# Sidorenko's conjecture, graph norms, and pseudorandomness




Joonkyung Lee

St. Hugh's College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity 2017

# Acknowledgements

I would like to thank my supervisor David Conlon for his invaluable support, encouragement, and inspiration. I have always felt more than lucky to be his DPhil student.

I would like to thank my colleagues, Jeong Han Kim and Choongbum Lee, together with whom I obtained many results in this thesis.

I am also grateful to my College advisor, Tom Sanders, for his warm support. Numerous people in the combinatorics group at Oxford inspired me through our conversations. I do appreciate all of them.

I would like to thank Sang-il Oum, who has constantly given helpful advice since my Masters programme at KAIST.

Lastly, I would like to thank ILJU Scholarship, whose support enabled me to pursue my research at Oxford without financial concerns.

# Abstract

This thesis is primarily concerned with correlation inequalities between the number of homomorphic copies of different graphs. In particular, many of the results relate to a beautiful conjecture of Sidorenko, which roughly states that the number of copies of a bipartite graph $H$ in a graph $G$ is asymptotically minimised when $G$ is the Erdős–Rényi random graph.

The first part of the thesis discusses recent approaches to attack Sidorenko's conjecture. We firstly prove that every graph that admits a special kind of tree decomposition satisfies the conjecture. The proof explicitly uses information theory, which also leads to a general tool for counting fixed graphs that are decomposable in analogous ways. A recursive approach to the conjecture, using Cartesian products of graphs, will also be discussed. We show that the class of graphs that satisfy the conjecture is closed under taking Cartesian products with either trees or even cycles.

The second part studies a conjecture of Kohayakawa, Nagle, Rödl, and Schacht, which states that we have a random-like lower count for any graph $H$ in $G$ whenever $G$ is locally dense, i.e., a fixed edge density is guaranteed for every large enough vertex subset. This conjecture is closely related to Sidorenko's conjecture in the sense that it is true for every bipartite graph $H$ that satisfies Sidorenko's conjecture. We prove a partial converse, stating that if $H$ satisfies the Kohayakawa–Nagle–Rödl–Schacht conjecture, then replacing edges of $H$ by internally disjoint paths of length two gives an instance of Sidorenko's conjecture. We also prove some new instances of the Kohayakawa–Nagle–Rödl–Schacht conjecture by adding extra ideas to the information theoretic approach previously discussed.

The last part obtains a large class of new graph norms and hence gives a partial answer to questions by Lovász and by Hatami. The *$H$-norm* $\|\cdot\|_H$ is a functional on the space of two-variable bounded symmetric measurable functions naturally defined by the (generalised) homomorphism density of $H$. We say that $H$ is *norming* if $\|\cdot\|_H$ is a semi-norm. The *absolute $H$-norm* $\|\cdot\|_{r(H)}$ can also be similarly defined, and $H$ is said to be *weakly norming* if $\|\cdot\|_{r(H)}$ is a norm. The (absolute) $H$-norms have played an important role in the theory of quasirandomness. In particular, the $C_4$-norm corresponds to the Gowers' octahedral norm for graphs.

We demonstrate that any graph whose edges percolate in an appropriate way under the action of a certain natural family of automorphisms is weakly norming. This result allows us to identify a much broader class of weakly norming graphs arising from finite reflection groups, which includes all previously known examples by Hatami and by Lovász. We also include several applications of our results. In particular, we define and compare a number of generalisations of Gowers' octahedral norms.

# Statement of Originality

This thesis does not contain any material which has been accepted in whole, or in part, for any other degree or diploma.

The chapters are based on joint work [28] with Jeong Han Kim and Choongbum Lee, [11, 12] with David Conlon, Jeong Han Kim, and Choongbum Lee, [13] with David Conlon, and the single-authored paper [32]. More specifically, Chapter 3 is based on [11, 12, 13, 32]. Chapter 4 consists of two parts based on [28] and [12], respectively. Chapter 5 is based on [12, 32] and Chapter 6 is based on [13].

# Contents


**1 Introduction** **1**

**2 Preliminaries** **10**

2.1 Graph theory . . . . . 10

2.1.1 Graph homomorphisms . . . . . 11

2.1.2 Spectral graph theory . . . . . 12

2.1.3 Quasirandomness . . . . . 13

2.1.4 Tree decompositions . . . . . 13

2.1.5 Tensor products . . . . . 16

2.1.6 Cartesian products . . . . . 17

2.2 Probability theory . . . . . 18

2.2.1 Conditional independence . . . . . 18

2.2.2 Entropy . . . . . 19

2.2.3 Large deviations . . . . . 20

**3 Information theory and Sidorenko's conjecture** **23**

3.1 Strongly tree-decomposable graphs . . . . . 24

3.1.1 Trees . . . . . 25

3.1.2 Branching random walks starting from a partial embedding . 28

3.1.3 Embedding strongly tree-decomposable graphs . . . . . 30

3.2 $J$-decomposable graphs . . . . . 35

3.2.1 Proof of Theorem 1.4 . . . . . 36

3.2.2 Applications of Theorem 1.4 and Theorem 3.6 . . . . . 39

3.3 Higher tree decompositions . . . . . 42

3.4 Concluding remarks . . . . . 48

**4 Cartesian products and Sidorenko's conjecture** **51**

4.1 Cartesian products with trees . . . . . 53

4.2 Cartesian products with even cycles . . . . . 56

4.3 Concluding remarks . . . . . 57

**5 On the conjecture of Kohayakawa, Nagle, Rödl, and Schacht** **60**

5.1 Subdivisions and Sidorenko's conjecture . . . . . 60

5.2 Counting tree-like graphs in locally dense graphs . . . . . 66

5.2.1 Counting $K(r_1, r_2, \cdots, r_\ell)$-decomposable graphs . . . . . 68

5.2.2 Counting $C_{2r+1}$-decomposable graphs . . . . . 72

5.3 Concluding remarks . . . . . 75

**6 Finite reflection groups and graph norms** **78**

6.1 A motivating example . . . . . 80

6.2 Cut involutions and Cauchy–Schwarz inequalities . . . . . 82

6.3 Preliminaries on finite reflection groups . . . . . 89

6.3.1 Root systems and simple reflections . . . . . 89

6.3.2 Chambers . . . . . 90

6.3.3 The length function . . . . . 92

6.3.4 Regular polytopes . . . . . 93

6.4 Euclidean embeddings of reflection graphs . . . . . 94

6.4.1 The Euclidean embedding . . . . . 94

6.4.2 Proof of Theorem 1.10 . . . . . 98

6.5 Generalisations and applications . . . . . 104

6.5.1 Hypergraph norms . . . . . 104

6.5.2 Domination between norms . . . . . 108

6.6 Concluding remarks . . . . . 117

**Bibliography** **120**

# List of Figures

3.1 A strong tree decomposition. . . . . 24
3.2 A triangle-vertex tree (left) and a triangle-edge tree (right). . . . . 40
3.3 An example of a 2-strong tree decomposition. . . . . 44

4.1 $J \square K_2$ in $J \square H$. . . . . 52
4.2 Blow-up via $\phi$. . . . . 59

5.1 The Goldner–Harary graph and its $K_4$-decomposition. . . . . 67

6.1 The Euclidean embedding of $C_6$. . . . . 99

# Chapter 1

# Introduction

Our main concern in this thesis will be with the following beautiful conjecture of Sidorenko, which occupies a central place in extremal graph theory:

**Conjecture 1.1** (Sidorenko's conjecture [49])**.** *Let $\mu$ be the Lebesgue measure on $[0,1]$ and let $f$ be a bounded, non-negative, symmetric, measurable function on $[0,1]^2$. Then, for every bipartite graph $H$ on $\{1,2,\cdots,k\}$,*

$$\int \prod_{ij\in E(H)} f(x_i,x_j)d\mu^{|V(H)|} \geq \left(\int f(x,y)\,d\mu^2\right)^{|E(H)|}. \tag{1.1}$$

Sidorenko [48, 49] noted that the functional on the left-hand side of the inequality (1.1) often appears in various fields of science: Feynman integrals in quantum field theory [53], Mayer integrals in classical statistical mechanics, and multicenter integrals in quantum chemistry [42].

The correlation inequality (1.1) resembles the famous FKG inequality [17], which asserts that increasing functions are positively correlated when the underlying measure is log-supermodular over a finite distributive lattice. Despite the similarity, it is unclear that the FKG inequality can be applied to show (1.1): indeed there exists a function $f$ and two edge disjoint subgraphs $H_1$ and $H_2$ of a bipartite graph such that $\prod_{ij\in E(H_1)} f(x_i,x_j)$ and $\prod_{ij\in E(H_2)} f(x_i,x_j)$ are not positively correlated [34]. However, it is unknown whether every bipartite graph $H$ can be decomposed into two edge disjoint non-empty subgraphs $H_1$ and $H_2$ (possibly depending on $f$) so that $\prod_{ij\in E(H_1)} f(x_i,x_j)$ and $\prod_{ij\in E(H_2)} f(x_i,x_j)$ are positively correlated.

An equivalent discrete form expresses the conjecture in terms of graph homomorphisms. For two graphs $H$ and $G$, a *homomorphism* from $H$ to $G$ is a mapping $g : V(H) \to V(G)$ such that $\{g(v), g(w)\}$ is an edge in $G$ whenever $\{v, w\}$ is an edge in $H$. Let $\mathrm{Hom}(H, G)$ denote the set of all homomorphisms from $H$ to $G$ and let $t_H(G)$ be the probability that a uniform random mapping from $H$ to $G$ is a homomorphism, i.e.,

$$t_H(G) = \frac{|\mathrm{Hom}(H, G)|}{|V(G)|^{|V(H)|}}.$$

The discrete form of Sidorenko's conjecture states that for every bipartite graph $H$,

$$t_H(G) \geq t_{K_2}(G)^{|E(H)|} \quad \text{for all graphs } G, \tag{1.2}$$

where $K_2$ is the single edge graph. If $G$ is the Erdős–Rényi random graph $G(n, p)$, the mean of $t_H(G(n, p))$ is $p^{|E(H)|}$ plus an error term of smaller order of magnitude provided $p$ is a fixed constant. Thus, (1.2) roughly asserts that $t_H(G)$ is minimised when $G$ is the Erdős–Rényi random graph of the same density.

Many statements in extremal graph theory can be rephrased as inequalities between certain graph homomorphism densities, especially when $H$ is a fixed graph and the target graph $G$ is sufficiently large. Furthermore, once rephrased, there are various topics in extremal graph theory that connect to Sidorenko's conjecture. For example, we say that, for fixed $p \in (0, 1)$, a graph sequence $(G_n)_{n=1}^{\infty}$ with $|V(G_n)| \to \infty$ is *quasirandom with density* $p$ if and only if

$$t_H(G_n) = (1 + o(1))p^{|E(H)|} \tag{1.3}$$

for every fixed graph $H$, i.e., the $H$-count is random-like in $G$. In particular, this implies that $t_{K_2}(G) = (1 + o(1))p$. A fundamental observation in the theory of quasirandom graphs, due to Thomason [56] and Chung, Graham, and Wilson [7], states that $(G_n)_{n=1}^{\infty}$ is quasirandom with density $p$ if and only if

$$t_{C_4}(G_n) = (1 + o(1))p^4 \text{ and } t_{K_2}(G) = (1 + o(1))p, \tag{1.4}$$

where $C_4$ is a cycle of length 4. It is natural to ask what makes $C_4$ special, that is, whether we can use different graphs. More precisely,

**Question 1.2.** *For which $H$ is the property that the homomorphism density is random-like, i.e.,*

$$t_H(G_n) = (1 + o(1))p^{|E(H)|},$$

*where $t_{K_2}(G) \to p$, equivalent to quasirandomness of $(G_n)_{n=1}^{\infty}$ with density $p$?*

Let us call such a graph $H$ *forcing*. In [7], it is shown that every even cycle is forcing, while in [52] the complete bipartite graph $K_{s,t}$ is shown to be forcing. As observed by Conlon, Fox, and Sudakov [9], Question 1.2 is closely related to Sidorenko's conjecture, essentially amounting to a stability result of that conjecture. Following Skokan and Thoma [52], they also conjectured that $H$ is forcing if and only if it is bipartite and contains a cycle, which is known as the *forcing conjecture*.

Another topic related to Sidorenko's conjecture is the classical theorem of Erdős and Stone [15] that generalises Turán's theorem [58]. The Erdős–Stone theorem states that, for $\varepsilon > 0$, there exists $\delta > 0$ such that $t_H(G) \geq \delta$ whenever $t_{K_2}(G) \geq 1 - \frac{1}{\chi(H)-1} + \epsilon$, where $\chi(H)$ is the *chromatic number* of $H$, i.e., the minimum integer $r > 0$ such that $\mathrm{Hom}(H, K_r)$ is non-empty. If true, Sidorenko's conjecture verifies the optimal dependence between $\varepsilon$ and $\delta$ above for bipartite graphs $H$. On the other hand, for the case when $H = K_r$, Lovász and Simonovits [37] conjectured that a natural sequence of graphs they constructed attains the optimal dependence between $\varepsilon$ and $\delta$ asymptotically. That is, for an integer $r \geq 3$ and $t_{K_2}(G) = d$ fixed, $t_{K_r}(G) \geq F(r, d) + O(|V(G)|^{-2})$ for a certain function $F(r, d)$ of $r$ and $d$. Razborov [43] proved the conjecture for $r = 3$ in 2008, and Nikiforov [41] proved it for $r = 4$ in 2011. Recently, Reiher [45] settled the Lovász–Simonovits conjecture for all values of $r$.

Erdős and Simonovits [16, 51] also made a similar conjecture to Sidorenko's conjecture in 1984. They conjectured that, if $H$ is a bipartite graph, then there exists a positive constant $c_H$ depending only on $H$ such that

$$t_H(G) \geq c_H t_{K_2}(G)^{|E(H)|} \tag{1.5}$$

for all graphs $G$.[1] It turns out that this version of Erdős–Simonovits conjecture is

[1] Their conjecture was originally stated in terms of injective homomorphisms but is equivalent to this form if $|E(G)|$ is sufficiently larger than $|V(G)|^{2-2/|E(H)|}$. See Chapter 2 for more details.

equivalent to Sidorenko's conjecture, as Sidorenko himself showed in [48, 49] using a tensor power trick.[2]

We say that a bipartite graph $H$ has *Sidorenko's property* if (1.2) holds for all graphs $G$. That paths have Sidorenko's property [3, 40] was proved around 1960, earlier than Sidorenko suggested the conjecture. Sidorenko [48, 49] showed that trees, even cycles, and complete bipartite graphs have Sidorenko's property. He also proved that, for a bipartite graph $H$ with a bipartition $A \cup B$, $H$ has Sidorenko's property if $|A| \leq 4$. Hatami [26] proved that hypercubes have Sidorenko's property by developing the concept of weakly norming graphs, which will also be discussed in this thesis. He proved that every weakly norming graph has Sidorenko's property, and that all hypercubes are weakly norming graphs. Conlon, Fox, and Sudakov [9] proved that if $H$ is a bipartite graph with a bipartition $A \cup B$ and there is a vertex in $A$ adjacent to all vertices in $B$, then $H$ has Sidorenko's property. Li and Szegedy [33] re-proved and generalised the result of Conlon, Fox, and Sudakov with a different approach using Jensen's inequality for logarithmic functions. Partly motivated by their work, Kim, Lee, and the author [28] also used logarithmic Jensen inequalities to prove that another class of graphs, the so-called tree-arrangeable graphs, has Sidorenko's property, which generalises Conlon, Fox, Sudakov's result in a slightly different way.

These applications of Jensen's inequality have been refined in terms of information-theoretic language in subsequent work by Szegedy [54] and independently by Conlon, Kim, Lee, and the author [12]. In Chapter 3, we discuss the information theoretic approach essentially based on [12].

**Theorem 1.3** ([12])**.** *Every strongly tree-decomposable graph has Sidorenko's property.*

Roughly speaking, a graph is strongly tree-decomposable if the graph admits a certain tree-like decomposition. We also remark that it is a common generalisation of reflection trees from [33] and tree-arrangeable graphs from [28].

We elaborate on this approach to obtain a general theorem on homomorphism densities. The following result is contained in the submitted work [32] of the author, though it was implicitly contained in earlier joint work with Conlon [13].

[2] The proof is given in Chapter 2.

**Theorem 1.4** ([32]). *Let $G$, $H$, and $J$ be graphs. Suppose that $H$ is $J$-decomposable and* $\mathrm{Hom}(J, G)$ *is non-empty. Fix a $J$-decomposition $(\mathcal{F}, \mathcal{T})$ of $H$. Then the following inequality holds:*

$$t_H(G) \geq \frac{t_J(G)^{|\mathcal{F}|}}{\prod_{XY \in E(\mathcal{T})} t_{H[X \cap Y]}(G)}. \tag{1.6}$$

Here a $J$-decomposable graph roughly means that the graph is decomposed in a certain tree-like way such that each part of the decomposition induces a copy of $J$. We postpone the precise definitions of strongly tree-decomposable graphs and $J$-decomposable graphs until Chapter 3.

In Chapter 4, we study a recursive approach to Sidorenko's conjecture using Cartesian product of graphs that appeared in [12, 28]. For two graphs $H_1$ and $H_2$, let the *Cartesian product* $H_1 \square H_2$ (also known as the *box product*) be the graph over the vertex set $V(H_1) \times V(H_2)$ such that two vertices $(u_1, u_2)$ and $(v_1, v_2)$ are adjacent if and only if (i) $u_1$ and $v_1$ are adjacent in $H_1$ and $u_2 = v_2$, or (ii) $u_2$ and $v_2$ are adjacent in $H_2$ and $u_1 = v_1$. We prove the following recursive statement:

**Theorem 1.5** ([12, 28]). *If $J$ is either a tree or an even cycle and $H$ is a bipartite graph with Sidorenko's property, then $J \square H$ also has Sidorenko's property.*

The key ingredient of the proof is a one-to-one correspondence between two sets of homomorphisms: $\mathrm{Hom}(J \square H, G)$ and $\mathrm{Hom}(H, G')$ for some auxiliary graph $G'$. A slightly weaker version of this correspondence was also the key observation used in [8] to prove different graph homomorphism inequalities.

In Chapter 5, a conjecture of Kohayakawa, Nagle, Rödl, and Schacht [29] will be discussed. Other than (1.4), there are a number of equivalent conditions for graph quasirandomness observed by Thomason [56] and by Chung, Graham, and Wilson [7]. One such example is that a graph sequence $G_n$ with $|V(G_n)| \to \infty$ is quasirandom if and only if

$$\text{every subset } X \subseteq V(G_n) \text{ spans } \frac{1}{2} t_{K_2}(G_n)|X|^2 + o(|V(G_n)|^2) \text{ edges.} \tag{1.7}$$

That is, we have uniform edge density everywhere up to an error dominated by $|V(G_n)|^2$.

We can also ask if some modifications of (1.3) or (1.7) imply variations of the other. For example, we say that a graph $G$ is $(\rho, d)$-*dense* if every vertex subset $X \subseteq V(G)$ of size at least $\rho|V(G)|$ contains at least $\frac{d}{2}|X|^2$ edges. This is a weaker condition than (1.7), as we do not have an upper bound for the number of edges spanned by a vertex subset $X$. This means that we cannot recover (1.3), but it is still plausible that we can recover the lower bound

$$t_H(G_n) \geq (1 - o(1))t_{K_2}(G_n)^{|E(H)|}$$

when $\rho$ is sufficiently small.

This question was formalised by Kohayakawa, Nagle, Rödl, and Schacht [29]. They conjectured that every $(o(1), d)$-dense graph $G$ gives a random-like count for the minimum number of any fixed graph $H$. More precisely,

**Conjecture 1.6** ([29])**.** *Let $H$ be a graph and let $\eta, d$ be positive. Then there exists $\rho = \rho(\eta, d, H) > 0$ such that*

$$t_H(G) \geq (1 - \eta)d^{|E(H)|} \tag{1.8}$$

*for every sufficiently large $(\rho, d)$-dense graph $G$.*

This conjecture is not an arbitrary variant of graph quasirandomness, but has natural applications to Ramsey theory. Indeed, the notion of $(\rho, d)$-dense graphs already appears in a paper of Graham, Rödl, and Ruciński [23], where they use it to bound the Ramsey number of sparse graphs. Roughly speaking, given a 2-edge-colouring of a complete graph, one colour is very dense on some vertex subset or the other colour is somewhat dense on all vertex subsets. In the first case, it is usually simple to embed a graph $H$, while in the second case, the problem reduces to an embedding problem in locally dense graphs. It is therefore of significant interest to understand when we can embed a graph $H$ in a locally dense graph and how many copies we obtain.

As (1.8) is reminiscent of the inequality (1.2), one may suspect that Conjecture 1.6 is closely related to Sidorenko's conjecture. This is true in two different ways. Firstly, every bipartite graph $H$ with Sidorenko's property trivially satisfies (1.8) with $\rho = 1$ for any $\eta > 0$. In [12], a partial converse of this connection is also obtained, which will be discussed in Chapter 5.

**Theorem 1.7** ([12]). *Replacing each edge of a graph $H$ that satisfies Conjecture 1.6 by a path with two edges gives a bipartite graph with Sidorenko's property.*

For non-bipartite graphs, it is folklore that complete multipartite graphs satisfy the conjecture, and Reiher [44] settled the case where $H$ is an odd cycle. A result by Conlon, Hàn, Person, and Schacht [10] on weakly quasirandom hypergraphs also implies that the line graph of a hypercube satisfies the conjecture.

We give a new class of graphs for which Conjecture 1.6 holds by using the information theoretic approach and the notion of $J$-decompositions defined in Chapter 3.

**Theorem 1.8.** *Conjecture 1.6 is true for every $J$-decomposable graph such that $J$ is either a complete multipartite graph or an odd cycle.*

Our concern in Chapter 6 will be with the question, proposed by Lovász [36], of deciding which graphs define norms. We say that a graph $H$ is *norming* if the functional defined by

$$\|f\|_H := \left| \int \prod_{ij \in E(H)} f(x_i, x_j) d\mu^{|V(H)|} \right|^{1/|E(H)|} \tag{1.9}$$

is a semi-norm, and *weakly norming* if

$$\|f\|_{r(H)} := \left( \int \prod_{ij \in E(H)} |f(x_i, x_j)| d\mu^{|V(H)|} \right)^{1/|E(H)|} \tag{1.10}$$

is a norm. As one might expect from the name, it is easy to check that every norming graph is weakly norming. While our focus will usually be on weakly norming graphs, we will often discuss how analogous results can be derived for norming graphs under an extra technical condition.

The first in-depth study of (weakly) norming graphs was undertaken by Hatami [26]. A moment's thought shows that $H$ is necessarily bipartite whenever it is weakly norming, because otherwise $\|\cdot\|_{r(H)}$ can be zero for strictly positive functions. In [26], Hatami showed that the $n$-dimensional hypercube $Q_n$ and the complete bipartite graph $K_{m,n}$ are weakly norming. He also observed that the functionals $\|\cdot\|_{C_{2k}}$ correspond to the classical Schatten–von Neumann norms and,

hence, even cycles are norming. In [36], Lovász showed that a similar method works to show that the complete bipartite graph $K_{n,n}$ minus a perfect matching is weakly norming.

We generalise these results, finding a much larger class of (weakly) norming graphs that includes all of the known examples. To give an indication of our results, suppose that $k$ and $r$ are integers with $k \leq r$ and $\mathcal{P}$ is a polytope. Consider the bipartite graph between $k$-faces and $r$-faces of $\mathcal{P}$ indicating their incidence. That is, we place an edge between a $k$-face and an $r$-face if one contains the other. We call this graph the $(k, r)$-*incidence graph* of the polytope $\mathcal{P}$. We then have the following theorem.

**Theorem 1.9** ([13])**.** $(k, r)$*-incidence graphs of regular polytopes are weakly norming.*

The definition of regular polytopes will be given in Section 6.3.4. For example, in an $n$-dimensional simplex, the $k$-faces and $r$-faces naturally correspond to $(k+1)$-element and $(r+1)$-element subsets of $[n]$. Therefore, the $(k, r)$-incidence graph of an $n$-simplex is the inclusion graph between $(k+1)$-sets and $(r+1)$-sets. In particular, the $(0, 1)$-incidence graph is the 1-subdivision of $K_n$, the $(0, n-2)$-incidence graph is $K_{n,n}$ minus a perfect matching, and the $(0, n-1)$-incidence graph is the star $K_{1,n}$, which by tensor powering shows that $K_{m,n}$ is also weakly norming. Even cycles $C_{2k}$ are the $(0, 1)$-incidence graphs of regular $k$-gons and thus are weakly norming. More generally, by considering the $(0, 1)$-incidence graph of any regular polytope, such as hypercubes or the icosahedron, we see that their 1-subdivisions are weakly norming.

We prove Theorem 1.9 as a corollary of a more general result showing how weakly norming graphs arise from finite reflection groups. A *finite reflection group* is a finite subgroup of $\mathbf{GL}(n, \mathbb{R})$ generated by a set of reflections across hyperplanes passing through the origin. Those readers who are not familiar with these groups may temporarily assume that, given a finite reflection group $W$, there exists a distinguished set of generators $S$ known as *simple reflections* (for further details, see Section 6.3). Fixing a generating set of simple reflections $S$ in a finite reflection group $W$, let $S_1$ and $S_2$ be subsets of $S$ and let $W_1$ and $W_2$ be the subgroups of $W$ generated by $S_1$ and $S_2$, respectively. Consider the bipartite graph between the

(left-)cosets of $W_1$ and $W_2$, where $wW_1$ and $wW_2$ are adjacent for every $w \in W$. We call this graph the $(S_1, S_2; W)$-*reflection graph* and say that a graph $H$ is a *reflection graph* if it is isomorphic to an $(S_1, S_2; W)$-reflection graph for some suitable choice of parameters. With these definitions, we may now state our main result.

**Theorem 1.10** ([13])**.** *Reflection graphs are weakly norming.*

This class includes the $(k, r)$-incidence graphs of regular polytopes, but also provides other simple examples, such as the hypercube (rather than its subdivision) and the graph obtained by replacing each edge of an octahedron with a cycle of length 4 (see Example 6.26 for a more formal description). It also opens the door to more exotic examples, coming from the exceptional reflection groups $E_6$, $E_7$, and $E_8$. An analogous result also holds for norming graphs, though under a slightly stronger condition.

**Theorem 1.11** ([13])**.** *Let $W$ be a finite reflection group and let $S$ be a generating set of simple reflections. Then, for any disjoint subsets $S_1$ and $S_2$ of $S$, the $(S_1, S_2; W)$-reflection graph is norming.*

As observed by Hatami [26], every weakly norming graph has Sidorenko's property. However, this is not the only application of our results to Sidorenko's conjecture. In Section 3.2.2, while obtaining Theorem 1.4, we will also prove that weakly norming graphs can be used as building blocks for constructing new graphs that satisfy the conjecture.

Finally, in Section 6.5, we will discuss generalisations of our results to hypergraphs, with the main result being that a suitably defined family of reflection hypergraphs are weakly norming. We will then show that every norm defined in this manner is equivalent, in some well-defined sense, to a corresponding cut-norm and then to an appropriate variant of the octahedral norms introduced by Gowers [21, 22] in his work on hypergraph regularity. Our methods also allow us to compare the relative strengths of these variants. These results generalise earlier work by Gowers [21, 22] and by Conlon, Hàn, Person, and Schacht [10] on equivalences between cut-norms and octahedral norms and, hence, the corresponding notions of pseudorandomness.

# Chapter 2

# Preliminaries

In this chapter, we give some preliminaries on graph theory and probability theory.

For two functions $f, g : \mathbb{N} \to \mathbb{R}^+$, $f = o(g)$ means $\lim_{n\to\infty} f(n)/g(n) = 0$. In particular, $f(n) = o(1)$ means $\lim_{n\to\infty} f(n) = 0$.

On a real vector space $V$, a functional $\|\cdot\| : V \to \mathbb{R}$ is a *semi-norm* if the following hold:

(i) $\|x\| = 0$ if $x = 0$;

(ii) $\|cx\| = |c| \, \|x\|$ for each $c \in \mathbb{R}$ and $x \in V$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$.

In (i) above, $\|x\| = 0$ does not imply $x = 0$. If this is also true, i.e., a semi-norm $\|\cdot\|$ satisfies $\|x\| = 0$ if and only if $x = 0$, then we say that $\|\cdot\|$ is a *norm*. The inequality (iii) is called the *triangle inequality*.

## 2.1 Graph theory

All graphs in this thesis are finite loopless undirected graphs without multiple edges. We denote by $K_r$ and $K_{r,s}$ the complete graph on $r$ vertices and the complete bipartite graph on the bipartition $A \cup B$, where $|A| = r$ and $|B| = s$, respectively. For $k \geq 3$, $C_k$ is the cycle on $k$ vertices and, for $\ell \geq 1$, $P_\ell$ means the path with $\ell$ edges and $\ell + 1$ vertices.

### 2.1.1 Graph homomorphisms

A function $f : V(H) \to V(G)$ is said to be a *graph homomorphism* from a graph $H$ to another graph $G$ if and only if $f(u)f(v)$ is an edge in $G$ whenever $uv$ is an edge in $H$.

The *graph homomorphism density* $t_H(G)$ is the probability that a uniform random map from $V(H)$ to $V(G)$ is a graph homomorphism, i.e.,

$$t_H(G) := \frac{|\mathrm{Hom}(H, G)|}{|V(G)|^{|V(H)|}}.$$

We may similarly define the *injective homomorphism density* $t_H^{\mathrm{inj}}(G)$ as the probability that a uniform random injective map from $V(H)$ to $V(G)$ is a graph homomorphism, i.e.,

$$t_H^{\mathrm{inj}}(G) := \frac{|\{f \in \mathrm{Hom}(H, G) : f \text{ is injective}\}|}{|V(G)|(|V(G)| - 1) \cdots (|V(G)| - |V(H)| + 1)}.$$

For a fixed graph $H$ and a graph sequence $(G_n)_{n=1}^{\infty}$ with $|V(G_n)| \to \infty$, the graph homomorphism density is asymptotically the same as the injective homomorphism density. More precisely, a random vertex map from $H$ to $G$ is a non-injective homomorphism if and only if there exist two vertices in $H$ mapped to the same vertex in $G$. Therefore,

$$|t_H(G_n) - t_H^{\mathrm{inj}}(G_n)| \leq \sum_{u,v \in V(H)} \mathbb{P}[x_u = x_v] \leq \frac{|V(H)|^2}{|V(G_n)|} = o(1), \tag{2.1}$$

where $(x_v)_{v \in V(H)}$ are random vertices chosen independently from $V(G)$. Thus, in extremal graph theory, classical results about finding a copy of $H$ in a sufficiently large graph $G$ can be rephrased in terms of graph homomorphism densities. For instance, the classical theorem of Turán [58] states, roughly speaking, that for all integers $r \geq 3$, every graph $G$ with more than $\left(1 - \frac{1}{r-1}\right) \frac{|V(G)|^2}{2}$ edges contains $K_r$ as a subgraph. In terms of graph homomorphisms, it may be re-stated as follows: for all integers $r \geq 3$, if $t_{K_2}(G) > 1 - \frac{1}{r-1}$, then there exists a homomorphism from $K_r$ to $G$. Since $\chi(H) = r$ means that there exists a homomorphism from $H$ to $K_r$, it follows that for all graphs $H$ with $\chi(H) = r$, there exists a homomorphism from $H$ to $G$ whenever $t_{K_2}(G) > 1 - \frac{1}{r-1}$.

We remark that the Erdős–Simonovits conjecture stated in terms of graph homomorphisms (1.5) is equivalent to the original statement in terms of injective homomorphisms unless the $o(1)$-term in (2.1) dominates $t_{K_2}(G)^{|E(H)|}$. More precisely, the equivalence remains valid if

$$\frac{C}{|V(G)|} \leq \left(\frac{2|E(G)|}{|V(G)|^2}\right)^{|E(H)|}$$

for a large enough constant $C$ only depending on $H$, which reduces to $|E(G)| \geq C|V(G)|^{2-1/|E(H)|}$. This is larger than a few known upper bounds for Turán numbers of bipartite graphs, e.g., Bondy–Simonovits [4] for even cycles and Kövári–Sós–Turán [31] for complete bipartite graphs, and hence (1.5) does not fully recover the supersaturation on the Turán threshold originally conjectured by Erdős and Simonovits.

### 2.1.2 Spectral graph theory

The *adjacency matrix* $A_G$ of a graph $G$ is the $|V(G)|\times|V(G)|$ symmetric 0-1 matrix such that the entry corresponding to the vertex pair $(u, v)$ is 1 if $uv \in E(G)$ and 0 otherwise. The *eigenvalues* of $G$ are the eigenvalues of its adjacency matrix. We order the eigenvalues in such a way that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. The following facts are immediate from diagonalisation of the adjacency matrix.

**Proposition 2.1.** *Let $G$ be an $n$-vertex graph and let $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ be its eigenvalues. Then*

*(i)* $t_{K_2}(G) = \frac{1}{n^2}\sum_{i=1}^n \lambda_i^2$ *and*

*(ii) for* $k > 2$, $t_{C_k}(G) = \frac{1}{n^k}\sum_{i=1}^n \lambda_i^k$.

These facts imply that the even cycle density $t_{C_{2k}}(G)$ exactly corresponds to the $2k$-th power of the $\ell_{2k}$-norm of the normalised spectrum $(\lambda_1/n, \lambda_2/n, \cdots, \lambda_n/n)$, which is usually referred to in the literature as a Schatten–von Neumann norm. As firstly observed by Hatami [26], this implies the fact that

$$\|f\|_{C_{2k}} := \left|\int \prod_{ij\in E(C_{2k})} f(x_i, x_j) d\mu^{|V(H)|}\right|^{1/2k}$$

defines a norm on the space of bounded symmetric measurable functions on $[0, 1]^2$.

### 2.1.3 Quasirandomness

A cornerstone in the theory of quasirandomness, due to Thomason [56] and Chung, Graham, and Wilson [7], is the following set of equivalences.

**Theorem 2.2.** *Let $G_n$ be a graph sequence with $|V(G_n)| \to \infty$, and let $p \in (0,1)$ be fixed. Suppose that $t_{K_2}(G_n) \to p$. Then the following are equivalent:*

*(i) for every fixed graph $H$, $t_H(G_n) = (1 \pm o(1))p^{|E(H)|}$.*

*(ii) $t_{C_4}(G_n) = (1 + o(1))p^4$.*

*(iii) every subset $X \subseteq V(G_n)$ spans $\frac{1}{2}p|X|^2 \pm o(|V(G_n)|^2)$ edges.*

*(iv) $\lambda_2(G_n) = o(|V(G_n)|)$.*

The equivalent conditions above define the notion of graph quasirandomness. In the introduction, we defined a graph sequence $(G_n)_{n=1}^{\infty}$ with $|V(G_n)| \to \infty$ to be *quasirandom with density $p$* if and only if, for every fixed graph $H$,

$$t_H(G_n) = (1 \pm o(1))p^{|E(H)|},$$

i.e., the homomorphism density is essentially what we expect from the Erdős–Rényi random graph with the same edge density $p$. Hence, a quasirandom graph sequence $(G_n)_{n=1}^{\infty}$ satisfies (i)–(iv) in Theorem 2.2.

### 2.1.4 Tree decompositions

A *tree decomposition* of a graph $H$, a concept introduced by Halin [24] and developed by Robertson and Seymour [47], is a pair $(\mathcal{F}, \mathcal{T})$ consisting of a family $\mathcal{F}$ of vertex subsets of $H$ and a tree $\mathcal{T}$ on vertex set $\mathcal{F}$ satisfying

(i) $\bigcup_{X \in \mathcal{F}} X = V(H)$,

(ii) for each $\{v, w\} \in E(H)$, there exists a set $X \in \mathcal{F}$ such that $v, w \in X$, and

(iii) for $X, Y, Z \in \mathcal{F}$, $X \cap Y \subseteq Z$ whenever $Z$ lies on the path from $X$ to $Y$ in $\mathcal{T}$.

We remark that the third condition, which can be seen as a 'Markov' property, will be crucial for our purpose.

We say that $\max_{F\in\mathcal{F}} |F|-1$ is the *width* of the tree decomposition $(\mathcal{F},\mathcal{T})$, and the minimum width over all possible tree decompositions of the given graph $H$ is called the *tree-width* of $H$.

The following observation will be used in Chapter 5.

**Lemma 2.3.** *There exists a tree decomposition $(\mathcal{F},\mathcal{T})$ of $H$ with the minimum width such that $|\mathcal{F}| \leq |E(H)|$.*

*Proof.* Let $(\mathcal{F},\mathcal{T})$ be a tree-decomposition of $H$ with the minimum width $w$. We may assume that, amongst those with the minimum width, $(\mathcal{F},\mathcal{T})$ attains the smallest $|\mathcal{F}|$. If a vertex bag $X \in \mathcal{F}$ contains an edge $e$ that is not contained in any other $F \in \mathcal{F}$, we say that $X$ *covers* $e$. Then each leaf $L$ of $\mathcal{T}$ covers at least one edge, since otherwise the leaf is redundant. While removing leaves successively, we map each leaf to an edge that is covered by the leaf. Suppose now that, after removing $X_1, X_2, \cdots, X_k \in \mathcal{F}$, a leaf $X$ of $\mathcal{T} \setminus \{X_1, X_2, \cdots, X_k\}$ covers no edge. Then by contracting $X$ with its parent $X'$ in $\mathcal{T}$, i.e., deleting $X$ in $\mathcal{T}$ and adding edges from $X'$ to the other neighbours of $X$ in $\mathcal{T}$, we obtain a tree decomposition $(\mathcal{F}',\mathcal{T}')$ of $H$ with the same width $w$ and $|\mathcal{F}'| = |\mathcal{F}| - 1$, which contradicts to the minimality of $|\mathcal{F}|$. Hence, we may always map each leaf successfully during the removal process, which gives $|\mathcal{F}| \leq |E(H)|$. □

Denote by $\mathcal{F}(v)$ the subfamily $\{X \in \mathcal{F} : v \in X\}$ of $\mathcal{F}$. The last condition in the definition of tree decompositions implies the following folklore fact:

**Lemma 2.4.** *If $(\mathcal{F},\mathcal{T})$ is a tree decomposition of $H$ and $v \in V(H)$, then $\mathcal{F}(v)$ always induces a subtree of $\mathcal{T}$.*

This lemma yields the following identity on the number of edges in $H$, which will be repeatedly used in what follows:

**Lemma 2.5.** *Let $(\mathcal{F},\mathcal{T})$ be a tree-decomposition of a graph $H$. Then*

$$|E(H)| = \sum_{X\in\mathcal{F}} |E(H[X])| - \sum_{XY\in E(\mathcal{T})} |E(H[X\cap Y])|. \tag{2.2}$$

*Proof.* Let $uv \in E(H)$ and let $\mathcal{F}(uv)$ be the collection of vertex sets $X \in \mathcal{F}$ containing both $u$ and $v$. Since $\mathcal{F}(uv) = \mathcal{F}(u) \cap \mathcal{F}(v)$ and $\mathcal{F}(uv) \neq \emptyset$ by (ii) of the definition of tree decompositions, $\mathcal{F}(uv)$ induces a non-empty subtree $\mathcal{T}_{uv}$ of $\mathcal{T}$. In (2.2), each $uv$ is counted $|\mathcal{F}(uv)|$ times in $\sum_{X \in \mathcal{F}} |E(H[X])|$ and $|E(\mathcal{T}_{uv})| = |\mathcal{F}(uv)| - 1$ times in $\sum_{XY \in E(\mathcal{T})} |E(H[X \cap Y])|$. Thus, every edge is counted once on both sides of (2.2), which proves the desired identity. □

**Lemma 2.6.** *Let $T$ be a tree and let $T_1, T_2, \cdots, T_k$ be subtrees of $T$ such that $\bigcap_{i=1}^{k} V(T_i)$ is non-empty. Suppose that a vertex subset $F \subseteq V(T)$ induces a subtree of $T$ and intersects each $V(T_i)$, $i = 1, 2, \cdots, k$. Then $F \cap \left(\bigcap_{i=1}^{k} V(T_i)\right)$ is non-empty.*

*Proof.* We use induction on $k$. The base case $k = 1$ is trivially true. Let $W := F \cap \left(\bigcap_{i=1}^{k-1} V(T_i)\right)$. Note that $W$ is non-empty by induction hypothesis, and moreover, $W$ induces a subtree of $T$. Suppose that $W$ and $V(T_k)$ are disjoint. Then there exists a vertex $u$ on the shortest path between the two disjoint sets $W$ and $V(T_k)$ such that $u \notin W \cup V(T_k)$. However, $u$ must be in $F$, since $F$ intersects both $W$ and $V(T_k)$ and $T[F]$ is a subtree of $T$. On the other hand, $u$ must also be in $\bigcap_{i=1}^{k-1} V(T_i)$ by exactly the same reason: $\cap_{i=1}^{k-1} V(T_i)$ intersects both $W$ and $V(T_k)$, and it induces a subtree. Thus, $u \in W$, which yields a contradiction. □

The following corollary of Lemma 2.4 will be used in Chapter 3.

