A characterization of functions with vanishing averages over products of disjoint sets

Hamed Hatami * Pooya Hatami † Yaqiao Li ‡

March 8, 2018

Abstract

Given $\alpha_1, \ldots, \alpha_m \in (0,1)$, we characterize all integrable functions $f : [0,1]^m \to \mathbb{C}$ satisfying $\int_{A_1\times\cdots\times A_m} f = 0$ for any collection of disjoint sets $A_1, \ldots, A_m \subseteq [0,1]$ of respective measures $\alpha_1, \ldots, \alpha_m$. We use this characterization to settle some of the conjectures in [S. Janson and V. Sós, More on quasi-random graphs, subgraph counts and graph limits, arXiv:1405.6808].

AMS Subject Classification: 05C99

Keywords: generalized Walsh expansion, graph limits, quasi-randomness.

1 Introduction

Answering a question of Janson and Sós [JS14, Problem 4.5], given $\alpha_1, \ldots, \alpha_m \in (0,1)$, we characterize all integrable functions $f : [0,1]^m \to \mathbb{C}$ that satisfy

$$\int_{A_1\times\cdots\times A_m} f = 0 \tag{1}$$

for every collection of disjoint sets $A_1, \ldots, A_m \subseteq [0,1]$ where $A_1, \ldots, A_m$ are of respective measures $\alpha_1, \ldots, \alpha_m$. While the question is very natural on its own, it also arises naturally in the study of certain quasi-random properties of graphs. Indeed this was the original motivation of Janson and Sós [JS14] for asking and studying this question.

Given a number $p \in (0,1)$, roughly speaking, a graph sequence $\{G_n\}_{n=1}^{\infty}$ is called p-quasi-random if, in the limit, it behaves similar to the sequence of Erdös-Rényi random graphs $G(|V(G_n)|, p)$. In the seminal works Thomason [Tho87a, Tho87b] and Chung, Graham and Wilson [CGW89] suggested a rigorous definition of a quasi-random graph sequence, and made a curious observation that many seemingly different definitions are equivalent, and thus lead to the same notion of quasi-randomness. One that turns out to be particularly useful is the following:

Definition 1.1. *A graph sequence $\{G_n\}_{n=1}^{\infty}$ is p-quasi-random if and only if $|V(G_n)| \to \infty$ and $N(F, G_n) = (p^{|E(F)|} + o(1))|V(G_n)|^{|V(F)|}$ for every graph F, where $N(F, G_n)$ denotes the number of* labeled *copies of F in G_n as a subgraph (not necessarily induced).*

*McGill University. hatami@cs.mcgill.ca. Supported by an NSERC grant.

†University of Chicago. pooya@cs.uchicago.edu.

‡McGill University. yaqiao.li@mail.mcgill.ca.

A graph sequence $\{G_n\}_{n=1}^{\infty}$ is called *convergent* [LS06] if the normalized subgraph counts $N(F,G_n)/|V_n|^{|V(F)|}$ converge for every graph F. The limit of a convergent graph sequence can be represented by a so called *graphon*, which is a symmetric measurable function $W:[0,1]^2 \to [0,1]$. More precisely, given a convergent graph sequence $\{G_n\}_{n=1}^{\infty}$, there always exists a graphon $W:[0,1]^2 \to [0,1]$ such that for every integer $m>0$ and every graph F with vertex set $\{1,\ldots,m\}$, we have

$$\lim_{n\to\infty} N(F,G_n)/|V(G_n)|^{|V(F)|} = \int_{[0,1]^m} \prod_{(i,j)\in E(F)} W(x_i,x_j)dx_1dx_2\cdots dx_m.$$

We denote the integral in the right-hand side by $t(F,W)$. Conversely, for every graphon W, one can construct a graph sequence that converges to W in the above sense. Note that every p-quasi-random graph sequence converges to the constant graphon $W=p$, where here and in the sequel when we say two functions are equal, we mean they are equal almost everywhere. Hence, often with a bit of work, one can translate various characterizations of p-quasi-random graph sequences to statements asserting that the constant graphon p is the unique graphon that satisfies a certain condition. For example, Chung, Graham and Wilson [CGW89] showed that it suffices to require the condition of Definition 1.1 only for two graphs $F=K_2$ and $F=C_4$. In the language of graph limits this corresponds to the fact that the graphon $W=p$ is the unique graphon that satisfies $t(K_2,W)=p$ and $t(C_4,W)=p^4$.

It is not difficult to see that there is no single graph F such that $t(W,F)=p^{|E(F)|}$ would imply $W=p$. As a substitute, Simonovits and Sós [SS97] considered the hereditary versions of the subgraph counts, and showed that in fact for every fixed graph F, the condition $N(F,G_n[U])=p^{|E(F)|}+o\left(|V(G_n)|^{|V(F)|}\right)$ is satisfied for all subsets $U\subseteq V(G_n)$ if and only if the sequence $\{G_n\}_{n=1}^{\infty}$ is p-quasi-random. Here $G_n[U]$ denotes the subgraph of G_n induced on U. In the language of graph limits this is equivalent to saying that given a graph F with vertices $\{1,\ldots,m\}$, the graphon $W=p$ is the only graphon that satisfies

$$\int_{A^m} \prod_{(i,j)\in E(F)} W(x_i,x_j)dx_1\cdots dx_m = p^{|E(F)|}\lambda(A)^m$$

for all measurable $A\subseteq[0,1]$. Yuster [Yus10] showed that given any $\alpha\in(0,1)$, it suffices to require this condition only for A of measure α. Shapira [Sha08], Yuster and Shapira [SY10], and Janson and Sós [JS14] considered the condition

$$\int_{A_1\times\cdots\times A_m} \prod_{(i,j)\in E(F)} W(x_i,x_j)dx_1\cdots dx_m = p^{|E(F)|}\prod_{i=1}^{m}\alpha_i \qquad (2)$$

for all disjoint sets $A_1,\ldots,A_m\subseteq[m]$ of respective measures $\alpha_1,\ldots,\alpha_m$. They studied the question that for which graphs F with vertex set $\{1,\ldots,m\}$ and sequences $\alpha_1,\ldots,\alpha_m\in(0,1)$ with $\sum_{i=1}^m \alpha_i \leqslant 1$, the graphon $W=p$ is the unique graphon that satisfies Eq. (2) for all disjoint sets $A_1,\ldots,A_m\subseteq[0,1]$ of respective measures $\alpha_1,\ldots,\alpha_m$. Following the notation of [JS14], in this case, we say that $\mathcal{P}(F,\alpha_1,\ldots,\alpha_m)$ is a *quasi-random property.* Note that this is equivalent to Eq. (1) with

$$f=\left(\prod_{(i,j)\in E(F)} W(x_i,x_j)\right)-p^{|E(F)|},$$

and thus naturally raises the question [JS14, Problem 4.5] that which integrable functions $f:[0,1]^m\to\mathbb{C}$ satisfy Eq. (1) for all disjoint sets $A_1,\ldots,A_m\subseteq[0,1]$ of respective measures $\alpha_1,\ldots,\alpha_m$. We solve this problem in Theorem 3.1.

As an application of our Theorem 3.1, in Corollary 3.4 and Theorem 3.5, we recover [JS14, Theorem 2.11], and furthermore show that when F contains twin vertices, $\mathcal{P}(F, \alpha_1, \ldots, \alpha_m)$ is a quasi-random property. The latter in particular answers [JS14, Problem 2.19] in the affirmative.

