \{x \underset{S}{\otimes} z : x \in \mathbb{F}_2^S\}.$$

We say that the *dimension* of the subcube $C(S,z)$ is $|S|$. Note that $C([n],\eta)$ where η is the empty string is precisely the hypercube Q_n. In Figure 2.1, we have two examples of subcubes.

We are also concerned about Boolean functions restricted to a subcube:

Definition 2.1.16. The *restriction* of $f : \mathbb{F}_2^n \to \{1,-1\}$ to the subcube $C(S,z)$ is the Boolean function $f|_{S,z} : \mathbb{F}_2^S \to \{1,-1\}$ defined by

$$f|_{S,z}(x) = f(x \underset{S}{\otimes} z)$$

If $S = \emptyset$, then $f|_{S,z}(x)$ is the constant function $f(z)$, and if $S = [n]$, then we recover f itself.

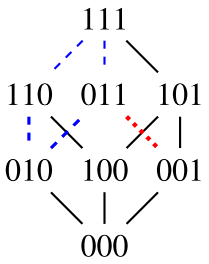

Figure 2.1. The blue dashed lines in the figure indicate the 2-dimensional subcube $C(\{1,3\},1)$, i.e., the set of all length 3 binary strings with a 1 in the second coordinate. The red dotted line indicates the 1-dimensional subcube $C(\{2\},01)$.

We will need the following result, translated into our notation.

Lemma 2.1.17. *[[70] Proposition 3.21] If $C(S,z)$ is a fixed subcube and $\gamma \in \mathbb{F}_2^S$, then*

$$\widehat{f|_{S,z}}(\gamma) = \sum_{\delta \in \mathbb{F}_2^{\overline{S}}} \widehat{f}\left(\gamma \underset{S}{\otimes} \delta\right) \chi_\delta(z).$$

2.2 Combinatorial aspects of Boolean functions

In the following subsections, Sections 2.2.1 to 2.2.4, we give several useful combinatorial interpretations of Boolean functions that are of interest in their own right. For two sets A, B, let $A \hookrightarrow B$ denote the set of all injective functions from A to B. Let $A \sqcup B$ denote the disjoint union of the sets A and B.

2.2.1 Cayley Graphs

Definition 2.2.1. Given a group G and a set $S \subseteq G$, *Cayley graph* of G generated by S is the graph with vertex set G and $a, b \in G$ adjacent if $ab^{-1} \in S$.

If $s \in S$ implies that $s^{-1} \in S$, then the Cayley graph with generating set S is an undirected graph.

Of the many ways to define a graph from a Boolean function, the following first comes to mind.

Definition 2.2.2. For a Boolean function $f : \mathbb{F}_2^n \to \{1, -1\}$, the *Cayley graph* of f, denoted $Cay(f)$, is the Cayley graph on $\mathbb{F}_2^n$ whose generating set is $f^{-1}(\{-1\})$.

As every element of $\mathbb{F}_2^n$ is its own additive inverse, it follows that $Cay(f)$ is an undirected graph.

The Cayley graph $Cay(f)$ appears in several of the extant theories of quasi-randomness we shall consider in Chapter 3, for instance [11].

2.2.2 Graph Homomorphisms

We will be interested in subgraph counts in $\mathrm{Cay}(f)$ which can be defined by graph homomorphisms.

Definition 2.2.3. A *graph homomorphism* from $H = (U, F)$ to $G = (V, E)$ is a map $\phi : V(H) \to V(G)$ such that

$$(u, v) \in F \implies (\phi(u), \phi(v)) \in E.$$

We will typically assume our graph homomorphisms are injective, and we denote the *normalized* number of injective graph homomorphisms via the following:

$$\hom(H, G) = \mathbb{E}_{\phi : V(H) \hookrightarrow V(G)} \prod_{(u,v) \in E(H)} [(\phi(u), \phi(v)) \in E(G)].$$

We will also make use of graph homomorphisms which may not be injective, and we denote the *normalized* number of such graph homomorphisms via the following:

$$\overline{\hom}(H,G) = \mathbb{E}_{\phi:V(H)\to V(G)} \prod_{(u,v)\in E(H)} [(\phi(u),\phi(v)) \in E(G)].$$

Note that the normalization factor in $\hom(H,G)$ is $\dfrac{1}{|V(G)|(|V(G)|-1)\dots(|V(G)|-|V(H)|+1)}$ whereas in $\overline{\hom}(H,G)$ the normalization factor is $\dfrac{1}{|V(G)|^{|V(H)|}}$.

2.2.3 Colored Multigraphs

The following definition is inspired by the work of Aharoni et al on rainbow extremal problems [2].

Definition 2.2.4. An *edge-colored multigraph M* with color set K is a multigraph with an edge-coloring using colors in K such that multiple edges between any two vertices u and v cannot have the same color.

We will typically think of the edges of an edge-colored multigraph as a subset of $V \times V \times K$.

Definition 2.2.5. For fixed $f : \mathbb{F}_2^n \to \{1,-1\}$ and $k \geq 1$, the *rainbow Hamming graph $RHG(k,f)$* is the colored multigraph on the vertex set $B_k(n,0)$ with color set $K = \mathbb{F}_2^n$ and edge set defined as

$$\{(u,v,x) \in V \times V \times K : f(u+x) = f(v+x)\}.$$

An explicit example of a rainbow Hamming graph is given in Figure 2.2.

2.2.4 Rainbow embeddings

We consider graph homomorphisms into a colored multigraph which agree with the coloring.

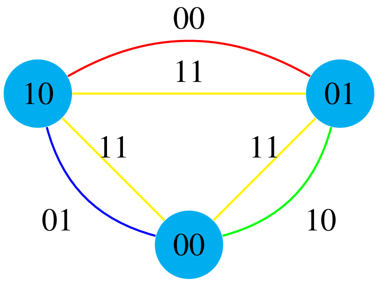

Figure 2.2. The rainbow Hamming graph $RHG(1,h)$ of the function $h(z) = (-1)^{1-z_1z_2}$ where $z_1, z_2 \in \mathbb{F}_2$. Each edge is labeled by the string in $\mathbb{F}_2^2$ which defines its color. Note that h encodes the NAND function.

Definition 2.2.6. Let M be a colored multigraph with color set K and let G be a fixed (simple) graph. A *rainbow embedding* of G into M is an injective coloring $\chi : E(G) \hookrightarrow K$ and an injective map $\phi : V(G) \hookrightarrow V(M)$ such that

$$(u,v) \in E(G) \implies (\phi(u), \phi(v), \chi((u,v))) \in E(M).$$

These embeddings are also considered in the work of Alon and Marshall [3].

For a fixed graph G, a fixed colored multigraph M with color set K, let

$$\mathrm{ch}(G,M) := \mathbb{E}_{\phi:V(G)\hookrightarrow V(M)} \mathbb{E}_{\chi:E(G)\hookrightarrow K} \prod_{(u,v)\in E(G)} [(\phi(u), \phi(v), \chi((u,v))) \in E(M)]$$

be the normalized count of rainbow embeddings of G into M. If we additionally fix the injection $\phi : V(G) \hookrightarrow V(M)$, let

$$\mathrm{ch}_\phi(G,M) := \mathbb{E}_{\chi:E(G)\hookrightarrow K} \prod_{(u,v)\in E(G)} [(\phi(u), \phi(v), \chi((u,v))) \in E(M)]$$

be the normalized count of rainbow embeddings with a fixed map ϕ.

2.3 Quasi-random Properties and the Equivalence Theorem

In this section, we describe a number of quasi-random properties of Boolean functions. Each property involves two parameters, denoted by d and ε, where ε indicates the error bound and d is often related to the rank or dimension of patterns or objects in the property. We will typically think of ε and d as constants, but our results sometimes hold when ε and d depend on n. The proofs of the equivalence of these properties will be given in sections §2.4 and §2.5.

We begin with a basic property regarding the density of our Boolean functions. A random Boolean function will be -1 and 1 about equally often, i.e., it has density close to $\frac{1}{2}$. We say that a Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ is **ε-balanced** if $\left|\text{dens}(f) - \frac{1}{2}\right| < \varepsilon$. Since the density $\text{dens}(f)$ is equal to $\frac{1-\widehat{f}(0)}{2}$, any ε-balanced function f satisfies $\left|\widehat{f}(0)\right| < 2\varepsilon$.

For the rest of the thesis, we consider the following weaker density property:

Property $\mathbf{P_0}$. *A Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ is **weakly balanced** if the density of f is at least $\frac{3}{10}$ and at most $\frac{7}{10}$.*

Equivalently, a weakly balanced function has $\left|\widehat{f}(0)\right| < \frac{2}{5}$. We remark that all of the quasi-random properties below will require a weakly balanced Boolean function. The assumption of weak balance is necessary, since there are Boolean functions which are not weakly balanced and satisfy some but not all of our quasi-random properties, as shown in Theorem 2.3.1. The specific value of $\frac{2}{5}$ is chosen for the sake of exposition and can be replaced by any constant strictly greater than $\frac{1}{2\sqrt{2}}$ and strictly less than $\frac{1}{2}$.

Our first property focuses on the directional influences defined in Section 2.1.5. If $f : \mathbb{F}_2^n \to \{1,-1\}$ is chosen uniformly at random, we expect that the γ-influence (see Definition Definition 2.1.12) should be close to $\frac{1}{2}$. Our first quasi-random property formalizes this notion for weakly balanced Boolean functions.

Property $\mathbf{P_1}$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the **Balanced Influences Property** $INF(d,\varepsilon)$ if the γ-Influence of f is close to $\frac{1}{2}$ for every nonzero γ in the Hamming* *ball of radius d centered at* 0*, i.e.,*

$$\left|\mathrm{I}_\gamma[f]-\frac{1}{2}\right|<\varepsilon$$

for every γ such that $1\leq|\gamma|\leq d$*.*

We remark that $INF(d,0)$ is also known as the dth-Order Avalanche Criterion as studied in cryptography [35].

It is natural to assume that the Balanced Influences Property implies weak balance, but the implication does not hold for $d=1$ and $d=2$ as we shall prove in Section 2.6.

Theorem 2.3.1. *For $d\in\mathbb{N}$ the following holds:*

- *For $d\geq 3$, if $f:\mathbb{F}_2^n\to\{1,-1\}$ satisfies the Balanced Influences Property $INF(d,\frac{2}{25}-2^{-d-1})$, then f is weakly balanced.*
- *For $d\leq 2$, there exists a Boolean function such that* $\mathrm{dens}(f)=\frac{1}{4}$ *but*

  $$\mathrm{I}_\gamma[f]=\frac{1}{2}$$

  for any $\gamma\in\mathbb{F}_2^n$ such that $0<|\gamma|\leq d$*.*

For $f:\mathbb{F}_2^n\to\{1,-1\}$ drawn uniformly from all Boolean functions, the expected spectral sample (see Definition Definition 2.1.14) is $\frac{1}{2^n}$ on each vector in $\mathbb{F}_2^n$. Rather than considering each vector in $\mathbb{F}_2^n$ individually, we will consider subcubes (see Definition Definition 2.1.15). In particular, the total weight of the uniform distribution on a subcube of dimension k is exactly 2^{k-n}. Our next quasi-random property states that the spectral sample $\mathscr{S}_f$ assigns similar weight to each subcube as the uniform distribution does.

Property $\mathbf{P_2}$. *A weakly balanced Boolean function $f:\mathbb{F}_2^n\to\{1,-1\}$ has the **Spectral Discrepancy Property** $SD(d,\varepsilon)$ if the spectral sample of f has total weight close to 2^{l-n} on every* *subcube of dimension l where $l \geq n-d$, i.e.,*

$$\left|\mathbb{P}_{z\sim\mathscr{S}_f}[z \in H] - 2^{\dim(H)-n}\right| < \varepsilon$$

for every subcube H of dimension at least $n-d$.

Next, we have a counting property on subcubes via the notion of restricted functions (see Definition Definition 2.1.16). As $f|_{S,z}$ is a map $\mathbb{F}^d \to \{1,-1\}$ for $|S| = d$, we can consider its Fourier coefficients. The next quasi-random property states that these Fourier coefficients are quite small on average.

Property P$_3$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Restriction Fourier Property*** *$RF(d,\varepsilon)$ if the average restriction of f is nearly a bent function on any subcube of dimension at most d, i.e.,*

$$\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\left[\widehat{f|_{S,z}}(\gamma)^2\right] - 2^{-\dim(C(S,z))}\right| < \varepsilon$$

for every subcube $C(S,z)$ of dimension at most d and every $\gamma \in \mathbb{F}_2^S$.

The next property states that we can control certain patterns in the restrictions of f.

Property P$_4$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Restriction Convolution Property*** *$RC(d,\varepsilon)$ if the average self-convolution of restrictions of f to subcubes of dimension at most d is close to the indicator function of the 0 vector, i.e.,*

$$\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}(f|_{S,z} * f|_{S,z})(x) - [x=0]\right| < \varepsilon$$

for every set $S \subseteq [n]$ of size at most d and every $x \in \mathbb{F}_2^S$.

Convolutions are closely related to influences, so we have an additional influences property pertaining to an average restricted function:

Property P$_5$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Restriction Influences Property*** *$RI(d,\varepsilon)$ if the γ-Influences of the average restriction to subcubes of dimension at most d are close to $\frac{1}{2}$, i.e.,*

$$\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\mathrm{I}_\gamma[f|_{S,z}] - \frac{1}{2}\right| < \varepsilon$$

for every set $S \subseteq [n]$ of size at most d and every nonzero $\gamma \in \mathbb{F}_2^S$.

The *directional derivative* of a Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ in the direction γ is $\Delta_\gamma f(x) = f(x+\gamma)f(x)$. Our next property states that pairs of multiplicative directional derivatives are equidistributed in the following sense:

Property P$_6$. *A Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Equidistributed Derivatives Property*** *$EQD(d,\varepsilon)$ if every pair of sufficiently close directional derivatives take each possible pair of values equally often, i.e., for every choice of $c_0, c_1 \in \{1,-1\}$ and for every $a,b \in \mathbb{F}_2^n$ such that $|a| \le d$, $|b| \le d$, and $0 < |a-b| \le d$, we have*

$$\left|\mathbb{E}_{x\in\mathbb{F}_2^n}[\Delta_a f(x) = c_0][\Delta_b f(x) = c_1] - \frac{1}{4}\right| < \varepsilon.$$

Next we consider some combinatorial properties. Our first few combinatorial properties focus on the Cayley graph of a Boolean function $\mathrm{Cay}(f)$ (see Definition Definition 2.2.2). For $v \in V(G)$, let $N_G(v)$ denote the neighborhood of v in G.

Property P$_7$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Local Strong Regularity Property*** *$LSR(d,\varepsilon)$ if any two vertices $u,v \in \mathbb{F}_2^n$ at Hamming distance most d have approximately the same number of common neighbors in the Cayley graph of f, i.e.,*

$$\left|\frac{\left|N_{\mathrm{Cay}(f)}(x) \cap N_{\mathrm{Cay}(f)}(y)\right|}{2^n} - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right| < \varepsilon$$

for every pair of vertices x,y in $\mathrm{Cay}(f)$ *such that* $0 < |x-y| \le d$.

We remark that the Local Strong Regularity Property is analogous to the co-degree property in Chung, Graham, and Wilson's work on quasi-random graphs [20]. Note that the term $\frac{\widehat{f}(0)}{2}$ allows for a range of edge densities in $\mathrm{Cay}(f)$, and in particular $\mathrm{Cay}(f)$ and $\mathrm{Cay}(-f)$ do not have the same edge density in general. Our next property states that nonetheless $\mathrm{Cay}(f)$ and $\mathrm{Cay}(-f)$ are somewhat interchangeable.

Property $\mathbf{P_8}$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Local Sameness Property*** *$SAME(d,\varepsilon)$ if for any two vertices $u,v \in \mathbb{F}_2^n$ at Hamming distance most d, approximately half of all other vertices are a common neighbor of u and v either in the Cayley graph of f or the Cayley graph of $-f$, i.e.,*

$$\left| \frac{\left|N_{\mathrm{Cay}(f)}(x) \cap N_{\mathrm{Cay}(f)}(y)\right| + \left|N_{\mathrm{Cay}(-f)}(x) \cap N_{\mathrm{Cay}(-f)}(y)\right|}{2^n} - \frac{1}{2} \right| < \varepsilon$$

for every pair of vertices $x \neq y$ in $\mathrm{Cay}(f)$ *such that* $0 < |x-y| \leq d$.