**Proposition 2.7.** *Let $(\mathcal{F}, \mathcal{T})$ be a tree decomposition of $H$ and let $U \subseteq V(H)$. Suppose $\cap_{u \in U} \mathcal{F}(u) = \emptyset$. Then there exists a subtree $\mathcal{T}'$ of $\mathcal{T}$ such that $\cup_{F \in V(\mathcal{T}')} F$ contains $U$ and every subtree $\mathcal{T}''$ of $\mathcal{T}$ such that $U \subseteq \cup_{F \in V(\mathcal{T}'')} F$ contains $\mathcal{T}'$ as a subtree.*

*Proof.* By Lemma 2.4, $\mathcal{F}(u) = \{X \in \mathcal{F} : u \in X\}$ induces a subtree of $\mathcal{T}$. Denote by $\mathcal{T}(u)$ the induced subtree $\mathcal{T}[\mathcal{F}(u)]$. Let $U = \{u_1, u_2, \cdots, u_t\}$. We use induction on $t$. For brevity, we say that a subtree $\mathcal{T}'$ of $\mathcal{T}$ is *good* for $U$ if $U \subseteq \cup_{F \in V(\mathcal{T}')} F$. Suppose $t = 2$ and $\mathcal{F}(u_1)$ and $\mathcal{F}(u_2)$ are disjoint. Then the shortest path from the two vertex-disjoint trees $\mathcal{T}(u_1)$ and $\mathcal{T}(u_2)$ is good for $U$.

For $t > 2$, suppose that $\cap_{i=1}^{t-1} \mathcal{F}(u_i) = \emptyset$. Then by induction hypothesis there exists a good subtree $\mathcal{T}''$ for $U \setminus \{u_t\}$ contained in every good subtree for $U \setminus \{u_t\}$.

If $V(\mathcal{T}'')$ intersects $\mathcal{F}(u_t)$ then let $\mathcal{T}' := \mathcal{T}''$. Then $\mathcal{T}'$ is good for $U$ too, and moreover, it is the minimum one, since every good subtree for $U$ is again good for $U \setminus \{u_t\}$. Otherwise if $V(\mathcal{T}'')$ and $\mathcal{F}(u_t)$ are disjoint, add the shortest path between $V(\mathcal{T}'')$ and $\mathcal{F}(u_t)$ to $\mathcal{T}''$. Then the extended tree $\mathcal{T}'$ is good for $U$. Moreover, every good subtree $\mathcal{T}_0$ for $U$ must contain $\mathcal{T}''$ and thus, $\mathcal{T}_0$ also contains the shortest path from $V(\mathcal{T}'')$ to $\mathcal{F}(u_t)$.

Suppose now that $\cap_{i=1}^{t-1}\mathcal{F}(u_i)$ is non-empty. Then $\cap_{i=1}^{t-1}\mathcal{F}(u_i)$ and $\mathcal{F}(u_t)$ must be disjoint. Let $\mathcal{T}'$ be the shortest path from $\cap_{i=1}^{t-1}\mathcal{F}(u_i)$ to $\mathcal{F}(u_t)$. We claim that $\mathcal{T}'$ is the desired minimum good subtree for $U$. To see this, suppose that $\mathcal{T}''$ is a good subtree for $U$. As the vertex set of each good subtree for $U$ intersects every $\mathcal{F}(u_i)$, $1 \leq i \leq t-1$, applying Lemma 2.6 with $k := t-1$, $T := \mathcal{T}$, $V(T_i) := \mathcal{F}(u_i)$, and $F := V(\mathcal{T}'')$ implies that $V(\mathcal{T}'')$ must intersect $\cap_{i=1}^{t-1}\mathcal{F}(u_i)$. Moreover, $V(\mathcal{T}'')$ also intersects $\mathcal{F}(u_t)$, since $\mathcal{T}''$ is good for $U$. Therefore, every good subtree for $U$ must contain $\mathcal{T}'$.

It remains to prove that $U$ is contained in $\cup_{F \in V(\mathcal{T}')} F$, as $\mathcal{T}'$ is already a path in $\mathcal{T}$. Firstly, the end vertex of the path $\mathcal{T}'$ is in $\mathcal{F}(u_t)$ and hence contains $u_t$. On the other hand, since every element of $\cap_{i=1}^{t-1}\mathcal{F}(u_i)$ contains $U \setminus \{u_t\}$, the starting vertex of the path $\mathcal{T}'$ contains $U \setminus \{u_t\}$, which implies $U \subseteq \cup_{F \in V(\mathcal{T}')} F$. □

### 2.1.5 Tensor products

Let $G$ and $H$ be graphs. The *tensor product* $G \otimes H$ is the graph on $V(G) \times V(H)$ such that $(u, v)$ and $(u', v')$ are adjacent if and only if $uu'$ and $vv'$ are edges in $G$ and $H$, respectively. Denote by $G^{\otimes k}$ the $k$-th tensor power of $G$, i.e., the tensor product of $k$ copies of $G$.

Tensor products are naturally related to graph homomorphism density. Namely,

$$t_H(G_1 \otimes G_2) = t_H(G_1) t_H(G_2).$$

In particular,

$$t_H(G^{\otimes k}) = t_H(G)^k.$$

Using these observations, we give a proof, originally due to Sidorenko [49], that the rephrased version (1.5) of the Erdős–Simonovits conjecture [16] in terms

of graph homomorphisms is equivalent to Sidorenko's conjecture. This lemma will also be used in subsequent chapters.

**Lemma 2.8** ([49]). *Let $H$ be a bipartite graph. If there exists a constant $c_H > 0$ such that*

$$t_H(G) \geq c_H t_{K_2}(G)^{|E(H)|}$$

*for all graphs $G$, then $t_H(G) \geq t_{K_2}(G)^{|E(H)|}$.*

*Proof.* Suppose that for some $G$ we have $t_H(G) = \beta t_{K_2}(G)^{|E(H)|}$, where $\beta < 1$. Taking the $k$-th tensor power gives

$$t_H(G)^k = t_H(G^{\otimes k}) \geq c_H t_{K_2}(G^{\otimes k})^{|E(H)|} = c_H t_{K_2}(G)^{k|E(H)|}.$$

However, as $t_H(G)^k = \beta^k t_{K_2}(G)^{k|E(H)|}$, $\beta^k > c_H$ holds for an arbitrary $k$, which gives a contradiction. □

We also remark that, in Hatami's work [26], an analogous tensor power trick is the key to obtain the equivalence between the inequality (6.2) and the triangle inequality.

### 2.1.6 Cartesian products

Let $G$ and $H$ be graphs. The *Cartesian product* $G \square H$ is the graph on $V(G) \times V(H)$ such that $(u, v)$ and $(u', v')$ are adjacent if and only if either $u = u'$ and $vv' \in E(H)$ or $uu' \in E(G)$ and $v = v'$.

For example, the $d$-dimensional hypercube $Q_d$ can be obtained by taking the successive Cartesian product of $d$ single-edge graphs. More generally, any $d$-dimensional grids are obtained by a successive Cartesian product of $d$ paths.

Let $G$ and $H$ be bipartite graphs on the bipartition $V \cup W$ and $A \cup B$, respectively. Then one may easily check that every edge in $G \square H$ has one end in $(V \times A) \cup (W \times B)$ and the other end in $(V \times B) \cup (W \times A)$. Thus, $G \square H$ is again bipartite. More generally, the chromatic number $\chi(G \square H)$ of $G \square H$ is equal to $\max\{\chi(G), \chi(H)\}$.

## 2.2 Probability theory

### 2.2.1 Conditional independence

For three random variables $X, Y$, and $Z$, we say that $X$ and $Z$ are *conditionally independent* given $Y$ if the distribution of $X$ conditioned on $(Y, Z) = (y, z)$ is identical to the distribution of $X$ conditioned on $Y = y$ for all possible values $y, z$. In particular, when $X, Y$, and $Z$ take values in a finite set,

$$\mathbb{P}[X = x | Y = y, Z = z] = \mathbb{P}[X = x | Y = y]$$

for all $x, y, z$.

A folklore lemma, essentially implied by the Kolmogorov extension theorem, will be necessary in Chapter 3. We give a simple proof when all the random variables take finitely many values, which suffices for our purpose. For a modern introduction to product measure spaces and the Kolmogorov extension theorem, we refer to [55].

**Lemma 2.9.** *Let $(X_1, X_2)$ and $(X_2', X_3)$ be random vectors taking values in a finite set. If $X_2$ and $X_2'$ are identically distributed, then there exists $(Y_1, Y_2, Y_3)$ such that $Y_1$ and $Y_3$ are conditionally independent given $Y_2$, $(X_1, X_2)$ and $(Y_1, Y_2)$ are identically distributed, and $(X_2', X_3)$ and $(Y_2, Y_3)$ are also identically distributed.*

*Proof.* Suppose $X_1, X_2$, and $X_3$ take values in finite sets $A, B$, and $C$, respectively. Note that $X_2'$ also takes values in $B$. Let $Y_1, Y_2$, and $Y_3$ be random variables defined by the joint distribution

$$\mathbb{P}[Y_1 = a, Y_2 = b, Y_3 = c] = \frac{\mathbb{P}[X_1 = a, X_2 = b]\ \mathbb{P}[X_2' = b, X_3 = c]}{\mathbb{P}[X_2 = b]} \tag{2.3}$$

for all possible values of $a, b$, and $c$. This is well-defined, since $\mathbb{P}[X_1 = a, X_2 = b] = 0$ whenever $\mathbb{P}[X_2 = b] = 0$. We claim that $(Y_1, Y_2, Y_3)$ is a random vector with desired properties. Summing (2.3) over all $a \in A$ gives

$$\mathbb{P}[Y_2 = b, Y_3 = c] = \mathbb{P}[X_2' = b, X_3 = c] \tag{2.4}$$

and the fact that $\mathbb{P}[X_2 = b] = \mathbb{P}[X_2' = b]$ gives that by symmetry

$$\mathbb{P}[Y_1 = a, Y_2 = b] = \mathbb{P}[X_1 = a, X_2 = b].$$

Moreover, by substituting (2.4) into (2.3), we obtain

$$\begin{aligned}\mathbb{P}[Y_1 = a|Y_2 = b, Y_3 = c] &= \frac{\mathbb{P}[Y_1 = a, Y_2 = b, Y_3 = c]}{\mathbb{P}[Y_2 = b, Y_3 = c]} \\ &= \frac{\mathbb{P}[X_1 = a, X_2 = b]}{\mathbb{P}[X_2 = b]} = \mathbb{P}[Y_1 = a|Y_2 = b],\end{aligned}$$

which implies the conditional independence of $Y_1$ and $Y_3$ given $Y_2$. □

### 2.2.2 Entropy

Throughout the thesis, logarithms will be understood to be base 2. Suppose that $X$ is a random variable taking values in some finite set $S$. The *entropy* of $X$ is defined by

$$\mathbb{H}(X) = \sum_{x \in S} \mathbb{P}[X = x] \log\left(\frac{1}{\mathbb{P}[X = x]}\right).$$

Given two random variables $X$ and $Y$ over the same space, the *conditional entropy* of $Y$ with respect to $X$ is

$$\mathbb{H}(Y|X) = \mathbb{H}(X, Y) - \mathbb{H}(X).$$

Here are some simple properties of entropy which will be crucial in what follows. Again, we assume that $X, Y$, and $Z$ take values in finite sets.

**Lemma 2.10.** *Let $X$, $Y$, and $Z$ be random variables and suppose that $X$ takes values in a set $S$. Then*

*(i)* $\mathbb{H}(X) \leq \log |S|$,

*(ii)* $\mathbb{H}(X|Y, Z) = \mathbb{H}(X|Y)$ *if $X$ and $Z$ are conditionally independent given $Y$.*

*Proof.* For the sake of brevity, denote by $p(x), p(y), p(z), p(x, y), p(y, z)$ and $p(x, y, z)$ the probability $\mathbb{P}[X = x], \mathbb{P}[Y = y], \mathbb{P}[Z = z], \mathbb{P}[X = x, Y = y], \mathbb{P}[Y = y, Z = z]$, and $\mathbb{P}[X = x, Y = y, Z = z]$, respectively. Moreover, define $p(x|y) := p(x, y)/p(y)$.

(i) By Jensen's inequality for the concave function $f(x) = \log x$,

$$\sum_{x \in S} p(x) \log\left(\frac{1}{p(x)}\right) \leq \log\left(\sum_{x \in S} p(x) \cdot \frac{1}{p(x)}\right) = \log |S|.$$

(ii) By definition,

$$\begin{aligned}\mathbb{H}(X|Y) &= \sum_{x,y} p(x,y)\log\left(\frac{1}{p(x,y)}\right) - \sum_{y} p(y)\log\left(\frac{1}{p(y)}\right) \\ &= \sum_{x,y} p(x,y)\log\left(\frac{1}{p(x|y)}\right) \\ &= \sum_{x,y,z} p(x,y,z)\log\left(\frac{1}{p(x|y)}\right). \end{aligned} \tag{2.5}$$

On the other hand, $\mathbb{H}(X|Y,Z) = \mathbb{H}(X,Y,Z) - \mathbb{H}(Y,Z)$ and hence,

$$\begin{aligned}\mathbb{H}(X|Y,Z) &= \sum_{x,y,z} p(x,y,z)\log\left(\frac{1}{p(x,y,z)}\right) - \sum_{y,z} p(y,z)\log\left(\frac{1}{p(y,z)}\right) \\ &= \sum_{x,y,z} p(x,y,z)\log\left(\frac{p(y,z)}{p(x,y,z)}\right). \end{aligned} \tag{2.6}$$

Conditional independence gives $p(x,y,z) = p(x|y)p(y,z)$ and thus (2.5) equals (2.6). □

### 2.2.3 Large deviations

In Chapter 5, we will need some concentration inequalities, which follow from a well-known large deviation inequality, called Azuma's inequality.

**Lemma 2.11** (Azuma's inequality, Corollary 7.2.2 in [1])**.** *Let $c = X_0, X_1, \cdots, X_m$ be a martingale with $|X_{i+1} - X_i| \leq 1$ for all $0 \leq i < m$. Then*

$$\mathbb{P}[|X_m - c| > \lambda\sqrt{m}] < 2e^{-\lambda^2/2}.$$

*In particular, by taking $\lambda = \frac{c}{3\sqrt{m}}$, we have*

$$\mathbb{P}[|X_m - c| > c/3] < 2e^{-c^2/18m}.$$

Let $W : [n]^2 \to [0,1]$ be a symmetric function with $W(u,u) = 0$ for every $u$. Let $\mathcal{G} = \mathcal{G}(n,W)$ be the (generalised) random graph on $[n]$ obtained by including each possible edge $uv$ with probability $W(u,v)$, independently of the others. For a fixed graph $H$, let

$$t_H(W) := \frac{1}{n^{|V(H)|}} \sum_{x_1,\ldots,x_h \in [n]} \prod_{ij \in E(H)} W(x_i, x_j).$$

Then Lemma 2.11 implies the following:

**Corollary 2.12.** *Let $H$ be a fixed graph and let $\mathcal{G} = \mathcal{G}(n, W)$ as above. Suppose that there exists $\varepsilon > 0$ independent of $n$ such that $t_H(W) \geq \varepsilon$, i.e., $t_H(W)$ does not vanish as $n$ grows. Then there exist positive constants $C = C(H)$ and $n_0 = n(\varepsilon, H)$ such that, for $n \geq n_0$,*

$$\mathbb{P}\left[|t_H(\mathcal{G}) - t_H(W)| > \frac{1}{2}t_H(W)\right] \leq Ce^{-\frac{n^2 t_H(W)^2}{72|E(H)|}}.$$

*Proof.* Let $m = \binom{n}{2}$. We use the standard martingale adapted to the edge exposure filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_m$, i.e., the edge exposure martingale $X_i := \mathbb{E}[t_H(\mathcal{G})|\mathcal{F}_i]$, $i = 0, 1, \cdots, m$. In particular, $X_m = t_H(\mathcal{G})$ and

$$X_0 = \mathbb{E}[t_H(\mathcal{G})] = \frac{1}{n^{|V(H)|}} \sum_{x_1,\ldots,x_h \in [n]} \mathbb{E}\left[\prod_{ij \in E(H)} \mathcal{G}(x_i, x_j)\right],$$

where $\mathcal{G}(x_i, x_j)$ is the indicator function of the event that $x_i$ and $x_j$ are adjacent in $\mathcal{G}$. If $\{x_i, x_j\}$, $ij \in E(H)$, are pairwise distinct, then $\mathcal{G}(x_i, x_j)$, $ij \in E(H)$, are mutually independent and hence, $\mathbb{E}\left[\prod_{ij \in E(H)} \mathcal{G}(x_i, x_j)\right] = \prod_{ij \in E(H)} W(x_i, x_j)$. Otherwise, $\mathbb{E}\left[\prod_{ij \in E(H)} \mathcal{G}(x_i, x_j)\right] \geq \prod_{ij \in E(H)} W(x_i, x_j)$. Thus,

$$t_H(W) \leq X_0 \leq t_H(W) + O(1/n). \tag{2.7}$$

In particular, $X_0/t_H(W) \to 1$ as $n$ tends to infinity, since $t_H(W) \geq \varepsilon$. As adding one edge makes at most $2|E(H)|n^{|V(H)|-2}$ new homomorphic copies of $H$, $|\mathbb{E}[t_H(\mathcal{G})|\mathcal{F}_{i+1}] - \mathbb{E}[t_H(\mathcal{G})|\mathcal{F}_i]| \leq 2|E(H)|/n^2$. Let $Y_i := \frac{n^2 X_i}{2|E(H)|}$, $i = 1, 2, \cdots, m$. Then $(Y_i)_{i=0}^m$ satisfies the Lipshitz condition $|Y_{i+1} - Y_i| \leq 1$ for $0 \leq i < m$. Now Lemma 2.11 with $c = Y_0$ and $m < n^2$ implies that, for a positive constant $C$ only depending on $|V(H)|$ and large enough $n$,

$$\begin{aligned}
\mathbb{P}\left[|t_H(\mathcal{G}) - t_H(W)| > \frac{1}{2}t_H(W)\right] &\leq \mathbb{P}\left[|t_H(\mathcal{G}) - X_0| + |X_0 - t_H(W)| > \frac{1}{2}t_H(W)\right] \\
&\leq \mathbb{P}\left[|X_m - X_0| > \frac{1}{3}X_0\right] = \mathbb{P}\left[|Y_m - Y_0| > \frac{1}{3}Y_0\right] \\
&\leq Ce^{-Y_0^2/18n^2},
\end{aligned}$$

where the second inequality follows from $t_H(W) \geq \frac{8}{9}X_0$ given by (2.7). Again by the inequality (2.7), we obtain $Y_0 \geq \frac{n^2 t_H(W)}{2|E(H)|}$, which concludes the proof. □

In particular, letting $H = K_2$ gives

$$\mathbb{P}\left[t_{K_2}(\mathcal{G}) < \frac{1}{2}t_{K_2}(W)\right] \leq Ce^{-n^2 t_{K_2}(W)^2/72}. \tag{2.8}$$

If $t_{K_2}(W) \geq d$ for some fixed $d > 0$, then $t_{K_2}(\mathcal{G}) \geq d/2$ with very high probability $1 - Ce^{-\Omega(n^2)}$. This implies the following corollary:

**Corollary 2.13.** *Let $d, \rho$ be positive. Suppose that, for every subset $U \subseteq [n]$ with $|U| \geq \rho n$, $\sum_{u,v\in U} W(u,v) \geq d|U|^2$. Then with probability $1 - o(1)$ as $n \to \infty$, $\mathcal{G} = \mathcal{G}(n, W)$ is $(\rho, d/2)$-dense.*

*Proof.* We say that a subset $U \subseteq [n]$ is *bad* if it spans less than $d|U|^2/4$ edges. Replacing $\mathcal{G}$ and $[n]$ in (2.8) by the induced subgraph $\mathcal{G}[U]$ and $U$, respectively, gives that $U$ is bad with probability at most $Ce^{-d^2|U|^2/72}$. Therefore,

$$\begin{aligned}\mathbb{P}[\text{ there exists a bad subset } U \text{ with } |U| \geq \rho n\ ] &\leq \sum_{|U|\geq\rho n} \mathbb{P}[U \text{ is bad }] \\ &\leq C2^n e^{-\rho^2 d^2 n^2/72} = o(1),\end{aligned}$$

which implies that there is no bad $U$ with $|U| \geq \rho n$ with probability $1 - o(1)$. That is, $\mathcal{G}$ is $(\rho, d/2)$-dense. □

# Chapter 3

# Information theory and Sidorenko's conjecture

Information theory has been applied to combinatorial problems for decades. One important example is Shearer's lemma [6], which has been a powerful tool to obtain 'projection inequalities' that resemble the classical Loomis–Whitney inequality [35]. We refer to Galvin's survey [19] for more example applications and Friedgut's article [18] for generalisations of Shearer's lemma.

The information theoretic approach discussed in this chapter however has a different motivation, and hence does not use Shearer-type projection inequalities. In [33], Li and Szegedy applied Jensen inequalities for logarithmic functions and re-proved the partial result by Conlon, Fox, Sudakov [9] on Sidorenko's conjecture. Partly motivated by their work, Kim, Lee, and the author [28] used the logarithmic Jensen inequalities to generalise Conlon, Fox, Sudakov's result in a slightly different direction. As entropy inequalities generalise the logarithmic Jensen inequalities used in [28, 33], it is natural to ask whether an information-theoretic approach can be taken directly. This intuition has been formalised by Conlon, Kim, Lee, and the author [12] and independently by Szegedy [54], and both of the works obtained a common generalisation of the results in [28, 33]. In this chapter, we discuss this information theoretic approach, including the recent variants appeared in [13, 32].

Though it has been slightly simplified by subsequent work, e.g., Theorem 3.6, the original argument in [12] has some virtue as it gives a constructive approach that explicitly generates a homomorphic copy of $H$ using a randomised algorithm.

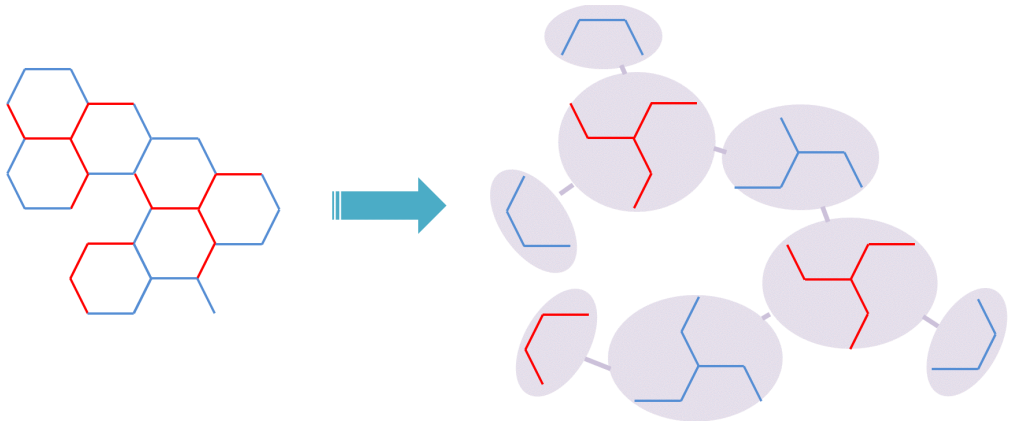


Figure 3.1: A strong tree decomposition.

Thus, in Section 3.1, we follow the algorithmic approach taken in [12], and further variants and refinements will be discussed in the subsequent sections.

## 3.1 Strongly tree-decomposable graphs

In this section, we prove Theorem 1.3, which we restate for the readers' convenience.

**Theorem 1.3** (Restated)**.** *Every strongly tree-decomposable graph has Sidorenko's property.*

The class of strongly tree-decomposable graphs consists of those graphs which are amenable to a certain embedding process using branching random walks.

For the definition of tree decompositions, see Section 2.1.4. A *strong tree decomposition* of a graph $H$ is a tree decomposition $(\mathcal{F}, \mathcal{T})$ of $H$ satisfying the following two extra conditions:

1. The induced subgraphs $H[X]$, $X \in \mathcal{F}$, are edge-disjoint trees.

2. For every pair $X, Y \in \mathcal{F}$ which are adjacent in $\mathcal{T}$, there is an isomorphism between the minimal subtrees of $H[X]$ and $H[Y]$ containing $X \cap Y$ that fixes $X \cap Y$.

A graph $H$ is *strongly tree-decomposable* if it allows a strong tree decomposition. One can easily verify that every strongly tree-decomposable graph is bipartite.

The family of strongly tree-decomposable graphs completely contains some other families for which the conjecture was previously known. For example, the family of reflection trees studied by Li and Szegedy [33] consists exactly of those strongly tree-decomposable graphs for which $\mathcal{T}$ is a star, while the family of tree-arrangeable graphs considered in [28] consists of those graphs $H$ such that $\mathcal{F} = \{N_H(a) \cup \{a\} \mid a \in A\}$ for an independent set $A$ in $V(H)$. However, our new family is broader than either of these, Figure 3.1 showing an example of a strongly tree-decomposable graph that is neither a reflection tree nor tree arrangeable.

The proof of Theorem 1.3 is based on a randomised algorithm that produces a homomorphism in $\mathrm{Hom}(H, G)$, with the desired bound for $|\mathrm{Hom}(H, G)|$ following from an appropriate estimate on the entropy of the random homomorphism.

### 3.1.1 Trees

We start with the case where $H = T$ is a tree. Let $r$ be a fixed vertex of $T$ and think of $T$ as a tree rooted at $r$. To generate a (not necessarily uniform) random homomorphism $\mathbf{w}_T$ from $T$ to $G$, we consider the *$T$-branching random walk*, or simply $T$-BRW, on $G$. If we fix an order $r = u_0, u_1, \cdots, u_t$ of all vertices in $T$ in which no child precedes its parent, this algorithm proceeds as follows:

1. For the root $r$, $\mathbf{w}_T(r)$ is a random vertex in $G$ chosen according to the stationary distribution, that is, with probability proportional to its degree,
$$\mathbb{P}[\mathbf{w}_T(r) = y] = \frac{\deg_G(y)}{2|E(G)|}.$$

2. Suppose that $u_i$, $i \geq 1$, is the first vertex that has not yet been embedded. Then, if $v_i$ is the parent of $u_i$, $\mathbf{w}_T(u_i)$ is a vertex chosen uniformly at random from all vertices that are adjacent to $\mathbf{w}_T(v_i)$ in $G$. (Note that $v_i = u_j$ for some $j < i$, as $v_i$ must precede $u_i$.)

   In others words, when the child $u_i$ of $v_i$ is born, it resides on a vertex of $G$ chosen uniformly at random from all vertices that are adjacent to $\mathbf{w}_T(v_i)$ in $G$.

3. Repeat Step 2 until all vertices have been embedded.

Unless $G$ has no edge, $\mathbf{w}_T$ is clearly a homomorphism from $T$ to $G$. Moreover, the distribution of $\mathbf{w}_T$ does not depend on either the choice of the root vertex $r$ or the order of the remaining vertices. To see this, fix $h \in \mathrm{Hom}(T, G)$ and note that

$$\begin{aligned}\mathbb{P}[\mathbf{w}_T = h] &= \frac{\deg_G(h(r))}{2|E(G)|}\left(\frac{1}{\deg_G(h(r))}\right)^{d_r}\prod_{v\in V(T)\setminus\{r\}}\left(\frac{1}{\deg_G(h(v))}\right)^{d_v-1}\\ &= \frac{1}{2|E(G)|}\prod_{v\in V(T)}\left(\frac{1}{\deg_G(h(v))}\right)^{d_v-1}, \qquad (3.1)\end{aligned}$$

where $d_v = \deg_T(v)$. As the last term depends only upon $T$ and $G$, the claim follows.

We will call the distribution of $\mathbf{w}_T$ on $\mathrm{Hom}(T, G)$ the *$T$-BRW distribution* and write $p_T : \mathrm{Hom}(T, G) \to [0, 1]$ for the probability distribution function of the $T$-BRW, i.e.,

$$p_T(h) = \mathbb{P}[\mathbf{w}_T = h].$$

Notice that (3.1) implies

$$p_T(h) > 0 \quad \text{for all } h \in \mathrm{Hom}(T, G). \qquad (3.2)$$

The $T$-BRW also has an interesting projection property: if $S$ is a fixed subtree of $T$, then, for an arbitrary homomorphism $h : V(T) \to V(G)$, the map $h$ restricted to $V(S)$, which is denoted by $h|_{V(S)}$, or simply $h|_S$, induces a homomorphism from $S$ to $G$. Hence, $\mathbf{w}_T|_S$ induces another probability distribution over $\mathrm{Hom}(S, G)$. The following proposition asserts that $\mathbf{w}_T|_S$ is identically distributed to $\mathbf{w}_S$. For a subset $X$ of $V(T)$ and $h : X \to V(G)$, let $\mathcal{G}_h(T)$ be the set of all homomorphisms from $T$ to $G$ that extend $h$, i.e., $\mathcal{G}_h(T) = \{g \in \mathrm{Hom}(T, G) : g = h \text{ on } X\}$.

**Proposition 3.1.** *Let $S$ be a subtree of a tree $T$ and $h \in \mathrm{Hom}(S, G)$. Then,*

$$\mathbb{P}\Big[\mathbf{w}_T = h \text{ on } S\Big] = p_S(h),$$

*where $\mathbf{w}_T$ is the $T$-BRW on $G$, or equivalently*

$$\sum_{g\in\mathcal{G}_h(T)} p_T(g) = p_S(h).$$

*Proof.* As the distribution of $\mathbf{w}_T$ is independent of the root as well as the order of the vertices, one may regard $\mathbf{w}_T$ as a $T$-BRW that maps vertices in $S$ before any other vertices in $T$. Then $\mathbf{w}_T|_S$ and $\mathbf{w}_S$ are constructed in the same way and the desired equality follows. □

For ordered edges $(u, v)$ and $(s, t)$ of $T$ and $G$, respectively, one may apply Proposition 3.1 to the tree $S$ consisting of the single edge $(u, v)$ and the homomorphism $h$ on $S$ with $(h(u), h(v)) = (s, t)$ to obtain

$$\mathbb{P}\Big[(\mathbf{w}_T(u), \mathbf{w}_T(v)) = (s, t)\Big] = p_S(h) = \frac{\deg_G(s)}{2|E(G)|}\frac{1}{\deg_G(s)} = \frac{1}{2|E(G)|}.$$

In other words, the $T$-BRW $\mathbf{w}_T$ restricted to any edge $(u, v)$ of $T$ has the uniform distribution on the set of all ordered edges in $G$. Therefore,

$$\mathbb{H}(\mathbf{w}_T(u), \mathbf{w}_T(v)) = \log(2|E(G)|). \tag{3.3}$$

We conclude this subsection by giving a lower bound for the entropy of the $T$-BRW $\mathbf{w}_T$, which in turn implies that trees satisfy Sidorenko's property.

**Proposition 3.2.** *For a tree $T$ and a graph $G$, the $T$-*BRW *$\mathbf{w}_T$ on $G$ satisfies*

$$\mathbb{H}(\mathbf{w}_T) \geq \log\left(|V(G)|^{|V(T)|}\left(\frac{2|E(G)|}{|V(G)|^2}\right)^{|E(T)|}\right)$$

*and therefore,*

$$|\mathrm{Hom}(T, G)| \geq |V(G)|^{|V(T)|}\left(\frac{2|E(G)|}{|V(G)|^2}\right)^{|E(T)|}.$$

*Proof.* The second part easily follows from the first part and Lemma 2.10(i), which implies that

$$\mathbb{H}(\mathbf{w}_T) \leq \log|\mathrm{Hom}(T, G)|.$$

For the first part, observe, by the definition of conditional entropy and Lemma 2.10(ii), that

$$\begin{aligned}\mathbb{H}(\mathbf{w}_T) &= \mathbb{H}(\mathbf{w}_T(r)) + \sum_{i=1}^{t}\mathbb{H}(\mathbf{w}_T(u_i)|\mathbf{w}_T(u_0, ..., u_{i-1})) \\ &= \mathbb{H}(\mathbf{w}_T(r)) + \sum_{i=1}^{t}\mathbb{H}(\mathbf{w}_T(u_i)|\mathbf{w}_T(v_i)),\end{aligned}$$

where $v_i$ is the parent of $u_i$. Denote by $C_v$ the set of all children of $v$. Rearranging the last sum, we have that

$$\begin{aligned}\mathbb{H}(\mathbf{w}_T) &= \mathbb{H}(\mathbf{w}_T(r)) + \sum_{v\in V(T)}\sum_{u\in C_v}\mathbb{H}(\mathbf{w}_T(u)|\mathbf{w}_T(v)) \\ &= \mathbb{H}(\mathbf{w}_T(r)) + \sum_{v\in V(T)}\sum_{u\in C_v}\left(\mathbb{H}(\mathbf{w}_T(u),\mathbf{w}_T(v)) - \mathbb{H}(\mathbf{w}_T(v))\right).\end{aligned}$$

Since, by (3.3),

$$\sum_{v\in V(T)}\sum_{u\in C_v}\mathbb{H}(\mathbf{w}_T(u),\mathbf{w}_T(v)) = \sum_{uv\in E(T)}\mathbb{H}(\mathbf{w}_T(u),\mathbf{w}_T(v)) = |E(T)|\log(2|E(G)|)$$

and, by Lemma 2.10(i),

$$\begin{aligned}-\mathbb{H}(\mathbf{w}_T(r)) + \sum_{v\in V(T)}\sum_{u\in C_v}\mathbb{H}(\mathbf{w}_T(v)) &= \sum_{v\in V(T)}(\deg_T(v)-1)\mathbb{H}(\mathbf{w}_T(v)) \\ &\le (2|E(T)| - |V(T)|)\log|V(G)|,\end{aligned}$$

we have that

$$\mathbb{H}(\mathbf{w}_T) \ge \log\Big(|V(G)|^{|V(T)|}\Big(\frac{2|E(G)|}{|V(G)|^2}\Big)^{|E(T)|}\Big),$$

as desired. □

### 3.1.2 Branching random walks starting from a partial embedding

In order to embed strongly tree-decomposable graphs, we will first need to understand $T$-BRWs starting from a partial embedding. Let $T$ be a tree, $G$ a graph with at least one edge, and $X$ a subset of $V(T)$. Let $S$ be the minimal subtree of $T$ that contains $X$. If $X = \emptyset$, then $S$ consists of a single fixed vertex of $T$. Take an ordering of the vertices of $T$ such that the vertices of $S$ precede all other vertices and, if the tree $T$ is seen as being rooted at the first vertex, no child precedes its parent. The *$T$-branching random walk* on $G$, or simply $T$-BRW, starting from a partial embedding $h : X \to V(G)$ is then defined as follows:

1. If $\mathcal{G}_h(S) := \{g \in \mathrm{Hom}(S,G) : g|_X = h\}$ is empty, then there is no $T$-BRW starting with $h$. Otherwise, $\mathbf{w}_T(\cdot\,|\,h) : V(T) \to V(G)$ restricted to $S$ is defined according to the distribution

$$\mathbb{P}\Big[\,\mathbf{w}_T(\cdot\,|\,h) = y \text{ on } S\,\Big] = \begin{cases} \frac{p_S(y)}{\sum_{g\in\mathcal{G}_h(S)}p_S(g)} & \text{if } y \in \mathcal{G}_h(S) \\ 0 & \text{if } y \notin \mathcal{G}_h(S)\ . \end{cases}$$

(Recall that $p_S$ is the probability distribution function of the $S$-BRW on $\mathrm{Hom}(S,G)$ and $p_S(g) > 0$ for all $g \in \mathrm{Hom}(S,G)$ as seen in (3.2) above).

2. Once $\mathbf{w}_T(\cdot \,|\, h)$ is defined on $S$, the embedding $\mathbf{w}_T(\cdot \,|\, h)$ may be defined in the same way that was used for the construction of the $T$-BRW. That is, for the first vertex $u$ that has not yet been embedded and its parent $v$, $\mathbf{w}_T(u \,|\, h)$ is a vertex chosen uniformly at random from all vertices adjacent to $\mathbf{w}_T(v \,|\, h)$ in $G$.

3. Repeat Step 2 until all vertices have been embedded.

If $X = \emptyset$, the $T$-BRW starting from $h : X \to V(G)$ is the usual $T$-BRW. Moreover, it is not difficult to show that the $T$-BRW starting from $h$ has the same distribution as the usual $T$-BRW conditioned on $\mathbf{w}_T|_X = h$. We conclude this subsection by proving this fact.

**Proposition 3.3.** *With the same notation as above and $x \in \mathrm{Hom}(T,G)$,*

$$\mathbb{P}\Big[\, \mathbf{w}_T(\cdot|h) = x \text{ on } S \,\Big] = \mathbb{P}\Big[\mathbf{w}_T = x \text{ on } S \;\Big|\; \mathbf{w}_T = h \text{ on } X\Big]$$

*and*

$$\mathbb{P}[\mathbf{w}_T(\cdot|h) = x] = \mathbb{P}\Big[\mathbf{w}_T = x \;\Big|\; \mathbf{w}_T = h \text{ on } X \,\Big].$$

*In particular, the distribution of $\mathbf{w}_T(\cdot|h)$ is independent of the order of vertices in $T$.*

*Proof.* As Proposition 3.1 gives

$$p_S(x|_S) = \mathbb{P}[\mathbf{w}_T = x \text{ on } S]$$

and

$$\sum_{g \in \mathcal{G}_h(S)} p_S(g) = \sum_{g \in \mathcal{G}_h(S)} \sum_{f \in \mathcal{G}_g(T)} p_T(f) = \sum_{f \in \mathcal{G}_h(T)} p_T(f) = \mathbb{P}[\mathbf{w}_T = h \text{ on } X],$$

we have that

$$\mathbb{P}\Big[\, \mathbf{w}_T(\cdot|h) = x \text{ on } S \,\Big] = \frac{p_S(x|_S)}{\displaystyle\sum_{g \in \mathcal{G}_h(S)} p_S(g)} = \frac{\mathbb{P}[\mathbf{w}_T = x \text{ on } S]}{\mathbb{P}[\mathbf{w}_T = h \text{ on } X]} = \mathbb{P}\Big[\mathbf{w}_T = x \text{ on } S \;\Big|\; \mathbf{w}_T = h \text{ on } X\Big].$$

Since $\mathbf{w}_T$ and $\mathbf{w}_T(\cdot|h)$ are constructed in the same way once their values on $S$ have been determined, we also have

$$\begin{aligned}\mathbb{P}[\mathbf{w}_T(\cdot|h) = x] &= \mathbb{P}\Big[\mathbf{w}_T(\cdot|h) = x \text{ on } S\Big]\mathbb{P}\Big[\mathbf{w}_T(\cdot|h) = x \,\Big|\, \mathbf{w}_T(\cdot|h) = x \text{ on } S\Big] \\ &= \mathbb{P}\Big[\mathbf{w}_T = x \text{ on } S \,\Big|\, \mathbf{w}_T = h \text{ on } X\Big]\mathbb{P}\Big[\mathbf{w}_T = x \,\Big|\, \mathbf{w}_T = x \text{ on } S\Big] \\ &= \mathbb{P}\Big[\mathbf{w}_T = x \,\Big|\, \mathbf{w}_T = h \text{ on } X\Big],\end{aligned}$$

as required. □

### 3.1.3 Embedding strongly tree-decomposable graphs

Recall that a graph $H$ is strongly tree decomposable if it has a tree decomposition $(\mathcal{F}, \mathcal{T})$ satisfying two extra conditions:

1. The induced subgraphs $H[X]$, $X \in \mathcal{F}$, are edge-disjoint trees.

2. For every pair $X, Y \in \mathcal{F}$ which are adjacent in $\mathcal{T}$, there is an isomorphism that fixes $X \cap Y$ between the minimal subtrees of $H[X]$ and $H[Y]$ containing $X \cap Y$.

Suppose that $G$ has at least one edge. We will attempt to construct a homomorphism $\mathbf{w}$ of $H$ into $G$ by regarding $\mathcal{T}$ as a rooted tree with root $R$, for a fixed $R \in \mathcal{F}$, and taking a fixed order $R = X_0, X_1, X_2, ..., X_t$ of all sets in $\mathcal{F}$ in which no child precedes its parent. We then proceed as follows:

1. Take $\mathbf{w}_{H[R]}$, the $H[R]$-BRW, and set $\mathbf{w} = \mathbf{w}_{H[R]}$ on $R$. Mark $R$ as explored.

2. Suppose $\mathbf{w}$ is defined on $Z_{i-1} := \bigcup_{j=0}^{i-1} X_i$, $i \geq 1$, and $X_i$ is the first set that has not yet been explored. Take $\mathbf{w}_{H[X_i]}(\cdot|h)$, the $H[X_i]$-BRW starting from $h = \mathbf{w}|_{X_i \cap Z_{i-1}}$, and set $\mathbf{w} = \mathbf{w}_{H[X_i]}(\cdot|h)$ on $X_i \setminus Z_{i-1}$, provided that there exists a $H[X_i]$-BRW starting from $h$. Mark $X_i$ as explored.

   If there is no $H[X_i]$-BRW starting from $h$, then the process stops and there is no output. (We will presently show that this case cannot occur.)

3. Repeat Step 2 until every set in $\mathcal{F}$ has been explored.

We first show that there is always an output $\mathbf{w}$ and it is a homomorphism from $H$ to $G$.