Finally as another application of our proof technique, in Theorem 3.6 we solve [JS14, Conjecture 9.4] regarding symmetric functions.

2 Notations and Preliminary Results

For every natural number m, denote $[m] := \{1, \ldots, m\}$. Let λ denote the Lebesgue measure on reals. For $x \in [0,1]^m$ and $S \subseteq [m]$, let $x_S \in [0,1]^S$ denote the restriction of x to the coordinates in S. For disjoint sets $S, T \subseteq [m]$, and $y \in [0,1]^S$ and $z \in [0,1]^T$, let (y, z) denote the unique element in $[0,1]^{S \cup T}$ satisfying $(y,z)_S = y$ and $(y,z)_T = z$. For a vector $x \in [0,1]^m$ and an index $i \in [m]$, $x(i) \in [0,1]$ denotes the i-th entry of x.

For $S \subseteq [m]$, we denote by $\overline{S} := [m] \setminus S$ the complement of S. Given a function $f : [0,1]^m \to \mathbb{C}$ and $y \in [0,1]^S$, we define $f_y : [0,1]^{\overline{S}} \to \mathbb{C}$ by $f_y : z \mapsto f(y,z)$ for every $z \in [0,1]^{\overline{S}}$. In the sequel, by an abuse of notation, we sometimes identify a function $f : [0,1]^S \to \mathbb{C}$ with its extension to $[0,1]^m$ defined as $x \mapsto f(x_S)$ for $x \in [0,1]^m$.

We say that a function $f : [0,1]^m \to \mathbb{C}$ is an *alternating function* with respect to the coordinates in $S \subseteq [m]$ if the interchange of the values of any two coordinates in S changes the sign of f.

Given $\alpha = (\alpha_1, \ldots, \alpha_m) \in [0,1]^m$ with $\sum_{i=1}^m \alpha_i = 1$, call a partition $A_1, \ldots, A_m$ of $[0,1]$ an *α-partition* if $\lambda(A_i) = \alpha_i$ for $i = 1, \ldots, m$ and the boundary of each A_i is of measure 0. Given subsets $A_1, \ldots, A_m \subseteq [0,1]$ and $S \subseteq [m]$, let A_S denote the product $\prod_{i \in S} A_i$.

For a positive integer m, let S_m denote the symmetric group of order m.

2.1 Generalized Walsh Expansion

Our proofs of Theorem 3.1 and Theorem 3.6 use the so called generalized Walsh expansion, which was first defined by Hoeffding in [Hoe48] (See also [ES81]).

Definition 2.1. *The generalized Walsh expansion of an integrable function $f : [0,1]^m \to \mathbb{C}$ is the expansion $f = \sum_{S \subseteq [m]} F_S$ that satisfies the following two properties.*

(i) For every $S \subseteq [m]$, the function F_S depends only on the coordinates in S, i.e. $F_S(x) = F_S(x_S)$;

(ii) $\int_{[0,1]} F_S(x) dx_i = 0$, for every $S \subseteq [m]$ and every $i \in S$.

We call a function F_S satisfying Definition 2.1 (i) and (ii) a *generalized Walsh function*. It is not difficult to see that the generalized Walsh expansion is unique and can be computed using the following formula

$$F_S(y) = \sum_{T \subseteq S} (-1)^{|S \setminus T|} \int_{[0,1]^{\overline{T}}} f(y_T, x_{\overline{T}}) dx_{\overline{T}}.$$

In the sequel, for the sake of brevity, we shall often drop the word "generalized" from the terms "generalized Walsh expansion" and "generalized Walsh function".

Given an integrable function $f : [0,1]^m \to \mathbb{C}$, for $0 \leqslant k \leqslant m$, we denote by $f^{\leqslant k} := \sum_{S \subseteq [m], |S| \leqslant k} F_S$ the projection of f to the first k "levels". The projections $f^{=k}$, $f^{\geqslant k}$, $f^{<k}$, and $f^{>k}$ are defined similarly.

3 Main Results

We are now ready to state our results formally. We start by stating our main theorem that characterizes all integrable functions $f : [0,1]^m \to \mathbb{C}$ satisfying Eq. (1) in the case where $\sum_{i=1}^m \alpha_i = 1$. We handle the case $\sum_{i=1}^m \alpha_i < 1$ as a consequence of this in Corollary 3.3.

Theorem 3.1 (Main Theorem)**.** *Let $\alpha_1, \ldots, \alpha_m \in (0,1)$ satisfy $\sum_{i=1}^m \alpha_i = 1$. An integrable function $f : [0,1]^m \to \mathbb{C}$ with the Walsh expansion $f = \sum_{S \subseteq [m]} F_S$ satisfies*

$$\int_{A_1 \times \cdots \times A_m} f = 0 \tag{3}$$

for all partitions of $[0,1]$ into disjoint sets $A_1, \ldots, A_m$ of respective measures $\alpha_1, \ldots, \alpha_m$ if and only if

(i) $F_\emptyset = 0$;

(ii) F_S *is an alternating function (with respect to the coordinates in S) for all $S \subseteq [m]$ with $|S| \geqslant 2$;*

(iii) For $S \subseteq [m]$, with $1 \leqslant |S| \leqslant m-1$, and $\ell \in [m] \backslash S$, we have

$$\frac{1}{\prod_{i \in S} \alpha_i} F_S(x) = \sum_{i \in S} \frac{1}{\prod_{j \in S_i} \alpha_j} F_{S_i}(x^{(i)}), \tag{4}$$

where $x^{(i)}$ is obtained from $x = (x_1, \ldots, x_m)$ by swapping x_ℓ and x_i, and $S_i := S \cup \{\ell\} \setminus \{i\}$.

We prove Theorem 3.1 in Section 4.

Remark 3.2. *Note that Theorem 3.1 (iii), applied to sets S of size 1, implies that there exists an integrable function $g : [0,1] \to \mathbb{C}$ with $\int_0^1 g(x)dx = 0$ such that $F_{\{i\}}(x) = \alpha_i g(x)$ for every $i \in [m]$.*

Theorem 3.1 provides a way to construct all functions f that satisfy Eq. (3). Indeed one can take any collection of alternating Walsh functions $F_S : [0,1]^m \to \mathbb{C}$ for $S \ni m$, and then use Theorem 3.1 (iii) with $\ell = m$ to define F_S for all other subsets $\emptyset \neq S \subseteq [m]$ accordingly (i.e. all $S \subseteq [m-1]$ of size at least one). Note that the resulting F_S will automatically be Walsh functions and satisfy Theorem 3.1 (i-iii), and thus the function $f = \sum_{S \subseteq [m]} F_S$ will satisfy Eq. (3).

Next we show how the case $\sum \alpha_i < 1$ follows from Theorem 3.1.

Corollary 3.3 ([JS14, Lemma 4.6])**.** *Let $\alpha_1, \ldots, \alpha_m \in (0,1)$, with $\sum_{i=1}^m \alpha_i < 1$. An integrable function $f : [0,1]^m \to \mathbb{C}$ satisfies*

$$\int_{A_1 \times \cdots \times A_m} f = 0 \tag{5}$$

for all disjoint $A_1, \ldots, A_m \subseteq [0,1]$ of respective measures $\alpha_1, \ldots, \alpha_m$ if and only if $f = 0$ almost everywhere. The same assertion holds if f is symmetric and $\sum_{i=1}^m \alpha_i = 1$, but $(\alpha_1, \ldots, \alpha_m) \neq (1/m, \ldots, 1/m)$.

