Received: 19 December 2020 | Accepted: 10 March 2023

DOI: 10.1112/jlms.12777

RESEARCH ARTICLE

Journal of the London Mathematical Society

Stability from graph symmetrisation arguments with applications to inducibility

Hong Liu[1] | **Oleg Pikhurko**[2] | **Maryam Sharifzadeh**[3] | **Katherine Staden**[4]

[1]Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea

[2]Mathematics Institute and DIMAP, University of Warwick, Coventry, UK

[3]Department of Mathematics and Mathematical Statistics, Umeå University, Umeå, Sweden

[4]School of Mathematics and Statistics, The Open University, Milton Keynes, UK

Correspondence

Katherine Staden, School of Mathematics and Statistics, The Open University, Milton Keynes, MK7 6AA, UK.
Email: katherine.staden@open.ac.uk

Funding information

UK Research and Innovation, Grant/Award Number: MR/S016325/1; Institute for Basic Science, Grant/Award Number: IBS-R029-C4; ERC, Grant/Award Number: 101020255; Leverhulme Trust, Grant/Award Number: RPG-2018-424; Engineering and Physical Sciences Research Council, Grant/Award Number: EP/V025953/1

Abstract

We present a sufficient condition for the stability property of extremal graph problems that can be solved via Zykov's symmetrisation. Our criterion is stated in terms of an analytic limit version of the problem. We show that, for example, it applies to the inducibility problem for an arbitrary complete bipartite graph B, which asks for the maximum number of induced copies of B in an n-vertex graph, and to the inducibility problem for $K_{2,1,1,1}$ and $K_{3,1,1}$, the only complete partite graphs on at most five vertices for which the problem was previously open.

MSC 2020

05C35 (primary)

1 | INTRODUCTION AND NOTATION

The notion of symmetrisation in graphs was introduced by Zykov in [40]. In its most basic form, symmetrisation is the process of considering two non-adjacent vertices x and y in a graph G, and replacing x by a clone of y, that is, a vertex y' whose neighbourhood is the same as that of y. Zykov used symmetrisation to reprove Turán's theorem [38], as follows. Let G be an n-vertex K_r-free graph with the maximum number of edges. Whenever there are non-adjacent vertices x, y with $d_G(x) \leqslant d_G(y)$, we symmetrise by replacing x by a clone of y. The graph obtained in this way is still K_r-free and has at least as many edges as G, and one can do this so that the final graph is complete partite. Standard convexity arguments imply that there are $r-1$ parts of almost equal size, recovering Turán's theorem. A variation of this approach was employed by Motzkin and Straus [28] also to reprove Turán's theorem.

Suppose that one seeks to maximise (or minimise) a graph parameter λ such that there is always a way to symmetrise any given non-adjacent pair in a graph without decreasing λ. Then it suffices to only consider 'totally symmetrised' (i.e. complete partite) graphs to determine the maximum value of λ. Bollobás [3] used symmetrisation to show that the parameter which counts any linear combination of cliques is symmetrisable, a special case of which provides a lower bound for the minimal number of cliques in a graph of given order and size.

In this paper, we are interested in more general graph parameters λ which do not decrease upon symmetrisation, in a specific sense we describe below. Like the example above, a symmetrisable λ is maximised (not necessarily uniquely) by a complete partite graph. Our main result gives a sufficient condition for stability for symmetrisable functions, namely that any graph which almost maximises λ looks very much like a complete partite graph. In fact, we prove the quantitatively sharper property of *perfect stability*, a strong form of stability which additionally implies an exact result.

1.1 | The statement of the main result

In order to define precisely what we mean by symmetrisable functions and perfect stability, we need to introduce some notation. We write $G = (V, E)$ for a graph with vertex set V and edge set E, and let $v(G) := |V|$ and $e(G) := |E|$. Given $X \subseteq V$, we write

$$G[X] := (X, \{xy \in E : x, y \in X\})$$

for the graph induced by G on X, and $G - X := G[V(G) \setminus X]$, and also $G - x := G - \{x\}$. Write $N_G(x) := \{y \in V : xy \in E\}$.

Fix a positive integer $k \geqslant 3$. Let $\mathcal{G}$ be the family of all finite graphs up to isomorphism and let $\mathcal{G}_n$ consist of graphs with n vertices. Let $\mathcal{P}_n \subseteq \mathcal{G}_n$ be the family of complete partite graphs on n vertices. Suppose we have a function $\gamma : \mathcal{G}_k \to \mathbb{R}$. For a graph $G = (V, E)$ with $v(G) \geqslant k$, define

$$\lambda(G) := \binom{n}{k}^{-1} \sum_{X \in \binom{V}{k}} \gamma(G[X]), \tag{1.1}$$

where $\binom{V}{k}$ is the collection of k-element subsets of V. Thus, $\lambda(G)$ is the expected value of $\gamma(G[X])$ where X is a random k-subset of V. We may also work with

$$\Lambda(G) := \sum_{X \in \binom{V}{k}} \gamma(G[X]) = \binom{n}{k} \lambda(G),$$

which may be more convenient in some calculations. For a vertex $x \in V(G)$, define

$$\Lambda(G,x) := \Lambda(G) - \Lambda(G-x) \;=\; \sum_{X \subseteq \binom{V}{k} : X \ni x} \gamma(G[X]),$$

$$\lambda(G,x) := \binom{n-1}{k-1}^{-1} \Lambda(G,x).$$

Thus $\lambda(G,x)$ is the conditional expectation of $\gamma(G[X])$ where X is a random k-subset of V conditioned on containing x.

Let $\lambda(n)$ be the maximum of $\lambda(G)$ over all n-vertex graphs G and define

$$\lambda_{\max} := \lim_{n\to\infty} \lambda(n).$$

One can easily show that the limit exists. Note that the minimisation problem reduces to a maximisation one just by negating γ, so we will always consider maximising λ here. We can now define what it means for λ to be symmetrisable.

Definition 1 (Symmetrisability). A function λ given by (1.1) is *symmetrisable* if for every $\varepsilon > 0$, there is $n_0 > 0$ such that the following two properties hold for every graph $G = (V, E)$ of order $n \geqslant n_0$:

- *(Sym1)* There is a sequence of graphs $G_0, G_1, \dots, G_m$ on V such that $G_0 = G$; G_m is complete partite and for every $i \in [m]$, we have $\lambda(G_{i-1}) \leqslant \lambda(G_i)$ and $|E(G_{i-1}) \triangle E(G_i)| < \varepsilon\binom{n}{2}$.
- *(Sym2)* If $G - z$ is complete partite with partite sets $V_1, \dots, V_t$, then there is a sequence of graphs $G_0, G_1, \dots, G_m$ on $V(G)$ such that $G_0 = G$; $G_i - z = G - z$; $\lambda(G_{i-1}) \leqslant \lambda(G_i)$; $|E(G_{i-1}) \triangle E(G_i)| \leqslant \varepsilon(n-1)$ for all $i \in [m]$; and for each $j \in [t]$, either $N_{G_m}(z) \supseteq V_j$ or $N_{G_m}(z) \cap V_j = \emptyset$.

Here is an example of a symmetrisable parameter. For graphs F, G with $v(F) \leqslant v(G)$, let $P(F,G)$ be the number of $v(F)$-subsets of $V(G)$ that induce a subgraph isomorphic to F. Let $p(F,G) = P(F,G)/\binom{v(G)}{v(F)}$ be the *induced density* of F in G. Let $\lambda(G) := \sum_{1\leqslant i\leqslant k} a_i p(K_i, G)$ for $a_1, \dots, a_k \in \mathbb{R}$. If we let $\gamma(F) = \sum_{1\leqslant i\leqslant k} a_i p(K_i, F)$ for $F \in \mathcal{G}_k$, then (1.1) holds. (Indeed, for $v(G) \geqslant k \geqslant i$, we have $p(K_i, G) = \sum_{F\in\mathcal{G}_k} p(K_i, F) p(F, G)$ which implies the statement.) As mentioned above, Bollobas [3] showed that $\lambda(n)$ is attained on a complete partite graph and his proof shows that every such λ is, in fact, symmetrisable (for more details and examples, see Section 6). In Section 1.2, we will see a generalisation of this parameter.

Secondly, we define perfect stability. The *edit* and *normalised edit* distances between graphs G and H of the same order n are given by

$$\hat{\Delta}_1(G,H) := \min_{\sigma\in S(G,H)} |E(H) \triangle E(\sigma(G))|, \quad \hat{\delta}_1(G,H) := \frac{2}{n^2}\,\hat{\Delta}_1(G,H),$$

where $S(G,H)$ is the set of bijections from $V(G)$ to $V(H)$. (We also write $S(X) := S(X,X)$.) We further define $\hat{\Delta}_1(G,\mathcal{H}) := \min_{H\in\mathcal{H}} \hat{\Delta}_1(G,H)$ for a family $\mathcal{H}$ of graphs of order n, and define $\hat{\delta}_1(G,\mathcal{H})$ analogously.

Definition 2 (Perfect stability). A graph parameter λ is *perfectly stable* if there exists $C > 0$ such that for every graph G of order $n \geqslant C$, there is a complete partite graph H of order n such that

$$\hat{\delta}_1(G, H) \leqslant C(\lambda(n) - \lambda(G)).$$

We say that a sequence $\boldsymbol{x} = (x_1, x_2, \ldots)$ with $x_1 \geqslant x_2 \geqslant \ldots \geqslant 0$ and $\sum_{i \geqslant 1} x_i \leqslant 1$ is a *maximiser* if there exists a sequence $(H_n)_n$ of complete partite graphs such that, as $n \to \infty$, we have $v(H_n) \to \infty$, $\lambda(H_n) \to \lambda_{\max}$ and for every $i \geqslant 1$, the number of vertices in the ith largest part of H_n is $(x_i + o(1))v(H_n)$. Let $\mathrm{OPT} = \mathrm{OPT}(\lambda)$ be the set of maximisers.

In Section 4, we will show that if OPT is a finite set, then there is $\beta > 0$ such that, for every $\boldsymbol{x} \in \mathrm{OPT}$ and every $i \geqslant 0$, the entry x_i is either 0 or at least β.

Observe that, if λ is perfectly stable, then the only graphs on which λ is maximised are complete partite. Perfect stability has already been proved in several contexts, most notably in Turán-type problems; for example, by Füredi [13], Norin and Yepremyan [30, 31], Pikhurko, Sliačan and Tyros [32] and Roberts and Scott [35].

Definition 3 (Realisation $G_{n,\boldsymbol{x}}$). Given $n \in \mathbb{N}$ and $\boldsymbol{x} = (x_1, x_2, \ldots)$ with $x_1 \geqslant x_2 \geqslant \ldots \geqslant 0$ and $x_0 := 1 - \sum_{i \geqslant 1} x_i \geqslant 0$, define a complete partite graph $G_{n,\boldsymbol{x}}$ with vertex set $[n]$, parts $V_1, \ldots, V_m$ for some m and a set V_0 of universal vertices, that is, $|V_0|$ singleton parts, as follows. If $x_0 = 0$, take a partition $[n] = V_1 \cup \ldots \cup V_m$ with $||V_i| - x_i n| < 1$ and let $V_0 = \emptyset$. Otherwise, for all $i \geqslant 1$ with $x_i n \geqslant 2$, let $|V_i| = \lfloor x_i n \rfloor$ and let V_0 consist of the remaining vertices in $[n]$.

We say that $G_{n,\boldsymbol{x}}$ is *the (n-vertex) realisation* of $\boldsymbol{x}$ and *has $\overline{\mathcal{P}}$-structure* $V_0, \ldots, V_m$.

If H is a graph obtained by adding a new vertex z to $G = G_{n,\boldsymbol{x}}$, we say that z is a *clone* of $u \in V(G)$ if $u \in V_0$ and $N_H(z) = V(G)$, or if $u \notin V_0$ and $N_H(z) = N_G(u)$. The following is one version of our main result, which is also stated as Theorem 3.3, in terms of limits. Roughly speaking, it states that a symmetrisable function λ is perfectly stable if it is 'strict', meaning that it is sensitive to small alterations in a graph.

Theorem 1.1. *Let λ be a symmetrisable function defined as above. Suppose $|\mathrm{OPT}| < \infty$. Suppose also that there exists $c > 0$ such that the following hold for all large n and maximisers $\boldsymbol{x} = (x_1, x_2, \ldots) \in \mathrm{OPT}$, where $G = G_{n,\boldsymbol{x}}$:*

(i) *For all distinct $x, y \in V(G)$ we have $\lambda(G) - \lambda(G \oplus xy) \geqslant cn^{-2}$, where $G \oplus xy$ has vertex set $V(G)$ and edge set $E(G) \triangle \{\{x, y\}\}$.*

(ii) *If G_v is obtained from G by adding a new vertex v which is complete or empty to each part of G (where each V_i, $i \in [m]$ is a part and we have $|V_0|$ singleton parts), then the minimum number of edits at v needed to make v a clone of some existing vertex of G is at most $n(\lambda(G) - \lambda(G_v, v))/c$.*

Then λ is perfectly stable.

As mentioned, see Theorem 3.3 for the 'limit version' of this statement, which concerns

$$\lambda(\boldsymbol{x}) := \lim_{n \to \infty} \lambda(G_{n,\boldsymbol{x}}).$$

One can easily show that this limit exists and that it does not depend on the choice of the part sizes $|V_i|$ in Definition 3 (only on the ratios x_i). The conditions in Theorem 1.1 become a series of inequalities that must be verified for maximisers $\boldsymbol{x}$, which are a finite collection of polynomial inequalities if the number of maximisers is finite and $x_0 = 0$, since, for example, given i, j, the quantity $\lambda(G) - \lambda(G \oplus xy)$ is identical for all $x \in V_i$ and $y \in V_j$. The value of the theorem is that, given the set of maximisers, the conditions are usually very easy to check, so in some sense, the 'combinatorial part' of the problem is solved. It remains to determine the set of maximisers, amounting to a polynomial optimisation, which is unfortunately difficult in general.

1.2 | Applications to inducibility

A large class of problems where symmetrisation was sucessfully applied is the inducibility problem for complete partite graphs. The *inducibility* problem for a graph F is to determine $i(F, n) := \max\{P(F, G) : v(G) = n\}$, the maximum number of induced copies of F that an order-n graph G can have. Note that $p(\overline{F}, \overline{G}) = p(F, G)$, where $\overline{G}$ denotes the complement of G, so $i(\overline{F}, n) = i(F, n)$. Also, consider

$$i(F) := \lim_{n\to\infty} \frac{i(F, n)}{\binom{n}{v(F)}};$$

the limit is known to exist and is, in fact, equivalent to the maximum density of induced copies of F in a graphon W. Brown and Sidorenko [7, Proposition 1] used symmetrisation to prove that if F is complete partite, then for every $n \in \mathbb{N}$, at least one $i(F, n)$-extremal graph is complete partite. Schelp and Thomason [36], also via symmetrisation, extended both the result of Brown and Sidorenko and a result of Bollobás [3] by showing that the same conclusion holds (at least one graph attaining $\lambda(n)$ is complete partite) if the objective function is $\lambda(G) = \sum_F c_F \cdot p(F, G)$, where each F is complete partite, including K_t and $\overline{K}_t$, and c_F is non-negative if F is not a clique. Their proof (which is essentially the same as that of Bollobás [3]) implies that this parameter is symmetrisable (see Section 6 for a proof).

Lemma 1.2 [36]. *The function $\lambda(G) := \sum_F c_F \cdot p(F, G)$ is symmetrisable, where each F is complete partite and $c_F \geqslant 0$ if F is not a clique.*

In particular, Theorem 1.1 applies to the inducibility problem for complete partite graphs. To the best of our knowledge, for every instance of this problem where the set of maximisers is known, we can prove perfect stability.

Pippenger and Golumbic [34] determined $i(K_{s,t}, n)$ for all s, t with $|s - t| \leqslant 1$, observing that the complete balanced bipartite graph is an extremal graph. Some of these results were independently reproved in [5]. Brown and Sidorenko [7] showed that $i(K_{s,t}, n)$ with $st \geqslant 2$ is attained by a complete bipartite graph, and that if $\binom{t-s}{2} \leqslant s \leqslant t$, then the unique maximiser is $(\frac{1}{2}, \frac{1}{2}, 0, \ldots)$. Perhaps surprisingly, this does not mean that $K_{\lfloor n/2\rfloor,\lceil n/2\rceil}$ is optimal for $i(K_{s,t}, n)$, and they show that if $3n = 4a^2 + 4$ for a large integer a, then $K_{n/2-a,n/2+a}$ is optimal for $K_{3,1}$. We prove a corresponding stability result for complete bipartite graphs.

Theorem 1.3. *Let $s, t \in \mathbb{N}$ with $st \geqslant 2$. Then $p(K_{s,t}, \cdot)$ is perfectly stable, $i(K_{s,t}) = \binom{s+t}{s} M_{s,t}$ and there is a unique maximiser $(\alpha, 1 - \alpha, 0, 0, 0, \ldots)$, where $\alpha \in [\frac{1}{2}, 1]$ maximises*

$$f_{s,t}(\alpha) := \alpha^s(1 - \alpha)^t + \alpha^t(1 - \alpha)^s$$

and $M_{s,t} := \max_{x\in[\frac{1}{2},1]} f_{s,t}(x)$ for $s \neq t$, and $M_{s,s} := \frac{1}{2} \max_{x\in[\frac{1}{2},1]} f_{s,s}(x)$.

Bollobás, Egawa, Harris and Jin [4] studied the inducibility problem for complete equipartite graphs. They showed that if the size t of each part is not too small compared to the number r of parts, then the complete balanced r-partite graph $T_r(n)$ is the *unique* extremal graph for each large n. This strengthened an earlier work of Brown and Sidorenko [7] which showed that $T_r(n)$ is an asymptotically extremal construction (without proving any uniqueness) — that is, $(\frac{1}{r}, \dots, \frac{1}{r}, 0, \dots)$ with $\frac{1}{r}$ repeated r times is an element of OPT. We prove a corresponding stability result.

Theorem 1.4. *Let $r, t \geqslant 2$ be integers and let $K_r(t)$ denote the complete r-partite graph with parts of size t. Suppose that $t > 1 + \log r$ (denoting the natural logarithm by $\log$). Then $p(K_r(t), \cdot)$ is perfectly stable, $i(K_r(t)) = \frac{(tr)!}{r!t!^r r^{tr}}$, and the unique maximiser is $(\underbrace{\frac{1}{r}, \dots, \frac{1}{r}}_{r}, 0, \dots)$.*

Interestingly, if the above lower bound on t in terms of r does not hold, then $(\frac{1}{r}, \dots, \frac{1}{r}, 0, \dots) \notin$ OPT (see [7]).

Finally, we obtain perfect stability for every previously unknown complete partite graph F on $k \leqslant 5$ vertices. For this, note that trivially K_k and $\overline{K_k}$ have unique maximisers $(0, 0, \dots)$, $(1, 0, \dots)$, respectively. If $F = K_{s,t}$ is bipartite, then Theorem 1.3 implies that the unique maximiser $(\alpha, 1-\alpha, 0, \dots)$ maximises $\alpha^s(1-\alpha)^t + \alpha^t(1-\alpha)^s$. Solving this, we see that $p(K_{s,t}, \cdot)$ has unique maximiser $(\frac{1}{2}, \frac{1}{2}, 0, \dots)$ for all $s + t \leqslant 5$ apart from $\{s, t\} = \{4, 1\}$, and here $p(K_{4,1}, \cdot)$ has unique maximiser $(\frac{4}{5}, \frac{1}{5}, 0 \dots)$. Pikhurko, Sliačan and Tyros [32] showed that $K_{2,1,1}$ is perfectly stable with unique maximiser $(\frac{1}{5}, \dots, \frac{1}{5}, 0, \dots)$, and that $K_{2,2,1}$ is perfectly stable with unique maximiser $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \dots)$ (we can also recover these results but do not provide proofs here). The remaining F are $K_{3,1,1}$ and $K_{2,1,1,1}$. Flag algebra calculations of Even-Zohar and Linial [10] give numerical upper bounds for these $i(F)$. Also, they provided lower bound constructions; these appear to match for both $K_{3,1,1}$ and $K_{2,1,1,1}$. They speculated that their lower bound constructions are tight in both cases. We confirm this and prove perfect stability for these F. (After this paper was submitted, Liu, Mubayi and Reiher [29, Theorem 1.13] determined the value of $i(K_t^-)$ for every t, where $K_t^- = K_{2,1,\dots,1}$ is the complete graph of order t minus one edge.)

Theorem 1.5. *$p(K_{2,1,1,1}, \cdot)$ is perfectly stable, $i(K_{2,1,1,1}) = \frac{525}{1024}$, and the unique maximiser is $(\frac{1}{8}, \dots, \frac{1}{8}, 0, \dots)$.*

Theorem 1.6. *$p(K_{3,1,1}, \cdot)$ is perfectly stable, $i(K_{3,1,1}) = \frac{216}{625}$, and the unique maximiser is $(\frac{3}{5}, 0, \dots)$.*

The latter is particularly interesting since the extremal graph contains a clique part: it is a clique with a clique of proportion $3/5$ removed. This demonstrates that allowing maximisers $\boldsymbol{x}$ with $x_0 = 1 - \sum_{i \geqslant 1} x_i > 0$ in our theory — which complicates matters somewhat — is essential in giving a full picture.

We remark that the case $\lambda(\cdot) = -p(\overline{K_3}, \cdot) - p(K_3, \cdot)$ (which is not a function as in Lemma 1.2) is given by a classical theorem of Goodman, who determined this value exactly. Here, asymptotically extremal graphs are those for which all but $o(n)$ vertices have degree $\frac{n}{2} + o(n)$ (including many graphs which are not complete partite). (Note that $p(\overline{K_3}, \cdot) + p(K_3, \cdot)$ is trivially maximised by the complete and empty graphs.) It remains a major open problem to determine $\lambda_{\max}$ for $\lambda(\cdot) = -p(\overline{K_4}, \cdot) - p(K_4, \cdot)$.

Pikhurko, Sliačan and Tyros [32] were able to prove perfect stability for $i(F, n)$ for several small graphs F via flag algebra calculations. The graphs they considered were $C_4 = K_{2,2}$, $K_{2,1,1}$, $K_{3,2}$, $K_{2,2,1}$, as well as the non-complete partite graphs $P_3 \cup K_2$, the 'Y' graph and the paw graph which we do not define. Their results extend inducibility results obtained in [7], [34], and by Hirst in [18]. Our Theorem 1.3 in particular reproves the cases $K_{2,2}$ and $K_{3,2}$ from [32].

Before stating the limit version of our main theorem in Section 3, we give here an illustration of it in the case $F = C_4$. (Perfect stability was already proved here in [32].) It is easy to see that OPT consists only of the unique vector $(\frac{1}{2}, \frac{1}{2}, 0, \ldots)$ with $\lambda_{\max} = \frac{3}{8}$. Thus, in order to apply our criterion, we have to check that, starting with $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$, the following two properties hold: (i) if we add an edge into a part or remove an edge across, then we decrease the number of induced copies of C_4 by $\Omega(n^2)$; (ii) if we add a new vertex v which is either isolated or connected to every other vertex, the number of induced copies of C_4 containing v is at most $(1 - \Omega(1))\, \frac{3}{8}\binom{n}{3}$. Both properties trivially hold, so the inducibility problem for C_4 is indeed perfectly stable by Theorem 1.1.

The following conjecture seems plausible.

Conjecture 1. *The inducibility problem for F is perfectly stable for every complete partite F.*

However, it is not the case that every problem with $\lambda = \sum_F c_F \cdot p(F, \cdot)$ is perfectly stable, where each F is complete partite, and $c_F \geqslant 0$ if F is not a clique. Indeed, if $k \geqslant 3$ and the sum is over all complete partite F on k vertices, and each $c_F = 1$, then every k-vertex subset of every complete partite graph contributes (the maximum value of) 1 to Λ, so OPT is the set of *all* $\boldsymbol{x}$ with $x_1 \geqslant x_2 \geqslant \ldots \geqslant 0$ and $\sum_{i \geqslant 1} x_i \leqslant 1$. Let us show that λ is not perfectly stable. Indeed, if it is, there is C such that for every graph G of order $n \geqslant C$, there is a complete partite H such that $\hat{\delta}_1(G, H) \leqslant C(\lambda(n) - \lambda(G))$. Choose $1/n \ll c \ll 1/C$. Starting with K_n, remove every edge with both endpoints inside a set A of size $5cn$ and add into A a blow-up of C_5 with each part $A_1, \ldots, A_5$ of size cn, to obtain an n-vertex graph G. Then $\hat{\delta}_1(G, H) = \Omega(c^2)$ for every complete partite H, but $\lambda(n) - \lambda(G) = 1 - \lambda(G) = O(c^3)$. Indeed, a subset of G is not complete partite only if it contains at least three vertices in A. So, the fraction of subsets inducing a non-complete partite graph is $O(c^3)$. This is a contradiction.

Finally, it would be remiss not to remark on the inducibility problem for non-complete partite graphs, for which the present paper does not apply, and which is in general wide open (see [10] for a list of known results of order up to 5). The outstanding open problem in the area is determining $i(P_4)$, the smallest unsolved case, for which there is not even a conjectured value. Hatami, Hirst and Norin proved that extremal graphs of large blow-ups are essentially blow-ups themselves [16]. Graphs with more interesting structure appear as extremal graphs for other F. An important longstanding conjecture of Pippenger and Golumbic [34] is that $i(C_k) = k!/(k^k - k)$ for $k \geqslant 5$, attained by the *iterated blow-up* of C_k. Balogh, Hu, Lidický and Pfender [1] proved this conjecture for $k = 5$: they obtained an exact result for $\lambda(\cdot) = p(C_5, \cdot)$ and showed that if n is a power of 5, then the unique graph attaining $i(C_5, n)$ is an iterated blow-up of a 5-cycle. There has recently been progress on the general conjecture in [17, 22]. Yuster [39] and independently Fox, Huang and Lee [11] proved that for almost all graphs F, the extremal graph is the iterated blow-up of F. Fox, Sauermann and Wei [12] considered graphs H obtained by removing a small number of vertices from a random Cayley graph $\tilde{H}$ of an abelian group, showing that here the extremal graph is the iterated blow-up of $\tilde{H}$ (not of H). Liu, Mubayi and Reiher [29] began a systematic study of the feasible region of induced graphs; that is, the set of points (x, y) in the unit square for which there is a graph of edge density approaching x with F-density approaching y. The inducibility problem seeks the maximum y-value of such a point.

