Combinatorics, Probability and Computing (2024), **33**, pp. 16–31
doi:10.1017/S0963548323000263

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Forcing generalised quasirandom graphs efficiently*

Andrzej Grzesik[1,†], Daniel Kráľ[2,‡] and Oleg Pikhurko[3,§]

[1]Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland, [2]Faculty of Informatics, Masaryk University, Brno, Czech Republic, and [3]Mathematics Institute and DIMAP, University of Warwick, Coventry, UK
Corresponding author: Oleg Pikhurko; Email: o.pikhurko@warwick.ac.uk

(Received 11 May 2023; revised 27 June 2023; accepted 20 July 2023; first published online 5 September 2023)

Abstract
We study generalised quasirandom graphs whose vertex set consists of q parts (of not necessarily the same sizes) with edges within each part and between each pair of parts distributed quasirandomly; such graphs correspond to the stochastic block model studied in statistics and network science. Lovász and Sós showed that the structure of such graphs is forced by homomorphism densities of graphs with at most $(10q)^q + q$ vertices; subsequently, Lovász refined the argument to show that graphs with $4(2q+3)^8$ vertices suffice. Our results imply that the structure of generalised quasirandom graphs with $q \geq 2$ parts is forced by homomorphism densities of graphs with at most $4q^2 - q$ vertices, and, if vertices in distinct parts have distinct degrees, then $2q+1$ vertices suffice. The latter improves the bound of $8q-4$ due to Spencer.

Keywords: Graph limits; Graphons; Homomorphism density; Quasirandomness

2020 MSC Codes: Primary: 05C60

1. Introduction

Quasirandom graphs play an important role in structural and extremal graph theory. The notion of quasirandom graphs can be traced to the works of Rödl [42], Thomason [46, 47] and Chung, Graham, and Wilson [9] in the 1980s and is also deeply related to Szemerédi's Regularity Lemma [44]. Indeed, the Regularity Lemma asserts that each graph can be approximated by partitioning it into a bounded number of quasirandom bipartite graphs. There is also a large body of literature concerning quasirandomness of various kinds of combinatorial structures such as groups [24], hypergraphs [5, 6, 22, 23, 29, 32, 41, 43], permutations [4, 10, 34, 35], Latin squares [11, 17, 20, 25], subsets of integers [8], tournaments [3, 7, 13, 14, 26, 28], etc. Many of these notions have been treated in a unified way in the recent paper by Coregliano and Razborov [15].

The starting point of our work is the following classical result on quasirandom graphs [9]: a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ is quasirandom with density p if and only if the homomorphism densities of the single edge K_2 and the 4-cycle C_4 in $(G_n)_{n\in\mathbb{N}}$ converge to p and p^4, that is, to their expected densities in the Erdős-Rényi random graph with density p. In particular, quasirandomness is forced by homomorphism densities of graphs with at most 4 vertices. In this paper, we consider a generalisation of quasirandom graphs, which corresponds to the stochastic block

*An extended abstract announcing the results presented in this paper has been published in the Proceedings of Eurocomb'23.
†Andrzej Grzesik: Supported by the National Science Centre grant number 2021/42/E/ST1/00193.
‡Daniel Kráľ: Supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University.
§Oleg Pikhurko: Supported by ERC Advanced Grant 101020255 and Leverhulme Research Project Grant RPG-2018-424.

model in statistics. In this model, the edge density of a (large) graph is not homogeneous as in the Erdős-Rényi random graph model, however, the graph can be partitioned into q parts such that the edge density is homogeneous inside each part and between each pair of the parts. Lovász and Sós [37] established that the structure of such graphs is forced by homomorphism densities of graphs with at most $(10q)^q + q$ vertices. Lovász [36, Theorem 5.33] refined this result by showing that homomorphism densities of graphs with at most $4(2q+3)^8$ vertices suffice. Our main result, which we state below (we refer to Section 2 for not yet defined notation), improves this bound: the structure of generalised quasirandom graphs with $q \geq 2$ parts is forced by homomorphism densities of graphs with at most $4q^2 - q$ vertices.

Theorem 1. *The following holds for every $q \geq 2$ and every q-step graphon W: if the density of each graph with at most $4q^2 - q$ vertices in a graphon W' is the same as in W, then the graphons W and W' are weakly isomorphic.*

We remark that our line of arguments to prove Theorem 1 substantially differs from that in [36, 37], with the exception of initial application of Lemma 2. In particular, the key steps in our proof are more explicit and so of a more constructive nature, which is of importance in relation to applications [2, 19, 30, 31].

Spencer [45] considered generalised quasirandom graphs with q parts with an additional assumption that vertices in distinct parts have distinct degrees and established that the structure of such graphs is forced by homomorphism densities of graphs with at most $8q - 4$ vertices. Addressing a question posed in [45], we show (Theorem 11) that graphs with at most $2q + 1$ vertices suffice in this restricted setting for any $q \geq 2$.

We present our results and arguments using the language of the theory of graph limits, which is introduced in Section 2. We remark that similarly to arguments presented in [36, 37], although not explicitly stated there, our arguments also apply in a more general setting of kernels in addition to graphons (see Section 2 for the definitions of the two notions). We present various auxiliary results in Section 3 and use them to prove our main result in Section 4. The case with the additional assumption that vertices in distinct parts have distinct degrees is analysed in Section 5.

2. Notation

We now introduce the notions and tools from the theory of graph limits that we need in our arguments; we refer the reader to the monograph by Lovász [36] for a more comprehensive introduction and further details. We also rephrase results concerning quasirandom graphs and generalised quasirandom graphs with q parts presented in Section 1 in the language of the theory of graph limits.

We start with fixing some general shorthand notation used throughout the paper. The set of the first q positive integers is denoted by $[q]$ and more generally the set of integers between a and b (inclusive) is denoted by $[a, b]$. If H and G are two graphs, the *homomorphism density* of H in G, denoted by $t(H, G)$, is the probability that a random function $f : V(H) \to V(G)$, with all $|V(G)|^{|V(H)|}$ choices being equally likely, is a *homomorphism* of H to G, that is, $f(u)f(v)$ is an edge of H for every edge uv of G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *convergent* if the number of vertices of G_n tends to infinity and the values of $t(H, G_n)$ converge for every graph H as $n \to \infty$. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *quasirandom with density p* if it is convergent and the limit of $t(H, G_n)$ is equal to $p^{|E(H)|}$ for every graph H, where $E(H)$ denotes the edge set of H. If the particular value of p is irrelevant or understood, we just say that a sequence of graphs is *quasirandom* instead of quasirandom with density p.

The theory of graph limits provides analytic ways of representing sequences of convergent graphs. A *kernel* is a bounded measurable function $U : [0, 1]^2 \to \mathbb{R}$ that is *symmetric*, that is,

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

$U(x, y) = U(y, x)$ for all $(x, y) \in [0, 1]^2$. A *graphon* is a kernel whose values are restricted to $[0, 1]$. The *homomorphism density* of a graph H in a kernel U is defined as follows:

$$t(H, U) = \int_{[0,1]^{V(H)}} \prod_{uv \in E(H)} U(x_u, x_v) \mathrm{d}x_{V(H)},$$

where $\mathrm{d}x_A$ for a set $A = \{a_1, \ldots, a_k\}$ is a shorthand for $\mathrm{d}x_{a_1} \ldots \mathrm{d}x_{a_k}$; we often just briefly say the *density* of a graph H in a kernel U rather than the homomorphism density of H in U. A graphon W is a *limit* of a convergent sequence $(G_n)_{n \in \mathbb{N}}$ of graphs if $t(H, W)$ is the limit of $t(H, G_n)$ for every graph H. Every convergent sequence of graphs has a limit graphon and every graphon is a limit of a convergent sequence of graphs as shown by Lovász and Szegedy [38]; also see [16] for relation to exchangeable arrays. Two kernels (or graphons) U_1 and U_2 are *weakly isomorphic* if $t(H, U_1) = t(H, U_2)$ for every graph H. Note that any two limits of the same convergent sequence of graphs are weakly isomorphic, and we refer particularly to [1] for results on the structure of weakly isomorphic graphons and more generally kernels.

