When is a Graphical Sequence Stable?

Mark Jerrum†, Brendan McKay‡, and Alistair Sinclair†

September 1989

ABSTRACT The function which maps each graphical sequence $\mathbf{d}$ to the number of graphs with degree sequence $\mathbf{d}$ is considered, with particular attention being directed at the stability of the function under small perturbations in $\mathbf{d}$. In some parts of its domain this function varies smoothly, and in other parts erratically. The boundary between these two behaviours is here sharply characterised in terms of the minimum, maximum, and average of the components of $\mathbf{d}$. The result clarifies the range of applicability of some efficient randomised algorithms which sample and count degree-constrained graphs. Furthermore, the result appears to set theoretical limits on the range of validity of asymptotic formulas for the number of graphs with given degree sequence.

---

† Department of Computer Science, University of Edinburgh, The King's Buildings, Edinburgh EH9 3JZ, UK.

‡ Computer Science Department, Australian National University, GPO Box 4, ACT 2601, Australia. This author was partially supported by SERC Visiting Fellowship Research Grant GR/F 18169

1. Introduction

Suppose $\mathbf{d}$ is a sequence of n non-negative integers. Denote by $\mathcal{G}(\mathbf{d})$ the set of labelled graphs with degree sequence $\mathbf{d}$, and let $\mathcal{N}(\mathbf{d}) = |\mathcal{G}(\mathbf{d})|$. The set $\mathcal{G}(\mathbf{d})$ has been the object of extensive study, two particular concerns being the development of techniques for random uniform generation of members of $\mathcal{G}(\mathbf{d})$ [9], and the derivation of asymptotic estimates for $\mathcal{N}(\mathbf{d})$ [1,2,5,6,7]. All the known results in this area require certain restrictions to be placed on the sequence $\mathbf{d}$; it is not enough for the sequence $\mathbf{d}$ merely to be graphical (i.e., to satisfy $\mathcal{N}(\mathbf{d}) > 0$). In this note, we suggest that these restrictions may not represent a deficiency in the known results. Rather, they may be a consequence of the fact that $\mathcal{N}(\mathbf{d})$, as a function of $\mathbf{d}$, is ill behaved over parts of its domain. It is only for those parts of the domain where $\mathcal{N}(\mathbf{d})$ is well behaved that results—analytic or algorithmic—have been obtained.

Our approach is to study how the function $\mathcal{N}(\mathbf{d})$ responds to small perburbations in its argument $\mathbf{d}$. The smallest and hence most natural perturbations to consider are the ones which modify just two components of $\mathbf{d}$, decrementing the first and incrementing the second, in both cases by one. Consider, for example, the sequence $\mathbf{d} = (k, k, \ldots, k)$, where $2 \le k \le n-2$. (Thus $\mathcal{G}(\mathbf{d})$ is the set of all graphs on n vertices which are regular of degree k.) Imagine that $\mathbf{d}$ is perburbed to yield a modified sequence $\mathbf{d}' = (k-1, k, k, \ldots, k, k, k+1)$. Intuitively, this perturbation does not change the value of the function $\mathcal{N}$ by very much, $\mathcal{N}(\mathbf{d}')$ remaining 'close' to $\mathcal{N}(\mathbf{d})$. This intuition is indeed correct. However, a contrasting behaviour is exhibited by the n-element sequence

$$\mathbf{d} = (1, 2, \ldots, k-1, k, k, k+1, \ldots, 2k-2, 2k-1)$$

(where $n = 2k$) and the perturbed sequence $\mathbf{d}'$ derived from $\mathbf{d}$ by incrementing the first component and decrementing the final component. In this instance, $\mathcal{N}(\mathbf{d}')$ is exponential in n, even though $\mathcal{N}(\mathbf{d}) = 1$.

We say that a class of sequences is *P-stable* [3] if the function $\mathcal{N}(\mathbf{d})$ varies smoothly (in a certain well defined sense to be explained later) as $\mathbf{d}$ ranges over the whole of that class. The main contribution of this note is a very sharp characterisation of P-stability in terms of δ and Δ, the minimum and maximum components of a sequence. We show, for example, that the class of sequences with $\delta \ge 1$ and $\Delta \le 2\sqrt{n} - 2$ is P-stable, as is the class with $\delta \ge \frac{1}{4}n$ and $\Delta \le \frac{3}{4}n - 1$. (These examples represent just two points in a general trade-off between δ and Δ.) The quoted bounds on δ and Δ are essentially the best possible. In the final section of the paper, using a more refined analysis, we obtain an improved characterisation of P-stability which also involves the *average* degree of the sequence.