The directed analogue of the inducibility problem is also actively studied, for example, for stars [20, 21], paths [9] and 4-vertex graphs [6, 8, 19].

1.3 | Structure of the paper

The rest of the paper is organised as follows. In Section 2, we introduce the partite limit space corresponding to the collection of limits of complete partite graphs which we will need to prove our main result. In Section 3, we define the notion of strictness in terms of elements of this space and give a limit version of our main result, Theorem 3.3. The main result of Section 4 is that when OPT is finite, all part ratios of extremal graphs are bounded away from 0. We prove Theorem 3.3 in Section 5. We present some applications of Theorem 3.3 to the inducibility problem (Theorems 1.3–1.6) in Section 6. Section 7 contains some concluding remarks.

We denote by $\mathbb{N} := \{1, 2, \dots\}$ and $\mathbb{N}_0 := \{0, 1, \dots\}$ the sets of respectively positive and non-negative integers.

2 | THE PARTITE LIMIT SPACE

We will work in a space $\overline{\mathcal{P}}$, the *partite limit space*, which is in some sense the completion of the set of complete partite graphs. The aim of this section is to define $\overline{\mathcal{P}}$ and a metric δ_{edit} on this set, which will essentially generalise edit distance in graphs. We prove that this yields a compact metric space upon which λ can be extended continuously (Lemma 2.5). Thus, the set OPT of maximisers of λ in $\overline{\mathcal{P}}$ is non-empty. We define

$$\overline{\mathcal{P}} := \left\{ \boldsymbol{x} = (x_1, x_2, \dots) : x_1 \geqslant x_2 \geqslant \dots \geqslant 0 \text{ and } \sum_{i \geqslant 1} x_i \leqslant 1 \right\}.$$

As usual, $\text{supp}(\boldsymbol{x}) := \{i \geqslant 1 : x_i > 0\}$, and we also define $\text{supp}^*(\boldsymbol{x}) := \text{supp}(\boldsymbol{x}) \cup \{0\}$ if $\sum_{i \geqslant 1} x_i < 1$, and $\text{supp}^*(\boldsymbol{x}) := \text{supp}(\boldsymbol{x})$ otherwise. For $\beta > 0$, we write

$$\overline{\mathcal{P}}_\beta := \{\boldsymbol{x} \in \overline{\mathcal{P}} : x_i \geqslant \beta \; \forall \, i \in \text{supp}^*(\boldsymbol{x})\}.$$

Write $\mathbf{0} := (0, 0, \dots)$. Given $\boldsymbol{x}, \boldsymbol{x}_n \in \overline{\mathcal{P}}$, we will always write $\boldsymbol{x} = (x_1, x_2, \dots)$ and $\boldsymbol{x}_n = (x_{n,1}, x_{n,2}, \dots)$ and correspondingly $x_0 := 1 - \sum_{i \geqslant 1} x_i$ and $x_{n,0} := 1 - \sum_{i \geqslant 1} x_{n,i}$. A complete partite graph $G = K(V_1, \dots, V_m)$ on vertex set $[n]$ with $|V_1| \geqslant \dots \geqslant |V_m|$ corresponds to the vector

$$\boldsymbol{x}_G := (|V_1|/n, \dots, |V_m|/n, 0, \dots).$$

We write $\mathcal{P}$ for the set of those elements $\boldsymbol{x}$ of $\overline{\mathcal{P}}$ with finitely many non-zero entries all of which are rational, thus corresponding to the set of complete partite graphs. Somewhat conversely, we have the construction $G_{n,\boldsymbol{x}}$ from Definition 3. For example, we have $G_{n,\mathbf{0}} \cong K_n$, $G_{n,(1,0,\dots)} \cong \overline{K}_n$ and (assuming $n = 2\ell$ is even) $G_{n,(\frac{1}{2},\frac{1}{2},0,\dots)} \cong K_{\ell,\ell}$, but we cannot take, say, any $K_{a,b,1}$ for $G_{n,(x,1-x,0,\dots)}$.

2.1 | The measure-theoretic and graphon perspectives

For each $\boldsymbol{x} \in \overline{\mathcal{P}}$, one can define a probability measure $\mu_{\boldsymbol{x}}$ on $\mathbb{N}_0$ by setting $\mu_{\boldsymbol{x}}(\{i\}) = x_i$ and then let

$$\mathcal{M} := \left\{\mu_{\boldsymbol{x}} : \boldsymbol{x} \in \overline{\mathcal{P}}\right\}.$$

It is very natural to define the corresponding collection of 'complete partite' graphons (which will be used in Section 4). A *graphon* is a quadruple $Q = (\Omega, \mathcal{B}, \mu, W)$, where $(\Omega, \mathcal{B}, \mu)$ is a standard probability space and $W : \Omega \times \Omega \to [0,1]$ is a symmetric measurable function. For every graph G, we define the corresponding graphon $Q_G = (V, 2^V, \mu, A_G)$ where μ is the uniform measure on the finite set V and $A_G : V \times V \to \{0,1\}$ is the adjacency function of G. For a graph F on $[k]$, we write

$$p(F,Q) := \frac{k!}{|\text{aut(F)}|} \int_{\Omega^k} \prod_{ij \in E(F)} W(x_i, x_j) \prod_{ij \in E(\overline{F})} (1 - W(x_i, x_j)) d\mu(x_1) \ldots d\mu(x_k),$$

where aut(F) is the group of automorphisms of F. In the literature, one usually encounters $t_{\text{ind}}(F,Q)$ which is the above without the normalisation factor. Two graphons Q, Q' are *equivalent* or *weakly isomorphic* if $p(F,Q) = p(F,Q')$ for every graph F. A sequence of graphons $(Q_n : n \in \mathbb{N})$ is said to *converge* to a graphon Q if $\lim_{n\to\infty} p(F, Q_n) = p(F,Q)$ for every graph F. A Q*-random graph* of order k is obtained by sampling k random points $v_1, \ldots, v_k \in (\Omega, \mu)$ uniformly and independently, and adding each edge $x_i x_j$ with probability $W(x_i, x_j)$.

Now let $Q_{\boldsymbol{x}} := (\mathbb{N}_0, 2^{\mathbb{N}_0}, \mu_{\boldsymbol{x}}, K)$ where $K(i,j) := 0$ if $i = j \geqslant 1$ and $K(i,j) := 1$ otherwise, that is, if $i \neq j$ or $i = j = 0$. Then define

$$\mathcal{Q} := \left\{Q_{\boldsymbol{x}} : \boldsymbol{x} \in \overline{\mathcal{P}}\right\}.$$

There are various characterisations of weak isomorphism (see [26, Theorem 13.10]). All we will need is the easy fact that for distinct $\boldsymbol{x}, \boldsymbol{y} \in \overline{\mathcal{P}}$, their graphons $Q_{\boldsymbol{x}}, Q_{\boldsymbol{y}}$ are not weakly isomorphic. Indeed, if $i \geqslant 1$ is the minimum integer with $x_i \neq y_i$, say $x_i > y_i$, then it is not hard to see directly that the edgeless graph of sufficiently large order n has strictly larger density in $\boldsymbol{x}$ than in $\boldsymbol{y}$.

The spaces $\overline{\mathcal{P}}$, $\mathcal{M}$ and $\mathcal{Q}$ are equivalent and one can take any of these perspectives, but in this paper, we mainly work with $\overline{\mathcal{P}}$ (and briefly use $\mathcal{Q}$ in Section 4). The space $\mathcal{Q}$ was used in [2] by Bennett, Dudek, Lidický and Pikhurko who determined the minimum C_5-density in graphs of edge density $\frac{k-1}{k}$ for integers k. They used $\mathcal{Q}$ to prove a corresponding stability result. Therefore, we hope that the theory concerning $\overline{\mathcal{P}}$ (and, by extension, $\mathcal{M}$ and $\mathcal{Q}$) developed in this section may be useful for other extremal problems where the extremal graphs are complete partite.

2.2 | The edit metric

We would like to define a metric on $\overline{\mathcal{P}}$ which will correspond to the edit distance between graphs. Firstly, we define edit distance between two graphs of possibly different orders, often called the *fractional edit distance*. Given a graph G, let $G^{(n)}$ be an n-vertex almost uniform blow-up of G, that is, we replace each vertex $x \in V(G)$ by an independent set I_x, where each $||I_x| - |I_y|| \leqslant 1$, and $\sum_{x\in V(G)} |I_x| = n$, and add every edge between I_x and I_y whenever $xy \in E(G)$. Then let

$$\delta_{\text{edit}}(G, H) := \lim_{n\to\infty} \hat{\delta}_1(G^{(n)}, H^{(n)}).$$

It is easy to see that the limit exists; in fact, its value can be computed via a linear program with $v(G) \times v(H)$ variables that considers all fractional overlays between the vertex sets of G and H, cf. for example, [33, Equation (3)]. We also define for a family $\mathcal{H}$ of graphs $\delta_{\text{edit}}(G, \mathcal{H}) := \lim_{n\to\infty} \hat{\delta}_1(G^{(n)}, \{H^{(n)} : H \in \mathcal{H}\})$. We define the distance between $\boldsymbol{x}, \boldsymbol{y} \in \overline{\mathcal{P}}$ to be

$$\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) := \lim_{n\to\infty} \hat{\delta}_1(G_{n,\boldsymbol{x}}, G_{n,\boldsymbol{y}}).$$

For a graph G, define also $\delta_{\text{edit}}(\boldsymbol{x}, G) := \lim_{n\to\infty} \hat{\delta}_1(G_{n,\boldsymbol{x}}, G^{(n)})$ and, for a family $\mathcal{H}$ of graphs, $\delta_{\text{edit}}(\boldsymbol{x}, \mathcal{H}))$ in the obvious way. Again, the existence of the limit in these definitions is easy to establish. Note that the normalisation factor $\frac{2}{n^2}$ in the 'usual' edit distance $\hat{\delta}_1$ is motivated by vertices of G corresponding to independent sets of relative size $\frac{1}{n}$. The distances $\hat{\delta}_1$ and δ_{edit} are not the same even for graphs of the same order, due to rounding; see examples of Matsliah (see Appendix B in [15]) and Pikhurko [33]. The following lemma implies that we are free to interchange δ_{edit} and $\hat{\delta}_1$ in matters of convergence, and that with respect to δ_{edit} we are free to interchange H and $\boldsymbol{x}_H$ when H is complete partite.

Lemma 2.1. *We have the following.*

(i) $\delta_{\text{edit}}(G, H) \leqslant \hat{\delta}_1(G, H) \leqslant 3\delta_{\text{edit}}(G, H)$ *for graphs G, H with the same order.*
(ii) $\delta_{\text{edit}}(H, \boldsymbol{x}_H) = 0$ *and* $\delta_{\text{edit}}(\boldsymbol{x}, H) = \delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{x}_H)$ *and* $\delta_{\text{edit}}(G, H) = \delta_{\text{edit}}(\boldsymbol{x}_G, \boldsymbol{x}_H)$ *for all $\boldsymbol{x} \in \overline{\mathcal{P}}$ and complete partite graphs G, H.*
(iii) δ_{edit} *satisfies the triangle inequality on $\overline{\mathcal{P}}$.*

Proof. The non-trivial inequality of part (i) was proved in [33, Lemma 14]. For (ii), let H have h vertices. Then $\boldsymbol{x}_{H^{(nh)}} = \boldsymbol{x}_H$ and $G_{nh,\boldsymbol{x}_H} = (G_{h,\boldsymbol{x}_H})^{(n)} = H^{(nh)}$ for any integer n. Since any subsequence of $(\hat{\delta}_1(H^{(m)}, G_{m,\boldsymbol{x}_H}))_m$ converges to $\delta_{\text{edit}}(H, \boldsymbol{x}_H)$, we have

$$\delta_{\text{edit}}(H, \boldsymbol{x}_H) = \lim_{n\to\infty} \hat{\delta}_1(H^{(nh)}, G_{nh,\boldsymbol{x}_H}) = \lim_{n\to\infty} \hat{\delta}_1(H^{(nh)}, H^{(nh)}) = 0.$$

The remaining parts of (ii) now follow from (iii) which is immediate since $\hat{\delta}_1$ satisfies the triangle inequality on the set of graphs of the same given order. □

This notion of edit distance is very natural, yet rather unwieldy to work with. The following easy facts concerning it will be useful. Recall first that x_0 is not an entry in $\boldsymbol{x} = (x_1, x_2, \ldots)$, so, for example, $\|\boldsymbol{x}\|_1 = \sum_{i\geqslant 1} |x_i| = 1 - x_0$.

Proposition 2.2. *For all $\boldsymbol{x}, \boldsymbol{y} \in \overline{\mathcal{P}}$, we have that*

(i) $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) \leqslant 2\|\boldsymbol{x} - \boldsymbol{y}\|_1$.
(ii) $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{0}) = \|\boldsymbol{x}\|_2^2$.
(iii) $\delta_{\text{edit}}(\boldsymbol{x}, (x_1, \ldots, x_M, 0, \ldots)) \leqslant \sum_{i>M} x_i^2$ *for all $M \geqslant 1$.*

Proof. For (i), consider large $n \in \mathbb{N}$ and $G_{n,\boldsymbol{x}}, G_{n,\boldsymbol{y}}$ with $\overline{\mathcal{P}}$-structures $V_0, V_1, \dots, V_m$ and $U_0, U_1, \dots, U_\ell$, respectively, where without loss of generality $\ell \leqslant m$. For convenience, let $U_{\ell+1} = \cdots = U_m = \emptyset$. Let $\sigma \in S([n])$ be a permutation (and recall that $V(G_{n,\boldsymbol{x}}) = V(G_{n,\boldsymbol{y}}) = [n]$). For all $0 \leqslant i, j \leqslant m$, let $X_{ij} = \sigma(V_i) \cap U_j$. A pair of vertices is included in the symmetric difference $E(G_{n,\boldsymbol{x}}) \triangle E(\sigma(G_{n,\boldsymbol{y}}))$ if and only if either it lies in V_i for some $i \geqslant 1$ but not in X_{ij} for any $j \in [m]$, or lies in U_j for some $j \geqslant 1$ but not in X_{ij} for any $i \in [m]$. Thus,

$$E(G_{n,\boldsymbol{x}}) \triangle E(\sigma(G_{n,\boldsymbol{y}})) = \sum_{i\in[m]}\left(\binom{|V_i|}{2} - \sum_{j\in[m]}\binom{|X_{ij}|}{2}\right) + \sum_{j\in[m]}\left(\binom{|U_j|}{2} - \sum_{i\in[m]}\binom{|X_{ij}|}{2}\right). \tag{2.1}$$

Take $\sigma \in S([n])$ so that for all $i \geqslant 0$, $\sigma(V_i) \subseteq U_i$ whenever $|V_i| \leqslant |U_i|$, and $\sigma(V_j) \supseteq U_j$ whenever $|U_j| \leqslant |V_j|$ (and σ is otherwise arbitrary). If $|U_k| < |V_k|$, then $|X_{kk}| = |U_k|$ and $|X_{ik}| = 0$ for all $i \neq k$ and thus

$$\begin{aligned}\binom{|V_k|}{2} - \sum_{j\in[m]}\binom{|X_{kj}|}{2} + \binom{|U_k|}{2} - \sum_{j\in[m]}\binom{|X_{jk}|}{2} &= \binom{|V_k|}{2} - \sum_{j\in[m]}\binom{|X_{jk}|}{2}\\ \leqslant \binom{|V_k|}{2} - \binom{|X_{kk}|}{2} &\leqslant \frac{1}{2}(|V_k|^2 - |U_k|^2) + |U_k|.\end{aligned}$$

Then

$$\begin{aligned}\hat{\Delta}_1(G_{n,\boldsymbol{x}}, G_{n,\boldsymbol{y}}) &\leqslant \sum_{k\in[m]}\left(\frac{1}{2}\Big|\,|V_k|^2 - |U_k|^2\,\Big| + \min\{|U_k|, |V_k|\}\right)\\ &\leqslant n\sum_{k\in[m]}\left(\big|\,|V_k| - |U_k|\,\big| + y_k + o(1)\right) \leqslant n^2\|\boldsymbol{x} - \boldsymbol{y}\|_1 + O(n).\end{aligned}$$

So $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) \leqslant 2\|\boldsymbol{x} - \boldsymbol{y}\|_1$, as required.

Parts (ii) and (iii) are clear. ☐

Note, however, that convergence in ℓ_1 does not give the same topology as pointwise convergence, by considering for each $n \in \mathbb{N}$ the sequence $\boldsymbol{x}_n$ given by $x_{n,i} = 1/n$ for all $i \in [n]$ and $x_{n,i} = 0$ otherwise. We have that $\|\boldsymbol{x}_n\|_1 = 1$ for all n, while $\boldsymbol{x}_n$ clearly converges pointwise to $\boldsymbol{0}$ and by Proposition 2.2(ii), we see that $\delta_{\text{edit}}(\boldsymbol{x}_n, \boldsymbol{0}) = \frac{1}{n} \to 0$ as $n \to \infty$. On the other hand, convergence in δ_{edit} is equivalent to pointwise convergence, as we show in the next lemma.

Proposition 2.3. *In the space $\overline{\mathcal{P}}$, convergence in edit distance is equivalent to pointwise convergence. That is, whenever $(\boldsymbol{x}_n)_n$ is a sequence in $\overline{\mathcal{P}}$ and $\boldsymbol{x} \in \overline{\mathcal{P}}$, we have that* $\lim_{n\to\infty}\delta_{\text{edit}}(\boldsymbol{x}_n, \boldsymbol{x}) = 0$ *if and only if for all $i \in \mathbb{N}$, we have that* $\lim_{n\to\infty}|x_{n,i} - x_i| = 0$.

Proof. Let $(\boldsymbol{x}_n)_n$ be a sequence in $\overline{\mathcal{P}}$ and let $\boldsymbol{x} \in \overline{\mathcal{P}}$. Fix an arbitrary $\varepsilon > 0$.

Suppose first that $\boldsymbol{x}_n \to \boldsymbol{x}$ pointwise. We need to show that $\delta_{\text{edit}}(\boldsymbol{x}_n, \boldsymbol{x}) < \varepsilon$ for sufficiently large n. Since $\sum_{i\geqslant 1} x_i \leqslant 1$ and $x_1 \geqslant x_2 \geqslant \dots \geqslant 0$, there exists an integer $M > 0$ such that $\sum_{i\geqslant M} x_i < \varepsilon/8$, in particular, $x_M < \varepsilon/8$. As $\boldsymbol{x}_n \to \boldsymbol{x}$ pointwise, there exists n_0 such that, for all $i \leqslant M$ and for all integers $n \geqslant n_0$, we have that $|x_{n,i} - x_i| < \varepsilon/(8M)$. In particular, since $x_{n,j}$ is non-increasing with j, we have for all integers $n \geqslant n_0$ and $j \geqslant M$ that $x_{n,j} < \varepsilon/4$. Let $\boldsymbol{y} := (x_1, \dots, x_M, 0, \dots)$ and, for each $n \in \mathbb{N}$, define $\boldsymbol{y}_n := (x_{n,1}, \dots, x_{n,M}, 0, \dots)$. Let $n \geqslant n_0$ be an integer. Then by Proposition 2.2(i), $\delta_{\mathrm{edit}}(\boldsymbol{x}, \boldsymbol{y}) \leqslant 2\|\boldsymbol{x} - \boldsymbol{y}\|_1 = 2\sum_{i>M} x_i < \frac{\varepsilon}{4}$. Similarly, by Proposition 2.2(ii) and (iii),

$$\delta_{\mathrm{edit}}(\boldsymbol{x}_n, \boldsymbol{y}_n) \leqslant \delta_{\mathrm{edit}}((x_{n,M+1}, x_{n,M+2}, \dots), \boldsymbol{0}) = \sum_{i>M} x_{n,i}^2 \leqslant \sup_{i>M} x_{n,i} \cdot \sum_{i>M} x_{n,i} \leqslant x_{n,M+1} \leqslant \frac{\varepsilon}{4}.$$

But also $\delta_{\mathrm{edit}}(\boldsymbol{y}, \boldsymbol{y}_n) \leqslant 2\|\boldsymbol{y} - \boldsymbol{y}_n\|_1 = 2\sum_{i\leqslant M} |x_{n,i} - x_i| < \frac{\varepsilon}{4}$. By Lemma 2.1(iii), δ_{edit} defined on $\overline{\mathcal{P}}$ satisfies the triangle inequality. Thus, we have $\delta_{\mathrm{edit}}(\boldsymbol{x}_n, \boldsymbol{x}) \leqslant \varepsilon$ whenever $n \geqslant n_0$. Thus, $\boldsymbol{x}_n \to \boldsymbol{x}$ in edit distance, as required.

Conversely, suppose now that $(\boldsymbol{x}_n)_n$ converges to $\boldsymbol{x}$ in edit distance δ_{edit}. Let $i \geqslant 1$. We need to show that there exists $n_0 > 0$ such that for all $n > n_0$, we have $|x_{n,i} - x_i| \leqslant \varepsilon$. Now, there exists $n_0 > 0$ such that for all $n > n_0$, there is a permutation $\sigma : [n] \to [n]$ such that

$$\hat{\Delta}_1(G_{n,\boldsymbol{x}_n}, G_{n,\boldsymbol{x}}) = |E(G_{n,\boldsymbol{x}_n}) \triangle E(\sigma(G_{n,\boldsymbol{x}}))| \leqslant (\varepsilon n/12)^2.$$

Let $n > n_0$. For $A \subseteq [n]$, denote by $\sigma(A)$ and $\sigma^{-1}(A)$ the image and pre-image of A, respectively. By definition, $G_{n,\boldsymbol{x}_n}$ has a vertex partition $V_{n,0} \cup V_{n,1} \cup \dots \cup V_{n,m}$, where $V_{n,0}$ is a clique, $V_{n,i}$ is an independent set for all $i \in [m]$ and $G_{n,\boldsymbol{x}_n}$ is complete between every distinct $V_{n,i}$ and $V_{n,j}$. Define $V_0 \cup V_1 \cup \dots \cup V_\ell$ analogously for $G_{n,\boldsymbol{x}}$. So,

$$|V_{n,i}| = x_{n,i}n + O(1) \quad \text{and} \quad |V_i| = x_i n + O(1) \quad \text{for all } i \geqslant 0. \tag{2.2}$$

Choose an ordering of the vertices of $G_{n,\boldsymbol{x}_n}$ so that a vertex $u \in V_{n,i}$ comes before a vertex $v \in V_{n,j}$ if $1 \leqslant i < j$; or if $i \neq 0$ and $j = 0$. Choose an analogous ordering for $V(G_{n,\boldsymbol{x}})$. Note the following trivial equality:

$$|\sigma(V_j) \triangle V_{n,i}| = |V_j \triangle \sigma^{-1}(V_{n,i})| \quad \text{for all} \quad i, j \in \mathbb{N}_0. \tag{2.3}$$

We first show that for each vertex part $V_{n,i}$ which is not too small, there is a unique part V_{j_i} such that σ maps most of $V_{n,i}$ to V_{j_i}. Given $i \in \{0, 1, \dots, m\}$ and $j \in \{0, 1, \dots, \ell\}$, we say that i is *j-good* if $|\sigma(V_j) \triangle V_{n,i}| < \varepsilon n/4$.

Claim 2.4. Let $A := \{i \in [m] : |V_{n,i}| \geqslant \varepsilon n/2\}$. Then there exists $B \subseteq [\ell]$ with $|A| = |B|$ and a bijection $\mu : A \to B$ such that, for every $i \in A$, we have that i is j-good if and only if $j = \mu(i)$.