We phrase the results concerning quasirandom graphs using the language of the theory of graph limits. Observe that a sequence of graphs is quasirandom if and only if it converges to the graphon equal to p everywhere. The following holds for every graphon W and every real $p \in [0, 1]$: a graphon W is weakly isomorphic to the constant graphon equal to p if and only if $t(K_2, W) = p$ and $t(C_4, W) = p^4$. This leads us to the following definition: a graphon U is *forced* by graphs contained in a set $\mathcal{H}$ if every graphon U' such that $t(H, U') = t(H, U)$ for every graph $H \in \mathcal{H}$ is weakly isomorphic to U. In particular, any constant graphon is forced by the graphs K_2 and C_4. We refer particularly to [12, 27, 33, 39] for results on the structure of graphons forced by finite sets of graphs. Similarly, we say that a kernel U is *forced* by graphs from a set $\mathcal{H}$ if every kernel U' such that $t(H, U') = t(H, U)$ for every graph $H \in \mathcal{H}$ is weakly isomorphic to U. We emphasise that our results actually concern forcing kernels (rather than graphons), which makes them formally stronger.

A *q-step kernel* U is a kernel such that $[0, 1]$ can be partitioned into q non-null measurable sets $A_1, \ldots, A_q$ such that U is constant on $A_i \times A_j$ for all $i, j \in [q]$ but there is no such partition with $q - 1$ parts. A *q-step graphon* is a q-step kernel that is also a graphon. If the number of parts is not important, we use a *step kernel* or a *step graphon* for brevity. Observe that step graphons correspond to stochastic block models and so to generalised quasirandom graphs discussed in Section 1. As mentioned in Section 1, Lovász and Sós [37, Theorem 2.3] showed that every q-step graphon W is forced by graphs with at most $(10q)^q + q$ vertices, and Lovász [36, Theorem 5.33] further improved the bound on the number of vertices to $4(2q + 3)^8$; we remark that the proof of either of the results can be adapted to the setting of step kernels. Our main result (Theorem 1) states that every q-step graphon is forced by graphs with at most $\max\{4q^2 - q, 4\}$ vertices; our arguments also apply in the setting of step kernels as stated in Theorem 10.

In the rest of this section, we introduce some technical notation needed to present our arguments. A *k-rooted graph* is a graph with k distinguished pairwise distinct vertices, and more generally an $(s_1, \ldots, s_q)$*-rooted graph* is an $(s_1 + \cdots + s_q)$-rooted graph whose roots are split into q groups, each of size s_i, $i \in [q]$. If H is a k-rooted graph with vertices $v_1, \ldots, v_n$ such that its roots are $v_1, \ldots, v_k$ then the *density* of H in a kernel U when $x_1, \ldots, x_k \in [0, 1]$ are chosen as the roots is defined as:

$$t_{x_1,\ldots,x_k}(H, U) = \int_{[0,1]^{n-k}} \prod_{v_i v_j \in E(H)} U(v_i, v_j) \mathrm{d}x_{[k+1,n]}.$$

By the Fubini–Tonelli Theorem, the integral exists for almost all choices of $x_1, \ldots, x_k$ and we will often ignore exceptional null sets in this paper. Note that for $k = 0$ this definition coincides with the definition of the density of an unrooted graph in a kernel. If the particular choice of the roots is understood, we write $t_\star(H, U)$ instead of $t_{x_1,\ldots,x_k}(H, U)$. We sometimes think of and refer to the

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

elements of $[0, 1]$ as *vertices* of a kernel, which justifies the definition of the density of a rooted graph in a kernel and leads to the following definition: the *degree* of a vertex $x \in [0, 1]$ in a kernel U is the density $t_x(K_2^\bullet, U) = \int_0^1 U(x, y)\mathrm{d}y$, where $K_2^\bullet$ is the 1-rooted graph obtained from K_2 by choosing one of its vertices as the root.

A *quantum graph* is a formal finite linear combination $Q = \sum_{i=1}^m c_i H_i$ of graphs; a graph H_i with $c_i \neq 0$ is called a *constituent* of Q. More generally a *quantum k-rooted graph* is a formal finite linear combination of k-rooted graphs such that their roots induce the same (k-vertex) subgraph in each of the constituents. The *density* of a (rooted) quantum graph Q in a kernel U is the corresponding linear combination of the densities of the constituents forming Q.

For a k-rooted graph H, let $[\![H]\!]$ be the underlying unrooted graph. Note that it holds for every kernel U that

$$t([\![H]\!], U) = \int_{[0,1]^k} t_{x_1,\ldots,x_k}(H, U)\mathrm{d}x_{[k]}.$$

If H and H' are k-rooted graphs such that every pair of corresponding roots is joined by an edge in at most one of the graphs H and H', we define the *product* $H \times H'$ as follows: let H'' be the k-rooted graph isomorphic to H' that has the same roots as H and is vertex disjoint otherwise, and let $H \times H'$ be the graph with the vertex set $V(H) \cup V(H'')$, the edge set $E(H) \cup E(H')$ and the same set of roots. Note that $H \times H'$ does not have parallel edges as each pair of corresponding roots is joined by an edge in at most one of the graphs H and H'. Also observe that $|V(H \times H')| = |V(H)| + |V(H')| - k$ and it holds for every choice of roots and every kernel U that

$$t_\star(H \times H', U) = t_\star(H, U) \cdot t_\star(H', U).$$

If $H = H'$, we may write H^2 instead of $H \times H$. The definition of the operator $[\![\cdot]\!]$ and that of the product extend to rooted quantum graphs by linearity. Observe that, for every k-rooted quantum graph Q and every kernel U, it holds that $t([\![Q^2]\!], U) \geq 0$ and the equality holds if and only if $t_\star(Q, U) = 0$ for almost every choice of roots.

3. Forcing step structure

We start with recalling a construction from [36, Proposition 14.44], which forces the structure of a step kernel with at most q parts. For $k \in \mathbb{N}$ and $1 \leq i < j \leq k$, let Q_k^{ij} be the following $(2k)$-rooted quantum graph with roots $v_1, \ldots, v_k$ and $v'_1, \ldots, v'_k$. The quantum graph Q_k^{ij} has four constituents, each with a single non-root vertex: the graph with the non-root vertex adjacent to v_i and v'_i and the graph with the non-root vertex adjacent to v_j and v'_j, both with coefficient $+1$, as well as the graph with the non-root vertex adjacent to v_i and v'_j and the graph with the non-root vertex adjacent to v_j and v'_i, both with coefficient -1. See Figure 1 for an example. Let Q_k be the following quantum graph with each constituent having $2k + 2\binom{k}{2} = k(k+1)$ vertices:

$$Q_k = \left[\!\!\left[\prod_{1 \leq i < j \leq k} \left(Q_k^{ij}\right)^2 \right]\!\!\right].$$

The graph Q_k is the graph obtained in the proof of [36, Proposition 14.44] through an application [36, Lemma 14.37]. This gives the following lemma, whose proof we sketch for completeness.

Lemma 2. *For every $q \in \mathbb{N}$ and every kernel U, the following holds: $t(Q_{q+1}, U) = 0$ if and only if U is weakly isomorphic to a step kernel with at most q parts.*

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

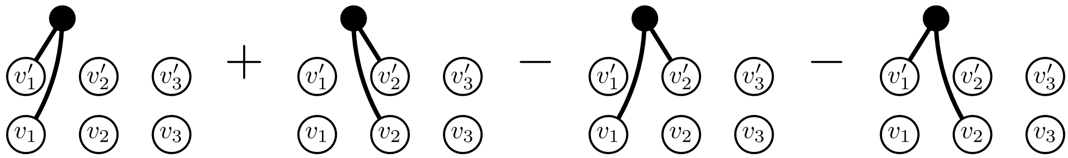

Figure 1. The 6-rooted quantum graph Q_3^{12}.

Proof. Observe that the value of $t(Q_{q+1}, U)$ for a kernel U is equal to

$$\int_{[0,1]^{2(q+1)}} \prod_{1\le i<j\le q+1} \left(\int_{[0,1]} \big(U(x_i,y) - U(x_j,y)\big)\big(U(x_i',y) - U(x_j',y)\big) \mathrm{d}y \right)^2 \mathrm{d}x_{[q+1]} \mathrm{d}x'_{[q+1]}. \quad (1)$$

If U is a step kernel with at most q parts, then for any choice of $x_1, \ldots, x_{q+1}$, there exist $1 \le i < j \le q+1$ such that x_i and x_j are from the same part of U and so $U(x_i, y) = U(x_j, y)$ for all $y \in [0, 1]$. Consequently, the product in (1) is zero for any choice of roots $x_1, \ldots, x_{q+1}$, which implies that $t(Q_{q+1}, U) = 0$.