Proof. First consider the case $\sum_{i=1}^m \alpha_i < 1$. Define $\alpha_{m+1} = 1 - \sum_{i=1}^m \alpha_i$ and apply Theorem 3.1 to the sequence $\alpha_1, \ldots, \alpha_{m+1}$ and the function $\tilde{f} : [0,1]^{m+1} \to \mathbb{C}$ defined as $\tilde{f} : (x_1, \ldots, x_{m+1}) \mapsto f(x_1, \ldots, x_m)$. The assertion now follows from Theorem 3.1 as in the Walsh expansion $\tilde{f} = \sum_{S \subseteq [m+1]} \tilde{F}_S$, we have $\tilde{F}_S = 0$ for every $S \subseteq [m+1]$ with $m+1 \in S$.

To prove the case where f is symmetric but $(\alpha_1, \dots, \alpha_m) \neq (1/m, \dots, 1/m)$ note that Theorem 3.1 (ii) and the symmetry of f imply that $F_S = 0$ for every S with $|S| > 1$. Finally, Remark 3.2 and the symmetry shows $F_S = 0$ for every S of size 1. $\square$

Following the notation of [JS14], we say that $\widetilde{\mathcal{P}}(F, \alpha_1, \dots, \alpha_m)$ is a *quasi-random property* if $W = p$ is the unique solution to

$$\frac{1}{m!} \sum_{\sigma \in S_m} \int_{A_{\sigma_1} \times \cdots \times A_{\sigma_m}} \prod_{(i,j) \in E(F)} W(x_i, x_j) dx_1 \cdots dx_m = p^{|E(F)|} \prod_{i=1}^{m} \alpha_i. \tag{6}$$

As it is noticed in [JS14], Corollary 3.3 has the following consequence.

Corollary 3.4 ([JS14, Theorem 2.11])**.** *Let F be a graph with vertex set $\{1, \dots, m\}$ that contains at least one edge, and let $0 < p \leqslant 1$. Furthermore, let $(\alpha_1, \dots, \alpha_m)$ be a vector of positive numbers with $\sum_{i=1}^m \alpha_i \leqslant 1$.*

(i) If $(\alpha_1, \dots, \alpha_m) \neq (1/m, \dots, 1/m)$, then $\widetilde{\mathcal{P}}(F, \alpha_1, \dots, \alpha_m)$ is a quasi-random property.

(ii) If $\sum_{i=1}^m \alpha_i < 1$ then $\mathcal{P}(F, \alpha_1, \dots, \alpha_m)$ is a quasi-random property.

We call two vertices in a graph *twins* if they share the same neighbors (and thus there is no edge between them). Next we use Theorem 3.1 to prove a theorem about graphs containing twin vertices. This in particular solves [JS14, Problem 2.19] regarding quasi-random properties of stars by noting that stars with at least three vertices always contain twins.

Theorem 3.5. *Let F be a graph containing twins, and let $0 < p \leqslant 1$, then $\mathcal{P}(F, \alpha_1, \dots, \alpha_m)$ is a quasi-random property for all $\alpha_1, \dots, \alpha_m \in (0, 1)$ with $\sum_{i=1}^m \alpha_i \leqslant 1$.*

Proof. The case $\sum \alpha_i < 1$ follows from Corollary 3.4. It remains to establish the case $\sum \alpha_i = 1$. Let $f := \prod_{(i,j) \in E(F)} W(x_i, x_j) - p^{|E(F)|}$. We will show that if $\int_{A_{[m]}} f = 0$ for all α-partitions, then $W = p$ almost everywhere.

Without loss of generality, assume v_{m-1}, v_m are twins in F, and $v_1, \dots, v_r$, $r \leqslant m-2$, are their common neighbors. Let $f = \sum_{S \subseteq [m]} F_S$ be the Walsh expansion of f. Therefore, f can be written in the following form

$$\begin{aligned} f &= \left(\prod_{\substack{(i,j) \in E(F) \\ i,j \in [m-2]}} W(x_i, x_j) \right) \Big(\prod_{i=1}^{r} W(x_i, x_{m-1}) W(x_i, x_m) \Big) - p^{|E(F)|} \\ &= \sum_{S \subseteq [m-2]} (F_S + F_{S \cup \{m-1\}} + F_{S \cup \{m\}} + F_{S \cup \{m-1,m\}}). \end{aligned} \tag{7}$$

We claim that for every $S \subseteq [m-2]$, $F_{S \cup \{m-1,m\}} = 0$ almost everywhere. Indeed since v_{m-1}, v_m are twins, $F_{S \cup \{m-1,m\}}$ is symmetric with respect to the two coordinates x_{m-1} and x_m, and on the other hand by Theorem 3.1 (ii), $F_{S \cup \{m-1,m\}}$ is also an alternating function with respect to those coordinates. Hence, $F_{S \cup \{m-1,m\}} = 0$ almost everywhere.

Fixing $x_1, \dots, x_{m-2}$ and integrating Eq. (7) with respect to $x_{m-1}, x_m \in [0, 1]$, we obtain that for almost every $x_1, \dots, x_{m-2}$,

$$p^{|E(F)|} + \sum_{S \subseteq [m-2]} F_S = \prod_{\substack{(i,j) \in E(F) \\ i,j \in [m-2]}} W(x_i, x_j) \left(\int_{[0,1]} \prod_{i=1}^{r} W(x_i, a) da \right)^2. \tag{8}$$

Next we would like to replace the same value a for both x_{m-1} and x_m in Eq. (7), and then integrate with respect to a. However, since $F_{S\cup\{m-1,m\}} = 0$ only almost everywhere, we need to consider the limit instead. More precisely we deduce the following from Eq. (7) and the fact that $F_{S\cup\{m-1,m\}} = 0$ almost everywhere: For almost all $x_1, \ldots, x_{m-2}, a \in [0,1]$, denoting $y := (x_{[m-2]}, a, a) \in [0,1]^m$, we have

$$\begin{aligned}\lim_{\varepsilon_1,\varepsilon_2\to 0^+} \frac{1}{\varepsilon_1\varepsilon_2}\int_{B_{\varepsilon_1/2}(a)}\int_{B_{\varepsilon_2/2}(a)} f\,dx_{m-1}dx_m &= \prod_{\substack{(i,j)\in E(F)\\ i,j\in[m-2]}} W(x_i,x_j)\prod_{i=1}^{r} W(x_i,a)^2 - p^{|E(F)|}\\ &= \sum_{S\subseteq[m-2]} F_S(y) + F_{S\cup\{m-1\}}(y) + F_{S\cup\{m\}}(y).\end{aligned}$$

Integrating this with respect to a, we obtain that for almost all $x_{[m-2]}$,

$$p^{|E(F)|} + \sum_{S\subseteq[m-2]} F_S = \prod_{\substack{(i,j)\in E(F)\\ i,j\in[m-2]}} W(x_i,x_j)\int_{[0,1]}\prod_{i=1}^{r} W(x_i,a)^2 da. \tag{9}$$

Hence (8) = (9) for almost all $x_{[m-2]}$, and then the equality condition of Cauchy-Schwarz implies that for almost all $x_{[m-2]}$, $\prod_{\substack{(i,j)\in E(F)\\ i,j\in[m-2]}} W(x_i,x_j)\prod_{i=1}^{r} W(x_i,a)$ does not depend on a. It follows that

$$\prod_{\substack{(i,j)\in E(F)\\ i,j\in[m-2]}} W(x_i,x_j)\prod_{i=1}^{r} W(x_i,x_{m-1})W(x_i,a)$$

does not depend on a for almost all $x_{[m-1]}$. Hence for every α-partition $A_{[m]}$ and every $B \subseteq A_m$ with $\lambda(B) > 0$, we have

$$\int_{A_{[m-1]}\times B}\prod_{(i,j)\in E(F)} W(x_i,x_j)dx_{[m]} = \frac{\lambda(B)}{\alpha_m}\int_{A_{[m]}}\prod_{(i,j)\in E(F)} W(x_i,x_j)dx_{[m]} = p^{|E(F)|}\lambda(B)\prod_{i=1}^{m-1}\alpha_i.$$

Now Corollary 3.4 (ii) implies that $W = p$ almost everywhere. □

Finally, we state our theorem about symmetric functions which in particular solves [JS14, Conjecture 9.4].