Proof. Let $i \in A$. Note first that i is not 0-good. Indeed, this follows from

$$(\varepsilon n/12)^2 \geqslant \hat{\Delta}_1(G_{n,\boldsymbol{x}_n}, G_{n,\boldsymbol{x}}) \geqslant \binom{|\sigma(V_0) \cap V_{n,i}|}{2}.$$

So, $\sum_{j\in\mathbb{N}} |\sigma(V_j) \cap V_{n,i}| = |V_{n,i}| - |\sigma(V_0) \cap V_{n,i}| > \varepsilon n/4$. Suppose now that i is not j-good for any $j \in [\ell]$. Then since $G_{n,\boldsymbol{x}}[V_j]$ and $G_{n,\boldsymbol{x}_n}[V_{n,i}]$ are empty graphs, and both $G_{n,\boldsymbol{x}}$ and $G_{n,\boldsymbol{x}_n}$ are complete partite graphs,

$$|E(G_{n,\boldsymbol{x}_n}) \triangle E(\sigma(G_{n,\boldsymbol{x}}))| \geqslant \sum_{j\in\mathbb{N}} |\sigma(V_j) \triangle V_{n,i}| \cdot |\sigma(V_j) \cap V_{n,i}|$$

$$\geqslant \varepsilon n/4 \cdot \sum_{j\in\mathbb{N}} |\sigma(V_j) \cap V_{n,i}| > (\varepsilon n/4)^2,$$

a contradiction. Thus, there is some $j_i \in \mathbb{N}$ for which i is j_i-good. We claim that we can set $\mu(i) := j_i$ and $B := \{\mu(i) : i \in A\}$. We first show that this is well defined, that is, j_i is unique. Fix an arbitrary $j' \in [\ell] \setminus \{j_i\}$. Since σ is a permutation, $\sigma(V_{j'}) \cap \sigma(V_{j_i}) = \emptyset$, and therefore,

$$|\sigma(V_{j'}) \triangle V_{n,i}| \geqslant |\sigma(V_{j_i}) \cap V_{n,i}| \geqslant |V_{n,i}| - |\sigma(V_{j_i}) \triangle V_{n,i}| > \varepsilon n/4,$$

that is, i is not j'-good. It remains to show that μ is injective, that is, that if $i' \in A \setminus \{i\}$, we have that i' is not j_i-good. By (2.3), it suffices to show that $|V_{j_i} \triangle \sigma^{-1}(V_{n,i'})| \geqslant \varepsilon n/4$. Since σ^{-1} is a permutation, $\sigma^{-1}(V_{n,i'}) \cap \sigma^{-1}(V_{n,i}) = \emptyset$, and therefore, $|V_{j_i} \triangle \sigma^{-1}(V_{n,i'})| \geqslant |V_{j_i} \cap \sigma^{-1}(V_{n,i})| > \varepsilon n/4$ as desired, where the last inequality follows from i being j_i-good and (2.3). This completes the proof of the claim. □

We are now ready to prove the desired conclusion that for all $i \in \mathbb{N}$, $|x_{n,i} - x_i| \leqslant \varepsilon$. Suppose that this is not true, and let k be the smallest integer i such that $|x_{n,i} - x_i| > \varepsilon$. Assume that $x_{n,k} > x_k + \varepsilon$ (the other case can be handled similarly). In particular, recalling (2.2), $|V_{n,k}| \geqslant \varepsilon n/2$, and so $[k] \subseteq A$. Since $x_1 \geqslant x_2 \geqslant \ldots$ and $x_{n,1} \geqslant x_{n,2} \geqslant \ldots$, we have for all $1 \leqslant i \leqslant k \leqslant i'$ that, neglecting $O(1/n)$ error terms, $x_{n,i} \geqslant x_{n,k} > x_k + \varepsilon \geqslant x_{i'} + \varepsilon$, so $|V_{n,i}| \geqslant |V_{i'}| + \varepsilon n/2$. Thus, for all positive integers $i \leqslant k$, we must have $\mu(i) < k$. In other words, $\mu([k]) \subseteq [k-1]$, which contradicts μ being a bijection. This completes the proof of the lemma. □

Remark 1. Lemma 3.8 in [2] proves that if $\boldsymbol{x}_n, \boldsymbol{x} \in \overline{\mathcal{P}}$ are such that $\boldsymbol{x}_n \to \boldsymbol{x}$ pointwise, then the corresponding graphons $Q_{\boldsymbol{x}_n}$ converge to $Q_{\boldsymbol{x}}$; that is, all the $p(F, Q_{\boldsymbol{x}_n})$ converge to $p(F, Q_{\boldsymbol{x}})$.

Lemma 2.5. *The space $\overline{\mathcal{P}}$ and distance δ_{edit} have the following properties.*

(i) *The space $(\overline{\mathcal{P}}, \delta_{\text{edit}})$ is a compact metric space.*
(ii) *The set of complete partite graphs $\mathcal{P}$ is dense in $(\overline{\mathcal{P}}, \delta_{\text{edit}})$.*
(iii) *The function λ can be extended to a continuous function on the whole of $\overline{\mathcal{P}}$, namely by defining*

$$\lambda(\boldsymbol{x}) := \lim_{n\to\infty} \lambda(G_{n,\boldsymbol{x}}), \quad \text{for } x \in \overline{\mathcal{P}}.$$

Proof. We begin with (i). From the definitions, it is clear that $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) = \delta_{\text{edit}}(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \overline{\mathcal{P}}$. By Lemma 2.1(iii), δ_{edit} defined on $\overline{\mathcal{P}}$ satisfies the triangle inequality. Finally, by definition, $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) = 0$ if and only if $\boldsymbol{x} = \boldsymbol{y}$. So, $(\overline{\mathcal{P}}, \delta_{\text{edit}})$ is a metric space. To show that it is compact, Proposition 2.3 implies that it suffices to show that $\overline{\mathcal{P}}$ is compact under the topology of pointwise convergence. For this, let $(\boldsymbol{x}_n)_n$ be an infinite sequence of elements of $\overline{\mathcal{P}}$. Then we can define its accumulation point $\boldsymbol{y}$ iteratively as follows. Initially, let $i = 0$. By passing to a subsequence of $(\boldsymbol{x}_n)_n$, we may assume that $(x_{n,i+1})_n$ converges to some $y_{i+1} \in \mathbb{R}$. If $y_{i+1} = 0$, then stop and output $\boldsymbol{y} := (y_1, \ldots, y_i, 0, 0, \ldots)$. Otherwise, increase i by one and continue. If the iteration does not terminate, output $\boldsymbol{y} := (y_1, y_2, \ldots)$. One can easily see that $\boldsymbol{y}$ is indeed an accumulation point of $(\boldsymbol{x}_{\underline{n}})_n$, completing the proof of (i). Alternatively, the compactness of $\overline{\mathcal{P}}$ follows from observing that $\overline{\mathcal{P}}$ is a closed subset of the compact space $[0,1]^{\mathbb{N}}$.

Part (ii) immediately follows since for every $\boldsymbol{x} \in \overline{\mathcal{P}}$, the sequence $(G_{n,\boldsymbol{x}})_n$ of complete partite graphs converges in edit distance to $\boldsymbol{x}$. Indeed, for each $n \in \mathbb{N}$, we have that $\boldsymbol{x}_{G_{n,\boldsymbol{x}}} \in \mathcal{P}$, and the definitions imply that $\boldsymbol{x}_{G_{n,\boldsymbol{x}}}$ converges pointwise to $\boldsymbol{x}$. By Proposition 2.3, it also converges in edit distance.

It remains to prove (iii). Recall that we fixed a function $\gamma : \mathcal{G}_k \to \mathbb{R}$ and for all $n \in \mathbb{N}$ and $G \in \mathcal{G}_n$, we defined $\lambda(G)$ as in (1.1). Let $\gamma_{\max} := \max\{|\gamma(F)| : F \in \mathcal{G}_k\}$ (which exists since the domain of γ is finite for fixed k). Let $n \in \mathbb{N}$, let $G \in \mathcal{G}_n$ and let xy be a pair in $V(G)$. Then

$$|\lambda(G) - \lambda(G \oplus xy)| \leqslant \binom{n}{k}^{-1} \sum_{\substack{X \in \binom{V(G)}{k}: \\ x,y \in X}} [\gamma(G[X]) - \gamma((G \oplus xy)[X])] \leqslant \frac{\binom{n-2}{k-2} \cdot 2\gamma_{\max}}{\binom{n}{k}} = \frac{2\binom{k}{2}\gamma_{\max}}{\binom{n}{2}}.$$

Therefore, using the triangle inequality, we have for any $G, H \in \mathcal{G}_n$ that

$$|\lambda(G) - \lambda(H)| \leqslant 2\binom{k}{2}\gamma_{\max} \cdot \hat{\delta}_1(G,H) + O(1/n) \leqslant 6\binom{k}{2}\gamma_{\max} \cdot \delta_{\mathrm{edit}}(G,H) + O(1/n), \tag{2.4}$$

where the final inequality follows from Lemma 2.1(i). Thus

$$|\lambda(G_{n,\boldsymbol{x}}) - \lambda(G_{n,\boldsymbol{y}})| \leqslant 6\binom{k}{2}\gamma_{\max} \cdot \delta_{\mathrm{edit}}(G_{n,\boldsymbol{x}}, G_{n,\boldsymbol{y}}) + O(1/n),$$

and by (i) we have that the function $\lambda : \overline{\mathcal{P}} \to \mathbb{R}$ given by $\lambda(\boldsymbol{x}) := \lim_{n\to\infty} \lambda(G_{n,\boldsymbol{x}})$ is well defined for all $\boldsymbol{x} \in \overline{\mathcal{P}}$ and is continuous with respect to δ_{edit}. ☐

Note that the extension in Part (iii) of Lemma 2.5 is unique since $\mathcal{P}$ is dense in $\overline{\mathcal{P}}$.

The lemma implies that $\lambda_{\max} := \lim_{n\to\infty} \lambda(n)$ defined in the introduction can equivalently be defined as $\lambda_{\max} := \max\{\lambda(\boldsymbol{x}) : \boldsymbol{x} \in \overline{\mathcal{P}}\}$. Moreover, for every $\boldsymbol{x} = (x_1, x_2, \ldots) \in \overline{\mathcal{P}}$, we have that $\lambda(\boldsymbol{x})$ has the following analytic formula. Let $\omega_1, \ldots, \omega_k$ be independent samples from $\Omega_{\boldsymbol{x}}$ which is the probability space on $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$ where the probability of i is x_i. Let the *random sample* $\mathbb{G}(\boldsymbol{x}, k)$ be equal to

$$G_{\boldsymbol{x}}^{\omega_1,\ldots,\omega_k} := \left([k], \binom{[k]}{2} \setminus \left(\bigcup_{i\in\mathbb{N}} \binom{\{j : \omega_j = i\}}{2}\right)\right),$$

which is the complete graph on $[k]$ except we do not connect two distinct indices $j, h \in [k]$ if $\omega_j = \omega_h \neq 0$. One can show using the Chernoff bound and the Borel–Cantelli lemma that $(\mathbb{G}(\boldsymbol{x}, n))_n$ converges to $\boldsymbol{x}$ in $\overline{\mathcal{P}}$ with probability 1 (see, e.g. the more general Proposition 11.32 in [26]). Clearly, we have that

$$\lambda(\boldsymbol{x}) = \mathbb{E}(\gamma(\mathbb{G}(\boldsymbol{x}, k))).$$

We let OPT consist of all maximisers $\boldsymbol{x} \in \overline{\mathcal{P}}$, that is,

$$\mathrm{OPT} = \mathrm{OPT}(\lambda) := \{\boldsymbol{x} \in \overline{\mathcal{P}} : \lambda(\boldsymbol{x}) \geqslant \lambda(\boldsymbol{y}) \text{ for all } \boldsymbol{y} \in \overline{\mathcal{P}}\} = \{\boldsymbol{x} \in \overline{\mathcal{P}} : \lambda(\boldsymbol{x}) = \lambda_{\max}\} \neq \emptyset.$$

The non-emptiness assertion follows from Lemma 2.5(i) and (iii). Let us see why the forward inclusion of the third equality is true. Take any $\boldsymbol{x} \in \overline{\mathcal{P}}$ with $\lambda(\boldsymbol{x}) \geqslant \lambda(\boldsymbol{y})$ for all $\boldsymbol{y} \in \overline{\mathcal{P}}$. For each $n \in \mathbb{N}$, since λ is symmetrisable, there is a complete partite graph F_n on n vertices such that $\lambda(F_n) = \lambda(n)$. Let $\boldsymbol{y}_n := \boldsymbol{x}_{F_n}$. For any $\boldsymbol{x} \in \overline{\mathcal{P}}$, we have $\lambda(G_{n,\boldsymbol{x}}) \leqslant \lambda(F_n) = \lambda(n)$. By passing to a subsequence, we may assume that $\boldsymbol{y}_{n_i}$ converges to some $\boldsymbol{y} \in \overline{\mathcal{P}}$. Then $\lambda(\boldsymbol{x}) \leqslant \lambda(\boldsymbol{y}) = \lim_{i\to\infty} \lambda(n_i) = \lambda_{\max}$. Thus, we must have $\lambda(\boldsymbol{x}) = \lambda_{\max}$, as desired.

This definition of OPT is equivalent to the one in the introduction. Indeed, let $\boldsymbol{a} = (a_1, a_2, \ldots) \in \overline{\mathcal{P}}$ be such that there exists a sequence $(H_n)_n$ of complete partite graphs such that, as $n \to \infty$, we have $v(H_n) \to \infty$, $\lambda(H_n) \to \lambda_{\max}$ and for every $i \geqslant 1$, the number of vertices in the ith largest part of H_n is $(a_i + o(1))v(H_n)$. Then $\boldsymbol{x}_{H_n} \to \boldsymbol{a}$ and $\lambda(\boldsymbol{a}) = \lim_{n\to\infty} \lambda(H_n) = \lambda_{\max}$, as required. On the other hand, let $\boldsymbol{x} \in \overline{\mathcal{P}}$ be such that $\lambda(\boldsymbol{x}) = \lambda_{\max}$. Then $(G_{n,\boldsymbol{x}})_n$ is the required sequence of graphs.

2.3 | Polynomials

We will be interested in various functions on $\overline{\mathcal{P}}$, in particular the extension of λ from the family of complete partite graphs to $\overline{\mathcal{P}}$. For these, we introduce a notion of polynomial on $\overline{\mathcal{P}}$ which will help us prove that functions related to λ are continuous.

Let $\Sigma(d) := \{(d_1, \ldots, d_t) \in \mathbb{N}^t : t \in \mathbb{N}_0 \text{ and } d_1 + \cdots + d_t = d\}$ be the set of ordered tuples of positive integers summing to d. Let $S_\emptyset(\boldsymbol{x}) := 1$ and for $t \in \mathbb{N}$ and $\boldsymbol{d} := (d_1, \ldots, d_t) \in \Sigma(d)$, define an *elementary symmetric polynomial* $S_{\boldsymbol{d}} : \{\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}} : \|\boldsymbol{x}\|_1 < \infty\} \to \mathbb{R}$ by

$$S_{\boldsymbol{d}}(\boldsymbol{x}) = S_{d_1,\ldots,d_t}(\boldsymbol{x}) := \sum_{\substack{\text{distinct}\\ i_1,\ldots,i_t \in \mathbb{N}}} \prod_{j=1}^{t} x_{i_j}^{d_j}. \tag{2.5}$$

Since $\sum_{\substack{\text{distinct}\\ i_1,\ldots,i_t \in \mathbb{N}}} \prod_{j=1}^{t} |x_{i_j}|^{d_j} \leqslant (\sum_{i\geqslant 1} |x_i|)^d < \infty$, each $S_{\boldsymbol{d}}(\boldsymbol{x})$ converges absolutely.

We say that a function $p : \overline{\mathcal{P}} \to \mathbb{R}$ is a *$\overline{\mathcal{P}}$-polynomial* if it can be written as a finite polynomial of $S_{\boldsymbol{d}}^I(\boldsymbol{x}) := S_{\boldsymbol{d}}(\boldsymbol{x}^I)$ for $I \subseteq \mathbb{N}$, where $\boldsymbol{x}^I \in \overline{\mathcal{P}}$ is obtained from $\boldsymbol{x}$ by removing every x_i with $i \in I$ and moving back remaining entries to fill in the 'gaps'. (Thus, $S_{\boldsymbol{d}}^I(\boldsymbol{x})$ is defined by the version of (2.5) where the sum is restricted to indices not in I.) So, for example, $x_0 = S_\emptyset(\boldsymbol{x}) - S_1(\boldsymbol{x})$, $x_i = S_1(\boldsymbol{x}) - S_1^{\{i\}}(\boldsymbol{x})$ for $i \in \mathbb{N}$ and $x_1 + x_3 + x_5 + x_7 + \cdots = S_1^{2\cdot\mathbb{N}}(\boldsymbol{x})$ are $\overline{\mathcal{P}}$-polynomials, while $x_1 + 2x_2 + 3x_3 + \cdots$ is not. Given any $\overline{\mathcal{P}}$-polynomial p, there is a finite partition $\mathbb{N} = I_1 \cup \ldots \cup I_s$ such that $p(x_1, x_2, \ldots) = p(y_1, y_2, \ldots)$ where $\boldsymbol{y}$ is any element of $\overline{\mathcal{P}}$ obtained from $\boldsymbol{x}$ by permuting indices within each part I_i. Indeed, one can obtain $I_1, \ldots, I_s$ by grouping together indices that belong to exactly the same sets I in the definition of p.

Take any $m \in \mathbb{N}$ and $\boldsymbol{d} = (d_1, \ldots, d_t) \in \mathbb{N}^t$. Consider $\frac{1}{h}(S_{\boldsymbol{d}}(\boldsymbol{x}') - S_{\boldsymbol{d}}(\boldsymbol{x}))$ where, for all $i \geqslant 1$, we have $x_i' = x_i$, except $x_m' = x_m + h$, and let $h \to 0$. Apply the binomial expansion to each $(x_m + h)^{d_j}$. As all series converge absolutely, we can change the order of summation and collect the same powers of h. We obtain

$$\frac{S_{\boldsymbol{d}}(\boldsymbol{x}') - S_{\boldsymbol{d}}(\boldsymbol{x})}{h} = \sum_{\substack{\text{distinct}\\ i_1,\ldots,i_t \in \mathbb{N}}} \frac{\partial}{\partial x_m}\left(\prod_{j=1}^{t} x_{i_j}^{d_j}\right) + \delta,$$

where δ is an error term satisfying $|\delta| \leqslant h \cdot 2^d$. So, we can define partial derivatives $\frac{\partial p}{\partial x_i}$ for $i = 1, 2, \ldots$ via term-by-term differentiation. Also, if $p = s(S_{\boldsymbol{d}} : \boldsymbol{d} \in \mathbb{N}^{\leqslant k})$ where s is a finite polynomial, then define $\frac{\partial p}{\partial x_0} := -\frac{\partial s}{\partial S_1}$. Thus, we can define partial derivatives of any $\overline{\mathcal{P}}$-polynomial, and each such derivative is itself a $\overline{\mathcal{P}}$-polynomial. For a complete partite graph G on n vertices with parts $V_1, \ldots, V_m$ of size at least 2 and clique part V_0, define for $I \subseteq \mathbb{N}$

$$S^I_{\boldsymbol{d}}(G) := (d!)^{-1}\binom{n}{d}^{-1} \sum_{\substack{\text{distinct} \\ i_1,\ldots,i_t \in [m]\setminus I}} \prod_{1 \leqslant j \leqslant t} d_j! \binom{|V_{i_j}|}{d_j} = \sum_{\substack{\text{distinct} \\ i_1,\ldots,i_t \in [m]\setminus I}} \prod_{1 \leqslant j \leqslant t} \left(\frac{|V_{i_j}| + O(1)}{n}\right)^{d_j}, \tag{2.6}$$

and let $S_{\boldsymbol{d}}(G) := S^{\emptyset}_{\boldsymbol{d}}(G)$. So $S^I_{\boldsymbol{d}}(G)$ is equal to $S_{\boldsymbol{d}}(G^I)$ up to a scaling factor, where $G^I := G - \bigcup_{i \in [m] \cap I} V_i$.

Lemma 2.6. *Let d be an integer and let* $\boldsymbol{d} = (d_1, \ldots, d_t) \in \Sigma(d)$. *Then*

(i) $S_{\boldsymbol{d}}$ *is uniformly continuous on* $(\overline{\mathcal{P}}, \delta_{\mathrm{edit}})$.
(ii) *Each* $\overline{\mathcal{P}}$*-polynomial is uniformly continuous on* $(\overline{\mathcal{P}}, \delta_{\mathrm{edit}})$.
(iii) *For all* $\boldsymbol{x} \in \overline{\mathcal{P}}$, *we have* $S_{\boldsymbol{d}}(\boldsymbol{x}) = \lim_{n\to\infty} S_{\boldsymbol{d}}(G_{n,\boldsymbol{x}})$.

Proof. We start with (i). By Proposition 2.3, convergence in edit distance and pointwise convergence induce the same topology on $\overline{\mathcal{P}}$. By Lemma 2.5(i), $\overline{\mathcal{P}}$ is compact. Therefore, it suffices to show that each $S_{\boldsymbol{d}}$ is continuous under pointwise convergence, which is, for example, given by the metric $d(\boldsymbol{x}, \boldsymbol{y}) := \sum_{i \geqslant 1} 2^{-i}|x_i - y_i|$. For this, let $\varepsilon > 0$ and let $\delta = 2^{-8d/\varepsilon}$. Let $\boldsymbol{x}, \boldsymbol{y} \in \overline{\mathcal{P}}$ satisfy $d(\boldsymbol{x}, \boldsymbol{y}) \leqslant \delta$. Choose $M = \lceil \log_2(\delta^{-1/2}) \rceil$ (so $1/M \leqslant \varepsilon/(4d)$) and let $\boldsymbol{x}' = (x_1, \ldots, x_M, 0, \ldots)$ and $\boldsymbol{y}' = (y_1, \ldots, y_M, 0, \ldots)$. Then $d(\boldsymbol{x}, \boldsymbol{x}') = \sum_{i > M} 2^{-i} x_i \leqslant 2^{-M} \leqslant \sqrt{\delta}$. So, $d(\boldsymbol{x}', \boldsymbol{y}') \leqslant 3\sqrt{\delta}$. Moreover,

$$S_{\boldsymbol{d}}(\boldsymbol{x}) - S_{\boldsymbol{d}}(\boldsymbol{x}') = \sum_{1 \leqslant s \leqslant t} \sum_{i > M} x_i^{d_s} S_{\boldsymbol{d}^{(s)}}(\boldsymbol{x}^{(i)}) \leqslant t x_{M+1} \leqslant d/M \leqslant \varepsilon/4,$$

where $\boldsymbol{d}^{(s)} = (d_1, \ldots, d_{s-1}, d_{s+1}, \ldots, d_t)$ and $\boldsymbol{x}^{(i)} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots)$. Similarly, $S_{\boldsymbol{d}}(\boldsymbol{y}) - S_{\boldsymbol{d}}(\boldsymbol{y}') \leqslant \varepsilon/4$. Now, $S_{\boldsymbol{d}}(\boldsymbol{x}')$ is a polynomial in at most M variables. For each $1 \leqslant i \leqslant M+1$, let $\boldsymbol{z}_i := (x_1, \ldots, x_{i-1}, y_i, y_{i+1}, \ldots, y_M, 0, \ldots)$. Then

$$|S_{\boldsymbol{d}}(\boldsymbol{x}') - S_{\boldsymbol{d}}(\boldsymbol{y}')| = |S_{\boldsymbol{d}}(\boldsymbol{z}_1) - S_{\boldsymbol{d}}(\boldsymbol{z}_{M+1})| \leqslant \sum_{i=1}^{M} |S_{\boldsymbol{d}}(\boldsymbol{z}_{i+1}) - S_{\boldsymbol{d}}(\boldsymbol{z}_i)|.$$

Now

$$S_{\boldsymbol{d}}(\boldsymbol{z}_{i+1}) - S_{\boldsymbol{d}}(\boldsymbol{z}_i) = \sum_{1 \leqslant s \leqslant t} (x_i^{d_s} - y_i^{d_s}) S_{\boldsymbol{d}^{(s)}}(\boldsymbol{z}_i^{(i)}) = p_i(x_i) - p_i(y_i),$$

where we view p_i as a polynomial in one variable. Thus, p_i is Lipschitz with constant at most $\max_{z \in [0,1]} |p_i'(z)| \leqslant d_1 + \cdots + d_t = d$. So, $|p_i(x_i) - p_i(y_i)| \leqslant d|x_i - y_i|$. Thus,

$$|S_{\boldsymbol{d}}(\boldsymbol{x}) - S_{\boldsymbol{d}}(\boldsymbol{y})| \leqslant \varepsilon/2 + d \sum_{i=1}^{M} |x_i - y_i| \leqslant \varepsilon/2 + d2^M d(\boldsymbol{x}, \boldsymbol{y}) \leqslant \varepsilon/2 + \sqrt{\delta} < 2\varepsilon/3,$$

completing the proof of (i).

Now (ii) follows immediately since every $S_{\boldsymbol{d}}$ is bounded, and sums and products of bounded uniformly continuous functions are uniformly continuous. For (iii), fix $\boldsymbol{x} \in \overline{\mathcal{P}}$. In $G_{n,\boldsymbol{x}}$, writing V_i^n for the ith part, we have each $(|V_i^n| + O(1))/n \to x_i$ as $n \to \infty$, so as $S_{\boldsymbol{d}}$ is continuous, we have $S_{\boldsymbol{d}}(G_{n,\boldsymbol{x}}) \to S_{\boldsymbol{d}}(\boldsymbol{x})$. ☐

3 | STRICTNESS AND A RESTATEMENT OF THE MAIN RESULT

In this section, we will finally define what it means for λ to be 'strict'. Very roughly speaking, it means that when an elementary change is made to a complete partite graph on which λ is maximised, the decrease in λ is as much as it possibly could be. An 'elementary change' is either 'flipping a pair' (changing a non-edge to an edge or vice versa); or adding a vertex which is either adjacent to every vertex in a part, or to no vertex in a part. It seems that it is more convenient to state this property in terms of limits rather than graphs (which is why the definition is deferred until now). We will first make the relevant definition and then discuss it further.

3.1 | Definitions and notation

Definition 4 ($\nabla^{\bullet\bullet}_{i_1 i_2}\lambda$ and $\nabla^{\bullet}_{b,\alpha}\lambda$). Given an n-vertex graph $G = (V, E)$ and a pair x, y of vertices of G, define

$$\nabla^{\bullet\bullet}_{xy}\lambda(G) := \frac{1}{\binom{n-2}{k-2}}(\Lambda(G) - \Lambda(G \oplus xy)).$$

Given $\boldsymbol{x} \in \overline{\mathcal{P}}$ and $i_1, i_2 \in \operatorname{supp}^*(\boldsymbol{x})$, define

$$\nabla^{\bullet\bullet}_{i_1 i_2}\lambda(\boldsymbol{x}) := \lim_{n\to\infty} \nabla^{\bullet\bullet}_{v_1 v_2}\lambda(G_{n,\boldsymbol{x}}),$$

where v_1, v_2 are distinct vertices of the vertex classes V_{i_1} and V_{i_2} of $G_{n,\boldsymbol{x}}$, respectively.