One motivation for the study of P-stability has already be mentioned: it provides insight into why asymptotic formulas for $\mathcal{N}(\mathbf{d})$ have been obtained only for a restricted class of sequences $\mathbf{d}$. It can now be seen, for example, that the range of validity of the formula of McKay and Wormald [5] could not be much extended without venturing into areas where the function $\mathcal{N}$ is ill-behaved. Another motivation is more positive in tone. Certain algorithmic questions concerning $\mathcal{G}(\mathbf{d})$ are known to have an efficient solution when the input $\mathbf{d}$ is restricted to a P-stable set [3]. For example, the P-stability assumption allows us, in time polynomial in n, to sample from the set $\mathcal{G}(\mathbf{d})$ with a probability distribution which is arbitrarily close to uniform. It also makes it possible, again in polynomial time, to estimate the quantity $\mathcal{N}(\mathbf{d})$ to any specified degree of accuracy. Good characterisations of P-stability yield good bounds on the range of validity of these algorithms.

In summary, P-stable classes of sequences are easy to handle: information about $\mathcal{G}(\mathbf{d})$ and $\mathcal{N}(\mathbf{d})$ can readily be obtained via polynomial-time algorithms and asymptotic formulas. In contrast, classes which are *not* P-stable may be inherently difficult to handle, simply because the function $\mathcal{N}(\mathbf{d})$ is ill behaved over such a class.

2. P-stability: definition and characterisation

All graphs will be undirected, with no loops or parallel edges. Let G be a graph with vertex set $V = \{1, \ldots, n\}$ and edge set E. The *degree*, d_i, of vertex $i \in V$ is the total number of edges incident at i; the sequence $(d_1, \ldots, d_n)$ is the *degree sequence* of G. For any finite sequence $\mathbf{d}$ of non-negative integers, denote by $\mathcal{G}(\mathbf{d})$ the set of all graphs with degree sequence $\mathbf{d}$. A sequence $\mathbf{d}$ is said to be *graphical* if $|\mathcal{G}(\mathbf{d})| > 0$.

We now proceed to define P-stability, the central concern of this paper. Let $\mathcal{G}'(\mathbf{d})$ be the union $\bigcup_{\mathbf{d}'} \mathcal{G}(\mathbf{d}')$, where $\mathbf{d}'$ ranges over all sequences $\mathbf{d}' = (d'_1, \dots, d'_n)$ which satisfy $\sum_{i=1}^n d_i = \sum_{i=1}^n d'_i$ and $\sum_{i=1}^n |d_i - d'_i| = 2$. (Informally, $\mathcal{G}'(\mathbf{d})$ is the set of labelled graphs whose degree sequences are very close to $\mathbf{d}$.) We say that a class of sequences is *P-stable* if there exists a polynomial p such that $|\mathcal{G}'(\mathbf{d})| \leq p(n)|\mathcal{G}(\mathbf{d})|$ for every sequence $\mathbf{d} = (d_1, \dots, d_n)$ in the class. Informally, a class of sequences is P-stable if the function $\mathcal{N}(\mathbf{d}) = |\mathcal{G}(\mathbf{d})|$ varies smoothly as $\mathbf{d}$ ranges over the whole of the class†.

If $u, v \in V$ are distinct vertices with $\{u, v\} \notin E$, then u and v are *co-adjacent* and the pair $\{u, v\}$ is a *co-edge* in G. Define an *alternating path* (of length l) in G to be a sequence of vertices $v_0, v_1, \dots, v_l$ such that $\{v_i, v_{i+1}\}$ is an edge when i is even, and a co-edge when i is odd; the path is said to be *edge-disjoint* if the edges and co-edges which compose it are all distinct.

Our main result is a sharp characterisation of P-stability in terms of the minimum and maximum components of the sequences which compose a class. In preparation for the main result we introduce the following technical lemma:

Lemma 1. Let G be an n-vertex graph with distinguished vertices s and t (not necessarily distinct), and suppose that the set of vertices adjacent to s is equal to the set of vertices adjacent to t. Suppose also that δ and Δ are natural numbers such that the degrees of all vertices in G other than s and t lie in the range $[\delta, \Delta]$, and such that s and t themselves have degree at least $\delta + 1$. If $(\Delta - \delta + 1)^2 \leq 4\delta(n - \Delta - 1)$, then there exists an edge-disjoint alternating path in G which starts at s, ends at t, and has length 1, 3, 5, or 7.

We move directly to the main result, deferring the proof of Lemma 1 until the end of the section.

Theorem 2. Let $\mathcal{F}$ be the class of all graphical sequences $\mathbf{d} = (d_1, \dots, d_n)$ satisfying $(\Delta - \delta + 1)^2 \leq 4\delta(n - \Delta - 1)$, where δ and Δ are the minimum and maximum components of $\mathbf{d}$. The class $\mathcal{F}$ is P-stable.

Proof. Suppose $\mathbf{d} = (d_1, \dots, d_n)$ is a sequence satisfying the conditions of the theorem. We show that $|\mathcal{G}'(\mathbf{d})|/|\mathcal{G}(\mathbf{d})| \leq n^{10}$. Let G be an element of $\mathcal{G}'(\mathbf{d})$, i.e., a graph whose degree sequence agrees with $\mathbf{d}$ except at two vertices s and t, the degree of s being one greater, and the degree of t one smaller than those specified by the sequence $\mathbf{d}$. We shall convert G to a graph with degree sequence *precisely* $\mathbf{d}$ by a short sequence of edge insertions/deletions.