Given the power of subgraph counts of 4-cycles in Chung, Graham, and Wilson's work on quasi-random graphs [20], we have an additional property regarding these 4-cycle counts. We say that a map $\psi : A \to \mathbb{F}_2^n$ has *diameter* at most k if $\left|\psi(u) - \psi(v)\right| \leq k$ for every $u,v \in A$.

Property $\mathbf{P_9}$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Local 4-Cycle Property*** *$L4C(d,\varepsilon)$ if in the Cayley graph* $\mathrm{Cay}(f)$*, for any two vertices $u,v \in \mathbb{F}_2^n$ at Hamming distance at most d, the expected number of 4-cycles with u and v as antipodal points is close to the expected value, i.e.,*

$$\left| \overline{\mathrm{hom}}_\phi(C_4, \mathrm{Cay}(f)) - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)^2 \right| < \varepsilon$$

via the definition of $\overline{\mathrm{hom}}_\phi(H,G)$ for any injection $\phi : L(C_4) \hookrightarrow \mathbb{F}_2^n$ of diameter at most d.

Here we assume the function f is weakly balanced, an assumption which will be crucial in the proof of Theorem 2.5.2. We remark that Chung, Graham, and Wilson give a global count of C_4's, whereas we give a stronger condition which controls local appearances of C_4. This intuitive connection will be expanded upon in Chapter 3 where we compare our properties to a number of previously known pseudo-random properties appearing in prior works.

Our final few combinatorial properties build on the Local 4-Cycle Property by giving strong control over local subgraph counts of an arbitrary graph. In particular, given a graph H, we fix the location of our desired subgraph in a larger graph derived from our Boolean function, and then ask for how many ways we can extend our choice of location to a homomorphism of H. To keep track of the extra information needed, these properties have a number of additional technical requirements and definitions.

We consider a count of rainbow embeddings in the rainbow Hamming graph (see sections §2.2.3 and §2.2.4).

Property $\mathbf{P_{10}}$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Rainbow Embeddings Property*** *$RAIN(d,\varepsilon)$ if for every fixed simple graph G with at most $\max\{\sqrt{\varepsilon}2^{n/2-1},1\}$ edges and every choice of injection ϕ from G to the rainbow Hamming graph of f, there are close to a $2^{-|E(G)|}$-fraction of colorings of G which become rainbow embeddings of G under ϕ, i.e., the Rainbow Embeddings Property holds if*

$$\left|\mathrm{ch}_\phi(G,RHG(d,f)) - 2^{-|E(G)|}\right| < \varepsilon$$

for every fixed graph G such that $|E(G)| \leq \max\{\sqrt{\varepsilon}2^{n/2-1},1\}$, and every $\phi : V(G) \hookrightarrow V(RHG(d,f))$ of diameter at most d.

Property $\mathbf{P_{11}}$. *A weakly balanced Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has the* ***Weak Rainbow Embeddings Property*** *$WRAIN(d,\varepsilon)$ if for every choice of injection ϕ from K_2 to the rainbow Hamming graph of f, there are close to a $\frac{1}{2}$-fraction of colorings of K_2 which become rainbow* *embeddings of G under ϕ, i.e., the Rainbow Embeddings Property holds if*

$$\left|\mathrm{ch}_{\phi}(K_2, RHG(d,f)) - \frac{1}{2}\right| < \varepsilon$$

for every $\phi : V(K_2) \hookrightarrow V(RHG(d,f))$ of diameter at most d.

A map $\Delta : \mathbb{N} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is a *loss function* if for each $d \in \mathbb{N}$, $\varepsilon < \varepsilon'$ implies that $\Delta(d,\varepsilon) \leq \Delta(d,\varepsilon')$. For properties $P(d,\varepsilon)$ and $Q(d,\varepsilon)$ and a loss function Δ, we say P Δ*-implies* Q, denoted $P(d,\varepsilon) \overset{\Delta}{\Longrightarrow} Q(d,\varepsilon)$, if for every $d \geq 1$, every $\varepsilon > 0$, every $n > 0$ and every Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$

$$P(d, \Delta(d,\varepsilon)) \implies Q(d,\varepsilon).$$

Notice that $\Delta(d,\varepsilon)$ does not depend on the function f or on the size of the domain, n. If

$$P(d,\varepsilon) \overset{\Delta_1}{\Longrightarrow} Q(d,\varepsilon) \text{ and } Q(d,\varepsilon) \overset{\Delta_2}{\Longrightarrow} P(d,\varepsilon)$$

for some loss functions Δ_1 and Δ_2, we say that P and Q are *equivalent*. Our main result is that Property $\mathbf{P_1}$, Property $\mathbf{P_2}$,...,Property $\mathbf{P_{11}}$ are all equivalent as stated below.

Theorem 2.3.2. *For any fixed d, the properties $INF(d,\varepsilon)$, $SD(d,\varepsilon)$, $RF(d,\varepsilon)$, $RC(d,\varepsilon)$, $RI(d,\varepsilon)$, $EQD(d,\varepsilon)$, $LSR(d,\varepsilon)$, $SAME(d,\varepsilon)$, $L4C(d,\varepsilon)$, $RAIN(d,\varepsilon)$, and $WRAIN(d,\varepsilon)$ are all equivalent.*

If a Boolean function f satisfies the Balanced Influences Property $INF(d,\varepsilon)$ for some d and ε, we say that f is *quasi-random* of rank d with error bound ε. Such a function f then satisfies all of the other properties in Theorem 2.3.2 with rank d and the appropriate value of ε.

We shall prove Theorem 2.3.2 via a series of theorems, each of which handles a specific implication between two properties. As we have a large number of properties and implications to prove, the proof of Theorem 2.3.2 is divided into two sections as follows:

$$\begin{tikzcd}
WRAIN(d,\varepsilon) \arrow[rrr, "\varepsilon", "\text{Thm 2.5.5}"'] & & & INF(d,\varepsilon) \arrow[rr, "\varepsilon/2", "\text{Thm 2.4.1}"'] \arrow[ddl, "2\varepsilon"', "\text{Thm 2.5.1}"] \arrow[ddddd, "\text{Thm 2.5.6}"', "2\varepsilon/3"] & & SD(d,\varepsilon) \arrow[dd, "\varepsilon"', "\text{Thm 2.4.2}"] \\
 & SAME(d,\varepsilon) \arrow[ul, "\varepsilon", "\text{Thm 2.5.4}"'] & & & & \\
 & & LSR(d,\varepsilon) \arrow[ul, "\varepsilon/2"', "\text{Thm 2.5.3}"] \arrow[dd, bend left, "\varepsilon/2"] & & & RF(d,\varepsilon) \arrow[dd, "\varepsilon/2^d"', "\text{Thm 2.4.3}"] \\
RAIN(d,\varepsilon) \arrow[uuu, "\text{Def.}", "\varepsilon"'] & & & & RI(d,\varepsilon) \arrow[uuul, "\varepsilon"', "\text{Thm 2.4.5}"] & \\
 & & LAC(d,\varepsilon) \arrow[uu, bend left, "\varepsilon/20", "\text{Thm 2.5.2}"'] & & & RC(d,\varepsilon) \arrow[ul, "\text{Thm 2.4.4}", "\varepsilon"'] \\
 & & & EQD(d,\varepsilon) \arrow[uulll, bend left, "\varepsilon/12"', "\text{Thm 2.5.7}"] & &
\end{tikzcd}$$

Figure 2.3. The implications in Theorem 2.3.2. Each edge gives the loss in ε and the reference to the theorem in which the implication is shown.

- Section 2.4 considers the properties $INF(d,\varepsilon)$, $SD(d,\varepsilon)$, $RF(d,\varepsilon)$, $RC(d,\varepsilon)$, and $RI(d,\varepsilon)$ which revolve around the Fourier expansion of a Boolean function.

- Section 2.5 considers the combinatorial properties $LSR(d,\varepsilon)$, $LAC(d,\varepsilon)$, $SAME(d,\varepsilon)$, $EQD(d,\varepsilon)$, $RAIN(d,\varepsilon)$, and $WRAIN(d,\varepsilon)$.

We can summarize the proof of our main theorem in Figure 2.3, where each arrow is labeled with the relevant theorem and error bound.

One can easily observe that $P(d+1,\varepsilon) \implies P(d,\varepsilon)$ for each property P and every d and ε. Our second main result, proven in Section 2.6, shows that these inclusions are strict, i.e., that there are functions which are quasi-random of rank d but not quasi-random of rank $d+1$.

Theorem 2.3.3. *For each $d \geq 1$ and any $0 < \varepsilon < \frac{1}{8}$, there exists an explicit weakly balanced function $f_d : \mathbb{F}_2^n \to \{1,-1\}$ such that*

- *f_d satisfies the Balanced Influences Property $INF(d,\varepsilon)$.*

- *f_d does not satisfy the Balanced Influences Property of rank $d+1$ for any $\varepsilon < \frac{1}{2}$.*

2.4 Proof of Equivalence of Analytical Properties

In this section, we shall prove the equivalence of a number of analytic properties in Theorem 2.3.2. Figure 2.4 provides an outline of the implications proven in this section.

$$\begin{tikzcd}
INF(d,\varepsilon) \arrow[rr, "\varepsilon/2", "\text{Thm 2.4.1}"'] & & SD(d,\varepsilon) \arrow[d, "\text{Thm 2.4.2}", "\varepsilon"'] \\
 & & RF(d,\varepsilon) \arrow[dd, "\text{Thm 2.4.3}", "\varepsilon/2^d"'] \\
 & RI(d,\varepsilon) \arrow[uul, "\varepsilon"', "\text{Thm 2.4.5}"] & \\
 & & RC(d,\varepsilon) \arrow[ul, "\varepsilon"', "\text{Thm 2.4.4}"]
\end{tikzcd}$$

Figure 2.4. The implications in Theorem 2.3.2 proven in Section 2.4. Each edge gives the loss in ε and the reference to the theorem in which the implication is shown.

First, we relate the Balanced Influences Property to the Spectral Discrepancy Property.

Theorem 2.4.1. *For any fixed integer $d \geq 1$ and any $\varepsilon > 0$, the Balanced Influences Property $INF(d,\varepsilon/2)$ implies the Spectral Discrepancy Property $SD(d,\varepsilon)$.*

Proof. Fix a subcube $C(S,z_0)$ where $|S| = n-k$ for $k \leq d$. Let $M \in \mathbb{F}_2^{\overline{S}\times[n]}$ be the projection matrix which sends $z \in \mathbb{F}_2^n$ to $z|_{\overline{S}}$.

We observe that the indicator function $[\gamma \in C(S,z_0)]$ can be written as

$$[\gamma \in C(S,z_0)] = \mathbb{E}_{v\in\mathbb{F}_2^{\overline{S}}}(-1)^{v\cdot(M\gamma - z_0)}. \tag{2.3}$$

Indeed, if $\gamma \in C(S,z_0)$, then $M\gamma = z_0$, and $\mathbb{E}_{v\in\mathbb{F}_2^{\overline{S}}}(-1)^{v\cdot(M\gamma-z_0)} = \mathbb{E}_{v\in\mathbb{F}_2^{\overline{S}}}1 = 1$. If $\gamma \notin C(S,z_0)$, then $\gamma_j \neq (z_0)_j$ for some $j \in \overline{S}$. Therefore, $\mathbb{E}_{v\in\mathbb{F}_2^{\overline{S}}}(-1)^{v\cdot(M\gamma-z_0)} = \mathbb{E}_{v\in\mathbb{F}_2^{\overline{S}}}(-1)^{v\cdot y}$ for some nonzero vector y. Hence, $\mathbb{E}_{v\in\mathbb{F}_2^{\overline{S}}}(-1)^{v\cdot(M\gamma-z_0)} = 0$. Let f be a function which satisfies the Balanced

Influence Property $INF(d, \varepsilon/2)$. We expand the definition of the spectral sample.

$$\begin{aligned}
\mathbb{P}_{\gamma\sim\mathscr{S}_f}\left[\gamma \in C(S,z_0)\right] &= \sum_{\gamma\in C(S,z_0)} \widehat{f}(\gamma)^2 \\
&= \sum_{\gamma\in\mathbb{F}_2^n} \widehat{f}(\gamma)^2 \left[\gamma \in C(S,z_0)\right] \\
&= \sum_{\gamma\in\mathbb{F}_2^n} \widehat{f}(\gamma)^2 \,\mathbb{E}_{v\in\mathbb{F}_2^{\bar{S}}}(-1)^{v\cdot(M\gamma - z_0)} \qquad (2.4)
\end{aligned}$$

where we use Equation (2.3)) in Equation (2.4)). Simplifying further, we have

$$\begin{aligned}
\mathbb{P}_{\gamma\sim\mathscr{S}_f}\left[\gamma \in C(S,z_0)\right] &= \mathbb{E}_{v\in\mathbb{F}_2^{\bar{S}}}(-1)^{v\cdot z_0} \sum_{\gamma\in\mathbb{F}_2^n} \widehat{f}(\gamma)^2 (-1)^{v\cdot M\gamma} \\
&= \mathbb{E}_{v\in\mathbb{F}_2^{\bar{S}}}(-1)^{v\cdot z_0} \sum_{\gamma\in\mathbb{F}_2^n} \widehat{f}(\gamma)^2 (-1)^{(M^\top v)\cdot\gamma} \\
&= \mathbb{E}_{v\in\mathbb{F}_2^{\bar{S}}}(-1)^{v\cdot z_0} \sum_{\gamma\in\mathbb{F}_2^n} \widehat{f}(\gamma)^2 \chi_\gamma(M^\top v) \qquad (2.5) \\
&= \mathbb{E}_{v\in\mathbb{F}_2^{\bar{S}}}(-1)^{v\cdot z_0} f*f(M^\top v) \qquad (2.6)
\end{aligned}$$

where we use the definition of χ_γ in Equation (2.5)) and Fourier expansion of $f*f$ in Equation (2.6)). Notice that $f*f(M^\top 0) = (f*f)(0) = 1$, and that $x = 0$ is the only solution to $M^\top x = 0$. Therefore, we can write

$$\begin{aligned}
\left|\mathbb{P}_{\gamma\sim\mathscr{S}_f}\left[\gamma \in C(S,z_0)\right] - 2^{-k}\right| &= \left|\sum_{v\in\mathbb{F}_2^k} (-1)^{v\cdot z_0} \frac{f*f(M^\top v)}{2^k} - 2^{-k}\right| \\
&= \left|\sum_{v\in\mathbb{F}_2^k\setminus\{0\}} (-1)^{v\cdot z_0} \frac{f*f(M^\top v)}{2^k}\right| \\
&\le \frac{1}{2^k} \sum_{v\in\mathbb{F}_2^k\setminus\{0\}} \left|f*f(M^\top v)\right| \\
&= \frac{1}{2^k} \sum_{v\in\mathbb{F}_2^k\setminus\{0\}} \left|1 - 2\mathrm{I}_{M^\top v}[f]\right|
\end{aligned}$$

where we use Lemma 2.1.13 in the final line. As $k \leq d$, we have $|v| \leq d$. Since M is a projection matrix, $\left|M^{\top} v\right| = |v| \leq d$. Therefore, we may apply $INF(d, \varepsilon/2)$ to find

$$\left|\mathbb{P}_{\gamma \sim \mathscr{S}_f}\left[\gamma \in C(S, z_0)\right] - 2^{-k}\right| \leq \frac{1}{2^k} \sum_{v \in \mathbb{F}_2^k \setminus \{0\}} \varepsilon \leq \varepsilon$$

As $C(S, z_0)$ is arbitrarily chosen, f satisfies the Spectral Discrepancy Property $SD(d, \varepsilon)$ as desired. □

Now we can relate the spectral sample to the Fourier coefficients of restricted functions.

