

MORE NON-BIPARTITE FORCING PAIRS

T. HUBAI, D. KRÁL', O. PARCZYK AND Y. PERSON

ABSTRACT. We study pairs of graphs (H_1, H_2) such that every graph with the densities of H_1 and H_2 close to the densities of H_1 and H_2 in a random graph is quasirandom; such pairs (H_1, H_2) are called forcing. Non-bipartite forcing pairs were first discovered by Conlon, Hàn, Person and Schacht [*Weak quasi-randomness for uniform hypergraphs*, Random Structures Algorithms **40** (2012), no. 1, 1–38]: they showed that (K_t, F) is forcing where F is the graph that arises from K_t by iteratively doubling its vertices and edges in a prescribed way t times. Reiher and Schacht [*Forcing quasirandomness with triangles*, Forum of Mathematics, Sigma. Vol. 7, 2019] strengthened this result for $t = 3$ by proving that two doublings suffice and asked for the minimum number of doublings needed for $t > 3$. We show that $\lceil (t + 1)/2 \rceil$ doublings always suffice.

1. INTRODUCTION AND RESULTS

The systematic study of quasirandom graphs has been initiated by Thomason [14, 15] and Chung, Graham and Wilson [1] in the 1980's. Since then, many properties of quasirandom graphs were described. We refer to the surveys [6, 7].

A key property of a quasirandom graph is an almost uniform edge distribution. A sequence $(G_n)_{n \in \mathbb{N}}$ of graphs is p -quasirandom if $e_{G_n}(U) = p \binom{|U|}{2} + o(|G_n|^2)$ for all subsets $U \subseteq V(G_n)$, where $|G_n|$ is the number of vertices of G_n and $e_{G_n}(U)$ is the number of edges of G_n with both end vertices in U . A particular graph G with n vertices is (ε, p) -quasirandom if $\left| e_G(U) - p \binom{|U|}{2} \right| \leq \varepsilon n^2$ for all subsets $U \subseteq V(G)$.

One of many equivalent characterizations of p -quasirandom sequences $(G_n)_{n \in \mathbb{N}}$ of graphs is the following: $(G_n)_{n \in \mathbb{N}}$ is quasirandom if and only if G_n has edge density p and contains $(p^4 + o(1))|G_n|^4$ labelled (non-induced) copies of C_4 . Equivalently, G_n contains asymptotically the expected number of copies of K_2 and C_4 as the Erdős-Rényi random graph $G(n, p)$. This leads to the definition of a forcing

Received June 7, 2019.

2010 *Mathematics Subject Classification*. Primary 05C35.

The first author was supported by the Engineering and Physical Sciences Research Council Standard Grant number EP/M025365/1. The second author was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 648509). This publication reflects only its authors' view; the European Research Council Executive Agency is not responsible for any use that may be made of the information it contains. The third and fourth authors were supported by DFG grant PE 2299/1-1.

pair of graphs given below. To give the definition, we need to introduce the following notation. If F and G are two graphs, then $t(F, G)$ is the number of graph homomorphisms from F to G , i.e. all maps $f: V(F) \rightarrow V(G)$ with $f(e) \in E(G)$ for all $e \in E(F)$. In addition, we write $e(F)$ for the number of edges of F .

Definition 1.1 (Forcing pairs). *A pair (F_1, F_2) is called forcing if for every $p \in (0, 1]$ and $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds. Every graph G with*

$$t(F_1, G) = (1 \pm \delta)p^{e(F_1)} \quad \text{and} \quad t(F_2, G) = (1 \pm \delta)p^{e(F_2)}$$

is (ε, p) -quasirandom.

In particular, the pair (K_2, C_4) is forcing. There are two exciting conjectures related to forcing pairs: Sidorenko's conjecture made independently by Sidorenko [11] and by Erdős and Simonovits [12], and the so-called forcing conjecture made by Skokan and Thoma [13]. While Sidorenko's conjecture asks whether the lower bound on $t(F_2, G)$ is always at least $p^{e(F_2)}$ where $p = t(K_2, G)$, the forcing conjecture states that any pair (K_2, F) , where F is a bipartite graph containing a cycle, is forcing. Due to their relation to Szemerédi's regularity lemma, these conjectures expedited tremendous amount of research in extremal combinatorics. Thus, additional forcing pairs were studied in [1, 2, 3, 4, 5, 9, 13]. The first non-bipartite forcing pairs were found by Conlon et al. in [3]. So far, all non-bipartite forcing pairs are obtained by the construction described below.