We now prove the other implication, that is, that if $t(Q_{q+1}, U) = 0$, then U is weakly isomorphic to a step kernel with at most q parts. Let U be a kernel such that $t(Q_{q+1}, U) = 0$. By (1), the following holds for almost all $x_{[q+1]} \in [0, 1]^{q+1}$ and $x'_{[q+1]} \in [0, 1]^{q+1}$:

$$\prod_{1\le i<j\le q+1} \int_{[0,1]} \big(U(x_i,y) - U(x_j,y)\big) \left(U(x_i',y) - U(x_j',y)\right) \mathrm{d}y = 0.$$

Using [36, Proposition 13.23], we get that the following holds for almost all $x_{[q+1]} \in [0, 1]^{q+1}$:

$$\prod_{1\le i<j\le q+1} \int_{[0,1]} \big(U(x_i,y) - U(x_j,y)\big)^2 \mathrm{d}y = 0. \quad (2)$$

Let us consider an equivalence relation on $[0, 1]$ defined as $x \equiv x'$ if $U(x, y) = U(x', y)$ for almost all $y \in [0, 1]$. Observe that (2) holds for $x_{[q+1]} \in [0, 1]^{q+1}$ if and only if there exist $1 \le i < j \le q+1$ such that $x_i \equiv x_j$. Hence, (2) holds for almost all $x_{[q+1]} \in [0, 1]^{q+1}$ if and only if the measure of the q largest equivalence classes of $\equiv$ is one, which is equivalent to U being weakly isomorphic to a step kernel with at most q parts. □

We next present two rather similar auxiliary lemmas; since their statements and constructions somewhat differ depending on the parity of q, we state them separately for readability.

Lemma 3. *For every even integer $q \ge 2$ and all integers $s_1, \ldots, s_q \in [q+2, 2q+2]$, there exists a graph G with vertex set formed by q disjoint sets $V_1, \ldots, V_q$ that satisfies the following:*

- *the size of V_i is s_i for each $i \in [q]$,*
- *the edge set of G can be partitioned into four sets $M_1, \ldots, M_4$ such that, for every $1 \le i < j \le q$, each of the sets M_1 and M_2 restricted to vertices of $V_i \cup V_j$, is a matching of size $q+2$, and each of the sets M_3 and M_4 is a matching of size q, and*
- *the chromatic number of G is q and the colour classes of every q-colouring of G are precisely the sets $V_1, \ldots, V_q$; in particular, each of the sets V_i, $i \in [q]$, is independent.*

Proof. Fix an even integer $q \ge 2$ and integers $s_1, \ldots, s_q \in [q+2, 2q+2]$. Let $V_i = \{i\} \times [s_i]$; note that the first coordinate of a vertex determines which of the sets contains the vertex. We now describe the graph G by listing the edges between V_i and V_j, $1 \le i < j \le q$, contained in the matchings $M_1, \ldots, M_4$, where we abbreviate $\{(a, b), (c, d)\}$ to $(a, b)(c, d)$.

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

- The matching M_1 consists of the edge $(i, 1)(j, 1)$, the edge $(i, q+2)(j, q+2)$, and the edges $(i, k)(j, k+1)$ and $(i, k+1)(j, k)$ for even values k between 2 and q.
- The matching M_2 consists of the edges $(i, k)(j, k+1)$ and $(i, k+1)(j, k)$ for odd values k between 1 and $q+1$.
- The matching M_3 consists of the edges $(i, k)(j, s_j - q + k)$ for all $k \in [q]$.
- The matching M_4 consists of the edges $(i, s_i - q + k)(j, k)$ for all $k \in [q]$.

Observe that the following edges are always present between V_i and V_j, $1 \leq i < j \leq q$:

- the edges $(i, 1)(j, 1)$,
- the edges $(i, k)(j, k+1)$ and $(i, k+1)(j, k)$ for $k \in [q+1]$, and
- the edges $(i, k)(j, s_j - q + k)$ and $(i, s_i - q + k)(j, k)$ for $k \in [q]$.

Since the sets $V_1, \ldots, V_q$ are independent, the chromatic number of G is at most q. On the other hand, the vertices $(i, 1)$, $i \in [q]$ form a complete graph of order q, which implies that the chromatic number of G is at least q and so it is equal to q.

Consider an arbitrary q-colouring of G and let W_i, $i \in [q]$, be the colour class containing the vertex $(i, 1)$. (Note that the vertices $(i, 1)$, $i \in [q]$, are coloured with distinct colours as they form a complete graph.) We prove the following statement by induction on k: for every $i \in [q]$, if $k \leq s_i$, then the vertex (i, k) belongs to W_i. If $k = 1$, the statement follows from the definition of the sets W_i. If $k \in [2, q+2]$, for every $i \in [q]$, the existence of the edges $(j, k-1)(i, k)$, $j \in [q] \setminus \{i\}$, and the induction assumption, which states that $(j, k-1)$ belongs to W_j for $j \neq i$, imply that the vertex (i, k) belongs to W_i. Finally, if $k \in [q+3, s_i]$, $i \in [q]$, the existence of the edges $(j, q+k-s_i)(i, k)$, $j \in [q] \setminus \{i\}$, implies that the vertex (i, k) belongs to W_i (note that $q + k - s_i \leq q$ and so $(j, q+k-s_i) \in W_j$ for $j \neq i$). Hence, the q-colouring of G is unique up to a permutation of colour classes. □

We next present the version of Lemma 4 for odd values of $q \geq 3$.

Lemma 4. *For every odd integer $q \geq 3$ and all integers $s_1, \ldots, s_q \in [q+2, 2q+2]$, there exists a graph G with vertex set formed by q disjoint sets $V_1, \ldots, V_q$ that satisfies the following:*

- *the size of V_i is s_i for each $i \in [q]$,*
- *the edge set of G can be partitioned into four sets $M_1, \ldots, M_4$ such that each of the sets $M_1, \ldots, M_4$ restricted to vertices of $V_i \cup V_j$, $1 \leq i < j \leq q$, is a matching of size $q+1$, and*
- *the chromatic number of G is q and the colour classes of every q-colouring of G are precisely the sets $V_1, \ldots, V_q$; in particular, each of the sets V_i, $i \in [q]$, is independent.*

Proof. Fix an odd integer $q \geq 3$ and integers $s_1, \ldots, s_q \in [q+2, 2q+2]$, and set $V_i = \{i\} \times [s_i]$. We describe the graph G by listing the edges between V_i and V_j, $1 \leq i < j \leq q$, contained in the matchings $M_1, \ldots, M_4$.

- The matching M_1 consists of the edge $(i, 1)(j, 1)$, the edge $(i, q+1)(j, q+1)$, and the edges $(i, k)(j, k+1)$ and $(i, k+1)(j, k)$ for even values k between 2 and $q-1$.
- The matching M_2 consists of the edges $(i, k)(j, k+1)$ and $(i, k+1)(j, k)$ for odd values k between 1 and q.
- The matching M_3 consists of the edges $(i, k)(j, s_j - q - 1 + k)$ for all $k \in [q+1]$ unless $s_j = q+2$; if $s_j = q+2$, then the matching M_3 consists of the edges $(i, q+1)(j, q+2)$, $(i, q+2)(j, 2)$ and $(i, k)(j, k+2)$ for $k \in [q-1]$.

- The matching M_4 consists of the edges $(i, s_i - q - 1 + k)(j, k)$ for all $k \in [q+1]$ unless $s_i = q+2$; if $s_i = q+2$, then the matching M_4 consists of the edges $(i, q+2)(j, q+1)$, $(i, 2)(j, q+2)$ and $(i, k+2)(j, k)$ for $k \in [q-1]$.

Observe that the following edges are always present between V_i and V_j, $1 \le i < j \le q$:

- the edges $(i, 1)(j, 1)$,
- the edges $(i, k)(j, k+1)$ and $(i, k+1)(j, k)$ for $k \in [q]$,
- the edges $(i, k)(j, s_j - q - 1 + k)$ for $k = 2q + 3 - s_j, \dots, q+1$, and
- the edges $(i, s_i - q - 1 + k)(j, k)$ for $k = 2q + 3 - s_i, \dots, q+1$.

The rest of the argument now follows exactly the lines of the corresponding part of the proof of Lemma 3. □

We are now ready to prove the main lemma of this section.