**Lemma 3.4.** *The above process always outputs a homomorphism from $H$ to $G$. Moreover, the distribution of the homomorphism does not depend on the choice of the order $X_1, ..., X_t$ of the sets in $\mathcal{F}$ provided that $X_0$ is a fixed set $R$.*

*Proof.* We will prove inductively that $\mathbf{w}$ is a homomorphism from $H[Z_i]$ to $G$ for $i = 0, 1, ..., t$. For $i = 0$, this follows since $G$ has at least one edge and $H[R]$ is a tree. Hence, the $H[R]$-BRW is well-defined and yields a homomorphism from $H[R]$ to $G$. For $i \geq 1$, suppose $\mathbf{w}|_{Z_{i-1}}$ is a homomorphism. Since $X_i$ is a leaf of the subtree $\mathcal{T}_i := \mathcal{T}[\{X_0, X_1, X_2, ..., X_i\}]$ of $\mathcal{T}$, every non-trivial path in $\mathcal{T}_i$ containing $X_i$ must contain its parent $Y_i$ and the third property of tree decompositions gives

$$X_i \cap Z_{i-1} = X_i \cap Y_i.$$

Since $H$ is strongly tree decomposable, there is an isomorphism that fixes $X_i \cap Y_i$ between the minimal subtrees of $H[X_i]$ and $H[Y_i]$ containing $X_i \cap Y_i$. Since $Y_i \subset Z_{i-1}$, the relevant subtree of $H[Y_i]$ has already been embedded by $\mathbf{w}$. In particular, there is a homomorphism, say $g$, from the relevant subtree of $H[X_i]$ to $G$ with $g = \mathbf{w}$ on $X_i \cap Y_i$. Hence, by the results of Section 3.1.2, there exists an $H[X_i]$-BRW starting from $\mathbf{w}|_{X_i \cap Y_i}$. This shows that $\mathbf{w}$ is a well-defined mapping from $Z_i = Z_{i-1} \cup X_i$ to $V(G)$.

To show that $\mathbf{w}$ is a homomorphism on $Z_i$, suppose $\{u, v\}$ is an edge of $H$ with $u, v \in Z_i$. Since both $\mathbf{w}|_{Z_{i-1}}$ and $\mathbf{w}|_{X_i}$ are homomorphisms, it is enough to show that either $u, v \in Z_{i-1}$ or $u, v \in X_i$, though not necessarily exclusively. If not, we may assume that $u \in X_i$ and $v \in X_j \subset Z_{i-1}$ for some $j < i$. We now take a set $U \in \mathcal{F}$ that contains both $u$ and $v$. Such a set exists by the second property of tree decompositions. If $U$ is a descendant of $X_i$ in $\mathcal{T}$, then the path from $U$ to $X_j$ in $\mathcal{T}$ must go through $X_i$. As $v \in U \cap X_j$, we have $v \in X_i$ and hence $u, v \in X_i$, contradicting our assumption. If $U$ is not a descendant of $X_i$ in $\mathcal{T}$, then the path from $X_i$ to $U$ in $\mathcal{T}$ must contain $Y_i$, the parent of $X_i$. Since $u \in X_i \cap U$, we have $u \in Y_i \subset Z_{i-1}$ and hence $u, v \in Z_{i-1}$, again contradicting our assumption.

Therefore, the process constructing $\mathbf{w}$ continues until $\mathbf{w}$ has been extended to all sets in $\mathcal{F}$. Since $\bigcup_{X \in \mathcal{F}} X = V(H)$ (by the first property of tree decompositions),

$\mathbf{w}$ is a homomorphism from $H$ to $G$. Moreover, using

$$\mathbb{P}\Big[\mathbf{w} = h\Big] = \mathbb{P}[\mathbf{w} = h \text{ on } X_0] \prod_{i=1}^{t} \mathbb{P}\left[\mathbf{w} = h \text{ on } X_i \;\middle|\; \mathbf{w} = h \text{ on } Z_{i-1}\right]$$

and

$$\mathbb{P}\left[\mathbf{w} = h \text{ on } X_i \;\middle|\; \mathbf{w} = h \text{ on } Z_{i-1}\right] = \mathbb{P}\left[\mathbf{w} = h \text{ on } X_i \;\middle|\; \mathbf{w} = h \text{ on } X_i \cap Y_i\right],$$

we easily see that if $C_Y$ is the set of children of $Y$ in $\mathcal{T}$, then

$$\begin{aligned}\mathbb{P}\Big[\mathbf{w} = h\Big] &= \mathbb{P}\Big[\mathbf{w} = h \text{ on } X_0\Big] \prod_{i=1}^{t} \mathbb{P}\left[\mathbf{w} = h \text{ on } X_i \;\middle|\; \mathbf{w} = h \text{ on } X_i \cap Y_i\right] \\ &= \mathbb{P}\Big[\mathbf{w} = h \text{ on } X_0\Big] \prod_{Y \in \mathcal{F}} \prod_{X \in C_Y} \mathbb{P}\left[\mathbf{w} = h \text{ on } X \;\middle|\; \mathbf{w} = h \text{ on } X \cap Y\right].\end{aligned}$$

Similarly, for any other order $R = X_0 = X_0', X_1', X_2', ..., X_t'$ in which no child precedes its parent,

$$\mathbb{P}\Big[\mathbf{w}' = h\Big] = \mathbb{P}\Big[\mathbf{w}' = h \text{ on } X_0'\Big] \prod_{Y \in \mathcal{F}} \prod_{X \in C_Y} \mathbb{P}\left[\mathbf{w}' = h \text{ on } X \;\middle|\; \mathbf{w}' = h \text{ on } X \cap Y\right].$$

On $X_0 = X_0' = R$, both $\mathbf{w}$ and $\mathbf{w}'$ are $H[R]$-BRWs and, hence,

$$\mathbb{P}[\mathbf{w}' = h \text{ on } X_0'] = \mathbb{P}[\mathbf{w} = h \text{ on } X_0].$$

Further, for all adjacent pairs $X, Y \in \mathcal{F}$, conditioned on $\mathbf{w}' = \mathbf{w} = h$ on $X \cap Y$, both $\mathbf{w}|_X$ and $\mathbf{w}'|_X$ are $H[X]$-BRWs starting from $h$ and

$$\mathbb{P}\left[\mathbf{w}' = h \text{ on } X \;\middle|\; \mathbf{w}' = h \text{ on } X \cap Y\right] = \mathbb{P}\left[\mathbf{w} = h \text{ on } X \;\middle|\; \mathbf{w} = h \text{ on } X \cap Y\right].$$

Thus, $\mathbb{P}[\mathbf{w} = h] = \mathbb{P}[\mathbf{w}' = h]$, which means that the distribution of $\mathbf{w}$ is independent of the order. □

In the following proposition, we show that the choice of root $R$ is also irrelevant.

**Proposition 3.5.** *The distribution of the homomorphism* $\mathbf{w}$ *described above does not depend on the choice of root* $R$*. In particular,* $\mathbf{w}|_X$ *is the* $H[X]$*-BRW for each* $X \in \mathcal{F}$*, i.e.,*

$$\mathbb{P}[\mathbf{w} = h \text{ on } X] = p_{H[X]}(h|_X)$$

*for* $h \in \mathrm{Hom}(H, G)$*.*

*Proof.* Since $X$ in the second part may be chosen as the root of $\mathcal{T}$, the equality easily follows from the first part. To prove the first part, let $\mathbf{w}_1$ and $\mathbf{w}_2$ be the random homomorphisms constructed using the above procedure with roots $Y_1, Y_2 \in \mathcal{F}$, respectively. We will show that $\mathbf{w}_1$ and $\mathbf{w}_2$ have the same distribution whenever $Y_1$ and $Y_2$ are adjacent in $\mathcal{T}$. The general case follows easily from repeated application of this special case.

Since the order of sets in $\mathcal{F}$ other than the root does not affect the distributions of $\mathbf{w}_1$ and $\mathbf{w}_2$, we may embed $Y_2$ immediately after $Y_1$ when $Y_1$ is the root. If, on the other hand, $Y_2$ is the root, then $Y_1$ is to be embedded immediately after $Y_2$. Order all remaining sets in $\mathcal{F}$ in the exact same order for both cases. It therefore suffices to prove that, for each $h \in \mathrm{Hom}(H, G)$,

$$\begin{aligned}&\mathbb{P}[\mathbf{w}_1 = h \text{ on } Y_1] \cdot \mathbb{P}[\mathbf{w}_1 = h \text{ on } Y_2 \mid \mathbf{w}_1 = h \text{ on } Y_1 \cap Y_2] \\ &= \mathbb{P}[\mathbf{w}_2 = h \text{ on } Y_2] \cdot \mathbb{P}[\mathbf{w}_2 = h \text{ on } Y_1 \mid \mathbf{w}_2 = h \text{ on } Y_1 \cap Y_2]. \end{aligned} \tag{3.4}$$

We will first show that if $S_1$ and $S_2$ are the minimal subtrees of $H[Y_1]$ and $H[Y_2]$ containing $Y_1 \cap Y_2$, respectively, then

$$\mathbb{P}[\mathbf{w}_1 = h \text{ on } Y_2 \mid \mathbf{w}_1 = h \text{ on } Y_1 \cap Y_2] = \frac{\mathbb{P}[\mathbf{w}_2 = h \text{ on } Y_2]}{\sum_{g \in \mathcal{G}_h(S_2)} p_{S_2}(g)}, \tag{3.5}$$

and, similarly,

$$\mathbb{P}[\mathbf{w}_2 = h \text{ on } Y_1 \mid \mathbf{w}_2 = h \text{ on } Y_1 \cap Y_2] = \frac{\mathbb{P}[\mathbf{w}_1 = h \text{ on } Y_1]}{\sum_{g \in \mathcal{G}_h(S_1)} p_{S_1}(g)}, \tag{3.6}$$

where $\mathcal{G}_h(S_i) = \{g \in \mathrm{Hom}(S_i, G) : g = h \text{ on } Y_1 \cap Y_2\}$. Note that

$$\begin{aligned}\mathbb{P}\Big[\mathbf{w}_1 = h \text{ on } Y_2 \;\Big|\; \mathbf{w}_1 = h \text{ on } Y_1 \cap Y_2\Big] &= \mathbb{P}\Big[\mathbf{w}_1 = h \text{ on } S_2 \;\Big|\; \mathbf{w}_1 = h \text{ on } Y_1 \cap Y_2\Big] \\ &\quad \times \mathbb{P}\Big[\mathbf{w}_1 = h \text{ on } Y_2 \;\Big|\; \mathbf{w}_1 = h \text{ on } S_2\Big] \\ &= \frac{p_{S_2}(h|_{S_2})}{\sum_{g \in \mathcal{G}_h(S_2)} p_{S_2}(g)} \mathbb{P}\Big[\mathbf{w}_1 = h \text{ on } Y_2 \;\Big|\; \mathbf{w}_1 = h \text{ on } S_2\Big],\end{aligned}$$

as $\mathbb{P}[\mathbf{w}_1 = h \text{ on } S_2, \mathbf{w}_1 = h \text{ on } Y_1 \cap Y_2] = \mathbb{P}[\mathbf{w}_1 = h \text{ on } S_2] = p_{S_2}(h|_{S_2})$. Hence, (3.5) follows since

$$\begin{aligned}\mathbb{P}\Big[\mathbf{w}_2 = h \text{ on } Y_2\Big] &= \mathbb{P}\Big[\mathbf{w}_2 = h \text{ on } S_2\Big] \mathbb{P}\Big[\mathbf{w}_2 = h \text{ on } Y_2 \;\Big|\; \mathbf{w}_2 = h \text{ on } S_2\Big] \\ &= p_{S_2}(h|_{S_2}) \mathbb{P}\Big[\mathbf{w}_1 = h \text{ on } Y_2 \;\Big|\; \mathbf{w}_1 = h \text{ on } S_2\Big],\end{aligned}$$

where, in the last equality, we used the fact that $\mathbf{w}_1$ and $\mathbf{w}_2$ are constructed on $Y_2$ in the same way once their values in $S_2$ are determined to be $h$. The same argument also works for (3.6).

The desired equality (3.4) follows since

$$\sum_{g\in\mathcal{G}_h(S_1)} p_{S_1}(g) = \sum_{g\in\mathcal{G}_h(S_2)} p_{S_2}(g),$$

(recall that $\mathcal{G}_h(S_i) = \{g \in \mathrm{Hom}(S_i, G) : g = h \text{ on } Y_1 \cap Y_2\}$), which, in turn, easily follows from the fact that there is an isomorphism from $S_1$ to $S_2$ that fixes $Y_1 \cap Y_2$. □

We are now in a position to prove Theorem 1.3.

*Proof of Theorem 1.3.* Let $H$ be a strongly tree-decomposable graph with tree decomposition $(\mathcal{F}, \mathcal{T})$ and let $G$ be a graph with at least one edge. Denote by $\mathbf{w}(X) := \mathbf{w}|_X$. Using

$$\mathbb{H}(\mathbf{w}) = \mathbb{H}(\mathbf{w}(R)) + \sum_{i=1}^{t} \mathbb{H}(\mathbf{w}(X_i)|\mathbf{w}(Z_{i-1}))$$

and, from Lemma 2.10(ii),

$$\mathbb{H}(\mathbf{w}(X_i)|\mathbf{w}(Z_{i-1})) = \mathbb{H}(\mathbf{w}(X_i)|\mathbf{w}(X_i \cap Z_{i-1})) = \mathbb{H}(\mathbf{w}(X_i)|\mathbf{w}(X_i \cap Y_i)),$$

where $Y_i$ is the parent of $X_i$ in the tree $\mathcal{T}$ rooted at $R$, we have that

$$\begin{aligned}\mathbb{H}(\mathbf{w}) &= \mathbb{H}(\mathbf{w}(R)) + \sum_{i=1}^{t} \mathbb{H}(\mathbf{w}(X_i)|\mathbf{w}(X_i \cap Y_i)) \\ &= \mathbb{H}(\mathbf{w}(R)) + \sum_{Y\in\mathcal{F}} \sum_{X\in C_Y} \mathbb{H}(\mathbf{w}(X)|\mathbf{w}(X \cap Y)),\end{aligned}$$

where $C_Y$ is the set of children of $Y$. As $\mathbb{H}(\mathbf{w}(X)|\mathbf{w}(X \cap Y)) = \mathbb{H}(\mathbf{w}(X)) - \mathbb{H}(\mathbf{w}(X \cap Y))$, we also have that

$$\begin{aligned}\mathbb{H}(\mathbf{w}) &= \mathbb{H}(\mathbf{w}(R)) + \sum_{Y\in\mathcal{F}} \sum_{X\in C_Y} \mathbb{H}(\mathbf{w}(X)) - \mathbb{H}(\mathbf{w}(X \cap Y)) \\ &= \sum_{X\in\mathcal{F}} \mathbb{H}(\mathbf{w}(X)) - \sum_{XY\in E(\mathcal{T})} \mathbb{H}(\mathbf{w}(X \cap Y)) \\ &\geq \sum_{X\in\mathcal{F}} \mathbb{H}(\mathbf{w}(X)) - \sum_{XY\in E(\mathcal{T})} |X \cap Y| \log |V(G)|, \qquad (3.7)\end{aligned}$$

where the inequality follows from Lemma 2.10(i). Proposition 3.5 implies that, for each $X \in \mathcal{F}$, the distribution of $\mathbf{w}|_X$ is the $H[X]$-BRW distribution and, therefore, by Proposition 3.2,

$$\mathbb{H}(\mathbf{w}(X)) \geq \log\Big(|V(G)|^{|X|}\Big(\frac{2|E(G)|}{|V(G)|^2}\Big)^{|E(H[X])|}\Big) = |E(H[X])|\log\frac{2|E(G)|}{|V(G)|^2} + |X|\log|V(G)|.$$

Since strong tree decomposability implies that

$$\sum_{X\in\mathcal{F}} |E(H[X])| = |E(H)| \quad \text{and} \quad \sum_{X\in\mathcal{F}} |X| - \sum_{XY\in E(\mathcal{T})} |X\cap Y| = |V(H)|,$$

it follows that

$$\begin{aligned} H(\mathbf{w}) &\geq \sum_{X\in\mathcal{F}} |E(H[X])|\log\frac{2|E(G)|}{|V(G)|^2} + \Big(\sum_{X\in\mathcal{F}} |X| - \sum_{XY\in E(\mathcal{T})} |X\cap Y|\Big)\log|V(G)| \\ &= |E(H)|\log\frac{2|E(G)|}{|V(G)|^2} + |V(H)|\log|V(G)| \\ &= \log\Big(|V(G)|^{|V(H)|}\Big(\frac{2|E(G)|}{|V(G)|^2}\Big)^{|E(H)|}\Big)\,. \end{aligned}$$

Since $H(\mathbf{w}) \leq \log|\mathrm{Hom}(H,G)|$, we finally have that

$$\log|\mathrm{Hom}(H,G)| \geq \log\Big(|V(G)|^{|V(H)|}\Big(\frac{2|E(G)|}{|V(G)|^2}\Big)^{|E(H)|}\Big),$$

or

$$|\mathrm{Hom}(H,G)| \geq |V(G)|^{|V(H)|}\Big(\frac{2|E(G)|}{|V(G)|^2}\Big)^{|E(H)|},$$

as desired. □

## 3.2 $J$-decomposable graphs

In this section, we simplify the information-theoretic approach taken in Section 3.1, leading to a more abstract, if less constructive, approach, comparable to Szegedy's independent work [54]. Our main general statement, Theorem 3.6, generalises Theorem 1.4 and we will also prove that it implies a result of Conlon and the author [13] and a result of Sidorenko [50]. We remark that Theorem 1.4 first appeared in [32], though was implicitly used in the earlier work [13].

### 3.2.1 Proof of Theorem 1.4

Let $\mathcal{F}$ be a family of subsets of $[k] := \{1, 2, \cdots, k\}$. Partly motivated by the notion of tree decompositions, a *Markov tree* on $[k]$ is a pair $(\mathcal{F}, \mathcal{T})$ with $\mathcal{T}$ a tree on vertex set $\mathcal{F}$ that satisfies

1. $\bigcup_{F \in \mathcal{F}} F = [k]$ and
2. for $A, B, C \in \mathcal{F}$, $A \cap B \subseteq C$ whenever $C$ lies on the path from $A$ to $B$ in $\mathcal{T}$.

Let $V$ be a finite set and for each $F \in \mathcal{F}$ let $\mathbf{X}_F = (X_{i;F})_{i \in F}$ be a random vector taking values in $V^F$. We are interested in such random vectors where 'local' information is 'globally' extendible. That is, there exist random variables $Y_1, Y_2, \cdots, Y_k$ such that, for each $F \in \mathcal{F}$, the two random vectors $(Y_i)_{i \in F}$ and $(X_{i;F})_{i \in F}$ are identically distributed over $V^F$. If such $Y_1, \cdots, Y_k$ exist, then $(X_{i;A})_{i \in A \cap B}$ and $(X_{j;B})_{j \in A \cap B}$ must be identically distributed. Our main theorem states that the converse is also true and, moreover, the maximum entropy under such constraints can always be attained.

**Theorem 3.6.** *Let $(\mathcal{F}, \mathcal{T})$ be a Markov tree on $[k]$. Let $V$ be a finite set and for each $F \in \mathcal{F}$ let $\mathbf{X}_F = (X_{i;F})_{i \in F}$ be a random vector taking values in $V^F$. If $(X_{i;A})_{i \in A \cap B}$ and $(X_{j;B})_{j \in A \cap B}$ are identically distributed whenever $AB \in E(\mathcal{T})$, then there exists $\mathbf{Y} = (Y_1, \cdots, Y_k)$ with entropy*

$$\mathbb{H}(\mathbf{Y}) = \sum_{F \in \mathcal{F}} \mathbb{H}(\mathbf{X}_F) - \sum_{AB \in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i \in A \cap B}) \tag{3.8}$$

*such that $(Y_i)_{i \in F}$ and $\mathbf{X}_F$ are identically distributed over $V^F$ for all $F \in \mathcal{F}$.*

Before getting into the proof, let us discuss how Theorem 3.6 relates to the classical Cauchy–Schwarz inequality. For a simple example, let $V = V(G)$ be the vertex set of a graph $G$, and let $(X, Y)$ and $(Y', Z)$ be two uniform random labelled edges. Since the distributions of $Y$ and $Y'$ are identical, Theorem 3.6 implies that there exists $(X_1, X_2, X_3)$ with entropy

$$\mathbb{H}(X_1, X_2, X_3) = \mathbb{H}(X, Y) + \mathbb{H}(Y', Z) - \mathbb{H}(Y).$$

Using basic facts on entropy (see Lemma 2.10), we have $\mathbb{H}(X, Y) = \mathbb{H}(Y', Z) = \log 2|E(G)|$, $\mathbb{H}(Y) \leq \log |V(G)|$, and $\mathbb{H}(X_1, X_2, X_3) \leq \log |\mathrm{Hom}(K_{1,2}, G)|$, which

implies $t_{K_{1,2}}(G) \geq t_{K_2}(G)^2$. This is also an easy consequence of the Cauchy–Schwarz inequality, and we may recover many graph homomorphism inequalities obtained using the Cauchy–Schwarz inequality by letting $|\mathcal{F}| = 2$ and $\mathcal{T}$ the single edge tree. Hence, (3.8) may be seen as a tree-like extension of the Cauchy–Schwarz inequality.

Indeed, one may recover Theorem 1.3 by using Theorem 3.6. Though the new proof may be less constructive, it is somewhat simpler and hence allows us to generalise Theorem 1.3. In particular, repeated proofs of invariance principles in Lemma 3.4 and in Proposition 3.5 can be omitted. We will give more details in Section 3.3.

*Proof of Theorem 3.6.* We use induction on $|\mathcal{F}|$. The base case $|\mathcal{F}| = 1$ is trivially true. Fix a leaf $L$ of $\mathcal{T}$ and let $\mathcal{T}'$ be the tree $\mathcal{T} \setminus L$ on $\mathcal{F}' := \mathcal{F} \setminus \{L\}$. By rearranging indices, we may assume that $L = \{t, t+1, \cdots, k\}$ for some $t \leq k$ and that $\mathcal{F}'$ satisfies $\cup_{F \in \mathcal{F}'} F = [\ell]$ for some $t \leq \ell \leq k$. Let $P$ be the neighbour of $L$ in $\mathcal{T}$. Then $\{t, t+1, \cdots, \ell\} = L \cap P$. By the inductive hypothesis, there is $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_\ell)$ such that $\mathbf{Y}_F := (Y_i)_{i \in F}$ and $\mathbf{X}_F := (X_{i;F})_{i \in F}$ are identically distributed for each $F \in \mathcal{F}'$ and moreover,

$$\mathbb{H}(\mathbf{Y}) = \sum_{F \in \mathcal{F}'} \mathbb{H}(\mathbf{X}_F) - \sum_{AB \in E(\mathcal{T}')} \mathbb{H}((X_{i;A})_{i \in A \cap B}). \tag{3.9}$$

Since $\mathbf{Y}_P$ and $\mathbf{X}_P$ are identically distributed and $\{t, t+1, \cdots, \ell\} \subseteq P$, $(Y_t, Y_{t+1}, \cdots, Y_\ell)$ and $(X_{t;L}, X_{t+1;L}, \cdots, X_{\ell;L})$ are identically distributed. Thus, we may apply Lemma 2.9 with

$$\begin{aligned} &X_1 = (Y_1, Y_2, \cdots, Y_{t-1}), \quad X_2 = (Y_t, Y_{t+1}, \cdots, Y_\ell), \\ &X_2' = (X_{t;L}, X_{t+1;L}, \cdots, X_{\ell;L}), \text{ and } X_3 = (X_{\ell+1;L}, X_{\ell+2;L}, \cdots, X_{k;L}). \end{aligned}$$

Then there exists $(Z_1, Z_2, \cdots, Z_k)$ such that $(Z_1, Z_2, \cdots, Z_{t-1})$ and $(Z_{\ell+1}, \cdots, Z_k)$ are conditionally independent given $(Z_t, Z_{t+1}, \cdots, Z_\ell)$, $(Z_1, \cdots, Z_\ell)$ and $\mathbf{Y}$ are identically distributed, and $(Z_t, Z_{t+1}, \cdots, Z_k)$ and $\mathbf{X}_L$ are identically distributed. By conditional independence, we obtain

$$\mathbb{H}(Z_1, Z_2, \cdots, Z_k) = \mathbb{H}(\mathbf{Y}) + \mathbb{H}(\mathbf{X}_L) - \mathbb{H}(Y_t, Y_{t+1}, \cdots, Y_\ell).$$

Now (3.9) and the fact that $\{t, t+1, \cdots, \ell\} = L \cap P$ implies (3.8). □

*Remark* 3.7. We say that $(\mathcal{F}', \mathcal{T}')$ is a *Markov subtree* if $\mathcal{T}'$ is a subtree of $\mathcal{T}$ induced on $\mathcal{F}' \subseteq \mathcal{F}$. One may check that a Markov subtree of a Markov tree is again Markov tree on the subset $K' := \cup_{F \in \mathcal{F}} F$ of $[k]$. Thus, replacing $(\mathcal{F}, \mathcal{T})$ by $(\mathcal{F}', \mathcal{T}')$ in Theorem 3.6 gives that there exists $(Y_i')_{i \in K'}$ such that $(Y_i')_{i \in F}$ and $(X_{i;F})_{i \in F}$ are identically distributed over $V^F$ for all $F \in \mathcal{F}'$ and

$$\mathbb{H}((Y_i')_{i \in K'}) = \sum_{F \in \mathcal{F}'} \mathbb{H}((X_{i;F})_{i \in F}) - \sum_{AB \in E(\mathcal{T}')} \mathbb{H}((X_{i;A})_{i \in A \cap B})$$

Moreover, the inductive proof above tells us that the random vectors $(Y_i)_{i \in K'}$ and $(Y_i')_{i \in K'}$ are identically distributed, that is, marginal distribution on $K'$ is preserved while extending $(Y_i')_{i \in K'}$ to $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_k)$. This generalises both Proposition 3.1 and Proposition 3.5.

Given a graph $J$, an *$J$-decomposition* of a graph $H$ is a tree decomposition $(\mathcal{F}, \mathcal{T})$ of $H$ satisfying the following two extra conditions:

1. The induced subgraphs $H[X]$, $X \in \mathcal{F}$, are each isomorphic to $J$.
2. For every pair $X, Y \in \mathcal{F}$ which are adjacent in $\mathcal{T}$, there is an isomorphism between the two copies of $J$ that fixes $X \cap Y$.

We say that a graph is *$J$-decomposable* if it allows a $J$-decomposition, i.e., it can be obtained by gluing copies of the graph $J$ in a tree-like way. We are now ready to prove Theorem 1.4.

**Theorem 1.4** (Restated)**.** *Let $G$, $H$, and $J$ be graphs. Suppose that $H$ is $J$-decomposable and* $\mathrm{Hom}(J, G)$ *is non-empty. Fix a $J$-decomposition $(\mathcal{F}, \mathcal{T})$ of $H$. Then the following inequality holds:*

$$t_H(G) \geq \frac{t_J(G)^{|\mathcal{F}|}}{\prod_{XY \in E(\mathcal{T})} t_{H[X \cap Y]}(G)}.$$

*Proof.* By definition, a $J$-decomposition $(\mathcal{F}, \mathcal{T})$ is a Markov tree on $V(H)$. Let $\mathbf{X}_F = (X_{i,F})_{i \in F}$ be the uniform random homomorphism in $\mathrm{Hom}(J, G)$. Then both $(X_{i,A})_{i \in A \cap B}$ and $(X_{j,B})_{j \in A \cap B}$ are supported on the set $\mathrm{Hom}(H[A \cap B], G)$ and moreover, they are identically distributed. This is because the distributions are projected from the uniform distribution on $\mathrm{Hom}(J, G)$ in the same way, as

there exists an isomorphism between $H[A]$ and $H[B]$ that fixes $A \cap B$. Thus, by Theorem 3.6, there exists $(Y_v)_{v \in V(H)}$ such that

$$\mathbb{H}((Y_v)_{v \in V(H)}) = |\mathcal{F}| \log |\mathrm{Hom}(J, G)| - \sum_{AB \in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i \in A \cap B}), \tag{3.10}$$

where each $(Y_u)_{u \in F}$ is identically distributed with $\mathbf{X}_F$. Since each $\mathbf{X}_F$ takes values in $\mathrm{Hom}(J, G)$, $(Y_u, Y_v)$ is always an ordered edge in $G$. Thus, $(Y_v)_{v \in V(H)}$ is a (not necessarily uniform) random homomorphism from $H$ to $G$. Now (3.10) gives

$$\begin{aligned} \log |\mathrm{Hom}(H, G)| &\geq |\mathcal{F}| \log |\mathrm{Hom}(J, G)| - \sum_{AB \in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i \in A \cap B}) \\ &\geq |\mathcal{F}| \log |\mathrm{Hom}(J, G)| - \sum_{AB \in E(\mathcal{T})} \log |\mathrm{Hom}(H[A \cap B], G)|. \end{aligned}$$

By rescaling by subtracting $|V(H)| \log |V(G)|$ on both sides and using the identity (2.2), we obtain the inequality (1.6). □

### 3.2.2 Applications of Theorem 1.4 and Theorem 3.6

When applying Theorem 1.4, cancelling the terms on the denominator in (1.6) is often crucial. If $J$ is weakly norming, then the following monotonicity holds for every subgraph $K$ of $J$:[1]

$$t_J(G) \geq t_K(G)^{|E(J)|/|E(K)|}. \tag{3.11}$$

This makes the cancellation in (1.6) for a $J$-decomposable graph straightforward.

**Theorem 3.8** ([13])**.** *If $J$ is weakly norming, any $J$-decomposable graph $H$ has Sidorenko's property.*

*Proof.* Let $(\mathcal{F}, \mathcal{T})$ be a $J$-decomposition of $H$ and let $G$ be the target graph in which we wish to embed $H$. By Theorem 1.4, we obtain the inequality

$$t_H(G) \geq \frac{t_J(G)^{|\mathcal{F}|}}{\prod_{XY \in E(\mathcal{T})} t_{H[X \cap Y]}(G)}. \tag{3.12}$$

The monotonicity property (3.11) implies that

$$t_{H[X \cap Y]}(G) \leq t_J(G)^{|E(H[X \cap Y])|/|E(J)|}.$$

Substituting this bound into (3.12), we have the bound $t_H(G) \geq t_J(G)^{|E(H)|/|E(J)|}$. As $J$ has Sidorenko's property, we conclude that $H$ does also. □

[1] For the proof, see the inequality (6.2) and the discussions that follow it.

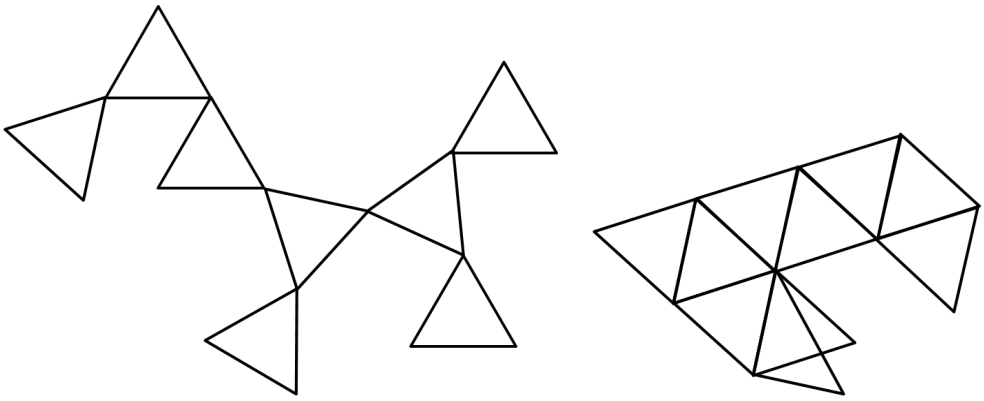


Figure 3.2: A triangle-vertex tree (left) and a triangle-edge tree (right).

Theorem 1.4 is also the key ingredient in proving Theorem 1.8. The proof of Theorem 1.8 however uses more sophisticated arguments relying on the log-convexity of certain graph homomorphism densities, and thus will be postponed until Chapter 5. Instead, we give a short proof of results by Sidorenko [50] on common graphs by using Theorem 3.6. A graph $H$ is *common* if every 2-edge-colouring of a complete graph $K_n$ contains at least $2^{1-|E(H)|}n^{|V(H)|}$ (labelled) monochromatic copies of $H$. A graph $H$ is a *triangle-edge tree* (resp. *triangle-vertex tree*) if it has a $K_3$-decomposition $(\mathcal{F}, \mathcal{T})$ with $|X \cap Y| = 2$ (resp. $|X \cap Y| = 1$) for each $XY \in E(\mathcal{T})$. For examples, see Figure 3.2.

We give an alternative proof of the following theorem of Sidorenko [50]:

**Theorem 3.9** ([50])**.** *Every triangle-edge tree or triangle-vertex tree is common.*

*Proof.* Let $(\mathcal{F}, \mathcal{T})$ be a $K_3$-decomposition of $H$ and assume that each edge of $K_n$ is coloured either red or blue. Suppose first that $H$ is a triangle-vertex tree, i.e., $|A \cap B| = 1$ whenever $AB \in E(\mathcal{T})$. For each $A \in \mathcal{F}$, let $\mathbf{X}_A = (X_{u;A})_{u \in A}$ be a random (labelled) blue triangle. As $\mathbf{X}_A$ and $\mathbf{X}_B$, $AB \in E(\mathcal{T})$, have the same marginals on the single coordinate, Theorem 3.6 implies that there exists $\mathbf{Y} = (Y_v)_{v \in V(H)}$ such that

$$\begin{aligned}\mathbb{H}(\mathbf{Y}) &= \sum_{F \in \mathcal{F}} \mathbb{H}(\mathbf{X}_F) - \sum_{AB \in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i \in A \cap B}) \\ &\geq |\mathcal{F}| \log(\Delta_B n^3) - |E(\mathcal{T})| \log n,\end{aligned}$$

where $\Delta_B n^3$ denotes the number of labelled blue triangles, and moreover, each $(Y_v)_{v \in F}$, $F \in \mathcal{F}$, is identically distributed to $\mathbf{X}_F$. Therefore, there are at least

$\Delta_B^{|\mathcal{F}|} n^{3|\mathcal{F}|-|E(\mathcal{T})|} = \Delta_B^{|\mathcal{F}|} n^{|V(H)|}$ blue copies of $H$. Let $\Delta_R n^3$ be the number of labelled red triangles. Then at least $\Delta_R^{|\mathcal{F}|} n^{|V(H)|}$ red copies of $H$ also exist by the same reason, and thus, the number of monochromatic copies of $H$ divided by $n^{|V(H)|}$ is at least $\Delta_B^{|\mathcal{F}|} + \Delta_R^{|\mathcal{F}|}$. By convexity, we obtain

$$\Delta_B^{|\mathcal{F}|} + \Delta_R^{|\mathcal{F}|} \geq 2\left(\frac{\Delta_B + \Delta_R}{2}\right)^{|\mathcal{F}|}.$$

Since a triangle is common, $\Delta_B + \Delta_R \geq 1/4$ and thus,

$$2\left(\frac{\Delta_B + \Delta_R}{2}\right)^{|\mathcal{F}|} \geq 2^{1-3|\mathcal{F}|}.$$

As $3|\mathcal{F}| = |E(H)|$, $H$ is common.

Suppose now that $H$ is a triangle-edge tree. For each $A \in \mathcal{F}$, let $\mathbf{X}_A = (X_u)_{u\in A}$ be a random (labelled) monochromatic triangle. Observe now that $\mathbf{X}_A$ and $\mathbf{X}_B$, $AB \in E(\mathcal{T})$, have the same marginals on every pair of coordinates, as $\mathbb{P}[(X_{u;A}, X_{v;A}) = (x, y)]$ is proportional to the number of monochromatic triangles containing the edge $xy$. Let $\Delta$ be the normalised number of monochromatic triangles, i.e., there are $\Delta n^3$ monochromatic triangles. Then Theorem 3.6 implies that there exists $\mathbf{Y} = (Y_v)_{v\in V(H)}$ such that

$$\begin{aligned}\mathbb{H}(\mathbf{Y}) &= \sum_{F\in\mathcal{F}} \mathbb{H}(\mathbf{X}_F) - \sum_{AB\in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i\in A\cap B}) \\ &\geq |\mathcal{F}| \log(\Delta n^3) - 2|E(\mathcal{T})| \log n,\end{aligned}$$

where the last inequality follows from the fact that $|A\cap B| = 2$ and Lemma 2.10 (i). Moreover, each $(Y_v)_{v\in F}$, $F \in \mathcal{F}$, is a uniform random monochromatic triangle. Note that $\mathbf{Y}$ induces a random homomorphic copy of $H$. Furthermore, the homomorphic copy of $H$ induced by $\mathbf{Y}$ must be monochromatic, since each pair $\mathbf{Y}_A$ and $\mathbf{Y}_B$, $AB \in E(\mathcal{T})$, of monochromatic triangles is in the same colour as they share an edge. Therefore, at least $\Delta^{|\mathcal{F}|} n^{3|\mathcal{F}|-2|E(\mathcal{T})|} = \Delta^{|\mathcal{F}|} n^{|V(H)|}$ monochromatic labelled copies of $H$ exist. Lemma 2.5 implies that

$$|E(H)| = \sum_{X\in\mathcal{F}} |E(H[X])| - \sum_{XY\in E(\mathcal{T})} |E(H[X\cap Y])| = 3|\mathcal{F}| - |E(\mathcal{T})| = 2|\mathcal{F}| + 1.$$

Since triangles are common, $\Delta \geq 1/4$ and thus,

$$\Delta^{|\mathcal{F}|} \geq 2^{-2|\mathcal{F}|} = 2^{1-|E(H)|}.$$

This implies that $H$ is common. □

Though the natural hypergraph generalisation of Sidorenko's conjecture is false even for tripartite linear 3-uniform hypergraph $H$ [48], it would still be interesting to find more applications of Theorem 1.4 to counting fixed hypergraphs.

## 3.3 Higher tree decompositions

In this section, we will define the notion of $k$-strong tree decompositions that generalises strong tree decompositions in Section 3.1, and thus obtain a generalisation of Theorem 1.3.

The definition of *$k$-strong tree decompositions* is recursively given as follows: for the base case, a 0-strong tree decomposition is an auxiliary tree decomposition $(\mathcal{F}, \mathcal{T})$ of a tree $T$, i.e., $\mathcal{F} = E(T)$ and $\mathcal{T}$ is a spanning tree of the line graph of $T$. Let $H$ be a graph and let $(\mathcal{F}, \mathcal{T})$ be a tree decomposition of $H$. For brevity, we say that $H$ is the *underlying graph* of $(\mathcal{F}, \mathcal{T})$. For positive integer $k$, a $k$-strong tree decomposition is $(\mathcal{F}, \mathcal{T})$ paired with a family $\mathcal{S} = (\mathcal{F}_X, \mathcal{T}_X, \mathcal{S}_X)_{X \in \mathcal{F}}$ of $(k-1)$-strong tree decompositions that satisfy the following:

1. Each $(\mathcal{F}_X, \mathcal{T}_X, \mathcal{S}_X)$ is a $(k-1)$-strong tree decomposition of the induced subgraph $H[X]$.

2. For each $XY \in E(\mathcal{T})$, $H[X \cap Y]$ is an induced forest.

3. For each $XY \in E(\mathcal{T})$, there exists an isomorphism between the two minimum sub-decompositions containing $X \cap Y$ of the $(k-1)$-strong tree decomposition $(\mathcal{F}_X, \mathcal{T}_X, \mathcal{S}_X)$ and $(\mathcal{F}_Y, \mathcal{T}_Y, \mathcal{S}_Y)$ that fixes $X \cap Y$.

To complete this recursive definition, we must define sub-decompositions of $k$-strong tree decompositions and isomorphisms between them. Firstly, a *retraction* of a tree decomposition $(\mathcal{F}, \mathcal{T})$ of a graph $H$ is a tree decomposition $(\mathcal{F}', \mathcal{T}')$ obtained by successively deleting leaves of $\mathcal{T}$ to obtain $\mathcal{T}'$ and letting $\mathcal{F}' = V(\mathcal{T}')$. In particular, $(\mathcal{F}', \mathcal{T}')$ is a tree decomposition of the subgraph of $H$ induced on $\cup_{X \in \mathcal{F}'} X$. A *sub-decomposition* of a 0-strong tree decomposition $(\mathcal{F}, \mathcal{T})$ is a retraction of $(\mathcal{F}, \mathcal{T})$. Thus, the underlying graph of a sub-decomposition of a 0-strong tree decomposition $(\mathcal{F}, \mathcal{T})$ is a subtree of the underlying tree of $(\mathcal{F}, \mathcal{T})$.

Let $(\mathcal{F}, \mathcal{T}, \mathcal{S})$ be a $k$-strong tree decomposition of a graph $H$ and let $U \subseteq V(H)$. Recall that $\mathcal{F}(u)$ denotes the subfamily $\{X \in \mathcal{F} : u \in X\}$ of $\mathcal{F}$. Suppose first that $\cap_{u \in U} \mathcal{F}(u)$ is empty. Then Proposition 2.7 implies that there exists the minimum $\mathcal{F}'$ such that $\cup_{F \in \mathcal{F}'} F$ contains $U$ and $\mathcal{F}'$ induces a subtree in $\mathcal{T}$. The minimum sub-decomposition containing $U$ is defined to be $(\mathcal{F}', \mathcal{T}', \mathcal{S}')$, where $(\mathcal{F}', \mathcal{T}')$ is a retraction of $(\mathcal{F}, \mathcal{T})$. Otherwise, if $\cap_{u \in U} \mathcal{F}(u)$ is non-empty, then every vertex subset $X \in \cap_{u \in U} \mathcal{F}(u)$ contains $U$. Define the minimum sub-decomposition of $(\mathcal{F}, \mathcal{T}, \mathcal{S})$ containing $U$ to be the minimum sub-decomposition of the $(k-1)$-strong tree decomposition $(\mathcal{F}_X, \mathcal{T}_X, \mathcal{S}_X)$ containing $U$, that is, we even retract 'inside' the single $(k-1)$-strong tree decomposition $(\mathcal{F}_X, \mathcal{T}_X, \mathcal{S}_X)$ while every element in $U$ is covered. In this case, $\mathcal{F}'$ may not be unique, but the isomorphism condition between $(k-1)$-strong tree decompositions gives the uniqueness up to isomorphisms, which makes no harm in the definition of $k$-strong tree decompositions. To this end, we must define isomorphism between $(k-1)$-strong tree decompositions.