Theorem 2.4.2. *For any fixed $d \geq 1$ and $\varepsilon > 0$ the Spectral Discrepancy Property $SD(d, \varepsilon)$ implies the Restriction Fourier Property $RF(d, \varepsilon)$.*

Proof. This proof is essentially the proof of Corollary 3.22 in [70], which we include here for completeness. Suppose $f : \mathbb{F}_2^n \to \{1, -1\}$ satisfies the Spectral Discrepancy Property $SD(d, \varepsilon)$. Let $C(S, z)$ be an arbitrary subcube of dimension k where $k \leq d$. Then for a fixed $\gamma \in \mathbb{F}_2^S$, Lemma 2.1.17 gives us

$$\begin{aligned}
\mathbb{E}_{z \in \mathbb{F}_2^{\overline{S}}} \widehat{f|_{S,z}}(\gamma)^2 &= \mathbb{E}_{z \in \mathbb{F}_2^{\overline{S}}} \left(\sum_{\delta \in \mathbb{F}_2^{\overline{S}}} \widehat{f}\left(\gamma \bigotimes_S \delta\right) \chi_\delta(z) \right)^2 \\
&= \sum_{\delta_1, \delta_2 \in \mathbb{F}_2^{\overline{S}}} \widehat{f}\left(\gamma \bigotimes_S \delta_1\right) \widehat{f}\left(\gamma \bigotimes_S \delta_2\right) \mathbb{E}_{z \in \mathbb{F}_2^{\overline{S}}} \chi_{\delta_1}(z) \chi_{\delta_2}(z) \\
&= \sum_{\delta \in \mathbb{F}_2^{\overline{S}}} \widehat{f}\left(\gamma \bigotimes_S \delta\right)^2 &\quad (2.7) \\
&= \mathbb{P}_{\eta \sim \mathscr{S}_f}\left[\eta \in C(\overline{S}, \gamma)\right] &\quad (2.8)
\end{aligned}$$

where we use the orthogonality of the Fourier characters in Equation (2.7)) and the definition of the spectral sample Equation (2.8)). As $k \leq d$, $\left|\overline{S}\right| = n - k \geq n - d$. Thus we can apply Property

$SD(d,\varepsilon)$ to $C(\overline{S},z)$ to find that

$$\left|\mathbb{P}_{\eta\sim\mathscr{S}_f}\left[\eta\in C(\overline{S},\gamma)\right]-2^{-k}\right|<\varepsilon$$

for every $\gamma\in\mathbb{F}_2^S$. Hence,

$$\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\widehat{f|_{S,z}}(\gamma)^2-2^{-k}\right|<\varepsilon$$

for every $\gamma\in\mathbb{F}_2^S$. As $C(S,z)$ is arbitrary, f satisfies the Restriction Fourier Property $RF(d,\varepsilon)$. □

With a bound on the Fourier coefficients of restricted functions, we can bound the convolution of a restricted function with itself.

Theorem 2.4.3. *For any fixed $d\geq 1$ and $\varepsilon>0$ the Restriction Fourier Property $RF(d,\varepsilon/2^d)$ implies the Restriction Convolution Property $RC(d,\varepsilon)$*

Proof. Let $f:\mathbb{F}_2^n\to\{1,-1\}$ have the Restriction Fourier Property $RF(d,\varepsilon/2^d)$, and note that f also satisfies $RF(k,\varepsilon/2^k)$ for every $k\leq d$. Fix $k\in\mathbb{N}$ such that $k\leq d$ and a set $S\subseteq[n]$ where $|S|=k$.

We have

$$\begin{aligned}\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}(f|_{S,z}*f|_{S,z})(x)&=\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\sum_{\delta\in\mathbb{F}_2^S}\widehat{f|_{S,z}}(\delta)^2\chi_\delta(x)\\&=\sum_{\delta\in\mathbb{F}_2^S}\left(\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\widehat{f|_{S,z}}(\delta)^2\right)\chi_\delta(x)\end{aligned}$$

Using the Fourier expansion of the indicator function $[x=0]$, we then have

$$\begin{aligned}
\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}(f|_{S,z} * f|_{Sz})(x) - [x=0]\right| &= \left|\sum_{\delta\in\mathbb{F}_2^S}\left(\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\widehat{f|_{S,z}}(\delta)^2 - \frac{1}{2^k}\right)\chi_\delta(x)\right| \\
&\leq \sum_{\delta\in\mathbb{F}_2^S}\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\widehat{f|_{S,z}}(\delta)^2 - \frac{1}{2^k}\right| \\
&\leq \sum_{\delta\in\mathbb{F}_2^S}\frac{\varepsilon}{2^k} \\
&\leq \varepsilon
\end{aligned}$$

where we use $RF(k,\varepsilon/2^k)$ in the penultimate line. Since k and S are arbitrary, we conclude that f satisfies the Restriction Convolution Property $RC(d,\varepsilon)$. □

Theorem 2.4.4. *For any fixed $d \geq 1$ and $\varepsilon > 0$ the Restriction Convolution Property $RC(d,2\varepsilon)$ implies the Restriction Influences Property $RI(d,\varepsilon)$.*

Proof. Suppose f satisfies the Restriction Convolution Property $RC(d,2\varepsilon)$. Applying Lemma 2.1.13 to $f|_{S,z}$, we have

$$\mathrm{I}_\gamma[f|_{S,z}] = \frac{1 - f|_{S,z} * f|_{S,z}(\gamma)}{2}$$

for any fixed S and z. Now fix $k\in\mathbb{N}$ such that $k\leq d$ and $S\subseteq[n]$ where $|S|=k$. Then,

$$\left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\mathrm{I}_\gamma[f|_{S,z}] - \frac{1}{2}\right| = \left|\left(\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\frac{1 - f|_{S,z} * f|_{S,z}(\gamma)}{2}\right) - \frac{1}{2}\right| = \left|\mathbb{E}_{z\in\mathbb{F}_2^{\overline{S}}}\frac{f|_{S,z} * f|_{S,z}(\gamma)}{2}\right|$$

If $\gamma \neq 0$, $RC(d,2\varepsilon)$ implies that the above is at most ε. Hence, f satisfies the Restriction Influences Property $RI(d,\varepsilon)$. □

Theorem 2.4.5. *For any fixed $d\geq 1$ and $\varepsilon > 0$, the Restriction Influences Property $RI(d,\varepsilon)$ implies the Balanced Influences Property $INF(d,\varepsilon)$.*

Proof. Suppose $f : \mathbb{F}_2^n \to \{1,-1\}$ satisfies the Restriction Influences Property $RF(d,\varepsilon)$. Fix $S \subseteq [n]$ with $|S| \le d$ and a nonzero $\gamma \in \mathbb{F}_2^S$. Then,

$$\begin{aligned}\mathbb{E}_{z\in\mathbb{F}_2^{\bar{S}}} \mathrm{I}_\gamma[f|_{S,z}] &= \mathbb{E}_{z\in\mathbb{F}_2^{\bar{S}}} \mathbb{E}_{x\in\mathbb{F}_2^S}[f|_{S,z}(x+\gamma) \neq f|_{S,z}(x)] \\ &= \mathbb{E}_{z\in\mathbb{F}_2^{\bar{S}}} \mathbb{E}_{x\in\mathbb{F}_2^S}[f(x \underset{S}{\otimes} z + \gamma \underset{S}{\otimes} 0)) \neq f(x \underset{S}{\otimes} z)]\end{aligned}$$

Let $y = x \underset{S}{\otimes} z$ and $\delta = \gamma \underset{S}{\otimes} 0$. Note that $|\delta| \le d$ as $|S| \le d$. Thus

$$\begin{aligned}\mathbb{E}_{z\in\mathbb{F}_2^{\bar{S}}} \mathrm{I}_\gamma[f|_{S,z}] &= \mathbb{E}_{y\in\mathbb{F}_2^n}[f(y+\delta) \neq f(y)] \\ &= \mathrm{I}_\delta[f]\end{aligned}$$

Since any vector of Hamming weight at most d can be represented as $\gamma \underset{S}{\otimes} 0$ for some set S with $|S| \le d$ and $\gamma \in \mathbb{F}_2^S$, f satisfies the Balanced Influences Property $INF(d,\varepsilon)$. □

2.5 Proof of Equivalence of Combinatorial Properties

In this section, we continue the proof of Theorem 2.3.2 and prove that several of our combinatorial properties are equivalent to the Balanced Influences Property. The diagram in Figure 2.5 summarizes the proofs found in this section.

We begin by considering the relationship between γ-Influences and the Local Strong Regularity Property.

Theorem 2.5.1. *For any fixed $d \ge 1$ and $\varepsilon > 0$, the Balanced Influences Property $INF(d, 2\varepsilon)$ implies the Local Strong Regularity Property $LSR(d,\varepsilon)$.*

Proof. Suppose $f : \mathbb{F}_2^n \to \{1,-1\}$ satisfies the Balanced Influences Property $INF(d, 2\varepsilon)$. Fix

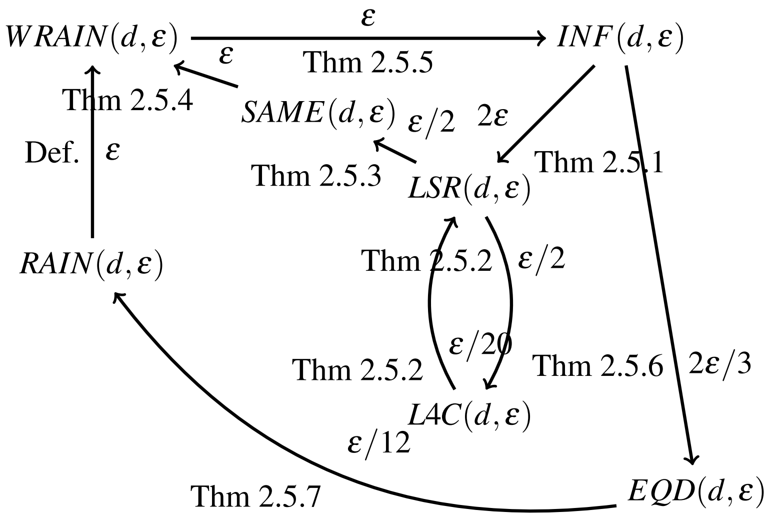

Figure 2.5. The implications in Theorem 2.3.2 proven in Section 2.5. Each edge gives the loss in ε and the reference to the theorem in which the implication is shown.

u,v in the Cayley graph of f such that $0 < |u-v| \leq d$. Then,

$$\begin{aligned}
\left|\frac{|N(u)\cap N(v)|}{2^n} - \frac{1}{4} + \frac{\widehat{f}(0)}{2}\right| &= \left|\mathbb{E}_{z\in\mathbb{F}_2^n}\frac{(1-f(u+z))(1-f(v+z))}{4} - \frac{1}{4} + \frac{\widehat{f}(0)}{2}\right| \\
&= \left|\frac{\widehat{f}(0)}{2} - \frac{\mathbb{E}_{z\in\mathbb{F}_2^n} f(u+z)+f(v+z)}{4} + \frac{\mathbb{E}_{z\in\mathbb{F}_2^n} f(u+z)f(v+z)}{4}\right| \\
&= \frac{1}{4}\left|f*f(u+v)\right| \\
&= \frac{1}{2}\mathrm{I}_{u+v}[f] \\
&\leq \varepsilon
\end{aligned}$$

where we use Lemma 2.1.13 in the penultimate line and $INF(d,2\varepsilon)$ in the ultimate line. It follows that f satisfies the Local Strong Regularity Property $LSR(d,\varepsilon)$. □

As Local Strong Regularity is a condition on common neighbors, we can use it to count 4-cycles.

Theorem 2.5.2. *For any fixed $d \geq 1$ and $\varepsilon > 0$, the Local 4-Cycle Property $L4C(d,\varepsilon/20)$ implies the Local Strong Regularity Property $LSR(d,\varepsilon)$ and the Local Strong Regularity Property*

$LSR(d,\varepsilon/2)$ implies the Local 4-Cycle Property $LAC(d,\varepsilon)$.

Proof. Let u,v be the vertices in the left part of C_4, and fix an injective map $\phi : \{u,v\} \hookrightarrow \mathbb{F}_2^n$. The key observation is the following:

$$\overline{\hom}_\phi(C_4,\mathrm{Cay}(f)) = \frac{\left|N(\phi(u)) \cap N(\phi(v))\right|^2}{2^{2n}}$$

Indeed, a (possibly non-injective) graph homomorphism of C_4 with a fixed left part is simply a choice of two vertices in the common neighborhood of $\phi(u)$ and $\phi(v)$ in $\mathrm{Cay}(f)$. Let $N(u,v) = \left|N(\phi(u)) \cap N(\phi(v))\right|$.

Hence,

$$\begin{aligned}\left|\overline{\hom}_\phi(C_4,\mathrm{Cay}(f)) - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)^2\right| &= \left|\frac{N(u,v)^2}{2^{2n}} - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)^2\right| \\ &= \left|\frac{N(u,v)}{2^n} + \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right|\left|\frac{N(u,v)}{2^n} - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right|\end{aligned} \tag{2.9}$$

Now we prove both of the implications in the theorem. Assume first that f satisfies the Local 4-Cycle Property $LAC(d,\varepsilon/20)$. By Equation (2.9)),

$$\frac{\varepsilon}{20} \geq \left|\overline{\hom}_\phi(C_4,\mathrm{Cay}(f)) - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)^2\right| \geq \frac{1}{20}\left|\frac{N(u,v)}{2^n} - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right| \tag{2.10}$$

where we use the fact that f is weakly balanced to show that

$$\left|\frac{N(u,v)}{2^n} + \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right| \geq \left|\frac{N(u,v)}{2^n} + \frac{1}{4}\right| - \left|\frac{\widehat{f}(0)}{2}\right| \geq \frac{1}{4} - \frac{1}{5} = \frac{1}{20}.$$

It follows that f satisfies the Local Strong Regularity Property $LSR(d,\varepsilon)$.

Now assume that f satisfies the Local Strong Regularity Property $LSR(d,\varepsilon/2)$, so that

$\left|\frac{N(u,v)}{2^n} - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right| < \varepsilon/2$. Again using Equation (2.9)), we find that

$$\begin{aligned}
\left|\overline{\hom}_\phi(C_4, \mathrm{Cay}(f)) - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)^2\right| &< \left|\frac{N(u,v)}{2^n} + \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right| \frac{\varepsilon}{2} \\
&\leq \left(\left|\frac{N(u,v)}{2^n} - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)\right| + 2\left|\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right|\right) \frac{\varepsilon}{2}
\end{aligned}$$

By $LSR(d, \varepsilon/2)$,

$$\begin{aligned}
\left|\overline{\hom}_\phi(C_4, \mathrm{Cay}(f)) - \left(\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right)^2\right| &\leq \left(\frac{\varepsilon}{2} + 2\left|\frac{1}{4} - \frac{\widehat{f}(0)}{2}\right|\right) \frac{\varepsilon}{2} \\
&\leq \left(\frac{\varepsilon}{2} + \frac{1}{2} + \left|\widehat{f}(0)\right|\right) \frac{\varepsilon}{2} \\
&= \left(\frac{\varepsilon}{2} + \frac{9}{10}\right) \frac{\varepsilon}{2} \\
&\leq \varepsilon
\end{aligned}$$

where we use the facts that f is weakly balanced and $\varepsilon \leq 1$. We conclude that f satisfies the Local 4-Cycle Property $L4C(d, \varepsilon)$. □

Local Strong Regularity also allows us to consider the Cayley graph $\mathrm{Cay}(-f)$.

Theorem 2.5.3. *For any fixed $d \geq 1$ and $\varepsilon > 0$, the Local Strong Regularity Property $LSR(d, \varepsilon/2)$ implies the Local Sameness Property $SAME(d, \varepsilon)$.*

Proof. Let $f : \mathbb{F}_2^n \to \{1, -1\}$ be a Boolean function which satisfies the Local Strong Regularity Property $LSR(d, \varepsilon/2)$. Fix $u, v \in \mathbb{F}_2^n$ such that $|u - v| \leq d$. Similarly to Theorem 2.5.2, let $N^+(u,v) = \left|N_{\mathrm{Cay}(f)}(u) \cap N_{\mathrm{Cay}(f)}(v)\right|$ and let $N^-(u,v) = \left|N_{\mathrm{Cay}(-f)}(u) \cap N_{\mathrm{Cay}(-f)}(v)\right|$.

We observe that

$$\begin{aligned}\frac{N^-(u,v)}{2^n} &= \mathbb{E}_{x\in\mathbb{F}_2^n}\frac{1+f(x+u)}{2}\frac{1+f(x+v)}{2}\\ &= \frac{1}{4}+\frac{\widehat{f}(0)}{2}+\mathbb{E}_{x\in\mathbb{F}_2^n}f(x+u)f(x+v)\\ &= \frac{N^+(u,v)}{2^n}+\widehat{f}(0)\end{aligned}$$

Hence,

$$\left|\frac{N^+(u,v)+N^-(u,v)}{2^n}-\frac{1}{2}\right| = \left|2\frac{N^+(u,v)}{2^n}-\frac{1}{2}+\widehat{f}(0)\right| = 2\left|\frac{N^+(u,v)}{2^n}-\frac{1}{4}+\frac{\widehat{f}(0)}{2}\right| \leq \varepsilon$$

where we use $LSR(d,\varepsilon/2)$ in the final line. Hence, f has the Local Sameness Property $SAME(d,\varepsilon)$. □

Since the rainbow Hamming graph has an edge uv with color x whenever $f(u+x) = f(v+x)$, the Local Sameness Property gives a natural way to control the rainbow Hamming graph.