For all $i \in \mathbb{N}_0$, we define e_i to be the function $e_i : \mathbb{N} \to \{0, 1\}$ with $e_i(j) = 0$ if and only if $j = i$ (so $e_0 \equiv 1$). Let $b : \mathbb{N} \to \{0, 1\}$ and $\alpha \in [0, 1]$. We write $G +_{b,\alpha} u$ for the graph obtained from G with $\overline{\mathcal{P}}$-structure $V_0, V_1, \ldots, V_m$ by adding a new vertex u and, for $i \geqslant 1$, adding every edge between u and V_i if $b(i) = 1$, and no edges otherwise; and adding $\lfloor \alpha |V_0| \rfloor$ edges between u and V_0. Define

$$\nabla^{\bullet}_{b,\alpha}\lambda(G) := \frac{1}{\binom{n}{k-1}}(\Lambda(G +_{e_1,1} u, u) - \Lambda(G +_{b,\alpha} u, u)) = \frac{1}{\binom{n}{k-1}}(\Lambda(G +_{e_1,1} u) - \Lambda(G +_{b,\alpha} u)),$$

where $u \notin V(G)$, and let

$$\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) := \lim_{n\to\infty} \nabla^{\bullet}_{b,\alpha}\lambda(G_{n,\boldsymbol{x}}) \quad \text{and} \quad \lambda(\boldsymbol{x}, (b, \alpha)) := \lim_{n\to\infty} \lambda(G_{n,\boldsymbol{x}} +_{b,\alpha} u, u).$$

By convention, take $\alpha = 1$ if $x_0 = 0$ (when $V_0 = \varnothing$).

Given $k_0 \in \mathbb{N}_0$ and a tuple $\boldsymbol{k} = (k_1, \ldots, k_t)$ of positive integers, define the graph $G^{k_0}_{\boldsymbol{k}}$ as follows. Let $G^{k_0}_{\boldsymbol{k}}$ be the complete partite graph with t parts $U_1, \ldots, U_t$ of size $k_1, \ldots, k_t$, respectively, together with an additional k_0 singletons $x_1, \ldots, x_{k_0}$, whose union is denoted by U_0.

Both limits in Definition 4 exist and each $\lambda, \nabla^{\bullet\bullet}_{i_1 i_2}\lambda$, $\nabla^{\bullet}_{b,\alpha}\lambda$, $\lambda(\cdot,(b,\alpha))$ is a $\overline{\mathcal{P}}$-polynomial. Indeed, since each $G_{n,\boldsymbol{x}}$ is a complete partite graph (with parts $V^n_0, V^n_1, \ldots$), the quantities $\lambda(G_{n,\boldsymbol{x}})$, $\nabla^{\bullet\bullet}_{v_1 v_2}\lambda(G_{n,\boldsymbol{x}})$ and $\nabla^{\bullet}_{b,\alpha}\lambda(G_{n,\boldsymbol{x}})$ are finite polynomials in variables $|V^n_0|$ and $S^I_{\boldsymbol{d}}(G_{n,\boldsymbol{x}})$ for $\boldsymbol{d} \in \Sigma(d)$ with $d \leqslant k$ and $I \subseteq \mathbb{N}$. Indeed, for λ, we need only $I = \emptyset$; for $\nabla^{\bullet\bullet}_{v_1 v_2}\lambda$, we could take only $I = \emptyset, \{v_1\}, \{v_2\}, \{v_1, v_2\}$ and their complements, and for $\nabla^{\bullet}_{b,\alpha}\lambda$, $I = \emptyset$, $\mathrm{supp}(b)$ and their complements. Thus, by Lemma 2.6, $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda$ and $\nabla^{\bullet}_{b,\alpha}\lambda$ are $\overline{\mathcal{P}}$-polynomials.

In fact, one can explicitly write these polynomials. For positive integers $b_1 \geqslant \ldots \geqslant b_r$, let $\mathrm{sym}(b_1, \ldots, b_r)$ be the number of permutations of $[r]$ that keep the sequence $(b_1, \ldots, b_r)$ unchanged. In other words, if we take $i_0 := 1 < i_1 < \ldots < i_q < r + 1 =: i_{q+1}$ such that $b_i = b_{i'}$ if and only if there is $j \in [q]$ such that $i_{j-1} \leqslant i, i' < i_j$, then $\mathrm{sym}(b_1, \ldots, b_r) = (i_1 - i_0)! \ldots (i_{q+1} - i_q)!$. Also, write $\binom{t}{t_1,\ldots,t_s} := t!(t_1! \ldots t_s!)^{-1}$ when $\sum_{i=1}^{s} t_i = t$. Consider $p(K_{a_1,\ldots,a_\ell}, \cdot)$, which is one instance of λ, where $a_1, \ldots, a_\ell$ are in non-increasing order, and let $t \in [\ell]$ be the largest integer such that $a_1, \ldots, a_t \geqslant 2$. Then we have the following analytic formula:

$$p(K_{a_1,\ldots,a_\ell}, \boldsymbol{x}) = p(K_{a_1,\ldots,a_\ell}, Q_{\boldsymbol{x}}) = \frac{\binom{a_1+\cdots+a_\ell}{a_1,\ldots,a_\ell}}{\mathrm{sym}(a_1, \ldots, a_\ell)} \sum_{0 \leqslant s \leqslant \ell - t} \binom{\ell - t}{s} x_0^s \cdot S_{a_1,\ldots,a_{\ell-s}}(\boldsymbol{x}). \tag{3.1}$$

Using (3.1), one can write $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda$ and $\nabla^{\bullet}_{b,\alpha}\lambda$ as $\overline{\mathcal{P}}$–polynomials.

The next proposition gives that for all $\boldsymbol{x} \in \mathrm{OPT}$, $\nabla^{\bullet}_{e_i,1}\lambda(\boldsymbol{x}) = 0$ for all $i \in \mathrm{supp}^*(\boldsymbol{x})$, which corresponds to saying that every vertex in the realisation of an optimal $\boldsymbol{x}$ contributes optimally to λ. Thus,

$$\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) = \lambda(\boldsymbol{x}) - \lambda(\boldsymbol{x}, (b,\alpha)) \quad \text{for all } (b,\alpha) \text{ and } \boldsymbol{x} \in \mathrm{OPT}.$$

Proposition 3.1. *Define k and λ as in (1.1). The following hold for all $\boldsymbol{x} \in \overline{\mathcal{P}}$.*

(i) *For all $i \in \mathrm{supp}^*(\boldsymbol{x})$, we have $\frac{1}{k} \cdot \frac{\partial \lambda(\boldsymbol{x})}{\partial x_i} = \lambda(\boldsymbol{x}, (e_i, 1))$.*

(ii) *If in addition $\boldsymbol{x} \in \mathrm{OPT}$, then for all $i \in \mathrm{supp}^*(\boldsymbol{x})$, we have $\frac{1}{k} \cdot \frac{\partial \lambda(\boldsymbol{x})}{\partial x_i} = \lambda(\boldsymbol{x})$.*

(iii) *The following pairs differ by $O(1/n)$ as $n \to \infty$: $\{\lambda_{\max}, \lambda(n)\}$, $\{\lambda(\boldsymbol{x}), \lambda(G_{n,\boldsymbol{x}})\}$ and $\{\lambda(G_{n,\boldsymbol{x}}, u), \lambda_{\max}\}$, the last pair for $\boldsymbol{x} \in \mathrm{OPT}$ and $u \in V(G_{n,\boldsymbol{x}})$.*

Proof. The equality in (i) can be checked directly.

For (ii), the theory of Lagrange multipliers implies that, for all $i, j \in \mathrm{supp}^*(\boldsymbol{x})$, we have $\frac{\partial \lambda(\boldsymbol{x})}{\partial x_i} = \frac{\partial \lambda(\boldsymbol{x})}{\partial x_j}$. Indeed, if we fix the rest of $\boldsymbol{x}$ apart from x_i, x_j, fix $s = x_i + x_j$ and vary x_i, x_j, then we can view λ as a polynomial in x_i, x_j (of degree at most k). Introducing a new variable μ, the Lagrangian is $\mathcal{L}(x_i, x_j, \mu) = \lambda(\boldsymbol{x}) - \mu(x_i + x_j - s)$. The stationary points of $\mathcal{L}$ occur when $(\frac{\partial \mathcal{L}}{\partial x_i}, \frac{\partial \mathcal{L}}{\partial x_j}, \frac{\partial \mathcal{L}}{\partial \mu}) = (0, 0, 0)$, that is, when $\frac{\partial \lambda(\boldsymbol{x})}{\partial x_i} - \mu = \frac{\partial \lambda(\boldsymbol{x})}{\partial x_j} - \mu$, as required. Since λ is a $\overline{\mathcal{P}}$-polynomial with each monomial having total degree k, we have for all $i \in \mathrm{supp}^*(\boldsymbol{x})$ that

$$\frac{\partial \lambda(\boldsymbol{x})}{\partial x_i} = \sum_{j \geqslant 0} x_j \frac{\partial \lambda(\boldsymbol{x})}{\partial x_j} = k \cdot \lambda(\boldsymbol{x}),$$

giving the required.

Let us turn to Part (iii). The inequality $|\lambda_{\max} - \lambda(n)| = O(1/n)$ follows from a standard blow-up trick, see, for example, [32, Lemma 2.2]. The claim for the second pair follows from the fact that each named function on $\overline{\mathcal{P}}$ is a $\overline{\mathcal{P}}$-polynomial, a finite polynomial of $S_{\boldsymbol{d}}(G^I_{n,\boldsymbol{x}})$ terms, so the error bound comes from (2.6) when applied to $G_{n,\boldsymbol{x}}$. For the last claim of Part (iii), a version of (2.6) implies that $|\lambda(G_{n,\boldsymbol{x}}, u) - \frac{1}{k} \cdot \frac{\partial \lambda(\boldsymbol{x})}{\partial x_i}| = O(1/n)$ where u is in the ith part of $G_{n,\boldsymbol{x}}$. Then Part (ii) gives the required. ☐

Corollary 3.2. *For every* $\varepsilon > 0$, *there exists* $\delta > 0$ *such for all* $\boldsymbol{x}, \boldsymbol{y} \in \overline{\mathcal{P}}$ *with* $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \delta$, *we have*

$$|\lambda(\boldsymbol{x}) - \lambda(\boldsymbol{y})|, |\nabla^{\bullet}_{b,\alpha}(\boldsymbol{x}) - \nabla^{\bullet}_{b,\alpha}(\boldsymbol{y})|, |\lambda(\boldsymbol{x}, (b,\alpha)) - \lambda(\boldsymbol{y}, (b,\alpha))| \leqslant \varepsilon$$

for all $b : \mathbb{N} \to \{0,1\}$ *and* $0 \leqslant \alpha \leqslant 1$, *and*

$$|\nabla^{\bullet\bullet}_{i_1 i_2} \lambda(\boldsymbol{x}) - \nabla^{\bullet\bullet}_{i_1 i_2} \lambda(\boldsymbol{y})| \leqslant \varepsilon$$

for all $i_1, i_2 \in \operatorname{supp}^*(\boldsymbol{x}) \cap \operatorname{supp}^*(\boldsymbol{y})$.

Proof. We have seen that each function $\lambda, \nabla^{\bullet\bullet}_{i_1 i_2}\lambda, \nabla^{\bullet}_{b,\alpha}\lambda, \lambda(\cdot, (b,\alpha))$ is a $\overline{\mathcal{P}}$-polynomial with degree at most k and with coefficients whose absolute values are bounded. Thus, Lemma 2.6 implies that the family of $\lambda, \nabla^{\bullet\bullet}_{i_1 i_2}\lambda, \nabla^{\bullet}_{b,\alpha}\lambda, \lambda(\cdot, (b,\alpha))$ over all i_1, i_2, b, α is uniformly equicontinuous, as required. ☐

The following crucial definition of the *strictness* property of a function λ requires that both $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda$ and $\nabla^{\bullet}_{b,\alpha}\lambda$ are bounded from below whenever (b,α) is not close to some $(e_i, 1)$. Roughly speaking, this means that λ is sensitive to small alterations in a graph.

Definition 5 (Strictness). We say that λ is *strict (with parameter* c*)* if there is $c = c(\lambda) > 0$ such that for each $\boldsymbol{x} \in \mathrm{OPT}$, we have

(Str1) $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda(\boldsymbol{x}) \geqslant c$ for all $i_1, i_2 \in \operatorname{supp}^*(\boldsymbol{x})$,
(Str2) $\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) \geqslant c((1-\alpha)x_0 + \min_{i \in \operatorname{supp}^*(\boldsymbol{x})} w_i)$, where

$$w_i := \mathbb{1}_{i>0} b_i x_i + \sum_{j \in \operatorname{supp}^*(\boldsymbol{x}) \setminus \{0,i\}} (1-b_j) x_j.$$

In the next two subsections, we will motivate these definitions, which appear somewhat complicated at first sight.

3.1.1 | $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda$: Flipping a pair of vertices

Take a complete partite graph G of large order n such that $\lambda(G) \approx \lambda_{\max}$ and let $G' = G \oplus xy$ be obtained by flipping the adjacency of an arbitrary pair $xy \in \binom{V}{2}$. Then the number of vertex subsets of size k which contain both x and y is $\binom{n-2}{k-2}$, so in the worst case, γ decreases by a constant for all such subsets, and thus λ decreases by $\Omega(\binom{n-2}{k-2}/\binom{n}{k}) = \Omega(1/n^2)$. Property (Str1) says that this worst-case behaviour is realised for every 'wrong' pair xy.

Observe that

$$\nabla^{\bullet\bullet}_{i_1 i_2}\lambda(\boldsymbol{x}) = \mathbb{E}_{\omega_3,\dots,\omega_k\sim\Omega_{\boldsymbol{x}}}[\gamma(G_{\boldsymbol{x}}^{i_1,i_2,\omega_3,\dots,\omega_k}) - \gamma(G_{\boldsymbol{x}}^{i_1,i_2,\omega_3,\dots,\omega_k} \oplus \{1,2\})],$$

that is, we look at the conditional expectation of the change in λ if we flip the pair $\{1,2\}$ in a random sample $\mathbb{G}(\boldsymbol{x},k)$ conditioned on $\omega_1 = i_1$ and $\omega_2 = i_2$.

3.1.2 | $\nabla^{\bullet}_{b,\alpha}\lambda$: Adding a new vertex

Again consider a complete partite graph G of large order n such that $\lambda(G) \approx \lambda_{\max}$ and obtain a graph G' from G by adding a new vertex u which, for each part of G, either connects to all or none of its vertices (here we are thinking of V_0, if it exists, as consisting of $|V_0|$ singleton parts). If the attachment of u mirrors an existing vertex, then its contribution to λ is approximately $\lambda_{\max}$ (and G' is the same as G in the limit). But, if not, as u lies in $\binom{n}{k-1}$ subsets, in the worst case, λ decreases by $\Omega(\binom{n}{k-1}/\binom{n+1}{k}) = \Omega(1/n)$. Property (Str2) says that this worst-case behaviour is realised for every u with 'wrong' attachment.

Suppose that $G_{n,\boldsymbol{x}}$ has $\overline{\mathcal{P}}$-structure $V_0, V_1, \dots, V_{m(n)}$. Then, for $0 \leqslant i \leqslant m(n)$, let W_i be the minimum number of edits needed to move the vertex u in $G_{n,\boldsymbol{x}} +_{b,\alpha} u$ into the ith part. So, each W_i being large corresponds to u being attached in an atypical manner, and some W_i small means that u behaves like an existing vertex. It is not hard to show that $\lim_{n\to\infty} W_i/n = w_i + (1-\alpha)x_0$, and, of course, if $b = e_i$ and $\alpha = 1$, then $w_i + (1-\alpha)x_0 = 0$ (since no edits are needed to move u to the ith part). So (Str2) requires that, whenever n is large, the contribution to λ lost by a vertex u in $G_{n,\boldsymbol{x}} +_{b,\alpha} u$ is a significant fraction of the number of edits needed to fit u into $G_{n,\boldsymbol{x}}$.

Observe that (using Proposition 3.1 and the remark immediately before it)

$$\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) := \mathbb{E}_{\omega_1,\dots,\omega_k\sim\Omega_{\boldsymbol{x}}}[\gamma(G_{\boldsymbol{x}}^{\omega_1,\dots,\omega_k})] - \mathbb{E}_{\omega_1,\dots,\omega_{k-1}\sim\Omega_{\boldsymbol{x}}}[\gamma(G_{\boldsymbol{x}}^{\omega_1,\dots,\omega_{k-1}} +_{b,\alpha} u)],$$

where $G_{\boldsymbol{x}}^{\omega_1,\dots,\omega_{k-1}} +_{b,\alpha} u$ is the random graph obtained by adding u to $\mathbb{G}(\boldsymbol{x},k-1)$ with ui an edge when $\omega_i \neq 0$ if and only if $b(\omega_i) = 1$, and ui an edge when $\omega_i = 0$ with probability α.

3.2 | Main result

We are now ready to precisely state the 'limit version' of our main result.

Theorem 3.3. *Let k be a positive integer and let $\gamma : \mathcal{G}_k \to \mathbb{R}$. Define $\lambda : \mathcal{G} \to \mathbb{R}$ by setting $\lambda(G) := \binom{n}{k}^{-1}\sum_{X\in\binom{V}{k}}\gamma(G[X])$ for all $G \in \mathcal{G}_n$ and $n \in \mathbb{N}$, and let $\lambda(n) := \max_{G\in\mathcal{G}_n}\lambda(G)$. Suppose that λ is symmetrisable and $|\mathrm{OPT}(\lambda)| < \infty$. Then λ has perfect stability if it is strict.*

The following corollary states that strict symmetrisable functions exhibit classical stability, in the sense that any sufficiently large graph which is sufficiently close to being optimal can be edited by changing an arbitrarily small fraction of its adjacencies to obtain a complete partite graph with the correct part sizes.

Corollary 3.4. *Define k and λ as in (1.1) and suppose that they satisfy the assumptions in Theorem 3.3, and suppose further that λ is strict. Then for all $\varepsilon > 0$, there exist $\delta, n_0 > 0$ such that for every graph G of order $n \geqslant n_0$ for which $\lambda(G) \geqslant \lambda_{\max} - \delta$, there is $\boldsymbol{x} \in \mathrm{OPT}(\lambda)$ for which $\delta_{\mathrm{edit}}(G, \boldsymbol{x}) \leqslant \varepsilon$.*

Proof. Let $c = c(\lambda) > 0$ be such that λ is strict with parameter c. Apply Theorem 3.3 to obtain C such that λ is perfectly stable with constant C. Suppose that the statement does not hold. Then there is a sequence of counterexamples $(G_n)_n$ with $v_n := v(G_n) \to \infty$ such that $\lambda(G_n) \geqslant \lambda_{\max} - 1/n$ but $\delta_{\mathrm{edit}}(G_n, \boldsymbol{x}) > \varepsilon$ for all $\boldsymbol{x} \in \mathrm{OPT}$. By taking a subsequence if necessary, we may assume that each $v_n \geqslant n$. Let n be sufficiently large. By Theorem 3.3, there is some $H_n \in \mathcal{P}_{v_n}$ for which

$$\hat{\delta}_1(G_n, H_n)/C \leqslant \lambda(v_n) - \lambda(G_n) \leqslant \lambda(v_n) - \lambda_{\max} + O(1/v_n) \leqslant O(1/n),$$

where we used Proposition 3.1(iii). But then by (2.4),

$$\lambda(H_n) \geqslant \lambda(G_n) - 2\binom{k}{2}\gamma_{\max}\hat{\delta}_1(G_n, H_n) - O(1/v_n) \geqslant \lambda_{\max} - O(1/n)\left(1 + 2C\binom{k}{2}\right).$$

So, writing $\boldsymbol{x}_n := \boldsymbol{x}_{H_n}$, and taking a subsequence if necessary, we see that $\boldsymbol{x}_n \to \boldsymbol{x} \in \mathrm{OPT}$. But then, when n is sufficiently large, using Lemma 2.1,

$$\delta_{\mathrm{edit}}(G_n, \boldsymbol{x}) \leqslant \delta_{\mathrm{edit}}(G_n, H_n) + \delta_{\mathrm{edit}}(\boldsymbol{x}_n, \boldsymbol{x}) \leqslant \hat{\delta}_1(G_n, H_n) + \delta_{\mathrm{edit}}(\boldsymbol{x}_n, \boldsymbol{x}) < \varepsilon,$$

a contradiction. □

4 | FINITELY MANY MAXIMISERS

We will need the following result which states that if the limit problem has finitely many optimisers, then all non-zero entries in them are separated from 0 by some constant $\beta > 0$.

Lemma 4.1. *If $|\mathrm{OPT}| < \infty$, then there is $\beta > 0$ such that $\mathrm{OPT} \subseteq \overline{\mathcal{P}}_{\beta}$.*

The rest of the section is dedicated to proving Lemma 4.1. Our proof is an adaptation of the proof of Glebov, Grzesik, Klimošová and Král' [14] who, in particular, worked on the finite forcibility of graphons which are a countable union of cliques. Recall notions related to graphons in Section 2.1. A graphon Q is *finitely forcible* if there are finitely many graphs $F_1, \ldots, F_\ell$ such that for every graphon Q', if $p(F_i, Q) = p(F_i, Q')$ for all $i \in [\ell]$, then Q and Q' are weakly isomorphic.

Firstly, we need the following result which is Lemma 11 in [14] (except it is obtained by complementing all graphs and using our language of partite limits).

Lemma 4.2. *If $\mathrm{OPT} = \{\boldsymbol{x}\}$ consists of a single element $\boldsymbol{x}$, then there is ℓ_0 (in fact, we can take $\ell_0 = k$ where k is as in the definition of λ) such that, for any $\boldsymbol{y} \in \overline{\mathcal{P}}$ with $y_0 = x_0$, if $p(\overline{K_i}, \boldsymbol{x}) = p(\overline{K_i}, \boldsymbol{y})$ for every $2 \leqslant i \leqslant \ell_0$, then $\boldsymbol{y} = \boldsymbol{x}$.*

Proof. Our $\boldsymbol{x}$ corresponds to a graphon $Q_{\boldsymbol{x}}$. The fact that $\boldsymbol{x}$ is the unique element of OPT is equivalent to saying that the equations $p(\overline{P_3}, Q) = 0$ (the induced density of triples spanning exactly one edge) and $\lambda(Q) = \lambda_{\max}$ force Q to be $Q_{\boldsymbol{x}}$ up to weak isomorphism in the space of graphons.

In particular, $Q_{\boldsymbol{x}}$ is finitely forcible. The constraint $p(\overline{P_3}, Q) = 0$ forces $Q \in \mathcal{Q}$ (i.e. to be a complete partite graphon) and thus automatically forces $p(F, Q) = 0$ for every graph F which is not complete partite, so we can ignore all such induced densities.

Thus, the equation $\lambda(Q) = \lambda_{\max}$ can be viewed as involving only induced densities of complete partite graphs on at most k vertices. We claim that it can be equivalently rewritten as some polynomial in x_0 and induced densities of independent sets of size at most k. Then, supposing that the claim is true, if $Q_{\boldsymbol{y}} \in \mathcal{Q}$ has $y_0 = x_0$ and the same induced densities of $\overline{K_2}, \dots, \overline{K_k}$ as $Q_{\boldsymbol{x}}$, then $Q_{\boldsymbol{y}}$ and $Q_{\boldsymbol{x}}$ are weakly isomorphic and thus $\boldsymbol{y} = \boldsymbol{x}$.

It remains to prove the claim. For this, it suffices to prove that for any complete partite graph $F = K_{a_1,\dots,a_\ell}$ with vertex set $[k]$ and with ℓ parts, for all $\boldsymbol{x} \in \overline{\mathcal{P}}$, we have that $p(F, Q_{\boldsymbol{x}})$ is some polynomial of x_0 and $p(\overline{K_2}, Q_{\boldsymbol{x}}), \dots, p(\overline{K_k}, Q_{\boldsymbol{x}})$. The claim is clear for $\ell = 1$ so assume $2 \leqslant \ell \leqslant k$. Assume that $a_1, \dots, a_\ell$ are in non-increasing order, and let $t \in [\ell]$ be the largest integer such that $a_1, \dots, a_t \geqslant 2$. Recall the analytic formula (3.1) for $p(F, Q_{\boldsymbol{x}})$. We have

$$\begin{aligned} S_{a_1,\dots,a_\ell}(\boldsymbol{x}) = S_{a_1}(\boldsymbol{x}) S_{a_2,\dots,a_\ell}(\boldsymbol{x}) - S_{a_1+a_2,a_3,\dots,a_\ell}(\boldsymbol{x}) - S_{a_2,a_1+a_3,\dots,a_\ell}(\boldsymbol{x}) - \dots \\ \dots - S_{a_2,\dots,a_{\ell-1},a_1+a_\ell}(\boldsymbol{x}), \end{aligned} \tag{4.1}$$

and for every $a \geqslant 2$, we have $p(\overline{K_a}, Q_{\boldsymbol{x}}) = S_a(\boldsymbol{x})$. The claim now follows by induction on ℓ. Indeed, every $S_{a_1,\dots,a_{\ell-s}}(\boldsymbol{x})$ can be expressed as a polynomial of $S_a(\boldsymbol{x})$ for $2 \leqslant a \leqslant k$, by (4.1) and induction, as required. ☐

We need the following easy generalisation of Lemma 4.2.

Lemma 4.3. *If* OPT *is finite, then there is ℓ_0 such that, for every $\boldsymbol{x} \in$ OPT and every $\boldsymbol{y} \in \overline{\mathcal{P}}$ with $y_0 = x_0$, if $\boldsymbol{x}$ and $\boldsymbol{y}$ have the same induced density of $\overline{K_i}$ for every $1 \leqslant i \leqslant \ell_0$, then $\boldsymbol{y} = \boldsymbol{x}$.*

Proof. For every pair $\boldsymbol{z}, \boldsymbol{z}' \in$ OPT, there is some graph F such that $p(F, \boldsymbol{z}) \neq p(F, \boldsymbol{z}')$. Indeed, since $\boldsymbol{z} \neq \boldsymbol{z}'$, their graphons $Q_{\boldsymbol{z}}, Q_{\boldsymbol{z}'}$ are not weakly isomorphic and thus have a different induced density of some graph F. Of course, this F has to be complete partite (otherwise its induced density in both $\boldsymbol{z}$ and $\boldsymbol{z}'$ is zero). Let $F_1, \dots, F_m$ be all such graphs F where $m \leqslant \binom{|\mathrm{OPT}|}{2}$. Let $\ell_0 := k + 2\max_{i \in [m]} v(F_i)$. Now let $\boldsymbol{x}$ and $\boldsymbol{y}$ be as in the lemma.