An *isomorphism* between two $k$-strong tree decompositions $(\mathcal{F}_1, \mathcal{T}_1, \mathcal{S}_1)$ of $H_1$ and $(\mathcal{F}_2, \mathcal{T}_2, \mathcal{S}_2)$ of $H_2$ is a graph isomorphism between the two underlying graphs $H_1$ and $H_2$ that is again an isomorphism between each corresponding pair of $(k-1)$-strong tree decompositions in $\mathcal{S}_1$ and $\mathcal{S}_2$. In particular, an isomorphism between two 0-strong tree decompositions is a graph isomorphism between the underlying trees. To summarise, sub-decompositions and isomorphisms of $k$-strong tree decompositions are recursively defined in terms of the same notions for $(k-1)$-strong tree decompositions.

We say that a graph is $k$-strongly tree-decomposable if it is the underlying graph of a $k$-strong tree decomposition. In particular, a graph is 0-strongly tree-decomposable if and only if it is a tree. Figure 3.3 describes an example of a 2-strong tree decomposition. The tree decomposition $(\mathcal{F}, \mathcal{T})$ of $H$ is represented by grey boxes, in each of which the coloured trees present the 1-strong tree decomposition of the corresponding induced subgraph. For example, the underlying graph of the minimal sub-decomposition containing $u, v$, and $w$ is the subgrid with two squares induced on the leftmost grey box containing them or the isomorphic one next to it, since both 1-strong decompositions in the two grey boxes contains $u$,$v$, and $w$, and there is no sub-decomposition of them containing $u$,$v$, and $w$. On

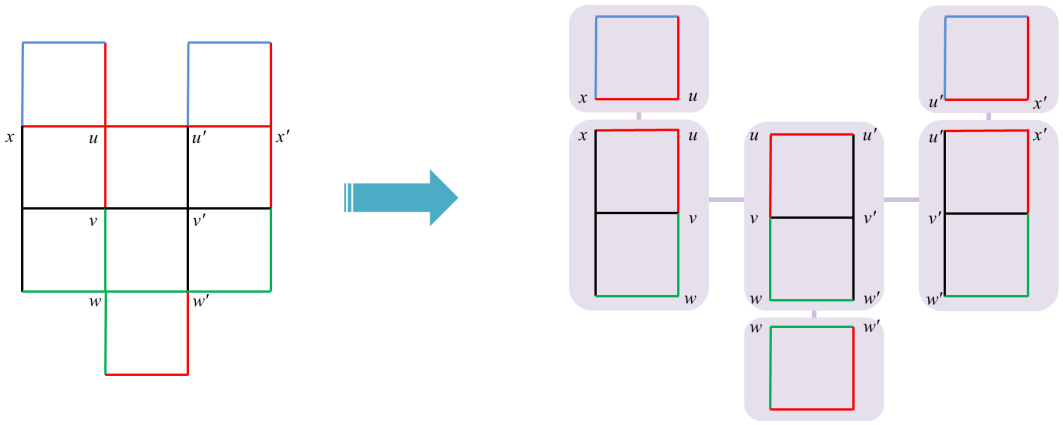



Figure 3.3: An example of a 2-strong tree decomposition.

the other hand, in the same box, the minimal sub-decomposition containing $x$,$u$, and $v$ is the trivial tree decomposition of the red tree induced on the three vertices with $\mathcal{F} = \{\{x, u\}, \{u, v\}\}$ and the single edge tree $\mathcal{T}$.

The main result of this section is:

**Theorem 3.10.** *For each $k \geq 0$, a $k$-strongly tree decomposable graph $H$ has Sidorenko's property.*

This generalises Theorem 1.3, as every strongly tree decomposable graph is 1-strongly tree decomposable. In fact, a 1-strong tree decomposition is a slightly more general concept than a strong tree decomposition: in the definition of strong tree decomposition in Section 3.1.3, $H[X \cap Y]$ must be an independent set whenever $X$ and $Y$ are adjacent vertices of $\mathcal{F}$, whereas we allow $H[X \cap Y]$ to be a forest in the definition of 1-strong tree decompositions. This technical improvement was first observed by Szegedy [54] in a different form: Lemma 3.1 in [54]. However, later we realised that the following lemma due to Fox (see acknowlegements in [28]), predating Szegedy's work, can also play the same role.

We prove a slightly stronger form of the lemma than what is sufficient for this chapter, as it will be necessary in subsequent chapters.

**Lemma 3.11** ([28])**.** *Let $H$ be a bipartite graph. If there exists a positive constant $c$, depending only on $H$, such that*

$$t_H(G) \geq c \cdot t_{K_2}(G)^{|E(H)|} \tag{3.13}$$

*for all graphs $G$ with maximum degree at most $\frac{8|E(G)|}{|V(G)|}$, then $H$ has Sidorenko's property.*

*Proof.* We may assume that $H$ has no isolated vertex, as adding an isolated vertex to a graph does not affect the value of $t_H(G)$ and $|E(H)|$. We may also assume that $G$ has no isolated vertex. To see this, let $G_0$ be the graph obtained from $G$ by deleting all the isolated vertices. Then $|\mathrm{Hom}(H,G)| = |\mathrm{Hom}(H,G_0)|$, as a homomorphism must map vertices of $H$ into non-isolated vertices. Note that

$$t_H(G) = t_H(G_0)\left(\frac{|V(G_0)|}{|V(G)|}\right)^{|V(H)|} \text{ and } t_{K_2}(G)^{|E(H)|} = t_{K_2}(G_0)^{|E(H)|}\left(\frac{|V(G_0)|}{|V(G)|}\right)^{2|E(H)|}.$$

Thus, it follows that

$$\frac{t_H(G)}{t_{K_2}(G)^{|E(H)|}} = \frac{t_H(G_0)}{t_{K_2}(G_0)^{|E(H)|}}\left(\frac{|V(G)|}{|V(G_0)|}\right)^{2|E(H)|-|V(H)|} \geq \frac{t_H(G_0)}{t_{K_2}(G_0)^{|E(H)|}},$$

where the last inequality follows from $2|E(H)| \geq |V(H)|$ and $|V(G)| \geq |V(G_0)|$. Therefore, it is enough to prove (3.13) for $G = G_0$.

Suppose that $H$ is a bipartite graph satisfying (3.13) without isolated vetices. Let $G$ be an arbitrary graph (not necessarily satisfying the maximum degree condition) without isolated vertices and let $D = \lceil \frac{2|E(G)|}{|V(G)|} \rceil$. Since there are no isolated vertices, the average degree $2|E(G)|/|V(G)|$ is at least 1, and hence $D \leq 4|E(G)|/|V(G)|$. Denote by $G'$ the graph obtained from $G$ by the following process. Fix an ordering of the vertices of $G$, and take vertices $v$ one at a time according to the ordering. Replace $v$ with $t = \lceil \frac{\deg(v)}{D} \rceil$ vertices $v_1, \cdots, v_t$ and choose the neighbours of these new vertices so that (i) $N(v_i) \subseteq N(v)$, (ii) $N(v_i) \cap N(v_j) = \emptyset$ for all distinct pairs $i, j$, and (iii) $\deg(v_i) \leq D$ for all $i$. Note that such a choice exists, as one can greedily assign the neighbours of $v$ to the vertices $v_i$ under the given constraints. Further note that during this process, $\deg(v)$ remains the same until $v$ is replaced, and the number of edges always remains the same as $|E(G)|$.

Define a function $\pi : V(G') \to V(G)$ as $\pi(v_i) = v$ for all $i$. Since $\pi$ is a homomorphism from $G'$ to $G$, we obtain a map $\phi : \mathrm{Hom}(H,G') \to \mathrm{Hom}(H,G)$ such that $\varphi(h) := \pi \circ h$. Further note that for an adjacent pair of vertices $v, w \in V(G)$, there exists a unique choice of $v' \in \pi^{-1}(v)$ and $w' \in \pi^{-1}(w)$ such that $v'$ and $w'$ are adjacent in $G'$. Therefore if $\pi \circ h_1 = \pi \circ h_2$ for some $h_1, h_2 \in \mathrm{Hom}(H,G')$,

then for each edge $\{x, y\}$ of $H$, we must have $h_1(x) = h_2(x)$ and $h_1(y) = h_2(y)$. Since $H$ has no isolated vertex, we see that $h_1(x) = h_2(x)$ for all $x \in V(H)$, i.e. $h_1 = h_2$. This implies that our map $\varphi$ from $\mathrm{Hom}(H, G')$ to $\mathrm{Hom}(H, G)$ is an injection. Therefore, $|\mathrm{Hom}(H, G)| \geq |\mathrm{Hom}(H, G')|$.

The graph $G'$ has the same number of edges as the graph $G$, and the number of vertices is at most

$$\begin{aligned} |V(G')| &= \sum_{v \in V(G)} \left\lceil \frac{\deg(v)}{D} \right\rceil \\ &\leq |V(G)| + \sum_{v \in V(G)} \frac{\deg(v)}{D} \\ &\leq |V(G)| + \frac{2|E(G)|}{2|E(G)|/|V(G)|} = 2|V(G)|. \end{aligned}$$

Combining this with the fact $|E(G')| = |E(G)|$, it follows that $G'$ has maximum degree $D \leq \frac{4|E(G)|}{|V(G)|} \leq \frac{8|E(G')|}{|V(G')|}$. Hence, $G'$ satisfies the given maximum degree condition, so

$$|\mathrm{Hom}(H, G)| \geq |\mathrm{Hom}(H, G')| \geq c|V(G')|^{|V(H)|} \left( \frac{2|E(G')|}{|V(G')|^2} \right)^{|E(H)|},$$

where the last inequality is from the assumption (3.13). Since $|V(G')| \leq 2|V(G)|$ and $|V(H)| \leq 2|E(H)|$, it follows that

$$|\mathrm{Hom}(H, G)| \geq c2^{|V(H)|-2|E(H)|}|V(G)|^{|V(H)|} \left( \frac{2|E(G)|}{|V(G)|^2} \right)^{|E(H)|}.$$

The inequality above can be written as $t_H(G) \geq c' t_{K_2}(G)^{|E(H)|}$, where $c' := c2^{|V(H)|-2|E(H)|}$. As $c'$ is a constant only depending on $H$, Lemma 2.8 concludes the proof. □

*Proof of Theorem 3.10.* The goal is to prove that, for each $k$-strongly tree decomposable graph $H$, there exists a distribution $Y = (Y_1, Y_2, \cdots, Y_{|V(H)|})$ on $\mathrm{Hom}(H, G)$ such that

$$\mathbb{H}(Y) \geq |E(H)| \log \left( \frac{2|E(G)|}{|V(G)|^2} \right) + |V(H)| \log |V(G)| + c_H, \tag{3.14}$$

where $c_H$ is a constant only depending on $H$. If so, then Lemma 2.10 implies that $t_H(G) \geq c \cdot t_{K_2}(G)^{|E(H)|}$. Thus, by Lemma 3.11, it suffices to prove that (3.14) holds for $G$ with maximum degree at most $8|E(G)|/|V(G)|$. We construct

a distribution on $\mathrm{Hom}(H, G)$, which we call $k$-good, satisfying (3.14) by induction on $k$. The base case $k = 0$ is when $H$ is a tree and hence proved by letting $\mathbf{Y} = \mathbf{w}_T$ in Proposition 3.2. Hence, the distribution given by the branching random walk, or explicitly by (3.1), is 0-good. Furthermore, we may further assume that if $(\mathcal{F}_1, \mathcal{T}_1, \mathcal{S}_1)$ and $(\mathcal{F}_2, \mathcal{T}_2, \mathcal{S}_2)$ are two isomorphic $(k-1)$-strong decompositions by an isomorphism $\phi$ with underlying graphs $H_1$ and $H_2$, respectively, then $\phi$ preserves the $(k-1)$-good distributions on $\mathrm{Hom}(H_1, G)$ and $\mathrm{Hom}(H_2, G)$. That is, for each $h \in \mathrm{Hom}(H_1, G)$,

$$\mathbb{P}[\mathbf{x}_1 = h] = \mathbb{P}[\mathbf{x}_2 = h \circ \phi],$$

where $\mathbf{x}_i$ is a random homomorphism generated by the $(k-1)$-good distributions on $\mathrm{Hom}(H_i, G)$, $i = 1, 2$. This is again true for the base case $k = 0$, as seen in (3.1).

Let $(\mathcal{F}, \mathcal{T}, \mathcal{S})$ be a $k$-strong tree decomposition of $H$. By definition $(\mathcal{F}, \mathcal{T})$ is a Markov tree on $V(H)$. Moreover, if $(\mathcal{F}', \mathcal{T}', \mathcal{S}')$ is a sub-decomposition of $(\mathcal{F}, \mathcal{T}, \mathcal{S})$, then $(\mathcal{F}', \mathcal{T}')$ is a Markov subtree of $(\mathcal{F}, \mathcal{T})$. For each $A \in \mathcal{F}$, let $(X_{i;A})_{i \in A}$ be a random homomorphism in $\mathrm{Hom}(H[A], G)$ with the $(k-1)$-good distribution. To apply Theorem 3.6, we claim that $(X_{i;A})_{i \in A \cap B}$ and $(X_{j;B})_{j \in A \cap B}$ are identically distributed whenever $AB \in E(\mathcal{T})$. By inductive hypothesis, we may further assume that the $(k-1)$-good distribution is also constructed by using Theorem 3.6.

Let $(\mathcal{F}_1, \mathcal{T}_1, \mathcal{S}_1)$ and $(\mathcal{F}_2, \mathcal{T}_2, \mathcal{S}_2)$ be the two minimum sub-decompositions of $(\mathcal{F}_A, \mathcal{T}_A, \mathcal{S}_A)$ and $(\mathcal{F}_B, \mathcal{T}_B, \mathcal{S}_B)$ containing $A \cap B$. Then by definition $(\mathcal{F}_1, \mathcal{T}_1, \mathcal{S}_1)$ and $(\mathcal{F}_2, \mathcal{T}_2, \mathcal{S}_2)$ are isomorphic by an isomorphism $\phi$ that fixes $A \cap B$. Denote by $H_i$ the underlying graph of $(\mathcal{F}_i, \mathcal{T}_i)$, $i = 1, 2$. Then $\phi$ preserves the $(k-1)$-good distributions on $\mathrm{Hom}(H_1, G)$ and $\mathrm{Hom}(H_2, G)$. Thus, since $(\mathcal{F}_1, \mathcal{T}_1)$ and $(\mathcal{F}_2, \mathcal{T}_2)$ are Markov subtrees of $(\mathcal{F}_A, \mathcal{T}_A)$ and $(\mathcal{F}_B, \mathcal{T}_B)$, respectively, Remark 3.7 implies that the two $(k-1)$-good distributions on $\mathrm{Hom}(H[A], G)$ and $\mathrm{Hom}(H[B], G)$ have the same marginals on $\mathrm{Hom}(H[A \cap B], G)$. This proves the claim and hence, we are now ready to use Theorem 3.6.

Denote by $e_A$ the number of edges induced on $A \subseteq V(H)$ for brevity. Theorem 3.6 gives that

$$\begin{aligned}\mathbb{H}(Y_1, \cdots, Y_k) &= \sum_{F \in \mathcal{F}} \mathbb{H}((X_{i;F})_{i \in F}) - \sum_{AB \in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i \in A \cap B}) \\ &\geq \log\left(\frac{2|E(G)|}{|V(G)|^2}\right) \sum_{F \in \mathcal{F}} e_F + \log|V(G)| \sum_{F \in \mathcal{F}} |F| \\ &\quad - \sum_{AB \in E(\mathcal{T})} \mathbb{H}((X_{i;A})_{i \in A \cap B}). \end{aligned} \tag{3.15}$$

Since $H[A \cap B]$ is always a forest, we obtain

$$\begin{aligned}|\mathrm{Hom}(H[A \cap B], G)| &\leq |V(G)|^{|A \cap B|} \left(\frac{\Delta(G)}{|V(G)|}\right)^{e_{A \cap B}} \\ &\leq |V(G)|^{|A \cap B|} \left(\frac{8|E(G)|}{|V(G)|^2}\right)^{e_{A \cap B}} \\ &= 2^{2e_{A \cap B}} |V(G)|^{|A \cap B|} \left(\frac{2|E(G)|}{|V(G)|^2}\right)^{e_{A \cap B}}.\end{aligned}$$

Thus, we obtain the upper bound

$$\mathbb{H}((X_{i;A})_{i \in A \cap B}) \leq |A \cap B| \log|V(G)| + e_{A \cap B} \log\left(\frac{2|E(G)|}{|V(G)|^2}\right) + 2e_{A \cap B}$$

for negative terms in (3.15). Substituting this into (3.15) yields the desired bound (3.14). Finally, the additional inductive hypothesis that an isomorphism between $k$-strong tree decompositions preserves $k$-good distributions is maintained, since we applied Theorem 3.6 with the same marginal constraints $(X_{i;F})_{F \in \mathcal{F}}$. □

## 3.4 Concluding remarks

**A stability result from branching random walks.** In the proof of Proposition 3.2, we only used two inequalities that follow from Lemma 2.10(i):

$$\mathbb{H}(\mathbf{w}_T) \leq \log|\mathrm{Hom}(T, G)| \quad \text{and} \quad \mathbb{H}(\mathbf{w}_T(v)) \leq \log|V(G)|.$$

In each inequality above, equality holds if and only if the distribution is uniform on the support, i.e., $\mathbf{w}_T$ is a uniform random homomorphism in $\mathrm{Hom}(T, G)$ and $\mathbf{w}_T(v)$ is a uniform random vertex in $G$. Thus, $t_T(G)$ attains the lower bound $t_{K_2}(G)^{|E(T)|}$ given by Proposition 3.2 if and only if every vertex in $G$ has the same degree. A slightly more careful analysis also gives a simpler proof of the

following theorem by Dellamonica, Haxell, Łuczak, Mubayi, Nagle, Person, Rödl, and Schacht [14] for dense graphs.

**Theorem 3.12** (Theorem 1 in [14], dense case)**.** *For all $\varepsilon, c \in (0,1]$ and integers $t \geq 3$, there exists $\delta > 0$ so that the following statement holds. Let $T$ be a $t$-vertex tree and let $G$ be an $n$-vertex graph such that $t_{K_2}(G) = p$ and $p > c$. If $t_T(G) \leq (1+\delta)p^{|E(T)|}$, then $\sum_{v \in V(G)} |\deg(v) - pn| \leq \varepsilon pn^2$.*

**Caveats from marginal constraints.** When using Theorem 3.6, the major caveat is how to find random variables $(X_{i;F})_{i \in F}$ that agree on the marginals $(X_{j;A})_{j \in A \cap B}$. The symmetry condition in the definition of the $J$-decomposition is in fact tailored to satisfy the marginal constraints. However, for non-isomorphic graphs $H_1$ and $H_2$ containing the same induced subgraph $J$, it is often hard to find distributions on $\mathrm{Hom}(H_1, G)$ and $\mathrm{Hom}(H_2, G)$ that agree on the natural projection to $\mathrm{Hom}(J, G)$.

For example, it is possible to generate a random copy of a tree in such a way that the projection of the distribution onto a subtree agrees with the distribution generated by the same algorithm, which leads to the definition of strongly tree-decomposable graphs. Another example is the theorem of Li and Szegedy [33] proving that, if $H$ has Sidorenko's property, we may assume that the projection of a uniform random homomorphic copy of $H$ onto a single edge is again uniform. Therefore, it is possible to glue graphs having Sidorenko's property on a single edge while preserving the property. Motivated by these instances, we ask the following:

**Question 3.13.** *Let $H$, $J$, and $K$ be graphs. Suppose that there exist induced subgraphs of $H$ and $J$ isomorphic to $K$. If $t_H(G) \geq t_K(G)^{|E(H)|/|E(K)|}$ and $t_J(G) \geq t_K(G)^{|E(J)|/|E(K)|}$ hold for any graph $G$, then do there exist distributions on $\mathrm{Hom}(H, G)$ and $\mathrm{Hom}(J, G)$ agree on the natural projection on $\mathrm{Hom}(K, G)$?*

**Homomorphism domination exponent.** Though graph homomorphism inequalities have appeared in many contexts for decades, Kopparty and Rossman [30] recently initiated the study of a general type of domination inequality between the number of graph homomorphisms. A normalised version of their main

questions in [30] is to determine an optimal exponent $c = c(H, J)$ for graphs $H$ and $J$ such that

$$t_H(G) \geq t_J(G)^c \tag{3.16}$$

holds for all graphs $G$. For example, weakly norming graphs give a wide range of such inequalities, and Lemma 5.16 gives the optimal domination exponent between an even path $H$ and its subgraph $J$.

In [30], Kopparty and Rossman obtained optimal exponents for some pairs of graphs for example some paths and 2-trees.[2] This type of question, or more generally, determining exponents $c_1, c_2, \cdots, c_k$ for $H$ and its subgraphs $J_1, J_2, \cdots, J_k$ such that

$$t_H(G) \geq t_{J_1}(G)^{c_1} t_{J_2}(G)^{c_2} \cdots t_{J_k}(G)^{c_k}, \tag{3.17}$$

is useful in deducing inequalities from Theorem 1.4, as done in Theorem 3.8 or Theorem 3.9. Hence, it will be interesting to see more inequalities of a similar type.

[2] In fact, many of their results are technically different from (3.16) because of the normalisation issue.

# Chapter 4

# Cartesian products and Sidorenko's conjecture

The main purpose of this chapter is to study how Cartesian products relate to Sidorenko's conjecture. More specifically, we will prove Theorem 1.5.

**Theorem 1.5** (Restated)**.** *If $J$ is either a tree or an even cycle and $H$ is a bipartite graph with Sidorenko's property, then $J\square H$ also has Sidorenko's property.*

Recall that the Cartesian product $H_1\square H_2$ of two graphs $H_1$ and $H_2$ is the graph on $V(H_1) \times V(H_2)$ where $(x_1, x_2)$ and $(y_1, y_2)$ are adjacent if and only if either (i) $x_1$ and $y_1$ are adjacent in $H_1$ and $x_2 = y_2$, or (ii) $x_2$ and $y_2$ are adjacent in $H_2$ and $x_1 = y_1$.

The following alternative description of the graph $J\square H$ provides insight into the proof of Theorem 1.5. Let $J_1, \cdots, J_{|V(H)|}$ be vertex-disjoint copies of the graph $J$. For each edge $\{a, b\}$ of $H$, place an edge between the copy of each $v \in V(T)$ in $J_a$ and $J_b$ so that $J_a$ and $J_b$ together form a copy of $J\square K_2$. It is not too difficult to check that the resulting graph is $J\square H$.

We wish to count the number of homomorphisms from $J\square H$ to a given graph $G$, through counting the number of homomorphisms from $H$ to an auxiliary graph constructed from $G$. For each vertex $v$ of $H$, there exists a copy $J_v$ of $J$ in $J\square H$ on the vertices $V(J) \times \{v\}$. Moreover, as seen above, for each edge $e = \{v, w\}$ of $H$, the two copies of $J_v$ and $J_w$ form a copy of $J\square K_2$ in $J\square H$ (see Figure 4.1). Thus, a copy of $J\square K_2$ in $G$ can be 'contracted' into an edge in the desired auxiliary graph of $G$. This motivates the following definition of the operation $\psi_J$ on $G$.

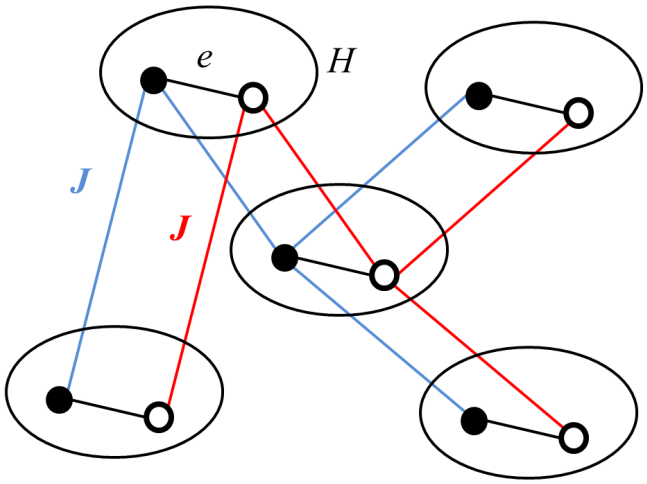



Figure 4.1: $J\square K_2$ in $J\square H$.

**Definition 4.1.** For given graphs $J$ and $G$, let $\psi_J(G)$ be the graph with vertex set $\mathrm{Hom}(J, G)$ such that two vertices $h_1, h_2 \in \mathrm{Hom}(J, G)$ are adjacent if and only if $h_1(v)$ and $h_2(v)$ are adjacent in $G$ for all $v \in V(J)$.

The observation above is essentially equivalent to saying that a copy of $J\square H$ in $G$ can be mapped to a copy of $H$ in $\psi_J(G)$, and the following lemma formalises this intuition.

**Lemma 4.2.** *For all graphs $J$, $H$ and $G$, there exists a one-to-one correspondence between* $\mathrm{Hom}(J\square H, G)$ *and* $\mathrm{Hom}(H, \psi_J(G))$. *In particular,*

$$|\mathrm{Hom}(J\square H, G)| = |\mathrm{Hom}(H, \psi_J(G))|.$$

*Proof.* We will define a map $\xi : \mathrm{Hom}(J\square H, G) \to \mathrm{Hom}(H, \psi_J(G))$ and a map $\varphi : \mathrm{Hom}(H, \psi_J(G)) \to \mathrm{Hom}(J\square H, G)$ such that $\varphi \circ \xi = id$. Then we will prove that $\varphi$ and $\xi$ are injective, which implies that $\xi = \varphi^{-1}$ is the desired one-to-one correspondence.

For each $h \in \mathrm{Hom}(J\square H, G)$ and each $v \in V(H)$, define $h_v : V(J) \to V(G)$ as $h_v(w) = h(w, v)$.Whenever $w, w' \in V(J)$ are adjacent, the vertices $h_v(w) = h(w, v)$ and $h_v(w') = h(w', v)$ are adjacent. Thus, $h_v \in \mathrm{Hom}(J, G)$ for all $v \in V(H)$. Moreover, if $v, v'$ are adjacent vertices of $H$, then $h_v(w) = h(w, v)$ and $h_{v'}(w) = h(w, v')$ are adjacent, and thus $h_v$ and $h_{v'}$ are adjacent in $\psi_J(G)$. Hence if we let $\xi(h) : V(H) \to \mathrm{Hom}(J, G)$ be defined by $\xi(h)(v) = h_v$, then $\xi$ is a map from $\mathrm{Hom}(J\square H, G)$ to $\mathrm{Hom}(H, \psi_J(G))$.

On the other hand, given a map $g \in \mathrm{Hom}(H, \psi_J(G))$, define $\varphi(g) : V(J) \times V(H) \to V(G)$ as $\varphi(g)(w, v) = g(v)(w)$ for each $v \in V(H)$ and $w \in V(J)$. We first prove that $\varphi(g) \in \mathrm{Hom}(J \Box H, G)$. For edges of the form $\{(w, v), (w', v)\}$, $\varphi(g)(w, v) = g(v)(w)$ and $\varphi(g)(w', v) = g(v)(w')$ are adjacent since $g(v) \in V(\psi_J(G)) = \mathrm{Hom}(J, G)$. For edges of the form $\{(w, v), (w, v')\}$, we have $\varphi(g)(w, v) = g(v)(w)$ and $\varphi(g)(w, v') = g(v')(w)$, and these two vertices are adjacent in $G$ since $g(v)$ and $g(v')$ are adjacent in $\psi_J(G)$, i.e., $\varphi(g) \in \mathrm{Hom}(J \Box H, G)$.

Note that $\varphi \circ \xi = id$ follows from the fact that, for $h \in \mathrm{Hom}(J \Box H, G)$, $v \in V(H)$ and $w \in V(J)$,

$$((\varphi \circ \xi)(h))\,(w, v) = ((\xi(h)(v))(w) = h_v(w) = h(w, v)$$

for the map $h_v$ defined as above.

Suppose $g$ and $g'$ are distinct elements in $\mathrm{Hom}(H, \psi_J(G))$. Then $g(v) \neq g'(v)$ for some $v \in V(H)$, and thus, $\phi(g)(w, v) = g(v)(w) \neq g'(v)(w) = \phi(g')(w, v)$ for some $w \in V(J)$. Hence, $\varphi$ is injective. Note that $\xi$ is also injective, since otherwise $\varphi \circ \xi = id$ is not injective. Therefore, $\xi = \varphi^{-1}$ is the desired one-to-one correspondence. □

## 4.1 Cartesian products with trees

We will prove the following half of Theorem 1.5, i.e., when $J$ is a tree.

**Theorem 4.3** ([28])**.** *If $T$ is a tree and $H$ is a bipartite graph having Sidorenko's property, then $T \Box H$ also has Sidorenko's property.*

By Lemma 4.2, we can now estimate the size of $\mathrm{Hom}(T \Box H, G)$ through estimating the size of $\mathrm{Hom}(H, \psi_T(G))$, where Sidorenko's property for $H$ provides a lower bound on the size of $\mathrm{Hom}(H, \psi_T(G))$. We can use this idea to show the simplest case of Theorem 4.3, i.e., when $T = K_2$. Here we give a full proof of this simple case, as the result will be used in the proof of Theorem 4.3.

**Theorem 4.4.** *If $H$ is a bipartite graph having Sidorenko's property, then $K_2 \Box H$ has Sidorenko's property.*

*Proof.* Let $G$ be a given graph and put $\psi(G) = \psi_{K_2}(G)$ for simplicity. By Lemma 4.2 and the fact that $H$ has Sidorenko's property, we have

$$\begin{aligned}|\mathrm{Hom}(K_2\Box H, G)| &= |\mathrm{Hom}(H, \psi(G))| \\ &\geq |V(\psi(G))|^{|V(H)|} \left(\frac{|\mathrm{Hom}(K_2, \psi(G))|}{|V(\psi(G))|^2}\right)^{|E(H)|} \\ &= |V(\psi(G))|^{|V(H)|-2|E(H)|} |\mathrm{Hom}(K_2, \psi(G))|^{|E(H)|}. \end{aligned} \tag{4.1}$$

We have that

$$|V(\psi(G))| = |\mathrm{Hom}(K_2, G)| = |V(G)|^2 t_{K_2}(G).$$

On the other hand, by Lemma 4.2 with $H = K_2$, we have

$$|\mathrm{Hom}(K_2, \psi(G))| = |\mathrm{Hom}(K_2\Box K_2, G)|,$$

where since $K_2\Box K_2$ is isomorphic to $C_4$, Sidorenko's property of $C_4$ implies that

$$|\mathrm{Hom}(K_2\Box K_2, G)| = |V(G)|^4 t_{C_4}(G) \geq |V(G)|^4 (t_{K_2}(G))^4.$$

Therefore, in (4.1), we get

$$\begin{aligned}|\mathrm{Hom}(K_2\Box H, G)| &\geq \left(|V(G)|^2 t_{K_2}(G)\right)^{|V(H)|-2|E(H)|} \cdot \left(|V(G)| t_{K_2}(G)\right)^{4|E(H)|} \\ &= |V(G)|^{2|V(H)|} t_{K_2}(G)^{|V(H)|+2|E(H)|}.\end{aligned}$$

Since $|V(K_2\Box H)| = 2|V(H)|$ and $|E(K_2\Box H)| = 2|E(H)| + |V(H)|$, we deduce that $K_2\Box H$ has Sidorenko's property. □

If one attempts to use the same idea as in the proof of Theorem 4.4 to prove Theorem 4.3 for general graphs $T$ other than $K_2$, then the inequality corresponding to (4.1) will be

$$|\mathrm{Hom}(T\Box H, G)| \geq |V(\psi_T(G))|^{|V(H)|-2|E(H)|} |\mathrm{Hom}(K_2, \psi_T(G))|^{|E(H)|}.$$

Thus, we need estimates on $|V(\psi_T(G))| = |\mathrm{Hom}(T, G)|$ and $|\mathrm{Hom}(K_2, \psi_T(G))| = |\mathrm{Hom}(K_2\Box T, G)|$. If $T$ has Sidorenko's property, then $K_2\Box T$ also has Sidorenko's property by Theorem 4.4. Hence in this case we have lower bound estimates on both $|V(\psi_T(G))|$ and $|\mathrm{Hom}(K_2, \psi_T(G))|$. Unfortunately, these bounds do not transfer to a lower bound on $|\mathrm{Hom}(T\Box H, G)|$, since such a lower bound requires an upper bound on $|V(\psi_T(G))|$ if $|V(H)| - 2|E(H)| \leq 0$. We will solve this problem when $T$ is a tree through Lemma 3.11.

*Proof of Theorem 4.3.* We may assume that $H$ has no isolated vertex, as adding an isolated vertex to a graph does not affect the value of $t_H(G)$ and $|E(H)|$.

Let $T$ be a tree with $\tau$ vertices, and let $G$ be a given graph. By Lemma 3.11, we may assume that $G$ has maximum degree at most $\frac{8|E(G)|}{|V(G)|} = 4|V(G)|t_{K_2}(G)$. By Lemma 4.2 and the fact that $H$ has Sidorenko's property, we have

$$\begin{aligned}|\mathrm{Hom}(T\square H, G)| &= |\mathrm{Hom}(H, \psi_T(G))| \\ &\geq |V(\psi_T(G))|^{|V(H)|} \left(\frac{|\mathrm{Hom}(K_2, \psi_T(G))|}{|V(\psi_T(G))|^2}\right)^{|E(H)|} \\ &= |V(\psi_T(G))|^{|V(H)|-2|E(H)|} |\mathrm{Hom}(K_2, \psi_T(G))|^{|E(H)|}. \end{aligned} \tag{4.2}$$

Recall that $V(\psi_T(G)) = \mathrm{Hom}(T, G)$. We can construct an element in $\mathrm{Hom}(T, G)$ by starting from an arbitrary vertex of $T$, defining its image in $V(G)$, and then extending the homomorphism one vertex at a time. By the condition on the maximum degree of $G$, we thus have

$$\begin{aligned}|V(\psi_T(G))| = |\mathrm{Hom}(T, G)| &\leq |V(G)|\Big(4|V(G)|t_{K_2}(G)\Big)^{\tau-1} \\ &= 2^{2\tau-2}|V(G)|^{\tau} t_{K_2}(G)^{\tau-1}. \end{aligned} \tag{4.3}$$

On the other hand, by Lemma 4.2 with $H = K_2$ and Theorem 4.4, we have

$$|\mathrm{Hom}(K_2, \psi_T(G))| = |\mathrm{Hom}(T\square K_2, G)| \geq |V(G)|^{2\tau} t_{K_2}(G)^{3\tau-2}. \tag{4.4}$$

Since $H$ has no isolated vertex, we have $|V(H)| \leq 2|E(H)|$, and thus in (4.2), we may use the bounds from (4.3) and (4.4) to obtain

$$\begin{aligned}&|\mathrm{Hom}(T\square H, G)| \\ &\geq \Big(2^{2\tau-2}|V(G)|^{\tau} t_{K_2}(G)^{\tau-1}\Big)^{|V(H)|-2|E(H)|} \Big(|V(G)|^{2\tau} t_{K_2}(G)^{3\tau-2}\Big)^{|E(H)|} \\ &= 2^{2(\tau-1)(|V(H)|-2|E(H)|)} \cdot |V(G)|^{\tau|V(H)|} t_{K_2}(G)^{(\tau-1)|V(H)|+\tau|E(H)|}. \end{aligned}$$

Since $|V(T\square H)| = \tau|V(H)|$ and $|E(T\square H)| = (\tau-1)|V(H)| + \tau|E(H)|$, by Lemma 3.11, we deduce that $T\square H$ has Sidorenko's property. □

Since an arbitrary $d$-dimensional grid can be obtained from the Cartesian product of $d$ paths, we obtain the following corollary.

**Corollary 4.5.** *For all $d \geq 1$, all $d$-dimensional grids have Sidorenko's property.*

## 4.2 Cartesian products with even cycles

In this section, we prove Theorem 1.5 when $J$ is an even cycle and hence complete the proof of Theorem 1.5.

**Theorem 4.6** ([12]). *If $J$ is an even cycle and $H$ is a bipartite graph having Sidorenko's property, then $J\square H$ also has Sidorenko's property.*

*Proof.* Let $k \geq 2$ be a given positive integer. It suffices to prove the theorem when $H$ has no isolated vertex, since deleting isolated vertices in $H$ only removes connected components of $J\square H$ each of which is isomorphic to $J$.

Let $G$ be a graph with $n$ vertices and let

$$p := \frac{2|E(G)|}{n^2} = \frac{|\mathrm{Hom}(K_2, G)|}{n^2}, \quad \alpha := \frac{|\mathrm{Hom}(C_4, G)|}{p^4 n^4}, \quad \beta_k := \frac{|\mathrm{Hom}(C_{2k}, G)|}{p^{2k} n^{2k}},$$

for $k \geq 2$. By Lemma 3.11, we may assume that $G$ has maximum degree at most $4pn$. Since $C_4$ and $C_{2k}$ have Sidorenko's property, $\alpha$ and $\beta_k$ are both at least 1. Let $\lambda := \max_i |\lambda_i|$. Then by Proposition 2.1 (ii), for $k \geq 2$,

$$\beta_k p^{2k} n^{2k} = \sum_{i=1}^{n} \lambda_i^{2k} \leq \lambda^{2k-4} \sum_{i=1}^{n} \lambda_i^4 = \alpha \lambda^{2k-4} p^4 n^4.$$

Since $\lambda$ is at most the maximum degree of $G$, we know that $\lambda \leq 4pn$. Thus, we obtain

$$\beta_k \leq 2^{4k-8} \alpha. \tag{4.5}$$

Writing $v = |V(H)|$ and $e = |E(H)|$, we apply Lemma 4.2 and then use the fact that $H$ has Sidorenko's property to obtain

$$|\mathrm{Hom}(H\square C_{2k}, G)| = |\mathrm{Hom}(H, \psi_{C_{2k}}(G))| \geq |V(\psi_{C_{2k}}(G))|^v \left( \frac{2|E(\psi_{C_{2k}}(G))|}{|V(\psi_{C_{2k}}(G))|^2} \right)^e.$$

As $|V(\psi_{C_{2k}}(G))| = |\mathrm{Hom}(C_{2k}, G)| = \beta_k p^{2k} n^{2k}$, it follows that

$$|\mathrm{Hom}(H\square C_{2k}, G)| \geq (\beta_k p^{2k} n^{2k})^v \left( \frac{2|E(\psi_{C_{2k}}(G))|}{(\beta_k p^{2k} n^{2k})^2} \right)^e. \tag{4.6}$$

Lemma 4.2 also yields

$$2|E(\psi_{C_{2k}}(G))| = |\mathrm{Hom}(K_2, \psi_{C_{2k}}(G))| = |\mathrm{Hom}(K_2 \square C_{2k}, G)| = |\mathrm{Hom}(C_{2k}, \psi_{K_2}(G))|.$$

Since $C_{2k}$ has Sidorenko's property, we obtain

$$\begin{aligned} 2|E(\psi_{C_{2k}}(G))| &\geq |V(\psi_{K_2}(G))|^{2k} \left(\frac{2|E(\psi_{K_2}(G))|}{|V(\psi_{K_2}(G))|^2}\right)^{2k} \\ &= \Big(\frac{2|E(\psi_{K_2}(G))|}{|\mathrm{Hom}(K_2,G)|}\Big)^{2k}. \end{aligned}$$

Applying Lemma 4.2 once more, we have

$$2|E(\psi_{K_2}(G))| = |\mathrm{Hom}(K_2, \psi_{K_2}(G))| = |\mathrm{Hom}(K_2 \Box K_2, G)| = |\mathrm{Hom}(C_4, G)| = \alpha p^4 n^4,$$

which, together with $|\mathrm{Hom}(K_2, G)| = 2|E(G)| = pn^2$, yields

$$2|E(\psi_{C_{2k}}(G))| \geq \Big(\frac{\alpha p^4 n^4}{pn^2}\Big)^{2k} = \alpha^{2k} p^{6k} n^{4k}.$$

Together with (4.6), this inequality implies that

$$|\mathrm{Hom}(H \Box C_{2k}, G)| \geq (\beta_k p^{2k} n^{2k})^v \Big(\frac{\alpha^{2k} p^{2k}}{\beta_k^2}\Big)^e = \alpha^{2ek} \beta_k^{v-2e} p^{2kv+2ke} n^{2kv}.$$

Since $H$ has no isolated vertex, we have $v - 2e \leq 0$ and hence, by (4.5), $\beta_k^{v-2e} \geq 2^{(4k-8)(v-2e)} \alpha^{v-2e}$. Therefore, since $\alpha \geq 1$,

$$|\mathrm{Hom}(H \Box C_{2k}, G)| \geq 2^{(4k-8)(v-2e)} \alpha^{2ek+(v-2e)} n^{2kv} p^{2kv+2ke} \geq 2^{(4k-8)(v-2e)} n^{2kv} p^{2kv+2ke}.$$

Since $H \Box C_{2k}$ has $2kv$ vertices and $2kv + 2ke$ edges, the result now follows from Lemma 3.11. □

As a corollary, we obtain a new class of graphs with Sidorenko's property.