Since the vertex t has a degree deficit, it must be co-adjacent to some vertex, say t'. Adding the edge $\{t, t'\}$ to the graph G yields a graph whose degree sequence agrees with $\mathbf{d}$ except at vertices s and t', which between them have a degree surplus of 2. If s and t' are the same vertex, or if they are distinct but adjacent to the same set of vertices, then we are in the situation of Lemma 1. The lemma guarantees the existence of an odd-length alternating path, starting at s and ending at t', which has length at most 7. Dualising along the path (changing edges to co-edges and vice versa) yields a graph with degree sequence *exactly* $\mathbf{d}$.

It remains to deal with the case when s and t' are distinct, and are adjacent to different sets of vertices. In this case, there must exist an alternating path of length 2 which joins s and t'. (The path may start at s and end at t' or vice versa.) Dualising along the path places us in the situation of Lemma 1.

The procedure described above associates each element of $\mathcal{G}'(\mathbf{d})$ with an element of $\mathcal{G}(\mathbf{d})$. Moreover, it can be verified that at most n^{10} elements of $\mathcal{G}'(\mathbf{d})$ may be associated with any particular element of $\mathcal{G}(\mathbf{d})$. (Roughly, this is because the inverse procedure which associates each element of $\mathcal{G}(\mathbf{d})$ with a subset of $\mathcal{G}'(\mathbf{d})$ can be viewed as a sequence of 10 nondeterministic choices, each involving the selection of one of the n vertices.) □

Remark. The condition $(\Delta - \delta + 1)^2 \leq 4\delta(n - \Delta - 1)$ in the statement of the theorem is invariant under complementation; i.e., a sequence $(d_1, \dots, d_n)$ satisfies the condition if and only if the complementary sequence $(n - d_1 - 1, \dots, n - d_n - 1)$ does. □

---

† This definition is not formally identically with the original one appearing in [3], but is equivalent and perhaps more natural.

In order to gain a feel for the range of sequences covered by Theorem 2, it is perhaps worth listing some special cases. Denote the minimum and maximum components of a sequence $\mathbf{d}$ by $\delta(\mathbf{d})$ and $\Delta(\mathbf{d})$, respectively. Then it is immediate from Theorem 2 that the following classes of sequences are P-stable:

(i) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) \geq 1$ and $\Delta(\mathbf{d}) \leq 2\sqrt{n} - 2$;

(ii) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) \geq n^{1/3}$ and $\Delta(\mathbf{d}) \leq 2n^{2/3} - n^{1/3} - 1$;

(iii) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) \geq \frac{1}{9}n$ and $\Delta(\mathbf{d}) \leq \frac{5}{9}n - 1$;

(iv) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) \geq \frac{1}{4}n$ and $\Delta(\mathbf{d}) \leq \frac{3}{4}n - 1$;

(v) all regular sequences.

It was observed in the Introduction that P-stable classes of degree sequences are algorithmically tractable, while more general classes appear not to be so. In order to make that claim precise we require some further definitions. For non-negative real numbers a, $\tilde{a}$, ϵ, we say that $\tilde{a}$ *approximates* a *within ratio* $1+\epsilon$ if $a(1+\epsilon)^{-1} \leq \tilde{a} \leq a(1+\epsilon)$. A *generation problem* is defined by a function, $\mathcal{S}$, which maps a set of *instances* to a set of possible *solutions*. For example, the function $\mathcal{S}$ may have as its domain the set of all undirected graphs, and may associate with each graph Γ the set of all perfect matchings in Γ. We shall assume that each problem instance x is has a well-defined *size*. An *almost uniform generator* for $\mathcal{S}$ is a probabilistic algorithm which, given an instance x and a positive real *bias* ϵ, outputs an element of $\mathcal{S}(x)$ such that the probability of each element appearing approximates $|\mathcal{S}(x)|^{-1}$ within ratio $1+\epsilon$. The generator is *fully polynomial* if its execution time is bounded by a polynomial in $\lg \epsilon^{-1}$ and the size of x. (For a fuller treatment of almost uniform generation see [4,8].) It is known [3, Thm 2] that a fully polynomial almost uniform generator for $\mathcal{G}(\mathbf{d})$ exists, provided the input $\mathbf{d}$ is restricted to a P-stable class. Combining this result with Theorem 2 above, we immediately have:

Corollary 3. Let $\mathcal{F}$ be the class of sequences described in Theorem 2. There exists a fully polynomial almost uniform generator for the set $\mathcal{G}(\mathbf{d})$, which is valid for all sequences $\mathbf{d} \in \mathcal{F}$. □