Theorem 3.6. *Let $\alpha \in (0,1)$, and $0 \leqslant r \leqslant m$ be an integer. A symmetric integrable function $f : [0,1]^m \to \mathbb{C}$ satisfies $\int_{A^{m-r}\times(\overline{A})^r} f = 0$ for every $A \subset [0,1]$ with $\lambda(A) = \alpha$ if and only if at least one of the following two cases holds.*

(i) $f = 0$ almost everywhere.

(ii) For $K := K(m,r,\alpha) = \{k \in [m] : \sum_{i=0}^{k}\binom{m-r}{k-i}\binom{r}{i}\left(\frac{-\alpha}{1-\alpha}\right)^i = 0\}$, we have

$$f(x_1,\ldots,x_m) = \sum_{k\in K}\sum_{S\subseteq[m],|S|=k} g_k(x_S)$$

where $g_k : [0,1]^k \to \mathbb{C}$ are symmetric functions satisfying $\int g_k(x_1,\ldots,x_k)dx_i = 0$ for every $i \in [k]$.

Note that Theorem 3.6 (ii) means that in the Walsh expansion $f = \sum_{S\subseteq[m]} F_S$, we have $F_S = 0$ if $|S| \notin K(m,r,\alpha)$ and $F_S(x) = g_k(x_S)$ if $|S| = k \in K(m,r,\alpha)$. We shall not venture to characterize the sets $K(m,r,\alpha)$. However we remark that these sets can contain more than one element, as for example, it is not difficult to see that $K(6,3,\frac{1}{2}) = \{1,3,5\}$. Thus, in general, the Walsh expansion of f can be supported on more than one "level".

The case $m = 3$ and $r = 1$ of Theorem 3.6 was conjectured in [JS14, Conjecture 9.4]. Note that if $r = 1$ in Theorem 3.6, then we have $K = \{k\}$ if $\alpha = \frac{m-k}{m}$ and $K(m, r, \alpha) = \emptyset$ if α is not of the form $\frac{m-k}{m}$. We state this case separately as Corollary 3.7

Corollary 3.7 ([JS14, Conjecture 9.4])**.** *Let $\alpha \in (0, 1)$. A symmetric integrable function $f : [0, 1]^m \to \mathbb{C}$ satisfies $\int_{A^{m-1}\times\overline{A}} f = 0$ for every $A \subset [0, 1]$ with $\lambda(A) = \alpha$ if and only if at least one of the following two cases holds.*

(i) $f = 0$ almost everywhere.

(ii) $\alpha = \frac{m-k}{m}$ for some $k \in [m - 1]$ and

$$f(x_1, \ldots, x_m) = \sum_{S\subseteq[m],|S|=k} g(x_S)$$

where $g : [0, 1]^k \to \mathbb{C}$ is a symmetric function satisfying $\int g(x_1, \ldots, x_k)dx_i = 0$ for every $i \in [k]$.

3.1 Proof Technique: A first variation argument

In this short section we prove the main step used in the proofs of Theorem 3.1 and Theorem 3.6. Let us recall the following form of the Lebesgue differentiation theorem.

Lemma 3.8. *Let $g : [0, 1] \to \mathbb{C}$ be Lebesgue integrable and let $x \in [0, 1]$ be a Lebesgue point of g. Then*

$$g(x) = \lim_{\varepsilon\to0^+} \frac{1}{\varepsilon} \int_{B_{\varepsilon/2}(x)} g(y)dy,$$

where $B_{\varepsilon/2}(x)$ is the ball of radius $\varepsilon/2$ around x.

Let $\alpha_1, \ldots, \alpha_m \in (0, 1)$ and suppose that $A_1, \ldots, A_m \subseteq [0, 1]$ are of measures $\alpha_1, \ldots, \alpha_m$, respectively. Consider $K \subseteq [m]$, and given $\mathbf{y}, \mathbf{z} \in [0, 1]^K$ and $t \geqslant 0$, set

$$A_i(t) := A_i \cup B_{t/2}(\mathbf{z}(i)) \setminus B_{t/2}(\mathbf{y}(i)) \tag{10}$$

for every $i \in K$, and $A_i(t) := A_i$ for $i \in \overline{K}$. Now consider an integrable function $f : [0, 1]^m \to \mathbb{C}$, and define

$$F(t) = \int_{A_1(t)\times\cdots\times A_m(t)} f.$$

It follows from Lemma 3.8 that for almost every $\mathbf{y} \in \mathrm{Int}(A_K)$ and $\mathbf{z} \in \mathrm{Int}(\prod_{i\in K} \overline{A_i})$, we have

$$\left.\frac{dF(t)}{dt}\right|_{0^+} = \sum_{i\in K} \int_{A_{[m]\setminus\{i\}}} \left(f_{\mathbf{z}(i)} - f_{\mathbf{y}(i)}\right) dx_{[m]\setminus\{i\}} = \sum_{i\in K} \frac{1}{\alpha_i} \int_{A_{[m]}} \left(f_{\mathbf{z}(i)} - f_{\mathbf{y}(i)}\right) dx_{[m]}. \tag{11}$$

Let us introduce the notation

$$\partial^i_{\mathbf{y},\mathbf{z}} f := \frac{f_{\mathbf{z}(i)} - f_{\mathbf{y}(i)}}{\alpha_i},$$

and

$$\partial^K_{\mathbf{y},\mathbf{z}} f := \sum_{i\in K} \partial^i_{\mathbf{y},\mathbf{z}} f = \sum_{i\in K} \frac{f_{\mathbf{z}(i)} - f_{\mathbf{y}(i)}}{\alpha_i},$$

so that

$$\left.\frac{dF(t)}{dt}\right|_{0^+} = \int_{A_{[m]}} \partial^K_{\mathbf{y},\mathbf{z}} f.$$

Further, suppose $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ and $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_k)$ where $\mathbf{y}_i, \mathbf{z}_i \in [0,1]^K$ for $i = 1, \ldots, k$. Define

$$\partial^K_{\mathbf{Y},\mathbf{Z}} f := \partial^K_{\mathbf{y}_k,\mathbf{z}_k} \cdots \partial^K_{\mathbf{y}_1,\mathbf{z}_1} f.$$

Note that when $g : [0,1]^m \to \mathbb{C}$ does not depend on the i-th coordinate, then $\partial^i_{\mathbf{y},\mathbf{z}} g = 0$. Combining this and the fact that $\partial^i_{\mathbf{y},\mathbf{z}} f$ does not depend on the i-th coordinate, we conclude that for any Walsh function F_S, and any $\mathbf{Y}, \mathbf{Z} \in ([0,1]^K)^k$, we have

$$\partial^K_{\mathbf{Y},\mathbf{Z}} F_S = \sum_{\substack{j_1,\ldots,j_k \in S \cap K \\ |\{j_1,\ldots,j_k\}| = k}} \partial^{j_k}_{\mathbf{y}_k,\mathbf{z}_k} \cdots \partial^{j_1}_{\mathbf{y}_1,\mathbf{z}_1} F_S.$$

Expanding this formula leads to the following lemma which is central to the proofs of both Theorem 3.1 and Theorem 3.6.