Consider the new optimisation problem where we maximise

$$\lambda'(\boldsymbol{z}) := \lambda(\boldsymbol{z}) - \sum_{i=1}^{m} (p(F_i, \boldsymbol{z}) - p(F_i, \boldsymbol{x}))^2.$$

Again, as in the proof of Lemma 4.2, λ' can be written as a polynomial of x_0 and induced densities of anticliques on at most ℓ_0 vertices. Also, clearly, $\boldsymbol{x}$ is the unique element of OPT(λ'). Apply Lemma 4.2 to OPT$(\lambda') = \{\boldsymbol{x}\}$. ☐

Proof of Lemma 4.1. Let ℓ_0 be as in Lemma 4.3. It is enough to show that, for every $\boldsymbol{x} \in$ OPT, there are at most $m := \ell_0$ distinct non-zero values among $x_1, x_2, \dots$ (then since $|\mathrm{OPT}| < \infty$, the lemma trivially follows).

Suppose on the contrary that $x_{i_1}, \dots, x_{i_{m+1}}$ are all positive and distinct for some $1 \leqslant i_1 < \dots < i_{m+1}$. Without loss of generality, assume that these are the smallest such indices we could have chosen. Consider unknown variables $y_{i_1},\dots,y_{i_{m+1}}$ and set $y_i := x_i$ for every other $i \geqslant 1$. We get a contradiction to our choice of ℓ_0 if we show that there is a choice of $y_{i_1},\dots,y_{i_{m+1}} > 0$ such that

$$\sum_{j=1}^{m+1} y_{i_j}^d = \sum_{j=1}^{m+1} x_{i_j}^d, \quad \text{for every } d = 1,\dots,m, \tag{4.2}$$

but the reordering $\boldsymbol{y}'$ of $\boldsymbol{y}$ (so that $y_1' \geqslant y_2' \geqslant \dots$ and $y_0' = y_0$) is not equal to $\boldsymbol{x}$. (Indeed, then $\boldsymbol{y}' \in \overline{\mathcal{P}}$ by the case $d = 1$ of (4.2) and it satisfies $p(\overline{K_d},\boldsymbol{y}') = p(\overline{K_d},\boldsymbol{x})$ for every $d = 2,\dots,\ell_0$ by the corresponding case of (4.2).)

Consider the map $g : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$ which sends $(z_1,\dots,z_{m+1})$ to $(\sum_{j=1}^{m+1} z_j^d)_{d=1}^m$. The Jacobian of $g(\cdot, x_{i_{m+1}}) : \mathbb{R}^m \to \mathbb{R}^m$, which sends $\boldsymbol{z} \in \mathbb{R}^m$ to $g(\boldsymbol{z}, x_{i_{m+1}})$, has non-zero determinant at $\boldsymbol{z}_0 := (x_{i_1},\dots,x_{i_m})$. Indeed, the (s,t)-entry of the Jacobian at $(z_1,\dots,z_m)$ is sz_t^{s-1}, so if we divide its sth row by s, we obtain the Vandermonde matrix of $z_1,\dots,z_m$, so its determinant is $m! \prod_{1\leqslant s<t\leqslant m}(z_s - z_t)$ which is non-zero at $\boldsymbol{z} = \boldsymbol{z}_0$.

Thus, the Jacobian of $g(\cdot, x_{i_{m+1}})$ is invertible. By the Implicit Function Theorem, for every choice of $y_{i_{m+1}}$ sufficiently close to $x_{i_{m+1}}$, there is a continuous choice of $(y_{i_1},\dots,y_{i_m})$ close to $(x_{i_1},\dots,x_{i_m})$ satisfying (4.2). Choose such a $y_{i_{m+1}}$ not equal to any x_j and such that $y_{i_1},\dots,y_{i_m}$ are all positive. Then the reordering $\boldsymbol{y}'$ of the obtained sequence $\boldsymbol{y}$ is not equal to $\boldsymbol{x}$, giving the desired contradiction. ☐

5 | THE PROOF OF THEOREM 3.3

In the first part of the proof, we find a suitable ‘hypothetical counterexample’ H on h vertices (Claim 5.2). This means that H is very close to being optimal ($\lambda(H)$ is almost as large as $\lambda(h)$), but it is comparatively far from being complete partite (though it is important that H is not *too far* from being complete partite, and also that H is very large). Using (Sym1), given a candidate for H which has too many imperfections, we can incrementally symmetrise it until this is no longer the case, and without decreasing λ.

In the second part of the proof (Claim 5.3), we use the strictness of λ to obtain a contradiction. We compare H with the graph H' obtained by removing all imperfections (roughly speaking H' is the closest complete partite graph to H). The ratios of part sizes of H' are necessarily close to some $\boldsymbol{x} \in \mathrm{OPT}$. The contradiction will come from the fact that $\lambda(H') - \lambda(H)$ is too large (which implies that H is actually far from optimal). We would like to argue that $\lambda(H) - \lambda(H')$ can be approximated looking at each *wrong pair* $e \in W := E(H) \triangle E(H')$ separately and summing its contribution to the function. This need not be true if e is incident to many other wrong pairs, so instead, we consider two families of wrong pairs: those incident to vertices in B, which are those with high degree in W, and the collection E' of remaining wrong pairs. The fact that each $e \in E'$ has a large contribution to $\lambda(H) - \lambda(H')$ will follow from (Str1): namely that $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda(\boldsymbol{x})$ is large, where i_1, i_2 are the indices of the parts where e lies. The fact that the edges incident to each $v \in B$ have a large contribution to $\lambda(H) - \lambda(H')$ is slightly more involved. For this we use (Sym2) to symmetrise the neighbourhood of v, and, depending on the attachment of v in the resulting graph, the required conclusion will follow from (Str1) (if it is ‘canonical’) and (Str2) (otherwise).

The following lemma will be useful when comparing λ evaluated on a complete partite graph with λ evaluated on the same graph with a few imperfections.

Lemma 5.1. *Let $c > 0$ and let $\gamma : \mathcal{G}_k \to \mathbb{R}$ be fixed. Let H, H' be graphs on the same vertex set of size h, where h is large and $H' \in \mathcal{P}$ has $\overline{\mathcal{P}}$-structure $V_0, V_1, \dots, V_m$. Write $R := E(H) \triangle E(H')$ and given $x \in V(H')$, write $p(x)$ for the index of the part of H' containing x. Define*

$$\xi_0 := k^2|R|c/h^2, \quad \xi_1 := 2\gamma_{\max}k^4|R|^2/h^4, \quad \xi_2 := 2\gamma_{\max}k^3|R|\Delta(R)/h^3.$$

Then $\lambda(H') - \lambda(H)$ is

- (i) *at least $\xi_0/2 - \xi_1 - \xi_2$ if $\nabla^{\bullet\bullet}_{p(x)p(y)}\lambda(\boldsymbol{x}_{H'}) \geqslant c$ for all $xy \in R$;*
- (ii) *at least $\xi_0/2 - \xi_2$ if $\nabla^{\bullet\bullet}_{p(x)p(y)}\lambda(\boldsymbol{x}_{H'}) \geqslant c$ for all $xy \in R$ and R is a star;*
- (iii) *at most $\xi_0 + \xi_1 + \xi_2$ if $\nabla^{\bullet\bullet}_{p(x)p(y)}\lambda(\boldsymbol{x}_{H'}) \leqslant c$ for all $xy \in R$.*

Proof. Write $S := \lambda(H') - \lambda(H) = \binom{h}{k}^{-1} \sum_{X \in \binom{V}{k}} (\gamma(H'[X]) - \gamma(H[X]))$ and

$$S_1 := \binom{h}{k}^{-1} \sum_{xy \in R} \sum_{X \in \binom{V}{k} : \{x,y\} \subseteq X} (\gamma(H'[X]) - \gamma((H' \oplus xy)[X]))$$

$$= \frac{\binom{h-2}{k-2}}{\binom{h}{k}} \sum_{xy \in R} \nabla^{\bullet\bullet}_{p(x)p(y)}\lambda(H') = \frac{\binom{k}{2}}{\binom{h}{2}} \sum_{xy \in R} \left(\nabla^{\bullet\bullet}_{p(x)p(y)}\lambda(\boldsymbol{x}_{H'}) + o(1)\right).$$

Then $\binom{h}{k}|S - S_1| \leqslant \sum_{X \in I_1} 2\gamma_{\max}$ where $I_1 := \{X \in \binom{V}{k} : |R \cap \binom{X}{2}| \geqslant 2\}$. The number of X that contain two disjoint pairs from R is at most $|R|^2 \cdot \binom{h-4}{k-4}$. The number of X containing two adjacent pairs from R is at most $|R| \cdot \Delta(R) \cdot \binom{h-3}{k-3}$. So,

$$\frac{|I_1|}{\binom{h}{k}} \leqslant \frac{|R|^2\binom{h-4}{k-4} + |R|\Delta(R)\binom{h-3}{k-3}}{\binom{h}{k}} \leqslant \frac{|R|^2k^4}{h^4} + \frac{|R|\Delta(R)k^3}{h^3}.$$

All three parts follow immediately, noting for (ii) that when R is a star, it has no disjoint pairs. □

We now have all the tools in place to prove our main theorem.

Proof of Theorem 3.3. Let λ be a symmetrisable graph parameter as in (1.1). Note that λ is not identically 0 (otherwise OPT is infinite). Lemma 4.1 implies that there exists $\beta > 0$ such that $\mathrm{OPT} \subseteq \overline{\mathcal{P}}_\beta$. So, $|\mathrm{supp}(\boldsymbol{x})| \leqslant 1/\beta$ for all $\boldsymbol{x} \in \mathrm{OPT}$.

Suppose that λ is strict with parameter $c > 0$. Without loss of generality, we may assume that $c \ll \beta, 1/\gamma_{\max}, 1/k$. We want to show that there exists a constant $C > 0$ such that for every graph G on at least $1/C$ vertices, there exists a complete partite graph H on the same vertex set such that $\hat{\delta}_1(G, H) \leqslant C(\lambda(v(G)) - \lambda(G))$. Suppose that this is false. That is, there exists a sequence of counterexamples $(G_n)_n$ with $v_n := v(G_n) \to \infty$, such that

$$1 \geqslant d_n := \hat{\delta}_1(G_n, \mathcal{P}_{v_n}) > n(\lambda(v_n) - \lambda(G_n)), \quad \text{so} \tag{5.1}$$

$$\lambda(v_n) \geqslant \lambda(G_n) > \lambda(v_n) - \frac{1}{n}, \tag{5.2}$$

and thus, $\lambda(G_n) - \lambda(v_n) \to 0$.

Using the graphs G_n, we now find a large graph H which is almost optimal and has a small but *comparatively large* number of imperfections.

Claim 5.2. For all $\varepsilon > 0$, there exists $\varepsilon' > 0$ such that the following holds. For all $N > 0$, there exist $\boldsymbol{x} \in \mathrm{OPT}$ and a graph H on vertex set $[h]$ such that $h > N$, $\delta_{\mathrm{edit}}(H, \boldsymbol{x}) \leqslant 2\varepsilon$ and $\lambda(H) \geqslant \lambda(h) - 1/N$. Further, $\delta_{\mathrm{edit}}(H, \mathcal{P}_h) \geqslant \min\{\varepsilon', N(\lambda(h) - \lambda(H))\}$.

Proof. We consider two cases depending on whether $(d_n)_n$ contains a subsequence converging to 0. If it does not, then our counterexamples are eventually always far from being complete partite. In this case, we perform an additional step of symmetrising each G_n to obtain a graph which has a controlled number of imperfections; this number will be a small fraction of v_n^2. In the other case, the counterexamples are becoming gradually more like complete partite graphs, so the number of imperfections could be subquadratic (in v_n).

Case 1: $(d_n)_n$ does not contain a subsequence converging to 0.

In this case, there exists $\xi > 0$ such that $d_n \geqslant \xi$ for all sufficiently large n. Since we are free to make ε and ξ smaller, we may assume without loss of generality that $\xi = \varepsilon$. Further, we may assume that $d_n \geqslant \varepsilon$ for all $n \in \mathbb{N}$.

Let $V_n := V(G_n)$. Property (Sym1) (applied with parameter ε) implies that there exists $n_0 = n_0(\varepsilon)$ such that for each $n \geqslant n_0$, we can find a sequence $G_{n,0}, G_{n,1}, \dots, G_{n,m(n)}$ of graphs on V_n such that $G_{n,0} := G_n$; $G_n' := G_{n,m(n)}$ is complete partite; for all $i \in [m(n)]$, we have $\lambda(G_{n,i-1}) \leqslant \lambda(G_{n,i})$; and $\delta_{\mathrm{edit}}(G_{n,i-1}, G_{n,i}) \leqslant \hat{\delta}_1(G_{n,i-1}, G_{n,i}) < \varepsilon$. By (5.2), we have for all $0 \leqslant i \leqslant m(n)$ that $\lambda(v_n) \geqslant \lambda(G_{n,i}) \geqslant \lambda(G_n) > \lambda(v_n) - 1/n$.

Let $\boldsymbol{y}_n := \boldsymbol{x}_{G_n'}$. By choosing a convergent subsequence since $(\overline{\mathcal{P}}, \delta_{\mathrm{edit}})$ is compact, we may assume that $\boldsymbol{y}_n$ converges to some $\boldsymbol{y} \in \overline{\mathcal{P}}$. But $\lambda(\boldsymbol{y}_n) \to \lambda_{\max}$, so $\boldsymbol{y} \in \mathrm{OPT}$ by the continuity of λ. By definition, $\delta_{\mathrm{edit}}(G_{n,0}, \mathrm{OPT}) \geqslant \delta_{\mathrm{edit}}(G_{n,0}, \mathcal{P}_{v_n}) = d_n \geqslant \varepsilon$ and $\delta_{\mathrm{edit}}(G_{n,m(n)}, \mathrm{OPT}) \to \delta_{\mathrm{edit}}(\boldsymbol{y}, \mathrm{OPT}) = 0$. Let t be the largest element of $[m(n)]$ such that $\delta_{\mathrm{edit}}(G_{n,t}, \mathrm{OPT}) \geqslant \varepsilon$, and let $J_n := G_{n,t}$. By increasing n_0, we can assume that $t < m(n)$. Then

$$\delta_{\mathrm{edit}}(J_n, \mathrm{OPT}) \leqslant \delta_{\mathrm{edit}}(G_{n,t}, G_{n,t+1}) + \delta_{\mathrm{edit}}(G_{n,t+1}, \mathrm{OPT}) < 2\varepsilon.$$

That is, $\delta_{\mathrm{edit}}(J_n, \mathrm{OPT}) \in [\varepsilon, 2\varepsilon]$. Let $\boldsymbol{x}_n \in \mathrm{OPT}$ be such that $\delta_{\mathrm{edit}}(J_n, \boldsymbol{x}_n) = \delta_{\mathrm{edit}}(J_n, \mathrm{OPT})$. We claim that there exists $\varepsilon' > 0$ for which $p_n := \delta_{\mathrm{edit}}(J_n, \mathcal{P}_{v_n}) \geqslant \varepsilon'$ for all sufficiently large n.

Indeed, if the claim is not true, then by passing to a subsequence, we may assume that $p_n \to 0$. For each n, pick a complete partite graph P_n on v_n vertices with $\delta_{\mathrm{edit}}(J_n, P_n) = p_n$. Let $\boldsymbol{z}_n := \boldsymbol{x}_{P_n} \in \overline{\mathcal{P}}$ be the sequence that encodes the part ratios of P_n. We can pass to a subsequence of n such that $\boldsymbol{z}_n$ converges to some $\boldsymbol{z} \in \overline{\mathcal{P}}$; then $\lambda(\boldsymbol{z}) = \lim_{n\to\infty} \lambda(P_n) = \lambda_{\max}$. Thus, $\boldsymbol{z} \in \mathrm{OPT}$. However, by Lemma 2.1,

$$\varepsilon \leqslant \delta_{\mathrm{edit}}(J_n, \mathrm{OPT}) \leqslant \delta_{\mathrm{edit}}(J_n, \boldsymbol{z}) \leqslant \delta_{\mathrm{edit}}(J_n, P_n) + \delta_{\mathrm{edit}}(\boldsymbol{z}_n, \boldsymbol{z}) \leqslant p_n + o(1) \to 0,$$

a contradiction.

This ε' satisfies the lemma. Indeed, for any given $N > 0$, choose $n > N$ sufficiently large so that $v_n > N$ and $\delta_{\mathrm{edit}}(J_n, \boldsymbol{x}_n) \in [\varepsilon, 2\varepsilon]$ and $\delta_{\mathrm{edit}}(J_n, \mathcal{P}_{v_n}) \geqslant \varepsilon'$. Then we can set $\boldsymbol{x} := \boldsymbol{x}_n$ and $H := J_n$ and $h := v_n$, since $\lambda(J_n) \geqslant \lambda(v_n) - 1/n \geqslant \lambda(h) - 1/N$. The claim is proved in this case.

Case 2: $(d_n)_n$ contains a subsequence $(d_{n_i})_i$ such that $d_{n_i} \to 0$ as $i \to \infty$.

Assume without loss of generality that $(d_n)_n \to 0$. Therefore, there exists a sequence $(\boldsymbol{x}_n)_n$ with $\boldsymbol{x}_n \in \overline{\mathcal{P}}$ such that $\delta_{\text{edit}}(G_n, \boldsymbol{x}_n) \to 0$. By choosing a convergent subsequence of $(\boldsymbol{x}_n)_n$, we may assume that the sequence itself converges to some $\boldsymbol{x} \in \overline{\mathcal{P}}$. Then for sufficiently large n, $\delta_{\text{edit}}(G_n, \boldsymbol{x}) \leqslant \delta_{\text{edit}}(G_n, \boldsymbol{x}_n) + \delta_{\text{edit}}(\boldsymbol{x}_n, \boldsymbol{x}) \to 0$. Then the continuity of λ with respect to δ_{edit} and (5.2) imply that $\boldsymbol{x} \in \text{OPT}$. We can choose n sufficiently large so that, by (5.1), $H := G_n$ satisfies all the required properties in Claim 5.2 (where, for concreteness, we let $\varepsilon' := 1$). This completes the proof of the claim. ☐

Choose an additional constant $0 < \eta \ll c$. Obtain $\varepsilon > 0$ by applying Corollary 3.2 with $\eta^2, 6\varepsilon$ playing the roles of ε, δ, respectively. We may assume that $\varepsilon \ll \eta$. Claim 5.2 furnishes us with an $\varepsilon' > 0$ which we may assume satisfies $\varepsilon' \ll \varepsilon$. Now choose $N \in \mathbb{N}$ such that $1/N \ll \varepsilon'$. We have the following hierarchy:

$$0 < 1/N \ll \varepsilon' \ll \varepsilon \ll \eta \ll c \ll \beta, 1/\gamma_{\max}, 1/k. \tag{5.3}$$

Apply Claim 5.2 to yield an $\boldsymbol{x} \in \text{OPT}$ and a graph H on $h \geqslant N$ vertices. Let us list some properties of $\boldsymbol{x}$ (which will be all we need from now on):

(P1) $m := |\text{supp}(\boldsymbol{x})| \leqslant 1/\beta$.

(P2) $\boldsymbol{x} \in \overline{\mathcal{P}}_\beta$.

(P3) $\nabla^{\bullet\bullet}_{i_1 i_2}\lambda(\boldsymbol{x}) \geqslant c$ for all $i_1, i_2 \in \text{supp}^*(\boldsymbol{x})$.

(P4) $\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) \geqslant c((1-\alpha)x_0 + \min_{i \in \text{supp}^*(\boldsymbol{x})} w_i)$, where

$$w_i = \mathbb{1}_{i>0} b_i x_i + \textstyle\sum_{j \in \text{supp}^*(\boldsymbol{x}) \setminus \{0,i\}} (1 - b_j) x_j,$$

for all $b : \mathbb{N} \to \{0,1\}$ and $\alpha \in [0,1]$.

(P5) Whenever $\boldsymbol{y} \in \overline{\mathcal{P}}$ satisfies $\delta_{\text{edit}}(\boldsymbol{x}, \boldsymbol{y}) \leqslant 6\varepsilon$ and $\text{supp}^*(\boldsymbol{y}) = \text{supp}^*(\boldsymbol{x}) =: S$, we have that $|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leqslant \eta^2$ for all choices of $i_1, i_2 \in S$, $b : \mathbb{N} \to \{0,1\}$, $\alpha \in [0,1]$ and $f \in \{\lambda, \nabla^{\bullet\bullet}_{i_1 i_2}\lambda, \lambda(\cdot, (b, \alpha)), \nabla^{\bullet}_{b,\alpha}\lambda\}$.

(P6) $\delta_{\text{edit}}(\boldsymbol{x}, H) \leqslant 2\varepsilon$.

Properties (P1) and (P2) follow immediately from $|\text{OPT}| < \infty$ and Lemma 4.1. Properties (P3) and (P4) follow since λ is strict with parameter c. Property (P5) follows from our choice of ε and the fact from Corollary 3.2 that any f in this family of functions is uniformly equicontinuous. Property (P6) is a direct consequence of Claim 5.2.

Let $\mathcal{H}$ be the family of h-vertex graphs with $\overline{\mathcal{P}}$-structure $(V_i : i \in \text{supp}^*(\boldsymbol{x}))$, that is, V_0 (if it exists) is a clique, V_i is a non-empty independent set for all $i \in [m] = \text{supp}^*(\boldsymbol{x}) \setminus \{0\}$ and (V_i, V_j) is complete for every distinct $i, j \in \text{supp}^*(\boldsymbol{x})$.

Among all graphs in $\mathcal{H}$, let H' be one whose edit distance δ_{edit} to H is minimised, with $\overline{\mathcal{P}}$-structure $(V_i : i \in \text{supp}^*(\boldsymbol{x}))$ as above, where $V(H') = [h] = \bigcup\{V_i : i \in \text{supp}^*(\boldsymbol{x})\}$. Define $W := ([h], E(H) \triangle E(H'))$, and call the edges of W *wrong*. By the definition of H', (P6) and Lemma 2.1, we have that $\delta_{\text{edit}}(H, H') \leqslant \delta_{\text{edit}}(H, \boldsymbol{x}) + O(1/h) \leqslant 3\varepsilon$. Consequently, $e(W) = \hat{\Delta}_1(H, H') \leqslant 3h^2/2 \cdot \delta_{\text{edit}}(H, H') \leqslant 5\varepsilon h^2$. Let $\boldsymbol{v}$ be the vector of part ratios in H', that is, $\boldsymbol{v} := \boldsymbol{x}_{H'}$.

Then

$$\delta_{\text{edit}}(\boldsymbol{v}, \boldsymbol{x}) = \delta_{\text{edit}}(H', \boldsymbol{x}) \leqslant \delta_{\text{edit}}(H, H') + \delta_{\text{edit}}(H, \boldsymbol{x}) \overset{(P6)}{\leqslant} 5\varepsilon. \tag{5.4}$$

Note that, by (P2), this implies $v_i = |V_i|/h \geqslant c/2$ for all $i \in [m]$. Call a vertex x *bad* if it is incident to at least ηh wrong pairs, that is, $d_W(x) \geqslant \eta h$. Let B consist of all bad vertices and $B^c = [h] \setminus B$

of all *good* (i.e. not bad) vertices. Let also $E' := E(W[B^c])$ and $e' := |E'|$. By definition of B and that $\varepsilon \ll \eta$, we have

$$e' \leqslant e(W) \leqslant 5\varepsilon h^2 \quad \text{and} \quad |B| \leqslant \frac{2e(W)}{\eta h} \leqslant \frac{10\varepsilon}{\eta} \cdot h \leqslant \sqrt{\varepsilon}h. \tag{5.5}$$

For a vertex v of H', let $H' \oplus v$ denote the graph obtained from H' by removing every edge containing v and then for all $y \in [h] \setminus \{v\}$ adding the edge vy if and only if $y \in N_H(v)$. The heart of the proof is the following claim.

Claim 5.3. The following statements hold.

(i) $\delta_{\text{edit}}(H', H) \leqslant 2(\frac{|B|}{h} + \frac{e'}{h^2})$.

(ii) For every $v \in B$, $\lambda(H') - \lambda(H' \oplus v) \geqslant \frac{kc\eta^{3/2}}{3h}$.

(iii) $\lambda(H') - \lambda(H) \geqslant \eta^2(\frac{|B|}{h} + \frac{e'}{h^2})$.

We first see how this claim completes the proof of Theorem 3.3. We have by Claim 5.2 that

$$\begin{aligned}\frac{1}{N} &\geqslant \lambda(h) - \lambda(H) \geqslant \lambda(H') - \lambda(H) \overset{(iii)}{\geqslant} \eta^2\left(\frac{|B|}{h} + \frac{e'}{h^2}\right) \overset{(i)}{\geqslant} \frac{\eta^2}{2}\delta_{\text{edit}}(H', H) \geqslant \frac{\eta^2}{2}\delta_{\text{edit}}(H, \mathcal{P}_h) \\ &\geqslant \frac{\eta^2}{2}\min\{\varepsilon', N(\lambda(h) - \lambda(H))\}.\end{aligned} \tag{5.6}$$

If $\varepsilon' \leqslant N(\lambda(h) - \lambda(H))$, then considering the first and last terms of (5.6) gives $1/N \geqslant \eta^2\varepsilon'/2$, a contradiction to our choice of N (i.e. (5.3)). If instead $\varepsilon' > N(\lambda(h) - \lambda(H))$, then considering the second and last terms of (5.6) gives $1 \geqslant \eta^2 N/2$, also a contradiction to our choice of N. Thus, Theorem 3.3 holds given Claim 5.3.