Efficient estimation of the *cardinality* of $\mathcal{G}(\mathbf{d})$ is also possible under similar conditions on $\mathbf{d}$. More precisely, let f be a function from problem instances to natural numbers. An *approximation scheme* for f is a probabilistic algorithm which, when presented with an instance x and a real number $\epsilon > 0$, outputs a number which, with probability at least $\frac{3}{4}$, approximates $f(x)$ within ratio $(1+\epsilon)$. The approximation scheme is *fully polynomial* if it runs in time polynomial in ϵ^{-1} and the size of x. It is known [3, Thm 4] that a fully polynomial randomised approximation scheme for $\mathcal{G}(\mathbf{d})$ exists, provided the input $\mathbf{d}$ is restricted to a P-stable class. Combining this result with Theorem 2 above, we immediately have:

Corollary 4. Let $\mathcal{F}$ be the class of sequences described in Theorem 2. There exists a fully polynomial randomised approximation scheme for the function $\mathcal{N}(\mathbf{d}) = |\mathcal{G}(\mathbf{d})|$, which is valid for all sequences $\mathbf{d} \in \mathcal{F}$. □

We conclude the section by presenting the deferred proof of Lemma 1. In the next section we shall study classes of sequences which *fail* to be P-stable. It will become apparent that the characterisation of Theorem 2 is sharp.

Proof of Lemma 1. Suppose, to the contrary, that no such alternating path exists. Define subsets X, Y, and Z of the vertex set V of G as follows:

- X is the set of all vertices which are adjacent to s (equivalently, the set of all vertices adjacent to t).
- Y is the set of all vertices which are co-adjacent to at least 2 vertices in X. (Alternatively, Y is the set of all vertices which are reachable from s or t via at least two distinct alternating paths of length 2.)
- Z is the set of all vertices which are adjacent to some vertex in Y, but not adjacent to s or t.

The absence of odd-length edge-disjoint alternating paths starting at s and ending at t has a number of important consequences for the sets X, Y, and Z:

(i) The set X is a clique (otherwise there would exist an alternating path of length 3 with the form $s \to X \to X \to t$).

(ii) Y is an independent set (otherwise there would exist an alternating path of length 5 with the form $s \to X \to Y \to Y \to X \to t$).

(iii) The set Z is a clique (otherwise there would exist an alternating path of length 7 with the form $s \to X \to Y \to Z \to Z \to Y \to X \to t$).

(iv) The bipartite graph induced by X and Z is complete (otherwise there would exist an alternating path of length 5 with the form $s \to X \to Y \to Z \to X \to t$).

(v) The sets $\{s,t\}$, X, Y, and Z are pairwise disjoint.

(Note that the assumption that each vertex in Y is reachable from s via *at least two* alternating paths of length 2 is crucial here: it ensures that the alternating paths referred to in observations (ii)–(iv) can be chosen to be edge-disjoint.) Let $K = X \cup Z$ and $R = V - K - Y$, and denote the cardinalities of the sets K, R, and Y by k, r, and y, respectively. Note that $k + r + y = n$, since $K \cup R \cup Y$ is a partition of V; note also, by observations (i), (iii), and (iv), that K is a clique.

For A, B disjoint subsets of V, denote by $E(A,B)$ the set of all edges in G which have one endpoint in A and one in B. It follows from the definitions of the sets X, Y, and Z that no vertex in Y can be adjacent to a vertex in R. From this fact, and observation (ii), we deduce the inequality

$$|E(Y,K)| \geq y\delta. \tag{1}$$

From the definition of Y we know that each vertex in R is co-adjacent to at most one vertex in X; indeed at least one vertex in R, namely s, is adjacent to *all* of X. Moreover, the conditions of the lemma ensure $|X| \geq \delta + 1$. (Observe, in this context, that s and t, if distinct, must be co-adjacent.) Thus we have the inequality

$$|E(R,K)| \geq |E(R,X)| > r(|X| - 1) \geq r\delta. \tag{2}$$

Now, since K is a clique,

$$|E(K,R)| + |E(K,Y)| = |E(K, R \cup Y)| \leq k(\Delta - k + 1).$$

Combining the latter inequality with inequalities (1) and (2) yields $r\delta + y\delta < k(\Delta - k + 1)$ or, equivalently, since $k + r + y = n$:

$$k^2 - (\Delta + \delta + 1)k + n\delta < 0. \tag{3}$$

But the discriminant of the left hand side of (3), viewed as a quadratic in k, is equal to $(\Delta + \delta + 1)^2 - 4n\delta = (\Delta - \delta + 1)^2 - 4\delta(n - \Delta - 1)$; by assumption, the latter expression is less than or equal to 0. So the inequality (3) has no solution in real numbers, a contradiction. □

3. Classes of sequences which are not P-stable

The class of all graphical sequences is *not* P-stable, a fact that can be read off from the following lemma.