Lemma 5. *For all integers $q \ge 2$ and $s_1, \dots, s_q \in [q+2, 2q+2]$, there exists an $(s_1, \dots, s_q)$-rooted quantum graph $P_{s_1,\dots,s_q}$ such that*

- *each constituent of $P_{s_1,\dots,s_q}$ has $2q(q-1)$ non-root vertices,*
- *the $s_1 + \cdots + s_q$ roots of $P_{s_1,\dots,s_q}$ form an independent set,*
- *for every q-step kernel U, there exists $d_0 = d_0(U) > 0$ that does not depend on $s_1, \dots, s_q$ such that $t_\star(P_{s_1,\dots,s_q}, U)$ is either 0 or d_0 for all choices of roots, and it is non-zero if and only if all roots from each of the q groups of roots of $P_{s_1,\dots,s_q}$ are chosen from the same part of U but the roots from different groups are chosen from different parts.*

Proof. For q and $s_1, \dots, s_q \in [q+2, 2q+2]$, let G be the graph from Lemma 3 or Lemma 4 (depending on the parity of q). Let $V_1, \dots, V_q$ be the sets forming the vertex set of G, and let $M_1, \dots, M_4$ be the sets forming the edge set of G as given by the lemma. We identify the vertices of V_i with the s_i roots in the i-th group. Let M_k^{ij}, for $1 \le i < j \le q$ and $k \in [4]$, consist of the edges of M_k between V_i and V_j, and let $\mathcal{M}_k^{ij}$ be the set of all $2^{\left|M_k^{ij}\right|}$ subsets of $V_i \cup V_j$ such that each set in $\mathcal{M}_k^{ij}$ contains exactly one vertex from each edge of M_k^{ij}. Next, if $W \subseteq V_1 \cup \cdots \cup V_q$, we write $P[W]$ for the $(s_1, \dots, s_q)$-rooted graph with a single non-root vertex such that the non-root vertex is adjacent to the roots in W. Finally, we define the $(s_1, \dots, s_q)$-rooted quantum graph $P_{s_1,\dots,s_q}$ as follows:

$$P_{s_1,\dots,s_q} = \prod_{1 \le i < j \le q} \prod_{k \in [4]} \sum_{W \in \mathcal{M}_k^{ij}} (-1)^{|W \cap V_i|} P[W].$$

Observe that each constituent of the quantum graph $P_{s_1,\dots,s_q}$ has exactly $4 \cdot \binom{q}{2} = 2q(q-1)$ non-root vertices, and the $s_1 + \cdots + s_q$ roots form an independent set. We remark that the $(s_1, \dots, s_q)$-rooted quantum graph

$$\sum_{W \in \mathcal{M}_k^{ij}} (-1)^{|W \cap V_i|} P[W] \tag{3}$$

from the definition of $P_{s_1,\dots,s_q}$ can also be obtained in the following alternative way, which gives additional insight into the definition of $P_{s_1,\dots,s_q}$. Let $P'[v]$ be the $(s_1, \dots, s_q, 1)$-rooted graph such that $P'[v]$ has no non-root vertices, v is a root contained in one of the first q groups of roots, and the only edge of $P'[v]$ is an edge joining the vertex v and the single root contained in the last group.

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

For $1 \le i < j \le q$ and $k \in [4]$, the $(s_1, \ldots, s_q)$-rooted quantum graph (3) can be obtained from the $(s_1, \ldots, s_q, 1)$-rooted graph

$$\prod_{vu \in M_k^{ij}} \left(P'[v] - P'[u]\right)$$

by changing the single root contained in the last group to a non-root vertex.

For the rest of the proof, fix a q-step kernel U and let z_i, $i \in [q]$, be any vertex of U contained in the i-th part of U. Consider a choice x_v, $v \in V(G)$, of roots. Suppose that G has an edge uv such that $u \in V_i$, $v \in V_j$, $1 \le i < j \le q$, $uv \in M_k$, $k \in [4]$, and the vertices x_u and x_v belong to the same part of the kernel U. Observe that

$$\begin{aligned}
&\sum_{W \in \mathcal{M}_k^{ij}} (-1)^{|W \cap V_i|} \prod_{w \in W} U(x_w, y) \\
&= \sum_{\substack{W \in \mathcal{M}_k^{ij} \\ u \in W}} (-1)^{|W \cap V_i|} \prod_{w \in W} U(x_w, y) + \sum_{\substack{W \in \mathcal{M}_k^{ij} \\ v \in W}} (-1)^{|W \cap V_i|} \prod_{w \in W} U(x_w, y) \\
&= \sum_{\substack{W \in \mathcal{M}_k^{ij} \\ u \in W}} (-1)^{|W \cap V_i|} \prod_{w \in W} U(x_w, y) + \sum_{\substack{W \in \mathcal{M}_k^{ij} \\ u \in W}} (-1)^{|W \cap V_i|-1} U(x_v, y) \prod_{w \in W \setminus \{u\}} U(x_w, y) \\
&= \sum_{\substack{W \in \mathcal{M}_k^{ij} \\ u \in W}} (-1)^{|W \cap V_i|} \left(U(x_u, y) - U(x_v, y)\right) \prod_{w \in W \setminus \{u\}} U(x_w, y) \\
&= 0.
\end{aligned}$$

It follows that $t_{x_{V(G)}}(P_{s_1,\ldots,s_q}, U) = 0$ if the colouring of the vertices of G such that v is coloured with the part containing x_v is not a proper colouring of G. Either Lemma 3 or Lemma 4 (depending on the parity of q) implies that $t_{x_{V(G)}}(P_{s_1,\ldots,s_q}, U) \neq 0$ only if all roots from each of the q groups of roots are chosen from the same part of U and the roots from different groups are chosen from different parts. If this is indeed the case and q is odd, the properties of the graph G given in Lemma 4 imply that

$$t_{x_{V(G)}}\left(P_{s_1,\ldots,s_q}, U\right) = \prod_{1 \le i < j \le q} \left(\int_{[0,1]} \left(U(z_i, y) - U(z_j, y)\right)^{q+1} \mathrm{d}y\right)^4. \tag{4}$$

This is positive since for every distinct $i, j \in [q]$ there is a positive measure of y with $U(z_i, y) \neq U(z_j, z)$ (as otherwise the i-th and j-th parts can be merged together contrary to the definition of a q-step kernel). Hence, the existence of d_0 follows and it is equal to the right-hand side of (4), which does not depend on the values of $s_1, \ldots, s_q$. Similarly, if q is even, the existence of d_0 follows from Lemma 3 and the definition of $P_{s_1,\ldots,s_q}$, and its value is

$$d_0 = \prod_{1 \le i < j \le q} \left(\int_{[0,1]} \left(U(z_i, y) - U(z_j, y)\right)^{q+2} \mathrm{d}y\right)^2 \left(\int_{[0,1]} \left(U(z_i, y) - U(z_j, y)\right)^{q} \mathrm{d}y\right)^2. \tag{5}$$

The proof of the lemma is now completed. □

We emphasise that the value of d_0 from the statement of Lemma 5 depends on the kernel U only, that is, it does not depend on $s_1, \ldots, s_q$; namely, d_0 is given by the right-hand side of (4) or (5) depending on the parity of q, the number of parts of the step kernel U.

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

4. Main result

We start with a construction of a quantum graph that restricts the *density* of each part A of a step kernel U, that is, the value of U on $A \times A$.

Lemma 6. *For all integers $q \geq 2$, $k \in [q]$ and reals $d_1, \ldots, d_k$, there exists a quantum graph $R_{d_1,\ldots,d_k}$ such that each constituent of $R_{d_1,\ldots,d_k}$ has $3q^2$ vertices and the following holds for every q-step kernel U: $t\left(R_{d_1,\ldots,d_k}, U\right) = 0$ if and only if the density of each part of U is one of the reals $d_1, \ldots, d_k$.*

Proof. Fix $q \geq 2$ and reals $d_1, \ldots, d_k$. Let $P_{q+2,\ldots,q+2}$ be the graph from Lemma 5. Note that $P_{q+2,\ldots,q+2}$ has $q(q+2) + 2q(q-1) = 3q^2$ vertices. For $m \in [0, 2k]$, we set $P^{(m)}_{q+2,\ldots,q+2}$ to be a graph obtained from $P_{q+2,\ldots,q+2}$ by adding arbitrary m edges among the roots in the first group (without creating parallel edges); note that this is possible since $2k \leq 2q \leq \binom{q+2}{2}$. Further, let $p(x)$ be the polynomial defined as

$$p(x) = \prod_{i=1}^{k} (x - d_i)^2,$$

and set $R_{d_1,\ldots,d_k}$ to be the quantum graph obtained from the expansion of $p(x)$ into monomials by replacing each monomial x^m, including x^0, with $\left[\!\left[P^{(m)}_{q+2,\ldots,q+2} \right]\!\right]$.