Theorem 2.5.4. *For any fixed $d \geq 1$ and $\varepsilon > 0$, the Local Sameness Property $SAME(d,\varepsilon)$ implies the Weak Rainbow Embeddings Property $WRAIN(d,\varepsilon)$.*

Proof. Suppose $f : \mathbb{F}_2^n \to \{1,-1\}$ satisfies the Local Sameness Property $SAME(d,\varepsilon)$. Fix $u,v \in B_d(n,0)$, and let ϕ be an injection from $V(K_2)$ to $\{u,v\}$. By definition of rainbow embeddings, we have

$$\mathrm{ch}_\phi(K_2,RHG(d,f)) = \mathbb{E}_{\chi:E(K_2)\to\mathbb{F}_2^n}[(\phi(u),\phi(v),\chi(e)) \in E(RHG(d,f))]$$

Setting $x = \chi(e)$ and applying the definition of the edge set of $RHG(f)$, we have

$$\begin{aligned}
\mathrm{ch}_\phi(K_2, RHG(d,f)) &= \mathbb{E}_{x \in \mathbb{F}_2^n}[f(\phi(u)+x) = f(\phi(v)+x)] \\
&= \mathop{\mathbb{E}}_{x \in \mathbb{F}_2^n}[f(\phi(u)+x) = f(\phi(v)+x) = 1] + \\
&\qquad [f(\phi(u)+x) = f(\phi(v)+x) = -1] \\
&= \frac{\left|N_{\mathrm{Cay}(-f)}(\phi(u)) \cap N_{\mathrm{Cay}(-f)}(\phi(v))\right| + \left|N_{\mathrm{Cay}(f)}(\phi(u)) \cap N_{\mathrm{Cay}(f)}(\phi(v))\right|}{2^n}
\end{aligned}$$

Let $X = \left|\mathrm{ch}_\phi(K_2, RHG(d,f)) - \frac{1}{2}\right|$. We then have,

$$\begin{aligned}
X &= \left| \frac{\left|N_{\mathrm{Cay}(-f)}(\phi(u)) \cap N_{\mathrm{Cay}(-f)}(\phi(v))\right| + \left|N_{\mathrm{Cay}(f)}(\phi(u)) \cap N_{\mathrm{Cay}(f)}(\phi(v))\right|}{2^n} - \frac{1}{2} \right| \\
&< \varepsilon
\end{aligned}$$

by the Local Sameness Property $SAME(d,\varepsilon)$. Hence, f satisfies the Weak Rainbow Embeddings Property $WRAIN(d,\varepsilon)$. □

Our next theorem is an immediate consequence of the Weak Rainbow Embeddings Property.

Theorem 2.5.5. *For any fixed $d \geq 1$ and $\varepsilon > 0$, the Weak Rainbow Embeddings Property $WRAIN(d,\varepsilon)$ implies the Balanced Influences Property $INF(d,\varepsilon)$.*

Proof. Suppose $f : \mathbb{F}_2^n \to \{1,-1\}$ satisfies the Weak Rainbow Embeddings Property $WRAIN(d,\varepsilon)$. Fix $u \in B_d(n,0)$, and let ϕ be an injection from $V(K_2)$ to $\{u,0\}$. By definition of rainbow embeddings, we have

$$\mathrm{ch}_\phi(K_2, RHG(d,f)) = \mathbb{E}_{\chi : E(K_2) \to \mathbb{F}_2^n}[(u, 0, \chi(e)) \in E(RHG(d,f))]$$

Setting $x = \chi(e)$ and applying the definition of the edge set of $RHG(f)$

$$\begin{aligned}
\mathrm{ch}_{\phi}(K_2, RHG(d,f)) &= \mathbb{E}_{x\in\mathbb{F}_2^n}[f(u+x) = f(x)] \\
&= \mathbb{P}_{x\in\mathbb{F}_2^n}\left[f(x+u) = f(x)\right] \\
&= 1 - \mathrm{I}_u[f]
\end{aligned}$$

By Property $WRAIN(d,\varepsilon)$, we have that $|\,\mathrm{ch}_{\phi}(K_2, RHG(d,f)) - \frac{1}{2}| < \varepsilon$. Hence, it follows that $\left|\mathrm{I}_u[f] - \frac{1}{2}\right| < \varepsilon$ and f satisfies the Balanced Influences Property $INF(d,\varepsilon)$. □

Theorem 2.5.6. *For any fixed $d \geq 1$ and $\varepsilon > 0$, the Balanced Influences Property $INF(d, 2\varepsilon/3)$ implies the Equidistributed Derivatives Property $EQD(d,\varepsilon)$.*

Proof. Fix $a, b \in \mathbb{F}_2^n$ such that $|a|, |b| \leq d$ and $0 < |a-b| \leq d$. Fix also $c_0, c_1 \in \{1,-1\}$. Let

$$X = \left|\mathbb{E}_{x\in\mathbb{F}_2^n}[\Delta_a f(x) = c_1][\Delta_b f(x) = c_0] - \frac{1}{4}\right|$$

We then have

$$\begin{aligned}
X &= \left|\mathbb{E}_{x\in\mathbb{F}_2^n}\left(\frac{1+c_1\Delta_a f(x)}{2}\right)\left(\frac{1+c_0\Delta_b f(x)}{2}\right) - \frac{1}{4}\right| \\
&= \frac{1}{4}\left|c_1\,\mathbb{E}_{x\in\mathbb{F}_2^n}\Delta_b f(x) + c_0\,\mathbb{E}_{x\in\mathbb{F}_2^n}\Delta_a f(x) + c_0c_1\,\mathbb{E}_{x\in\mathbb{F}_2^n}\Delta_a f(x)\Delta_b f(x)\right| \\
&= \frac{1}{4}\left|\mathbb{E}_{x\in\mathbb{F}_2^n}\big(c_1 f(x)f(x+b) + c_0 f(x+a)f(x)\big) + c_0c_1\,\mathbb{E}_{x\in\mathbb{F}_2^n}\Delta_a f(x)\Delta_b f(x)\right|
\end{aligned}$$

Note that $\Delta_a f(x)\Delta_b f(x) = f(x+a)f(x)f(x+b)f(x) = f(x+a)f(x+b)$ as $f(x) \in \{1,-1\}$.

Therefore

$$\begin{aligned}
X &= \frac{1}{4}\left|c_1 f * f(b) + c_0 f * f(a) + c_0 c_1 f * f(a+b)\right| \\
&\leq \frac{1}{4}\left(\left|f * f(b)\right| + \left|f * f(a)\right| + \left|f * f(a+b)\right|\right) \\
&= \frac{1}{2}\left(\mathrm{I}_b[f] + \mathrm{I}_a[f] + \mathrm{I}_{a+b}[f]\right) \\
&\leq \varepsilon
\end{aligned}$$

where we use Lemma 2.1.13 and $INF(d, 2\varepsilon/3)$ thrice in the final line. Thus f satisfies the Equidistributed Derivatives property $EQD(d, \varepsilon)$. □

Our final and most technical result connects equidistributed derivatives and rainbow embeddings.

Theorem 2.5.7. *For any fixed $d \geq 1$ and $1 \geq \varepsilon > 0$, the Equidistributed Derivatives Property $EQD(d, \varepsilon/12)$ implies the Rainbow Embeddings Property $RAIN(d, \varepsilon)$.*

Proof. Let G be a fixed graph with at most $\max\{\sqrt{\varepsilon}2^{n/2-1}, 1\}$ edges. Let $\phi : V(G) \hookrightarrow B_d(n, 0)$ be an injection of diameter at most d.

We first consider the case where 1 maximizes the above. Let (u, v) be the single edge in G. By the definition of $RHG(f)$, we have

$$\begin{aligned}
\mathrm{ch}_\phi(G, RHG(d, f)) &= \mathbb{E}_{\chi : E(G) \hookrightarrow \mathbb{F}_2^n}[(\phi(u), \phi(v), \chi((u,v))) \in E(RHG(d, f))] \\
&= \mathbb{E}_{x \in \mathbb{F}_2^n}[f(\phi(u) + x) = f(\phi(v) + x)] \\
&= \mathbb{E}_{x \in \mathbb{F}_2^n}[f(\phi(u) + x) f(x) = f(\phi(v) + x) f(x)] \\
&= \mathbb{E}_{x \in \mathbb{F}_2^n}[\Delta_{\phi(u)} f(x) = \Delta_{\phi(v)} f(x)] \\
&= \mathbb{E}_{x \in \mathbb{F}_2^n}[\Delta_{\phi(u)} f(x) = 1][\Delta_{\phi(v)} f(x) = 1] + \\
&\qquad \mathbb{E}_{x \in \mathbb{F}_2^n}[\Delta_{\phi(u)} f(x) = -1][\Delta_{\phi(v)} f(x) = -1]
\end{aligned}$$

By $EQD(d,\varepsilon/12)$, we have

$$\begin{aligned}\left|\mathrm{ch}_\phi(G,RHG(d,f))-\frac{1}{2}\right| &\le \left|\mathbb{E}_{x\in\mathbb{F}_2^n}[\Delta_{\phi(u)}f(x)=1][\Delta_{\phi(v)}f(x)=1]-\frac{1}{4}\right|+\\ &\quad \left|\mathbb{E}_{x\in\mathbb{F}_2^n}[\Delta_{\phi(u)}f(x)=-1][\Delta_{\phi(v)}f(x)=-1]-\frac{1}{4}\right|\\ &\le \frac{\varepsilon}{6}\\ &\le \varepsilon\end{aligned}$$

so we turn to the case where G has more than one edge, but at most $\sqrt{\varepsilon}2^{n/2-1}$ edges. Let $m=\left|E(G)\right|$

Recall that $\mathrm{ch}_\phi(G,RHG(d,f))$ counts the normalized number of colorings χ such that ϕ becomes a rainbow embedding of G with the coloring χ in the rainbow Hamming graph $RHG(d,f)$. More formally, we have

$$\begin{aligned}\mathrm{ch}_\phi(G,RHG(d,f)) &= \mathbb{E}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in E(G)}[(\phi(u),\phi(v),\chi((u,v)))\in E(RHG(d,f))]\\ &= \mathbb{E}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in E(G)}[f(\phi(u)+\chi((u,v)))=f(\phi(v)+\chi((u,v)))]\end{aligned}$$

We observe that the event

$$f(\phi(u)+\chi((u,v)))=f(\phi(v)+\chi((u,v))) \iff \Delta_{\phi(u)}f(\chi((u,v)))=\Delta_{\phi(v)}f(\chi((u,v)))$$

Let $P_\chi(u,v)$ denote the event that $\Delta_{\phi(u)}f(\chi((u,v)))=1$ and $\Delta_{\phi(v)}f(\chi((u,v)))=1$. Let $N_\chi(u,v)$ denote the event that $\Delta_{\phi(u)}f(\chi((u,v)))=-1$ and $\Delta_{\phi(v)}f(\chi((u,v)))=-1$. We then have

$$\mathrm{ch}_\phi(G,RHG(d,f))=\mathbb{E}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in E(G)}\left([N_\chi(u,v)]+[P_\chi(u,v)]\right)$$

Hence, we have

$$
\begin{aligned}
\left|\mathrm{ch}_{\phi}(G,RHG(d,f))-2^{-m}\right| &= \left|\mathbb{E}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in E(G)}\left([N_{\chi}(u,v)]+[P_{\chi}(u,v)]\right)-2^{-m}\right| \\
&= \left|\mathop{\mathbb{E}}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in E(G)}\left([N_{\chi}(u,v)]+[P_{\chi}(u,v)]-\frac{1}{2}+\frac{1}{2}\right)-2^{-m}\right| \\
&= \left|\sum_{\emptyset\neq R\subseteq E(G)}\frac{\mathop{\mathbb{E}}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in R}\left([N_{\chi}(u,v)]+[P_{\chi}(u,v)]-\frac{1}{2}\right)}{2^{\left|E(G)\setminus R\right|}}\right| \\
&\leq \sum_{\emptyset\neq R\subseteq E(G)}\frac{\left|\mathop{\mathbb{E}}_{\chi:E(G)\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in R}\left([N_{\chi}(u,v)]+[P_{\chi}(u,v)]-\frac{1}{2}\right)\right|}{2^{\left|E(G)\setminus R\right|}}
\end{aligned}
$$

For $R\subseteq E(G)$, let $X_R=\sum_{\chi:R\hookrightarrow\mathbb{F}_2^n}\prod_{(u,v)\in R}\left([N_{\chi}(u,v)]+[P_{\chi}(u,v)]-\frac{1}{2}\right)$. Let Y_R be the analogous version of X_R which sums over all functions, not just injections, i.e.,

$$
Y_R=\sum_{\chi:R\to\mathbb{F}_2^n}\prod_{(u,v)\in R}\left([N_{\chi}(u,v)]+[P_{\chi}(u,v)]-\frac{1}{2}\right).
$$

We then have

$$
\begin{aligned}
\left|\mathrm{ch}_{\phi}(G,RHG(d,f))-2^{-m}\right| &\leq \sum_{\emptyset\neq R\subseteq E(G)}2^{-\left|E(G)\setminus R\right|}\frac{1}{(2^n)_{|R|}}\left|X_R\right| \\
&\leq \sum_{\emptyset\neq R\subseteq E(G)}2^{-\left|E(G)\setminus R\right|}\frac{1}{(2^n)_{|R|}}\left(\left|X_R-Y_R\right|+\left|Y_R\right|\right)
\end{aligned}
$$

Fix $R \subseteq E(G)$.

$$\begin{aligned}
|X_R - Y_R| &= \left| \sum_{\substack{\chi : R \to \mathbb{F}_2^n \\ \chi \text{not injective}}} \prod_{(u,v) \in R} \left([N_\chi(u,v)] + [P_\chi(u,v)] - \frac{1}{2} \right) \right| \\
&\leq \sum_{\substack{\chi : R \to \mathbb{F}_2^n \\ \chi \text{not injective}}} \prod_{(u,v) \in R} \left| \left([N_\chi(u,v)] + [P_\chi(u,v)] - \frac{1}{2} \right) \right|
\end{aligned}$$

As $N_\chi(u,v)$ and $P_\chi(u,v)$ cannot occur simultaneously, we have

$$\begin{aligned}
|X_R - Y_R| &\leq \sum_{\substack{\chi : R \to \mathbb{F}_2^n \\ \chi \text{not injective}}} \left(\frac{1}{2} \right)^{|R|} \\
&\leq \left(2^{n|R|} - (2^n)_{|R|} \right) \left(\frac{1}{2} \right)^{|R|} \\
&= (2^n)_{|R|} \left(\frac{2^{n|R|}}{(2^n)_{|R|}} - 1 \right) \left(\frac{1}{2} \right)^{|R|}
\end{aligned}$$

Observe that $|R|^2 \leq \left|E(G)\right|^2 \leq \varepsilon 2^{n-2}$. Thus $\frac{|R|^2}{2^n} \leq \frac{\varepsilon}{4} \leq \frac{1}{4}$. We have

$$\begin{aligned}
\frac{2^{n|R|}}{(2^n)_{|R|}} &\leq \left(\frac{2^n}{2^n - |R|} \right)^{|R|} \\
&= \left(1 - \frac{|R|}{2^n} \right)^{-|R|} \\
&\leq \exp\left(2 \frac{|R|^2}{n} \right) \qquad (2.11) \\
&\leq 1 + 2\frac{|R|^2}{n} + \left(2\frac{|R|^2}{n} \right)^2 \qquad (2.12)
\end{aligned}$$

where we use the fact that $e^{-2x} \leq 1 - x$ for $x \in [0, 1.59]$ in Equation (2.11)) and the fact that

$e^x \leq 1+x+x^2$ for $x \in [0,1.79]$ in Equation (2.12)). As $\frac{|R|^2}{2^n} \leq \frac{\varepsilon}{4}$, it follows that

$$\frac{2^{n|R|}}{(2^n)_{|R|}} \leq 1+\frac{\varepsilon}{2}$$

and thus

$$|X_R - Y_R| \leq (2^n)_{|R|} \left(\frac{1}{2}\right)^{|R|} \frac{\varepsilon}{2}$$