Lemma 5. Let $k \geq 3$ be an integer, and $\mathbf{d}$ be the sequence

$$\mathbf{d} = (1, 2, \ldots, k-1, k, k, k+1, \ldots, 2k-2, 2k-1).$$

Then $|\mathcal{G}(\mathbf{d})| = 1$ and $|\mathcal{G}'(\mathbf{d})| \geq 2 \times 3^{k-2}$.

Proof. Observe that the selection of a graph with any given degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$ can be viewed recursively as follows:

(1) select a set of d_n neighbours for vertex n;

(2) recursively select a graph with degree sequence $\mathbf{d}' = (d'_1, \ldots, d'_{n-1})$, where $\mathbf{d}'$ is obtained from $\mathbf{d}$ by deleting the final component and decrementing by 1 the components which correspond to the d_n chosen neighbours for vertex n.

Adopting this view, it is clear that there is a unique graph with degree sequence $\mathbf{d}$.

Now modify $\mathbf{d}$ by incrementing the first component, and decrementing the final component. Note that graphs on the modified degree sequence are members of $\mathcal{G}'(\mathbf{d})$. Applying the recursive selection procedure we therefore have

$$\begin{aligned}
\left|\mathcal{G}'(\mathbf{d})\right| &\geq \left|\mathcal{G}(2,2,3,\ldots,k-1,k,k,k+1,\ldots,2k-3,2k-2,2k-2)\right| \\
&\geq 2 \times \left|\mathcal{G}(1,2,2,3,\ldots,k-2,k-1,k-1,k,\ldots,2k-4,2k-3)\right| \\
&\geq 2 \times \left|\mathcal{G}(1,1,1,2,\ldots,k-3,k-2,k-2,k-1,\ldots,2k-6,2k-5)\right| \\
&\geq 2 \times 3 \times \left|\mathcal{G}(1,1,1,2,\ldots,k-4,k-3,k-3,k-2,\ldots,2k-8,2k-7)\right| \\
&\geq 2 \times 3^2 \times \left|\mathcal{G}(1,1,1,2,\ldots,k-5,k-4,k-4,k-3,\ldots,2k-10,2k-9)\right| \\
&\;\;\vdots \\
&\geq 2 \times 3^{k-3} \times |\mathcal{G}(1,1,1,1)| \\
&= 2 \times 3^{k-2}.
\end{aligned}$$

The factor of 3 arises at each stage from the freedom to choose one of three degree-1 vertices to be adjacent to the vertex of largest degree. (This construction is a slight modification of one which originally appeared in [3].) □

In the previous lemma, the gap between δ and Δ, the minimum and maximum degrees, is a large as it possibly can be. However, by refining the construction, the gap can be narrowed until it approaches the limit set by Theorem 2.

Theorem 6. Let n, δ, Δ be positive integers satisfying $1 \leq \delta \leq \Delta \leq n-1$, and let k be the largest integer not exceeding $\frac{1}{2}(\Delta - \delta + 1) - \left\lceil \sqrt{\delta(n-\Delta-1)}\, \right\rceil$. If $k \geq 3$, there exists a graphical sequence $\mathbf{d} = (d_1, \ldots, d_n)$, with minimum and maximum components δ and Δ, such that $|\mathcal{G}'(\mathbf{d})|/|\mathcal{G}(\mathbf{d})| \geq 2 \times 3^{k-2}$.

Proof. Set $s = \left\lceil \sqrt{\delta(n-\Delta-1)}\, \right\rceil$. We consider graphs on vertex set $X \cup Y \cup R$, where X, Y, and R are disjoint sets with cardinalities $|X| = \delta + s$, $|Y| = n - \Delta - 1 + s$, and $|R| = \Delta - \delta + 1 - 2s$. Note that the conditions of the theorem enforce $|R| \geq 2k$.

Let $D = s\Delta - s^2 + n\delta - \delta$. We claim that it is possible to define an integer weight function $w : X \to [\delta, \Delta]$ on the vertex set X in such a way that total weight of X is D. That this can be done is clear from the two inequalities

$$|X|\Delta - D = (\delta + s)\Delta - (s\Delta - s^2 + n\delta - \delta) = s^2 - \delta(n - \Delta - 1) \geq 0,$$

and

$$D - |X|\delta = (s\Delta - s^2 + n\delta - \delta) - (\delta + s)\delta = s(\Delta - \delta - s) + \delta(n - \delta - 1) \geq 0.$$

(The final inequality here follows from the observation that $\Delta - \delta - s = |R| + s - 1 > 0$.) Let G be a graph on vertex set $X \cup Y \cup R$ whose degrees are constrained as follows. The degree of every vertex in Y is δ, and the degree of each vertex in X is specified by the weight function w; the degrees of vertices in R are left unconstrained. As before, we use $E(X,Y)$ to denote the set of all edges of G which span the two sets X and Y. By considering in turn the total number of edges which are incident at Y and X, we can bound $|E(X,Y)|$ from above and below:

$$\begin{aligned}
|E(X,Y)| &\leq |Y|\delta = (n - \Delta - 1 + s)\delta; \\
|E(X,Y)| &\geq D - |X|(|R| + |X| - 1) = (n - \Delta - 1 + s)\delta.
\end{aligned}$$

Clearly, *equality* must hold in both instances. Thus a large part of the structure of G is completely determined by the degree constraints: the set X must be a clique, Y must be an independent set, the edge set $E(Y,R)$ must be empty, and the bipartite graph induced by the sets X and R must be complete.