Let F be a t -partite graph and $V(F) = V_1(F) \dot{\cup} \dots \dot{\cup} V_t(F)$ a fixed t -coloring of F . The doubling $\mathcal{T}(F)$ on $V_1(F)$ is the graph obtained by taking two identical disjoint copies F_1 and F_2 of F and identifying the corresponding vertices in $V_1(F_1)$ and $V_1(F_2)$. In this way, we obtain a t -coloring of $\mathcal{T}(F)$ given by the sets $V_1(F_1) = V_1(F_2)$, $V_2(F_1) \dot{\cup} V_2(F_2)$, \dots , $V_t(F_1) \dot{\cup} V_t(F_2)$. For $k \leq t$, the k -fold doubling $\mathcal{T}_k(F)$ is defined as the doubling $\mathcal{T}(F)$ on $V_1(F)$ for $k = 1$ and the doubling $\mathcal{T}(\mathcal{T}_{k-1}(F))$ on $V_k(\mathcal{T}_{k-1}(F))$ for $k \geq 2$. The order of the doublings has no influence on $\mathcal{T}_k(F)$, i.e. we could permute $V_1(F), \dots, V_k(F)$ arbitrarily. Observe that $\mathcal{T}_2(K_2) = C_4$.

The pair $(K_2, C_4) = (K_2, \mathcal{T}_2(K_2))$ is forcing. The result from [3] states that the pair $(K_t, \mathcal{T}_t(K_t))$ is also forcing for any $t \geq 3$. Hàn et al. [4] generalized this result for any t -colorable graph F in a similar way. Reiher and Schacht [9] improved the result from [3] for $t = 3$ by showing that the pair $(K_3, \mathcal{T}_2(K_3))$ is forcing. We generalize this result for $t > 3$; we note that the same was independently proven by Reiher and Schacht [10].

Theorem 1.2. *The pair $(K_t, \mathcal{T}_{\lceil (t+1)/2 \rceil}(K_t))$ is forcing for any $t \geq 2$.*

We will present the proof in the language of graph limits, which we now introduce, since this makes the arguments particularly short and transparent. Let W be a *kernel*, i.e. a bounded symmetric Lebesgue measurable function from $[0, 1]^2$. We write $W \equiv p$ if W is equal to p almost everywhere; a kernel with $W \equiv p$ is called *p -quasirandom*. The homomorphism density extends in a natural way for a

graph F and a kernel W :

$$t(F, W) = \int_{[0,1]^{V(F)}} \prod_{uv \in E(F)} W(x_u, x_v) \prod_{u \in V(F)} dx_u.$$

A *graphon* is a kernel W with values in $[0, 1]$. A pair of graphs (F_1, F_2) is called *forcing* if for every real $p \in (0, 1]$, every graphon W with $t(F_1, W) = p^{e(F_1)}$ and $t(F_2, W) = p^{e(F_2)}$ is p -quasirandom. This definition, see [8, Chapter 16], coincides with the definition of a forcing pair given earlier.

2. PROOF OF THEOREM 1.2 FOR $t = 4$

In this section, we give a proof of Theorem 1.2 for K_4 . We need the following lemma.

Lemma 2.1 (Lemma 10 from [3]). *Let W be a graphon and $p \in (0, 1]$ such that*

$$t(K_4, W) = p^6 \quad \text{and} \quad t(\mathcal{T}_3(K_4), W) = p^{48}.$$

Then

$$W(x_1, x_2)W(x_1, x_3)W(x_2, x_3) \int_{[0,1]} W(x_1, y)W(x_2, y)W(x_3, y)dy = p^6$$

for almost all $(x_1, x_2, x_3) \in [0, 1]^3$. □

The proof of this lemma is given in [3] in the language of quasirandom (hyper)graphs, and we sketch the line of arguments here for completeness. It can be proved by repeatedly applying Cauchy-Schwarz inequality starting to $t(K_4, W)$ three times. This series of applications of Cauchy-Schwarz inequality yields that $t(\mathcal{T}_3(K_4), W) \geq p^{48}$. Since it holds $t(\mathcal{T}_3(K_4), W) = p^{48}$ by the assumption of the lemma, it follows that for almost all values of $(x_1, x_2, x_3) \in [0, 1]^3$ one has

$$W(x_1, x_2)W(x_1, x_3)W(x_2, x_3) \int_{[0,1]} W(x_1, y)W(x_2, y)W(x_3, y)dy = p^6.$$

The next lemma together with Lemma 2.1 readily implies Theorem 1.2 for $k = 4$.