Lemma 3.9. *Consider $S, K \subseteq [m]$, and let $F_S : [0,1]^m \to \mathbb{C}$ depend only on the coordinates in S. Given any $\mathbf{Y}, \mathbf{Z} \in ([0,1]^K)^k$, we have*

$$\partial^K_{\mathbf{Y},\mathbf{Z}} F_S = \sum_{\substack{D \subseteq S \cap K \\ |D| = k}} \frac{1}{\prod_{i \in D} \alpha_i} \sum_{\pi : D \xrightarrow{1:1} [k]} \sum_{B \subseteq [k]} (-1)^{|B|} F_S(\mathbf{w}, \cdot)$$

where $\mathbf{w} = \mathbf{w}_{B,\pi} \in [0,1]^D$ defined as

$$\mathbf{w}(t) = \begin{cases} \mathbf{y}_{\pi(t)}(t) & \pi(t) \in B, \\ \mathbf{z}_{\pi(t)}(t) & \pi(t) \in [k] \setminus B. \end{cases}$$

In particular, $\partial^K_{\mathbf{Y},\mathbf{Z}} F_S$ is equal to zero if $|S \cap K| < k$, and is a constant if $|S| = |S \cap K| = k$.

In the sequel, $\partial_{\mathbf{Y},\mathbf{Z}}$ and $\partial_{\mathbf{y},\mathbf{z}}$ are respectively short forms for $\partial^{[m]}_{\mathbf{Y},\mathbf{Z}}$ and $\partial^{[m]}_{\mathbf{y},\mathbf{z}}$.

4 Proof of Theorem 3.1

Throughout this section fix $\alpha = (\alpha_1, \ldots, \alpha_m) \in (0,1)^m$. Given a subset $S \subseteq [m]$, denote $\alpha(S) := \prod_{i \in S} \alpha_i$. Let us recall the statement of Theorem 3.1.

Theorem 3.1 (restated). *Let $\alpha_1, \ldots, \alpha_m \in (0,1)$ satisfy $\sum_{i=1}^m \alpha_i = 1$. An integrable function $f : [0,1]^m \to \mathbb{C}$ with the Walsh expansion $f = \sum_{S \subseteq [m]} F_S$ satisfies*

$$\int_{A_1 \times \cdots \times A_m} f = 0 \tag{12}$$

for all partitions of $[0,1]$ into disjoint sets $A_1, \ldots, A_m$ of respective measures $\alpha_1, \ldots, \alpha_m$ if and only if

(i) $F_\emptyset = 0$;

(ii) F_S is an alternating function (with respect to the coordinates in S) for all $S \subseteq [m]$ with $|S| \geqslant 2$;

(iii) For $S \subseteq [m]$, with $1 \leqslant |S| \leqslant m-1$, and $\ell \in [m] \setminus S$, we have

$$\frac{1}{\prod_{i \in S} \alpha_i} F_S(x) = \sum_{i \in S} \frac{1}{\prod_{j \in S_i} \alpha_j} F_{S_i}(x^{(i)}), \tag{13}$$

where $x^{(i)}$ is obtained from $x = (x_1, \ldots, x_m)$ by swapping x_ℓ and x_i, and $S_i := S \cup \{\ell\} \setminus \{i\}$.

We divide the proof into two sections, the "if", and the "only if" parts. Firstly, we prove the following lemma which will be useful in both directions. Recall that $f^{=k} := \sum_{\substack{S\subseteq[m]\\|S|=k}} F_S$.

Lemma 4.1. *Given any fixed $1 \leqslant k \leqslant m$, assume that Theorem 3.1 (ii) and (iii) hold for all F_S such that $|S| = k$. Then for all α-partitions $A_1, \ldots, A_m$, we have*

$$\int_{A_{[m]}} f^{=k} = 0. \tag{14}$$

Note that (ii) is void in the case of $k = 1$ and thus holds trivially.

Proof. Consider an α-partition $A_1, \ldots, A_m$. For the given k, for any $S \subseteq [m]$ with $m \in S$ and $|S| = k$, because $\sum_{i=1}^{m} \int_{A_i} F_S(x)dx_m = \int_{[0,1]} F_S(x)dx_m = 0$, we have

$$\int_{A_{[m]}} F_S = -\sum_{i=1}^{m-1} \int_{A_i} \left(\int_{A_{[m-1]}} F_S dx_{[m-1]} \right) dx_m.$$

Theorem 3.1 (ii) says F_S is an alternating function, which implies $\int_{A_i}\int_{A_i} F_S dx_i dx_m = 0$ if $i \in S$. Hence

$$\int_{A_{[m]}} F_S = -\sum_{i\notin S} \int_{A_i} \int_{A_{[m-1]}} F_S dx_{[m-1]} dx_m.$$

Consequently

$$\begin{aligned}
\frac{1}{\alpha([m])} \int_{A_{[m]}} \sum_{\substack{S\subseteq[m]\\|S|=k}} F_S dx_{[m]} &= \sum_{\substack{S\subseteq[m-1]\\|S|=k}} \frac{1}{\alpha([m])} \int_{A_{[m]}} F_S dx_{[m]} + \sum_{\substack{S\subseteq[m],S\ni m\\|S|=k}} \frac{1}{\alpha([m])} \int_{A_{[m]}} F_S dx_{[m]} \\
&= \sum_{\substack{S\subseteq[m-1]\\|S|=k}} \frac{1}{\alpha(S)} \int_{A_S} F_S dx_S - \sum_{\substack{S\subseteq[m],S\ni m\\|S|=k}} \frac{1}{\alpha([m])} \sum_{i\notin S} \int_{A_i} \int_{A_{[m-1]}} F_S dx_{[m-1]} dx_m \\
&= \sum_{\substack{S\subseteq[m-1]\\|S|=k}} \frac{1}{\alpha(S)} \int_{A_S} F_S dx_S - \sum_{\substack{S\subseteq[m],S\ni m\\|S|=k}} \frac{1}{\alpha(S)} \sum_{i\notin S} \int_{A_i} \int_{A_{S\setminus\{m\}}} F_S dx_{S\setminus\{m\}} dx_m \\
&= \sum_{\substack{S\subseteq[m-1]\\|S|=k}} \frac{1}{\alpha(S)} \int_{A_S} F_S(x) dx_S - \sum_{\substack{S\subseteq[m],S\ni m\\|S|=k}} \frac{1}{\alpha(S)} \sum_{i\notin S} \int_{A_{S\setminus\{m\}}\times A_i} F_S(x) dx_{S\setminus\{m\}} dx_m \\
&= \sum_{\substack{S\subseteq[m-1]\\|S|=k}} \left(\frac{1}{\alpha(S)} \int_{A_S} F_S(x) dx_S - \sum_{i\in S} \frac{1}{\alpha(S_i)} \int_{A_S} F_{S_i}(x^{(i)}) dx_S \right) = 0,
\end{aligned}$$

where the last equality uses Theorem 3.1 (iii), with $S \subseteq [m-1]$ and $\ell = m$. □

4.1 Proof of Theorem 3.1, the "if" part

Using Lemma 4.1, $\int_{A_{[m]}} f^{=k} = 0$ for all $1 \leqslant k \leqslant m$ by (ii) and (iii). Hence (i-iii) imply $\int_{A_{[m]}} f = 0$.