Proof of Claim 5.3. For (i), we see that

$$\delta_{\text{edit}}(H', H) \leqslant \hat{\delta}_1(H', H) \leqslant \frac{\sum_{v \in B} d_W(v) + e'}{h^2/2} \leqslant \frac{2|B|}{h} + \frac{2e'}{h^2}.$$

For (ii), fix an arbitrary $v \in B$, and let $p(v) \in \operatorname{supp}^*(\boldsymbol{x})$ be such that $v \in V_{p(v)}$. Let $\mathcal{H}$ consist of all graphs G on $[h] = V(H')$ with $G - v = H' - v$ and for each $i \in [m]$, either $N_G(v) \supseteq V_i \setminus \{v\}$ or $N_G(v) \cap (V_i \setminus \{v\}) = \emptyset$ (and with arbitrary attachment to V_0). That is, either v is adjacent to every vertex or no vertices in each part $V_1, \dots, V_m$. For brevity, let $H'_v := H' \oplus v$.

Apply (Sym2) with parameter ε to H'_v at v to obtain a sequence of graphs $H'_v =: H_0, H_1, \dots, H_r \in \mathcal{H}$, such that $H_i - v = H'_v - v$ for all $i \in [r]$; $\lambda(H_{i-1}) \leqslant \lambda(H_i)$; and $\hat{\Delta}_1^v(H_{i-1}, H_i) \leqslant \varepsilon(h-1)$ for all $i \in [r]$, where here for any two graphs J, J' which differ only at a vertex v, we define $\hat{\Delta}_1^v(J, J')$ to be the minimum number of edits of pairs containing v to make J equal to J'.

By the definition of $\mathcal{H}$, there are $b : [m] \to \{0,1\}$ and $0 \leqslant \alpha \leqslant 1$ such that $b(i) = 1$ if $N_{H_r}(v) \supseteq V_i \setminus \{v\}$ and $b(i) = 0$ if $N_{H_r}(v) \cap (V_i \setminus \{v\}) = \emptyset$; and $d_{H_r}(v, V_0) = \lfloor \alpha |V_0| \rfloor$ (if $V_0 = \emptyset$ we let $\alpha := 1$). We consider two cases depending on (b, α): in Case 1, the attachment of v in H_r is very different to any vertex in $H' - v$, and in Case 2, it is similar.

Case 1: At least one of the following holds: (a) $x_0 > 0$ and $\alpha < 1 - \eta/2$; (b) $|b^{-1}(0)| \geqslant 2$; (c) $x_0 = 0$ and $b^{-1}(0) = \emptyset$.

We will first show that

$$\frac{h}{k}\big(\lambda(H')-\lambda(H_r)\big) \geqslant \nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) - 3\eta^2. \tag{5.7}$$

For this, let $\boldsymbol{y}$ be the vector of part ratios of $H'_r - v = H' - v$, that is, $y_i = |V_i|/(h-1)$ if $i \in \{0,\dots,m\}\setminus\{p(v)\}$ and $y_{p(v)} = (|V_{p(v)}|-1)/(h-1)$. Then $H_r = (H'-v) +_{b,\alpha} v = G_{h-1,\boldsymbol{y}} +_{b,\alpha} v$ and so

$$\begin{aligned}\frac{h}{k}\big(\lambda(H')-\lambda(H_r)\big) &= \frac{h}{k}\cdot\binom{h}{k}^{-1}\sum_{X\in\binom{V}{k}:X\ni v}\big(\gamma(H'[X]) - \gamma((G_{h-1,\boldsymbol{y}} +_{b,\alpha} v)[X])\big)\\ &= \frac{h}{k}\cdot\frac{\binom{h-1}{k-1}}{\binom{h}{k}}\big(\lambda(H',v) - \lambda(G_{h-1,\boldsymbol{y}} +_{b,\alpha} v, v)\big)\\ &= \lambda(\boldsymbol{v},(e_{p(v)},1)) - \lambda(\boldsymbol{y},(b,\alpha)) + O(1/h).\end{aligned}$$

Now $\delta_{\mathrm{edit}}(\boldsymbol{x},\boldsymbol{y}), \delta_{\mathrm{edit}}(\boldsymbol{x},\boldsymbol{v}) \leqslant 5\varepsilon$. So, we have

$$\lambda(\boldsymbol{v},(e_{p(v)},1)) - \lambda(\boldsymbol{y},(b,\alpha)) \overset{(P5)}{\geqslant} \lambda(\boldsymbol{x},(e_{p(v)},1)) - \lambda(\boldsymbol{x},(b,\alpha)) - 2\eta^2 = \nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) - 2\eta^2,$$

where the final equality follows from Proposition 3.1(i). This proves (5.7). Now,

$$\lambda(H') - \lambda(H'\oplus v) \geqslant \lambda(H') - \lambda(H_r) \overset{(5.7)}{\geqslant} \frac{k}{h}\Big(\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) - 3\eta^2\Big)$$

so to complete the proof of the claim, it suffices to show that $\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) \geqslant c\eta^{3/2}/2$.

We will use the lower bound on $\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x})$ given by (P4), and that $x_i \geqslant \beta > c$ for all $i \in \mathrm{supp}^*(\boldsymbol{x})$ from (P2). Suppose first that (a) holds. Since each term in the expression for w_i is non-negative, we have for all $i \in \mathrm{supp}^*(\boldsymbol{x}) = \{0,\dots,m\}$ that $\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{x}) \geqslant c(1-\alpha)x_0 \geqslant c\beta\eta/2 > c\eta^{3/2}/2$, as required. Suppose secondly that (b) holds. Then for every $i \in [m]$, either $b_i = 1$ or $b_j = 0$ for some $j \in [m]\setminus\{i\}$. So, $w_i \geqslant \min_{j\in[m]} x_j \geqslant \beta$ for all $0 \leqslant i \leqslant m$, as required. Finally, if (c) holds, $x_0 = 0, b = (1,1,\dots)$, $\mathrm{supp}^*(\boldsymbol{x}) = [m]$ and $w_i = x_i \geqslant \beta$ for all $i \in [m]$, as required.

Case 2: Either (a) $x_0 = 0$ and $|b^{-1}(0)| = 1$, or (b) $x_0 > 0$, $\alpha > 1-\eta/2$, and $|b^{-1}(0)| \leqslant 1$.

Notice that Cases 1 and 2 are the only possible outcomes (recalling that if $x_0 = 0$, then $\alpha = 1$). For all $0 \leqslant i \leqslant r$, let

$$d_i := \min_{j\in\mathrm{supp}^*(\boldsymbol{x})} \hat{\Delta}^v_1(H_i, (H'-v) +_{e_j,1} v),$$

that is, the smallest number of edits at v needed to move v into some part in H_i. Now, $d_0 = d_W(v) \geqslant \eta h$. On the other hand, d_r is comparatively small: if (a) holds, then $d_r = 0$, and if (b) holds, then $d_r \leqslant \eta h/2$. So, we can choose the largest integer $0 \leqslant t < r$ such that $d_t \geqslant \eta h$, and let $H^* := H_t$ and $d^* := d_t$. Let $j^* \in \mathrm{supp}^*(\boldsymbol{x})$ be such that $d^* = \hat{\Delta}^v_1(H_t, \widetilde{H})$ where $\widetilde{H} := (H'-v) +_{e_{j^*},1} v$. So, $\lambda(H^*) \geqslant \lambda(H'\oplus v)$ and additionally

$$\eta h \leqslant d^* \leqslant \hat{\Delta}^v_1(H_t, H_{t+1}) + d_{t+1} \leqslant \varepsilon(h-1) + \eta h \leqslant 2\eta h.$$

So, one must make between ηh and $2\eta h$ edits to H^* at v to move v to the j^*th part, and the (complete partite) graph obtained in this way is $\widetilde{H}$. Since $\delta_{\mathrm{edit}}(\widetilde{H}, H')\cdot h^2/2 \leqslant$

$\hat{\Delta}_1(\widetilde{H},H') \leqslant h$, we have by (5.4) that $\delta_{\mathrm{edit}}(\boldsymbol{x}_{\widetilde{H}},\boldsymbol{x}) \leqslant \delta_{\mathrm{edit}}(\widetilde{H},H') + \delta_{\mathrm{edit}}(H',\boldsymbol{x}) \leqslant 5\varepsilon + 2/h \leqslant 6\varepsilon$. Now, (P5) and (P3) imply that for each of the d^* vertices u for which uv was flipped, we have $\nabla^{\bullet\bullet}_{p(u)p(v)}\lambda(\boldsymbol{x}_{\widetilde{H}}) \geqslant \nabla^{\bullet\bullet}_{p(u)p(v)}\lambda(\boldsymbol{x}) - \eta^2 \geqslant c/2$. Lemma 5.1(ii) implies that $\lambda(\widetilde{H}) - \lambda(H^*) \geqslant k^2 d^* c/2h^2 - 2\gamma_{\max} k^3 (d^*)^2/h^3$. So,

$$\lambda(H') - \lambda(H' \oplus v) \geqslant \lambda(\widetilde{H}) - \lambda(H^*) + O(1/h) \geqslant k^2\eta^{3/2}c/3h,$$

as required for (ii).

For (iii), our task is to obtain a suitable lower bound on $T := \lambda(H') - \lambda(H)$. Notice that the only k-sets X contributing to T are those containing some $e \in W$. Let

$$T_0 := \textstyle\sum_{v\in B}(\lambda(H') - \lambda(H' \oplus v)) \quad \text{and} \quad T' := \lambda(H') - \lambda(H' \bigtriangleup E').$$

In a similar fashion to part (ii), we will first give lower bounds for T_0, T', respectively, and then show that T is well approximated by $T_0 + T'$. First consider T_0. By Claim 5.3(ii), we have $T_0 \geqslant |B|k\eta^{3/2}c/(3h)$. Now consider T'. Again, $\nabla^{\bullet\bullet}_{p(x)p(y)}\lambda(\boldsymbol{x}_{H'}) \geqslant c/2$ for all $xy \in E'$, so Lemma 5.1 and (5.5) imply that

$$T' \geqslant k^2 e'/h^2 \cdot (c/2 - 50\gamma_{\max}k^2\varepsilon^2 - 2\gamma_{\max}k\eta) \geqslant \frac{k^2ce'}{4h^2}.$$

For the final step, note that $\binom{h}{k}|T - T_0 - T'| \leqslant \sum_{X\in I_0} 2\gamma_{\max}$, where

$$I_0 = \left\{ X \in \binom{V}{k} \,:\, |X \cap B| \geqslant 2 \text{ or } |X \cap B|, e(W[X \setminus B]) \geqslant 1 \right\}.$$

But

$$\frac{|I_0|}{\binom{h}{k}} \leqslant \frac{|B|^2\binom{h-2}{k-2} + |B|e'\binom{h-3}{k-3}}{\binom{h}{k}} \overset{(5.5)}{\leqslant} \frac{|B|}{h}\left(2k^2\sqrt{\varepsilon} + 5k^3\varepsilon\right) \leqslant \varepsilon^{1/3}T_0. \tag{5.8}$$

Thus,

$$T \geqslant T_0 + T' - \frac{2\gamma_{\max}}{\binom{h}{k}}|I_0| \overset{(5.8)}{\geqslant} T_0/2 + T' \geqslant \frac{|B|k\eta^{3/2}c}{6h} + \frac{k^2ce'}{4h^2} \geqslant \eta^2\left(\frac{|B|}{h} + \frac{e'}{h^2}\right),$$

as desired. This completes the proof of Claim 5.3. ☐

Thus, we complete the proof of Theorem 3.3. ☐

6 | APPLICATIONS TO INDUCIBILITY

Firstly we prove Lemma 1.2 which is essentially Theorem 1 in [36].

Proof of Lemma 1.2. In fact, we can require that $|E(G_{i-1}) \bigtriangleup E(G_i)|$ is at most $n-1$ (resp. at most 1) in (Sym1) (resp. (Sym2)) for every graph G of every order $n \geqslant k$.

Let us show (Sym1). Initially, let $H := G$ and let $\mathcal{V} = \{V_1, \ldots, V_n\}$ be the partition of $V(H)$ into singletons. At each stage, every part of $\mathcal{V}$ will consist of twin vertices, that is, vertices with identical neighbourhoods (in particular, every part is an independent set). We will modify the current graph H and the current partition $\mathcal{V} = \{V_1, \dots, V_s\}$ so that at each step, λ does not decrease, while the affected edges are incident to a single vertex.

If for each $1 \leqslant i < j \leqslant s$, $H[V_i, V_j]$ is complete bipartite, then H is a complete partite graph so we stop. Otherwise, pick $i < j$, $x \in V_i$ and $y \in V_j$ such that $xy \notin E(H)$. Let $X = N_H(x)$ and $Y = N_H(y)$. Fix a complete partite graph F. Note that every $A \subseteq V$ with $H[A] \cong F$ is one of the following four kinds: (1) $x \in A, y \notin A$; (2) $x \notin A, y \in A$; (3) $x \in A, y \in A$ and (4) $x \notin A, y \notin A$. Given $H - x - y$, we can thus write

$$p(F, H) = f_F(X) + f_F(Y) + g_F(X \cap Y, V \setminus (X \cup Y)) + C_F$$

for some constant $C_F > 0$ and functions f_F and g_F. Here $f_F(X)$ (resp. $f_F(Y)$) counts the number of copies of F of type (1) (resp. type (2)) as this depends only on X (resp. Y). For disjoint U, W, we define $g_F(U, W)$ to be the number of copies of any graph J with $V(J) \subseteq U \cup W$ such that by adding two new vertices z, z' to J and adding edges $\{uz, uz' : u \in U\}$ to J, we obtain a copy of F. Observe that if $\{x, y\} \cup V(J)$ induces a copy of F in H as above, then x and y are in the same partite set, $U \cap V(J) \subseteq X \cap Y$ and $W \cap V(J) \cap (X \cup Y) = \emptyset$. Thus, $g_F(X \cap Y, V \setminus (X \cup Y))$ counts type (3) copies. The type (4) count is a constant depending only on $H - x - y$. Then, letting $f = \sum_F c_F \cdot f_F$ and defining g, C similarly, we have

$$\lambda(H) = f(X) + f(Y) + g(X \cap Y, V \setminus (X \cup Y)) + C. \tag{6.1}$$

Notice that $g(\cdot, \cdot)$ is non-decreasing in both arguments, that is,

$$g(U, W) \leqslant g(U', W'), \quad \forall\, U \subseteq U', W \subseteq W'. \tag{6.2}$$

Indeed, if F is a clique, then no copy of F contains both x and y, and $c_F \geqslant 0$ otherwise.

Suppose that $f(X) \geqslant f(Y)$, let H_{xy} be the graph obtained from H by making y a clone of x. Let $H' = H_{xy}$ and let $\mathcal{V}'$ be obtained from $\mathcal{V}$ by moving y to the part containing x. It satisfies all the claimed properties as

$$\begin{aligned} \lambda(H_{xy}) &= 2f(X) + g(X, V \setminus X) + C \\ &\overset{(6.1),(6.2)}{\geqslant} f(X) + f(Y) + g(X \cap Y, V \setminus (X \cup Y)) + C = \lambda(H). \end{aligned}$$

Finally, it remains to argue that one can avoid infinite cycles. The rule for breaking ties $f(X) = f(Y)$ with, for example, $|V_i| \geqslant |V_j|$ is to take $H' = H_{xy}$. This strictly increases $\sum_{V \in \mathcal{V}} |V|^2 \in [n, n^3]$ so that are at most n^3 steps where λ stays constant. (In fact, one can bound the total number of steps by $1 + 2 + \cdots + n - 1 = \binom{n}{2}$: if there are currently $i \geqslant 2$ groups and we merge one group entirely into another, then we can do this by moving at most $n - i + 1$ vertices.)

Let us show (Sym2). Given G and z as in the property, we have a partition consisting of all partite sets in $G - z$ and z will always stay a single part. Given any partite set V_i of $G - z$, we can partition vertices $V_i = V_i' \cup V_i''$ depending on their adjacency to z, say $V_i' \subseteq N(z)$. Start with this initial partition into parts V_i' and V_i''. Fix arbitrary non-adjacent vertices $x \in V_i', y \in V_i''$, note that (6.1) and (6.2) still hold. If $f(X) > f(Y)$, take $G' = G_{xy}$. If $f(X) < f(Y)$, take $G' = G_{yx}$. The rule for breaking ties is again to clone the vertex from the larger part: if $f(X) = f(Y)$ and, say, $|V_i'| \geqslant |V_i''|$, take $G' = G_{xy}$. Otherwise, take $G' = G_{yx}$. Note that G' differs from G only in one pair. As before, λ has not decreased. Then redefine V_i', V_i'' and repeat the process. The final graph has $N(z) \subseteq V_i$ or $N(z) \cap V_i = \emptyset$. (Note that each tie $f(X) = f(Y)$ strictly increases $|V_i'|^2 + |V_i''|^2$ so as before there are at most n^3 steps where λ stays constant, so there are no infinite cycles.) Repeating this for all i, we make at most n steps in total, and the resulting graph is as desired. □

6.1 | Proofs of Theorems 1.3–1.6

Since by Lemma 1.2, $p(F, \cdot)$ is symmetrisable whenever F is complete partite, to prove Theorems 1.3–1.6, it suffices to determine OPT (if it is not already known), and then check that $p(F, \cdot)$ is strict. The result then follows from Theorem 3.3.

In all cases, OPT consists of a single point, and checking strictness is generally straightforward (it is slightly more involved for $F = K_{1,t}$). However, determining OPT where it is not already known, for $F = K_{2,1,1,1}$ and $F = K_{3,1,1}$, is challenging and we are required to solve a polynomial optimisation problem. We use computer-assisted semidefinite programming to solve the last problem.

Proof of Theorem 1.3. Assume $s \leqslant t$. Firstly we collect some facts about the function $f_{s,t}$ defined on [0,1] given by $f_{s,t}(\alpha) = \alpha^s(1-\alpha)^t + \alpha^t(1-\alpha)^s$, recalling that for $s < t$, $f_{s,t}(\alpha) = \binom{t+s}{t}^{-1} p(K_{s,t}, (\alpha, 1-\alpha, 0, \ldots))$ for $\alpha \in [\frac{1}{2}, 1]$, and $f_{s,s}$ can similarly be expressed with a factor of $\frac{1}{2}$ on the right-hand side:

(i) If $s \geqslant \binom{t-s}{2}$, then the unique maximum of $f_{s,t}$ in $[\frac{1}{2}, 1]$ is $\frac{1}{2}$.
(ii) If $s < \binom{t-s}{2}$, then $f'_{s,t}$ has a single root in $(\frac{1}{2}, 1)$, which corresponds to a maximum, has $\frac{1}{2}$ a root corresponding to a minimum and has no other roots in $[\frac{1}{2}, 1)$.
(iii) If $\alpha \in [\frac{1}{2}, 1]$ maximises $f_{1,t}$, then $1 - \alpha > \frac{1}{t+1}$.
(iv) $\max_{\alpha \in [0,1]} (t+1)\alpha^t(1-\alpha) = (\frac{t}{t+1})^t$, attained uniquely at $\frac{t}{t+1}$.

Note that $f_{s,t}$ is symmetric about $\alpha = \frac{1}{2}$. For (i) and (ii), we just follow the proof of [7, Theorem 3]. We have

$$f'_{s,t}(\alpha) = \alpha^t(1-\alpha)^{s-1} h\left(\frac{1-\alpha}{\alpha}\right) \quad \text{where } h(x) = sx^{t-s+1} - tx^{t-s} + tx - s.$$

Assume first that $s \geqslant \binom{t-s}{2}$. Setting $x = 1 + \varepsilon$ for $\varepsilon \geqslant 0$, one can show that $h(x) > 0$, so $f_{s,t}$ is non-decreasing in $[0, \frac{1}{2}]$, and thus, the unique maximum of $f_{s,t}$ in [0,1] is at $\frac{1}{2}$, as required for (i). (In the calculation in [7, Theorem 3], it is shown that $h(x) \geqslant 0$, but there is equality in the first inequality only if $t - s = 1$, but in this case, the final inequality is strict.) Note that [7] uses $(t, s+t)$ and (a, b) instead of our (s, t). Assume secondly that $s < \binom{t-s}{2}$. Following the remarks after [7, Theorem 5], it suffices to show that h has a single root in (0,1). This is a consequence of $h(0) < 0$, $h(1) = 0$, $h'(1) < 0$ and $h''(x) < 0$ for all $x \in (0, 1)$, as required for (ii).

Next we show that (iii) holds. If $t = 2, 3$ (i.e. $1 \geqslant \binom{t-1}{2}$), then (i) implies that $1 - \alpha = \frac{1}{2}$, as required. If $t \geqslant 4$, then by (ii), $f'_{1,t}$ has a unique root (i.e. $1 - \alpha$) in $(0, \frac{1}{2})$, corresponding to a maximum and $\frac{1}{2}$ is a root corresponding to a minimum. Thus, $f'_{1,t}(x) > 0$ for $x \in (0, 1-\alpha)$ and $f'_{1,t}(x) < 0$ for $x \in (1-\alpha, \frac{1}{2})$. One can check that $f'(\frac{1}{t+1}) > 0$, which gives $1 - \alpha > \frac{1}{t+1}$. This proves (iii). Property (iv) can be easily checked via differentiation: indeed, $\frac{d}{d\alpha}\alpha^t(1-\alpha) =$

$\alpha^{t-1}(t-(t+1)\alpha)$ is strictly positive for $0<\alpha<\frac{t}{t+1}$, equals 0 at $\alpha=\frac{t}{t+1}$ and is strictly negative for $\alpha>\frac{t}{t+1}$.

Now we show that OPT $=\{(\alpha,1-\alpha,0,\ldots)\}$ with $\alpha\geqslant\frac{1}{2}$ (where $f_{s,t}(\alpha)$ is uniquely maximised). This was essentially proved by Brown and Sidorenko [7]. They do not prove the uniqueness of the optimal element but this can be extracted from their proof, so we only give a sketch of how to do this here.

Firstly we claim that if G is a complete multipartite graph on n vertices whose two largest parts of sizes n_r, n_{r-1} satisfy $n_r, n_{r-1}=\Omega(n)$ and $n-n_r-n_{r-1}=\Omega(n)$, then by merging parts, we increase the number of induced copies of $K_{s,t}$ by $\Omega(n^{s+t})$. Indeed, to see the claim, fix $\varepsilon>0$ and suppose $G=K_{n_1,n_2,\ldots,n_r}$ with $r\geqslant 3$ and $n_1\leqslant n_2\leqslant\ldots\leqslant n_r$ with $\sum_{i\in[r]} n_i=n$, $n_{r-1}\geqslant\varepsilon n$, $\sum_{i\in[r-2]} n_i\geqslant\varepsilon n$ and $G'=K_{n_1+n_2,n_3,\ldots,n_r}$. It is shown in [7, Proposition 2] that merging the two smallest parts in any complete multipartite graph with at least three parts does not decrease the number of induced copies of $K_{s,t}$. Thus, in G, we can successively merge two smallest parts until we obtain a graph G'' with exactly three parts, of sizes $m_1\leqslant m_2\leqslant m_3$ with $m_1\geqslant\varepsilon n$. Now merge the parts of size m_1 and m_2 to obtain a complete bipartite graph G'''. Then

$$\begin{aligned}
&I(K_{s,t},G''')-I(K_{s,t},G)\\
&\quad\geqslant I(K_{s,t},G''')-I(K_{s,t},G'')\\
&\quad=\binom{m_1+m_2}{s}\binom{m_3}{t}+\binom{m_1+m_2}{t}\binom{m_3}{s}\\
&\qquad-\binom{m_1}{s}\binom{m_3}{t}-\binom{m_1}{t}\binom{m_3}{s}-\binom{m_2}{s}\binom{m_3}{t}\\
&\qquad-\binom{m_2}{t}\binom{m_3}{s}-\binom{m_1}{s}\binom{m_2}{t}-\binom{m_1}{t}\binom{m_2}{s}\\
&\quad\geqslant\binom{m_1+m_2}{s}\binom{m_2}{t}+\binom{m_1+m_2}{t}\binom{m_2}{s}-\binom{m_1}{s}\binom{m_2}{t}-\binom{m_1}{t}\binom{m_2}{s}\\
&\qquad-\binom{m_2}{s}\binom{m_2}{t}-\binom{m_2}{t}\binom{m_2}{s}-\binom{m_1}{s}\binom{m_2}{t}-\binom{m_1}{t}\binom{m_2}{s}\\
&\quad=\frac{1}{t!s!}\left(m_2^t((m_1+m_2)^s-m_1^s-m_2^s-m_1^s)+m_2^s((m_1+m_2)^t-m_1^t-m_2^t-m_1^t)\right)+O(n^{s+t-1}).
\end{aligned}$$

To prove the claim, it suffices to prove that this is at least $O(\varepsilon)n^{s+t}$ for all (s,t). Neglecting the $O(n^{s+t-1})$ error terms, the quantity in the last line is at least

$$\begin{cases}
\frac{1}{t!s!}\left(m_1^s(t-2)m_1^t+m_1^t(s-2)m_1^s\right)\geq\frac{1}{t!s!}m_1^{s+t} & \text{if } s+t\geq 5\\
\frac{1}{3!}\left(2m_1m_2^3+3m_1^2m_2^2-m_1^3m_2\right)\geq\frac{2}{3}m_1^4 & \text{if } (s,t)=(1,3)\\
\frac{1}{2!2!}2m_2^2(2m_1m_2-m_1^2)\geq\frac{1}{2}m_1^4 & \text{if } (s,t)=(2,2)\\
\frac{1}{2}(m_2^2m_1-m_1^2m_2)\geq\frac{1}{2}(m_2-m_1)m_1^2 & \text{if } (s,t)=(1,2).
\end{cases}$$

Since $m_1\geqslant\varepsilon n$, this proves the claim unless $(s,t)=(1,2)$ and $m_2-m_1<\varepsilon n$. In this case, we have $I(K_{1,2},G''')-I(K_{1,2},G'')=m_1^2m_3-m_1^3+O(\varepsilon)n^3=\mu^2(1-3\mu)n^3+O(\varepsilon)n^3$ where $\mu :=$

m_1/n. We are done if $\mu < \frac{1}{3} - \varepsilon$. If not, G'' has three parts of size $\frac{1}{3} \pm \varepsilon$ which is far from optimal, by comparing to the complete balanced bipartite graph. This completes the proof of the claim.