Now let $r = |R|$ and $\mathbf{e} = (e_1, \ldots, e_r)$ be any graphical sequence. While retaining the existing degree constraints on $|X|$ and $|Y|$, add degree constraints $|X| + e_1, |X| + e_2, \ldots, |X| + e_r$ to the vertices in R. (It is a straightforward matter to check that all of these degree constraints lie in the range $[\delta, \Delta]$.) The process of selecting a graph G which satisfies *all* the specified degree constraints can be divided into two independent stages:

(1) choose the set of edges $E(X, Y)$;

(2) choose the subgraph of G which is induced by the set R.

In stage (1) of the process, the number of possible choices, N say, is independent of the sequence $\mathbf{e}$. In stage (2), the number of possible choices is $|\mathcal{G}(\mathbf{e})|$. Thus the total number of graphs G which satisfy all the degree constraints is $N \times |\mathcal{G}(\mathbf{e})|$. The theorem follows by specialising $\mathbf{e}$ to the degree sequence described in Lemma 5. (If r is odd, it will be necessary to pad the degree sequence of Lemma 5 with a single 0.) □

When Lemma 1 was introduced, the decision to concentrate on alternating paths of length 1, 3, 5, or 7 may have seemed somewhat arbitrary. One might imagine that considering paths of length 9, for example, might lead to an improved characterisation of P-stability in terms of δ and Δ. Theorem 6 shows that this is not the case, and that paths of length 1, 3, 5 and 7 are the right ones to consider.

From Theorem 6, classes of graphical sequences which fail to be P-stable can be read off. These classes are only slightly larger than the P-stable classes guaranteed by Theorem 2, as can be appreciated from the following examples. Let $\omega(n)$ be any function which tends (however slowly) to ∞, as n increases. As before, denote the minimum and maximum components of a sequence $\mathbf{d}$ by $\delta(\mathbf{d})$ and $\Delta(\mathbf{d})$ respectively. Then the following classes of sequences are *not* P-stable:

(i) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) = 1$ and $\Delta(\mathbf{d}) = \lceil 2\sqrt{n} + \omega(n) \log n \rceil$;

(ii) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) = \lfloor n^{1/3} \rfloor$ and $\Delta(\mathbf{d}) = \lceil 2n^{2/3} - n^{1/3} + \omega(n) \log n \rceil$;

(iii) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) = \lfloor \frac{1}{9} n \rfloor$ and $\Delta(\mathbf{d}) = \lceil \frac{5}{9} n + \omega(n) \log n \rceil$;

(iv) all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ with $\delta(\mathbf{d}) = \lfloor \frac{1}{4} n \rfloor$ and $\Delta(\mathbf{d}) = \lceil \frac{3}{4} n + \omega(n) \log n \rceil$.

4. A more refined characterisation of P-stability

We have observed that Theorem 2 is essentially the best possible characterisation of P-stability in terms of minimum and maximum degrees. However, it is easy to construct P-stable families which lie well outside the conditions of Theorem 2. A simple example is the family of n-component sequences of the form $(1, 1, \frac{1}{2}n, \frac{1}{2}n, \ldots, \frac{1}{2}n)$, where n ranges over the even numbers. This and other examples can be handled by a more refined characterisation which involves the *average* degree in addition to the extreme degrees. The main result follows after a technical lemma.

Lemma 7. Let G be an n-vertex, m-edge graph with distinguished vertices s and t (not necessarily distinct), and suppose that the set of vertices adjacent to s is equal to the set of vertices adjacent to t. Suppose also that δ and Δ are natural numbers such that the degrees of all vertices in G other than s and t lie in the range $[\delta, \Delta]$, and such that s and t themselves have degree at least $\delta + 1$. Set $d = (\Delta - \delta + 1)^2 - 4\delta(n - \Delta - 1)$. If $d \geq 0$ and $4m \leq n(n-1) - (n - \delta - 1)(n - \delta) + \Delta(\Delta + 1) - (\Delta - \delta)\sqrt{d}$, then there exists an edge-disjoint alternating path in G which starts at s, ends at t, and has length 1, 3, 5, or 7.

We defer the proof of Lemma 7 until the end of the section, and move directly to the main theorem.

Theorem 8. Let $\mathcal{F}$ be the class of all graphical sequences $\mathbf{d} = (d_1, \ldots, d_n)$ which satisfy

$$(2m - n\delta)(n\Delta - 2m) \leq (\Delta - \delta)\bigl\{(2m - n\delta)(n - \Delta - 1) + (n\Delta - 2m)\delta\bigr\},$$

where δ and Δ are the minimum and maximum components of $\mathbf{d}$, and $2m = \sum_{i=1}^n d_i$. The class $\mathcal{F}$ is P-stable.