Now we turn to Y_R.

$$\begin{aligned}
|Y_R| &= \left| \sum_{\chi: E(G)\to\mathbb{F}_2^n} \prod_{(u,v)\in R} \left([N_\chi(u,v)] + [P_\chi(u,v)] - \frac{1}{2} \right) \right| \\
&= \left| \prod_{(u,v)\in R} \left(\left(\sum_{\chi:\{(u,v)\}\to\mathbb{F}_2^n} [N_\chi(u,v)] - \frac{1}{4} \right) + \left(\sum_{\chi:\{(u,v)\}\to\mathbb{F}_2^n} [P_\chi(u,v)] - \frac{1}{4} \right) \right) \right| \\
&= 2^{n|R|} \left| \prod_{(u,v)\in R} \left(\left(\mathbb{E}_{\chi:\{(u,v)\}\to\mathbb{F}_2^n} [N_\chi(u,v)] - \frac{1}{4} \right) + \left(\mathbb{E}_{\chi:\{(u,v)\}\to\mathbb{F}_2^n} [P_\chi(u,v)] - \frac{1}{4} \right) \right) \right|
\end{aligned}$$

By definition,

$$\begin{aligned}
\mathbb{E}_{\chi:\{(u,v)\}\to\mathbb{F}_2^n}[P_\chi(u,v)] &= \mathbb{E}_{\chi:\{(u,v)\}\to\mathbb{F}_2^n}[\Delta_{\phi(u)} f(\chi((u,v))) = 1][\Delta_{\phi(v)} f(\chi((u,v))) = 1] \\
\mathbb{E}_{\chi:\{(u,v)\}\to\mathbb{F}_2^n}[N_\chi(u,v)] &= \mathbb{E}_{\chi:\{(u,v)\}\to\mathbb{F}_2^n}[\Delta_{\phi(u)} f(\chi((u,v))) = -1][\Delta_{\phi(v)} f(\chi((u,v))) = -1]
\end{aligned}$$

By assumption, ϕ is a map of diameter at most d from $V(G)$ to $B_d(n,0)$. Thus, $\left|\phi(u)\right| \leq d$, $\left|\phi(v)\right| \leq d$, and $\left|\phi(u)-\phi(v)\right| \leq d$ for every $(u,v) \in E(G)$. Hence, we may apply $EQD(d,\varepsilon/12)$ to find that

$$Y_R \leq 2^{n|R|} \left| \prod_{(u,v)\in R} \left(\frac{\varepsilon}{12} + \frac{\varepsilon}{12} \right) \right| \leq 2^{n|R|} \left(\frac{\varepsilon}{6}\right)^{|R|} \leq (2^n)_{|R|} \left(1+\frac{\varepsilon}{2}\right) \left(\frac{\varepsilon}{6}\right)^{|R|}$$

where we use the same bound on $2^{n|R|}$ as above. Now we can put everything back together as follows:

$$\begin{aligned}\left|\mathrm{ch}_{\phi}(G,RHG(d,f))-2^{-m}\right| &\le \sum_{\emptyset\neq R\subseteq E(G)} 2^{-\left|E(G)\setminus R\right|}\left(\left(\frac{1}{2}\right)^{|R|}\frac{\varepsilon}{2}+\left(1+\frac{\varepsilon}{2}\right)\left(\frac{\varepsilon}{6}\right)^{|R|}\right)\\ &=\frac{\varepsilon}{2}\left(1-2^{-m}\right)+\left(1+\frac{\varepsilon}{2}\right)\left(\left(\frac{1}{2}+\frac{\varepsilon}{6}\right)^{m}-\frac{1}{2}^{m}\right)\end{aligned}$$

As $\frac{\varepsilon}{3}\le\frac{1}{2}$, we have the following:

$$\begin{aligned}\left(\frac{1}{2}+\frac{\varepsilon}{6}\right)^{m}-\frac{1}{2}^{m} &= \frac{1}{2}^{m}\left(\left(1+\frac{\varepsilon}{3}\right)^{m}-1\right)\\ &\le \frac{1}{2}^{m}\left(\frac{\varepsilon}{3}m\left(1+\frac{\varepsilon}{3}\right)^{m-1}\right) && (2.13)\\ &\le \frac{\varepsilon}{6}m\left(\frac{3}{4}\right)^{m-1} && (2.14)\\ &\le \frac{\varepsilon}{3} && (2.15)\end{aligned}$$

where we use the fact that $(1+x)^m \le 1+mx(1+x)^{m-1}$ in Equation (2.13)), the fact that $\frac{\varepsilon}{3}\le\frac{1}{2}$ in Equation (2.14)), and the numerical fact that $m(3/4)^{m-1}\le 2$ for every $m\ge 1$ in Equation (2.15)). Therefore,

$$\left|\mathrm{ch}_{\phi}(G,RHG(d,f))-2^{-m}\right| \le \frac{\varepsilon}{2}+\left(1+\frac{\varepsilon}{2}\right)\frac{\varepsilon}{3}\le\varepsilon$$

as $\varepsilon\le 1$. Thus f also satisfies the Rainbow Embeddings Property $RAIN(d,\varepsilon)$. □

Remark 2.5.8. Chapter 2, in full, is a reprint of the material as it appears in Quasi-random Boolean Functions, Fan Chung and Nicholas Sieger, which is in review at the Electronic Journal of Combinatorics. The dissertation author was the primary investigator and author of this paper.

2.6 Constructions of quasi-random Functions and Separation of the Hierarchy

In this section, we construct a large class of functions which separate the Balanced Influences Property $INF(d+1,\varepsilon)$ from $INF(d,\varepsilon')$.

An $[n,k,d]$-*binary linear code* is a subspace $\mathscr{C} \subseteq \mathbb{F}_2^n$ of dimension k such that the minimum Hamming distance between elements of $\mathscr{C}$ is d. An $[n,k,d]$-binary linear code may be specified by its *parity check matrix* $M \in \mathbb{F}_2^{(n-k)\times n}$ which has the property that $x \in \mathscr{C} \iff Mx = 0$. Note that the parity check matrix has rank $n-k$. We will need the following elementary fact regarding linear codes of distance d.

Lemma 2.6.1. *[[50], Proposition 2.3.5] If M is the parity check matrix of a code with distance strictly greater than d, then any nonzero* $x \in \ker(M)$ *must have* $|x| > d$.

Example 2.6.2. Let $\mathscr{C}$ be the $[8,4]$-Extended Hamming code with parity check matrix H

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

One can check that no vector of Hamming weight 3 or less can be an element of the kernel, as every set of 3 columns has at least one row with an odd number of 1's

The goal of this section is to demonstrate that a bent function composed with the parity check matrix of a distance d linear code is quasi-random of rank d with error ε for any $\varepsilon > 0$.

Proof of Theorem 2.3.3. Let $\mathscr{C}$ be an $[n,k,d+1]$-binary linear code such that $n-k$ is even and $n \geq k+4$. Let $H \in \mathbb{F}_2^{(n-k)\times n}$ be a parity check matrix for $\mathscr{C}$. Let $g : \mathbb{F}_2^{n-k} \to \{1,-1\}$ be a bent function, and define$f : \mathbb{F}_2^n \to \{1,-1\}$ by

$$f(x) := g(Hx).$$

We claim that

$$\mathrm{I}_\gamma[f] = \begin{cases} \frac{1}{2} & \gamma \notin \ker(H) \\ 0 & \gamma \in \ker(H) \end{cases}$$

Indeed, by Lemma 2.1.13, we have

$$\begin{aligned} \mathrm{I}_\gamma[f] &= \frac{1}{2} - \frac{1}{2} f * f(\gamma) \\ &= \frac{1}{2} - \frac{1}{2} \mathbb{E}_{\delta \in \mathbb{F}_2^n} g(H\delta) g(H(\delta + \gamma)) \\ &= \frac{1}{2} - \frac{1}{2} \mathbb{E}_{\eta \in \mathrm{Range}(H)} g(\eta) g(\eta + H\gamma) \qquad (2.16) \end{aligned}$$

where in Equation (2.16)) we use the fact that $H\delta$ is uniformly distributed on Range(H) when δ is uniformly distributed on $\mathbb{F}_2^n$. As the parity check matrix is a surjective linear map from $\mathbb{F}_2^n \to \mathbb{F}_2^{n-k}$, we have

$$\begin{aligned} \mathrm{I}_\gamma[f] &= \frac{1}{2} - \frac{1}{2} \mathbb{E}_{\eta \in \mathbb{F}_2^{n-k}} g(\eta) g(\eta + H\gamma) \\ &= \begin{cases} \frac{1}{2} & \gamma \notin \ker(H) \\ 0 & \gamma \in \ker(H) \end{cases} \qquad (2.17) \end{aligned}$$

where we use the fact that a $g * g(x) = 0$ for $x \neq 0$ (see Proposition 2.1.6) in Equation (2.17)). Now we can apply Lemma 2.6.1 to conclude that if $|\gamma| \leq d$, $\gamma \notin \ker(H)$. It follows that $\mathrm{I}_\gamma[f] = \frac{1}{2}$ for every $\gamma \in \mathbb{F}_2^n$ with $0 < |\gamma| \leq d$.

Similarly, as $\mathscr{C}$ has distance $d+1$, there is some $\gamma' \in \mathbb{F}_2^n$ with Hamming weight $d+1$ such that $H\gamma' = 0$. Hence, $\mathrm{I}_{\gamma'}[f] = 0$ by Equation (2.17)) above. Thus $INF(d+1, \varepsilon)$ cannot hold for f unless $\varepsilon \geq \frac{1}{2}$.

It remains to show that $\left|\widehat{f}(0)\right| < \frac{2}{5}$, i.e., that f is weakly balanced. To that end we observe

$$\widehat{f}(0) = \mathbb{E}_{x \in \mathbb{F}_2^n} g(Hx) = \mathbb{E}_{y \in \mathbb{F}_2^{n-k}} g(y)$$

by the same reasoning as in Equation (2.16)) above. Since g is bent, it follows that

$$\left|\widehat{f}(0)\right| = \left|\widehat{g}(0)\right| = 2^{-\frac{n-k}{2}}.$$

As $n \geq k+4$, we conclude that $\left|\widehat{f}(0)\right| \leq \frac{1}{4} < \frac{2}{5}$ and thus $INF(d,\varepsilon)$ holds for f for any $\varepsilon > 0$.

□

Finally, we show that the Balanced Influences Property implies weak balance. We will need the following lemma:

Lemma 2.6.3. *If $f : \mathbb{F}_2^n \to \{1,-1\}$ has the Balanced Influences property $INF(d,\varepsilon/2)$, then $\widehat{f}(\gamma)^2 \leq 2^{-d} + \varepsilon$ for every $\gamma \in \mathbb{F}_2^n$.*

Proof. By Theorem 2.4.1, if f has the Balanced Influences Property $INF(d,\varepsilon/2)$, then f has the Spectral Discrepancy Property $SD(d,\varepsilon)$. Fix $\gamma \in \mathbb{F}_2^n$ and let $C(S,z)$ be a subcube of dimension $n-d$ which contains γ. By $SD(d,\varepsilon)$,

$$\widehat{f}(\gamma)^2 \leq \sum_{\delta \in C(S,z)} \widehat{f}(\delta)^2 \leq 2^{-d} + \varepsilon.$$

Therefore, $\left|\widehat{f}(\gamma)\right| \leq \sqrt{2^{-d}+\varepsilon}$ for every $\gamma \in \mathbb{F}_2^n$. □

Remark 2.6.4. The functions constructed in the proof of Theorem 2.3.3 show that the bound in Lemma 2.6.3 is tight. Indeed, these function have the property that every subcube of dimension $n-d$ contains exactly one nonzero Fourier coefficient of weight $2^{-d/2}$. Thus both of the above inequalities are tight for such functions.

Proof of Theorem 2.3.1. Fix $d \geq 3$. By Lemma 2.6.3, if $f : \mathbb{F}_2^n \to \{1, -1\}$ satisfies $INF(d, \frac{2}{25} - 2^{-d-1})$ (note that $\frac{2}{25} > \frac{1}{16}$, so this expression is positive when $d \geq 3$), then

$$\left|\widehat{f}(\alpha)\right| < \sqrt{2^{-d} + \left(\frac{4}{25} - 2^{-d}\right)} = \frac{2}{5}$$

for every $\alpha \in \mathbb{F}_2^n$. Hence,

$$\left|\frac{\left|f^{-1}(\{-1\})\right|}{2^n} - \frac{1}{2}\right| = \frac{1}{2}\left|\widehat{f}(0)\right| < \frac{1}{5}$$

and f is weakly balanced.

For the second part of the Theorem, consider the function $f : \mathbb{F}_2^2 \to \{1, -1\}$ which is -1 if and only if its input is 11. One can easily verify that $\mathrm{I}_{10}[f] = \mathrm{I}_{01}[f] = \mathrm{I}_{11}[f] = \frac{1}{2}$, and $\frac{f^{-1}(\{-1\})}{2^2} = \frac{1}{4}$ by construction. $\square$

Chapter 3

Comparison with Extant Theories of quasi-randomness

There are various quasi-randomness theorems for Boolean functions implicitly or explicitly considered in several related works ranging from hypergraphs to analysis of Boolean functions. Typically, these theories capture global properties of a Boolean function while the quasi-random properties defined in Section 2.3 are local. We will discuss an incomplete list of these extant theories and compare them with some of our local quasi-random properties.

3.1 Cycle-Regularity

3.1.1 Roth's Theorem and Fourier Analysis over $\mathbb{Z}/n\mathbb{Z}$.

A collection of pseudo-random properties of Boolean functions appears implicitly in Chung and Graham's work on quasi-random subsets of $\mathbb{Z}/N\mathbb{Z}$ [19]. Their work generalizes the pseudo-random property of small Fourier coefficients appearing in Roth's Theorem. By identifying the set of binary strings with the binary expansions of integers mod n, their work implicitly produces a theory of quasi-random Boolean functions. While both their Theorem and ours use Fourier analysis, it is important to note that Chung and Graham are using Fourier analysis over $\mathbb{Z}/N\mathbb{Z}$, and the Fourier characters of $\mathbb{Z}/N\mathbb{Z}$ are wildly different than those of $\mathbb{F}_2^n$.