**Corollary 4.7.** *For all $d \geq 1$ and positive integers $k_1, k_2, \cdots, k_d$, the torus-shaped graph $C_{2k_1} \Box C_{2k_2} \Box \cdots \Box C_{2k_d}$ has Sidorenko's property.*

## 4.3 Concluding remarks

**An alternative proof of Theorem 4.6**. After submitting the paper [12] that contains the above proof of Theorem 4.6, we found a simpler proof. Roughly speaking, the proof of Theorem 4.6 relies on the fact that the normalised 4-cycle count $t_{C_4}(G)$ is 'large' whenever $t_{C_{2k}}(G)$, $k \geq 3$, is 'large'. More precisely, one may show that, with the notation used in the proof of Theorem 4.6,

$$\beta_k \leq \alpha^{k/2}, \tag{4.7}$$

which can be an alternative of the key observation (4.5). Without invoking Lemma 3.11, (4.7) can be shown directly by the inequality

$$t_{C_{2k}}(G)^{1/2k} = \beta_k^{1/2k} p \leq \alpha^{1/4} p = t_{C_4}(G)^{1/4}. \tag{4.8}$$

The proof of the above inequality is again spectral. Let $\overline{\lambda}_i = \lambda_i/n$ be the normalised eigenvalues of the $n$-vertex graph $G$. By Proposition 2.1 (ii),

$$t_{C_{2k}}(G) = \sum_{i=1}^{n} \overline{\lambda}_i^{2k}.$$

Since the $\ell_4$-norm on $\mathbb{R}^n$ dominates the $\ell_{2k}$-norm[1] we obtain (4.8). Once we have (4.7), the rest of the proof remains the same if we replace $2^{2k-4}\alpha$ by $\alpha^{k/2}$. This approach however does not generalise to arbitrary bipartite graphs. For instance, the complete bipartite graph $K_{2,t}$, $t \geq 2$, satisfies the homomorphism inequality

$$t_{K_{2,t}}(G)^{1/2t} \geq t_{C_4}(G)^{1/4}.$$

This is essentially the reverse of (4.8) and thus, we cannot expect the same strategy to work. Hence, we propose the following open question:

**Question 4.8.** *Is Sidorenko's property preserved under taking Cartesian products? That is, does $J\Box H$ have Sidorenko's property whenever both $J$ and $H$ do? In particular, prove (or disprove) that $K_{2,3}\Box H$ has Sidorenko's property whenever $H$ does.*

**Extension of Cartesian product to non-bipartite graphs**. For a given (not necessarily bipartite) graph $H$, define a bipartite graph $\phi(H)$ as follows: The bipartition of $\phi(H)$ consists of two disjoint copies of $V(H)$. Two vertices in distinct parts are adjacent in $\phi(H)$ if they are copies of the same vertex in $H$, or two adjacent vertices in $H$. In particular, $\phi(H)$ has $2|E(H)| + |V(H)|$ edges.

It is not too difficult to see that for bipartite graphs $H$, we have $\phi(H) = K_2\Box H$. Hence, $\phi(H)$ has more restricted meaning than Cartesian products $J\Box H$ for general graphs $J$ when considering bipartite graphs $H$. However, the operation $\phi$ has the advantage of being applicable to non-bipartite graphs. For example, since

[1] In general, if $p < q$ then $(\sum_{i=1}^n |x_i|^q)^{1/q} \leq (\sum_{i=1}^n |x_i|^p)^{1/p}$. This follows from the homogeneity of the inequality and the simple fact $|x_i|^q \leq |x_i|^p$ whenever $|x_i| \leq 1$.

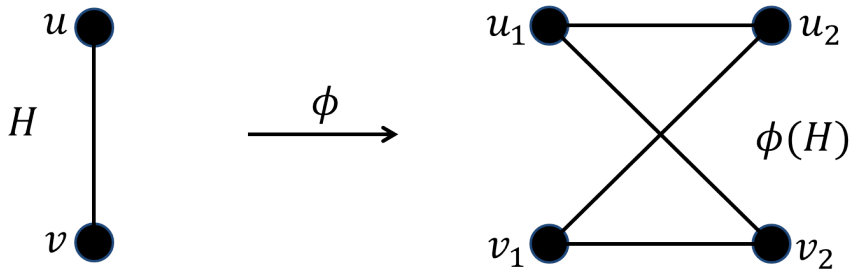



Figure 4.2: Blow-up via $\phi$.

$\phi(K_k) = K_{k,k}$, we know that $\phi(K_k)$ has Sidorenko's property for all $k \geq 2$. Thus, $\phi(H)$ may have Sidorenko's property even if $H$ is a non-bipartite graph. Also note that $\phi(C_5)$ is $K_{5,5} \setminus C_{10}$ which is the minimal bipartite graph unknown to satisfy Sidorenko's conjecture. The operation $\phi$ provides many interesting graphs for which Sidorenko's conjecture is not known to be true.

We conclude this chapter with some open problems regarding the operator $\phi$. We believe that the family $\{\phi(C_{2k+1})\}_{k\geq 1}$ could be an interesting starting point in further studying Sidorenko's conjecture. The only graph in this family known to have Sidorenko's property is $\phi(C_3)$.

**Question 4.9.** *Does there exist an integer $k \geq 2$ such that $\phi(C_{2k+1})$ has Sidorenko's property?*

Since $\phi(H) = K_2 \Box H$ for bipartite graphs, Theorem 4.3 implies that $\phi(H)$ has Sidorenko's property as long as $H$ does. Hence $\phi(H)$ is 'more likely' to have Sidorenko's property than $H$. For example, since $\phi(K_r) = K_{r,r}$ for integers $r \geq 1$, we know that $\phi(K_r)$ has Sidorenko's property, while $K_r$ is not even a bipartite graph. (Recall that a graph $H$ with odd cycles cannot satisfy Sidorenko's property as $t_H(G) = 0$ for bipartite graphs $G$). Thus, the following question may be posed.

**Question 4.10.** *For a (not necessarily bipartite) graph $H$, does there exist a non-negative integer $k = k_H$ such that $\phi^k(H)$ has Sidorenko's property?*

If Sidorenko's conjecture is true, then it certainly implies that the answers to the questions above are both yes. However, even if Sidorenko's conjecture turns out to be false, it is possible that the answers are still positive.

# Chapter 5

# On the conjecture of Kohayakawa, Nagle, Rödl, and Schacht

In this chapter, we prove the two main results, Theorem 1.7 and 1.8, on the Kohayakawa–Nagle–Rödl–Schacht conjecture.

## 5.1 Subdivisions and Sidorenko's conjecture

Recall that if $H$ is a bipartite graph with Sidorenko's property, it clearly satisfies Conjecture 1.6 with $\rho = 1$. However, the conjecture is also known for some non-bipartite graphs, including complete graphs, complete multipartite graphs, the line graph of the cube [10], and odd cycles [44]. Here we show that any graph for which this conjecture holds may be subdivided to produce a graph that satisfies Sidorenko's conjecture. The *subdivision* of a graph $H$ is the graph obtained from $H$ by replacing its edges with paths of length two that are internally vertex disjoint. The following is the main result of this section:

**Theorem 5.1.** *Suppose that $H$ is a graph for which there exists a function $f : (0,1] \to (0,1]$ and a constant $a = a_H > 0$ such that any $(f(d), d)$-dense graph $G$ contains at least*

$$a|V(G)|^{|V(H)|}d^{|E(H)|}$$

*homomorphic copies of $H$. Then the subdivision of $H$ has Sidorenko's property.*

Theorem 1.7 is an easy consequence of this theorem.

*Proof of Theorem 1.7.* Suppose $H$ satisfies the Kohayakawa–Nagle–Rödl–Schacht conjecture, i.e., for each $\eta > 0$, there exists $\rho = \rho(\eta, d, H)$ such that

$$t_H(G) \geq (1-\eta)d^{|E(H)|} \tag{5.1}$$

whenever $G$ is a sufficiently large $(\rho, d)$-dense graph. We may always assume $\rho < 1$, as decreasing $\rho$ only makes the $(\rho, d)$-denseness condition stronger. Then there exists a function $f : (0,1] \to (0,1]$ defined by $f(d) := \rho(1/2, d, H)$ and a constant $a = 1/2$ such that any $(f(d), d)$-dense graph $G$ contains at least

$$\frac{1}{2}|V(G)|^{|V(H)|}d^{|E(H)|}$$

homomorphic copies of $H$. Thus, by Theorem 5.1, the subdivision of $H$ has Sidorenko's property. □

It will therefore suffice to prove Theorem 5.1. For our proof, it will be convenient to have a version of the hypothesis that applies to weighted graphs as well as graphs.

**Lemma 5.2.** *Suppose that $H$ is a graph for which there exists a function $f : (0,1] \to (0,1]$ and a constant $a = a(H) > 0$ such that any $(f(d), d)$-dense graph $G$ contains at least*

$$a|V(G)|^{|V(H)|}d^{|E(H)|}$$

*homomorphic copies of $H$. Then there exist positive constants $c = c(H)$ and $n_0 = n_0(d, H)$ such that the following holds: If $n \geq n_0$ and $W : [n]^2 \to [0,1]$ is a symmetric function such that $W(v,v) = 0$ for each $v \in [n]$ and $\sum_{u,u' \in U} W(u,u') \geq d|U|^2$ for all $U \subseteq [n]$ with $|U| \geq f(d/2)n$, then*

$$t_H(W) = \frac{1}{n^{|V(H)|}} \sum_{x_1,\ldots,x_h \in [n]} \prod_{ij \in E(H)} W(x_i, x_j) \geq c \cdot d^{|E(H)|}. \tag{5.2}$$

*Proof.* Let $J$ be the graph on $[n]$ obtained by including the edge $uv$ if and only $W(u,v) \geq d/2$. Let $e_J(U) = |E(J[U])|$ for each $U \subseteq [n]$. If $|U| \geq f(d/2)n$, then

$$d|U|^2 \leq \sum_{u,u' \in U} W(u,u') \leq 2e_J(U) + \frac{d}{2}|U|^2.$$

Thus, $e_J(U) \geq \frac{d}{4}|U|^2$, which means $J$ is a $(f(d/2), d/2)$-dense graph on $V(G)$. Therefore, there exist at least $an^{|V(H)|}(d/2)^{|E(H)|}$ homomorphic copies of $H$. For each homomorphic copy of $H$ induced on $\{x_i \in [n] : i \in V(H)\}$, we have

$$\prod_{ij \in E(H)} W(x_i, x_j) \geq \left(\frac{d}{2}\right)^{|E(H)|}.$$

Thus, we obtain the bound

$$\sum_{x_1,\ldots,x_h \in [n]} \prod_{ij \in E(H)} W(x_i, x_j) \geq a 2^{-2|E(H)|} n^{|V(H)|} d^{2|E(H)|}, \tag{5.3}$$

which is still weaker than the desired bound (5.2). However, the following randomisation enables us to obtain (5.2). Let $\mathcal{G} = \mathcal{G}(n, W)$ be the random graph on $[n]$ obtained by including the edge $uv$ with probability $W(u, v)$. Then by Corollary 2.13, $\mathcal{G}$ is $(f(d/2), d/2)$-dense with probability $1 - o(1)$. Thus, by the assumption, we obtain that

$$t_H(\mathcal{G}) \geq a' d^{|E(H)|}, \tag{5.4}$$

where $a' = 2^{-|E(H)|}a$ with probability $1 - o(1)$, if $n$ is large enough. The bound (5.3) yields that $t_H(W) \geq b \cdot d^{2|E(H)|}$ for a constant $b > 0$, and hence, we may apply Corollary 2.12 to obtain

$$\mathbb{P}\left[|t_H(\mathcal{G}) - t_H(W)| \geq \frac{1}{2} t_H(W)\right] \leq C e^{-n^2 t_H(W)/18} \tag{5.5}$$

for a positive constant $C = C(H)$. Note that, again by $t_H(W) \geq b \cdot d^{2|E(H)|}$, $e^{-n^2 t_H(W)} = o(1)$ if $n$ is large enough depending on $d$. By combining (5.4) and (5.5), there exists a random instance $\mathcal{G}$ such that

$$a' d^{|E(H)|} \leq t_H(\mathcal{G}) \leq \frac{3}{2} t_H(W).$$

Hence, taking $c = 2a'/3 = 2^{-|E(H)|+1}a/3$ finishes the proof. □

By using Lemma 3.11, we may assume that the maximum degree of $G$ is at most four times the average degree $d$ in proving Theorem 5.1. We will now argue that any graph of this type contains a large subgraph where the minimum degree is at least $d/4$.

**Lemma 5.3.** *If $G$ is a graph on $n$ vertices with average degree $d$ and maximum degree at most $4d$, there exists a subgraph $G'$ with at least $2^{-9}n$ vertices such that the minimum degree of $G'$ is at least $d/4$.*

*Proof.* We will prove the result by removing one vertex at a time to form graphs $G_0 = G, G_1, \ldots, G_t$, eventually arriving at the required graph. Since the maximum degree of $G$ is at most $4d$, we see that $d(G_i) \leq 4d$ for all $i$, where $d(G_i)$ is the average degree of $G_i$. Suppose now that we have formed the graph $G_i$, but it does not satisfy the required condition. That is, there is a vertex $v_{i+1}$ of degree less than $d/4$. We remove this vertex to form $G_{i+1}$. Since we start from a graph of average degree at least $d$, it follows that $d(G_i) \geq d$ for all $i$. Note that

$$\begin{aligned} d(G_{i+1}) = \frac{2|E(G_{i+1})|}{|V(G_{i+1})|} = \frac{2|E(G_i)| - 2d(v_{i+1})}{|V(G_i)| - 1} &\geq \frac{2|E(G_i)| - d/2}{|V(G_i)| - 1} \\ = d(G_i) + \frac{d(G_i) - d/2}{|V(G_i)| - 1} &\geq d(G_i) + \frac{d}{2(|V(G_i)| - 1)}. \end{aligned}$$

Therefore,

$$d(G_t) \geq d + \frac{d}{2}\left(\frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{n-t}\right) > d + \frac{d}{2}\ln\left(\frac{n}{n-t}\right),$$

which contradicts our observation that $d(G_t) \leq 4d$ when $t = (1 - 2^{-9})n$. Hence, for some $i < t$, $G_i$ satisfies the required condition, completing the proof. □

The key observation in our proof is the next lemma, which says that if a graph has large minimum degree, then the average codegree, taken over all pairs of vertices, in any large set $U$ is itself large.

**Lemma 5.4.** *Suppose that $G$ is a graph with $n$ vertices and minimum degree $\delta$. Then, provided $\delta|U| \geq 2n$,*

$$\sum_{\substack{u,u' \in U \\ u \neq u'}} d(u,u') \geq \frac{\delta^2}{2n}|U|^2$$

*where $d(u,u')$ is the codegree of $u$ and $u'$.*

*Proof.* If we write $d_U(v)$ for the number of neighbors of a vertex $v$ in $U$, we see, by convexity, that

$$\sum_{\substack{u,u' \in U \\ u \neq u'}} d(u,u') = \sum_{v \in V(G)} d_U(v)(d_U(v)-1) \geq n \left( \frac{1}{n} \sum_{v \in V(G)} d_U(v) \right) \left( \frac{1}{n} \sum_{v \in V(G)} d_U(v) - 1 \right)$$
$$= \frac{1}{n} \left( \sum_{u \in U} d(u) \right) \left( \sum_{u \in U} d(u) - n \right) \geq \frac{\delta^2}{2n} |U|^2,$$

as required. □

We now have all the necessary ingredients for proving Theorem 5.1.

*Proof of Theorem 5.1.* Let $H_1$ be the subdivision of $H$. Suppose that $G$ is a graph with maximum degree at most $\frac{8|E(G)|}{|V(G)|}$. By Lemma 3.11, it will be enough to prove that there exists a constant $c$, depending only on $H$, such that

$$|\mathrm{Hom}(H_1, G)| \geq c|V(G)|^{|V(H)|+|E(H)|} \left( \frac{2|E(G)|}{|V(G)|^2} \right)^{2|E(H)|}. \tag{5.6}$$

By Lemma 5.3, $G$ has a subgraph $G'$ with at least $2^{-9}|V(G)|$ vertices and minimum degree $\delta \geq \frac{|E(G)|}{2|V(G)|}$. Now define a weighted graph $W$ on $V(G')$ by $W(u,v) = d(u,v)/|V(G')|$. By Lemma 5.4, the density of $W$ on every set $U \subseteq V(G')$ with $|U| \geq \frac{4|V(G)|}{|E(G)|}|V(G')|$ is at least

$$\frac{1}{|U|^2} \sum_{\substack{u,u' \in U \\ u \neq u'}} W(u,u') \geq \frac{1}{|U|^2} \sum_{\substack{u,u' \in U \\ u \neq u'}} \frac{d(u,u')}{|V(G')|} \geq \frac{|E(G)|^2}{8|V(G)|^4} = \frac{1}{32} t_{K_2}(G)^2. \tag{5.7}$$

From our assumption on $H$, if $J$ is a $(f(d), d)$-dense graph, then

$$t_H(J) \geq a d^{|E(H)|}.$$

Let $d = t_{K_2}(G)^2/32$. If $|V(G)|$ is large enough,[1] then $\frac{4|V(G)|}{|E(G)|} = \frac{8}{t_{K_2}(G)|V(G)|}$ is smaller than $f(d/2)$. Thus, by Lemma 5.2,

$$\sum_{x_1,\ldots,x_h \in V(G')} \prod_{ij \in E(H)} W(x_i, x_j) \geq c|V(G')|^{|V(H)|} \left( \frac{|E(G)|^2}{8|V(G)|^4} \right)^{|E(H)|}$$

[1] We may assume so, since a blow-up $G'$ of $G$, where each vertex and edge is replaced by an equal-sized independent set and a complete bipartite graph, respectively, satisfies $t_F(G) = t_F(G')$ for any fixed graph $F$.

Expanding out $W$, this is the same as

$$\sum_{x_1,\ldots,x_h \in V(G')} \prod_{ij \in E(H)} d(x_i, x_j) \geq c|V(G')|^{|V(H)|+|E(H)|} \left(\frac{|E(G)|}{2^{3/2}|V(G)|^2}\right)^{2|E(H)|}$$
$$\geq 2^{-9|V(H)|-14|E(H)|} c|V(G)|^{|V(H)|+|E(H)|} \left(\frac{2|E(G)|}{|V(G)|^2}\right)^{2|E(H)|}.$$

But the first expression counts homomorphisms from $H_1$ to $G'$, while the last expression is of the required type (5.6). This completes the proof. □

As an application of Theorem 5.1, we may prove the following more general theorem. We define the *$K_{2,t}$-replacement* of a graph $H$ to be the graph obtained from $H$ by replacing each of its edges with a copy of $K_{2,t}$, identifying the edge's endpoints with the two vertices on one side of the corresponding copy of $K_{2,t}$. Note that a $K_{2,1}$-replacement is equivalent to a subdivision.

**Theorem 5.5.** *Suppose that $H$ is a graph for which there exists a function $f : (0,1] \to (0,1]$ and a constant $a > 0$ such that any $(f(d), d)$-dense graph $G$ contains at least*

$$a|V(G)|^{|V(H)|} d^{|E(H)|}$$

*copies of $H$. Then the $K_{2,t}$-replacement of $H$ has Sidorenko's property.*

*Proof.* For a graph $G$ and an ordered sequence of $h = |V(H)|$ vertices $x_1, \ldots, x_h$, let $H_1(x_1, \ldots, x_h)$ be the number of homomorphisms from the subdivision $H_1$ of $H$ to $G$ such that the $i$th vertex of $H$ is mapped to $x_i$. By Theorem 5.1,

$$\sum_{x_1,\ldots,x_h} H_1(x_1, \ldots, x_h) \geq |V(G)|^{|V(H)|+|E(H)|} \left(\frac{2|E(G)|}{|V(G)|^2}\right)^{2|E(H)|}.$$

Therefore, by convexity,

$$\sum_{x_1,\ldots,x_h} H_1(x_1, \ldots, x_h)^t \geq |V(G)|^{|V(H)|} \left(\frac{\sum_{x_1,\ldots,x_h} H_1(x_1, \ldots, x_h)}{|V(G)|^{|V(H)|}}\right)^t$$
$$\geq |V(G)|^{|V(H)|+t|E(H)|} \left(\frac{2|E(G)|}{|V(G)|^2}\right)^{2t|E(H)|}.$$

Since the first expression counts homomorphisms from the $K_{2,t}$-replacement of $H$ to $G$, this completes the proof. □

## 5.2 Counting tree-like graphs in locally dense graphs

In this section, we prove Theorem 1.8 using the tools developed in Chapter 3. We follow an analogous strategy to Theorem 3.8, which proves that $J$-decomposable graphs have Sidorenko's property whenever $J$ is weakly norming.

Firstly, we will prove Theorem 1.8 for the case when $J$ is a complete multipartite graph.

**Theorem 5.6.** *Let $r_1, r_2, \cdots, r_\ell$ be non-negative integers. Then Conjecture 1.6 is true if $H$ is a $K(r_1, r_2, \cdots, r_\ell)$-decomposable graph.*

Let $H$ be a graph with tree-width $r$ and let $(\mathcal{F}, \mathcal{T})$ be the tree decomposition that attains the minimum width. Then by adding edges and vertices to each $X \in \mathcal{F}$ if necessary, we may obtain a graph $H'$ with the same tree decomposition $(\mathcal{F}, \mathcal{T})$ such that each $H'[X]$ is isomorphic to the complete graph $K_{r+1}$ on $(r+1)$ vertices. Such a graph $H'$ is called an *$r$-tree*. By the symmetry of complete graphs, a graph is an $r$-tree if and only if it is $K_{r+1}$-decomposable. Hence, the following theorem is a corollary of Theorem 5.6.

**Theorem 5.7.** *Conjecture 1.6 is true if $H$ is an $r$-tree.*

In particular, the Goldner–Harary graph shown in Figure 5.1, which is a typical example of a 3-tree, satisfies Conjecture 1.6.

We also prove the following result, building upon the theorem of Reiher [44] on odd cycles.

**Theorem 5.8.** *Conjecture 1.6 is true if $H$ is a $C_{2k+1}$-decomposable graph.*

This theorem together with Theorem 5.6 completes the proof of Theorem 1.8. Moreover, it is already shown in Theorem 3.8 that Conjecture 1.6 is true for $C_{2k}$-decomposable graphs. Thus, Theorem 5.8 in fact implies that every $C_k$-decomposable graph satisfies Conjecture 1.6.

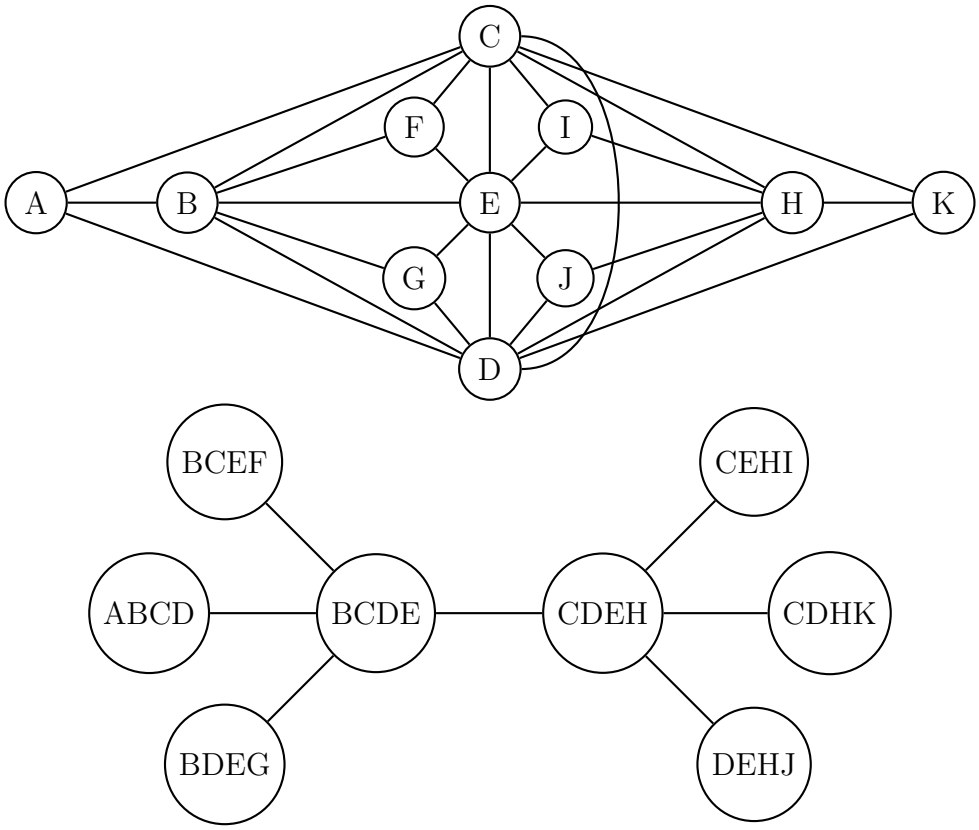



Figure 5.1: The Goldner–Harary graph and its $K_4$-decomposition.

### 5.2.1 Counting $K(r_1, r_2, \cdots, r_\ell)$-decomposable graphs

As a warm-up, we begin by proving the folklore fact that Conjecture 1.6 is true for the case $H = K_r$. It will also be technically helpful in what follows.

**Theorem 5.9.** *Given $\eta, d > 0$ and positive integer $r$, there exists $\rho = \rho(\eta, d, r) > 0$ such that*

$$t_{K_r}(G) \geq (1-\eta)d^{r(r-1)/2}$$

*for every sufficiently large $(\rho, d)$-dense graph $G$.*

This follows immediately from a recursive statement, which states that we may add an apex vertex to any graph satisfying Conjecture 1.6 to obtain another:

**Theorem 5.10.** *Let $\widehat{H}$ be the graph obtained by adding a vertex to $H$ which is adjacent to all vertices in $H$. If Conjecture 1.6 is true for $H$, then it is also true for $\widehat{H}$.*

*Proof.* Let $\rho > 0$ be such that

$$t_H(J) \geq (1-\eta/2)d^{|E(H)|}$$

whenever $J$ is a sufficiently large $(\rho, d)$-dense graph. We may assume that $\rho \leq \eta d/(2|V(H)|)$. Let $G$ be a $(\rho^2, d)$-dense graph on $n$ vertices. Denote by $U$ the set of vertices $v$ in $G$ such that $\deg(v) \geq \rho n$. Let $a$ be the apex vertex in $\widehat{H}$ and let $c(v)$ be the number of homomorphisms $\phi$ from $\widehat{H}$ to $G$ such that $\phi(a) = v$. Observe that for any $W \subseteq V(G)$ of size $|W| \geq \rho n$, the induced subgraph $G[W]$ is $(\rho, d)$-dense. Thus,

$$\begin{aligned}
|\mathrm{Hom}(\widehat{H}, G)| &= \sum_{v \in V(G)} c(v) \geq \sum_{u \in U} |\mathrm{Hom}(H, G[N(u)])| \\
&\geq \sum_{u \in U} (1-\eta/2)d^{|E(H)|}|N(u)|^{|V(H)|} \\
&\geq \frac{(1-\eta/2)d^{|E(H)|}}{|U|^{|V(H)|-1}} \left( \sum_{u \in U} |N(u)| \right)^{|V(H)|}, \qquad (5.8)
\end{aligned}$$

where the last inequality follows from convexity. Note that

$$\sum_{u\in U}|N(u)| = 2|E(G)| - \sum_{v\notin U}|N(v)| \tag{5.9}$$

$$\geq (d-\rho)n^2 \geq \left(1-\frac{\eta}{2|V(H)|}\right)dn^2, \tag{5.10}$$

where the last inequality follows from $\rho \leq \eta d/(2|V(H)|)$. By combining (5.8) and (5.9), we obtain

$$\begin{aligned}|\mathrm{Hom}(\widehat{H},G)| &\geq (1-\eta/2)\left(1-\frac{\eta}{2|V(H)|}\right)^{|V(H)|} d^{|V(H)|+|E(H)|}n^{|V(H)|+1}\\ &\geq (1-\eta)d^{|E(\widehat{H})|}n^{|V(\widehat{H})|}.\end{aligned}$$

□

However, the classical approach above does not give a tight relation between $t_{K_{r+1}}(G)$ and $t_{K_r}(G)$. When applying Theorem 1.4, the main difficulty often lies in bounding the terms $t_{H[X\cap Y]}(G)$ from above in terms of $t_J(G)$. The following lemma gives the control needed to prove Theorem 5.6.

**Lemma 5.11.** *Given $\delta > 0$ and positive integers $\ell, r_1, r_2, \cdots, r_\ell$, let $r = \sum_{i=1}^{\ell} r_i$. Then there exists $\rho = \rho(\delta, d, r_1, r_2, \cdots, r_\ell)$ such that*

$$t_{K(r_1,r_2,\cdots,r_\ell)}(G) \geq (1-\delta)d^{r-r_1}t_{K(r_1-1,r_2,\cdots,r_\ell)}(G)$$

*for every sufficiently large $(\rho, d)$-dense graph $G$.*

*Proof.* Suppose $r_1 \geq 2$. Then the complete $\ell$-partite graph $K(r_1, r_2, \cdots, r_\ell)$ can be obtained by gluing two copies of $K(r_1-1, r_2, \cdots, r_\ell)$ along their subgraphs induced on each vertex set minus a vertex in the first colour, which is isomorphic to $K(r_1-2, r_2, \cdots, r_\ell)$. Hence, by the Cauchy–Schwarz inequality or Theorem 1.4 with $|\mathcal{F}| = 2$, we have

$$t_{K(r_1,r_2,\cdots,r_\ell)}(G) \geq \frac{t_{K(r_1-1,r_2,\cdots,r_\ell)}(G)^2}{t_{K(r_1-2,r_2,\cdots,r_\ell)}(G)}.$$

Here we do not worry about the case $t_{K(r_1-2,r_2,\cdots,r_\ell)}(G) = 0$, because by Theorem 5.9 it is always positive. Repeating this inequality gives the following log-convexity:

$$\frac{t_{K(r_1,r_2,\cdots,r_\ell)}(G)}{t_{K(r_1-1,r_2,\cdots,r_\ell)}(G)} \geq \frac{t_{K(r_1-1,r_2,\cdots,r_\ell)}(G)}{t_{K(r_1-2,r_2,\cdots,r_\ell)}(G)} \geq \cdots \geq \frac{t_{K(1,r_2,\cdots,r_\ell)}(G)}{t_{K(r_2,\cdots,r_\ell)}(G)}. \tag{5.11}$$

Thus, the goal reduces to the case $r_1 = 1$. We claim that, given $\eta > 0$, there exists $\rho = \rho(\eta, d, r_2, \cdots, r_\ell) > 0$ such that

$$t_{K(1,r_2,\cdots,r_\ell)}(G) \geq (1-\eta) d^{r_2} t_{K(r_2+1,\cdots,r_\ell)}(G) \tag{5.12}$$

whenever $G$ is $(\rho, d)$-dense. If the claim is true, then using (5.12) repeatedly to reduce the number of colour classes and applying (5.11) to reduce the number of vertices in each class yields

$$\begin{aligned}\frac{t_{K(r_1,r_2,\cdots,r_\ell)}(G)}{t_{K(r_1-1,r_2,\cdots,r_\ell)}(G)} &\geq \frac{(1-\eta) d^{r_2} t_{K(r_2+1,\cdots,r_\ell)}(G)}{t_{K(r_2,\cdots,r_\ell)}(G)} \\ &\geq \cdots \geq (1-(\ell-1)\eta) d^{r_2+\cdots+r_\ell} \frac{t_{K(r_\ell+1)}(G)}{t_{K(r_\ell)}(G)}.\end{aligned}$$

Since both $K(r_\ell + 1)$ and $K(r_\ell)$ consist of isolated vertices, taking $\eta = \delta/(\ell-1)$ is enough to conclude.

Thus, it remains to prove (5.12). For brevity, let $H = K(r_2+1, r_3, r_4, \cdots, r_\ell)$ and let $h = |V(H)|$. Consider $\rho > 0$ such that $\rho^{1+r_2} \leq \frac{1}{2}\eta d^{r_2+h(h-1)/2}$ and $t_{K_h}(G') \geq \frac{1}{2} d^{h(h-1)/2}$ whenever $G'$ is a sufficiently large $(\rho, d)$-dense graph. Such $\rho$ exists by Theorem 5.9. For a homomorphism $\phi$ from $K(r_3, r_4, \cdots, r_\ell)$ to $G$, define $C_\phi$ to be the set of vertices $v$ such that adding $v$ to $\phi(K(r_3, \cdots, r_\ell))$ extends $\phi$ as a homomorphism from $K(1, r_3, r_4, \cdots, r_\ell)$ to $G$. Denote by $\Phi$ the set of homomorphisms $\phi \in \mathrm{Hom}(H, G)$ such that $|C_\phi| \geq \rho |V(G)|$. Then we have

$$|\mathrm{Hom}(K(1, r_2, r_3, \cdots, r_\ell), G)| \geq \sum_{\phi \in \Phi} |\mathrm{Hom}(K_{1,r_2}, G[C_\phi])| \geq d^{r_2} \sum_{\phi \in \Phi} |C_\phi|^{r_2+1}, \tag{5.13}$$

where the last inequality follows from the fact $t_{K_{1,r}}(J) \geq t_{K_2}(J)^r$ for any graph $J$. Note now that

$$\begin{aligned}\sum_{\phi \in \Phi} |C_\phi|^{r_2+1} &= |\mathrm{Hom}(H, G)| - \sum_{\phi \notin \Phi} |C_\phi|^{r_2+1} \\ &\geq |\mathrm{Hom}(H, G)| - \rho^{1+r_2} |V(G)|^{1+r_2} |\mathrm{Hom}(K(r_3, \cdots, r_\ell), G)| \\ &\geq |\mathrm{Hom}(H, G)| - \rho^{1+r_2} |V(G)|^h.\end{aligned}$$

Substituting this into (5.13) gives

$$|\mathrm{Hom}(K(1, r_2, r_3, \cdots, r_\ell), G)| \geq d^{r_2} |\mathrm{Hom}(H, G)| - \rho^{1+r_2} |V(G)|^{|V(H)|}.$$

Therefore, by normalising both sides by $|V(G)|^{|V(H)|}$,

$$t_{K(1,r_2,r_3,\cdots,r_\ell)}(G) \geq d^{r_2} t_H(G) - \rho^{1+r_2} \geq (1-\eta) d^{r_2} t_H(G),$$

where the last inequality follows from

$$\rho^{1+r_2} \leq \frac{1}{2}\eta d^{r_2+h(h-1)/2} \leq \eta d^{r_2} t_{K_h}(G) \leq \eta d^{r_2} t_H(G). \qquad \square$$

An immediate corollary is that we may compare $t_{K(r_1,r_2,\cdots,r_\ell)}(G)$ and $t_{K(s_1,s_2,\cdots,s_\ell)}(G)$ by repeatedly applying Lemma 5.11 whenever $r_i \geq s_i \geq 0$ for $i = 1, 2, \cdots, \ell$.

**Corollary 5.12.** *Suppose $\delta > 0$ and $\ell, r_1, \cdots, r_\ell$, and $s_1, s_2, \cdots, s_\ell$ are positive integers with $r_i \geq s_i$, $i = 1, 2, \cdots, \ell$. Let $r = |E(K(r_1, r_2, \cdots, r_\ell))|$ and $s = |E(K(s_1, s_2, \cdots, s_\ell))|$. Then there exists a positive $\rho = \rho(\delta, d, r_1, \cdots, r_\ell, s_1, \cdots, s_\ell)$ such that*

$$t_{K(r_1,r_2,\cdots,r_\ell)}(G) \geq (1-\delta) d^{r-s} t_{K(s_1,s_2,\cdots,s_\ell)}(G),$$

*whenever $G$ is a sufficiently large $(\rho, d)$-dense graph.* $\square$

We now turn to the proof of Theorem 5.6.

*Proof of Theorem 5.6.* . Let $K = K(r_1, r_2, \cdots, r_\ell)$ for brevity and let $(\mathcal{F}, \mathcal{T})$ be a $K$-decomposition of a graph $H$. For each edge $e = XY \in E(\mathcal{T})$, denote by $H(e)$ the complete multipartite subgraph of $H$ induced on $X \cap Y$. By Theorem 5.9 we know that $\mathrm{Hom}(H(e), G)$ is non-empty. Hence, Theorem 1.4 gives

$$t_H(G) \geq \frac{t_K(G)^{|\mathcal{F}|}}{\prod_{e \in E(\mathcal{T})} t_{H(e)}(G)} = \frac{t_K(G)^{|E(\mathcal{T})|+1}}{\prod_{e \in E(\mathcal{T})} t_{H(e)}(G)}, \tag{5.14}$$

as $|E(\mathcal{T})| = |\mathcal{F}| - 1$. Using the bound $t_K(G)/t_{H(e)}(G) \geq (1-\delta) d^{|E(K)|-|E(H(e))|}$ from Corollary 5.12, we obtain

$$\begin{aligned} t_H(G) &\geq t_K(G) \prod_{e \in E(\mathcal{T})} (1-\delta) d^{|E(K)|-|E(H(e))|} \\ &\geq (1-\delta) d^{|E(K)|} \prod_{e \in E(\mathcal{T})} (1-\delta) d^{|E(K)|-|E(H(e))|}. \end{aligned}$$

By Lemma 2.5, $|E(H)| = |\mathcal{F}||E(K)| - \sum_{e \in E(\mathcal{T})} |E(H(e))|$, and thus,

$$t_H(G) \geq (1-\delta)^{|\mathcal{F}|} d^{|E(H)|}.$$

Taking $\delta = \eta/|\mathcal{F}|$ concludes the proof. $\square$

### 5.2.2 Counting $C_{2r+1}$-decomposable graphs

In proving Theorem 5.8, we will again follow the same proof strategy. In order to apply Theorem 1.4, the key will be to prove appropriate graph homomorphism inequalities between odd cycles and the paths they contain. In what follows, denote by $P_\ell$ the path of length $\ell$, that is, with $\ell$ edges. The key lemma in this subsection is:

**Lemma 5.13.** *Given $d, \delta > 0$ and positive integers $\ell$ and $r$ with $\ell \leq 2r$, there exists $\rho = \rho(\delta, d, r, \ell)$ such that*

$$t_{C_{2r+1}}(G) \geq (1-\delta)d \cdot t_{P_\ell}(G)^{2r/\ell} \tag{5.15}$$

*whenever $G$ is sufficiently large and $(\rho, d)$-dense.*

There are two main ingredients in the proof. The first is the following lemma of Reiher [44]:

**Lemma 5.14** (Lemma 2.1 in [44])**.** *Let $G$ be a $(\rho, d)$-dense graph on $n$ vertices, and let $f : V(G) \to [0,1]$ be a function satisfying $\sum_{v \in V(G)} f(v) \geq \rho n$. Then*

$$\sum_{uv \in E(G)} f(u)f(v) \geq \frac{d}{2}\left(\sum_{v \in V} f(v)\right)^2 - n.$$

This lemma essentially says that the $(\rho, d)$-denseness condition bootstraps itself to a continuously relaxed version, which will be used to deduce (5.15) for the case $\ell = 2r$.

We will also need the following folklore fact for convexity:

**Lemma 5.15.** *Let $a_0, a_1, \cdots, a_r$ be a real sequence such that, for every $i = 1, 2, \cdots, r-1$,*

$$a_i \leq \frac{a_{i-1} + a_{i+1}}{2}.$$

*Then $a_k \leq \frac{1}{r}(ka_r + (r-k)a_0)$.*

*Proof.* Note that $b_i := a_i - a_{i-1}$ is increasing. As $a_k - a_0 = b_k + b_{k-1} + \cdots + b_1$,

$$\begin{aligned} r(a_k - a_0) &= k(b_1 + \cdots + b_k) + (r-k)(b_1 + \cdots + b_k) \\ &\leq k(b_1 + \cdots + b_k) + k(b_{k+1} + b_{k+2} + \cdots + b_r) = k(a_r - a_0), \end{aligned}$$

which completes the proof. □

The second ingredient in the proof of Theorem 5.8 is the following lemma that enables us to reduce all the cases with $\ell < 2r$ to the case $\ell = 2r$.

**Lemma 5.16.** *For any graph $G$ and positive integers $\ell < 2r$, the following homomorphism inequality holds:*

$$t_{P_\ell}(G) \le t_{P_{2r}}(G)^{\ell/2r}. \tag{5.16}$$

*Proof.* We will repeatedly use the inequality

$$t_{P_{k+t}}(G) \le t_{P_{2k}}(G)^{1/2} t_{P_{2t}}(G)^{1/2}$$

that follows from the Cauchy–Schwarz inequality. If (5.16) holds for all paths of even length, then by $t_{P_{2k+1}}(G) \le t_{P_{2k}}(G)^{1/2} t_{P_{2k+2}}(G)^{1/2}$ we are done. Thus, we may assume that our path is of even length $2\ell$. We claim that the sequence $t_{P_{2k}}(G)$, $k = 1, 2, \cdots, r$, is log-convex. Observe that the Cauchy–Schwarz inequality gives log-convexity for adjacent terms, i.e.,

$$t_{P_{2k}}(G) \le t_{P_{2(k+1)}}(G)^{1/2} t_{P_{2(k-1)}}(G)^{1/2}.$$

Now letting $a_k := \log t_{P_{2k}}(G)$ in Lemma 5.15 gives $a_k \le \frac{1}{r}(ka_r + (r-k)a_0)$. Therefore,

$$t_{P_{2\ell}}(G) \le t_{P_0}(G)^{\frac{r-\ell}{r}} t_{P_{2r}}(G)^{\frac{\ell}{r}}.$$

Since $P_0$ is the single vertex graph, this completes the proof. □

By Lemma 5.16, if we prove that there exists $\rho > 0$ such that

$$t_{C_{2r+1}(G)} \ge (1-\delta) d \cdot t_{P_{2r}}(G)$$

whenever $G$ is a sufficiently large $(\rho, d)$-dense graph, then Lemma 5.13 follows. The proof of this inequality is essentially the same as that of Reiher [44] proving Conjecture 1.6 for odd cycles, but we include it for completeness.