Lemma 2.2. *Let W be a graphon and $p \in (0, 1]$. If it holds that*

$$(1) \quad W(x_1, x_2)W(x_1, x_3)W(x_2, x_3) \int_{[0,1]} W(x_1, y)W(x_2, y)W(x_3, y)dy = p^6$$

for almost all $(x_1, x_2, x_3) \in [0, 1]^3$, then W is p -quasirandom.

Before presenting the proof, we recall the definition of the *essential supremum* $\text{ess sup}(f)$ of a (Lebesgue) measurable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$. It is the infimum over all $y \in \mathbb{R}$ with $f(x) \leq y$ for almost all $x \in \mathbb{R}^n$, i.e. $\text{ess sup}(f) := \inf \{y: \lambda(\{f \geq y\}) = 0\}$, where λ is the Lebesgue measure. The *essential infimum* of a function is defined analogously.

Proof of Lemma 2.2. Let $f: [0, 1] \rightarrow [0, 1]$ be defined as

$$f(x) = \sup_{y_1, y_2 \in [0, 1]} (W(x, y_1) - W(x, y_2)).$$

Observe that f is a measurable function and set $c := \text{ess sup } f \in [0, 1]$. If $c = 0$, then W is p -quasirandom. Thus, we assume that $c > 0$.

The definition of f and c implies that there exist reals $a, b \in \mathbb{R}$ with $c = b - a$ satisfying the following. For any $\eta > 0$, there exist $x_1, x_2, x_3, x_4 \in [0, 1]$ such that

$$(2) \quad \begin{aligned} W(x_1, x_2) &\geq b - \eta, & W(x_1, x_3) &\geq b - \eta, \\ W(x_2, x_4) &\leq a + \eta, & W(x_3, x_4) &\leq a + \eta, \end{aligned}$$

and (1) holds for (x_1, x_2, x_3) and (x_2, x_3, x_4) . In addition, we can assume that $a > 0$, $W(x_2, x_3) > 0$, $\text{ess sup}_{y \in [0, 1]} W(x_2, y)W(x_3, y) > 0$ (because $p > 0$), and that

$$(3) \quad W(x_1, y) \geq a - \eta \quad \text{and} \quad W(x_4, y) \leq b + \eta \quad \text{for almost all } y \in [0, 1].$$

We get from (1) that

$$\begin{aligned} &W(x_1, x_2)W(x_1, x_3)W(x_2, x_3) \int_{[0, 1]} W(x_1, y)W(x_2, y)W(x_3, y)dy \\ &= p^6 = W(x_2, x_4)W(x_3, x_4)W(x_2, x_3) \int_{[0, 1]} W(x_4, y)W(x_2, y)W(x_3, y)dy \end{aligned}$$

Using (2) and (3) we can lower bound the left hand side by

$$(b - \eta)^2(a - \eta)W(x_2, x_3) \int_{[0, 1]} W(x_2, y)W(x_3, y)dy$$

and similarly we can upper bound the right hand side by

$$(a + \eta)^2(b + \eta)W(x_2, x_3) \int_{[0, 1]} W(x_2, y)W(x_3, y)dy.$$

As $W(x_2, x_3) \int_{[0, 1]} W(x_2, y)W(x_3, y)dy$ is non-zero, we obtain that

$$(b - \eta)^2(a - \eta) \leq (a + \eta)^2(b + \eta)$$

and then deduce that $b \leq a + 10\eta(b/a)$. Together with the assumption that $b > a$ this implies that $W(x, y) \in [p \pm 20\eta(b/a)]$ for almost all $(x, y) \in [0, 1]^2$. Since this holds for every $\eta > 0$, the lemma follows. \square

3. PROOF OF THEOREM 1.2 – GENERAL CASE

The proof of the general case is based on the same idea as used in the previous section and follows from the next two lemmas. The first lemma can be proven by repeated applications of the Cauchy-Schwarz inequality similarly to the proof of Lemma 2.1.