Consider any q-step kernel U and let $d_0 = d_0(U) > 0$ be the constant from Lemma 5. Observe that

$$t\left(\left[\!\left[P^{(m)}_{q+2,\ldots,q+2} \right]\!\right], U\right) = d_0 (q-1)! \left(\prod_{i=1}^{q} a_i^{q+2} \right) \left(\sum_{i=1}^{q} p_i^m \right),$$

where a_i is the measure and p_i is the density of the i-th part of U, $i \in [q]$; note that the term $(q-1)!$ counts possible choices of parts of U for the second, third, etc. group of roots while the choices of the part for the first group of roots are accounted for by the last sum in the expression. It follows that

$$t\left(R_{d_1,\ldots,d_k}, U\right) = d_0 (q-1)! \left(\prod_{i=1}^{q} a_i^{q+2} \right) \left(\sum_{i=1}^{q} p(p_i) \right),$$

which, using $p(x) \geq 0$ for all $x \in \mathbb{R}$, is equal to zero if and only if $p(p_i) = 0$ for every $i \in [q]$. The latter holds if and only if each p_i is one of the reals $d_1, \ldots, d_k$ (note that $p(x) > 0$ unless $x \in \{d_1, \ldots, d_k\}$), and so the quantum graph $R_{d_1,\ldots,d_k}$ has the properties given in the statement of the lemma. □

The next lemma provides a quantum graph restricting densities between pairs of parts of a step kernel; its proof is similar to that of Lemma 6, however, we include it for completeness.

Lemma 7. *For all integers $q \geq 2$, $k \in [q(q-1)/2]$ and reals $d_1, \ldots, d_k$, there exists a quantum graph $S_{d_1,\ldots,d_k}$ with $3q^2$ vertices such that the following holds for every q-step kernel U: $t\left(S_{d_1,\ldots,d_k}, U\right) = 0$ if and only if the density between each pair of distinct parts of U is one of the reals $d_1, \ldots, d_k$.*

Proof. Fix $q \geq 2$ and reals $d_1, \ldots, d_k$. Let $P_{q+2,\ldots,q+2}$ be the graph from Lemma 5. Recall that $P_{q+2,\ldots,q+2}$ has $q(q+2) + 2q(q-1) = 3q^2$ vertices. For $m \in [0, 2k]$, we set $P^{(m)}_{q+2,\ldots,q+2}$ to be a graph obtained from $P_{q+2,\ldots,q+2}$ by adding arbitrary m edges joining a root in the first group and a root in the second group without creating parallel edges; note that this is possible since $2k \leq q(q-1) \leq (q+2)^2$. Further, let $p(x)$ be the polynomial defined as

$$p(x) = \prod_{i=1}^{k} (x - d_i)^2,$$

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

and set $S_{d_1,\dots,d_k}$ to be the quantum graph obtained from the expansion of $p(x)$ by replacing x^m with $\left[\!\left[P^{(m)}_{q+2,\dots,q+2}\right]\!\right]$.

Consider a q-step kernel U and let $d_0 = d_0(U) > 0$ be the constant from Lemma 5. Observe that

$$t\left(\left[\!\left[P^{(m)}_{q+2,\dots,q+2}\right]\!\right], U\right) = 2d_0(q-2)! \left(\prod_{i=1}^{q} a_i^{q+2}\right)\left(\sum_{1\le i<j\le q} p_{ij}^m\right),$$

where a_i is the measure of the i-th part of U, $i \in [q]$, and p_{ij} is the density between the i-th and j-th part of U, $1 \le i < j \le q$. It follows that

$$t\left(S_{d_1,\dots,d_k}, U\right) = 2d_0(q-2)! \left(\prod_{i=1}^{q} a_i^{q+2}\right)\left(\sum_{1\le i<j\le q} p\left(p_{ij}\right)\right),$$

which (by $p \ge 0$) is equal to zero if and only if $p(p_{ij}) = 0$ for all $1 \le i < j \le q$. The latter holds if and only if each p_{ij}, $1 \le i < j \le q$, is one of the reals $d_1, \dots, d_k$, and so the quantum graph $S_{d_1,\dots,d_k}$ has the properties given in the statement of the lemma. □

We next present a construction of a rooted quantum graph that "tests" whether there is a permutation of parts of a step kernel matching densities in a given matrix D. As the value of d_0 in Lemma 5, the value of c_0 in Lemma 8 does not depend on $s_1, \dots, s_q$, namely, it depends on the matrix D and the kernel U only.

Lemma 8. *For all integers $q \ge 2$, $s_1, \dots, s_q \in [q+2, 2q+2]$ and a symmetric real matrix $D \in \mathbb{R}^{q\times q}$, there exists an $(s_1, \dots, s_q)$-rooted quantum graph $T_{s_1,\dots,s_q}$ satisfying the following. Each constituent of $T_{s_1,\dots,s_q}$ has $2q(q-1)$ non-root vertices, and if U is a q-step kernel such that*

- *the density of each part of U is one of the diagonal entries of D, and*
- *the density between each pair of the parts of U is one of the off-diagonal entries of D,*

then there exists $c_0 = c_0(D, U) \ne 0$, which does not depend on $s_1, \dots, s_q$, such that $t_\star\left(T_{s_1,\dots,s_q}, U\right)$ is either 0 or c_0 for all choices of roots and it is non-zero if and only if

- *all roots from each of the q groups of roots are chosen from the same part of U,*
- *roots from different groups are chosen from different parts of U,*
- *D_{ii} is the density of the part of U that the i-th group of roots is chosen from, and*
- *D_{ij} is the density between the parts of U that the i-th and j-th groups of roots are chosen from.*

Proof. Fix integers $q \ge 2$, $s_1, \dots, s_q \in [q+2, 2q+2]$, and a matrix D. Let Z_1 be the set containing the values of diagonal entries of D and Z_2 the set containing the values of off-diagonal entries of D. We next define a polynomial p, whose $\binom{q+1}{2}$ are variables are indexed by pairs ij with $1 \le i \le j \le q$, as follows:

$$p\left(x_{11}, x_{12}, \dots, x_{q-1,q}, x_{q,q}\right) = \left(\prod_{i=1}^{q} \prod_{z\in Z_1\setminus\{D_{ii}\}} (x_{ii}-z)\right)\left(\prod_{1\le i<j\le q} \prod_{z\in Z_2\setminus\{D_{ij}\}} (x_{ij}-z)\right).$$

Let $P_{s_1,\dots,s_q}$ be the graph from Lemma 5. For $m_{ii} \in [0, |Z_1|]$, $i \in [q]$, and $m_{ij} \in [0, |Z_2|]$, $1 \le i < j \le k$, let $P^{m_{11},m_{12},\dots,m_{q,q}}_{s_1,\dots,s_q}$ be an $(s_1, \dots, s_q)$-rooted quantum graph obtained from $P_{s_1,\dots,s_q}$ by adding arbitrary m_{ij} edges joining roots in the i-th group and with the roots in the j-th group for $1 \le i \le j \le q$ (without creating parallel edges). The $(s_1, \dots, s_q)$-rooted quantum graph $T_{s_1,\dots,s_q}$ is

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

obtained from the expansion of $p\left(x_{11}, x_{12}, \ldots, x_{q,q}\right)$ into monomials by replacing each monomial $x_{11}^{m_{11}} x_{12}^{m_{12}} \cdots x_{q,q}^{m_{q,q}}$ with $P_{s_1,\ldots,s_q}^{m_{11},m_{12},\ldots,m_{q,q}}$ (including the monomial $x_{11}^0 \cdots x_{q,q}^0$).

Fix a q-step kernel U such that

- the density of each part of U belongs to Z_1, and
- the density between each pair of the parts of U belongs to Z_2.