Proof. The stated condition on δ, Δ, and m is equivalent, by straightforward algebraic manipulation, to the inequality

$$\bigl\{(2m - n\delta) - (\Delta - \delta)\delta\bigr\}\bigl\{(n\Delta - 2m) - (\Delta - \delta)(n - \Delta - 1)\bigr\} \leq (\Delta - \delta)^2\delta(n - \Delta - 1).$$

This inequality is of the form $(A + B)(A - B) \leq (\Delta - \delta)^2\delta(n - \Delta - 1)$ where $2A = (\Delta - \delta)(\Delta - \delta + 1)$ and $2B = 4m - n(n-1) + (n - \delta - 1)(n - \delta) - \Delta(\Delta + 1)$. Rearranging the inequality in the form $4B^2 \geq 4A^2 - 4(\Delta - \delta)^2\delta(n - \Delta - 1)$, substituting for A and B, and simplifying the right-hand side yields

$$\bigl\{4m - n(n-1) + (n - \delta - 1)(n - \delta) - \Delta(\Delta + 1)\bigr\}^2 \geq (\Delta - \delta)^2 d, \tag{4}$$

where $d = (\Delta - \delta + 1)^2 - 4\delta(n - \Delta - 1)$.

The proof now closely follows that of Theorem 2. Suppose $\mathbf{d} = (d_1, \ldots, d_n)$ is a sequence satisfying the conditions of the current theorem. We show that $|\mathcal{G}'(\mathbf{d})|/|\mathcal{G}(\mathbf{d})| \leq n^{10}$. Let G be an element of $\mathcal{G}'(\mathbf{d})$, i.e., a graph whose degree sequence agrees with $\mathbf{d}$ except at two vertices s and t, the degree of s being one greater, and the degree of t one smaller than those specified by the sequence $\mathbf{d}$. The case $d \leq 0$ has already been dealt with in the proof of Theorem 2, so suppose $d > 0$. From inequality (4) we see that two cases arise:

(i) $4m \leq n(n-1) - (n - \delta - 1)(n - \delta) + \Delta(\Delta + 1) - (\Delta - \delta)\sqrt{d}$, or
(ii) $4m \geq n(n-1) - (n - \delta - 1)(n - \delta) + \Delta(\Delta + 1) + (\Delta - \delta)\sqrt{d}$.

Suppose the first of these obtains. As before, we shall convert G to a graph with degree sequence *precisely* $\mathbf{d}$ by a short sequence of edge insertions/deletions. Using the manipulations of Theorem 2, we may, using at most two edge insertions and one edge deletion, place ourselves in the situation of Lemma 7. Then, dualising along the alternating path whose existence is guaranteed by that lemma, we obtain a graph with degree sequence exactly $\mathbf{d}$. As in Theorem 2, we deduce $|\mathcal{G}'(\mathbf{d})|/|\mathcal{G}(\mathbf{d})| \leq n^{10}$.

In remains to deal with case (ii). Let

$$\overline{\mathbf{d}} = (\overline{d}_1, \overline{d}_2, \ldots, \overline{d}_n) = (n - d_1 - 1, n - d_2 - 1, \ldots, n - d_n - 1)$$

be the sequence dual to $\mathbf{d}$. Let $\overline{\delta}$, $\overline{\Delta}$ be the minimum and maximum components of $\overline{\mathbf{d}}$, and $2\,\overline{m} = \sum_{i=1}^n \overline{d}_i$. Note that $\overline{\delta} = n - \Delta - 1$, $\overline{\Delta} = n - \delta - 1$, and $\overline{m} = \frac{1}{2}n(n-1) - m$. Now the inequalities which define cases (i) and (ii) above are dual; that is to say, a sequence $\mathbf{d}$ satisfies the first inequality if and only if the complementary sequence $\overline{\mathbf{d}}$ satisfies the second, and vice versa. (To see this, observe that the substitutions $\delta \leftarrow n - \overline{\Delta} - 1$, $\Delta \leftarrow n - \overline{\delta} - 1$, and $m \leftarrow \frac{1}{2}n(n-1) - \overline{m}$ map the inequality defining case (i) onto a primed version of the inequality defining case (ii), and vice versa.) We complete the treatment of case (ii) by observing that $|\mathcal{G}'(\mathbf{d})|/|\mathcal{G}(\mathbf{d})| = |\mathcal{G}'(\overline{\mathbf{d}})|/|\mathcal{G}(\overline{\mathbf{d}})| \leq n^{10}$, as required. □