To apply their work to Boolean functions, we can identify the set of binary strings with elements of $\mathbb{Z}/2^n\mathbb{Z}$. Then a Boolean function can be identified with the set of elements of $\mathbb{Z}/2^n\mathbb{Z}$

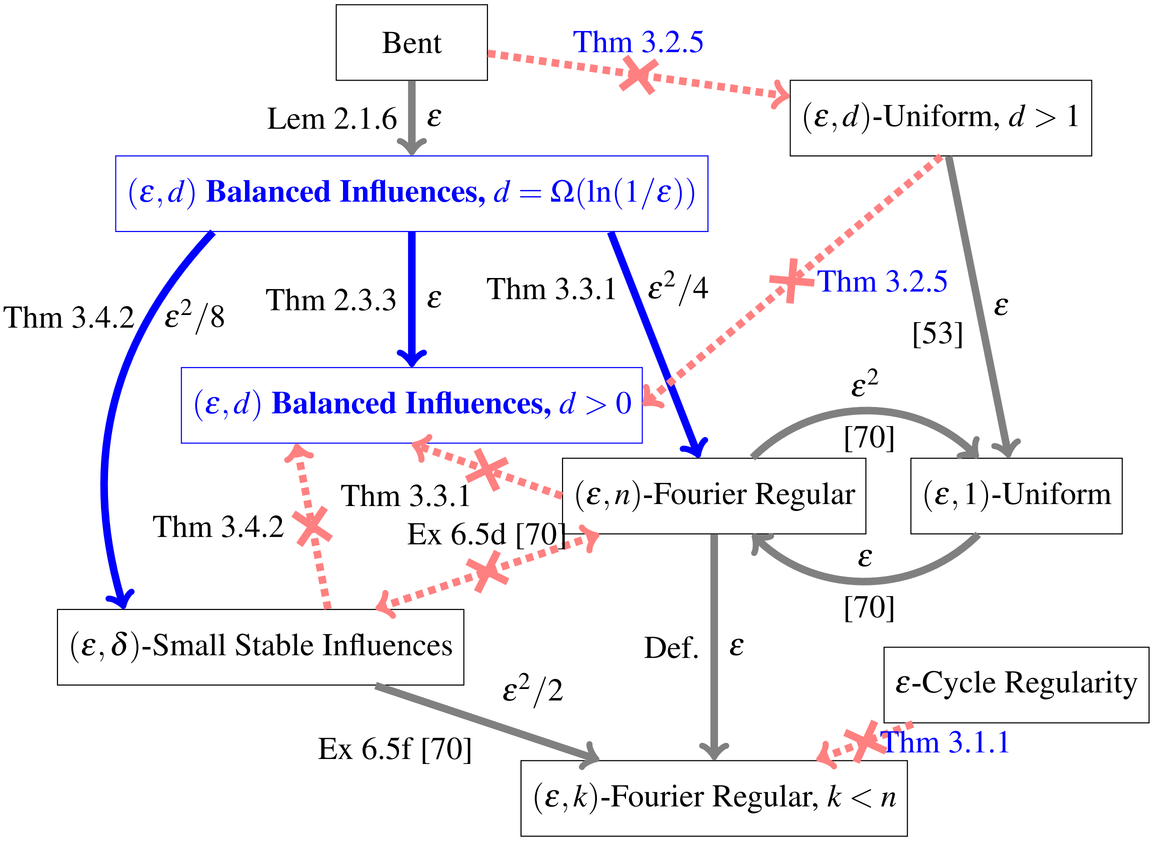
Bent
Thm 3.2.5
(ε,d)-Uniform, $d>1$
Lem 2.1.6
ε
(ε,d) Balanced Influences, $d=\Omega(\ln(1/\varepsilon))$
Thm 3.4.2
$\varepsilon^2/8$
Thm 2.3.3
ε
Thm 3.3.1
$\varepsilon^2/4$
Thm 3.2.5
[53]
ε
(ε,d) Balanced Influences, $d>0$
ε^2
[70]
(ε,n)-Fourier Regular
$(\varepsilon,1)$-Uniform
Thm 3.3.1
Thm 3.4.2
Ex 6.5d [70]
ε
[70]
(ε,δ)-Small Stable Influences
Def.
ε
ε-Cycle Regularity
$\varepsilon^2/2$
Thm 3.1.1
Ex 6.5f [70]
(ε,k)-Fourier Regular, $k<n$

Figure 3.1. The relationships between different theories of quasi-randomness. Each box is a distinct theory of quasi-randomness. Each arrow is a strict implication. Beside each arrow we give a reference to the proof of the implication and the loss function. The results of this thesis are in bold blue text and blue arrows. Non-implications are red dotted lines with an X in the middle, with a citation for each result.

on which it takes the value -1. Their key pseudo-random property is the following:

Property $\mathbf{P_{12}}$. *A Boolean function $f:\mathbb{Z}/2^n\mathbb{Z}\to\{1,-1\}$ is ε-**Cycle Regular** if f has correlation at most ε with all nonzero characters of $\mathbb{Z}/2^n\mathbb{Z}$, i.e., for every nonzero $j\in\mathbb{Z}/2^n\mathbb{Z}$,*

$$\left|\mathbb{E}_{z\in\mathbb{Z}/2^n\mathbb{Z}}\, f(z)\exp(2\pi i jz/2^n)\right| < \varepsilon.$$

As shown by Chung and Graham [19], ε-Cycle Regularity controls the correlations of a function f with a shifted copy of itself much like our Balanced Influences Property Property $\mathbf{P_1}$. However, the arithmetic operations considered in ε-Cycle Regularity are carried out over $\mathbb{Z}/2^n\mathbb{Z}$

rather than $\mathbb{F}_2^n$ as in the Balanced Influences Property.

3.1.2 Relationship between Balanced Influences and Cycle-Regularity

We prove the following theorem:

Theorem 3.1.1. *For any $\delta > 0$ there is a δ-Cycle Regular function which is not $(\varepsilon, k+1)$-Fourier Regular for any $\varepsilon < 1$ where $k = C_0 \ln(1/\delta)$ for some absolute constant C_0.*

The relationship between ε-Cycle Regularity and the other theories is more intricate than our other theories of quasi-randomness, largely due to the algebraic differences between $\mathbb{Z}/2^n\mathbb{Z}$ and $\mathbb{F}_2^n$. As Boolean functions in the sense of ε-Cycle Regularity are not functions on $\mathbb{F}_2^n$, we have the following definition to transfer results between these two theories:

Definition 3.1.2. Given $z \in \mathbb{Z}/2^n\mathbb{Z}$, let $z^* \in \mathbb{F}_2^n$ denote the binary expansion of z, i.e., the vector such that

$$z_i^* = a_i$$

where $z = \sum_{i=1}^n a_i 2^{i-1}$ is the binary expansion of z.

Chung and Graham [[19], Prop. 6.2] prove the following result.

Lemma 3.1.3. *[19] Let $g : \mathbb{Z}/2^n\mathbb{Z} \to \{1, -1\}$ be the function which is -1 if and only its input has an odd number of ones in its binary expansion. There is an absolute constant C such that g is ε-Cycle Regular where $\varepsilon = C \left(\frac{\sqrt{2+\sqrt{2}}}{2} \right)^n \approx 0.92^n$.*

In our notation, the function g considered in Lemma 3.1.3 can written as the composition of the binary expansion function defined in definition Definition 3.1.2 with the Fourier character $\chi_{\mathbf{1}}$. As $\chi_{\mathbf{1}}$ is a Fourier character, $\chi_{\mathbf{1}}$ cannot be (ε, n)-Fourier Regular for any $\varepsilon < 1$. Thus for any $\delta > 0$, δ-Cycle Regularity does not imply (ε, n)-Fourier Regularity for any $\varepsilon < 1$. Here we generalize Lemma 3.1.3 to show that there is a Fourier character χ_γ where $|\gamma|$ is much smaller than n which is ε-Cycle Regular for any $\varepsilon > 0$. As a consequence, we will show that for any

$\delta > 0$, δ-Cycle Regularity cannot even imply (ε, k)-Fourier Regularity for a wide range of $k < n$ and $\varepsilon < 1$.

Proof of Theorem 3.1.1. Set $k = \lceil C_0 \ln(1/\delta) \rceil$ for some absolute constant C_0 to be defined later. Define $S = \{1, \ldots, k\}$. Define $\gamma \in \mathbb{F}_2^n$ by $\gamma := \mathbf{1} \underset{S}{\otimes} 0$ where $\mathbf{1} \in \mathbb{F}_2^S$ is the all-ones vector and $0 \in \mathbb{F}_2^{\overline{S}}$ is the zero vector. We show that χ_γ is δ-Cycle Regular.

Define $\omega_n := \exp\left(\frac{2\pi i}{2^n}\right)$. Now let $c \in \mathbb{Z}/2^n\mathbb{Z} \setminus \{0\}$ be arbitrary, and via the Euclidean algorithm, write $c = 2^{n-k}a + b$ where $0 \le b < 2^{n-k}$. For $z \in \mathbb{Z}_{2^n}$, we write $z^* = y^* \otimes_S x^*$. We then have $\chi_\gamma(z^*) = \chi_{\mathbf{1}}(y^*)\chi_0(x^*) = \chi_{\mathbf{1}}(y^*)$ by the definition of γ. Then,

$$\begin{aligned}
\mathbb{E}_{z \in \mathbb{Z}/2^n\mathbb{Z}}\, \chi_\gamma(z^*)\,\omega_n^{-cz} &= \mathbb{E}_{0 \le y < 2^k}\, \mathbb{E}_{0 \le x < 2^{n-k}}\, \chi_\gamma(y^* \underset{S}{\otimes} x^*)\,\omega_n^{-(2^{n-k}a+b)(2^kx+y)} \\
&= \mathbb{E}_{0 \le y < 2^k}\, \mathbb{E}_{0 \le x < 2^{n-k}}\, \chi_{\mathbf{1}}(y^*)\,\omega_n^{-2^nax-2^kxb-2^{n-k}ay-by} \\
&= \mathbb{E}_{0 \le y < 2^k}\, \mathbb{E}_{0 \le x < 2^{n-k}}\, \chi_{\mathbf{1}}(y^*)\,\omega_n^{-2^kxb-2^{n-k}ay-by} \\
&= \mathbb{E}_{0 \le y < 2^k}\, \chi_{\mathbf{1}}(y^*)\,\omega_n^{-2^{n-k}ay-by}\, \mathbb{E}_{0 \le x < 2^{n-k}}\, \omega_n^{-2^kxb} \\
&= \mathbb{E}_{0 \le y < 2^k}\, \chi_{\mathbf{1}}(y^*)\,\omega_n^{-2^{n-k}ay-by}\, \mathbb{E}_{0 \le x < 2^{n-k}}\, \omega_{n-k}^{-xb} \\
&= \begin{cases} 0 & b \neq 0 \\ \mathbb{E}_{0 \le y < 2^k}\, \chi_{\mathbf{1}}(y^*)\,\omega_n^{-2^{n-k}ay} & b = 0 \end{cases} \\
&= \begin{cases} 0 & b \neq 0 \\ \mathbb{E}_{0 \le y < 2^k}\, \chi_{\mathbf{1}}(y^*)\,\omega_k^{ay} & b = 0 \end{cases}
\end{aligned}$$

Observe that $\chi_{\mathbf{1}}(y^*)$ is precisely the function considered in Lemma 3.1.3 on the group $\mathbb{Z}_{2^k}$. Hence, we may apply Lemma 3.1.3 to conclude that

$$
\begin{aligned}
\left|\mathbb{E}_{z\in\mathbb{Z}/2^n\mathbb{Z}}\chi_\gamma(z^*)\omega_n^{-cz}\right| &\leq C\left(\frac{\sqrt{2+\sqrt{2}}}{2}\right)^k \\
&\leq C\left(\frac{\sqrt{2+\sqrt{2}}}{2}\right)^{-C_0\ln(1/\delta)} \\
&\leq \delta
\end{aligned}
$$

where C and C_0 are sufficiently large absolute constants. Thus $z \to \chi_\gamma(z^*)$ is δ-Cycle Regular. However, $|\gamma| = k$, and so χ_γ cannot be $(\varepsilon, k+1)$-Fourier Regular for any $\varepsilon < 1$. Thus δ-Cycle Regularity does not imply $(\varepsilon, k+1)$-Fourier Regularity for any $\varepsilon < 1$. □

3.2 The Gowers Norms and (ε, k)-Uniformity

3.2.1 Gowers Norms and $\mathbb{F}_2$-Degree

The aforementioned Semeredi's Theorem proved that any subset S of the integers such that $\lim_{n\to\infty}\frac{|S\cap[n]|}{n}$ was positive contained arbitrarily long arithmetic progressions. Szemeredi's original proof is famously intricate, whereas Roth's theorem is in many ways the prototypical example of a "structure vs pseudo-randomness" argument. However, Roth's Theorem fails to capture progression of length 4 or higher. Gowers [48], extended Roth's argument to arbitrarily long progressions using the quasi-randomness found in Chung and Graham's work on hypergraphs. The same ideas appear in Chung and Tetali's work on the relationship between k-uniform hypergraphs and Boolean functions in [23] and in papers by Castro-Silva [11, 12]. These works convert a Boolean function to a k-uniform hypergraph via the following construction. Given a Boolean function $f : \mathbb{F}_2^n \to \{1, -1\}$, its *Cayley hypergraph H* has the vertex set $\mathbb{F}_2^n$ and hyper-edges $\{x_1, \ldots, x_k\} \in E(H) \iff f(x_1 + \cdots + x_k) = -1$. In the hypergraph setting, one considers the *deviation* in [15, 23, 24] which counts even and odd partial octahedra in k-uniform hypergraphs.

By translating the definitions of deviation and the Cayley hypergraph, we arrive at the main definition for our purposes:

Definition 3.2.1. For $k \geq 1$, the *k-th Gowers uniformity norm* of a function $f : \mathbb{F}_2^n \to \{1,-1\}$, denoted $\|f\|_{U(k)}$, is defined as

$$\|f\|_{U(k)} := \left(\mathbb{E}_{x \in \mathbb{F}_2^n} \mathbb{E}_{v_1,\ldots,v_k \in \mathbb{F}_2^n} \prod_{\alpha_1,\ldots,\alpha_k \in \{0,1\}} f(x + \alpha_1 v_1 + \cdots + \alpha_k v_k) \right)^{2^{-k}}$$

We will typically use the following equivalent formula

$$\|f\|_{U(k)} = \left(\mathbb{E}_{x \in \mathbb{F}_2^n, M \in \mathbb{F}_2^{n \times k}} \prod_{v \in \mathbb{F}_2^k} f(x + Mv) \right)^{2^{-k}}$$

(see [53]).

For these theories, the key pseudo-random property is the following:

Property $\mathbf{P_{13}}$. *A Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ is (ε, d)-**Uniform** if*

$$\|f\|_{U(d+1)} < \varepsilon$$

As shown in Castro-Silva's monograph [11], $(\varepsilon, k+1)$-Uniformity implies (ε, k)-Uniformity with no loss in ε, and the implication is strict. Hence, just as we have a hierarchy of quasi-random properties in our Theorem 2.3.3, we can view (ε, k)-Uniformity as a similar hierarchy indexed by k. As shown in [53], the $k+1$-st Gowers norm controls correlation of f with functions of $\mathbb{F}_2$-degree at most k (see Section 2.1.2 for the definition of $\mathbb{F}_2$-degree).

Remark 3.2.2. Directional derivatives provide a third means of defining the Gowers uniformity norms [53], so one might then think that the Equidistributed Derivatives Property Property $\mathbf{P_6}$ will have a close relationship with (δ, k)-Uniformity. However, the Equidistributed Derivatives Property only considers derivatives along vectors of Hamming weight at most k, whereas the

Gowers uniformity norms consider all possible directional derivatives. As we shall see in the proof of Theorem 3.2.5 below, the Spectral Discrepancy Property Property $\mathbf{P_2}$ is more applicable in comparing our work and the theory of (ε, k)-Uniformity.

3.2.2 Lemmas

To connect our work on balanced influences to (ε, k)-Uniformity, we will need two lemmas.

Lemma 3.2.3. *For even n, there is a function $f : \mathbb{F}_2^n \to \{1, -1\}$ which has $INF(d, \varepsilon)$ for every $d \leq n$ and $\varepsilon > 0$, but $\|f\|_{U(3)} = 1$.*

Proof. We consider the Inner Product function $IP(x) : \mathbb{F}_2^n \to \{1, -1\}$ defined in Definition 2.1.7. As shown in Lemma 2.1.9, IP is a bent function and therefore $\left|\widehat{IP}(\gamma)\right| = 2^{-n/2}$ for every $\gamma \in \mathbb{F}_2^n$. By Proposition Proposition 2.1.6 and Lemma 2.1.13. IP has the property $INF(d, \varepsilon)$ for every $1 \leq d \leq n$ and $\varepsilon > 0$. However, IP has $\mathbb{F}_2$-Degree 2. Since $\|g\|_{U(d+1)} = 1$ if g has $\mathbb{F}_2$-degree d (see [53]), we conclude that $\|IP\|_{U(3)} = 1$. □

Lemma 3.2.4. *Let $g : \mathbb{F}_2^n \to \{1, -1\}$ be a Boolean function. Let $M \in \mathbb{F}_2^{n\times(n+1)}$ be the projection matrix which sends $x \in \mathbb{F}_2^{n+1}$ to its first n coordinates, and let $w \in \mathbb{F}_2^{n+1}$ be the vector with a single* 1 *in the $n+1$st coordinate. Let $f : \mathbb{F}_2^{n+1} \to \{1, -1\}$ be defined by $f(x) = g(Mx)$. If g is (ε, k)-Uniform, then*

- *f is (ε, k)-Uniform*
- $\mathrm{I}_w[f] = 0$.