*Proof of Lemma 5.13.* Let $|V(G)| = n$ and let $q(v)$ be the normalised number of walks of length $r$ starting from $v$, i.e., we divide the number of walks by $n^{r-1}$. Denote by $U := \{u : q(u) > \rho n\}$ the set of vertices with large $q(u)$. Then

$$\frac{1}{n^{2r-2}} |\mathrm{Hom}(P_{2r}, G)| = \sum_{u \in U} q(u)^2 + \sum_{u \notin U} q(u)^2,$$

and hence

$$\sum_{u\in U} q(u)^2 \geq |\mathrm{Hom}(P_{2r}, G)|/n^{2r-2} - \rho^2 n^3.$$

On the other hand, let $f_u(v)$ be the normalised number of walks of length $r$ from $u$ to $v$. Then by definition $q(u) = \sum_{v\in V(G)} f_u(v)$, and $\sum_{v\in V(G)} f_u(v) > \rho n$ whenever $u$ is in $U$. For each $u \in U$, Lemma 5.14 gives

$$2\sum_{vw\in E(G)} f_u(v)f_u(w) \geq d\left(\sum_{v\in V(G)} f_u(v)\right)^2 - 2n \geq dq(u)^2 - 2n.$$

Summing this inequality over all $u \in U$ is at most the normalised number of homomorphisms $|\mathrm{Hom}(C_{2r+1}, G)|/n^{2r-2}$. Hence, we have

$$\frac{1}{n^{2r-2}}|\mathrm{Hom}(C_{2r+1}, G)| \geq \frac{d}{n^{2r-2}}|\mathrm{Hom}(P_{2r}, G)| - \rho^2 n^3 - 2n^2.$$

Dividing both sides by $n^3$ gives

$$t_{C_{2r+1}}(G) \geq d \cdot t_{P_{2r}}(G) - \rho^2 - 2/n.$$

By taking $\rho = \sqrt{\delta d^{2r+1}}/4$ and $n > 4/\delta d^{2r+1}$, the inequality $t_{P_{2r}}(G) \geq d^{2r}$ finishes the proof. □

*Proof of Theorem 5.8.* Let $(\mathcal{F}, \mathcal{T})$ be a $C_{2r+1}$-decomposition of $H$ and set $\epsilon = \eta/|\mathcal{F}|$. By Reiher's theorem [44] on odd cycles,[2] there exists $\rho = \rho(\delta, d, \epsilon) > 0$ such that

$$t_{C_{2r+1}}(G) \geq (1-\epsilon)d^{2r+1} \tag{5.17}$$

whenever $G$ is a sufficiently large $(\rho, d)$-dense graph. Let $e_{XY}$ be the number of edges in $H[X \cap Y]$ for $XY \in E(\mathcal{T})$. Each $H[X \cap Y]$, $XY \in E(\mathcal{T})$, is a vertex-disjoint union of paths, and thus by Lemma 5.16 we obtain the upper bound

$$t_{H[X\cap Y]}(G) \leq t_{P_{2r}}(G)^{e_{XY}/2r}.$$

[2] Obviously, it also follows from Lemma 5.13 for the case $\ell = 2r$, which rephrases Reiher's argument.

Combining this bound with Theorem 1.4, we obtain

$$t_H(G) \geq \frac{t_{C_{2r+1}}(G)^{|\mathcal{F}|}}{\prod_{XY\in E(\mathcal{T})} t_{H[X\cap Y]}(G)} \geq \frac{t_{C_{2r+1}}(G)^{|\mathcal{F}|}}{t_{P_{2r}}(G)^{\frac{1}{2r}\sum_{XY\in E(\mathcal{T})} e_{XY}}}.$$

The bound $t_{C_{2r+1}}(G) \geq (1-\epsilon)d \cdot t_{P_{2r}}(G)$ from Lemma 5.13 now gives

$$t_H(G) \geq (1-\epsilon)^{|\mathcal{F}|} d^{|\mathcal{F}|} t_{P_{2r}}(G)^{|\mathcal{F}|-\frac{1}{2r}\sum_{XY\in E(\mathcal{T})} e_{XY}}. \tag{5.18}$$

By Lemma 2.5,

$$|E(H)| = (2r+1)|\mathcal{F}| - \sum_{XY\in E(\mathcal{T})} e_{XY} = |\mathcal{F}| + 2r\left(|\mathcal{F}| - \frac{1}{2r}\sum_{XY\in E(\mathcal{T})} e_{XY}\right),$$

and thus, (5.18) yields

$$\begin{aligned} t_H(G) &\geq (1-|\mathcal{F}|\epsilon) d^{|\mathcal{F}|} t_{P_{2r}}(G)^{|\mathcal{F}|-\frac{1}{2r}\sum_{XY\in E(\mathcal{T})} e_{XY}} \\ &\geq (1-\eta) d^{|\mathcal{F}|} t_{P_{2r}}(G)^{\frac{1}{2r}(|E(H)|-|\mathcal{F}|)} \geq (1-\eta) d^{|E(H)|}. \end{aligned}$$ □

## 5.3 Concluding remarks

**Asymmetric $K_{2,t}$-replacements.** It is also possible to assign different $K_{2,t}$-replacements for distinct edges in Theorem 5.5 when the replaced graph $H$ is edge-transitive. For example, suppose that $H$ is a triangle and consider

$$X = \mathbb{E}_{x,y,z} d(x,y)^r d(y,z)^s d(z,x)^t.$$

This is a normalised count for the number of homomorphisms to $G$ of the graph formed by replacing the edges of the triangle by a $K_{2,r}$, a $K_{2,s}$, and a $K_{2,t}$, respectively. Note, by symmetry, that $X^3$ can be expressed in the form

$$X^3 = \mathbb{E}[d(x,y)^r d(y,z)^s d(z,x)^t]\mathbb{E}[d(x,y)^s d(y,z)^t d(z,x)^r]\mathbb{E}[d(x,y)^t d(y,z)^r d(z,x)^s].$$

Hence, by Hölder's inequality with exponent 3,

$$X \geq \mathbb{E}[(d(x,y)d(y,z)d(z,x))^{(r+s+t)/3}] \geq \mathbb{E}[d(x,y)d(y,z)d(z,x)]^{(r+s+t)/3},$$

where the last inequality follows by convexity. Therefore, the graph obtained by replacing the edges of the triangle by a $K_{2,r}$, a $K_{2,s}$, and a $K_{2,t}$ has Sidorenko's

property. This argument easily generalises to all complete graphs and other edge-transitive graphs.

**An approximate version of Conjecture 1.6 for graphs with bounded tree-width.** For each $m$-edge graph with tree-width $t$ and a tree decomposition $(\mathcal{F}, \mathcal{T})$ with the minimum width, it is enough to add $\binom{t+1}{2}|\mathcal{F}|$ edges to obtain a $t$-tree. Lemma 2.3 implies $\binom{t+1}{2}|\mathcal{F}| \leq \binom{t+1}{2}m$ and this gives an approximate version of Conjecture 1.6 for graphs with bounded tree-width.

**Corollary 5.17.** *Let $H$ be a graph with $m$ edges and tree-width $t$. Given $\eta > 0$, there exists $\rho = \rho(\eta, d, m, t) > 0$ such that*

$$t_H(G) \geq (1-\eta)d^{\left(\frac{t(t+1)}{2}+1\right)m}$$

*for every sufficiently large $(\rho, d)$-dense graph $G$.*

**Homomorphism inequalities for locally dense graphs.** To guarantee the existence of an exponent $c$ such that $t_H(G) \geq t_J(G)^c$, it is often assumed that there exists a homomorphism from $H$ to $J$. However, as seen in Lemma 5.11 or in Lemma 5.13, this condition is not necessary when proving analogous inequalities provided $G$ is locally dense. Hence, it might be interesting to find more homomorphism inequalities specialised for locally dense graphs.

**Open cases for Conjecture 1.6.** Using Theorems 5.6, 5.8, 5.9, 5.10, and the theorem of Reiher [44] on odd cycles, one may check that Conjecture 1.6 is true for $H$ with at most 4 vertices. Many graphs on 5 vertices can also be verified to satisfy the conjecture, but we do not know how to handle a 5-cycle with a chord, which is perhaps the simplest open case.

**Question 5.18.** *Is Conjecture 1.6 true for $H$ isomorphic to $C_5$ with a chord?*

In Chapter 6, we will prove an analogous result that bipartite graphs obtained naturally from faces of a regular polytope are (weakly) norming. It might also be possible to verify Conjecture 1.6 for highly symmetric graphs.

**Question 5.19.** *Is Conjecture 1.6 true for $H$ isomorphic to the 1-skeleton of a regular polytope? In particular, is it true for an icosahedron or an dodecahedron?*

As discussed before proposing Question 3.13, if it is possible to generate random copies of $K_r$ and $K_s$ that have the same marginals on $K_t$ for every $t \leq \min(r, s)$ in locally dense graphs, then it may be possible to prove Conjecture 1.6 for all chordal graphs.

# Chapter 6

# Finite reflection groups and graph norms

Recall that, from (1.9) and (1.10), a graph $H$ is norming if the functional on the space of bounded Lebesgue measurable symmetric functions defined by

$$\|f\|_H := \left| \int \prod_{ij \in E(H)} f(x_i, x_j) d\mu^{|V(H)|} \right|^{1/|E(H)|}$$

is a semi-norm, and weakly norming if

$$\|f\|_{r(H)} := \left( \int \prod_{ij \in E(H)} |f(x_i, x_j)| d\mu^{|V(H)|} \right)^{1/|E(H)|}$$

is a norm.

When proving that $\|\cdot\|_{r(H)}$ is a norm, all of the difficulties lie in proving the triangle inequality. Hatami's work in [26] started from the observation that a Hölder-like inequality is equivalent to the triangle inequality for $\|\cdot\|_{r(H)}$. To state his condition, Theorem 2.8 in [26], we have to introduce some notation generalising (1.9) and (1.10). Let $m = |E(H)|$ and let $\chi : E(G) \to [m]$ be a (not necessarily proper) edge colouring of $H$. Consider a sequence $\mathcal{F} = (f_i)_{i=1}^m$ of bounded symmetric measurable functions on $[0,1]^2$ and define

$$\langle \mathcal{F}; \chi \rangle_H := \int \prod_{e=ij \in E(H)} f_{\chi(e)}(x_i, x_j) d\mu^{|V(H)|}. \tag{6.1}$$

Note that if $f_i = |f|$ for all $i = 1, 2, \cdots, m$, then $\langle \mathcal{F}; \chi \rangle_H = \|f\|_{r(H)}^{|E(H)|}$, while if $f_i = f$, then $|\langle \mathcal{F}; \chi \rangle_H| = \|f\|_H^{|E(H)|}$. Hatami's result now says that the triangle

inequality holds for $\|\cdot\|_{r(H)}$ if and only if

$$\langle \mathcal{F}; \chi \rangle_H \leq \prod_{e \in E(H)} \left\| f_{\chi(e)} \right\|_{r(H)} \tag{6.2}$$

for all choices of $\mathcal{F}$ and $\chi$. Furthermore, $\|\cdot\|_H$ is a semi-norm if and only if the analogous inequality obtained by replacing $\left\|f_{\chi(e)}\right\|_{r(H)}$ with $\left\|f_{\chi(e)}\right\|_H$ holds. We remark that to prove (6.2), it is enough to prove it for injective maps $\chi$.

We will think of (6.2) in the following terms: regard the functions $f_1, \cdots, f_m$ as $m$ distinct graphs on the same vertex set[1] and imagine each edge of $f_i$ has the colour $i$. Then $\langle \mathcal{F}; \chi \rangle_H$ is the number of (homomorphic) copies of $H$ which are coloured according to $\chi$, i.e., each edge $e \in E(H)$ receives the colour $\chi(e)$. In particular, if $\chi$ is a one-to-one map then $\langle \mathcal{F}; \chi \rangle_H$ counts the number of 'rainbow' copies of $H$, while $\|f_i\|_{r(H)}^{|E(H)|}$ counts the number of monochromatic copies of $H$ in colour $i$. Thus, (6.2) is equivalent to the statement that the number of rainbow copies of $H$ is bounded above by the geometric mean of the number of monochromatic copies in each colour.

The proof that (6.2) holds for reflection graphs has two steps. In the first step, discussed in Section 6.2, we show that any graph whose edges percolate in an appropriate way under the action of a certain natural family of automorphisms is weakly norming. This statement, Theorem 6.5, is more general than Theorem 1.10, and may be of independent interest, but the resulting condition needs to be verified by hand for any particular graph or class of graphs. Accordingly, the second step in our proof, discussed in Section 6.4, is to find a general argument that verifies this condition for all reflection graphs. It is here that we use results from the theory of finite reflection groups.

Suppose now that $H$ is a weakly norming graph, $f$ is a bounded symmetric measurable function on $[0,1]^2$, and $e^*$ is an edge of $H$. If we put $f_1 = |f|$, $f_2 = f_3 = \cdots = f_m = 1$, and make $\chi$ one-to-one with $\chi(e^*) = 1$, then $\langle \mathcal{F}; \chi \rangle_H = \|f\|_{r(K_2)}$, where $K_2$ is just a single edge, so (6.2) implies that

$$\|f\|_{r(K_2)} \leq \|f\|_{r(H)} . \tag{6.3}$$

[1]There are large graphs approximating $f_1, \cdots, f_m$ by the limit theory of dense graphs [38].

That is, when $H$ is weakly norming, $H$ has Sidorenko's property (1.2). More generally, given a subgraph $J$ of a weakly norming graph $H$, letting $f_{\chi(e)} = 1$ for all $e \notin E(J)$ in inequality (6.2) gives that the following monotonicity holds for every subgraph $J$ of a weakly norming graph $H$:

$$t_H(G) \geq t_J(G)^{|E(H)|/|E(J)|}. \tag{6.4}$$

Using this inequality together with Theorem 1.4, another application to Sidorenko's conjecture has already been given in Section 3.2.2.

Finally, in Section 6.5, some hypergraph generalisations and applications to the theory of quasirandom hypergraphs will be discussed.

## 6.1 A motivating example

It is already non-trivial to show that (6.3) holds, even for graphs as simple as paths [3] or trees [49, 48], so it should not be surprising that it is more difficult to prove the strictly stronger inequality (6.2). To motivate what follows, we will prove that $C_6$, the cycle of length six, is norming without invoking spectral graph theory or Schatten–von Neumann norms. Let $H = C_6$ be the graph with vertex set $\{1, 2, \cdots, 6\}$, where $i$ and $i+1$ are adjacent for all $i$ (with addition taken modulo 6), and let $\mathcal{F} = (f_1, f_2, \cdots, f_6)$ be a sequence of six bounded symmetric measurable functions on $[0,1]^2$.

To show that (6.2) holds, we may assume that $\chi$ is a one-to-one map, that is, a rainbow colouring, as it is in the most general case. Without loss of generality, put $\chi(e_i) = i$ if $e_i$ is the edge between $i$ and $i+1$ modulo 6. Define functions $g$ and $h$ by

$$g(x_1, x_4) = \int f_1(x_1, x_2) f_2(x_2, x_3) f_3(x_3, x_4)\ dx_2 dx_3, \text{ and}$$
$$h(x_1, x_4) = \int f_4(x_4, x_5) f_5(x_5, x_6) f_6(x_6, x_1)\ dx_5 dx_6,$$

so that $\int gh\ dx_1 dx_4 = \langle \mathcal{F}; \chi \rangle_{C_6}$. The function $g$ gives the (normalised) count of rainbow walks from $x_1$ to $x_4$ which are coloured, in order, with the colours 1, 2, and 3. The function $h$ can be similarly interpreted, but using the colours 6, 5, and

4. A simple application of the Cauchy–Schwarz inequality gives

$$\langle \mathcal{F}; \chi \rangle_{C_6}^2 \leq \int g^2 \int h^2, \tag{6.5}$$

where here and throughout this chapter we suppress the variables of integration if they are clear from context. Observe now that

$$\int g^2 = \langle \mathcal{F}; \chi_L \rangle_{C_6},$$

where $\chi_L$ is the colouring obtained by doubling the 'left half' of the rainbow cycle, represented by the vector $(1, 2, 3, 3, 2, 1)$. A similar equation also holds for $h$, namely, $\int h^2 = \langle \mathcal{F}; \chi_R \rangle_{C_6}$, where $\chi_R = (6, 5, 4, 4, 5, 6)$ is the colouring obtained by doubling the 'right half' of the rainbow cycle. We have therefore bounded $\langle \mathcal{F}; \chi \rangle_{C_6}$ from above by the geometric mean of two functions of the same form, but simpler in the sense that they both contain fewer colours. Repeating this procedure twice more, by first applying the Cauchy–Schwarz inequality to each of $\langle \mathcal{F}; \chi_L \rangle_{C_6}$ and $\langle \mathcal{F}; \chi_R \rangle_{C_6}$ with respect to the variables $(x_2, x_5)$ and then applying the Cauchy–Schwarz inequality to each of the four resulting terms with respect to the variables $(x_3, x_6)$, we find that

$$\langle \mathcal{F}; \chi \rangle_{C_6} \leq \prod_{i=1}^{8} \langle \mathcal{F}; \chi_i \rangle_{C_6}^{1/8},$$

where

$$\begin{aligned} &\chi_1 = (1,1,1,1,1,1),\ \chi_2 = (2,1,1,2,2,2),\ \chi_3 = (2,2,2,2,3,3),\ \chi_4 = (3,3,3,3,3,3),\\ &\chi_5 = (6,6,6,6,6,6),\ \chi_6 = (5,6,6,5,5,5),\ \chi_7 = (5,5,5,5,4,4),\ \chi_8 = (4,4,4,4,4,4). \end{aligned} \tag{6.6}$$

Recall that $\langle \mathcal{F}; \chi_1 \rangle_{C_6} = \|f_1\|_{C_6}^6$ and hence it gives one of the terms in the desired upper bound, though with an incorrect exponent.

We now iterate this 3-step process of applying the Cauchy–Schwarz inequality along different vertex cuts, noting that the more often we repeat the process the more monochromatic colourings appear in our upper bound. In particular, as in (6.6), after each repetition, the sum of the exponents of the non-monochromatic forms reduces by at least a factor of two. Hence, if we iterate $k$ times, we get the bound

$$\langle \mathcal{F}; \chi \rangle_{C_6} \leq \prod_{i=1}^{2^{3k}} \langle \mathcal{F}; \chi_{i,k} \rangle_{C_6}^{1/2^{3k}},$$

where at least $1 - 2^{-k}$ of the $2^{3k}$ colourings $\chi_{i,k}$ are monochromatic. We may rewrite this inequality as

$$\langle \mathcal{F}; \chi \rangle_{C_6} \leq \prod_{i=1}^{6^6} \langle \mathcal{F}; \chi_i' \rangle_{C_6}^{\alpha_{i,k}},$$

where each $\chi_i'$ represents one of the $6^6$ possible edge-colourings of $C_6$ with 6 colours. In particular, we assume that $\chi_1', \chi_2', \cdots, \chi_6'$ represent the monochromatic colourings in $1, 2, \cdots, 6$, respectively. By the argument above, $\alpha_{1,k} + \cdots + \alpha_{6,k} \geq 1 - 2^{-k}$ and it is also easy to see that $\alpha_{i,k}$ is non-decreasing in $k$ for all $i = 1, 2, \cdots, 6$. Therefore, taking the limit as $k$ tends to infinity, we have

$$\langle \mathcal{F}; \chi \rangle_{C_6} \leq \prod_{i=1}^{6} \langle \mathcal{F}; \chi_i' \rangle_{C_6}^{\alpha_i} = \prod_{i=1}^{6} \|f_i\|_{C_6}^{6\alpha_i}, \tag{6.7}$$

where $\alpha_1 + \alpha_2 + \cdots + \alpha_6 = 1$. Therefore, if $\alpha_1, \alpha_2, \cdots, \alpha_6$ were equal, this would be the desired inequality. If they are not equal, then starting from $\chi' = (2, 3, 4, 5, 6, 1)$ we obtain the inequality (addition is again taken modulo 6)

$$\langle \mathcal{F}; \chi \rangle_{C_6} = \langle \mathcal{F}; \chi' \rangle_{C_6} \leq \prod_{i=1}^{6} \|f_{i+1}\|_{C_6}^{6\alpha_i},$$

where $\langle \mathcal{F}; \chi' \rangle_{C_6} = \langle \mathcal{F}; \chi \rangle_{C_6}$ follows from the symmetry of the 6-cycle. Repeating the same idea four more times, we can cyclically permute the exponents in the inequality to all six possible positions. Taking the product of these six inequalities then completes the proof.

## 6.2 Cut involutions and Cauchy–Schwarz inequalities

There were three steps to the proof given in the previous section: firstly, we showed how to apply the Cauchy–Schwarz inequality along certain vertex cuts; then we showed that a monochromatic edge-colouring of $H$ can be obtained through a sequence of such Cauchy–Schwarz inequalities; finally, we used a limiting argument and the edge transitivity of $C_6$ to complete the proof. In this section, we generalise these arguments. The first step will be generalised through the use of a natural class of graph automorphisms which we call cut involutions, while an appropriate

generalisation of the last step is relatively straightforward. Generalising the second step, that is, finding a monochromatic edge-colouring, proves more difficult, so in this section we reduce it to a simpler question which we will resolve in Section 6.4 for the special case of reflection graphs.

To generalise the Cauchy–Schwarz inequality (6.5), we need to find a vertex cut which splits the graph into two isomorphic pieces, which we refer to, informally, as the left and right halves. To be more precise, given a connected graph $H$, say that an automorphism $\phi$ of $H$ is a *cut involution* if the following hold:

(i) $\phi$ is an involution, i.e., $\phi = \phi^{-1}$;

(ii) the fixed point set $F_\phi = \{v \in V(H) : \phi(v) = v\}$ is a vertex cut of $H$;

(iii) there exists a partition $L_\phi \cup R_\phi$ of $V(H) \setminus F_\phi$ such that no edge crosses between $L_\phi$ and $R_\phi$ and $\phi(L_\phi) = R_\phi$.

Let $W_H$, the *cut involution group* of $H$, be the subgroup of the automorphism group of $H$ generated by the cut involutions.

Since it is possible for $H \setminus F_\phi$ to have more than two components, the left and right halves of a cut involution may not be well-defined. In reflection graphs, we can avoid this ambiguity by focusing on certain special cut involutions. For now, we will simply assume that the left and right halves are somehow fixed for each cut involution $\phi$. That is, we will assume that there is a fixed pair of disjoint subsets $L_\phi$ and $R_\phi$ of $V(H) \setminus F_\phi$ that are unions of connected components in $H \setminus F_\phi$ and which are mapped to each other under $\phi$.

Define the *left-folding map* $\phi^+ : V(H) \to V(H)$ of a cut involution $\phi$ by

$$\phi^+(v) = \begin{cases} \phi(v) & \text{if } v \in R_\phi \\ v & \text{if } v \in L_\phi \cup F_\phi, \end{cases}$$

and, similarly, define the *right-folding map* $\phi^-$ of $\phi$ by swapping the roles of $L_\phi$ and $R_\phi$. The following lemma formalises applications of the Cauchy–Schwarz inequality with respect to a cut involution.

**Lemma 6.1.** *Let $H$ be an $m$-edge graph and let $\mathcal{F} = (f_1, f_2, \cdots, f_m)$ be a sequence of non-negative bounded measurable symmetric functions on $[0,1]^2$. For each edge-colouring $\chi : E(H) \to \{1, 2, \cdots, m\}$ and cut involution $\phi$,*

$$\langle \mathcal{F}; \chi \rangle_H \leq \left\langle \mathcal{F}; \chi \circ \phi^+ \right\rangle_H^{1/2} \left\langle \mathcal{F}; \chi \circ \phi^- \right\rangle_H^{1/2}. \tag{6.8}$$

*Proof.* As $\phi$ will be fixed throughout the proof, let $F := F_\phi$, $L := L_\phi$, and $R := R_\phi$ for brevity. Let $E_F := E(H[F])$, $E_L := E(H[L \cup F]) \setminus E_F$, and $E_R := E(H[R \cup F]) \setminus E_F$. Note that $E_F$, $E_L$, and $E_R$ partition $E(H)$, and also that $E_L$ and $E_R$ are mapped to each other by $\phi$. For brevity, let $\mathbf{x}_F := (x_v)_{v \in F}$. Define

$$\begin{aligned} g_L(\mathbf{x}_F) &:= \int \prod_{ij \in E_L} f_{\chi(ij)}(x_i, x_j) \prod_{k \in L} dx_k, \\ g_R(\mathbf{x}_F) &:= \int \prod_{ij \in E_R} f_{\chi(ij)}(x_i, x_j) \prod_{k \in R} dx_k, \text{ and} \\ g_F(\mathbf{x}_F) &:= \prod_{ij \in E_F} f_{\chi(ij)}(x_i, x_j). \end{aligned}$$

If $E_F$ is empty, then let $g_F := 1$. Since no edge crosses between $L$ and $R$, integrating variables $x_v$, $v \in L \cup R$, gives that

$$\langle \mathcal{F}; \chi \rangle_H = \int \prod_{ij \in E(H)} f_{\chi(ij)}(x_i, x_j) d\mu^{|V(H)|} = \int g_L g_R g_F(\mathbf{x}_F) d\mu^{|F|}.$$

The Cauchy–Schwarz inequality now implies

$$\int g_L g_R g_F(\mathbf{x}_F) d\mu^{|F|} \leq \left( \int g_L^2 g_F(\mathbf{x}_F) d\mu^{|F|} \right)^{1/2} \left( \int g_R^2 g_F(\mathbf{x}_F) d\mu^{|F|} \right)^{1/2}. \quad (6.9)$$

Since $\phi$ is an isomorphism between the two induced subgraphs $H[F \cup R]$ and $H[F \cup L]$ that fixes $F$,

$$\begin{aligned} g_L^2 g_F(\mathbf{x}_F) &= \left( \int \prod_{ij \in E_L} f_{\chi(ij)}(x_i, x_j) \prod_{k \in L} dx_k \right)^2 g_F(\mathbf{x}_F) \\ &= \left( \int \prod_{ij \in E_L} f_{\chi(ij)}(x_i, x_j) \prod_{k \in L} dx_k \right) \left( \int \prod_{ij \in E_R} f_{\chi(\phi(ij))}(x_i, x_j) \prod_{k \in R} dx_k \right) g_F(\mathbf{x}_F) \end{aligned}$$

and hence,

$$\int g_L^2 g_F(\mathbf{x}_F) d\mu^{|F|} = \left\langle \mathcal{F}; \chi \circ \phi^+ \right\rangle_H.$$

Similarly, $\int g_R^2 g_F(\mathbf{x}_F) d\mu^{|F|} = \langle \mathcal{F}; \chi \circ \phi^- \rangle_H$. Thus, (6.9) reduces to (6.8). □

In (6.8), $\chi \circ \phi^+$ is the colouring of $H$ where the colour of the edges in $R_\phi \cup F_\phi$ copies the corresponding edges in $L_\phi \cup F_\phi$. Similarly, $\chi \circ \phi^-$ is the colouring where the colour of the edges in $L_\phi \cup F_\phi$ copies the corresponding edges in $R_\phi \cup F_\phi$.

To obtain (6.8), the non-negativity of the functions in $\mathcal{F}$ is used in (6.9) only if there is an edge inside $F_\phi$. Conversely, if there is no edge of $H$ fixed by $\phi$, then the non-negativity of functions in $\mathcal{F}$ is unnecessary. This observation leads naturally to the definition of a *stable involution*. This is a cut involution $\phi$ such that the fixed point set $F_\phi$ contains no edge of $H$, that is, it is an independent (or stable) set in $H$ as well as a vertex cut. This leads us to the following variant of Lemma 6.1:

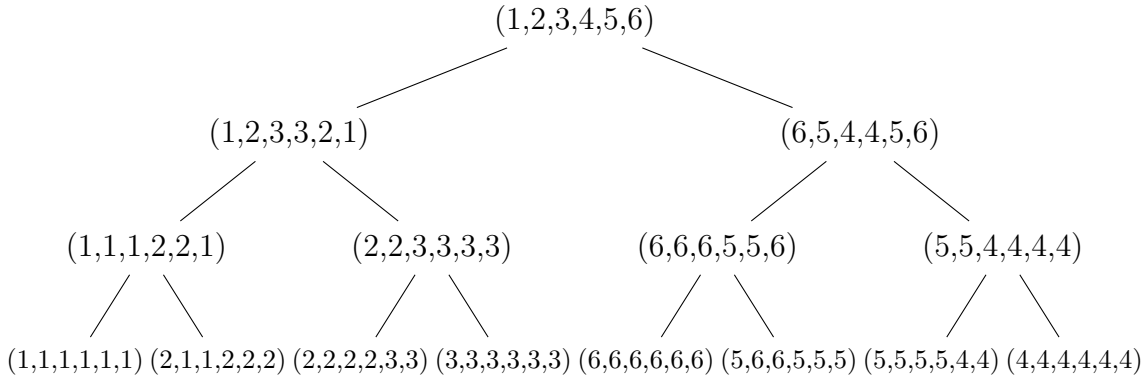
**Lemma 6.2.** *Let $H$ be an $m$-edge graph and let $\mathcal{F} = (f_1, f_2, \cdots, f_m)$ be a sequence of bounded measurable symmetric functions on $[0,1]^2$. For each edge-colouring $\chi : E(H) \to \{1, 2, \cdots, m\}$ and stable involution $\phi$,*

$$\langle \mathcal{F}; \chi \rangle_H \leq \left\langle \mathcal{F}; \chi \circ \phi^+ \right\rangle_H^{1/2} \left\langle \mathcal{F}; \chi \circ \phi^- \right\rangle_H^{1/2}.$$

The subgroup of the automorphism group of $H$ generated by the stable involutions is then called the *stable involution group* of $H$ and is denoted $S_H$. In what follows, we will show that graphs whose edges percolate under the action of the cut involution group are weakly norming. Using the simple observation above, it is possible to prove analogous results for the stronger norming property by substituting the stable involution group for the cut involution group. We will not always be explicit about this below, choosing to focus instead on cut involutions and the weakly norming property. Nevertheless, it is worth bearing in mind.

Consider now a sequence $\phi_1, \phi_2, \cdots, \phi_d$ of cut involutions and suppose that we wish to apply a sequence of Cauchy–Schwarz inequalities of the form (6.8), first with $\phi = \phi_1$, then with $\phi = \phi_2$, and so on. Let $\mathcal{T}$ be the rooted binary tree of depth $d$ encoding which colourings have been obtained through such iterations: the vertices are labelled by $m$-edge-colourings of $H$, the root is labelled by the initial colouring $\chi_0$, and each vertex at depth $i < d$ labelled with $\chi$, say, has two children with labels $\chi \circ \phi_{i+1}^+$ and $\chi \circ \phi_{i+1}^-$. We call this tree the *Cauchy–Schwarz tree* associated with $(\phi_i)_{i=1}^d$ rooted at $\chi_0$. As here, we will often abuse notation by identifying a vertex and its label.

**Example 6.3.** The 3-step process described in the previous section corresponds to the following Cauchy–Schwarz tree of depth 3:




Roughly speaking, a Cauchy–Schwarz tree shows how colours spread under applications of Cauchy–Schwarz inequalities of the form (6.8). With this terminology, the following theorem generalises part of what was proven in the previous section.

**Theorem 6.4.** *Let $H$ be an edge-transitive $m$-edge graph and $(\phi_i)_{i=1}^d$ be a finite sequence of cut involutions of $H$. Let $\chi : E(G) \to [m]$ be a rainbow edge-colouring. If the Cauchy–Schwarz tree $\mathcal{T}$ associated with $(\phi_i)_{i=1}^d$ rooted at $\chi$ contains a leaf labelled with a monochromatic colouring $\chi'$, then $H$ is weakly norming. Moreover, if $(\phi_i)_{i=1}^d$ are stable involutions with the same property, then $H$ is norming.*

*Proof.* Consider the Cauchy–Schwarz tree $\mathcal{T}$ associated with $(\phi_i)_{i=1}^d$ rooted at $\chi$, such that amongst its leaves $\chi_{1,d}, \cdots, \chi_{2^d,d}$ there exists a leaf labelled with the monochromatic colouring $\chi'$. By applying (6.8) repeatedly, we arrive at the upper bound

$$\langle \mathcal{F}; \chi \rangle_H \leq \prod_{i=1}^{2^d} \langle \mathcal{F}; \chi_{i,d} \rangle_H^{1/2^d} .$$

We now iterate the whole $d$-step process $k$ times. Consider the Cauchy–Schwarz tree $\mathcal{T}'$ of depth $dk$ associated with the repeated sequence rooted at $\chi$. For $i = 1, 2, \cdots, m$, let $\chi_i'$ be the monochromatic colouring that only uses the colour $i$. Observe that all descendants of a vertex in $\mathcal{T}'$ that is labelled with $\chi_i'$ are again labelled with $\chi_i'$. Moreover, by our main assumption, every iteration of the $d$-step process makes at least a $1/2^d$ proportion of the non-monochromatic leaves

monochromatic. We may rearrange the upper bound obtained from $\mathcal{T}'$ as

$$\langle \mathcal{F}; \chi \rangle_H \leq \prod_{i=1}^{m^m} \langle \mathcal{F}; \chi_i' \rangle_H^{\alpha_{i,k}} ,$$

where each $\chi_i'$ represents one of the $m^m$ possible edge-colourings of $H$ with $m$ colours. By the remarks above, $\alpha_{1,k} + \cdots + \alpha_{m,k} \geq 1 - (1 - 1/2^d)^k$ and each $\alpha_{i,k}$ with $1 \leq i \leq m$ is non-decreasing in $k$. Therefore, taking the limit as $k$ tends to infinity, we get

$$\langle \mathcal{F}; \chi \rangle_H \leq \prod_{i=1}^{m} \langle \mathcal{F}; \chi_i' \rangle_H^{\alpha_i} ,$$

where $\alpha_1 + \cdots + \alpha_m = 1$. It remains to prove that we may take $\alpha_i = 1/m$. Let $\sigma$ be a permutation on $[m]$. Consider an edge colouring $\chi_\sigma := \sigma \circ \chi$. Then this new colouring $\chi_\sigma$ colours the edge $e$ with the permuted colour $\sigma(\chi(e))$. Repeating the same argument as above, but with a Cauchy–Schwarz tree rooted at $\chi_\sigma$, we have the inequality

$$\langle \mathcal{F}; \chi_\sigma \rangle = \langle \mathcal{F}; \sigma \circ \chi \rangle \leq \prod_{i=1}^{m} \langle \mathcal{F}; \chi_i' \rangle_H^{\beta_i} ,$$

where $\beta_i = \alpha_j$ if $i = \sigma(j)$. Let $\Sigma_H$ be the set of all permutations $\sigma$ of $[m]$ such that $\chi$ and $\chi_\sigma$ colour $H$ in an isomorphic way. That is, there exists an automorphism $\phi$ of $H$ such that $\chi(e) = \chi_\sigma(\phi(e))$ for each $e \in E(H)$. Taking the geometric mean of these upper bounds over all permutations $\sigma \in \Sigma_H$, we get

$$\langle \mathcal{F}; \chi \rangle = \langle \mathcal{F}; \chi_\sigma \rangle \leq \prod_{i=1}^{m} \langle \mathcal{F}; \chi_i' \rangle_H^{\gamma_i} ,$$

where

$$\gamma_i = \frac{1}{|\Sigma_H|}(\alpha_1 |A_{1\to i}| + \alpha_2 |A_{2\to i}| + \cdots + \alpha_m |A_{m\to i}|).$$

Here $A_{j\to i}$ denotes the set of permutations in $\Sigma_H$ sending $j$ to $i$. Note that all $A_{j\to i}$ have the same size, as they are all cosets of the subgroup $A_{i\to i}$. Therefore, $\gamma_i = 1/m$ for all $i$, as required. □

The theorem above shows that if we want to prove that a graph $H$ is weakly norming, it is sufficient to show that there exists a Cauchy–Schwarz tree with a

rainbow root and a monochromatic leaf. We say that a vertex map $\phi$ from $H$ into itself is a *folding map* if it is a left- or right-folding map of a cut involution of $H$. Suppose now that $J$ is an edge subset of $H$ and $\psi$ a folding map of $H$. Now consider

$$\psi^{-1}(J) = \{e \in E(H) : \psi(e) \in J\}.$$

That is, we copy the edges of $J$ from the left half onto the right half, if $\psi$ is a left-folding map of a cut involution, and the other way around if it is a right-folding map. We will look at sequences $J_0, J_1, J_2, \cdots$ of edge subsets of $H$ such that, for each $i \geq 0$, there exists a folding map $\psi$ that satisfies $J_{i+1} = \psi^{-1}(J_i)$, where $J_0 = \{e_0\}$ for some edge $e_0$ and $J_N = E(H)$ for some finite $N$. We call such a sequence $\{e_0\} = J_0, J_1, \cdots, J_N = E(H)$ a *percolating sequence*.

In other words, a percolating sequence exists if and only if there exists a sequence $\psi_1, \psi_2, \cdots, \psi_N$ of folding maps such that there exists a fixed $e_0 \in E(H)$ satisfying $\psi_1\psi_2\cdots\psi_N(e) = e_0$ for every $e \in E(H)$. If such a sequence of folding maps exists, then we say that $H$ *collapses* under its cut involutions and that the sequence $\psi_1, \psi_2, \cdots, \psi_N$ of folding maps is a *collapsing sequence* of $H$. Since each folding map is either an identity or a cut involution when restricted to a single edge, the image $\psi_1\psi_2\cdots\psi_N(e)$ under a sequence of folding maps $(\psi_i)_{i=1}^N$ reduces to some $\phi_1\phi_2\cdots\phi_k(e)$, where each $\phi_j$, $1 \leq j \leq k$, is a cut involution. Thus, if $H$ collapses under its cut involution group $W_H$, then it must be edge-transitive under $W_H$. With this terminology, we may rephrase Theorem 6.4.

**Theorem 6.5.** *Suppose that a graph $H$ collapses under its cut involutions. Then $H$ is weakly norming.*

We say that a graph $H$ *collapses under its stable involutions* if there exists a collapsing sequence that consists of folding maps of stable involutions. The analogue of Theorem 6.5 for the full norming property is then as follows.

**Theorem 6.6.** *Suppose that a graph $H$ collapses under its stable involutions. Then $H$ is norming.*

The key question now is to find a collapsing sequence or, equivalently, a percolating sequence. It is often convenient to switch between the two equivalent

goals. Although each cut involution $\phi$ gives enough flexibility when choosing between $(\phi^+)^{-1}(J)$ and $(\phi^-)^{-1}(J)$ to guarantee that the number of edges does not decrease, this is far from guaranteeing that a percolating sequence exists. Indeed, as we choose between $(\phi^+)^{-1}(J)$ and $(\phi^-)^{-1}(J)$, we may lose as well as gain edges, making it difficult to keep track of the changes. As shown by the following example, even a subset of cut involutions that acts transitively on edges may not produce a percolating sequence:

**Example 6.7.** Let $H = K_{1,4}$ and let $a, b, c$, and $d$ be leaves of $H$ that are incident to the edges $e_a, e_b, e_c$, and $e_d$, respectively. Consider the two cut involutions $\phi$ and $\psi$ such that $\phi(a) = c$, $\psi(a) = d$, $L_\phi = L_\psi = \{a, b\}$, and $R_\phi = R_\psi = \{c, d\}$. Then $H$ is edge-transitive under the subgroup generated by $\phi$ and $\psi$. Let $J = \{e_a, e_c\}$. Then $(\phi^+)^{-1}(J) = (\phi^-)^{-1}(J) = J$, $(\psi^+)^{-1}(J) = \{e_a, e_d\}$, and $(\psi^-)^{-1}(J) = \{e_b, e_c\}$. No matter how we repeat this process, we have exactly one edge in $F_\phi \cup L_\phi$ and another in $F_\phi \cup R_\phi$. Thus, no percolating sequence exists.

Nevertheless, one may easily check that any known example of a weakly norming graph collapses if the whole cut involution group can be used. The main purpose of Section 6.4.1 is to find a common generalisation for these ad hoc arguments.

## 6.3 Preliminaries on finite reflection groups

In this section, we state some preliminary facts about reflection groups, which will be used in what follows. To make the discussion more concise, we focus on the case of finite groups, which will also be enough for our purpose. For more details, we refer the reader to [2] and [27].

### 6.3.1 Root systems and simple reflections

Let $W$ be a finite reflection group in $\mathbf{GL}(n, \mathbb{R})$ and let $T$ be the family of reflections in $W$. Denote by $\Phi$ the set of unit vectors orthogonal to the reflection hyperplanes, where each hyperplane gives rise to two vectors $\alpha$ and $-\alpha$. This set of unit vectors $\Phi$ is called a *root system* and each element $\alpha \in \Phi$ is called a *root*. Fixing an ordered basis $\{u_1, u_2, \cdots, u_n\}$ of $\mathbb{R}^n$, we say that a root $\alpha$ is *positive* if $\alpha = \sum c_i u_i$ and

$c_k > 0$, where $k$ is the smallest index $i$ for which $c_i \neq 0$. Otherwise, a root is said to be *negative*. Clearly, the set $\Phi^+$ of positive roots and $\Phi^-$ of negative roots partition $\Phi$ and are of equal size. Let $\Delta$ be a minimal subset of $\Phi^+$ such that each $\alpha \in \Phi^+$ is a linear combination of positive roots in $\Delta$ with non-negative coefficients. Such a minimal subset always exists, since $\Phi^+$ itself already satisfies the condition. We call this $\Delta$ a *simple system* and its elements are called *simple roots*. Given a simple root $\alpha$, the hyperplane orthogonal to $\alpha$ is called a *simple reflection hyperplane*. We denote by $s_\alpha$ the reflection through the simple reflection hyperplane orthogonal to $\alpha$ and refer to such reflections as *simple reflections*.