Let $d_0 = d_0(U) > 0$ be the constant from Lemma 5. Note that $t_\star\left(T_{s_1,\ldots,s_q}, U\right) = 0$ unless

- all roots from each of the q groups of roots are chosen from the same part of U,
- roots from different groups are chosen from different parts of U,
- D_{ii} is the density of the part of U that the i-th group of roots is chosen from, and
- D_{ij} is the density between the parts of U that the i-th and j-th groups of roots are chosen from,

and if $t_\star\left(T_{s_1,\ldots,s_q}, U\right) \neq 0$, then it is equal to

$$c_0 = d_0 \left(\prod_{i=1}^{q} \prod_{z \in Z_1 \setminus \{D_{ii}\}} (D_{ii} - z) \right) \left(\prod_{1 \le i < j \le q} \prod_{z \in Z_2 \setminus \{D_{ij}\}} (D_{ij} - z) \right) \neq 0.$$

Hence, the $(s_1, \ldots, s_q)$-rooted quantum graph $T_{s_1,\ldots,s_q}$ has the properties given in the statement of the lemma. □

To prove the main result of this paper, we need the following well-known result, which we state explicitly for reference.

Lemma 9. *For every $q \ge 1$ and reals $z_1, \ldots, z_q$, the following system of equations has at most one solution $x_1, \ldots, x_q \in \mathbb{R}$ (up to a permutation of the values):*

$$\begin{array}{ccccccc} x_1 & + & \cdots & + & x_q & = & z_1 \\ x_1^2 & + & \cdots & + & x_q^2 & = & z_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ x_1^q & + & \cdots & + & x_q^q & = & z_q. \end{array}$$

Proof. The system of equations gives the first q power sums of $x_1, \ldots, x_q$. By Newton's identities (see e.g., [40, Equation (2.11′)]), this determines the first q elementary symmetric polynomials, which are the coefficients of the polynomial $\prod_{i=1}^{q} (x + x_i)$. Therefore any other solution $y_1, \ldots, y_q$ of the system satisfies that $\prod_{i=1}^{q} (x + x_i) = \prod_{i=1}^{q} (x + y_i)$, which yields the statement of the lemma because of the uniqueness of polynomial factorisation. □

We are now ready to prove our main result, which implies Theorem 1 stated in Section 1.

Theorem 10. *The following holds for every $q \ge 2$ and every q-step kernel U: if the density of each graph with at most $4q^2 - q$ vertices in a kernel U' is the same as in U, then the kernels U and U' are weakly isomorphic.*

Proof. Fix a q-step kernel U. Let $a_1, \ldots, a_q$ be the measures of the q parts. Further let $D \in \mathbb{R}^{q \times q}$ be the matrix such that D_{ii} is the density of the i-th part of U and D_{ij}, $i \neq j$, is the density between the i-th and j-th part.

Consider a kernel U' such that $t(H, U) = t(H, U')$ for all graphs with at most $4q^2 - q$ vertices. Since each constituent of the quantum graphs Q_q and Q_{q+1} from Lemma 2 has $q(q + 1)$ and

$(q+1)(q+2) \leq 4q^2 - q$ vertices, respectively, it holds that $t\big(Q_q, U'\big) \neq 0$ and $t\big(Q_{q+1}, U'\big) = 0$ (as they are the same as the corresponding densities in U). We conclude using Lemma 2 that U' is a q-step kernel.

Let $R_{D_{11},\ldots,D_{qq}}$ be the quantum graph from the statement of Lemma 6; note that each constituent of $R_{D_{11},\ldots,D_{qq}}$ has $3q^2 \leq 4q^2 - q$ vertices. Since $t(R_{D_{11},\ldots,D_{qq}}, U') = 0$ (as the value is the same as for the kernel U), Lemma 6 yields that the density of each part of U' is equal to one of the diagonal entries of D. Similarly, Lemma 7 applied with the off-diagonal entries of D yields that the density between any pair of parts of U' is equal to one of the off-diagonal entries of D. In addition, the $(q+2, \ldots, q+2)$-rooted quantum graph $T_{q+2,\ldots,q+2}$ from Lemma 8 applied with the matrix D satisfies $t\big([\![T_{q+2,\ldots,q+2}]\!], U\big) \neq 0$; thus it holds that $t\big([\![T_{q+2,\ldots,q+2}]\!], U'\big) \neq 0$. Hence, we derive using Lemma 8 that, after possibly permuting the parts of U', the density of the i-th part of U' is D_{ii} and the density between the i-th and j-th parts of U' is D_{ij}.

Let $d_0 = d_0(U) > 0$ be the constant from Lemma 5 for the kernel U. Observe that, for each $k \in [0, q]$, the following holds for the rooted quantum graph $P_{q+k+2,q+2,\ldots,q+2}$ from Lemma 5:

$$t\big([\![P_{q+k+2,q+2,\ldots,q+2}]\!], U\big) = d_0 (q-1)! \left(\prod_{j=1}^{q} a_j^{q+2}\right)\left(\sum_{i=1}^{q} a_i^k\right).$$

It follows that the following holds for every $k \in [q]$:

$$\sum_{i=1}^{q} a_i^k = \frac{q \cdot t\big([\![P_{q+k+2,q+2,\ldots,q+2}]\!], U\big)}{t\big([\![P_{q+2,q+2,\ldots,q+2}]\!], U\big)}.$$

Similarly, with a'_i denoting the measure of the i-th part of U', we obtain that

$$\sum_{i=1}^{q} \big(a'_i\big)^k = \frac{q \cdot t\big([\![P_{q+k+2,q+2,\ldots,q+2}]\!], U'\big)}{t\big([\![P_{q+2,q+2,\ldots,q+2}]\!], U'\big)}.$$

Hence, Lemma 9 and the assumption that the homomorphism densities of all graphs with at most $q(q+2) + q + 2q(q-1) = 3q^2 + q \leq 4q^2 - q$ vertices are the same in U and U' implies that the multisets $a_1, \ldots, a_q$ and $a'_1, \ldots, a'_q$ are the same.

Let $c_0 = c_0(D, U) \neq 0$ be the constant from Lemma 8 for the kernel U and let Π_D be the set of all permutations π of the parts of U such that the densities inside the parts and between the parts in U and after applying π to the parts of U are still as given by D. Observe that it holds that

$$t\Big([\![T_{s_1,\ldots,s_q}]\!], U\Big) = c_0 \sum_{\pi \in \Pi_D} \prod_{i=1}^{q} a_{\pi(i)}^{s_i}.$$

Let $p(x_1, \ldots, x_q)$ be the polynomial defined as

$$p(x_1, \ldots, x_q) = \left(\prod_{j=1}^{q} x_j^{q+2}\right)\left(\prod_{i=1}^{q} \prod_{a \in \{a_1,\ldots,a_q\} \setminus \{a_i\}} (x_i - a)\right).$$

Note that each variable in each monomial of p has degree between $q+2$ and $2q+1$. Since each a_i is non-zero, we have for all q-tuples $(x_1, \ldots, x_q)$ of reals with $\{x_1, \ldots, x_q\} \subseteq \{a_1, \ldots, a_q\}$ that $p(x_1, \ldots, x_q) = 0$ if and only if there exists $i \in [q]$ such that $x_i \neq a_i$. Let T be the quantum graph obtained from the polynomial p by expanding it and then replacing each monomial $x_1^{s_1} \cdots x_q^{s_q}$ with

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

$[\![T_{s_1,\dots,s_q}]\!]$ (including the monomial $x_1^0 \cdots x_1^0$). Note that the number of vertices of each constituent of T is at most $q(2q+1)+2q(q-1)=4q^2-q$ and

$$t(T,U)=c_0\sum_{\pi\in\Pi_D}p(a_{\pi(1)},\dots,a_{\pi(q)}).$$

In particular, it holds that $t(T,U)\neq 0$ and so $t(T,U')\neq 0$. Along the same lines, we obtain that

$$t(T,U')=c_0'\sum_{\pi\in\Pi_D'}p\left(a'_{\pi(1)},\dots,a'_{\pi(q)}\right),$$

where $c_0'=c_0(D,U')\neq 0$ is the constant from Lemma 8 for the kernel U' and Π_D' is the set of all permutations π of the parts of U' such that the densities of the parts and between the parts after applying π are as given by D. Since it holds that $t(T,U')\neq 0$, the set Π_D' is non-empty. It follows that Π_D' contains a permutation π such that $a'_{\pi(i)}=a_i$ for all $i\in[q]$, which implies that the kernels U and U' are weakly isomorphic. □

5. Parts with different degrees

In this section, we show that a q-step kernel such that its vertices contained in different parts have different degrees is forced by graphs with at most $2q+1$ vertices.