Proof. We first show that f is (ε,k)-Uniform. To that end, we have

$$\begin{aligned}
\|f\|_{U(k)} &= \left(\mathbb{E}_{x\in\mathbb{F}_2^{n+1}}\,\mathbb{E}_{N\in\mathbb{F}_2^{(n+1)\times k}}\prod_{v\in\mathbb{F}_2^k} f(x+Nv)\right)^{2^{-k}}\\
&= \left(\mathbb{E}_{x\in\mathbb{F}_2^{n+1}}\,\mathbb{E}_{N\in\mathbb{F}_2^{(n+1)\times k}}\prod_{v\in\mathbb{F}_2^k} g(M(x+Nv))\right)^{2^{-k}}\\
&= \left(\mathbb{E}_{x\in\mathbb{F}_2^{n+1}}\,\mathbb{E}_{N\in\mathbb{F}_2^{(n+1)\times k}}\prod_{v\in\mathbb{F}_2^k} g(Mx+MNv)\right)^{2^{-k}}
\end{aligned}$$

We write $y=Mx$ and $P=MN$. Since M is a projection matrix and x is uniformly distributed on $\mathbb{F}_2^{n+1}$, y is uniformly distributed on $\mathbb{F}^n$. Similarly, P is a uniformly distributed matrix in $\mathbb{F}_2^{n\times k}$. Hence,

$$\begin{aligned}
\|f\|_{U(k)} &= \left(\mathbb{E}_{y\in\mathbb{F}_2^{n}}\,\mathbb{E}_{P\in\mathbb{F}_2^{n\times k}}\prod_{v\in\mathbb{F}_2^k} g(y+Pv)\right)^{2^{-k}}\\
&= \|g\|_{U(k)}\\
&\le \varepsilon \qquad (3.1)
\end{aligned}$$

where we use our assumption on g in Equation (3.1)). Thus f is (ε,k)-Uniform.

For the second claim, we observe that $f(x+w)=f(x)$ for every x. Therefore, $\mathrm{I}_w[f]=0$. □

3.2.3 Relationship between Balanced Influences and (ε,k)-Uniformity

We show the following theorem:

Theorem 3.2.5. *For any $\varepsilon>0$, a Boolean function $f:\mathbb{F}_2^n\to\{1,-1\}$ with (δ,d)-Balanced Influences is also $(\varepsilon,1)$-Uniform by setting $\delta=\frac{\varepsilon^4}{4}$ and $d\ge 1+\lceil\frac{4\ln(1/\varepsilon)}{\ln(2)}\rceil$.*

Furthermore, (ε,d)-Balanced Influences and (ε,k)-Uniformity are incomparable for any $d\le n$ and $k\ge 2$. More precisely, *(1) There is a function $f : \mathbb{F}_2^n \to \{1,-1\}$ with (δ,n)-Balanced Influences for any $\delta > 0$, yet f is not $(\varepsilon,2)$-Uniform for any $\varepsilon < 1$.*

(2) For any $k \geq 2$ and any $\delta > 0$, there is a function $g : \mathbb{F}_2^n \to \{1,-1\}$ which is (δ,k)-Uniform and does not have $INF(d,\varepsilon)$ for any rank $d \geq 1$ or $\varepsilon < \frac{1}{2}$.

Proof. We have three claims to prove. First, we consider the relationship between (ε,n) Balanced Influences and $(\varepsilon,1)$-Uniformity. Let $f : \mathbb{F}_2^n \to \{1,-1\}$ satisfy $INF(d,\varepsilon^4/4)$ where $d \geq 1 + \lceil \frac{4\ln(1/\varepsilon)}{\ln(2)} \rceil$. By Lemma 2.6.3, f is $(\sqrt{2^{-d}+\varepsilon^4/2},n)$-Fourier Regular. Using the bound on d, we find that

$$2^{-d} + \varepsilon^4/4 \leq \frac{1}{2}\exp\big(-4\ln(1/\varepsilon)\big) + \varepsilon^4/2 = \varepsilon^4 \tag{3.2}$$

Thus f is (ε^2,n)-Fourier Regular. By Proposition 6.7 in O'Donnell's book [70], $(\sqrt{\varepsilon},n)$-Fourier Regularity implies $(\varepsilon,1)$-Uniformity. Thus, f is $(\varepsilon,1)$-Uniform.

Next we show that (ε,k)-Balanced Influences is incomparable with (ε,d)-Uniformity for $d > 1$ and any k. Lemma 3.2.3 provides a function f which possesses $INF(k,\varepsilon)$ for any $k \in \mathbb{N}$ with $1 \leq k \leq n$ and any $\varepsilon > 0$ yet has $\|f\|_{U(3)} = 1$.

Now we can show that (ε,d)-Uniformity cannot imply (ε,k)-Balanced Influences for any $k \geq 1$. Let $g : \mathbb{F}_2^n \to \{1,-1\}$ be a uniformly random Boolean function. For any $\varepsilon > 0$, there is n sufficiently large such that g is (ε,k)-Uniform. By Lemma 3.2.4 if $f : \mathbb{F}_2^{n+1} \to \{1,-1\}$ is g composed with a projection matrix, f is (ε,k)-Uniform yet there is a vector $w \in \mathbb{F}_2^{n+1}$ such that $\mathrm{I}_w[f] = 0$ and $|w| = 1$. Thus, f cannot have the Balanced Influences Property $INF(k,\varepsilon)$ for any $k \geq 1$ and $\varepsilon < \frac{1}{2}$.

It follows that (ε,k)-Uniformity and quasi-randomness of rank d with error ε are incomparable for $k \geq 2$ and $d \geq 1$. □

3.3 Fourier Expansion

3.3.1 History

O'Donnell presents several pseudo-random properties in [70] which center on the Fourier expansion defined in Section 2.1.1. In this section and the next we will treat the relationship of our work to his. The relevant properties date back to the work of Siegenthaler on cryptography [78]. Seigenthaler observed that a random-like cipher function $f : \mathbb{F}_2^n \to \{0,1\}$ has little to no weight on Fourier coefficients $\widehat{f}(\gamma)$ where $|\gamma| \leq k$. This naturally leads to a notion of a pseudo-random Boolean function.

It is easy to observe that *every* Fourier coefficient of a random Boolean function is small with high probability. We can combine these two ideas with the following definition, which also appears in a work of Xiao and Massey [87]:

Property $\mathbf{P_{14}}$. *A Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ is (ε,d)-**Fourier Regular** if*

$$\left|\widehat{f}(\gamma)\right| < \varepsilon$$

for every $\gamma \in \mathbb{F}_2^n$ with $|\gamma| \leq d$.

By definition, $(\varepsilon, d+1)$-Fourier Regularity ε-implies (ε,d)-Fourier Regularity, and a Fourier character χ_γ where $|\gamma| = d+1$ shows that the implication is strict. Hence, just as we have a hierarchy of quasi-random properties in our Theorem 2.3.3, (ε,k)-Fourier Regularity can be viewed as forming an increasing hierarchy of pseudo-random properties indexed by k. Furthermore, (ε,n)-Fourier Regularity and $(\varepsilon,1)$-Uniformity are equivalent as is shown in Proposition 6.7 of [70].

O'Donnell notes that the (ε,k)-Fourier Regularity is equivalent to several combinatorial notions for small values of k. For instance, an (ε,k)-Fourier Regular function has small correlation with *k-juntas*, i.e., functions whose output is determined by at most k inputs. The support sets of an (ε,k)-Fourier Regular Boolean function can be used to produce a k-wise independent distribution on $\mathbb{F}_2^n$, i.e., a distribution such any choice of k coordinates are mutually independent (although larger sets might not be independent).

3.3.2 Relationship between Balanced Influences and (ε, k)-Regularity

As for the relationship between (ε,k)-Fourier regularity and our properties, we show the following theorem:

Theorem 3.3.1. *For any $\varepsilon > 0$, a Boolean Function $f : \mathbb{F}_2^n \to \{1,-1\}$ with (δ,d)-Balanced Influences is also (ε,k)-Fourier Regular for any $k \le n$ by setting $\delta = \frac{\varepsilon^2}{4}$ and $d = 1 + \lceil \frac{2\ln(1/\varepsilon)}{\ln(2)} \rceil$.*

Conversely, for any $\delta > 0$ there is a function which is (δ,n)-Fourier Regular which does not have (ε,k) Balanced Influences for any rank $k \ge 1$ or error bound $\varepsilon < \frac{1}{2}$.

Proof. Assume that f satisfies $INF(d,\varepsilon^2/2)$ where $d \ge 1 + \lceil \frac{2\ln(1/\varepsilon)}{\ln(2)} \rceil$. Lemma 2.6.3 implies that if $f : \mathbb{F}_2^n \to \{1,-1\}$ has $INF(d,\varepsilon)$, then f is also $(2\sqrt{2^{-d} + \varepsilon^2/2}, n)$-Fourier Regular. By the bound on d,

$$2^{-d} + \varepsilon^2/2 \le \frac{1}{2}\exp\left(-2\ln(1/\varepsilon)\right) + \varepsilon^2/2 = \varepsilon^2$$

Thus, f is (ε,n)-Fourier Regular. If a function g is (δ,k)-Fourier Regular then g is also $(\delta,k-1)$-Fourier Regular by definition. Hence, if $f : \mathbb{F}_2^n \to \{1,-1\}$ has $(\varepsilon^2/2,d)$-Balanced Influences then f is (ε,k)-Fourier Regular for any $k \le n$.

For the second claim, we must show that (ε,k)-Fourier Regularity cannot imply (ε,d)-Balanced Influences for any $k \le n$, $d \ge 1$ or $\varepsilon < 1$. Consider the inner product function $IP : \mathbb{F}_2^{2n} \to \{1,-1\}$ defined in Example Equation (2.1). By applying Lemma 3.2.4 to IP, we find a function $f : \mathbb{F}_2^{2n+1} \to \{1,-1\}$ which is $(2^{-n/2},n)$-Fourier Regular and yet does not have $INF(d,\varepsilon)$ for any $d \ge 1$ and $\varepsilon < \frac{1}{2}$. As (ε,n)-Fourier Regularity implies (ε,k)-Fourier Regularity for $k < n$, IP is (ε,k)-Fourier Regular for any $k \le n$. It follows that (ε,k)-Fourier Regularity does not imply $INF(d,\varepsilon)$ for any choice of $k \le n$, $d \ge 1$, and $\varepsilon < \frac{1}{2}$. □

3.4 Small Stable Influences

3.4.1 Stable Influences

The influences of a Boolean function have These functions have been of great interest in the context of thresholds in random structures thanks to Friedgut's theorem [38] and the KKL theorem on voting rules. O'Donnell gives four different notions of pseudo-randomness for Boolean functions based on influences.

$$\begin{tikzcd}
 & \text{Small Influences} \arrow[dl, "{\text{[70], Ex 6.5(a) } \varepsilon^2}"'] \arrow[dr, "{\text{[70], Ex 6.5(d) } \varepsilon}"] & \\
\text{Regular} \arrow[dr, "{\text{[70], Ex 6.5(f) } \varepsilon}"'] & & \text{Small Stable Influences} \arrow[dl, "{\text{[70], Def 6.11 } \varepsilon}"] \\
 & \text{Low-Degree Regular} &
\end{tikzcd}$$

Figure 3.2. The various pseudo-random properties considered in O'Donnell's work.

Here, small influences is equivalent to the property that $\mathrm{I}_\gamma[f] < \varepsilon$ for each vector $\gamma \in \mathbb{F}_2^n$ of Hamming weight 1, low-degree regularity and regularity are the same as the (ε, k)-Fourier regularity from Section 3.3. Curiously, a random function does not have small influences as we noted in defining Property $\mathbf{P_1}$ and O'Donnell himself notes [70]. Thus Small Influences are not a true pseudo-random property for an arbitrary function. Stable influences capture the idea that individual coordinates are unimportant while still applying to random functions. Stable influences are defined as follows.

Definition 3.4.1. For a coordinate i and a parameter $\rho \in [0,1]$, the **ρ-stable influence** of a

Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ is

$$\mathrm{I}_i^{\rho}[f] := \sum_{\substack{\gamma \in \mathbb{F}_2^n \\ \gamma_i = 1}} \rho^{|\gamma|-1} \widehat{f}(\gamma)^2 .$$

Stable influences have been of great importance in computer science ever since the KKL Theorem [57] on the existence, and much of O'Donnell's test is focused on functions which small stable influences [70].

The key pseudo-random property is:

Property $\mathbf{P_{15}}$. *A Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ has (ε,ρ)-**Small Stable Influences** for some $\varepsilon \in \mathbb{R}_{\geq 0}$ and $\rho \in [0,1]$ if*

$$\mathrm{I}_i^{1-\rho}[f] < \varepsilon$$

for every $i \in [n]$.

As shown by O'Donnell [70], ρ-Small Stable Influences measure the expected change in the function if the input bits are changed via a particular noise model. Thus, (ε,ρ) Small Stable Influences implies a form of noise stability [67].

Much of O'Donnell's work is focused on *monotone Boolean functions* where changing an input bit from 0 to 1 cannot make the function change from true to false. These are precisely the functions considered in Freidgut's Theorem and are of great interest in applications in Hardness of Approximation [52,59,67,68]. All four notions of influence are pseudo-random properties for monotone functions, and are equivalent in that setting [70].

3.4.2 Relationship between Balanced Influences and Small Stable Influences

We show the following theorem:

Theorem 3.4.2. *For any $\varepsilon > 0$ and $1 > \rho \geq 2 - \sqrt{2}$, a Boolean function $f : \mathbb{F}_2^n \to \{1,-1\}$ with (δ,d)-Balanced Influences also has (ε,ρ)-Small Stable Influences by setting $\delta = \frac{\varepsilon^2}{8}$ and*

$d = \lceil \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2-\rho)} \rceil$.

Conversely, there is a function which has $((1-\delta)^{n-1}, \delta)$*-Small Stable Influences for any* $\delta < 1$ *but does not have* (ε, k)*-Balanced Influences for any* k *and any* $\varepsilon < \frac{1}{2}$.

Roughly speaking, (ε, ρ)-Small Stable Influences indicates that a function has little Fourier weight on vectors of small Hamming weight, whereas (ε, d)-Balanced Influences indicates that the Fourier weight is spread evenly over all of $\mathbb{F}_2^n$.

Proof of Theorem 3.4.2. Assume f satisfies $INF(d, \varepsilon^2/8)$ for $d = \lceil \frac{\ln\left(2/\varepsilon\right)}{\ln(2-\rho)} \rceil$. By Theorem 2.4.1, f also satisfies $SD(d, \varepsilon^2/4)$ for any $d \geq \lceil \frac{\ln(2/\varepsilon)}{\ln(2-\rho)} \rceil$.

Recall that $1 > \rho \geq 2 - \sqrt{2} \approx 0.58$. We want to show that $\mathrm{I}_i^{1-\rho}[f] < \varepsilon$ for each $i \in [n]$ via the Spectral Discrepancy Property. We observe that the set of $\gamma \in \mathbb{F}_2^n$ with $\gamma_i = 1$ is precisely the $n-1$-dimensional subcube $C(\overline{\{i\}}, 1)$, and the same subcube may be divided into 2^{d-1} subcubes of dimension $n-d$ as follows. Pick a set S of size d which contains i. Then, $C(\overline{\{i\}}, 1) = \bigsqcup_{\substack{z \in \mathbb{F}_2^S \\ z_i = 1}} C(\overline{S}, z)$. Therefore,

$$\begin{aligned}
\mathrm{I}_i^{1-\rho}[f] &= \sum_{\substack{\gamma \in \mathbb{F}_2^n \\ \gamma_i = 1}} (1-\rho)^{|\gamma|-1} \widehat{f}(\gamma)^2 \\
&= \sum_{\substack{z \in \mathbb{F}_2^S \\ z_i = 1}} \left(\sum_{\gamma \in C(\overline{S}, z)} (1-\rho)^{|\gamma|-1} \widehat{f}(\gamma)^2 \right) \\
&\leq \sum_{\substack{z \in \mathbb{F}_2^S \\ z_i = 1}} \left(\max_{\gamma \in C(\overline{S}, z)} (1-\rho)^{|\gamma|-1} \right) \left(\sum_{\gamma \in C(S, z)} \widehat{f}(\gamma)^2 \right) \\
&\leq \left(2^{-d} + \varepsilon^2/4\right) \sum_{\substack{z \in \mathbb{F}_2^S \\ z_i = 1}} \left(\max_{\gamma \in C(\overline{S}, z)} (1-\rho)^{|\gamma|-1} \right)
\end{aligned}$$

where we use $SD(d, \varepsilon^2/4)$ in the ultimate line. Now we can simplify further:

$$\begin{aligned}
\mathrm{I}_i^{1-\rho}[f] &\leq \left(2^{-d} + \varepsilon^2/4\right) \sum_{\substack{z \in \mathbb{F}_2^S \\ z_i = 1}} (1-\rho)^{|z|-1} \\
&= \left(2^{-d} + \varepsilon^2/4\right) \left(\sum_{j=0}^{d-1} \binom{d-1}{j} (1-\rho)^j \right) \\
&= \left(2^{-d} + \varepsilon^2/4\right) (2-\rho)^{d-1}
\end{aligned}$$

Since $d = \lceil \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2-\rho)} \rceil$, we have that $d \leq \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2-\rho)} + 1$. As $\rho < 1$, we have

$$(2-\rho)^{d-1} \leq (2-\rho)^{\frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2-\rho)}} = \frac{2}{\varepsilon}$$

Since $\rho \geq 2 - \sqrt{2}$, $d = \lceil \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2-\rho)} \rceil \geq \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2-\rho)} \geq \frac{2\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2)}$. Therefore,

$$2^{-d} \leq 2^{-\frac{2\ln\left(\frac{2}{\varepsilon}\right)}{\ln(2)}} = \frac{\varepsilon^2}{4}$$

Thus,

$$\mathrm{I}_i^{1-\rho}[f] \leq \left(\frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{4}\right) \frac{2}{\varepsilon} = \varepsilon$$

as desired.