The following theorem states some important facts about simple systems. For a proof, we refer the reader to Sections 1.3 and 1.5 of [27].

**Theorem 6.8.** *A simple system $\Delta$ has the following properties:*

*(i) it is unique with respect to $\Phi^+$;*

*(ii) it consists of linearly independent vectors;*

*(iii) the set of all simple reflections generates $W$.*

### 6.3.2 Chambers

For each positive root $\alpha$, let $H_\alpha$ be the reflection hyperplane orthogonal to $\alpha$. Consider the collection $\mathcal{C}$ of connected components of $\mathbb{R}^n \setminus \cup_{\alpha \in \Phi^+} H_\alpha$. Each component $C$ in $\mathcal{C}$ is called an *open chamber* of the reflection group $W$. Denote by $\overline{\mathcal{C}} = \{\overline{C} : C \in \mathcal{C}\}$ the set of closures of open chambers, which we call *closed chambers*. Each open chamber consists of those vectors $v$ in $\mathbb{R}^n$ with a certain fixed sign for $\langle v, \alpha \rangle$ for each positive root $\alpha$. Conversely, if we fix a sign for $\langle v, \alpha \rangle$ for each positive root $\alpha$, then, provided these choices are consistent, the collection of such vectors is an open chamber. In particular, there is a unique open chamber consisting of all vectors $v$ satisfying $\langle v, \alpha \rangle > 0$ for all positive roots $\alpha$, since it is possible to have $\langle v, \alpha \rangle > 0$ for all simple roots $\alpha$ and any such vector must have a positive inner product with every positive root. We call this chamber $C_0$ (or its closure $\overline{C}_0$) the *fundamental open (or closed) chamber*. Note that the fundamental closed chamber $\overline{C}_0$ is a cone given by the intersection of closed half-spaces

obtained from simple roots:

$$\overline{C}_0 = \bigcap_{\alpha \in \Delta} \{v \in \mathbb{R}^n : \langle v, \alpha \rangle \geq 0\}.$$

In other words, it is a closed cone surrounded by simple reflection hyperplanes, so a point $v \in \overline{C}_0$ must be contained in either a simple reflection hyperplane or the open chamber $C_0$.

An important fact is that the action of $W$ on $\mathcal{C}$ is simply transitive. We refer the reader to Sections 1.6 and 1.7 of [27] for a proof.

**Theorem 6.9.** *The action of $W$ on $\mathcal{C}$ is simply transitive. In particular, the identity is the only element in $W$ that fixes the fundamental open chamber.*

It immediately follows that the action of $W$ on the set $\bar{\mathcal{C}}$ of all closed chambers is also simply transitive. For each reflection $t \in T$, let $H(t)$ be the hyperplane in $\mathbb{R}^n$ defining $t$ and let $D^+(t)$ be the open half-space $\{x \in \mathbb{R}^n : \langle x, \alpha \rangle > 0\}$, where $\alpha$ is the positive root orthogonal to $H(t)$. Let $S \subset T$ be the set of all simple reflections and $I$ be a subset of $S$. Define the cone $C(I)$ by

$$C(I) := \left( \bigcap_{s \in I} H(s) \right) \cap \left( \bigcap_{s \in S \setminus I} D^+(s) \right).$$

Observe that the fundamental open chamber $C_0$ is $C(\emptyset)$ and $C(I)$ and $C(J)$ are disjoint whenever $I \neq J$. Moreover, the fundamental closed chamber can be expressed as

$$\overline{C}_0 = \bigcup_{I \subset S} C(I).$$

The simple transitivity of the $W$-action may be extended as follows:

**Proposition 6.10** (Theorem 3A8, [39])**.** *The cone $C(I)$ is mapped into the fundamental closed chamber by $w \in W$ if and only if $w$ is in the subgroup of $W$ generated by $I$. In particular, the stabiliser of $C(I)$ is precisely the subgroup generated by $I$.*

Although we refer the reader to Theorem 3A8 in [39] for more information, we remark that one direction of the proposition above is easy to see: if $w = s_1 s_2 \cdots s_k$ for $s_i \in I$, then $w$ fixes each point in $C(I)$ since the subspace $\bigcap_{s \in I} H(s)$ is fixed under each $s_i$. Note, therefore, that if $C(I)$ is mapped into the fundamental closed chamber by an element of $W$, it is in fact mapped to itself.

**Example 6.11.** Let $W \cong S_n$ be the reflection group in $\mathbf{GL}(n, \mathbb{R})$ with reflection hyperplanes

$$H_{ij} := \{(x_1, x_2, \cdots, x_n) : x_i = x_j\}$$

for $1 \leq i < j \leq n$. Note that the reflection $t_{ij}$ through $H_{ij}$ is the map swapping $x_i$ and $x_j$. With respect to the standard ordered basis $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$, the set of positive roots is

$$\Phi^+ = \left\{ \frac{1}{\sqrt{2}}(\mathbf{e}_i - \mathbf{e}_j) : 1 \leq i < j \leq n \right\}.$$

The simple system $\Delta$ is then given by

$$\Delta = \left\{ \frac{1}{\sqrt{2}}(\mathbf{e}_i - \mathbf{e}_{i+1}) : 1 \leq i \leq n-1 \right\}$$

and each permutation $\sigma \in S_n$ corresponds to an open chamber

$$C_\sigma = \{(x_1, \cdots, x_n) : x_{\sigma(1)} < x_{\sigma(2)} < \cdots < x_{\sigma(n)}\}.$$

It is clear that $W$ is a simply transitive action on the set $\mathcal{C} = \{C_\sigma : \sigma \in S_n\}$ of all open chambers.

### 6.3.3 The length function

Each $w \in W$ can be expressed as $w = s_1 s_2 \cdots s_k$, where $s_1, s_2, \cdots, s_k \in S$. Define the *length* $\ell(w)$ of $w \in W$ to be the minimum length $k$ over all such expressions. This purely combinatorial concept of length has a geometric interpretation, which will play a crucial role in Section 6.4.

**Theorem 6.12.** *Let $w \in W$ and $v \in \mathbb{R}^n$ be a point in the fundamental open chamber. For each positive root $\alpha$ and the reflection $t_\alpha \in T$ induced by $\alpha$, $\ell(t_\alpha w) > \ell(w)$ holds if and only if $\langle w(v), \alpha \rangle$ is positive.*

The theorem above says that $t_\alpha w$ has a greater length than $w$ if and only if $w(v)$ lies on the 'positive side' of $H_\alpha$, i.e., the same side as the positive root $\alpha$, whereas its image $t_\alpha w(v)$ under the reflection $t_\alpha$ lies on the other side. Observe now that a root $\alpha$ is positive if and only if $\langle v, \alpha \rangle > 0$ whenever $v$ is chosen from the fundamental open chamber. Thus, the 'positive side' with respect to the hyperplane $H_\alpha$ is

exactly the component of $\mathbb{R}^n \setminus H_\alpha$ containing the fundamental open chamber. To summarise, Theorem 6.12 implies that the length of $t_\alpha w$ becomes greater than $w$ if and only if by applying $t_\alpha$ we move $w(v)$ from the same side of $H_\alpha$ as the fundamental open chamber to the opposite side. Theorem 6.12 is a key fact in the theory of Coxeter groups, so proofs can be found in many places: for instance, Sections 1.6 and 1.7 in [27] or Proposition 4.4.6 in [2].

### 6.3.4 Regular polytopes

In this subsection, some brief definitions and elementary facts on regular polytopes to prove Theorem 1.9 will be given, following [39]. Again for simplicity, we focus on finite polytopes.

A *n-dimensional polytope* $\mathcal{P}$ (or simply *n-polytope*) is a family of subsets of a convex set $K \subseteq \mathbb{R}^n$ defined as follows:

- $\mathcal{P}$ is a lattice under the partial order given by inclusion;
- each element of $\mathcal{P}$ is called a *face* and every maximal chain of faces is called a *flag* of $\mathcal{P}$. A chain has *length* $i$ if it consists of $i+1$ faces;
- the *least face* $F_{-1}$ is the empty set;
- the affine hull of the *greatest face* $F_n = K$ is $n$-dimensional;
- a *k-dimensional face* is a convex set whose affine hull is a $k$-dimensional subspace.
- each flag has length $n+1$, i.e., consists of exactly $n+2$ faces including $F_{-1}$ and $F_n$.

A *homomorphism* of a polytope is a map between the set of faces that preserves incidence and an *isomorphism* is a homomorphic bijection with a homomorphic inverse.

Two flags $\Phi$ and $\Psi$ are *j-adjacent* if they differ only at the $j$-faces, or simply *adjacent* if they are $j$-adjacent for some $j$. A polytope $\mathcal{P}$ is *strongly connected* if, for any flags $\Phi$ and $\Psi$, there exists a sequence $\Phi = \Phi_0, \Phi_1, \cdots, \Phi_k = \Psi$ of flags such that each $\Phi_i$ contains $\Phi \cap \Psi$ and $\Phi_j$ and $\Phi_{j+1}$ are adjacent.

The *symmetry group* $G(\mathcal{P})$ of $\mathcal{P}$ is the set of the isometries $f$ of $\mathbb{R}^n$ such that $f(\mathcal{P}) = \mathcal{P}$. We say that a polytope $\mathcal{P}$ is *regular* if $G(\mathcal{P})$ acts transitively on the set of flags.

If $F$ and $G$ are incident faces of dimension $i-1$ and $i+1$, respectively, then there exist exactly two $i$-faces $H$ such that $F < H < G$. Hence, a flag $\Phi$ has a unique $j$-adjacent flag $\Phi^j$ for each $j = 0, 1, \cdots, n-1$. In a regular polytope, the involutary isometry $s_{j+1}$ that maps $\Phi$ to $\Phi^j$ is a reflection, which motivates the following fact:

**Theorem 6.13** (Theorem 1B1, [39])**.** *Let $\mathcal{P}$ be a regular $n$-polytope and let a flag $\Phi$ be fixed. Denote by $(s_j)_{j=1}^n$ the reflections such that each $s_{j+1}$ maps $\Phi$ to $\Phi^j$. Then the symmetry group $G(\mathcal{P})$ of $\mathcal{P}$ is a finite reflection group generated by the reflections $S := \{s_1, s_2, \cdots, s_n\}$. Moreover, one may take a basis of $\mathbb{R}^n$ such that $S$ is the set of simple reflections.*

The following theorem essentially asserts that the $j$-faces in $\Phi$ can be identified as cosets of the subgroup generated by $S \setminus \{s_{j+1}\}$, which will be used to deduce Theorem 1.9.

**Theorem 6.14** (Theorem 3D7, [39])**.** *Let $\mathcal{P}$ be a regular $n$-polytope and $W := G(\mathcal{P})$ be the symmetry group generated by simple reflections $S = \{s_1, \cdots, s_n\}$ given by Theorem 6.13. For $k = 0, 1, \cdots, n-1$, define*

$$F_k := \bigcup_{w \in \langle s_1, s_2, \cdots, s_k \rangle} w\overline{C}(\{k+2, k+3, \cdots, n\})$$

*and set $F_n = \mathbb{R}^n, F_{-1} = \emptyset$. Then $\hat{\mathcal{P}} := \{wF_k : -1 \le k \le n, w \in W\}$ is a regular polytope isomorphic to $\mathcal{P}$.*

# 6.4 Euclidean embeddings of reflection graphs

## 6.4.1 The Euclidean embedding

We begin with a motivating example:

**Example 6.15.** Let $H$ be the 1-subdivision of $K_4$. This graph is *biregular*, that is, it is a bipartite graph such that all vertices on the same side of the bipartition

have the same degree. In this case, all vertices on one side have degree three, while all vertices on the other side have degree two. One may check that the cut involution group $W_H$ of $H$ is isomorphic to the symmetric group $S_4$. Note also that $S_4$ may be represented as the reflection group of the tetrahedron. There is a natural way to embed $H$ into $\mathbb{R}^4$ that uses these observations: consider the map $\eta : V(H) \to \mathbb{R}^4$ which sends the vertices $v_1, v_2, v_3$, and $v_4$ of degree 3 to $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and $\mathbf{e}_4$, respectively, and the unique common neighbour $u_{ij}$ of $v_i$ and $v_j$, $i \neq j$, to the midpoint $\frac{1}{2}(\mathbf{e}_i + \mathbf{e}_j)$ of their images under $\eta$. Then every reflection in $S_4$ becomes a cut involution of $H$ and, conversely, every cut involution is represented by a reflection.

This example shows that some graphs may be embedded in Euclidean space so that cut involutions are represented by genuine reflections. Our aim in this subsection will be to confirm the existence of a similar embedding for every reflection graph.

We will continue to use notation from Section 6.3. That is, we let $W \subset \mathbf{GL}(n, \mathbb{R})$ be a finite reflection group, $T$ be the family of reflections in $W$, and $S \subset T$ be the set of simple reflections with respect to some fixed ordered basis of $\mathbb{R}^n$. Fix also subsets $S_1$ and $S_2$ of $S$ and let $W_1$ and $W_2$ be the subgroups of $W$ generated by $S_1$ and $S_2$, respectively. For brevity in what follows, we will write $H$ for the $(S_1, S_2; W)$-reflection graph. Recall that this is the bipartite graph between the (left-)cosets of $W_1$ and $W_2$, where $wW_1$ and $wW_2$ are adjacent for every $w \in W$.

In Example 6.15, we mapped the vertices of $H$ to points in $\mathbb{R}^n$ so as to visualise the graph more easily. However, to avoid unnecessary ambiguity in the general case, we will instead construct a map $\eta$ that sends vertices of $H$ to cones in $\mathbb{R}^n$. Indeed, let $\eta$ be the map from $V(H)$ to the family $\{wC(I) : w \in W, I \subset S\}$ of cones in $\mathbb{R}^n$ such that, for $i = 1, 2$ and each $w \in W$,

$$\eta(wW_i) := wC(S_i).$$

In particular, $\eta(W_1) = C(S_1)$ and $\eta(W_2) = C(S_2)$, i.e., they are mapped into subcones of the closed fundamental chamber. Since $uW_i = wW_i$ if and only if

$w^{-1}u \in W_i$ and the cone $C(S_i)$ is fixed by each $g \in W_i$, we have

$$\eta(uW_i) = uC(S_i) = w(w^{-1}u)C(S_i) = wC(S_i) = \eta(wW_i),$$

and hence the map $\eta$ is well-defined. We call the map $\eta$ the *Euclidean embedding* of $H$.

**Example 6.16.** It will be instructive to revisit the case where $H$ is the 1-subdivision of $K_4$. Following the notation in Example 6.11 with $n = 4$, let $W \cong S_4$ be the reflection group in $\mathbf{GL}(4, \mathbb{R})$ with the reflection $t_{ij}$ swapping $x_i$ and $x_j$, $1 \leq i < j \leq 4$. Let $S_1 = \{t_{23}, t_{34}\}$ and $S_2 = \{t_{12}, t_{34}\}$. One may check that the $(S_1, S_2; W)$-reflection graph $H$ is isomorphic to the 1-subdivision of $K_4$ by listing all the cosets of $W_1$ and $W_2$ and their incidence, but the Euclidean embedding gives a more geometric way to see this. Note first that the Euclidean embedding $\eta$ is then given by

$$\eta(W_1) = \{(x_1, x_2, x_3, x_4) : x_1 > x_2 = x_3 = x_4\},$$
$$\eta(W_2) = \{(x_1, x_2, x_3, x_4) : x_1 = x_2 > x_3 = x_4\},$$

and the other values are given by permuting some coordinates of $\eta(W_1)$ and $\eta(W_2)$. Thus, there are 4 and 6 vertices, respectively, on each side of the bipartition. Observe that $W_1$ and $W_2$ are the only vertices such that their images under $\eta$ are subsets of the fundamental closed chamber $\overline{C}_0 = \{(x_1, x_2, x_3, x_4) : x_1 \geq x_2 \geq x_3 \geq x_4\}$, since for $\eta(wW_i)$ to be contained in $\overline{C}_0$, the 'large' coordinates need to come first. Moreover, as the edges are of the form $(wW_1, wW_2)$ for some $w \in W$, $u$ and $v$ are adjacent if and only if $\eta(u)$ and $\eta(v)$ are contained in the same closed chamber $wC_0$. For instance, both $\eta(W_2)$ and $\eta(t_{12}W_1) = t_{12}\eta(W_1)$ are contained in the closed chamber $C = \{(x_1, x_2, x_3, x_4) : x_2 \geq x_1 \geq x_3 \geq x_4\}$, and one may check that $W_1$ and $t_{12}W_1$ are the only vertices adjacent to $W_2$. Repeating this argument gives that the $(S_1, S_2; W)$-reflection graph $H$ is isomorphic to the 1-subdivision of $K_4$. In fact, the points used in Example 6.15 to embed vertices are 'typical' points from the corresponding cones.

The key properties of the Euclidean embedding are spelled out in the following lemma.

**Proposition 6.17.** *Let $H$ be the $(S_1, S_2; W)$-reflection graph. Then the Euclidean embedding $\eta$ of $H$ has the following properties:*

*(i) for two vertices $a$ and $a'$ on the same side of the bipartition of $H$, $\eta(a) = \eta(a')$ if and only if $a = a'$;*

*(ii) every closed chamber $\overline{C}$ contains exactly one image of a vertex under $\eta$ in each colour class. Moreover, if both $\eta(a)$ and $\eta(b)$ are subsets of $\overline{C}$, then $ab \in E(H)$;*

*(iii) there exists a closed chamber $\overline{C}$ containing both $\eta(a)$ and $\eta(b)$ whenever $ab$ is an edge in $H$.*

*Proof.* (i) Suppose that $a = wW_1$, $a' = uW_1$ and $\eta(a) = \eta(a')$. By the definition of $\eta$, we have $\eta(a) = wC(S_1)$ and $\eta(a') = uC(S_1)$. Since $\eta(a) = \eta(a')$, it follows that $uw^{-1}$ fixes $C(S_1)$. By Proposition 6.10, this implies that $uw^{-1} \in W_1$, and hence $a = a'$.

(ii) It follows from the construction of $\eta$ that the fundamental closed chamber $\overline{C}_0$ contains $\eta(W_1)$ and $\eta(W_2)$. Again by Proposition 6.10, $\eta(W_1)$ is mapped to a cone in $\overline{C}_0$ by $w \in W$ if and only if $w \in W_1$ and hence every point in $\eta(W_1)$ is mapped to itself. Since $\eta$ is one-to-one on each side of the bipartition by (i), $W_1$ and $W_2$ are the only vertices of $H$ that are mapped to subsets of $\overline{C}_0$. Moreover, each closed chamber $w\overline{C}_0$ contains exactly one pair of cones $(\eta(wW_1), \eta(wW_2))$ that are images of an adjacent pair $(wW_1, wW_2)$, since the action of $W$ is simply transitive on $\mathcal{C}$ by Theorem 6.9 and the fundamental closed chamber contains $\eta(W_1)$ and $\eta(W_2)$.

(iii) Since $ab \in E(H)$, there exists $w \in W$ such that $a = wW_1$ and $b = wW_2$. Then $\eta(a)$ and $\eta(b)$ must be contained in $w\overline{C}_0$, where $\overline{C}_0$ is the fundamental closed chamber. □

The next corollary highlights the fact that we have identified a class of cut involutions in $H$, namely, those corresponding to genuine reflections in the Euclidean embedding.

**Corollary 6.18.** *Let $H$ be the $(S_1, S_2; W)$-reflection graph. Then every reflection $t \in T$ is a cut involution of $H$.*

*Proof.* It is straightforward to check that every $t \in T$ acts as an involutary graph automorphism on $H$. Indeed, the fixed point set $F_t$ consists of the vertices in $H$ that are mapped to the cones contained in the reflection hyperplane $H(t)$. Let $L_t$ and $R_t$ be the vertex subsets mapped into the open half-spaces $D^+(t)$ and $D^-(t)$, respectively. Then clearly $L_t$ and $R_t$ are mapped to each other by $t$. Suppose that there is an edge between $u \in L_t$ and $v \in R_t$, Then by part (iii) of the proposition above there exists a closed chamber $\overline{C}$ containing both $\eta(u)$ and $\eta(v)$. Thus, $\overline{C}$ must intersect both $D^+(t)$ and $D^-(t)$, but this is impossible since $C$ is a connected component of $\mathbb{R}^n \setminus \cup_{t \in T} H(t)$. □

One may wonder if the converse of the corollary above also holds, that is, whether every cut involution of a reflection graph $H$ becomes a reflection under the Euclidean embedding. However, this is not true in general, as evidenced by the following example:

**Example 6.19.** Suppose that $W \cong S_4$, $S = \{t_{12}, t_{23}, t_{34}\}$, and $S_1 = \{t_{23}, t_{34}\}$, as in Example 6.16, but take $S_2 = S$ instead. Then the $(S_1, S_2; W)$-reflection graph is isomorphic to the star $K_{1,4}$. However, it is not possible to render all cut involutions so that they correspond to reflections. This is because $K_{1,4}$ has a cut involution which is a product of two other cut involutions and, therefore, the determinants of all three of the corresponding linear transformations cannot be $-1$.

We have shown that every reflection $t \in T$ induces a cut involution on $H$, mapping the vertex $wW_i$ to the vertex $twW_i$. Since $T$ generates $W$, the graph $H$ is edge-transitive under the action of these cut involutions. Thus, in order to apply Theorem 6.5, it remains to show that there exists a collapsing sequence. This will be the topic of the next subsection.

### 6.4.2 Proof of Theorem 1.10

Proposition 6.17 allows us to lift the notion of folding sequences to a group theoretic setting. Here is an illustrative example:

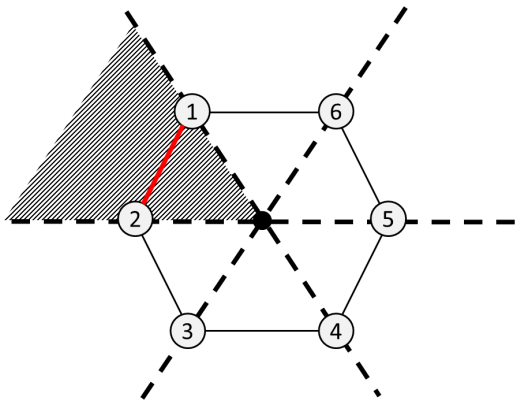



Figure 6.1: The Euclidean embedding of $C_6$.

**Example 6.20.** Recalling the example of $C_6$ on vertex set $\{1, 2, \cdots, 6\}$ discussed in Section 6.1, suppose that we have oriented each cut involution in such a way that the component containing the edge $\{1, 2\}$ is on the left. Then the folding sequence starting from the single edge set $J_0 := \{\{1, 2\}\}$ with $J_{i+1} = (\phi_i^+)^{-1}(J_i)$ at the $i$-th step, where the $\phi_i$ are chosen as in Section 6.1, is a percolating sequence. Now consider $C_6$ as the $(S_1, S_2; W)$-reflection graph, where $W = I_2(3) \subset \mathbf{GL}(2, \mathbb{R})$ is the dihedral group, $S = \{s_1, s_2\}$ is a generating set of simple reflections of $I_2(3)$ with angle $\pi/3$ between the reflection lines defining $s_1$ and $s_2$, $S_1 = \{s_1\}$, and $S_2 = \{s_2\}$. Figure 6.1 shows what $\eta$ looks like if we label $W_1$ and $W_2$ with 1 and 2, respectively: we have chosen a basis of $\mathbb{R}^2$ in such a way that the fundamental open chamber $C_0$, represented by the grey region, is the unique chamber whose closure contains $\eta(1)$ and $\eta(2)$ as subsets and each vertex is mapped to a ray from the origin as marked. The percolating sequence described above naturally gives a percolating sequence on chambers, allowing us to spread the grey colour on the fundamental chamber across all chambers. For instance, we may spread the colour on the grey chamber in Figure 6.1 to the chamber containing the edge $\{2, 3\}$ by applying the reflection through the horizontal line through 2 and 5. In this subsection, we will approach the general question in the opposite direction, using the algebraic tools built up over the previous subsections to show the existence of a collapsing sequence.

Let $\eta$ be the Euclidean embedding of the $(S_1, S_2; W)$-reflection graph $H$. For each reflection $t \in T$ and the associated positive root $\alpha$, let $t^+$ and $t^-$ be the left

and right-folding maps defined by

$$t^+(x) := \begin{cases} t(x) & \text{if } \langle x, \alpha \rangle < 0 \\ x & \text{if } \langle x, \alpha \rangle \geq 0 \end{cases} \text{ and } t^-(x) := \begin{cases} x & \text{if } \langle x, \alpha \rangle \leq 0 \\ t(x) & \text{if } \langle x, \alpha \rangle > 0, \end{cases}$$

respectively. We say that $\tau : \mathbb{R}^n \to \mathbb{R}^n$ is a *folding map* of $W$ if it is a left or right-folding map of a reflection $t \in W$. As every folding map $\tau$ maps an open chamber to another, we abuse the notation slightly to regard $\tau$ as a map from the set $\mathcal{C}$ of all open chambers into itself. Given a family $\mathcal{J}$ of open chambers, we are again interested in whether the inverse image

$$\tau^{-1}(\mathcal{J}) = \{C \in \mathcal{C} : \tau(C) \in \mathcal{J}\}$$

'expands' to the set of all open chambers $\mathcal{C}$. As for the graph case, we say that the group $W$ *collapses* by folding maps $\tau_1, \tau_2, \cdots, \tau_N$ if $\tau_1 \tau_2 \cdots \tau_N C = C_0$ for any open chamber $C$, or simply $W$ *collapses* if such a sequence $\tau_1, \tau_2, \cdots, \tau_N$ of folding maps exists. Note that $W$ collapses if and only if there exists a finite sequence $\mathcal{J}_0, \mathcal{J}_1, \cdots, \mathcal{J}_N$ of families of open chambers such that

$$\mathcal{J}_{i+1} = \tau^{-1}(\mathcal{J}_i),\ \mathcal{J}_0 = \{C_0\},\ \text{and } \mathcal{J}_N = \mathcal{C}.$$

Recall that each reflection $t \in T$ acts in two different ways: it acts on the set of chambers $\mathcal{C}$ group theoretically and on the vertex set $V(H)$ consisting of all cosets of $W_1$ and $W_2$ as a cut involution. To distinguish these two actions, denote by $\phi_t$ the cut involution of $H$ that corresponds to $t \in T$. If $\tau = t^+$ is a folding map, then denote by $\psi_\tau = \phi_t^+$, i.e., $\tau$ and $\psi_\tau$ are folding maps on $\mathcal{C}$ and $V(H)$, respectively. Define $\psi_\tau$ for $\tau = t^-$ in the symmetric way.

The following lemma proves that, whenever $W$ collapses by a sequence of folding maps, the sequence of corresponding folding maps on $V(H)$ sends every vertex in the same side to the same vertex, namely the one embedded into the fundamental closed chamber.

**Lemma 6.21.** *Suppose $W$ collapses by $\tau_1, \tau_2, \cdots, \tau_N$. Let $\psi_i := \psi_{\tau_i}$. Then for each $i = 1, 2$, $\psi_1 \psi_2 \cdots \psi_N(wW_i) = W_i$.*

*Proof.* It suffices to show for the case $i = 1$. For each open chamber $C$, let $v_C$ be the vertex in the side $\{wW_1 : w \in W\}$ such that $\eta(v_C) \subset \overline{C}$, whose uniqueness is

guaranteed by Proposition 6.17 (ii). In particular, $v_{C_0} = W_1$. We claim that

$$v_{\tau(C)} = \psi_\tau(v_C). \tag{6.10}$$

Suppose $\tau = t^+$ for a reflection $t \in T$. Recall that we orient each cut involution $\phi_t$ in such a way that

$$L_{\phi_t} := \{x \in V(H) : \eta(x) \subset D^+(t))\} \text{ and } R_{\phi_t} := \{x \in V(H) : \eta(x) \subset D^-(t)\}.$$

For brevity, we write $L_t := L_{\phi_t}$, $R_t := R_{\phi_t}$, and $F_t := V(H) \setminus (L_t \cup R_t)$. Suppose that $C \subseteq D^+(t)$. Then $\tau(C) = t^+(C) = C$, and thus $v_{\tau(C)} = v_C$. On the other hand, as $\eta(v_C)$ is a cone in $\overline{C}$ disjoint from $D^-(t)$, $v_C$ is in $L_t \cup F_t$, which implies that $v_C$ is fixed under $\psi_\tau = \phi_t^+$. Hence, $v_{\tau(C)} = v_C = \psi_\tau(v_C)$. Suppose instead that $C \subseteq D^-(t)$. Then $\tau(C) = t^+(C) = tC$, which implies $v_{\tau(C)} = v_{tC}$. In this case $\eta(v_C)$ is disjoint from $D^+(t)$, and thus, $v_C \in R_t \cup F_t$. Therefore, $\psi_\tau(v_C) = \phi_t(v_C)$, which again proves $v_{tC} = \psi_\tau(v_C)$. The case $\tau = t^-$ follows by symmetry.

Now let $\tau_1, \tau_2, \cdots, \tau_N$ be a sequence of folding maps such that $\tau_1\tau_2\cdots\tau_N(C) = C_0$ for every $C \in \mathcal{C}$. By Proposition 6.17 (iii), for each $v \in V(H)$, there exists $C \in \mathcal{C}$ such that $v = v_C$. Then by the correspondence (6.10),

$$\psi_{\tau_1}\psi_{\tau_2}\cdots\psi_{\tau_N}(v_C) = v_{\tau_1\tau_2\cdots\tau_N(C)} = v_{C_0} = W_1. \qquad \square$$

Now we are ready to prove the key correspondence principle:

**Proposition 6.22.** *If $W$ collapses, then the $(S_1, S_2; W)$-reflection graph $H$ collapses.*

*Proof.* Suppose $W$ collapses by $\tau_1, \tau_2, \cdots, \tau_N$. Then by Lemma 6.21, there exists a sequence $\psi_1, \psi_2, \cdots, \psi_N$ of folding maps on $H$ such that $\psi_1\psi_2\cdots\psi_N(wW_i) = W_i$. Thus, every edge $(wW_1, wW_2)$ is mapped to the fixed edge $(W_1, W_2)$ by $\psi_1\psi_2\cdots\psi_N$, which means $H$ collapses. $\square$

To prove Theorem 1.10, it remains to show that a finite reflection group $W$ always collapses, i.e., there exists a sequence $\tau_1, \tau_2, \cdots, \tau_N$ of folding maps such that $\tau_1\tau_2\cdots\tau_N C = C_0$ for all $C \in \mathcal{C}$.

We will now prove that every finite reflection group $W$ collapses by reducing to a purely group-theoretic framework. There is a natural one-to-one correspondence

between open chambers and elements of $W$ that maps $C \in \mathcal{C}$ to the unique element $w \in W$ such that $C = wC_0$. For example, the identity element of $W$ corresponds to the fundamental open chamber $C_0$. For each reflection $t$, by slightly abusing notations, we may define the left and right-folding maps $t^+$ and $t^-$ on $W$ so that

$$(t^+(w))C_0 = t^+(wC_0) \text{ and } (t^-(w))C_0 = t^-(wC_0)$$

hold. Recall that $t^+(C) = tC$ if and only if $\langle x, \alpha \rangle < 0$ for all $x \in C$, where $\alpha$ is the positive root that induces $t$. In other words, $t^+w = tw$ if and only if for every $x \in wC_0$ we have $\langle x, \alpha \rangle < 0$. By Theorem 6.12, this is equivalent to the purely algebraic property $\ell(tw) < \ell(w)$. Thus, one may check that

$$t^+(w) = \begin{cases} tw & \text{if } \ell(tw) < \ell(w) \\ w & \text{if } \ell(tw) \geq \ell(w) \end{cases} \quad \text{and} \quad t^-(w) = \begin{cases} w & \text{if } \ell(tw) \leq \ell(w) \\ tw & \text{if } \ell(tw) > \ell(w) \end{cases} \tag{6.11}$$

hold.[2] We say that a map $\tau : W \to W$ is a *folding map* if it is a left or right-folding map of a reflection in $W$.

With this framework, $W$ collapses if and only if there exists a finite sequence $\tau_1, \tau_2, \cdots, \tau_N$ of folding maps on $W$ such that $\tau_1\tau_2 \cdots \tau_N(w) = e$ for all $w \in W$, where $e$ is the identity element. The existence of such a sequence follows from the next theorem. In fact, it shows something stronger: we may always choose $\tau = s^+$ with $s$ a simple reflection.

**Theorem 6.23.** *Let $W$ be a finite reflection group and let $e$ be the identity of $W$. Then there exists a finite sequence $\tau_1, \tau_2, \cdots, \tau_N$ such that $\tau_1\tau_2 \cdots \tau_N(w) = e$ for all $w \in W$. Moreover, for each $i = 1, 2, \cdots, N$ there is a simple reflection $s$ such that $\tau_i = s^+$.*

*Proof.* Let $S = \{s_1, s_2, \cdots, s_n\}$ be the set of all simple reflections and let $\sigma_i = s_i^+$. We claim that $\ell(\sigma_1\sigma_2 \cdots \sigma_n(w)) < \ell(w)$ unless $w = e$. If the claim is true, then taking $\tau_1, \tau_2, \cdots, \tau_N$ to be the $L$-fold repetition of $\sigma_1, \sigma_2, \cdots, \sigma_n$ concludes the proof, where $L$ is the maximum length of an element in $W$. By (6.11), $\ell(t^+(w)) \leq \ell(w)$ whenever $t \in W$ is a reflection, and moreover, the length remains the same if and only if $t^+(w) = w$. Let $i$ be the largest amongst $1, 2, \cdots, n$ such that $\sigma_i(w) \neq$

[2] In fact, it is impossible to have $\ell(tw) = \ell(w)$. This follows from the simple fact that $w \mapsto \det(w) = (-1)^{\ell(w)}$ is a group homomorphism from $W$ to the multiplicative group $\{\pm 1\}$.

$w$. Such $i$ always exists because, if $w = s'_1 s'_2 \cdots s'_\ell$ is a shortest word for $w$ of length $\ell$, then $\ell(s'_1 w) < \ell(w)$. Then $\sigma_{i+1}\sigma_{i+2}\cdots\sigma_n(w) = w$ but $\sigma_i\sigma_{i+1}\cdots\sigma_n(w) \neq w$, which implies $\ell(\sigma_i\sigma_{i+1}\cdots\sigma_n(w)) < \ell(w)$. Since the length never increases as we apply $\sigma_i$, the claim follows. □

This completes the proof of Theorem 1.10, since we have just shown the existence of a collapsing sequence in $H$. Theorem 1.9, concerning incidence graphs of regular polytopes, follows as a simple application.

*Proof of Theorem 1.9.* Let $\mathcal{P}$ be a regular polytope with symmetry group $W = G(\mathcal{P})$ and let $\Psi$ be a fixed flag in $\mathcal{P}$. By Theorem 6.14, we may assume that the $k$-face $F_k$ in $\Psi$ is fixed under all reflections in $S$ but $s_{k+1}$, i.e., $s_{k+1}$ is the only active mirror for $F_k$ in $S$, and that all other $k$-faces are of the form $wF_k$ for some $w \in W$. Hence, it follows that the $(S_1, S_2; W)$-reflection graph with $S_1 = S \setminus \{s_{k+1}\}$ and $S_2 = S \setminus \{s_{r+1}\}$ is isomorphic to the $(k, r)$-incidence graph of $\mathcal{P}$. □

There are many more examples than those coming from regular polytopes. For instance, the vertex-edge incidence graphs of many quasiregular, i.e., vertex and edge-transitive, polytopes can also be written as $(S_1, S_2; W)$-reflection graphs. Moreover, it is possible to obtain interesting examples by considering the orbits of faces of the same rank:

**Example 6.24.** Let $D_n$ be the $n$-demicube group, i.e., the reflection group of order $2^{n-1}n!$ generated by permutations of the coordinate axes and reflections that flip the signs of *two* coordinate axes. That is, the generating reflection hyperplanes are $x_1 + x_2 = 0$ and $x_i = x_{i+1}$ for $i = 1, 2, \cdots, n-1$.

The action of $D_n$ on the set of vertices of an $n$-dimensional hypercube is not transitive, but there exist two orbits, each of which corresponds to a colour class of the hypercube when considered as a bipartite graph. There is a generating set of simple reflections $S = \{s_1, s_2, \cdots, s_n\}$ for $D_n$ such that the order of $s_1 s_3$ and $s_i s_{i+1}$ for $i \geq 2$ is 3, while all other $s_i s_j$, $i \neq j$, are of order 2. Namely, $s_1$ is the reflection along the hyperplane $x_1 + x_2 = 0$ and $s_i$ is the reflection along $x_{i-1} = x_i$, $i = 2, 3, \cdots, n$. Then one may check that the $n$-dimensional hypercube is isomorphic to the $(S_1, S_2; D_n)$-reflection graph with $S_1 = S \setminus \{s_1\}$

and $S_2 = S \setminus \{s_2\}$. It follows that the hypercube is weakly norming, a fact first proved in [26].

By Theorem 1.9, we already know that the vertex-edge incidence graph of an octahedron, i.e., the 1-subdivision of an octahedron, is weakly norming, but we may also prove that it is norming through a judicious choice of reflection group.

**Example 6.25.** Let $D_3$ be the 3-dimensional demicube group and $S$ the same generating set described in Example 6.24. Take $S_1 = \{s_1, s_2\}$ and $S_2 = \{s_3\}$. Then $(S_1, S_2; D_3)$-reflection graph is isomorphic to the 1-subdivision of an octahedron, as the latter is the face-edge incidence graph of the cube. Therefore, the 1-subdivision of the octahedron is norming.

Let us mention another example of a norming graph. In Chapter 5, a graph is said to be a $K_{2,t}$-replacement of $H$ if each edge of $H$ is replaced with a copy of $K_{2,t}$ by identifying the two vertices of the edge with the two vertices on the smaller side of $K_{2,t}$.

**Example 6.26.** Let $B_3$ be the cube group, that is, the symmetry group of the cube. Then there is a generating set $S = \{s_1, s_2, s_3\}$ such that $s_1s_2$, $s_2s_3$, and $s_1s_3$ are of orders 4, 3, and 2, respectively. Let $S_1 = \{s_1, s_2\}$ and $S_2 = \{s_3\}$. Then one may check that the $(S_1, S_2; B_3)$-reflection graph is isomorphic to the $K_{2,2}$-replacement of the octahedron graph, and thus is norming.

## 6.5 Generalisations and applications

### 6.5.1 Hypergraph norms

As noted in Hatami's PhD thesis [25], the concepts of norming and weakly norming graphs generalise in the obvious way to hypergraphs, with Gowers' octahedral norms [21, 22] serving as standard examples. In this short subsection, we discuss the appropriate generalisation of Theorem 1.10.

We say that an automorphism $\phi$ of a $k$-graph $H$ is a *cut involution* of $H$ if the following hold:

(i) $\phi$ is an involution, i.e., $\phi = \phi^{-1}$;

(ii) the fixed point set $F_\phi = \{v \in V(H) : \phi(v) = v\}$ is a vertex cut of $H$;

(iii) there exists a partition $L_\phi \cup R_\phi$ of $V(H) \setminus F_\phi$ such that $\phi(L_\phi) = R_\phi$ and no edge intersects both $L_\phi$ and $R_\phi$.