4.2 Proof of Theorem 3.1, the "only if" part

For $\mathbf{y} = (y_1, \ldots, y_m) \in [0,1]^m$ and $\sigma \in S_m$, let $\mathbf{y}_\sigma = (y_{\sigma(1)}, \ldots, y_{\sigma(m)}) \in [0,1]^m$ be obtained by rearranging the coordinates of $\mathbf{y}$ according to σ. Given $\sigma \in S_m$, define $K_\sigma := \{i \in [m] : i \neq \sigma(i)\}$. Let us denote $\partial^j_{\mathbf{y},\sigma} := \partial^j_{\mathbf{y},\mathbf{y}_\sigma}$, $\partial_{\mathbf{y},\sigma} := \partial^{K_\sigma}_{\mathbf{y},\mathbf{y}_\sigma}$ and $\partial_{\mathbf{Y},\sigma} := \partial^{K_\sigma}_{\mathbf{Y},\mathbf{Y}_\sigma}$, where for $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k) \in ([0,1]^m)^k$, we have $\mathbf{Y}_\sigma = ((\mathbf{y}_1)_\sigma, \ldots, (\mathbf{y}_k)_\sigma)$.

To prove the theorem we will use induction on $|S|$ to show that F_S satisfies Theorem 3.1 (ii) and (iii) for all S with $|S| \geqslant 1$. Theorem 3.1 (i) then follows from Theorem 3.1 (ii-iii) and Eq. (13). Let $k \geqslant 1$, and assume (ii) and (iii) hold for all F_S such that $k+1 \leqslant |S| \leqslant m$. By Lemma 4.1 we have

$$\int_{A_{[m]}} f^{\leqslant k} = \int_{A_{[m]}} f = 0, \tag{15}$$

for every α-partition $A_1, \ldots, A_m$.

Consider an α-partition $A_1, \ldots, A_m$, and let $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ where $\mathbf{y}_i = (y_{i1}, \ldots, y_{im}) \in \mathrm{Int}(A_{[m]})$ are generic points for $i = 1, \ldots, k$.

Note that for every sufficiently small $t > 0$, the sets $A_1(t), \ldots, A_m(t)$ defined as in Eq. (10), with $\mathbf{y} = \mathbf{y}_1$ and $\mathbf{z} = (\mathbf{y}_1)_\sigma$, and any $\sigma \in S_m$ form an α-partition of $[0,1]$. Hence by Eq. (15), we have $F(t) = \int_{A_1(t)\times\cdots\times A_m(t)} f^{\leqslant k} = 0$ for sufficiently small t. Consequently $\frac{dF(t)}{dt}\Big|_{0^+} = 0$ which in turn implies that $\int_{A[m]} \partial_{\mathbf{y}_1,\sigma} f^{\leqslant k} = 0$. Replacing $f^{\leqslant k}$ with $\partial_{\mathbf{y}_1,\sigma} f^{\leqslant k}$ and repeating the above argument we conclude that

$$\int_{A_{[m]}} \partial_{\mathbf{Y},\sigma} f^{\leqslant k} = \int_{A_{[m]}} \partial_{\mathbf{y}_k,\sigma} \cdots \partial_{\mathbf{y}_1,\sigma} f^{\leqslant k} = 0, \tag{16}$$

for every $\sigma \in S_m$, every α-partition $A_1, \ldots, A_m$, and almost every set of points $\mathbf{y}_1, \ldots, \mathbf{y}_k \in \mathrm{Int}(A_{[m]})$.

Theorem 3.1 (ii): Let $S \subseteq [m]$ be of size $k \geqslant 2$. Without loss of generality assume that $S = [k]$. Setting $\sigma = (1\ 2\ \cdots\ k) \in S_m$, we have $K_\sigma = [k] = S$. By Lemma 3.9, Eq. (16) simplifies to

$$0 = \int_{A_{[m]}} \partial_{\mathbf{Y},\sigma} f^{\leqslant k} = \int_{A_{[m]}} \partial_{\mathbf{Y},\sigma} F_S = \frac{\alpha([m])}{\alpha([k])} \sum_{\pi:[k]\xrightarrow{1:1}[k]} \sum_{B\subseteq[k]} (-1)^{|B|} F_S(\mathbf{w}),$$

where $\mathbf{w} = \mathbf{w}_{B,\pi} \in [0,1]^{[k]}$ is defined as

$$\mathbf{w}(t) = \begin{cases} y_{\pi(t)t}, & \pi(t) \in B, \\ y_{\pi(t)\sigma(t)}, & \pi(t) \in [k] \setminus B. \end{cases}$$

Hence, for almost every $\mathbf{y}_1, \ldots, \mathbf{y}_k \in [0,1]^m$, we have

$$\sum_{\pi:[k]\xrightarrow{1:1}[k]} \sum_{B\subseteq[k]} (-1)^{|B|} F_S(\mathbf{w}) = 0. \tag{17}$$

Let us fix the k entries $\{y_{12}, y_{22}, y_{33}, \ldots, y_{kk}\}$ among $m \times k$ entries in $\mathbf{Y}$. We claim that in (17), there are only two terms containing these k entries simultaneously: $F_S(y_{12}, y_{22}, y_{33}, \ldots, y_{kk})$ and $F_S(y_{22}, y_{12}, y_{33}, \ldots, y_{kk})$, corresponding respectively to $(\pi(1), \ldots, \pi(k)) := (1, \ldots, k), B = [k]\backslash\{1\}$ and $(\pi(1), \pi(2), \ldots, \pi(k)) := (2, 1, 3, 4 \ldots, k), B = [k]\backslash\{2\}$. In particular, the cardinalities of these two Bs are the same, hence these two terms are of the same sign.

To see the claim, observe that 2 appears twice as the column index in the k entries $\{y_{12}, y_{22}, y_{33}, \ldots, y_{kk}\}$, and hence by the definition of $\mathbf{w}(t)$, we must have either $\mathbf{w}(1) = y_{12}, \mathbf{w}(2) = y_{22}$ or $\mathbf{w}(1) = y_{22}, \mathbf{w}(2) = y_{12}$. It is then easy to see that, by our choice of $\sigma = (1\ 2\ \cdots\ k)$, the values for the remaining entries of $\mathbf{w}(t)$ are uniquely determined as $\mathbf{w}(t) = y_{tt}, 3 \leqslant t \leqslant k$. The permutation π and the set B are then determined accordingly.

Thus by Definition 2.1 (ii), integrating Eq. (17) over all the variables $\{y_{ij} : i \in [m], j \in [k]\}\backslash\{y_{12}, y_{22}, y_{33}, \ldots, y_{kk}\}$ we get

$$F_S(y_{12}, y_{22}, y_{33}, \ldots, y_{kk}) + F_S(y_{22}, y_{12}, y_{33}, \ldots, y_{kk}) = 0.$$

This shows that F_S is an alternating function with respect to the first two coordinates. The condition with respect to the other coordinates can be shown similarly.