Lemma 3.1. *Let W be a graphon, $p \in (0, 1]$, $t \geq 3$, and $k := \lceil (t + 1)/2 \rceil$ such that*

$$t(K_t, W) = p^{e(K_t)} \quad \text{and} \quad t(\mathcal{T}_k(K_t), W) = p^{e(\mathcal{T}_k(K_t))}.$$

Then

$$p^{e(K_t)} = \left(\prod_{1 \leq i < j \leq k} W(x_i, x_j) \right) \cdot \int_{[0,1]^{t-k}} \left(\prod_{1 \leq i < j \leq t-k} W(y_i, y_j) \right) \left(\prod_{i \in [k], j \in [t-k]} W(x_i, y_j) \right) \prod_{i \in [t-k]} dy_i.$$

for almost all $(x_1, \dots, x_k) \in [0, 1]^k$.

Lemma 3.2. *Let W be a graphon, $p \in (0, 1]$, $t \geq 3$, and $k := \lceil (t + 1)/2 \rceil$. If it holds that*

$$(4) \quad p^{e(K_t)} = \left(\prod_{1 \leq i < j \leq k} W(x_i, x_j) \right) \cdot \int_{[0,1]^{t-k}} \left(\prod_{1 \leq i < j \leq t-k} W(y_i, y_j) \right) \left(\prod_{i \in [k], j \in [t-k]} W(x_i, y_j) \right) \prod_{i \in [t-k]} dy_i.$$

for almost all $(x_1, \dots, x_k) \in [0, 1]^k$, then W is p -quasirandom.

Proof. Let $f: [0, 1] \rightarrow [0, 1]$ be defined as

$$f(x) = \sup_{y_1, y_2 \in [0,1]} (W(x, y_1) - W(x, y_2)).$$

Again, f is a measurable function and we set $c := \text{ess sup } f \in [0, 1]$. If $c = 0$, then W is p -quasirandom. So, we assume that $c > 0$ and consider positive reals $a, b \in \mathbb{R}$ with $c = b - a$ such that the following holds. For any $\eta > 0$, there exist $x_1, \dots, x_{k+1} \in [0, 1]$ such that

$$(5) \quad W(x_1, x_i) \geq b - \eta \quad \text{and} \quad W(x_{k+1}, x_i) \leq a + \eta \quad \text{for all } i = 2, \dots, k,$$

for (x_1, \dots, x_k) and (x_2, \dots, x_{k+1}) equation (4) holds,

$$(6) \quad W(x_1, y), W(x_{k+1}, y) \in [a - \eta, b + \eta] \text{ for almost all } y \in [0, 1],$$

and Q is non-zero, where

$$Q = \left(\prod_{2 \leq i < j \leq k} W(x_i, x_j) \right) \cdot \int_{[0,1]^{t-k}} \left(\prod_{1 \leq i < j \leq t-k} W(y_i, y_j) \right) \left(\prod_{2 \leq i \leq k, j \in [t-k]} W(x_i, y_j) \right) \prod_{i \in [t-k]} dy_i.$$

We get from (4) that

$$\begin{aligned} & \left(\prod_{1 \leq i < j \leq k} W(x_i, x_j) \right) \\ & \cdot \int_{[0,1]^{t-k}} \left(\prod_{1 \leq i < j \leq t-k} W(y_i, y_j) \right) \left(\prod_{1 \leq i \leq k, j \in [t-k]} W(x_i, y_j) \right) \prod_{i \in [t-k]} dy_i \\ & = p^{e(K_t)} = \left(\prod_{2 \leq i < j \leq k+1} W(x_i, x_j) \right) \\ & \cdot \int_{[0,1]^{t-k}} \left(\prod_{1 \leq i < j \leq t-k} W(y_i, y_j) \right) \left(\prod_{2 \leq i \leq k+1, j \in [t-k]} W(x_i, y_j) \right) \prod_{i \in [t-k]} dy_i. \end{aligned}$$

Using (5) and (6) we can lower bound the left hand side by

$$(7) \quad (b - \eta)^{k-1} (a - \eta)^{t-k} Q$$

and similarly we can upper bound the right hand side by

$$(8) \quad (a + \eta)^{k-1} (b + \eta)^{t-k} Q.$$

As Q is non-zero this gives

$$(b - \eta)^{k-1} (a - \eta)^{t-k} \leq (a + \eta)^{k-1} (b + \eta)^{t-k}.$$

If t is even, we have $k - 1 > t - k$ and we can finish the proof similar to the case $t = 4$. If t is odd, i.e. $k - 1 = t - k$, a more refined argument is needed.