Theorem 11. *The following holds for every $q\geq 2$ and every q-step kernel U such that the degrees of vertices in different parts are different: if the density of each graph with at most $2q+1$ vertices in a kernel U' is the same as in U, then the kernels U and U' are weakly isomorphic.*

Proof. Fix $q\geq 2$, a q-step kernel U and a kernel U' such that $t(H,U)=t(H,U')$ for every graph H with at most $2q+1$ vertices. For $i\in[q]$, let A_i be the i-th part of U, a_i be the measure of A_i, and let d_i be the common degree of the vertices contained in A_i.

Let $K_1^\bullet$ and $K_2^\bullet$ be the 1-rooted graphs obtained from K_1 and K_2, respectively, by choosing one of their vertices to be the root. Note that $t_x(K_2^\bullet-dK_1^\bullet,V)=0$ if and only if the degree of x in a kernel V is d. It follows that a kernel V satisfies that

$$t\left(\left[\!\!\left[\prod_{i\in[q]}\left(K_2^\bullet-d_iK_1^\bullet\right)^2\right]\!\!\right],V\right)=0 \tag{6}$$

if and only if the degree of almost every vertex of V is one of the numbers $d_1,\dots,d_q$. Next observe that if a kernel V satisfies (6) and

$$t\left(\left[\!\!\left[\prod_{i\in[q]\setminus\{k\}}\left(K_2^\bullet-d_iK_1^\bullet\right)\right]\!\!\right],V\right)=a_k\prod_{i\in[q]\setminus\{k\}}(d_k-d_i) \tag{7}$$

for $k\in[q]$, then the measure of the set of vertices of V with degree equal to d_k is a_k. Since U satisfies (6) and (7) for every $k\in[q]$ and the graphs appearing in (6) and (7) have at most $2q+1$ and q vertices, respectively, the vertex set of the kernel U' can be partitioned into q (measurable) sets $A_1',\dots,A_q'$ and a null set A_0' such that the measure of A_k' is a_k and all vertices contained in A_k' have degree equal to d_k for every $k\in[q]$.

Let $G^{\bullet\bullet}$, $G^{\circ\bullet}$ and $G^{\bullet\circ}$ be the following 2-rooted graphs: $G^{\bullet\bullet}$ consists of two isolated roots only, $G^{\circ\bullet}$ is obtained from $G^{\bullet\bullet}$ by adding a non-root vertex adjacent to the first root, and $G^{\bullet\circ}$ is obtained

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

from $G^{\bullet\bullet}$ by adding a non-root vertex adjacent to the second root. For $k, \ell \in [q]$, let $H_{k\ell}$ be the 2-rooted quantum graph defined as

$$H_{k\ell} = \left(\prod_{i \in [q] \setminus \{k\}} (G^{\circ\bullet} - d_i G^{\bullet\bullet}) \right) \times \left(\prod_{j \in [q] \setminus \{\ell\}} (G^{\bullet\circ} - d_j G^{\bullet\bullet}) \right),$$

and observe that

$$t_{xy}\left(H_{k\ell}, U'\right) = \left(\prod_{i \in [q] \setminus \{k\}} (d_k - d_i) \right) \left(\prod_{j \in [q] \setminus \{\ell\}} (d_\ell - d_j) \right)$$

if the degree of x is d_k and the degree of y is d_ℓ, and $t_{xy}\left(H_{k\ell}, U'\right) = 0$ otherwise. In particular, it follows that

$$t\left([\![H_{k\ell}]\!], U'\right) = a_k a_\ell \left(\prod_{i \in [q] \setminus \{k\}} (d_k - d_i) \right) \left(\prod_{j \in [q] \setminus \{\ell\}} (d_\ell - d_j) \right). \tag{8}$$

Note that each constituent of the 2-rooted quantum graph $H_{k\ell}$ has at most $2q$ vertices. Let $H'_{k\ell}$ be the 2-rooted quantum graph obtained from $H_{k\ell}$ by joining the two roots in each of its constituents by an edge. Similarly as above, one can show that

$$t\left([\![H'_{k\ell}]\!], U'\right) = \left(\int_{A'_k \times A'_\ell} U'(x, y)\mathrm{d}x\mathrm{d}y \right) \left(\prod_{i \in [q] \setminus \{k\}} (d_k - d_i) \right) \left(\prod_{j \in [q] \setminus \{\ell\}} (d_\ell - d_j) \right). \tag{9}$$

Using (8) and (9), we obtain that

$$\frac{\int_{A'_k \times A'_\ell} U'(x, y)\mathrm{d}x\mathrm{d}y}{a_k a_\ell} = \frac{t\left([\![H'_{k\ell}]\!], U'\right)}{t([\![H_{k\ell}]\!], U')} = \frac{t\left([\![H'_{k\ell}]\!], U\right)}{t([\![H_{k\ell}]\!], U)} = \frac{\int_{A_k \times A_\ell} U(x, y)\mathrm{d}x\mathrm{d}y}{a_k a_\ell},$$

that is, the average density between the parts A'_k and A'_ℓ in the kernel U' is the same as the density between the parts A_k and A_ℓ in the kernel U.

We now recall that each step kernel is the unique (up to weak isomorphism) minimiser of the density of C_4 among all kernels with the same number of parts of the same measures and the same density between them. This statement for step graphons with parts of equal measure appears in [12, Lemma 11] and the same proof applies for kernels with parts not necessarily having the same sizes; also see [36, Propositions 14.13 and 14.14] for related results. Since $t(C_4, U) = t(C_4, U')$, it follows that U' is weakly isomorphic to U. □

6. Concluding remarks

Theorem 10 asserts that every q-step kernel is forced by graphs with at most $4q^2 - q$ vertices. We do not know whether it suffices to consider homomorphism densities of graphs with $o(q^2)$ vertices, both in the case of kernels and in the more restrictive case of graphons. We leave this as an open problem.

We finish by establishing that it is necessary to consider graphs with the number of vertices linear in q. The argument is similar to that used in analogous scenarios, for example, in [18, 21]. For reals $a_1, \ldots, a_q \in (0, 1)$ such that $a_1 + \cdots + a_q < 1$, let $U_{a_1,\ldots,a_q}$ be the $(q+1)$-step graphon with parts of measures $a_1, \ldots, a_q$ and $1 - a_1 - \cdots - a_q$ such that the graphon $U_{a_1,\ldots,a_q}$ is equal to one within each of the first q parts and to zero elsewhere. Observe that if H is a graph, which

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

consists of k components with $n_1, \ldots, n_k$ vertices after the removal of isolated vertices, then

$$t\left(H, U_{a_1,\ldots,a_q}\right) = \prod_{i=1}^{k} \sum_{j=1}^{q} a_j^{n_i} = \prod_{i=1}^{k} t\left(K_{n_i}, U_{a_1,\ldots,a_q}\right).$$

It follows that if

$$t\left(K_{\ell+1}, U_{a_1,\ldots,a_q}\right) = t\left(K_{\ell+1}, U_{a'_1,\ldots,a'_q}\right) \quad \text{for every } \ell \in [q-1], \tag{10}$$

then the homomorphism density of every graph with at most q vertices is the same in $U_{a_1,\ldots,a_q}$ and in $U_{a'_1,\ldots,a'_q}$. When $f(a_1, \ldots, a_q) = \left(t\left(K_{\ell+1}, U_{a_1,\ldots,a_q}\right)\right)_{\ell=1}^{q-1}$ is viewed as a function of $a_1, \ldots, a_q$, then its Jacobian matrix J with respect to the first $q-1$ coordinates is

$$\begin{bmatrix} 2a_1 & \cdots & 2a_{q-1} \\ 3a_1^2 & \cdots & 3a_{q-1}^2 \\ \vdots & & \vdots \\ qa_1^{q-1} & \cdots & qa_{q-1}^{a-1} \end{bmatrix} = \begin{bmatrix} 2 & & \\ & \ddots & \\ & & q \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_{q-1} \\ \vdots & & \vdots \\ a_1^{q-2} & \cdots & a_{q-1}^{a-2} \end{bmatrix} \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_{q-1} \end{bmatrix}. \tag{11}$$

Fix any distinct positive reals $a_1, \ldots, a_q$ with sum less than 1. Note that the middle matrix in (11) is the Vandermonde matrix of $(a_1, \ldots, a_{q-1})$ and thus the Jacobian matrix J is non-singular. By the Implicit Function Theorem, for every a'_q sufficiently close to a_q there is a vector $\left(a'_1, \ldots, a'_{q-1}\right)$ close to $(a_1, \ldots, q_q)$ such that (10) holds. By making a'_q sufficiently close but not equal to a_q, we can ensure that $a'_q \notin \{a_1, \ldots, a_q\}$ and that all elements a'_i are positive and sum to less than 1. Thus we obtain two $(q+1)$-step graphons, namely $U_{a_1,\ldots,a_q}$ and $U_{a'_1,\ldots,a'_q}$, that have the same homomorphism density of every graph with at most q vertices but are not weakly isomorphic; the latter can be established by, for example, applying the proof of Theorem 10 to these two graphons (alternatively, it also follows from the general analytic characterisation of weakly isomorphic kernels [36, Theorem 13.10]).