Theorem 3.1 (iii): It remains to show that F_S satisfies Theorem 3.1 (iii). By symmetry it suffices to prove the statement for $\ell = m$. Again without loss of generality assume $S = [k]$, and now let $\rho = (1, 2, \ldots, k, m) \in S_m$. Since $K_\rho = \{1, \ldots, k\} \cup \{m\}$, denoting $S_0 := S$ and defining $S_1, \ldots, S_k$ as in Theorem 3.1 (iii), Eq. (16) reduces to

$$\begin{aligned} 0 &= \int_{A_{[m]}} \partial_{\mathbf{Y},\rho} f^{\leqslant k} = \int_{A_{[m]}} \partial_{\mathbf{Y},\rho} \sum_{i=0}^{k} F_{S_i} \\ &= \alpha([m]) \sum_{i=0}^{k} \sum_{\pi : S_i \xrightarrow{1:1} [k]} \sum_{B \subseteq [k]} \frac{1}{\alpha(S_i)} (-1)^{|B|} F_{S_i}(\mathbf{w}), \end{aligned}$$

where $\mathbf{w} = \mathbf{w}_{i,B,\pi} \in [0,1]^{S_i}$ is defined as

$$\mathbf{w}(t) = \begin{cases} y_{\pi(t)t}, & \pi(t) \in B, \\ y_{\pi(t)\rho(t)}, & \pi(t) \in S_i \setminus B. \end{cases}$$

Hence for almost every $\mathbf{y}_1, \ldots, \mathbf{y}_k \in [0,1]^m$, we have

$$\sum_{i=0}^{k} \sum_{\pi : S_i \xrightarrow{1:1} [k]} \sum_{B \subseteq [k]} \frac{1}{\alpha(S_i)} (-1)^{|B|} F_{S_i}(\mathbf{w}) = 0, \tag{18}$$

where $\mathbf{w} = \mathbf{w}_{i,B,\pi} \in [0,1]^{S_i}$ is as above.

This time we fix the k variables $y_{11}, y_{22}, \ldots, y_{kk}$ among the $k \times m$ entries of $\mathbf{Y}$. In Eq. (18), using the definition of ρ and a similar argument as for the previous claim, those terms containing exactly these k points as their coordinates are as follows:

- The term: $(-1)^k F_{S_0}(y_{11}, y_{22}, \ldots, y_{kk})$, corresponding to $(\pi(1), \ldots, \pi(k)) = (1, \ldots, k)$ and $B = [k]$.
- For each $1 \leqslant i \leqslant k$, there is one such term: $(-1)^{k-i} F_{S_i}(\mathbf{w})$ with

  $$\mathbf{w}(j) = \begin{cases} y_{j+1,j+1}, & 1 \leqslant j \leqslant i-1, \\ y_{jj}, & i+1 \leqslant j \leqslant k, \\ y_{11}, & j = m, \end{cases}$$

  for $j \in S_i$, corresponding to $B = \{i+1, \ldots, k\}$, and π defined as $\pi(m) = 1$, $\pi(j) = j+1$ for $1 \leqslant j \leqslant i-1$, and $\pi(j) = j$ for $i+1 \leqslant j \leqslant k$. Since F_{S_i} is an alternating function,

  $$(-1)^{k-i} F_{S_i}(\mathbf{w}) = (-1)^{k-i}(-1)^{i-1} F_{S_i}(\mathbf{w}') = (-1)^{k-1} F_{S_i}(\mathbf{w}'),$$

  where for $j \in S_i$, $\mathbf{w}'$ is defined as

  $$\mathbf{w}'(j) = \begin{cases} y_{jj}, & j \neq m, \\ y_{ii}, & j = m. \end{cases}$$

Hence fixing $y_{11}, \ldots, y_{kk}$ and integrating with respect to the other $(m-1)k$ entries of $\mathbf{Y}$, by Definition 2.1 (ii), Eq. (18) reduces to Theorem 3.1 (iii).

5 Proof of Theorem 3.6

In this section we will prove Theorem 3.6.

Theorem 3.6 (restated). *Let $\alpha \in (0,1)$, and $0 \leqslant r \leqslant m$ be an integer. A symmetric integrable function $f : [0,1]^m \to \mathbb{C}$ satisfies $\int_{A^{m-r}\times(\overline{A})^r} f = 0$ for every $A \subset [0,1]$ with $\lambda(A) = \alpha$ if and only if at least one of the following two cases holds.*

(i) $f = 0$ almost everywhere.

(ii) For $K := K(m,r,\alpha) = \{k \in [m] : \sum_{i=0}^{k} \binom{m-r}{k-i}\binom{r}{i}\left(\frac{-\alpha}{1-\alpha}\right)^i = 0\}$, we have

$$f(x_1,\dots,x_m) = \sum_{k\in K}\sum_{S\subseteq[m],|S|=k} g_k(x_S)$$

where $g_k : [0,1]^k \to \mathbb{C}$ are symmetric functions satisfying $\int g_k(x_1,\dots,x_k)dx_i = 0$ for every $i \in [k]$.

Since f is symmetric, the Walsh expansion $f = \sum_{S\subseteq[m]} F_S$ has the following structure. Every F_S is symmetric with respect to the coordinates in S, and furthermore for every $0 \leqslant k \leqslant m$ and every $S \subseteq [m]$ with $|S| = k$, we have

$$F_S(a_1,\dots,a_k) = F_{[k]}(a_1,\dots,a_k).$$

Note that

$$\int_{A^{m-r}\times(\overline{A})^r} F_S = (-1)^{|S\cap\{m-r+1,\dots,m\}|}\left(\frac{1-\alpha}{\alpha}\right)^{r-|S\cap\{m-r+1,\dots,m\}|}\int_{A^m} F_S.$$

We conclude that

$$\int_{A^{m-r}\times(\overline{A})^r} f = \left(\frac{1-\alpha}{\alpha}\right)^r \sum_{k=0}^{m}\sum_{i=0}^{k}\binom{m-r}{k-i}\binom{r}{i}\left(\frac{-\alpha}{1-\alpha}\right)^i\int_{A^m} F_{[k]}.$$

This verifies the "if" part of Theorem 3.6. It remains to prove the "only if" part.

Let

$$F =: \sum_{k=0}^{m}\left(\sum_{i=0}^{k}\binom{m-r}{k-i}\binom{r}{i}\left(\frac{-\alpha}{1-\alpha}\right)^i\right)F_{[k]}.$$

Under the assumption of the theorem, we have $\int_{A^m} F = 0$ for every $A \subseteq [0,1]$ with $\lambda(A) = \alpha$. Now similar to Section 3.1 and the proof of Theorem 3.1, we use the fact that the integral needs to remain 0 under small modifications of A that do not change its measure.

Fix $A \subset [0,1]$ with $\lambda(A) = \alpha$ and nonempty interior and exterior. In order to use the notation of Section 3.1, define $A_1 := \cdots := A_m := A$. Consider $\mathbf{a} = (a_1,\dots,a_k) \in \mathrm{Int}(A)^k$ and $b = (b_1,\dots,b_k) \in \mathrm{Int}(\overline{A})^k$, and let $\mathbf{y}_i := (a_i,\dots,a_i) \in [0,1]^m$ and $\mathbf{z}_i := (b_i,\dots,b_i) \in [0,1]^m$. We conclude that for almost every $a_1,\dots,a_k \in \mathrm{Int}(A)$ and $b_1,\dots,b_k \in \mathrm{Int}(\overline{A})$, we have

$$\int_{A_{[m]}} \partial_{\mathbf{y}_k,\mathbf{z}_k}\cdots\partial_{\mathbf{y}_1,\mathbf{z}_1} F = 0. \tag{19}$$

Claim 5.1. *We have*

$$\partial_{\mathbf{y}_k,\mathbf{z}_k}\cdots\partial_{\mathbf{y}_1,\mathbf{z}_1} F_{[k]} = \alpha^{-k}k!\sum_{B\subseteq[k]}(-1)^{|B|}F_{[k]}(\mathbf{a}_B,\mathbf{b}_{[k]\setminus B}).$$

Furthermore, $\partial_{\mathbf{y}_k,\mathbf{z}_k}\cdots\partial_{\mathbf{y}_1,\mathbf{z}_1} F_{[\ell]} = 0$ if $\ell < k$.