Remarks. (i) The condition in the statement of Theorem 8 is invariant under complementation. (ii) From the equivalent form of the condition, presented as inequality (4), it is clear that Theorem 8 is a generalisation of Theorem 2. (iii) Theorem 8 is also strictly more powerful than a criterion for P-stability previously presented by two of the authors in [3]. (iv) The following example will help to give the flavour of the general result. By setting $\delta = \frac{4}{25}n$ and $\Delta = \frac{21}{25}n - 1$ in Theorem 8 we may read off the following fact: the class of all graphical sequences with minimum degree not less than $\frac{4}{25}n$, maximum degree not exceeding $\frac{21}{25}n - 1$, and average degree not exceeding $\frac{37}{125}n - \frac{1}{5}$, is P-stable. Note that the minimum and maximum degrees, taken in isolation, are insufficient to ensure P-stability. (v) Naturally enough, Corollaries 3 and 4 continue to hold when $\mathcal{F}$ is the extended class of sequences specified in Theorem 8. □

We conclude the section by presenting a proof of the technical lemma.

Proof of Lemma 7. As in the proof of Lemma 1, we assume that such a path does *not* exist and obtain a contradiction. Recall the notation of the proof of Lemma 1, in particular the partition the vertex set of G into three sets K, R, and Y, with respective cardinalities k, r, and y. Combining inequalities (1) and (2), we obtain a lower bound on the number of edges which have exactly one endpoint in the set K:

$$|E(K, R \cup Y)| = |E(K,R)| + |E(K,Y)| > r\delta + y\delta = (n-k)\delta.$$

Furthermore, the set K, being a clique, has $\frac{1}{2}k(k-1)$ internal edges. Thus the total number of edges in G is greater than $(n-k)\delta + \frac{1}{2}k(k-1)$, leading to the inequality

$$k^2 - (2\delta+1)k + (2n\delta - 2m) < 0. \tag{5}$$

The statement of the lemma imposes a bound on m which is of the form $4m \le A - (\Delta - \delta)\sqrt{d}$, where A is an expression involving δ, Δ, and n; this inequality may be recast in the form

$$\begin{aligned} 8m &\le 2A + (\Delta - \delta - \sqrt{d}\,)^2 - (\Delta-\delta)^2 - d \\ &= 2A + (\Delta - \delta - \sqrt{d}\,)^2 - (\Delta-\delta)^2 - (\Delta - \delta + 1)^2 + 4\delta(n - \Delta - 1), \end{aligned}$$

which, on substituting for A and collecting terms, simplifies to

$$8m \le (\Delta - \delta - \sqrt{d}\,)^2 + 8n\delta - (2\delta+1)^2. \tag{6}$$

Eliminating m between inequalities (5) and (6), we obtain

$$(2k - 2\delta - 1)^2 - (\Delta - \delta - \sqrt{d}\,)^2 < 0. \tag{7}$$

If we were assured, at this point, that $\sqrt{d} \le \Delta - \delta$, then we would have an immediate contradiction: inequality (7) would imply $k < \frac{1}{2}\{\Delta + \delta + 1 - \sqrt{d}\}$, whereas inequality (3) implies $k > \frac{1}{2}\{\Delta + \delta + 1 - \sqrt{d}\}$. However, the definition of d provides only the (slightly) weaker guarantee that $\sqrt{d} \le \Delta - \delta + 1$.

We complete the proof by dealing with this pathological case, which is characterised by the condition $-1 \le \Delta - \delta - \sqrt{d} < 0$. On the one hand, from inequality (7), we have $k < \frac{1}{2}\big\{(2\delta+1) - (\Delta - \delta - \sqrt{d}\,)\big\} \le \delta + 1$. On the other, from inequality (3), we have $k > \frac{1}{2}\{\Delta + \delta + 1 - \sqrt{d}\} > \delta$. Clearly, there is no *integral* k which satisfies this pair of inequalities. □

References

[1] E. A. Bender and E. R. Canfield, The asymptotic number of labelled graphs with given degree sequence, *Journal of Combinatorial Theory (A)* **24** (1978), pp. 296–307.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal of Combinatorics* **1** (1980), pp. 311–316.

[3] M. R. Jerrum and A. J. Sinclair, *Fast uniform generation of regular graphs*, Internal report CSR-281-88, Department of Computer Science, University of Edinburgh, November 1988. (To appear in *Theoretical Computer Science.*)

[4] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani, Random generation of combinatorial structures from a uniform distribution, *Theoretical Computer Science* **43** (1986), pp. 169–188.

[5] B. D. McKay, and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs with degrees* $o(n^{1/2})$, Preprint, 1988.

[6] B. D. McKay and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, Preprint, 1988.

[7] B. D. McKay and N. C. Wormald, *Uniform generation of random regular graphs of moderate degree.* To appear in *Journal of Algorithms.*

[8] A. J. Sinclair, *Randomised Algorithms for Counting and Generating Combinatorial Structures*, Ph.D. Thesis, University of Edinburgh, 1988.

[9] N. C. Wormald, Generating random regular graphs, *Journal of Algorithms* **5** (1984), pp. 247–280.