Conversely, one can easily verify that $\chi_{\mathbf{1}}$ has $((1-\rho)^{n-1}, \rho)$-Small Stable Influences, but $\mathrm{I}_\gamma[\chi_{\mathbf{1}}] = 1$ for every $\gamma \in \mathbb{F}_2^n$ with Hamming weight 1. Thus $\chi_{\mathbf{1}}$ does not have $INF(d, \varepsilon)$ for any $d \geq 1$ unless $\varepsilon = \frac{1}{2}$. □

Remark 3.4.3. Chapter 3, in full, is a reprint of the material as it appears in Quasi-random Boolean Functions, Fan Chung and Nicholas Sieger, which is in review at the Electronic Journal of Combinatorics. The dissertation author was the primary investigator and author of this paper.

Bibliography

[1] Dimitris Achlioptas and Ehud Friedgut. A Sharp Threshold for k-Colorability. *Random Structures & Algorithms*, 14(1):63–70, 1999. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291098-2418%281999010%2914%3A1%3C63%3A%3AAID-RSA3%3E3.0.CO%3B2-7.

[2] Ron Aharoni, Matt DeVos, Sebastián González Hermosillo de la Maza, Amanda Montejano, and Robert Šámal. A rainbow version of Mantel's Theorem. *arXiv preprint arXiv:1812.11872*, 2018.

[3] N Alon and T.H. Marshall. Homomorphisms of Edge-Colored Graphs and Coxeter Groups. Technical report, 1998. Publication Title: Journal of Algebraic Combinatorics Volume: 8.

[4] Noga Alon. Explicit Ramsey graphs and orthonormal labelings. *The Electronic Journal of Combinatorics*, 1(1):R12–R12, October 1994. Publisher: The Electronic Journal of Combinatorics.

[5] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. John Wiley & Sons, January 2016. Google-Books-ID: Iu5RCgAAQBAJ.

[6] John Banzhaf. Weighted Voting doesn't Work: A mathematical analysis. *Rugers Law Review*, 19(2):317–343, 1965.

[7] Michael Ben-Or and Nathan Linial. Collective coin flipping, robust voting schemes and minima of Banzhaf values. In *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*, pages 408–416, October 1985. ISSN: 0272-5428.

[8] Eli Ben-Sasson, Madhu Sudan, Salil Vadhan, and Avi Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, STOC '03, pages 612–621, New York, NY, USA, June 2003. Association for Computing Machinery.

[9] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. Robust PCPs of Proximity, Shorter PCPs, and Applications to Coding. *SIAM Journal on Computing*, 36(4):889–974, January 2006. Publisher: Society for Industrial and Applied Mathematics.

[10] Matija Bucić, Eoin Long, Asaf Shapira, and Benny Sudakov. Tournament Quasirandomness from Local Counting. *Combinatorica*, 41(2):175–208, April 2021.

[11] Davi Castro-Silva. Quasirandomness in additive groups and hypergraphs, 2021.

[12] Davi Castro-Silva. Quasirandom additive sets and Cayley hypergraphs, May 2023. arXiv:2305.03012 [math].

[13] Timothy F. N. Chan, Daniel Král', Jonathan A. Noel, Yanitsa Pehova, Maryam Sharifzadeh, and Jan Volec. Characterization of quasirandom permutations by a pattern sum. *Random Structures & Algorithms*, 57(4):920–939, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20956.

[14] F R K Chung and R L Graham. Quasi-random hypergraphs. *Random Structures and Algorithms*, 1(1):105–124, 1990.

[15] F. R. K. Chung and R. L. Graham. Quasi-Random Set Systems. *Journal of the American Mathematical Society*, 4(1):151, January 1991. Publisher: JSTOR.

[16] F R K Chung and R L Graham. Quasi-random tournaments. *Journal of Graph Theory*, 15(2):173–198, November 1991. Publisher: John Wiley & Sons, Ltd.

[17] F. R. K. Chung and R. L. Graham. Cohomological aspects of hypergraphs. *Transactions of the American Mathematical Society*, 334(1):365–388, January 1992. Publisher: American Mathematical Society (AMS).

[18] F R K Chung and R L Graham. Cohomological aspects of hypergraphs. 334(1), 1992.

[19] F R K Chung and R L Graham. Quasi-random subsets of Z_n. *Journal of Combinatorial Theory Series A*, 61(1):64–86, November 1992. Publisher: Academic Press, Inc. PUB20 Orlando, FL, USA.

[20] F. R.K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, December 1989. Publisher: Springer-Verlag.

[21] Fan Chung and Ronald Graham. Sparse Quasi-Random Graphs. *Combinatorica*, 22(2):217–244, April 2002.

[22] Fan R K Chung. Regularity lemmas for hypergraphs and quasi-randomness. *Random Structures and Algorithms*, 2(2):241–252, 1991.

[23] Fan R K Chung and Prasad Tetali. Communication complexity and quasi-randomness. *SIAM Journal on Discrete Mathematics*, 6(1):110–123, 1993. Publisher: SIAM.

[24] Fan R.K. Chung. Quasi-random classes of hypergraphs. *Random Structures & Algorithms*, 1(4):363–382, 1990.

[25] James Samuel Coleman. Control of Collectivities and the Power of a Collectivity to Act. Technical report, RAND Corporation, August 1968.

[26] David Conlon, Hiêp Hàn, Yury Person, and Mathias Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures & Algorithms*, 40(1):1–38, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20389.

[27] Joshua Cooper and Aaron Dutle. Spectra of uniform hypergraphs. *Linear Algebra and Its Applications*, 436(9):3268–3292, May 2012.

[28] Joshua N Cooper. Quasirandom permutations. *Journal of Combinatorial Theory, Series A*, 106:123–143, 2004.

[29] Irit Dinur. The PCP theorem by gap amplification. *Journal of the ACM*, 54(3):12–es, June 2007.

[30] Sean Eberhard, Freddie Manners, and Rudi Mrazović. Transversals in quasirandom latin squares. *Proceedings of the London Mathematical Society*, 127(1):84–115, 2023. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1112/plms.12538.

[31] Paul Erdos and András Hajnal. On spanned subgraphs of graphs. *Graphentheorie und Ihre Anwendungen (Oberhof, 1977), www. renyi. hu/p erdos/1977-19. pdf*, 1977.

[32] P. Erdös. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.

[33] P. Erdős and Vera T. Sós. On Ramsey—Turán type theorems for hypergraphs. *Combinatorica*, 2(3):289–295, September 1982.

[34] Dvir Falik and Alex Samorodnitsky. Edge-Isoperimetric Inequalities and Influences. *Combinatorics, Probability and Computing*, 16(5):693–712, September 2007.

[35] Réjane Forrié. The Strict Avalanche Criterion: Spectral Properties of Boolean Functions and an Extended Definition. In Shafi Goldwasser, editor, *Advances in Cryptology — CRYPTO' 88*, pages 450–468, New York, NY, 1990. Springer New York.

[36] P. Frankl and V. Rödl. The Uniformity Lemma for hypergraphs. *Graphs and Combinatorics*, 8(4):309–312, December 1992. Publisher: Springer-Verlag.

[37] Ehud Friedgut. Hunting for sharp thresholds. *Random Structures & Algorithms*, 26(1-2):37–51, 2005. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20042.

[38] Ehud Friedgut and Appendix By Jean Bourgain. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American Mathematical Society*, 12(4):1017–1054, May 1999.

[39] K. Friedl and M. Sudan. Some improvements to total degree tests. pages 0190–0190. IEEE Computer Society, January 1995.

[40] Joel Friedman and Avi Wigderson. On the second eigenvalue of hypergraphs. *Combinatorica*, 15(1):43–65, March 1995.

[41] Tobias Friedrich and Ralf Rothenberger. Sharpness of the Satisfiability Threshold for Non-uniform Random k-SAT. In Olaf Beyersdorff and Christoph M. Wintersteiger, editors, *Theory and Applications of Satisfiability Testing – SAT 2018*, pages 273–291, Cham, 2018. Springer International Publishing.

[42] Alan Frieze and Michał Karoński. *Introduction to Random Graphs*. Cambridge University Press, Cambridge, 2015.

[43] Frederik Garbe, Robert Hancock, Jan Hladky, and Maryam Sharifzadeh. Limits of Latin squares. *Discrete Analysis*, July 2023.

[44] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 339–348, October 1996. ISSN: 0272-5428.

[45] W. T. Gowers. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Combinatorics Probability and Computing*, 15(1-2):143–184, January 2006.

[46] W. T. Gowers. Quasirandom groups. *Combinatorics Probability and Computing*, 17(3):363–387, May 2008. arXiv: 0710.3877.

[47] W T Gowers. GENERALIZATIONS OF FOURIER ANALYSIS, AND HOW TO APPLY THEM. *BULLETIN (New Series) OF THE AMERICAN MATHEMATICAL SOCIETY*, 54(1):1–44, 2017.

[48] W.T. Gowers. A new proof of Szemerédi's theorem. *GAFA Geometric And Functional Analysis*, 11(3):465–588, August 2001.

[49] Simon Griffiths. Quasi-Random Oriented Graphs. *Journal of Graph Theory*, 74(2):198–209, October 2013. Publisher: John Wiley & Sons, Ltd.

[50] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. *Essential Coding Theory*. 2019.

[51] Robert Hancock, Adam Kabela, Daniel Král', Taísa Martins, Roberto Parente, Fiona Skerman, and Jan Volec. No additional tournaments are quasirandom-forcing. *European Journal of Combinatorics*, 108:103632, February 2023.

[52] Hamed Hatami. A structure theorem for Boolean functions with small total influences. *Annals of Mathematics*, 176(1):509–533, 2012. Publisher: Annals of Mathematics.

[53] Hamed Hatami, Pooya Hatami, and Shachar Lovett. Higher-order Fourier Analysis and Applications. *Foundations and Trends® in Theoretical Computer Science*, 13(4):247–448, 2019. Publisher: Now Publishers Inc.

[54] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Moreira, Balázs Ráth, and Rudini Menezes Sampaio. Limits of permutation sequences. *Journal of Combinatorial Theory, Series B*, 103(1):93–113, January 2013.

[55] Johan Håstad. Testing of the long code and hardness for clique. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*, pages 11–19, Philadelphia, Pennsylvania, United States, 1996. ACM Press.

[56] Johan Håstad. Clique is hard to approximate within n1-{ϵ}. *Acta Mathematica*, 182(1):105–142, January 1999. Publisher: Institut Mittag-Leffler.

[57] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *[Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science*, pages 68–80, October 1988.

[58] Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. Global hypercontractivity and its applications, 2021. arXiv: 2103.04604.

[59] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, January 2007. Publisher: Society for Industrial and Applied Mathematics.

[60] Y. Kohayakawa and V. Rödl. Szemerédi's Regularity Lemma and Quasi-randomness. In Bruce A. Reed and Cláudia L. Sales, editors, *Recent Advances in Algorithms and Combinatorics*, pages 289–351. Springer New York, New York, NY, 2003.

[61] Yoshiharu Kohayakawa, Brendan Nagle, Vojtěch Rödl, and Mathias Schacht. Weak hypergraph regularity and linear hypergraphs. *Journal of Combinatorial Theory. Series B*, 100(2):151–160, March 2010.

[62] Yoshiharu Kohayakawa, Vojtěch Rödl, and Jozef Skokan. Hypergraphs, Quasi-randomness, and conditions for regularity. *Journal of Combinatorial Theory. Series A*, 97(2):307–352, 2002. Publisher: Academic Press Inc.

[63] Daniel Král' and Oleg Pikhurko. Quasirandom permutations are characterized by 4-point densities. *Geometric and Functional Analysis*, 23(2):570–579, April 2013.

[64] John Lenz and Dhruv Mubayi. EIGENVALUES AND LINEAR QUASIRANDOM HYPERGRAPHS. *Forum of Mathematics, Sigma*, 3:26, January 2015. arXiv: 1208.4863 Publisher: Cambridge University Press.

[65] John Lenz and Dhruv Mubayi. Eigenvalues of non-regular linear quasirandom hypergraphs. *Discrete Mathematics*, 340(2):145–153, February 2017.

[66] László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, Providence, Rhode Island, December 2012.

[67] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 21–30, October 2005. ISSN: 0272-5428.

[68] Elchanan Mossel, Ryan O'Donnell, and Rocco A. Servedio. Learning functions of k relevant variables. *Journal of Computer and System Sciences*, 69(3):421–434, November 2004.

[69] Brendan Nagle, Vojtěch Rödl, and Mathias Schacht. The counting lemma for regular k-uniform hypergraphs. *Random Structures and Algorithms*, 28(2):113–179, November 2006.

[70] Ryan O'Donnell. *Analysis of Boolean Functions*, volume 1. Cambridge University Press, Pittsburgh, PA, 2014. arXiv: 1205.0314.

[71] Michal Parnas, Dana Ron, and Alex Samorodnitsky. Testing Basic Boolean Formulae. *SIAM Journal on Discrete Mathematics*, 16(1):20–46, January 2002. Publisher: Society for Industrial and Applied Mathematics.

[72] L. S. Penrose. The Elementary Statistics of Majority Voting. *Journal of the Royal Statistical Society*, 109(1):53–57, 1946. Publisher: [Wiley, Royal Statistical Society].

[73] K. F. Roth. On Certain Sets of Integers. *Journal of the London Mathematical Society*, s1-28(1):104–109, 1953. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1112/jlms/s1-28.1.104.

[74] O S Rothaus. On "bent" functions. *Journal of Combinatorial Theory, Series A*, 20(3):300–305, May 1976. Publisher: Academic Press.

[75] Ronitt Rubinfeld and Madhu Sudan. Robust Characterizations of Polynomials with Applications to Program Testing. *SIAM Journal on Computing*, 25(2):252–271, April 1996. Publisher: Society for Industrial and Applied Mathematics.

[76] Vojtěch Rödl and Jozef Skokan. Counting subgraphs in quasi-random 4-uniform hypergraphs. *Random Structures and Algorithms*, 26(1-2):160–203, 2005.

[77] Alex Samorodnitsky. Low-degree tests at large distances. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, STOC '07, pages 506–515, New York, NY, USA, June 2007. Association for Computing Machinery.

[78] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications (Corresp.). *IEEE Transactions on Information Theory*, 30(5):776–780, September 1984. Conference Name: IEEE Transactions on Information Theory.

[79] Endre Szemerédi. On sets of integers containing no \(k\) elements in arithmetic progression. *Acta Arithmetica*, 27:199–245, 1975.

[80] Terence Tao. The dichotomy between structure and randomness, arithmetic progressions, and the primes. In *Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30,2006 : invited lectures, Vol. 1, 2006, ISBN 978-3-03719-022-7, págs. 581-608*, pages 581–608, 2006. Section: Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30,2006 : invited lectures.

[81] Terence Tao. Obstructions to Uniformity and Arithmetic Patterns in the Primes. *Pure and Applied Mathematics Quarterly*, 2(2):395–433, March 2006. Publisher: International Press of Boston.

[82] Andrew Thomason. Pseudo-random graphs. *North-Holland Mathematics Studies*, 144(C):307–331, 1987.

[83] Erik Thörnblad. Decomposition of tournament limits. *European Journal of Combinatorics*, 67:96–125, January 2018.

[84] Henry Towsner. {σ}-algebras for quasirandom hypergraphs. *Random Structures & Algorithms*, 50(1):114–139, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20641.

[85] Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1-3):1–336, 2011.

[86] Richard M. Wilson. Cyclotomy and difference families in elementary abelian groups. *Journal of Number Theory*, 4(1):17–47, February 1972.

[87] G.-Z. Xiao and J.L. Massey. A spectral characterization of correlation-immune combining functions. *IEEE Transactions on Information Theory*, 34(3):569–571, May 1988. Conference Name: IEEE Transactions on Information Theory.