Acknowledgement

The authors would like to thank both anonymous reviewers for their comments that have improved the presentation of the results in the paper.

References

[1] Borgs, C., Chayes, J. and Lovász, L. (2010) Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** 1597–1619.
[2] Borgs, C., Chayes, J. T., Cohn, H. and Ganguly, S. (2021) Consistent nonparametric estimation for heavy-tailed sparse graphs. *Ann. Stat.* **49** 1904–1930.
[3] Bucić, M., Long, E., Shapira, A. and Sudakov, B. (2021) Tournament quasirandomness from local counting. *Combinatorica* **41** 175–208.
[4] Chan, T., Král', D., Noel, J. A., Pehova, Y., Sharifzadeh, M. and Volec, J. (2020) Characterization of quasirandom permutations by a pattern sum. *Random Struct. Algorithms* **57** 920–939.
[5] Chung, F. R. K. and Graham, R. L. (1990) Quasi-random hypergraphs. *Random Struct. Algorithms* **1** 105–124.
[6] Chung, F. R. K. and Graham, R. L. (1991) Quasi-random set systems. *J. Am. Math. Soc.* **4** 151–196.
[7] Chung, F. R. K. and Graham, R. L. (1991) Quasi-random tournaments. *J. Graph Theory* **15** 173–198.
[8] Chung, F. R. K. and Graham, R. L. (1992) Quasi-random subsets of Z_n. *J. Comb. Theory Ser. A* **61** 64–86.
[9] Chung, F. R. K., Graham, R. L. and Wilson, R. M. (1989) Quasi-random graphs. *Combinatorica* **9** 345–362.
[10] Cooper, J. N. (2004) Quasirandom permutations. *J. Comb. Theory Ser. A* **106** 123–143.

https://doi.org/10.1017/S0963548323000263 Published online by Cambridge University Press

[11] Cooper, J. W., Král', D., Lamaison, A. and Mohr, S. (2022) Quasirandom Latin squares. *Random Struct. Algorithms* **61** 298–308.
[12] Cooper, J. W., Král', D. and Martins, T. (2018) Finitely forcible graph limits are universal. *Adv. Math.* **340** 819–854.
[13] Coregliano, L. N., Parente, R. F. and Sato, C. M. (2019) On the maximum density of fixed strongly connected subtournaments. *Electron. J. Comb.* **26** P1.44.
[14] Coregliano, L. N. and Razborov, A. A. (2017) On the density of transitive tournaments. *J. Graph Theory* **85** 12–21.
[15] Coregliano, L. N. and Razborov, A. A. (2020) Natural quasirandomness properties, preprint arXiv: 2012.11773.
[16] Diaconis, P. and Janson, S. (2008) Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* **28** 33–61.
[17] Eberhard, S., Manners, F. and Mrazović, R. (2022) Transversals in quasirandom Latin squares, preprint arXiv: 2209.02180.
[18] Erdős, P., Lovász, L. and Spencer, J. (1979) Strong independence of graphcopy functions. In *Graph Theory and Related Topics*. Academic Press, pp. 165–172.
[19] Gao, C., Lu, Y. and Zhou, H. H. (2015) Rate-optimal graphon estimation. *Ann. Stat.* **43** 2624–2652.
[20] Garbe, F., Hancock, R., Hladký, J. and Sharifzadeh, M. (2020) Limits of Latin squares, preprint arXiv: 2010.07854.
[21] Glebov, R., Grzesik, A., Klimošová, T. and Král', D. (2015) Finitely forcible graphons and permutons. *J. Comb. Theory Ser. B* **110** 112–135.
[22] Gowers, W. T. (2006) Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Comb. Probab. Comput.* **15** 143–184.
[23] Gowers, W. T. (2007) Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. Math.* **166** 897–946.
[24] Gowers, W. T. (2008) Quasirandom groups. *Comb. Probab. Comput.* **17** 363–387.
[25] Gowers, W. T. and Long, J. (2020) Partial associativity and rough approximate groups. *Geom. Funct. Anal.* **30** 1–65.
[26] Grzesik, A., Il'kovič, D., Kielak, B. and Král', D. (2022) Quasirandom forcing orientations of cycles, preprint arXiv: 2212.09343.
[27] Grzesik, A., Král', D. and Lovász, L. M. (2020) Elusive extremal graphs. *Proc. Lond. Math. Soc.* **121** 1685–1736.
[28] Hancock, R., Kabela, A., Král', D., Martins, T., Parente, R., Skerman, F. and Volec, J. (2023) No additional tournaments are quasirandom-forcing. *Eur. J. Comb.* **108** 103632.
[29] Haviland, J. and Thomason, A. (1989) Pseudo-random hypergraphs. *Discrete Math.* **75** 255–278.
[30] Klopp, O., Tsybakov, A. B. and Verzelen, N. (2017) Oracle inequalities for network models and sparse graphon estimation. *Ann. Stat.* **45** 316–354.
[31] Klopp, O. and Verzelen, N. (2019) Optimal graphon estimation in cut distance. *Probab. Theory Relat. Fields* **174** 1033–1090.
[32] Kohayakawa, Y., Rödl, V. and Skokan, J. (2002) Hypergraphs, quasi-randomness, and conditions for regularity. *J. Comb. Theory Ser. A* **97** 307–352.
[33] Král', D., Lovász, L. M., Noel, J. A. and Sosnovec, J. (2020) Finitely forcible graphons with an almost arbitrary structure. *Discrete Anal.* **2020** Article 9, 36pp.
[34] Král', D. and Pikhurko, O. (2013) Quasirandom permutations are characterized by 4-point densities. *Geom. Funct. Anal.* **23** 570–579.
[35] Kureĉka, M. (2022) Lower bound on the size of a quasirandom forcing set of permutations. *Comb. Probab. Comput.* **31** 304–319.
[36] Lovász, L. (2012) *Large Networks and Graph Limits*, Vol. **60** of *Colloquium Publications*, American Mathematical Society.
[37] Lovász, L. and Sós, V. (2008) Generalized quasirandom graphs. *J. Comb. Theory Ser. B* **98** 146–163.
[38] Lovász, L. and Szegedy, B. (2006) Limits of dense graph sequences. *J. Comb. Theory Ser. B* **96** 933–957.
[39] Lovász, L. and Szegedy, B. (2011) Finitely forcible graphons. *J. Comb. Theory Ser. B* **101** 269–301.
[40] Macdonald, I. G. (1995) Symmetric Functions and Hall Polynomials, 2nd edition. Oxford University Press.
[41] Nagle, B., Rödl, V. and Schaht, M. (2006) The counting lemma for regular k-uniform hypergraphs. *Random Struct. Algorithms* **28** 113–179.
[42] Rödl, V. (1986) On universality of graphs with uniformly distributed edges. *Discrete Math.* **59** 125–134.
[43] Rödl, V. and Skokan, J. (2004) Regularity lemma fork-uniform hypergraphs. *Random Struct. Algorithms* **25** 1–42.
[44] Simonovits, M. and Sós, V. T. (1991) Szemerédi's partition and quasirandomness. *Random Struct. Algorithms* **2** 1–10.
[45] Spencer, J. (2010) Quasirandom multitype graphs. In *An Irregular Mind*, Vol. **21** of *Bolyai Society Mathematical Studies*, Springer, pp. 607–617.
[46] Thomason, A. (1987) Pseudo-random graphs. *Ann. Discrete Math.* **144**, 307–331.
[47] Thomason, A. (1987) Random graphs, strongly regular graphs and pseudo-random graphs. In *Surveys in Combinatorics*, Vol. **123** of *London Mathematical Society Lecture Note Series*, pp. 173–196.

Cite this article: Grzesik A, Král' D, and Pikhurko O (2024). Forcing generalised quasirandom graphs efficiently. *Combinatorics, Probability and Computing* **33**, 16–31. https://doi.org/10.1017/S0963548323000263