Proof. The claim is an easy consequence of Lemma 3.9. By this lemma, $\partial_{\mathbf{y}_k,\mathbf{z}_k}\cdots\partial_{\mathbf{y}_1,\mathbf{z}_1}F_T = 0$ if $|T| < k$, and moreover

$$\partial_{\mathbf{y}_k,\mathbf{z}_k}\cdots\partial_{\mathbf{y}_1,\mathbf{z}_1}F_{[k]} = \frac{1}{\alpha^k}\sum_{\pi:[k]\xrightarrow{1:1}[k]}\sum_{B\subseteq[k]}(-1)^{|B|}F_{[k]}(\mathbf{w}), \tag{20}$$

where $\mathbf{w} = \mathbf{w}_{B,\pi} \in [0,1]^k$ is defined as

$$\mathbf{w}(t) = \begin{cases} a_{\pi(t)} & \pi(t) \in B, \\ b_{\pi(t)} & \pi(t) \in [k] \setminus B. \end{cases}$$

which by the symmetry of $F_{[k]}$ simplifies to the desired

$$\alpha^{-k}\sum_{B\subseteq[k]}(-1)^{|B|}k!F_{[k]}(\mathbf{a}_B,\mathbf{b}_{[k]\setminus B}).$$

□

Note that in particular we have

$$\int_{[0,1]^k}\left(\partial_{\mathbf{y}_k,\mathbf{z}_k}\cdots\partial_{\mathbf{y}_1,\mathbf{z}_1}F_{[k]}\right)da_1\cdots da_k = \alpha^{-k}k!F_{[k]}(\mathbf{b}). \tag{21}$$

Suppose for the sake of contradiction that the statement of the theorem is not true. Then there exists a largest $k \in [m]\setminus K(m,r,\alpha)$ such that $F_{[k]}$ is not zero almost everywhere. By Eq. (19) and Eq. (21) we have

$$\alpha^{m-k}k!\left(\sum_{i=0}^{k}\binom{m-r}{k-i}\binom{r}{i}\left(\frac{-\alpha}{1-\alpha}\right)^i\right)F_{[k]}(\mathbf{b}) = 0,$$

for almost all $\mathbf{b}$, which then implies that $\sum_{i=0}^{k}\binom{m-r}{k-i}\binom{r}{i}\left(\frac{-\alpha}{1-\alpha}\right)^i = 0$ as $F_{[k]}$ is not zero almost everywhere. But this means that $k \in K(m,r,\alpha)$, which is a contradiction.

6 Concluding Remarks

One of the main problems studied in the paper of Janson and Sós [JS14] is determining for which $(F,\alpha_1,\ldots,\alpha_m)$, the property $\mathcal{P}(F,\alpha_1,\ldots,\alpha_m)$ is always (i.e. for every $p \in (0,1]$) a *quasi-random property.* The only known example for which this is *not* the case is $\mathcal{P}(K_2,\frac{1}{2},\frac{1}{2})$. This fact was already observed by Chung and Graham in [CG92]. In the same paper, they also showed that $\mathcal{P}(K_2,\alpha,1-\alpha)$ is a quasi-random property for every $\alpha \in (0,1)\setminus\{\frac{1}{2}\}$.

Conjecture 6.1 (See [JS14, Conjecture 2.13 and Problem 2.16])**.** *Let $F \neq K_2$ be a non-empty graph with vertex set $\{1,\ldots,m\}$, and let $\alpha_1,\ldots,\alpha_m \in (0,1)$ satisfy $\sum_{i=1}^{m}\alpha_i \leqslant 1$. Then $\mathcal{P}(F,\alpha_1,\ldots,\alpha_m)$ is a quasi-random property for every $p \in (0,1]$.*

When $\sum_{i=1}^{m}\alpha_i < 1$, Conjecture 6.1 is verified in Corollary 3.4 (originally proved in [JS14, Theorem 2.11]). The case where $\alpha_1 = \cdots = \alpha_m = \frac{1}{m}$ and F is a regular graph, a star, or a disconnected graph with at least one edge is verified by Janson and Sós in [JS14, Theorem 2.12].

Our Theorem 3.5 settles the case when α_i are arbitrary and F contains twin vertices. Prior to our work this was unknown even for the path on 3 vertices and was stated as an open problem in [JS14, Problem 2.19].

Acknowledgements

The authors are grateful to Svante Janson for his invaluable comments and corrections to an earlier version of the paper. The second author would like to thank Alexander Razborov and Madhur Tulsiani for their helpful comments on the presentation of our main result.

References

[CG92] F. R. K. Chung and R. L. Graham, *Maximum cuts and quasirandom graphs*, Random graphs, Vol. 2 (Poznań, 1989), Wiley-Intersci. Publ., Wiley, New York, 1992, pp. 23–33. MR 1166604 (93b:05149) 13

[CGW89] F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, Combinatorica **9** (1989), no. 4, 345–362. MR 1054011 (91e:05074) 1, 2

[ES81] B. Efron and C. Stein, *The jackknife estimate of variance*, Ann. Statist. **9** (1981), no. 3, 586–596. MR 615434 (82k:62074) 3

[Hoe48] Wassily Hoeffding, *A class of statistics with asymptotically normal distribution*, Ann. Math. Statistics **19** (1948), 293–325. MR 0026294 (10,134g) 3

[JS14] Svante Janson and Vera T Sós, *More on quasi-random graphs, subgraph counts and graph limits*, arXiv preprint arXiv:1405.6808 (2014). 1, 2, 3, 4, 5, 6, 7, 13

[LS06] László Lovász and Balázs Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957. MR 2274085 (2007m:05132) 2

[Sha08] Asaf Shapira, *Quasi-randomness and the distribution of copies of a fixed graph*, Combinatorica **28** (2008), no. 6, 735–745. MR 2488748 (2010b:05156) 2

[SS97] Miklós Simonovits and Vera T. Sós, *Hereditarily extended properties, quasi-random graphs and not necessarily induced subgraphs*, Combinatorica **17** (1997), no. 4, 577–596. MR 1645698 (99j:05115) 2

[SY10] Asaf Shapira and Raphael Yuster, *The effect of induced subgraphs on quasi-randomness*, Random Structures Algorithms **36** (2010), no. 1, 90–109. MR 2591049 (2011a:05320) 2

[Tho87a] Andrew Thomason, *Pseudorandom graphs*, Random graphs ’85 (Poznań, 1985), North-Holland Math. Stud., vol. 144, North-Holland, Amsterdam, 1987, pp. 307–331. MR 930498 (89d:05158) 1

[Tho87b] ______, *Random graphs, strongly regular graphs and pseudorandom graphs*, Surveys in combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 173–195. MR 905280 (88m:05072) 1

[Yus10] Raphael Yuster, *Quasi-randomness is determined by the distribution of copies of a fixed graph in equicardinal large sets*, Combinatorica **30** (2010), no. 2, 239–246. MR 2676838 (2011j:05302) 2