Given a finite reflection group $W$, the set of all simple reflections $S$, and subsets $S_1, S_2, \cdots, S_k$ of $S$, let the $(S_1, \cdots, S_k; W)$-reflection hypergraph be the $k$-partite $k$-graph whose parts are the cosets of the subgroup $W_i$ generated by $S_i$ for each $i = 1, \cdots, k$, with an edge for every $k$-tuple of the form $(wW_1, wW_2, \cdots, wW_k)$ with $w \in W$. We say that a $k$-graph is a *reflection hypergraph* if it is isomorphic to the $(S_1, \cdots, S_k; W)$-reflection hypergraph for a suitable choice of parameters. We have the following hypergraph generalisation of Theorem 1.10:

**Theorem 6.27.** *A $k$-partite $k$-graph $H$ is weakly norming whenever it is a reflection hypergraph. Moreover, if there is no mirror containing a hyperedge, i.e., $\bigcap_{i=1}^k S_i = \emptyset$, then $H$ is norming.*

The proof of Theorem 6.27 is a straightforward generalisation of arguments used to prove Theorem 1.10. Firstly, we generalise the analytic part, namely Theorem 6.5. As folding maps are vertex maps defined by cut involutions, the definitions of folding maps, percolating sequences, and collapsing sequences generalise in a straightforward way, once we use the generalised notion of cut involutions. To this end, it is essential to generalise Lemma 6.1 as follows:

**Lemma 6.28.** *Let $H$ be an $m$-edge $k$-graph and let $\mathcal{F} = (f_1, f_2, \cdots, f_m)$ be a sequence of non-negative bounded measurable symmetric functions on $[0,1]^k$. For each edge-colouring $\chi : E(H) \to \{1, 2, \cdots, m\}$ and a cut involution $\phi$,*

$$\langle \mathcal{F}; \chi \rangle_H \leq \left\langle \mathcal{F}; \chi \circ \phi^+ \right\rangle_H^{1/2} \left\langle \mathcal{F}; \chi \circ \phi^- \right\rangle_H^{1/2}.$$

*Proof.* Let $E_F := E(H[F])$, $E_L := E(H[L \cup F]) \setminus E_F$, and $E_R := E(H[R \cup F]) \setminus E_F$. Again, $E_F$, $E_L$, and $E_R$ partition $E(H)$, and moreover, $E_L$ and $E_R$ are mapped

to each other by $\phi$. Let $\mathbf{x}_F = (x_v)_{v\in F}$ and define

$$g_L(\mathbf{x}_F) := \int \prod_{e\in E_L} f_{\chi(e)}((x_i)_{i\in e}) \prod_{k\in L} dx_k,$$
$$g_R(\mathbf{x}_F) := \int \prod_{e\in E_R} f_{\chi(e)}((x_i)_{i\in e}) \prod_{k\in R} dx_k, \text{ and}$$
$$g_F(\mathbf{x}_F) := \prod_{e\in E_F} f_{\chi(e)}((x_i)_{i\in e}),$$

where $g_F := 1$ if $E_F$ is empty. With these notations, the proof follows from the same argument as in Lemma 6.1. □

Now we may replace each term by the generalised notion in the statement of Theorem 6.4 to obtain its hypergraph generalisation. Hence, Theorem 6.27 reduces to the following:

**Theorem 6.29.** *If a hypergraph $H$ collapses under its cut involution group, then $H$ is weakly norming.*

The proof of Theorem 6.29 is identical to that of Theorem 6.5 modulo replacing graphs by $k$-graphs. The definition of the Euclidean embedding $\eta$ remains the same, i.e., for $1 \leq i \leq k$,

$$\eta(wW_i) := wC(S_i).$$

The following generalises Proposition 6.17:

**Proposition 6.30.** *Let $H$ be the $(S_1, \cdots, S_k; W)$-reflection graph. Then the Euclidean embedding $\eta$ of $H$ has the following properties:*

*(i) for two vertices $a$ and $a'$ on the same colour class of the $k$-partition of $H$, $\eta(a) = \eta(a')$ if and only if $a = a'$;*

*(ii) every closed chamber $\overline{C}$ contains exactly one image of a vertex under $\eta$ in each colour class. Moreover, if $\eta(v_1), \eta(v_2), \cdots, \eta(v_k)$ are subsets of $\overline{C}$, then $(v_1, v_2, \cdots, v_k) \in E(H)$;*

*(iii) for every $(v_1, v_2, \cdots, v_k) \in E(H)$, there exists a closed chamber $\overline{C}$ containing all $\eta(v_1), \eta(v_2), \cdots, \eta(v_k)$.*

*Proof.* (i) The proof is identical to that of Proposition 6.17 (i).

(ii) By Proposition 6.17 (ii), $W_1, W_2, \cdots, W_k$ are the only vertices of $H$ that are mapped to subsets of $\overline{C}_0$. Moreover, each closed chamber $w\overline{C}_0$ contains exactly one $k$-tuple of cones $(\eta(wW_1), \eta(wW_2), \cdots, \eta(wW_k))$ that are images of a $k$-edge $(wW_1, wW_2, \cdots, wW_k)$, since the action of $W$ is simply transitive on $\mathcal{C}$.

(iii) Since $(v_1, v_2, \cdots, v_k) \in E(H)$, there exists $w \in W$ such that $v_i = wW_i$, $i = 1, 2, \cdots, k$. Then $\eta(v_1), \eta(v_2), \cdots, \eta(v_k)$ must be contained in $w\overline{C}_0$, where $\overline{C}_0$ is the fundamental closed chamber. □

The proof of Corollary 6.18 adapts to show that every every reflection $t \in W$ is a cut involution of the $(S_1, \cdots, S_k; W)$-reflection hypergraph. The key point is that, once we define $L_t = \eta^{-1}(D^+(t))$ and $R_t = \eta^{-1}(D^-(t))$ for each $t \in T$, there is no edge that intersects both $L_t$ and $R_t$.

**Corollary 6.31.** *Let $H$ be the $(S_1, \cdots, S_k; W)$-reflection hypergraph. Then every reflection $t \in W$ is a cut involution of $H$.* □

Finally, we generalise Proposition 6.22.

**Proposition 6.32.** *If $W$ collapses, then the $(S_1, \cdots, S_k; W)$-reflection hypergraph $H$ collapses.*

*Proof.* Suppose $W$ collapses by $\tau_1, \tau_2, \cdots, \tau_N$. Then by Lemma 6.21, there exists a sequence $\psi_1, \psi_2, \cdots, \psi_N$ of folding maps on $H$ such that $\psi_1\psi_2\cdots\psi_N(wW_i) = W_i$. Thus, every edge $(wW_1, wW_2, \cdots, wW_k)$ is mapped to the fixed edge $(W_1, \cdots, W_k)$ by $\psi_1\psi_2\cdots\psi_N$, which means $H$ collapses. □

Now Theorem 6.23 concludes the proof of Theorem 6.27.

With this framework, we may easily recover Gowers' octahedral norms [21, 22].

**Example 6.33.** Let $W = W_1 \times W_2 \times \cdots \times W_k$, where each $W_i$ is a reflection group generated by a single reflection $s_i$. Then the set $S$ of simple reflections consists of $k$ orthogonal reflections. If we let $S_i = S \setminus \{s_i\}$ and $H_k$ be the $(S_1, \cdots, S_k; W)$-reflection hypergraph, then Theorem 6.27 implies that $H_k$, consisting of the $(k-1)$-faces of the $k$-dimensional octahedron, is a norming $k$-graph.

Another example of a hypergraph norm comes from the work by Conlon, Hàn, Person, and Schacht [10] on weak quasirandomness.

**Example 6.34.** In [10], a $k$-graph $M_k$ is constructed recursively as follows. Given a $k$-partite $k$-graph $M$ on $A_1 \cup \cdots \cup A_k$, we write $\mathrm{db}_i(M)$ for the $k$-graph obtained by gluing two vertex-disjoint copies of $M$ so that the corresponding vertices in each copy of $A_i$ are identified. Note that $\mathrm{db}_i(M)$ contains $2|V(M)| - |A_i|$ vertices and $2|E(M)|$ edges. Starting from the graph $H_0$ with a single edge, define

$$M_k := \mathrm{db}_k(\mathrm{db}_{k-1}(\cdots \mathrm{db}_1(H_0) \cdots)).$$

In other words, $M_k$ is the $k$-graph on the edge set of the $k$-dimensional hypercube, where each $k$-edge is a collection of edges (in the hypercube) sharing a vertex. In particular, $M_2$ is the 4-cycle. Letting $S_i = \{s_i\}$ and using the same reflection group $W$ as in Example 6.33, Theorem 6.27 implies that the $(S_1, \cdots, S_k; W)$-reflection hypergraph is norming. Since this hypergraph is isomorphic to $M_k$, we see that $M_k$ defines a semi-norm.

The examples above are very natural: both of them use the simplest reflection group with $k$ generators and the resulting Cauchy–Schwarz trees are very symmetric and easy to analyse. Our new framework gives a much larger class of hypergraph norms, each of which defines a certain notion of quasirandomness. We will discuss the relations between these notions in the next subsection.

### 6.5.2 Domination between norms

In this subsection, we will be interested in relations between (hyper)graph norms. To begin, we will study the question of determining whether the (absolute) $H$-norm dominates another (absolute) $J$-norm. That is, for any bounded measurable function $f : [0,1]^2 \to \mathbb{R}$, we would like to know if

$$\|f\|_J \le \|f\|_H \quad \Big(\text{or } \|f\|_{r(J)} \le \|f\|_{r(H)}\Big) \tag{6.12}$$

holds. This question remains valid even if $H$ and $J$ are not (weakly) norming, but weakly norming graphs can be regarded as local maxima for such comparisons, as shown in (6.4).

An immediate consequence of this inequality is Sidorenko's conjecture, which essentially states that the absolute $H$-norm dominates the single-edge norm whenever $H$ is a bipartite graph as already seen in (6.3). Hence, every weakly norming graph $H$ has Sidorenko's property. Similar arguments show that if $H$ is weakly norming and contains a cycle, then it is forcing and hence gives a partial answer to Question 1.2 (or the so-called forcing conjecture). We refer the interested reader to [9], where the forcing conjecture is formalised, for further information about this conjecture and its relationship with graph norms.

We remark that one (hyper)graph norm may dominate another even when the first (hyper)graph does not contain the second. For example, if $m \geq n$, then

$$\|f\|_{C_{2m}} \leq \|f\|_{C_{2n}},$$

which immediately follows from Proposition 2.1 (also see (4.8)). Note that the cut involution groups of different even cycles are always non-isomorphic. In general, it is hard to compare (hyper)graph norms coming from non-isomorphic reflection groups, but we can say something if the two (hyper)graphs have isomorphic cut involution groups of a certain type.

**Proposition 6.35.** *Let $W = W_1 \times W_2 \times \cdots \times W_n$, where each $W_i$ is a reflection group generated by a single reflection $s_i$, and let $S = \{s_1, \cdots, s_n\}$ be the set of simple reflections in $W$. Suppose $S_1, \cdots, S_k$ and $S'_1, \cdots, S'_k$ are subsets of $S$ such that $S'_i \subset S_i$ for each $i = 1, 2, \cdots, k$. Let $H$ and $H'$ be the $(S_1, \cdots, S_k; W)$-reflection hypergraph and the $(S'_1, \cdots, S'_k; W)$-reflection hypergraph, respectively. Then both $H$ and $H'$ are norming and the $H$-norm dominates the $H'$-norm.*

Our proof is quite heavy on notation, so we firstly give an example that conveys the rough idea of the proof.

**Example 6.36.** Consider the special case of Proposition 6.35 where $n = 3$, $S_i = S \setminus \{s_i\}$ for $i = 1, 2, 3$, $S'_2 = S_2$, $S'_3 = S_3$, and $S'_1 = \{s_3\}$. We may assume that $s_1$, $s_2$, and $s_3$ are reflections in the planes $x = 0$, $y = 0$, and $z = 0$, respectively, in $\mathbb{R}^3$. We use the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to make the fundamental open chamber the first octant. Then $H$ is the octahedral 3-graph on $\{X_0, X_1, Y_0, Y_1, Z_0, Z_1\}$ with

edges $\{X_iY_jZ_k : i,j,k = 0,1\}$, where

$$X_0 = \{(x,0,0) : x > 0\},\ X_1 = \{(x,0,0) : x < 0\},$$
$$Y_0 = \{(0,y,0) : y > 0\},\ Y_1 = \{(0,y,0) : y < 0\},$$
$$Z_0 = \{(0,0,z) : z > 0\},\ Z_1 = \{(0,0,z) : z < 0\}.$$

As noted in Example 6.33, the $H$-norm is the Gowers' octahedral norm for 3-graphs. On the other hand, $H'$ is the 3-graph on eight vertices

$$\{X_{00}, X_{01}, X_{10}, X_{11}, Y_0, Y_1, Z_0, Z_1\}$$

with eight edges $\{X_{ij}Y_iZ_k : i,j,k = 0,1\}$. Here $Y_i, Z_j$ are exactly the same as in $H$ since $S_2 = S_2'$ and $S_3 = S_3'$, and

$$X_{00} = \{(x,y,0) : x > 0, y > 0\},\ X_{10} = \{(x,y,0) : x < 0, y > 0\},$$
$$X_{01} = \{(x,y,0) : x > 0, y < 0\},\ X_{11} = \{(x,y,0) : x < 0, y < 0\}.$$

Observe that the induced subgraphs of $H$ on $\{X_0, X_1, Y_i, Z_0, Z_1\}$, for $i = 0,1$, and the induced subgraphs of $H'$ on $\{X_{0j}, X_{1j}, Y_j, Z_0, Z_1\}$, for $j = 0,1$, are all isomorphic to the kite-shaped 3-graph with 4 edges, say $K$. Moreover, $H$ is obtained by gluing two vertex-disjoint copies of $K$, identifying the copies of $X_0$, $X_1$, $Z_0$, and $Z_1$ in each graph. We can build $H'$ similarly but now we only identify the copies of $Z_0$ and $Z_1$. Therefore, setting

$$g(x_0,x_1,z_0,z_1) = \int \prod_{i,j=0,1} f(x_i,y,z_j)dy,$$

whose average is the $4^{th}$ power of the $K$-norm, we have

$$\|f\|_H = \int g^2 dx_0dx_1dz_0dz_1 \text{ and } \|f\|_{H'} = \int \left(\int g dx_0dx_1\right)^2 dz_0dz_1.$$

Thus, the Cauchy–Schwarz inequality implies that $\|f\|_H \geq \|f\|_{H'}$. We shall generalise this idea later to prove Proposition 6.35.

For the proof of Proposition 6.35, we first describe the structure of the given hypergraphs $H$ and $H'$. Let $W \subset \mathbf{GL}(n,\mathbb{R})$ be a reflection group with simple

reflections $S = \{s_1, s_2, \cdots, s_n\}$ such that each $s_i$ is the reflection along the hyperplane

$$H_i := \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n : x_i = 0\},$$

i.e., $s_i$ flips the sign of the $i$-th coordinate. Then the $(S_1, \cdots, S_k; W)$-reflection hypergraph $H$ is the $k$-partite graph with $k$-partition $\mathcal{A}_1 \cup \cdots \cup \mathcal{A}_k$, where $\mathcal{A}_i$ consists of the $2^{n-|S_i|}$ vertices of the form

$$v_i(\xi) := \{(x_1, \cdots, x_n) : \xi_j x_j > 0 \text{ for all } j \notin S_i,\ x_j = 0 \text{ for all } j \in S_i\},$$

where $\xi = (\xi_1, \cdots, \xi_n) \in \{\pm 1\}^n$. Here each $\xi \in \{\pm 1\}^n$ represents the closed chamber

$$\{(x_1, \cdots, x_n) : \xi_j x_j \geq 0 \text{ for all } j = 1, 2, \cdots, n\},$$

so a $k$-tuple of vertices in $\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_k$ is an edge of $H$ if and only if it can be written as $(v_1(\xi), v_2(\xi), \cdots, v_k(\xi))$ with the same $\xi \in \{\pm 1\}^n$. Let $H^+$ be the subgraph of $H$ induced on cones in the half-space $\{(x_1, \cdots, x_n) : x_n \geq 0\}$ induced by the hyperplane $H_n$. Using the fact that $s_n$ is a cut involution of $H$, we have

$$\|f\|_H^{|E(H)|} = \int \left( \int g \; dy_{V(H^+) \setminus F(H)} \right)^2 dy_{F(H)}, \tag{6.13}$$

where $g = \prod_{(u_1, \cdots, u_k) \in E(H^+)} f(y_{u_1}, \cdots, y_{u_k})$, $y_I = (y_i)_{i \in I}$ for $I \subset V(H)$, and $F(H)$ is the set of vertices $\{v_j(\xi) \in V(H) : s_n \in S_j, \xi \in \{\pm 1\}^n\}$ that are fixed by $s_n$. Furthermore, $H^+$ is isomorphic to the $(S_1^+, \cdots, S_k^+; W)$-reflection hypergraph with $S_i^+ = S_i \cup \{s_n\}$ for all $i$, since adding $s_n$ to all $S_i$ makes the sign of $x_n$ fixed.

*Proof of Proposition 6.35.* If $\bigcap_{i=1}^k S_i$ is empty, then by Theorem 6.27 $H$ is norming. Suppose instead that there exists a simple reflection $s \in \bigcap_{i=1}^k S_i$. Let $S_i^- := S_i \setminus \{s\}$ and let $W_i^-$ be the subgroup generated by $S_i^-$. Then the $(S_1^-, \cdots, S_k^-; W)$-reflection hypergraph $H^-$consists of two vertex-disjoint copies of $H$ mapped to each other by $s$. More precisely, the homomorphisms $\phi, \psi : V(H) \to V(H^-)$ given by $\phi(wW_i) = wW_i^-$ and $\psi(wW_i) = swW_i^-$, respectively, where $W_i^-$ denotes the subgroup generated by $S_i^-$, map $H$ to two isomorphic connected components of $V(H^-)$. Then the $H^-$-norm takes exactly the same value as the $H$-norm and, hence, by induction on $|\bigcap_{i=1}^k S_i|$, we are done.

To prove that the $H$-norm dominates the $H'$-norm, we may assume $S_i' = S_i \setminus \{s_n\}$, $s_n \in S_i$, and $S_j' = S_j$ for $j \neq i$. Both $H$ and $H'$ have the same induced subgraph on cones in the half-space $\{(x_1, \cdots, x_n) : x_n \geq 0\}$, because it must be isomorphic to the $(S_1^+, \cdots, S_k^+; W)$-reflection hypergraph with $S_j^+ = S_j \cup \{s_n\}$ for all $j$. When writing $\|f\|_{H'}$ as in (6.13), the fixed point set $F(H')$ under $s_n$ becomes

$$F(H') = \{v_j(\xi) \in V(H') : s_n \in S_j', \xi \in \{\pm 1\}^n\} = F(H) \setminus \mathcal{A}_i,$$

since $s_n$ is removed from $S_i$ . Therefore, the $H'$-norm can be expressed, with the same $g$ as in (6.13), by

$$\|f\|_{H'}^{|E(H')|} = \int \left( \int g \, dy_{V(H^+) \setminus F(H)} dy_{\mathcal{A}_i} \right)^2 dy_{F(H) \setminus \mathcal{A}_i}, \tag{6.14}$$

and the Cauchy–Schwarz inequality implies that this is less than or equal to (6.13). □

Another way to compare (hyper)graph norms is to ask that they be polynomially related, in the sense that if one norm is large, say at least $c$, then the other norm is at least $c^k$ for some appropriate $k$. Formally, we say that a (semi-)norm $\|\cdot\|$ *polynomially dominates* another (semi-)norm $\|\cdot\|_0$ if, for any $|f| \leq 1$,

$$\|f\| \leq c \text{ implies } \|f\|_0 \leq c^\delta,$$

where $\delta > 0$ is a constant independent from $f$. It is also possible for two norms to polynomially dominate each other, so in this case we say that the two norms are *polynomially equivalent*. For the rest of this subsection, we shall state our results only for norming graphs, but analogous statements for weakly norming graphs hold if we replace $f$ by $|f|$ and graph norms by absolute graph norms. Following Gowers' approach [20, 21, 22] to quasirandomness, but borrowing notation from graph limit theory [36, 38], we define the cut-norm $\|f\|_\square$ for a bounded measurable function $f$ on $[0,1]^2$ by

$$\|f\|_\square := \sup_{u,v} \left| \int f(x,y)u(x)v(y)dxdy \right|,$$

where the supremum is taken over all measurable $u, v : [0,1] \to [-1,1]$. By applying inequality (6.2) with $C_4$, we see that

$$\int f(x,y)u(x)v(y)dxdy \leq \|f\|_{C_4} \|u\|_{C_4} \|v\|_{C_4} \leq \|f\|_{C_4} \tag{6.15}$$

holds for all measurable $u, v : [0,1] \to [-1,1]$, and hence $\|f\|_\square \leq \|f\|_{C_4}$. Conversely, if

$$\|f\|_{C_4}^4 = \int f(x,y)f(x,y')f(x',y)f(x',y') > c^4,$$

then there exist some $x^*, y^* \in [0,1]$ such that

$$\left| \int f(x,y)f(x,y^*)f(x^*,y)f(x^*,y^*)dxdy \right| \geq c^4. \tag{6.16}$$

Taking $u(x) = f(x^*,y^*)f(x,y^*)$ and $v(y) = f(x^*,y)$ gives $\|f\|_\square \geq c^4$, so the $C_4$-norm and the cut-norm are polynomially equivalent. Similarly, using (6.2) to deduce (6.15) and averaging to obtain (6.16), it is easy to see that any $H$-norm is polynomially equivalent to the cut-norm whenever $H$ is norming and contains a cycle. Hence, if two norming graphs $H$ and $J$ contain cycles, then the $H$-norm and the $J$-norm are always polynomially equivalent. Conversely, suppose $H$ is a norming graph isomorphic to a tree. Then $H$ must be bi-regular by Theorem 2.10 in [26] and, hence, is isomorphic to a star $K_{1,t}$. By taking $f$ to be the balanced function of a non-quasirandom but regular graph, one may obtain an example with small $K_{1,t}$-norm but large cut-norm. Hence, star-norms are polynomially dominated by the cut-norm, but not vice versa.

Our aim now is to decide whether two different $H$-norms, where each $H$ is an $(S_1, \cdots, S_k; W)$-reflection hypergraph, are polynomially equivalent. The result below says that we may 'forget' the angles between the reflection hyperplanes of a reflection group $W$ if $H$ is an $(S_1, \cdots, S_k; W)$-reflection hypergraph and still obtain a polynomially equivalent norm:

**Proposition 6.37.** *Let $W$ and $W'$ be two reflection groups with the same number of simple reflections, i.e., $S = \{s_1, \cdots, s_n\}$ and $S' = \{s'_1, \cdots, s'_n\}$ are the set of simple reflections of $W$ and $W'$, respectively. Suppose, for $S_1, \cdots, S_k \subset S$ and $S'_1, \cdots, S'_k \subset S'$, that the indices covered by $S_i$ and $S'_i$ are always the same. That is, $\{j : s_j \in S_i\} = \{j : s'_j \in S'_i\}$ for each $i = 1, 2, \cdots, k$. Then the $(S_1, \cdots, S_k; W)$-reflection hypergraph $H$ and the $(S'_1 \cdots, S'_k; W')$-reflection hypergraph $H'$ give polynomially equivalent hypergraph semi-norms, provided that both hypergraphs are norming.*

For example, the proposition above allows us to prove that if a reflection hypergraph $H$ consists of 'tight triples', then the $H$-norm is polynomially equivalent to Gowers' octahedral norm for 3-graphs.

**Example 6.38.** In Example 6.16, the reflection group $W$ is isomorphic to the symmetric group on 4 elements and the set of simple reflections $S = \{t_{12}, t_{23}, t_{34}\}$, where $t_{ij}$ is a reflection swapping the $i$-th and $j$-th coordinates. Let $S_1 = S \setminus \{t_{12}\}$, $S_2 = S \setminus \{t_{23}\}$, $S_3 = S \setminus \{t_{34}\}$, and $H$ be the $(S_1, S_2, S_3; W)$-reflection hypergraph. Observe that $H$ is a tripartite 3-graph isomorphic to the vertex-edge-face incidence 3-graph of a tetrahedron, i.e., a vertex, an edge, and a face form an edge if they can be extended to a flag. The fact that $S_1 \cap S_2 \cap S_3 = \emptyset$ implies that $H$ is norming and, by Proposition 6.37 and Example 6.36, the $H$-norm is polynomially equivalent to Gowers' octahedral norm for 3-graphs.

The key to proving Proposition 6.37 is to show a polynomial equivalence between hypergraph norms and suitably generalised cut-norms. For hypergraphs, Gowers [21, 22] proved that there is a polynomial equivalence between the octahedral norms defined in Example 6.33 and certain generalised cut-norms. Later, Conlon, Hàn, Person, and Schacht [10] showed that the norms discussed in Example 6.34 exist and are polynomially equivalent to certain weaker cut-norms. More recently, Reiher, Rödl, and Schacht [46] and Towsner [57] proposed further cut-norms lying between these two extremes. As we shall see, the result below gives a suitable $H$-norm which is polynomially equivalent to each of these cut-norms.

Let $\mathcal{M} = (M_1, M_2, \cdots, M_n)$ be an $n$-tuple of subsets of $[k] = \{1, 2, \cdots, k\}$ and define the hypergraph cut-norm $\|\cdot\|_{\square,\mathcal{M}}$ with respect to $\mathcal{M}$ by

$$\|f\|_{\square,\mathcal{M}} := \sup_{u_1,\cdots,u_n} \left| \int f(x_1, \cdots, x_k) \prod_{i=1}^{n} u_i(x_{M_i}) \right|,$$

where $f$ is a bounded measurable function on $[0,1]^k$, $x_I$ for $I \subset [k]$ is the vector $(x_i)_{i \in I}$ of variables with indices in $I$, and the supremum is taken over all measurable functions $u_1, \cdots, u_k$ taking values between $[-1, 1]$. For example, $n = 2$, $M_1 = \{1\}$, and $M_2 = \{2\}$ is the graph cut-norm. The cut-norms described by Gowers [21, 22] correspond to the case when $n = k$ and $M_i = [k] \setminus \{i\}$. In particular, for $n = 3$,

we have

$$\|f\|_{\square,\mathcal{M}} = \sup_{u,v,w:[0,1]^2\to[-1,1]} \left| \int f(x,y,z)u(x,y)v(y,z)w(z,x) \right|.$$

The lemma below states that whenever $H$ is the $(S_1, \cdots, S_k; W)$-reflection hypergraph, the $H$-norm is polynomially equivalent to the generalised cut-norm with respect to a suitably chosen $\mathcal{M}$.

**Lemma 6.39.** *Let $W$ be a reflection group with simple reflections $S = \{s_1, \cdots, s_n\}$ and let $S_1, \cdots, S_k$ be subsets of $S$. Suppose $\mathcal{M} = (M_1, \cdots, M_n)$ is the $n$-tuple of subsets of $[k]$ such that $M_i = \{j \in [k] : s_i \in S_j\}$ for $i = 1, \cdots, n$. Then, whenever $H$ is an $(S_1, \cdots, S_k; W)$-reflection hypergraph that is norming, the $H$-norm is polynomially equivalent to the cut-norm $\|\cdot\|_{\square,\mathcal{M}}$.*

*Proof.* Let $V_1V_2 \cdots V_k$ be the edge of the $(S_1, S_2, \cdots, S_k; W)$-reflection hypergraph $H$ contained in the fundamental closed chamber. Observe that $M_i$ is the collection of indices $j$ such that $V_j$ is fixed under the simple reflection $s_i$, since each collection of inactive mirrors $S_j$ defines the cone $V_j$. It follows that there is another edge of $H$, formed by reflecting the edge $V_1V_2 \ldots V_k$ in $s_i$, that contains the set $\{V_j : j \in M_i\}$. The remainder of the proof is now similar to the $C_4$ case discussed earlier.

To see that $\|f\|_{\square,\mathcal{M}} \leq \|f\|_H$, observe that, since $H$ is norming,

$$\int f(x_1, \cdots, x_k) \prod_{i=1}^{n} u_i(x_{M_i}) \leq \|f\|_H \prod_{i=1}^{n} \|u_i\|_H \leq \|f\|_H$$

whenever $u_i : [0,1]^{|M_i|} \to [-1,1]$. To prove the converse, note that if $\|f\|_H > c$, then we may assign values to the variables corresponding to the vertices other than $V_1, V_2, \cdots, V_k$ to get

$$\left| \int f(x_1, \cdots, x_k) \prod_{i=1}^{n} u_i(x_{M_i}) \right| \geq c^{|E(H)|},$$

as desired. □

Since the choice of $\mathcal{M}$ is independent of the group structure of $W$, Proposition 6.37 follows as a corollary. Note that each $M_i$ corresponds to the $i$-th column vector of the $k \times n$ incidence matrix between $k$ sets $S_1, S_2, \cdots, S_k$ and the simple

reflections $s_1, s_2, \cdots, s_n$. Thus, given a cut-norm $\|\cdot\|_{\square,\mathcal{M}}$ with an ordered collection $\mathcal{M}$ of index subsets of $k$, we may always construct a $k$-graph norm that is polynomially equivalent to it. To give some examples, we will now describe hypergraph norms that are polynomially equivalent to the weak cut-norms introduced in [46].

**Example 6.40.** In [46], two different cut-norms for 3-graphs were studied. For 3-graphs, we always have $k = 3$ and if we let $n = 2$, $M_1 = \{1\}$, and $M_2 = \{2, 3\}$, we get the cut-norm

$$\|f\|_{\square,\mathcal{M}} = \sup_{u,v,w:[0,1]^2\to[-1,1]} \left| \int f(x, y, z)u(x)v(y, z) \right|.$$

To ensure that $M_i = \{j : s_i \in S_j\}$ holds, let $S_1 = \{s_1\}$ and $S_2 = S_3 = \{s_2\}$. Considering the simplest reflection group $W = \langle s_1\rangle \times \langle s_2\rangle$ generated by $s_1$ and $s_2$, the $(S_1, S_2, S_3; W)$-reflection hypergraph is isomorphic to the 3-graph $H$ with

$$V(H) = \{x_0, x_1, y_0, y_1, z_0, z_1\} \text{ and } E(H) = \{x_i y_j z_j : i, j = 0, 1\}.$$

The second cut-norm is the case when $n = 2$, $M_1 = \{1, 2\}$, and $M_2 = \{2, 3\}$, i.e.,

$$\|f\|_{\square,\mathcal{M}} = \sup_{u,v,w:[0,1]^2\to[-1,1]} \left| \int f(x, y, z)u(x, y)v(y, z) \right|.$$

Then we have $S_1 = \{s_1\}, S_2 = \{s_1, s_2\}$, and $S_3 = \{s_2\}$. With the same group $W = \langle s_1\rangle \times \langle s_2\rangle$, the $(S_1, S_2, S_3; W)$-reflection hypergraph is isomorphic to the kite-shaped 3-graph with 5 vertices denoted by $K$ in Example 6.36.

By combining Proposition 6.37 and Proposition 6.35, we also have the following weak domination result between hypergraph norms with the same cut involution group.

**Corollary 6.41.** *Let $W$ be a reflection group with $S = \{s_1, \cdots, s_n\}$ the set of simple reflections and let $S_1, \cdots, S_k, S'_1, \cdots, S'_k$ be subsets of $S$ such that $S'_i \subset S_i$ for each $i = 1, \cdots, k$. Suppose $H$ and $H'$ are the $(S_1, \cdots, S_k; W)$-reflection hypergraph and the $(S'_1, \cdots, S'_k; W)$-reflection hypergraph, respectively, and suppose that they are both norming. Then the $H$-norm polynomially dominates the $H'$-norm.*

## 6.6 Concluding remarks

**Characterising weakly norming graphs.** Although Theorem 1.10 gives a fairly large class of weakly norming graphs, it is still an open problem to characterise them all. For our arguments to work on a particular graph $H$, it was necessary that $H$ be edge-transitive under its cut involution group. We also insisted on the existence of a percolating sequence in $H$, but we suspect that this condition may not be needed.

**Conjecture 6.42.** *A bipartite graph $H$ is weakly norming if it is edge-transitive under its cut involution group.*

Underlying this conjecture is a further suspicion that reflection graphs may be the only graphs which are edge-transitive under their cut involution groups. Were this indeed the case, Conjecture 6.42 would follow immediately from Theorem 1.10.

In [26], Hatami showed that weakly norming graphs must be balanced and bi-regular and that the class of weakly norming graph is closed under taking tensor products. There are analogues of these results for reflection graphs: any reflection graph must be bi-regular since it is vertex-transitive on each side, while the tensor product of two reflection graphs is again a reflection graph. It is not too difficult to prove that the same properties hold for graphs that are edge-transitive under the cut involution group. These coincidences suggest that the converse to Conjecture 6.42 may also be true, though this is significantly more tentative than Conjecture 6.42.

**Conjecture 6.43.** *A bipartite graph $H$ is weakly norming only if it is edge-transitive under its cut involution group.*

It would already be very interesting to show that weakly norming graphs are necessarily edge-transitive.

**Norming graphs versus weakly norming graphs.** When can we guarantee that a weakly norming graph is also norming? Theorem 1.11 gives a partial answer to this natural question. However, we again suspect, in analogy with Conjectures 6.42 and 6.43, that edge-transitivity under the stable involution group is a necessary and sufficient property.

**Conjecture 6.44.** *A bipartite graph $H$ is norming if and only if it is edge-transitive under its stable involution group.*

This is reminiscent of another conjecture, proposed in [5], about determining those graphs $H$ for which $t_H(f)$ is always positive. It might be interesting to investigate the connection.

**Conjecture 6.45** (Positive graph conjecture [5])**.** *$t_H(f)$ is positive for all $f : [0,1]^2 \to \mathbb{R}$ if and only if there is a stable involution of $H$.*

**Complex-valued functions.** Our results on norming (hyper)graphs, Theorem 1.11 and Theorem 6.27, can also be used to define (hyper)graph norms for complex-valued bounded measurable symmetric functions. Suppose $H$ is the $(S_1, \cdots, S_k; W)$-reflection hypergraph, where $W$ is a reflection group and $\bigcap_{i=1}^k S_i = \emptyset$. Note that each reflection in $W$ acts as a stable involution and there exists a one-to-one correspondence between the edges of $H$ and the set $\mathcal{C}$ of open chambers, since each edge is contained in exactly one closed chamber. By the one-to-one correspondence between $W$ and $\mathcal{C}$, there also exists a one-to-one correspondence $\xi : W \to E(H)$ between $W$ and $E(H)$. Let $W_+ := \{w \in W : \ell(w) \text{ is even}\}$ and $W_- := \{w \in W : \ell(w) \text{ is odd}\}$. We may write $W_+$ and $W_-$ as the inverse images of 1 and $-1$ under the group homomorphism $w \mapsto \det(w) = (-1)^{\ell(w)}$ from $W$ to the multiplicative group $\{\pm 1\}$. It follows that $W_+$ and $W_-$ partition $W$ into parts of equal size and each reflection $t \in W$ maps $w \in W_+$ to $tw \in W_-$ and vice versa. For a complex-valued bounded measurable symmetric function $f$ on $[0,1]^k$, define

$$\|f\|_H := \left| \int \prod_{e \in \xi(W_+)} f(x_e) \prod_{e' \in \xi(W_-)} \overline{f(x_{e'})} \right|^{1/|E(H)|}, \tag{6.17}$$

where $x_e = (x_{u_1}, x_{u_2}, \cdots, x_{u_k})$ for $e = (u_1, u_2, \cdots, u_k)$. This is well-defined, even without taking the absolute value of the integral above, since

$$\int \prod_{e \in \xi(W_+)} f(x_e) \prod_{e' \in \xi(W_-)} \overline{f(x_{e'})} = \int g\bar{g}, \tag{6.18}$$

where

$$g(x_{F_\phi}) = \int \prod_{e \in \xi(W_+) \cap (L_\phi \cup F_\phi)^k} f(x_e) \prod_{e' \in \xi(W_-) \cap (L_\phi \cup F_\phi)^k} \overline{f(x_{e'})}\, dx_{L_\phi}$$

for a stable involution $\phi$ that corresponds to a reflection in $W$. When $W = \langle s_1 \rangle \times \cdots \times \langle s_n \rangle$, $s_i \in S$, this gives complex-valued versions of Gowers' octahedral norms, but it also allows some more exotic examples. For example, the hypergraph described in Example 6.38 can now easily be modified to define a norm on the vector space of complex-valued functions $f : [0,1]^3 \to \mathbb{C}$.

# Bibliography


[1] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008.

[2] A. Björner and F. Brenti. *Combinatorics of Coxeter Groups*. Graduate Texts in Mathematics. Springer Berlin Heidelberg, 2005.

[3] G. Blakley and P. Roy. Hölder type inequality for symmetrical matrices with non-negative entries. *Proc. Amer. Math. Soc.*, 16:1244–1245, 1965.

[4] J. A. Bondy and M. Simonovits. Cycles of even length in graphs. *J. Combin. Theory Ser. B*, 16:97–105, 1974.

[5] O. A. Camarena, E. Csóka, T. Hubai, G. Lippner, and L. Lovász. Positive graphs. To appear in European Journal of Combinatorics.

[6] F. R. K. Chung, R. L. Graham, P. Frankl, and J. B. Shearer. Some intersection theorems for ordered sets and graphs. *J. Combin. Theory Ser. A*, 43(1):23–37, 1986.

[7] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9:345–362, 1989.

[8] E. Cohen, P. Csikvári, W. Perkins, and P. Tetali. The Widom–Rowlinson model, the hard-core model and the extremality of the complete graph. *Europ. J. Combin.*, 62:70–76, 2017.

[9] D. Conlon, J. Fox, and B. Sudakov. An approximate version of Sidorenko's conjecture. *Geom. Funct. Anal.*, 20(6):1354–1366, 2010.

[10] D. Conlon, H. Hàn, Y. Person, and M. Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures Algorithms*, 40:1–38, 2012.

[11] D. Conlon, J. H. Kim, C. Lee, and J. Lee. Sidorenko's conjecture for higher tree decompositions. Unpublished note.

[12] D. Conlon, J. H. Kim, C. Lee, and J. Lee. Some advances on Sidorenko's conjecture. Submitted.

[13] D. Conlon and J. Lee. Finite reflection groups and graph norms. *Adv. Math.*, 315:130–165, 2017.

[14] D. Dellamonica Jr, P. Haxell, T. Łuczak, D. Mubayi, B. Nagle, Y. Person, V. Rödl, and M. Schacht. Tree-minimal graphs are almost regular. *J. Comb*, 3:49–62, 2012.

[15] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[16] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983.

[17] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22:89–03, 1971.

[18] E. Friedgut. Hypergraphs, entropy, and inequalities. *Amer. Math. Monthly*, 111(9):749–760, 2004.

[19] D. Galvin. Three tutorial lectures on entropy and counting. arXiv preprint arXiv:1406.7872, 2014.

[20] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[21] W. T. Gowers. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Combin. Probab. Comput.*, 15(1-2):143–184, 2006.

[22] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math. (2)*, 166(3):897–946, 2007.

[23] R. L. Graham, V. Rödl, and A. Ruciński. On bipartite graphs with linear Ramsey numbers. *Combinatorica*, 21(2):199–209, 2001.

[24] R. Halin. S-functions for graphs. *J. Geom.*, 8:171–186, 1976.

[25] H. Hatami. *On Generalizations of Gowers Norms*. PhD thesis, University of Toronto, 2009.

[26] H. Hatami. Graph norms and Sidorenko's conjecture. *Israel J. Math.*, 175(1):125–150, 2010.

[27] J. Humphreys. *Reflection Groups and Coxeter Groups*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1992.

[28] J. H. Kim, C. Lee, and J. Lee. Two approaches to Sidorenko's conjecture. *Trans. Amer. Math. Soc.*, 368(7):5057–5074, 2016.

[29] Y. Kohayakawa, B. Nagle, V. Rödl, and M. Schacht. Weak hypergraph regularity and linear hypergraphs. *J. Combin. Theory Ser. B*, 100(2):151–160, 2010.

[30] S. Kopparty and B. Rossman. The homomorphism domination exponent. *Europ. J. Combin.*, 32(7):1097–1114, 2011.

[31] T. Kövári, V. Sós, and P. Turán. On a problem of k. zarankiewicz. *Colloquium Math.*, 3:50–57, 1954.

[32] J. Lee. Counting tree-like graphs in locally dense graphs. Submitted.

[33] J. Li and B. Szegedy. On the logarithmic calculus and Sidorenko's conjecture. To appear in Combinatorica.

[34] D. London. Two inequalities in nonnegative symmetric matrices. *Pacific J. Math.*, 16(3):515–536, 1966.

[35] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc.*, 55(10):961–962, 1949.

[36] L. Lovász. *Large Networks and Graph Limits*. Amer. Math. Soc. Colloq. Publ. American Mathematical Society, 2012.

[37] L. Lovász and M. Simonovits. On the number of complete subgraphs of a graph. II. In *Studies in pure mathematics*, pages 459–495. Birkhäuser, Basel, 1983.

[38] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[39] P. McMullen and E. Schulte. *Abstract Regular Polytopes*. Cambridge University Press, 2002.

[40] H. Mulholland and C. Smith. An inequality arising in genetical theory. *Amer. Math. Monthly*, 66, 1959.

[41] V. Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.*, 363(3):1599–1618, 2011.

[42] C. M. R. Daudel, R. Lefebvre. *Quantum chemistry: Methods and applications*. Interscience Publishers, New York-London, 1959.

[43] A. A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.*, 17(4):603–618, 2008.

[44] C. Reiher. Counting odd cycles in locally dense graphs. *J. Combin. Theory Ser. B*, 105:1–5, 2014.

[45] C. Reiher. The clique density theorem. *Ann. of Math.*, 184(3):683–707, 2016.

[46] C. Reiher, V. Rödl, and M. Schacht. On a Turán problem in weakly quasirandom 3-uniform hypergraphs. Preprint.

[47] N. Robertson and P. D. Seymour. Graph minors III: Planar tree-width. *J. Combin. Theory Ser. B*, 36:49–64, 1984.

[48] A. Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.*, 9(2-4):201–204, 1993.

[49] A. Sidorenko. Inequalities for functionals generated by bipartite graphs. *Discrete Math. Appl.*, 2:489–504, 1993.

[50] A. Sidorenko. An analytic approach to extremal problems for graphs and hypergraphs. In *Extremal Problems for Finite Sets*, volume 3 of *Bolyai Society Mathematical Studies*, pages 425–457. 1994.

[51] M. Simonovits. Extremal graph problems, degenerate extremal problems and super-saturated graphs. In *Progress in graph theory (Waterloo, Ont., 1982).* Academic Press, Toronto, ON, 1984.

[52] J. Skokan and L. Thoma. Bipartite subgraphs and quasi-randomness. *Graphs Combin.*, 20:255–262, 2004.

[53] G. Stell. Generating functionals and graphs. In *Graph Theory and Theoretical Physics*. Academic Press, London, 1967.

[54] B. Szegedy. An information theoretic approach to Sidorenko's conjecture. arXiv:1406.6738 [math.CO].

[55] T. Tao. *An introduction to measure theory*, volume 126. American Mathematical Society, 2011.

[56] A. G. Thomason. Pseudorandom graphs. *North-Holland Math. Stud.*, 144:307–331, 1987. in Random graphs '85 (Poznań, 1985).

[57] H. Towsner. Sigma-algebras for quasirandom hypergraphs. *Random Struct. Alg.*, 50(1):114–139, 2017.

[58] P. Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok*, 48:436452, 1941.