Thus, if $s \geqslant 2$, then every $\boldsymbol{x} \in \mathrm{OPT}$ has exactly two non-zero entries which sum to 1, as required. We want to show that this also holds for $s = 1$. For this, we only need to show that there is no $0 < x \leqslant 1$ for which $\boldsymbol{x} = (x, 0, \ldots)$ is optimal. Indeed, $\lambda(\boldsymbol{x}) = p(K_{1,t}, \boldsymbol{x}) = (t+1)x^t(1-x) \leqslant (\frac{t}{t+1})^t$ by (iv). But $\lambda(K_{1,t}, (\frac{t}{t+1}, \frac{1}{t+1}, 0, \ldots)) = (t+1)f_{1,t}(\frac{t}{t+1}) = 2(\frac{t}{t+1})^t$, so $\boldsymbol{x} \notin \mathrm{OPT}$.

We have shown that every element of OPT is of the form $(\alpha, 1-\alpha, 0, \ldots)$ for some $\alpha \in [\frac{1}{2}, 1]$. By (i) and (ii), $f_{s,t}$ has a unique maximum in $[\frac{1}{2}, 1]$. Thus, OPT contains a unique element $\boldsymbol{x} := (\alpha, \beta, 0, \ldots)$ (where from now on we write $\beta := 1 - \alpha$). It remains to show that there is $c = c(\lambda) > 0$ such that (Str1) and (Str2) hold, where $\lambda(\cdot) := p(K_{s,t}, \cdot)$.

Let $G := G_{n,\boldsymbol{x}}$. First we check (Str1). A non-edge between two partite sets is not contained in any induced copy of $K_{s,t}$, nor is an edge within a partite set. So, $\nabla^{\bullet\bullet}_{xy}\lambda(G)$ is the number of copies of F in G containing the pair xy divided by $\binom{n-2}{s+t-2}$. Rather roughly, this is always at least $\beta^{s+t-2} + o(1)$ as $n \to \infty$.

Now we check (Str2). We have $x_0 = 0$ and $\mathrm{supp}^*(\boldsymbol{x}) = \{1, 2\}$. Since $x_0 = 0$, given any $(b(1), b(2)) =: (b_1, b_2) \in \{0,1\}^2$, we are required to show that

$$\nabla^{\bullet}_{b,1}\lambda(\boldsymbol{x}) = \lambda(\boldsymbol{x}) - \lim_{n\to\infty} \lambda(G_{n,\boldsymbol{x}} +_b u, u) \geqslant c \min\{b_1\alpha + (1-b_2)\beta, b_2\beta + (1-b_1)\alpha\}.$$

Recall that as usual the right-hand side equals 0 if $(b_1, b_2) \in \{(0,1), (1,0)\}$, so the inequality is trivially true. If $(b_1, b_2) = (0, 0)$, then u lies in no copies of $K_{s,t}$ in $G_{n,\boldsymbol{x}} +_{b,1} u$; similarly, if $(b_1, b_2) = (1, 1)$ and $s \geqslant 2$. So, we may assume that $(b_1, b_2, s) = (1, 1, 1)$, and we need to show $\nabla^{\bullet}_{b,1}\lambda(\boldsymbol{x}) \geqslant c\beta$ (recalling $\alpha \geqslant \beta$). We have

$$\lambda(\boldsymbol{x}) = (t+1)(\alpha\beta^t + \alpha^t\beta), \quad \lambda(\boldsymbol{x}, (b, 1)) = \alpha^t + \beta^t.$$

Recall that $\frac{\partial\lambda(x)}{\partial x_i} = (t+1)\lambda(\boldsymbol{x})$ for $i = 1, 2$ by Proposition 3.1(ii). We have

$$\frac{\partial\lambda(\boldsymbol{x})}{\partial x_1} = (t+1)(\beta^t + t\alpha^{t-1}\beta) \quad \text{and} \quad \frac{\partial\lambda(\boldsymbol{x})}{\partial x_2} = (t+1)(t\alpha\beta^{t-1} + \alpha^t), \text{ and so}$$

$$\begin{aligned} 2\nabla^{\bullet}_{b,1}\lambda(\boldsymbol{x}) = 2\lambda(\boldsymbol{x}) - 2\lambda(\boldsymbol{x}, (b,1)) &= \frac{1}{t+1}\left(\frac{\partial\lambda(\boldsymbol{x})}{\partial x_1} + \frac{\partial\lambda(\boldsymbol{x})}{\partial x_2}\right) - 2(\alpha^t + \beta^t) \\ &= \alpha^{t-1}(t\beta - \alpha) + \beta^{t-1}(t\alpha - \beta) = \alpha^{t-1}((t+1)\beta - 1) + \beta^{t-1}((t+1)\alpha - 1). \end{aligned}$$

It suffices to show that $\beta > \frac{1}{t+1}$, since then writing $\varepsilon = \beta - \frac{1}{t+1}$, we have $\nabla^{\bullet}_{b,1}\lambda(\boldsymbol{x}) \geqslant (t+1)^{1-t}\varepsilon$. This follows from (iii), completing the proof. □

Proof of Theorem 1.4. Firstly we show that for $F := K_r(t)$ with $t > 1 + \log r$, we have $\mathrm{OPT} = \{\boldsymbol{x}\}$ where $\boldsymbol{x} = (\frac{1}{r}, \ldots, \frac{1}{r}, 0, \ldots)$ and $x_0 = 0$. This essentially follows from [7] where it is proved that $\boldsymbol{x}$ lies in OPT (but without proving uniqueness), and [4, Theorem 13] where it is proved that the Turán graph with r parts is the unique extremal graph; but as in the proof of Theorem 1.3, we again need to make some modifications. Write $\lambda(\cdot) := p(K_r(t), \cdot)$, and observe that $\lambda(\boldsymbol{y}) = (tr)!/(r!(t!)^r) \cdot S_{t,\ldots,t}(\boldsymbol{y})$ where t is repeated r times. The method of Lagrange multipliers [7, Proposition 7] shows that every $\boldsymbol{y}$ which maximises λ has exactly r non-zero entries (which sum to

1, since $t > 1$). Thus, it suffices to show that $S_r^t(x_1, \dots, x_r) := x_1^t \dots x_r^t$ over all $x_1, \dots, x_r > 0$ with $x_1 + \dots + x_r = 1$ is uniquely maximised by $(\frac{1}{r}, \dots, \frac{1}{r})$. This is easy to see for $r = 2$. Suppose that it is not true for some $r \geqslant 3$, and so $x_1 \neq x_2$, say. Then $S_r^t(x_1, \dots, x_r) > S_r^t(\frac{x_1+x_2}{2}, \frac{x_1+x_2}{2}, x_3, \dots, x_r)$, a contradiction.

We have proved that $\mathrm{OPT} = \{\boldsymbol{x}\}$. It remains to show that there is $c = c(\lambda) > 0$ such that (Str1) and (Str2) hold.

Note that (Str1) is immediate as a non-edge between two partite sets is not contained in any induced copy of F, and an edge within a partite set is not contained in any induced copy of F. As in the proof of Theorem 1.3, this means that every $\nabla^{\bullet\bullet}_{ij}\lambda(\boldsymbol{x}) = \lambda(\boldsymbol{x})$, which is always at least $(\frac{1}{r})^{tr-2} + o(1)$. Similarly, (Str2) is immediate since any vertex in $G_{n,\boldsymbol{x}}$ without neighbours in at least two parts does not have a K_{r-1} in its neighbourhood so does not lie in any copies of $K_r(t)$, and any dominating vertex clearly lies in no copies of $K_r(t)$. So again, $\nabla^{\bullet}_{b,1}\lambda(\boldsymbol{x}) = \lambda(\boldsymbol{x})$ whenever $b \neq e_i$ for any $i \in [r]$. $\square$

Proof of Theorem 1.5. Let us show that $\lambda_{\max}$ is $\lambda_0 := \frac{525}{1024}$ and the vector $\boldsymbol{a} = (\frac{1}{8}, \dots, \frac{1}{8}, 0, \dots)$, which is the limit of $K^8_{n/8,\dots,n/8}$, is the unique maximiser. (Here, $K^{\ell}_{n_1,\dots,n_\ell}$ is the complete ℓ-partite graph with parts of size $n_1, \dots, n_\ell$.)

Let $\boldsymbol{x} \in \mathrm{OPT}$ be arbitrary. At some places, it will be convenient to use the language of finite graphs. So, let n be large and let $G = G_{n,\boldsymbol{x}}$ be a realisation of $\boldsymbol{x}$ with $\overline{\mathcal{P}}$-structure $V_0, \dots, V_m$.

Let us show that there are $\ell \in \mathbb{N}$, $p := (1 - x_0)/\ell$ and a sequence $\boldsymbol{x} = \boldsymbol{x}_0, \dots, \boldsymbol{x}_t = (p, \dots, p, 0, \dots) \in \mathrm{OPT}$ where for all $j \in [t]$, $x_{j,0} = x_0$, the entries of $\boldsymbol{x}_j$ are obtained by replacing some non-zero x_{j-1,i_1}, x_{j-1,i_2} in the entries of $\boldsymbol{x}_{j-1}$ with $z_j := x_{j-1,i_1} + x_{j-1,i_2}$, and any $\boldsymbol{y}_j$ obtained by replacing x_{j-1,i_1}, x_{j-1,i_2} in the entries of $\boldsymbol{x}_{j-1}$ with non-negative reals that sum to z_j is also in OPT.

Indeed, suppose that there are non-zero $x_i \neq x_j$, and let $i + j$ be minimal with this property. Each copy of $F = K_{2,1,1,1}$ intersects each $V_i \cup V_j$ in at most three vertices. Thus, if we fix the rest of $\boldsymbol{x}$, fix $s = x_i + x_j$ and vary $a = x_i/s$ between 0 and 1, then the number of copies of F is given by a polynomial $p(a)$ of degree at most 2 which is symmetric around $\frac{1}{2}$: $p(a) = p(1-a)$. (Note that the number of copies of F having $2 + 1$ vertices in $V_i \cup V_j$ is a constant times $a^2(1-a) + (1-a)^2 a = a - a^2$, which has no a^3 term, that is, is also a quadratic polynomial.) If p is not constant, then by symmetry, it follows that $p'(\frac{1}{2}) = 0$ and since p' is a linear function of a, this is the only root. Thus, p is maximised at $0, 1$ or $\frac{1}{2}$ and we can strictly increase p, a contradiction. Thus, p is constant, and any $\boldsymbol{z}$ obtained from $\boldsymbol{x}$ by replacing x_i and x_j by one or two new entries whose sum of sizes is $x_i + x_j$ is in OPT (corresponding to taking any value $0 \leqslant a \leqslant 1$). We let $i_1 = i$ and $i_2 = j$, giving $\boldsymbol{x}_1$. Then $\boldsymbol{x}_1$ and any $\boldsymbol{y}_1$ as described lie in OPT. If we cannot take $t = 1$, then $\boldsymbol{x}_1$ has unequal non-zero entries and we can repeat the above. It remains to check that this process terminates. If not, since we can always merge the largest part with the next (non-equal non-zero) largest part, for all $\varepsilon > 0$, there is some $m = m(\varepsilon) > 0$ such that $x_{m,1} > 1 - x_0 - \varepsilon$ (recalling $x_{m,0} = x_0$). Then $p(F, \boldsymbol{x}_m) = 5!(1-x_0)^2/2 \cdot x_0^3/6 + O(\varepsilon)$ which is maximised when $x_0 = \frac{3}{5}$, with value $\frac{216}{625} < \frac{525}{1024}$, a contradiction.

Let $\boldsymbol{y} = \boldsymbol{x}_t$, so $\boldsymbol{y}$ has ℓ equal non-clique parts each of ratio $p = (1-y)/\ell$ for some $y \in [0,1]$. Thus, $p(F, \boldsymbol{y})$ is equal to $h_\ell(y)$, where

$$h_\ell(y) := 5!\,\ell\,\frac{p^2}{2}\left(\frac{(1-p)^3}{3!} - (\ell-1)\frac{p^2}{2}(1-2p) - (\ell-1)\frac{p^3}{3!}\right).$$

Indeed, we first choose one part V_i where two non-adjacent vertices go ($\ell\binom{pn}{2}$ choices). Then the other three vertices of F have to go outside of V_i ($\binom{(1-p)n}{3}$ choices) except we have to rule out two (exclusive) cases: exactly two of them are in some V_j ($(\ell-1)\binom{pn}{2}(n-2pn)$ choices) and all three of them are in some V_j ($(\ell-1)\binom{pn}{3}$ choices). We have $h'_\ell(y)=\frac{10}{\ell^4}(y-1)q(y)$ where

$$\begin{aligned} q(y) = {} & -30+49\ell-21\ell^2+2\ell^3+90y-123\ell y+33\ell^2 y-90y^2 \\ & +99\ell y^2-12\ell^2 y^2+30y^3-25\ell y^3. \end{aligned}$$

We claim that for each $\ell \geqslant 8$, the function h_ℓ is strictly monotone decreasing (i.e. the optimal y is 0 meaning that the clique part is empty). So, it suffices to show that $q(y)>0$ for $\ell \geqslant 8$ and $y \in [0,1]$. We have that q is positive at its endpoints: $q(0)=42+97(\ell-8)+27(\ell-8)^2+2(\ell-8)^3$ and $q(1)=2\ell^3$. So, if $q(y)<0$ for some $y\in[0,1]$, then q' has a root in [0,1]. However, the quadratic polynomial q' has a negative coefficient at y^2 and is positive at endpoints: $q'(0)=1218+405(\ell-8)+33(\ell-8)^2$ and $q'(1)=9\ell^2$, so there is no such root. (These symbolic calculations can be found in `2111.nb` in the ancillary folder of the arXiv version of this paper [23].)

We claim that $k \in \mathbb{R}[\ell]$ given by $k(\ell)=h_\ell(0)$ is decreasing for $\ell \geqslant 8$. That is, out of all $\boldsymbol{y}$ with at least eight non-zero entries which are all equal, the unique extremal $\boldsymbol{y}$ is $\boldsymbol{a}$. Indeed, $k'(\ell)=-10j(\ell)/\ell^5$ where $j(\ell)=(\ell-9)^3+15(\ell-9)^2+60(\ell-9)+30$ so $k'(\ell)$ is decreasing for all $\ell\geqslant 9$, and also $k(9)=\frac{1120}{2187}<\frac{525}{1024}=k(8)$. (See `2111.nb`.)

Let us show that none of $\ell\in[7]$ is optimal. Fix such an ℓ. Direct calculations show that $h_\ell(0)<\lambda_0$ (while $y=1$ gives K_n which has zero density of F). So, it remains to investigate critical points, that is, $y\in(0,1)$ such that h_ℓ has derivative zero at y. Thus, $q(y)=0$.

Introduce a new variable z and define $p_1(y):=h'_\ell(y)$ and $p_2(y,z):=z-h_\ell(y)$. Thus, if y is a critical point with $\lambda_{\max}=h_\ell(y)$ and we define $z:=h_\ell(y)$, then (y,z) belongs to the variety $V=V(I)\subseteq\mathbb{R}^2$ defined by the ideal $I:=\langle p_1,p_2\rangle$ generated by the polynomials p_1,p_2. By applying Buchberger's algorithm to I (where we eliminate the variable y), we see that J, the intersection of I with $\mathbb{R}[z]$ (the set of polynomials that depend on z only), is generated by one polynomial q_ℓ, explicitly computed in `2111.nb` for every $\ell\in[7]$. We actually need only a part of the above claim, namely that there are polynomials $f_1,f_2\in\mathbb{R}[y,z]$ such that we have a polynomial identity $q_\ell(z)=f_1(y,z)p_1(y)+f_2(y,z)p_2(y,z)$, that is, all terms on the right-hand side depending on y cancel each other.

We have $q_1(z)=z(625z-216)$ which has roots at $z=0,\frac{216}{625}$, and $h_1(z_0)<\lambda_{\max}$ for both roots z_0. For each $2\leqslant l\leqslant 7$, the polynomial q_ℓ on inspection has the following properties: we have $q_\ell(z)=zr_\ell(z)$ where r_ℓ has degree at most 3, the coefficient of the leading term of r_ℓ is positive, and furthermore, we have $r_\ell(0)\geqslant 0$, $r_\ell(\lambda_0)<0$ and $r_\ell(1)<0$. This implies that r_ℓ has no roots in $[\lambda_0,1]$, and hence, q_ℓ has no roots in (0,1]. That is, it is impossible to have $\lambda_{\max}>\lambda_0$ (because, as a graph density, $\lambda_{\max}$ is at most 1). Thus, $\lambda_{\max}=\lambda_0$ and none of the polynomials can achieve $\lambda_{\max}$ except when $\ell=8$ (with $\boldsymbol{y}=(\frac{1}{8},\ldots,\frac{1}{8},0,\ldots)$ being the unique maximiser among $\boldsymbol{y}\in\overline{\mathcal{P}}$ with $y_i=y_j$ for all $i,j\in[8]$).

If $\boldsymbol{y}=\boldsymbol{x}_t\neq\boldsymbol{x}_0$, then $\boldsymbol{x}_{t-1}$ exists, and it is of the form $\boldsymbol{x}_{t-1}:=(\frac{1}{8},\ldots,\frac{1}{8},a,\frac{1}{8}-a,0,\ldots)$ for some $\frac{1}{16}\leqslant a<\frac{1}{8}$, where $\frac{1}{8}$ is repeated seven times, and moreover, the element of $\overline{\mathcal{P}}$ obtained by setting $a=\frac{1}{16}$, say, lies in OPT. A routine calculation shows this to be a contradiction (see `2111.nb`).

Thus, OPT $=\{\boldsymbol{a}\}$, where $\boldsymbol{a}:=(\frac{1}{8},\ldots,\frac{1}{8},0,\ldots)$ with $a_0=0$.

Finally, it remains to check strictness. Let us check (Str1). First let x,y be in different parts V_i,V_j of $G=G_{n,\boldsymbol{a}}$. Write $p=\frac{1}{8}$ and assume $8|n$. Consider copies of $K_{2,1,1,1}$ that contain both x

and y, with A denoting the two-element part. Then in G, the edge xy can be such that x lies in A ($pn \cdot \binom{6}{2} \cdot (pn)^2$) choices), y lies in A (the same), or neither x nor y lie in A ($6 \cdot \binom{pn}{2} \cdot 5pn$ choices). In $G \oplus xy$, xy is a non-edge and so we can only have $\{x, y\}$ playing the role of A in F, so the number of such copies of F is $\binom{6}{3}(pn)^3$. Thus, for distinct $i, j \in [8]$, we have

$$\nabla^{\bullet\bullet}_{ij}\lambda(\boldsymbol{a}) = 3!p^3\left(2\binom{6}{2} + 15 - \binom{6}{3}\right) = \frac{150}{512}.$$

Now let x, y be in the same part V_i of G. Then in G, the non-edge xy lies in $\binom{7}{3}(pn)^3$ copies of F. In $G \oplus xy$, the edge xy lies in $7\binom{pn}{2} \cdot 6pn$ copies of F. So,

$$\nabla^{\bullet\bullet}_{ii}\lambda(\boldsymbol{a}) = 3!p^3\left(\binom{7}{3} - \frac{7 \cdot 6}{2}\right) = \frac{84}{512},$$

as required.

For (Str2), let $b : [8] \to \{0, 1\}$ be such that $|\mathrm{supp}(b)| = k$. Then

$$\lambda(\boldsymbol{a}, (b, 1)) = \lim_{n\to\infty} \binom{n}{4}^{-1} \left(\frac{n}{8}\right)^4 \left((8-k)\binom{k}{3} + \binom{k}{2}(k-2) \cdot \frac{1}{2}\right) = \frac{4!}{8^4} \cdot \binom{k}{3}\left(\frac{19}{2} - k\right).$$

Indeed, counting induced copies of F in $G_{n,\boldsymbol{a}} +_{b,1} u$ containing u: if u plays the role of a vertex in A, then we choose the other vertex from this set among any of the $8 - k$ parts not adjacent to u, and then choose three distinct parts of the k adjacent to u to contain the other vertices. If u plays the role of a singleton, we choose two among k parts for the other two singletons, and another for A (dividing by two for both orders). Routine calculations show that this is uniquely maximised (with value λ_0) when $k = 7$, as required. ☐

Proof of Theorem 1.6. We will show that $\mathrm{OPT} = \{(\frac{3}{5}, 0, \ldots)\}$ and $\lambda_{\max} = \frac{216}{625}$. Let G be a complete partite graph on n vertices which maximises the number of $F := K_{3,1,1}$. Comparing G to $(\frac{3}{5}, 0, 0, \ldots)$, we have $P(F, G) \geqslant \frac{216}{625}\binom{n}{5} + O(n^4)$. Suppose that Y, Z are the two largest parts of G, with $|Y| = yn$, $|Z| = zn$ and $y \geqslant z$. Let $S := V(G) \setminus (Y \cup Z)$.

First, let us derive a contradiction from assuming that $z \geqslant \frac{2}{5}$. Let $s := 1 - y - z$, so $s \leqslant \frac{1}{5}$. The number of copies of F with at least three vertices in S is at most

$$n^5\left(\frac{s^5}{5!} + \frac{s^4(1-s)}{4!} + \frac{s^3}{3!} \cdot yz\right) \leqslant n^5\left(\frac{s^5}{5!} + \frac{s^4}{4!} + \frac{s^3}{4!}\right) \leqslant n^5\frac{151}{600 \cdot 625}.$$

The number of copies of F with exactly two vertices in S is at most

$$\binom{sn}{2}\left(\binom{yn}{3} + \binom{n - sn - yn}{3}\right) = \frac{n^5 s^2}{12}(y^3 + (1 - s - y)^3) + O(n^4). \tag{6.3}$$

We have $y \leqslant 1 - s - \frac{2}{5}$ (since $z \geqslant \frac{2}{5}$) and $y > 1 - s - y$ (since $y \geqslant z$). For fixed s, the expression $y^3 + (1 - s - y)^3$ is maximised when y is as large as possible. Indeed, $y^3 + (1 - s - y)^3$ for $y \in \mathbb{R}$ is a quadratic polynomial whose coefficient at y^2 is positive and whose minimum is at $\frac{1-s}{2}$, and we have $y > \frac{1-s}{2}$. So, the expression in (6.3) is at most

$$r(s)n^5 + O(n^4) \quad \text{where} \quad r(s) = \frac{s^2}{12}\left(\left(1 - s - \frac{2}{5}\right)^3 + \left(\frac{2}{5}\right)^3\right).$$

We claim that r' has no roots in $(0,\frac{1}{5}]$, which implies that $r(s)$ attains its maximum at $s=\frac{1}{5}$, of value $\frac{160}{600\cdot 625}$. Indeed, $r'(s)=\frac{s}{300}t(s)$ where $t(s)=-125s^3+180s^2-81s+14$. Furthermore, $t(1)<0<t(\frac{4}{5})$ so t has at least one root in $[\frac{4}{5},1]$. If the claim does not hold, then t has three real roots, which are interlaced by the roots of the quadratic $t'(s)=-3(5s-3)(25s-9)$. The smallest root of t' is $\frac{9}{25}>\frac{1}{5}$, and the coefficient of s^3 in t is negative, so t has a root in $(0,\frac{1}{5}]$ only if $t(\frac{9}{25})<0$, a contradiction.

Every other copy of F has exactly four vertices in $Y\cup Z$. So, writing $q:=\frac{1-s}{y}$, their number is

$$\binom{yn}{3}zn\cdot sn+\binom{zn}{3}\cdot yn\cdot sn=\frac{n^5}{6}(q^3(1-q)+(1-q)^3q)s(1-s)^4+O(n^4),$$

which, for $s\in[\frac{4}{5},1]$, is maximised when $(s,q)=(\frac{4}{5},\frac{1}{2})$, with value $\frac{640}{600}\cdot\frac{n^5}{625}+O(n^4)$. So, when $z\geqslant\frac{2}{5}$, we have $p(F,G)\leqslant\frac{5!}{600\cdot 625}(151+160+640)<\lambda_0$, and we obtain the desired contradiction.

Assume from now on that $z<2/5$. Fix $v\in Z$. Let $p(F,G,v)$ be the number of copies of F containing v. Then

$$P(F,G,v)\leqslant p(v):=\binom{zn-1}{2}\left(\binom{(1-z)n}{2}-\binom{yn}{2}\right)+\binom{yn}{3}(1-y-z)n$$
$$+\frac{1}{3}\sum_{w\in S}\binom{n-1-d(w)}{2}(d(w)-zn).$$

We would like a good upper bound for the last term. Since Z is the second largest part, we have that $(1-z)n\leqslant d(w)\leqslant n$ for all $w\in S$. Now $f(x)=\frac{1}{2}(1-x)^2(x-z)$ is maximised when $x=x_0:=\frac{1}{3}(1+2z)$ and is decreasing on the interval $[x_0,1]$. Since $z\leqslant\frac{2}{5}$, we have $x_0\leqslant 1-z$, so f defined in the range $[1-z,1]$ is maximised at $x=1-z$. So, the last term divided by n^4 is at most

$$\frac{1}{3}\sum_{w\in S}f(d(w))n^{-1}+O(1/n)=\frac{1}{3}(1-y-z)f(1-z)+O(1/n).$$

Define

$$h(y,z):=12\left(\frac{z^2}{4}((1-z)^2-y^2)+\frac{y^3}{6}(1-y-z)+\frac{1}{3}(1-y-z)f(1-z)-\frac{9}{625}\right)$$
$$=2y^3-2y^4-2y^3z+5z^2-2yz^2-3y^2z^2-12z^3+4yz^3+7z^4-\frac{108}{625}.$$

By the above, $h(y,z)\geqslant 12(p(v)n^{-4}+O(1/n)-\frac{9}{625})\geqslant O(1/n)$, that is, $h(y,z)\geqslant 0$ for all $0\leqslant z\leqslant y$ with $z+y\leqslant 1$ and $z\leqslant\frac{2}{5}$. Let

$$R:=\{(y,z)\in[0,1]^2\,:\,y\geqslant z,y+z\leqslant 1\}.$$

Claim 6.1. For every $(y,z)\in R$ with $h(y,z)\geqslant 0$, we have that $y\geqslant\frac{3}{5}$.