Since W is a graphon, the difference of (7) and (8) is at most $2t\eta$. In particular, the estimates used to derive (7) cannot be too wasteful and it follows that

$$W(x_1, y) \leq a + 2t\eta \quad \text{and} \quad W(x_{k+1}, y) \geq b - 2t\eta \quad \text{for almost all } y \in [0, 1].$$

Together with (6) we get

$$W(x_1, y) \in [a \pm 2t\eta] \quad \text{and} \quad W(x_{k+1}, y) \in [b \pm 2t\eta] \quad \text{for almost all } y \in [0, 1].$$

It follows that $b - a \leq 2(t + 1)\eta$ and consequently $W(x, y) \in [p \pm 4(t + 1)\eta]$ for almost all $(x, y) \in [0, 1]^2$. \square

REFERENCES

1. Chung F. R. K., Graham R. L. and Wilson R. M., *Quasi-random graphs*, *Combinatorica* **9** (1989), 345–362.
2. Conlon D., Fox J. and Sudakov B., *An approximate version of Sidorenko's conjecture*, *Geom. Funct. Anal.* **20** (2010), 1354–1366.
3. Conlon D., Hàn H., Person Y. and Schacht M., *Weak quasi-randomness for uniform hypergraphs*, *Random Structures Algorithms* **40** (2012), 1–38.
4. Hàn H., Person Y. and Schacht M., *Note on forcing pairs*, *Electron. Notes Discrete Math.* **38** (2011), 437–442.
5. Hatami H., *Graph norms and Sidorenko's conjecture*, *Israel J. Math.* (2010), 125–150.
6. Komlós J., Shokoufandeh A., Simonovits M. and Szemerédi E., *The regularity lemma and its applications in graph theory*, in: *Theoretical aspects of computer science (Tehran, 2000)*, *Lecture Notes in Comput. Sci.* 2292, Springer, Berlin, 2002.

7. Krivelevich M. and Sudakov B., *Pseudo-random graphs*, in: More sets, graphs and numbers, Bolyai Soc. Math. Stud., vol. 15, Springer, Berlin, 2006.
8. Lovász L., *Large Networks and Graph limits*, vol. 60, Amer. Math. Soc., 2012.
9. Reiher C. and Schacht M., *Forcing quasirandomness with triangles*, in: Forum of Mathematics, Sigma. Vol. 7, 2019.
10. Schacht M., personal communication.
11. Sidorenko A., *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), 201–204.
12. Simonovits M., *Extremal graph problems, degenerate extremal problems, and supersaturated graphs*, in: Progress in Graph Theory (Waterloo, Ont., 1982), Academic Press, Toronto, ON, 1984.
13. Skokan J. and Thoma L., *Bipartite subgraphs and quasi-randomness*, Graphs Combin. **20** (2004), 255–262.
14. Thomason A., *Pseudorandom graphs*, in: Random graphs '85 (Poznań, 1985), North-Holland Math. Stud. 144, North-Holland, 1987.
15. Thomason A., *Random graphs, strongly regular graphs and pseudorandom graphs*, in: Surveys in Combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser. 123, Cambridge Univ. Press, Cambridge, 1987.

T. Hubai, Institute of Mathematics, Eötvös Loránd University, Budapest, Hungary
Previous affiliation: Department of Computer Science and DIMAP, University of Warwick, Coventry, UK,
e-mail: htamas@cs.elte.hu

D. Král', Faculty of Informatics, Masaryk University, Brno, Czech Republic,
Mathematics Institute, DIMAP and Department of Computer Science, University of Warwick, Coventry, UK,
e-mail: dkral@fi.muni.cz

O. Parczyk, Institut für Mathematik, Technische Universität Ilmenau, Ilmenau, Germany,
e-mail: olaf.parczyk@tu-ilmenau.de

Y. Person, Institut für Mathematik, Technische Universität Ilmenau, Ilmenau, Germany,
e-mail: yury.person@tu-ilmenau.de