Suppose that the claim holds. Since G is optimal, Proposition 3.1 implies that v has optimal attachment in G; that is, $P(F,G,v)=\binom{n-1}{4}\lambda(G,v)=\binom{n-1}{4}\lambda_{\max}+O(n^3)\geqslant\frac{9}{625}n^4+O(n^3)$. Thus, $h(y,z)\geqslant 0$ for the y,z corresponding to Y,Z, since, as we have shown, $z\leqslant\frac{2}{5}$. So $y\geqslant\frac{3}{5}$. Consider the graph H obtained by replacing Z by a clique. Then we lose every copy of F containing the 3-independent set in Z (and lose no other copies), while we gain copies of F with the 3-independent set in Y and the two other vertices in Z. So,

$$\begin{aligned}\frac{P(F,G)-P(F,H)}{120n^5} &\leqslant \frac{z^3}{3!}\frac{(1-z)^2}{2}-\frac{(3/5)^3}{3!}\frac{z^2}{2}+o(1)=\frac{z^2}{2}\left(\frac{z(1-z)^2}{6}-\frac{3^3}{6\cdot 5^3}\right)+o(1)\\ &\leqslant \frac{z^2}{2}\left(\frac{4}{6\cdot 3^3}-\frac{3^3}{6\cdot 5^3}\right)+o(1)\leqslant \frac{-229}{40\,500}z^2+o(1).\end{aligned}$$

This is a contradiction to the optimality of G if $z=\Omega(1)$. Thus, $z=o(1)$ and, up to $o(n^2)$ edits, G consists of an independent set of size yn and $(1-y)n$ universal vertices. So $p(F,G)=120(\frac{y^3}{3!}\frac{(1-y)^2}{2})+o(1)$. Ignoring the error term, this is uniquely maximised when $y=\frac{3}{5}$, with value $\frac{216}{625}$. Then $\mathrm{OPT}=\{\boldsymbol{a}\}$, where $\boldsymbol{a}=(\frac{3}{5},0,\dots)$.

So, in order to determine OPT, it remains to prove Claim 6.1.

Proof of Claim 6.1. Firstly we consider (y,z) on the boundary of R. If $z=0$, then $h(y,0)=\frac{-108}{625}+1250y^3(1-y)$ which is uniquely maximised when $y=\frac{3}{5}$. If $y=z$, then $h(y,y)=y^2(2y-1)(2y-5)-\frac{108}{625}$ which is negative for $y\in[0,1]$.

Now we consider (y,z) in the interior of R. Let (y_0,z_0) in the interior of R be such that $h(y_0,z_0)\geqslant 0$ and y_0 is minimal with this property (such a y_0 exists by compactness of R and continuity of h). Since (y_0,z_0) is in the interior of R, we have $h(y_0,z_0)=0$ and $\frac{\partial h}{\partial z}(y_0,z_0)=0$ (otherwise we can find $z'\approx z_0$ with $h(y_0,z')>h(y_0,z_0)=0$ and by the continuity of h, $y'<y_0$ and $h(y',z')\geqslant 0$, contradicting the minimality of y_0). Applying Buchberger's algorithm to eliminate z, we obtain a degree-12 polynomial q such that y satisfies $h(y,z)=0=\frac{\partial h}{\partial z}(y,z)$ only if $q(y)=0$ (see `311.nb`):

$$\begin{aligned}q(y):=&-2\,500\,858\,044+14\,506\,020\,000y-18\,911\,610\,000y^2-85\,830\,803\,750y^3\\ &+545\,884\,288\,750y^4-1\,430\,659\,375\,000y^5+4\,001\,212\,109\,375y^6\\ &-12\,503\,827\,343\,750y^7+30\,477\,566\,015\,625y^8-54\,597\,656\,250\,000y^9\\ &+64\,171\,142\,578\,125y^{10}-42\,002\,929\,687\,500y^{11}+12\,102\,539\,062\,500y^{12}.\end{aligned}$$

Let $\alpha:=\frac{272}{1000}$ and $R':=\{(y,z)\in(0,1]^2\,:\,z\leqslant y\leqslant\alpha\}$. We claim that $p(y):=q(y+\alpha)$ is a positive polynomial. Then $q(y)>0$ for all $y\in R\setminus R'$, and hence, $(y_0,z_0)\in R'$. For this, it suffices to show that there are polynomials $r_1(y),r_2(y)$ with non-negative coefficients satisfying $p(y)r_1(y)=r_2(y)$. Once one fixes the degree d of r_1, this amounts to solving a linear program, where a_k is the kth coefficient of p and b_k is the kth (unknown) coefficient of r_1:

$$\begin{aligned}&\text{minimise} && \sum_{0\leqslant k\leqslant d} b_k\\ &\text{subject to} && \sum_{\substack{j+k=i:\\ 0\leqslant j\leqslant 12;\\ 0\leqslant k\leqslant d}} a_j b_k>0, \quad && i=0,1,\dots,d+12,\\ & && b_k>0, && k=0,1,\dots,d.\end{aligned}$$

In fact, we only need a feasible solution, not an optimal one, so the objective function can be anything. For degrees $d = 1, 2, \ldots$, we attempted this (using python) until we obtained a numerical solution for $d = 16$. The following degree-16 polynomial was obtained by multiplying this solution by a fairly large power of 10 and rounding.

$$\begin{aligned} r_1(y) = {} & 405\,631\,585\,336x^{16} + 291\,048\,000\,156x^{15} + 172\,228\,102\,580x^{14} + 76\,577\,243\,592x^{13} \\ & + 32\,501\,733\,953x^{12} + 13\,576\,227\,809x^{11} + 5\,344\,727\,909x^{10} + 1\,954\,537\,506x^{9} \\ & + 73\,709\,7269x^{8} + 264\,696\,828x^{7} + 90\,984\,085x^{6} + 30\,184\,081x^{5} + 10\,472\,958x^{4} \\ & + 3\,090\,485x^{3} + 1\,000\,538x^{2} + 206\,609x + 108\,298. \end{aligned}$$

Clearly, its coefficients are positive and one can check (see `311.nb`) that the degree-28 polynomial $p(y)r_1(y)$ also has positive coefficients, as required.

Suppose we can find non-negative polynomials $s_0, \ldots, s_3$ in y, z and positive $t \in \mathbb{Q}$ such that

$$-h(y,z) - t - zs_1 - (y-z)s_2 - (\alpha - y)s_3 = s_0,$$

where a polynomial $p \in \mathbb{R}[y,z]$ is non-negative if $p(y,z) \geqslant 0$ whenever $y, z \geqslant 0$. Then $-h(y,z) > 0$ on R'. This will complete the proof of the claim. Let $\underline{x} := (1, y, z, y^2, yz, z^2)^{\intercal}$. To ensure that the s_i are non-negative, it suffices to find positive semidefinite 6×6 matrices Q_i such that $s_i(y,z) = \underline{x}^{\intercal} Q_i \underline{x}$. For this, a sum-of-squares solver (we used the YALMIP Matlab toolbox [24, 25] with SeDuMi [37]) numerically maximises t such that the above equality holds; that is, we obtain $t' \approx 0.02$ and real matrices $Q'_0, \ldots, Q'_3$ such that $-h(y,z) - t' - zs'_1 - (y-z)s'_2 - (\alpha - y)s'_3 \approx s'_0$, where $s'_i = \underline{x}^{\intercal} Q'_i \underline{x}$. Now let Q_i be a (symmetric) rational approximation to Q'_i for $i \in [3]$ and let R_0 be a rational approximation to Q'_0. We obtain

$$R_0 = \begin{pmatrix} \frac{47\,560\,627}{605\,583\,685} & -\frac{27\,288\,737}{128\,683\,162} & -\frac{5\,823\,553}{403\,766\,228} & -\frac{22\,660\,833}{166\,625\,377} & \frac{64\,761\,638}{445\,638\,833} & -\frac{10\,092\,851}{42\,370\,543} \\ -\frac{27\,288\,737}{128\,683\,162} & \frac{412\,450\,960}{208\,083\,677} & -\frac{154\,126\,052}{222\,170\,865} & -\frac{123\,333\,398}{74\,059\,181} & -\frac{45\,208\,772}{76\,054\,353} & \frac{29\,997\,552}{77\,062\,243} \\ -\frac{5\,823\,553}{403\,766\,228} & -\frac{154\,126\,052}{222\,170\,865} & \frac{56\,961\,038}{76\,246\,587} & \frac{75\,134\,651}{68\,479\,911} & -\frac{114\,623\,437}{74\,768\,701} & \frac{68\,436\,686}{157\,424\,595} \\ -\frac{22\,660\,833}{166\,625\,377} & -\frac{123\,333\,398}{74\,059\,181} & \frac{75\,134\,651}{68\,479\,911} & \frac{231\,222\,579}{42\,9\,11\,653} & -\frac{33\,046\,138}{90\,840\,815} & -\frac{27\,557\,233}{25\,108\,228} \\ \frac{64\,761\,638}{445\,638\,833} & -\frac{45\,208\,772}{76\,054\,353} & -\frac{114\,623\,437}{74\,768\,701} & -\frac{33046\,138}{90\,840\,815} & \frac{142\,375\,474}{17\,195\,129} & -\frac{204\,334\,483}{99\,244\,906} \\ -\frac{10\,092\,851}{42\,370\,543} & \frac{29\,997\,552}{77\,062\,243} & \frac{68\,436\,686}{157\,424\,595} & -\frac{27\,557\,233}{25\,108\,228} & -\frac{204\,334\,483}{99\,244\,906} & \frac{152\,251\,273}{45\,491\,357} \end{pmatrix} \succcurlyeq 0$$

$$Q_1 = \begin{pmatrix} \frac{113\,823\,133}{103\,564\,772} & -\frac{153\,720\,698}{116\,964\,597} & -\frac{514\,694\,857}{175\,951\,034} & -\frac{26\,958\,123}{134\,065\,612} & -\frac{5\,214\,837}{679\,601\,578} & \frac{424\,549\,711}{451\,760\,648} \\ -\frac{153\,720\,698}{116\,964\,597} & \frac{98\,271\,451}{22\,705\,510} & \frac{108\,839\,271}{102\,671\,668} & -\frac{37\,652\,132}{76\,331\,505} & -\frac{98\,556\,781}{98\,719\,039} & -\frac{86\,545\,565}{156\,277\,133} \\ -\frac{514\,694\,857}{175\,951\,034} & \frac{108\,839\,271}{102\,671\,668} & \frac{178\,543\,136}{16\,280\,101} & -\frac{13\,588\,975}{554\,452\,603} & -\frac{66\,382\,289}{197\,496\,474} & -\frac{31\,5010\,733}{72\,953\,806} \\ -\frac{26\,958\,123}{134\,065\,612} & -\frac{37\,652\,132}{76\,331\,505} & -\frac{13\,588\,975}{55\,4452\,603} & \frac{127\,914\,572}{23\,010\,911} & -\frac{93\,779\,957}{771\,873\,704} & -\frac{258\,311\,971}{316\,622\,401} \\ -\frac{5\,214\,837}{679\,601\,578} & -\frac{98\,556\,781}{9\,8719\,039} & -\frac{66\,382\,289}{197\,496\,474} & -\frac{93\,779\,957}{771\,873\,704} & \frac{183\,401\,329}{33\,290\,110} & -\frac{60\,904\,303}{161\,208\,591} \\ \frac{424\,549\,711}{451\,760\,648} & -\frac{86\,545\,565}{156\,277\,133} & -\frac{315\,010\,733}{72\,953\,806} & -\frac{258\,311\,971}{316\,622\,401} & -\frac{60\,904\,303}{161\,208\,591} & \frac{502\,508\,117}{78\,490\,640} \end{pmatrix} \succcurlyeq 0$$

$$Q_2 = \begin{pmatrix} \frac{21\,520\,940}{25\,577\,879} & -\frac{56\,020\,343}{32\,074\,003} & \frac{4\,0731\,578}{75\,1516\,279} & -\frac{46\,544\,963}{139\,367\,268} & -\frac{41\,177\,990}{108\,764\,983} & -\frac{26\,606\,007}{46\,612\,636} \\ -\frac{56\,020\,343}{32\,074\,003} & \frac{112\,841\,678}{19\,842\,961} & -\frac{139\,240\,153}{172\,670\,104} & -\frac{45\,501\,317}{43\,903\,809} & \frac{64\,055\,491}{88\,725\,341} & \frac{21\,288\,583}{30\,121\,110} \\ \frac{40\,731\,578}{75\,1516\,279} & -\frac{139\,240\,153}{172\,670\,104} & \frac{68\,362\,401}{21\,097\,442} & -\frac{168\,386\,141}{819\,717\,774} & -\frac{155\,286\,027}{198\,655\,888} & \frac{30\,506\,956}{19\,158\,511} \\ -\frac{46\,544\,963}{139\,367\,268} & -\frac{45\,501\,317}{43\,903\,809} & -\frac{168\,386\,141}{819\,717\,774} & \frac{166\,235\,485}{28\,138\,938} & \frac{15\,677\,552}{218\,059\,291} & -\frac{15\,992\,364}{2\,5871\,383} \\ -\frac{41\,177\,990}{108\,764\,983} & \frac{64\,055\,491}{88\,725\,341} & -\frac{155\,286\,027}{198\,655\,888} & \frac{15\,677\,552}{218\,059\,291} & \frac{253\,525\,900}{46\,511\,459} & \frac{95\,613\,053}{837\,681\,775} \\ -\frac{26\,606\,007}{46\,612\,636} & \frac{21\,288\,583}{30\,121\,110} & \frac{30\,506\,956}{19\,158\,511} & -\frac{15\,992\,364}{25\,871\,383} & \frac{95\,613\,053}{837\,681\,775} & \frac{267\,687\,310}{41\,812\,157} \end{pmatrix} \succcurlyeq 0$$

$$Q_3 = \begin{pmatrix} \frac{29\,877\,454}{113\,194\,375} & -\frac{110\,018\,062}{390\,364\,861} & -\frac{39\,492\,021}{93\,889\,856} & -\frac{44\,736\,353}{260\,223\,501} & -\frac{27\,286\,543}{148\,218\,452} & -\frac{211\,317\,628}{549\,271\,497} \\ -\frac{110\,018\,062}{390\,364\,861} & \frac{168\,343\,502}{34\,876\,437} & -\frac{63\,781\,869}{56\,201\,314} & \frac{813\,722\,845}{556\,876\,698} & \frac{1\,719\,950}{4\,084\,346\,189} & \frac{20\,149\,420}{711\,586\,093} \\ -\frac{39\,492\,021}{93\,889\,856} & -\frac{63\,781\,869}{56\,201\,314} & \frac{293\,980\,380}{89\,098\,241} & -\frac{16\,659\,683}{50\,114\,131} & \frac{24\,295\,714}{57\,792\,167} & \frac{20\,062\,513}{28\,511\,329} \\ -\frac{44\,736\,353}{260\,223\,501} & \frac{813\,722\,845}{556\,876\,698} & -\frac{16\,659\,683}{50\,114\,131} & \frac{166\,235\,485}{28\,138\,938} & -\frac{91\,733\,513}{894\,919\,007} & -\frac{11\,949\,058}{15\,299\,253} \\ -\frac{27\,286\,543}{148\,218\,452} & \frac{1\,719\,950}{4\,084\,346\,189} & \frac{24\,295\,714}{57\,792\,167} & -\frac{91\,733\,513}{894\,919\,007} & \frac{187\,073\,509}{34\,708\,874} & -\frac{1\,990\,762}{36\,615\,949} \\ -\frac{211\,317\,628}{549\,271\,497} & \frac{20\,149\,420}{711\,586\,093} & \frac{20\,062\,513}{28\,511\,329} & -\frac{11\,949\,058}{15\,299\,253} & -\frac{1\,990\,762}{36\,615\,949} & \frac{192\,697\,280}{35\,564\,393} \end{pmatrix} \succcurlyeq 0.$$

At this stage, it does not matter (for the purposes of a verifiable proof) where R_0, Q_1, Q_2, Q_3 came from; it suffices to show that they are positive semidefinite and that the polynomial

$$\varepsilon(y,z) := -h(y,z) - zs_1 - (y-z)s_2 - (\alpha - y)s_3 - r_0$$

is positive on $[0,1]^2$, where $r_0 = \underline{x}^{\intercal} R_0 \underline{x}$. To check positive semi-definiteness of a matrix $A = (a_{ij})_{i,j\in[m]}$, we first check that A is symmetric, then we use Sylvester's criterion, which says that a Hermitian matrix A is positive semi-definite if and only if $A^{(k)} = (a_{ij})_{i,j\in[k]}$ has positive determinant for all $k \in [m]$. We bound $\varepsilon(y,z)$ from below by its constant term minus the sum of the absolute value of its other coefficients (see `311.nb`) to see that $\varepsilon(y,z) \geqslant \frac{1}{50}$ in the required region. This completes the proof of the claim. ☐

Since Claim 6.1 implies that $\mathrm{OPT} = \{(\frac{3}{5}, 0, \ldots)\}$, it remains to check that $p(K_{3,1,1}, \cdot)$ is strict. Consider $G = G_{n,\boldsymbol{a}}$ which has a clique part V_0 of size $\frac{2n}{5} + O(1)$ and another part V_1 of size $\frac{3n}{5} + O(1)$ which is an independent set. Now (Str1) is immediate as $G \oplus xy$ has no induced copy of F containing both x and y.

Now we check (Str2). Let $c := \frac{108}{125}$. We have $\mathrm{supp}^*(\boldsymbol{a}) = \{0,1\}$, so given any $b : \{1\} \to \{0,1\}$ and $\alpha \in [0,1]$ it is enough to show that

$$\nabla^{\bullet}_{b,\alpha}\lambda(\boldsymbol{a}) = \lambda(\boldsymbol{a}) - \lambda(\boldsymbol{a},(b,\alpha)) = \lambda(\boldsymbol{a}) - \lim_{n\to\infty} \lambda(G_{n,\boldsymbol{a}} +_{b,\alpha} u, u) \geqslant \frac{2}{5}c(1-\alpha) = \lambda_{\max}(1-\alpha),$$

that is, $\lambda(\boldsymbol{a},(b,\alpha)) \leqslant \lambda_{\max}\alpha$. If $b(1) = 0$, then u lies in a copy of $K_{3,1,1}$ only if it lies in the 3-set with two vertices in V_1 and the two singletons are in $N(u) \cap V_0$, so $\lambda(\boldsymbol{a},(b,\alpha)) = \binom{4}{2,2}(\frac{2\alpha}{5})^2(\frac{3}{5})^2 = \lambda_{\max}\alpha^2$, as required. If $b(1) = 1$, then u lies in a copy of $K_{3,1,1}$ only if the 3-set is in V_1 and the other singleton is in $N(u) \cap V_0$, so $\lambda(\boldsymbol{a},(b,\alpha)) = \binom{4}{1,3}(\frac{2\alpha}{5})(\frac{3}{5})^3 = \lambda_{\max}\alpha$, as required.

This completes the proof of the theorem. ☐

7 | CONCLUDING REMARKS

In this paper, we have shown how to obtain stability from results in extremal graph theory which use symmetrisation. We have applied our general theory to the inducibility problem for complete partite graphs. It would be interesting to solve other instances of the polynomial optimisation problem which amounts to determining $i(F)$.

It would be particularly interesting to find other extremal graph theory problems to which our theory applies.

ACKNOWLEDGEMENTS

The first author was supported by the UK Research and Innovation Future Leaders Fellowship MR/S016325/1 and the Institute for Basic Science Grant IBS-R029-C4. The second author was supported by ERC Advanced Grant 101020255 and Leverhulme Research Project Grant RPG-2018-424. The fourth author was supported by Engineering and Physical Sciences Research Council Fellowship EP/V025953/1.

JOURNAL INFORMATION

The *Journal of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. J. Balogh, P. Hu, B. Lidický, and F. Pfender, *Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle*, Europ. J. Combin. **52** (2016), 47–58.
2. P. Bennett, A. Dudek, B. Lidický, and O. Pikhurko, *Minimizing the number of 5-cycles in graphs with given edge-density*, Combin. Probab. Comput. **29** (2020), 44–67.
3. B. Bollobás, *On complete subgraphs of different orders*, Math. Proc. Cambridge Philos. Soc. **79** (1976), 19–24.
4. B. Bollobás, Y. Egawa, A. Harris, and G. Jin, *The maximal number of induced r-partite subgraphs*, Graphs Combin. **11** (1995), 1–19.
5. B. Bollobás, C. Nara, and S. Tachibana, *The maximal number of induced complete bipartite graphs*, Discrete Math. **62** (1986), no. 3, 271–275.
6. Ł. Bożyk, A. Grzesik and B. Kielak, *On the inducibility of oriented graphs on four vertices*, Discrete Math. **345** (2022), no. 20, 112874.
7. J. I. Brown and A. Sidorenko, *The inducibility of complete bipartite graphs*, J. Graph Theory **18** (1994), 629–645.
8. D. Burke, B. Lidický, F. Pfender, and M. Phillips, *Inducibility of 4-vertex tournaments*, Manuscript.
9. I. Choi, B. Lidický, and F. Pfender, *Inducibility of directed paths*, Discrete Math. **343** (2020), no. 10, 112015.
10. C. Even-Zohar and N. Linial, *A note on the inducibility of 4-vertex graphs*, Graphs Combin. **31** (2015), 1367–1380.
11. J. Fox, H. Huang, and C. Lee, *A solution to the inducibility problem for almost all graphs*, Manuscript.
12. J. Fox, L. Sauermann, and F. Wei, *On the inducibility problem for random Cayley graphs of abelian groups with a few deleted vertices*, Random Struct. Alg. **59** (2021), 554–615.
13. Z. Füredi, *A proof of the stability of extremal graphs, Simonovits' stability from Szemerédi's regularity*, J. Combin. Theory (B) **115** (2015), 66–71.
14. R. Glebov, A. Grzesik, T. Klimošová, and D. Král', *Finitely forcible graphons and permutons*, J. Combin. Theory (B) **110** (2015), 112–135.
15. O. Goldreich, M. Krivelevich, I. Newman, and E. Rozenberg, *Hierarchy theorems for property testing*, Technical Report TR08-097, Electronic Colloquium on Computational Complexity, 2008.

16. H. Hatami, J. Hirst, and S. Norin, *The inducibility of blow-up graphs*, J. Combin. Theory Ser. B **109** (2014), 196–212.
17. D. Hefetz and M. Tyomkyn, *On the inducibility of cycles*, J. Combin. Theory Ser. B **133** (2018), 243–258.
18. J. Hirst, *The inducibility of graphs on four vertices*, J. Graph Theory **75** (2014), 231–243.
19. P. Hu, B. Lidický, F. Pfender, and J. Volec, *Inducibility of orientations of* C_4, Manuscript.
20. P. Hu, J. Ma, S. Norin, and H. Wu, *Inducibility of oriented stars*, arXiv:2008.05430.
21. H. Huang, *On the maximum induced density of directed stars and related problems*, SIAM J. Discrete Math. **28** (2014), no. 1, 92–98.
22. D. Král', S. Norin, and J. Volec, *A bound on the inducibility of cycles*, J. Combin. Theory Ser. A **161** (2019), 359–363.
23. H. Liu, O. Pikhurko, M. Sharifzadeh, and K. Staden, *Stability from symmetrisation arguments*, E-print arxiv:2012.10731, 2020.
24. J. Löfberg, *Yalmip: a toolbox for modeling and optimization in Matlab*, Proceedings of the CACSD Conference, Taipei, Taiwan, 2004.
25. J. Löfberg, *Pre- and post-processing sum-of-squares programs in practice*, IEEE Trans. Automat. Control **54** (2009), no. 5, 1007–1011.
26. L. Lovász, *Large networks and graph limits*, American Mathematical Society, vol. 60 of Colloquium Publications, 2012.
27. L. Lovász and B. Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957.
28. T. S. Motzkin and E. G. Straus, *Maxima for graphs and a new proof of a theorem of Turán*, Canad. J. Math. **17** (1965), 533–540.
29. X. Liu, D. Mubayi, and C. Reiher, *The feasible region of induced graphs*, J. Combin. Theory (B) **158** (2023), 105–135.
30. S. Norin and L. Yepremyan, *Turán number of generalized triangles*, J. Combin. Theory (A) **146** (2017), 312–343.
31. S. Norin and L. Yepremyan, *Turán numbers of extensions*, J. Combin. Theory (A) **155** (2018), 476–492.
32. O. Pikhurko, J. Sliacan, and K. Tyros, *Strong forms of stability from flag algebra calculations*, J. Combin. Theory (B) **135** (2019), 129–178.
33. O. Pikhurko, *An analytic approach to stability*, Discrete Math. **310** (2010), no. 21, 2951–2964.
34. N. Pippenger and M. C. Golumbic, *The inducibility of graphs*, J. Combin. Theory (B) **19** (1975), 189–203.
35. A. Roberts and A. Scott, *Stability results for graphs with a critical edge*, Europ. J. Combin. **74** (2018), 27–38.
36. R. H. Schelp and A. G. Thomason, *A remark on the number of complete and empty subgraphs*, Combin. Probab. Comput. **7** (1998), 217–220.
37. J. F. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optim. Methods Softw. **11–12** (1999), 625–653. Version 1.05 available from http://fewcal.kub.nl/sturm.
38. P. Turán, *On an external problem in graph theory*, Mat. Fiz. Lapok **48** (1941), 436–452.
39. R. Yuster, *On the exact maximum induced density of almost all graphs and their inducibility*, J. Combin. Theory Ser. B **136** (2019), 81–109.
40. A. A. Zykov, *On some properties of linear complexes*, Matematičeskiĭ Sbornik N.S. **24** (1952), no. 66, 163–188.