Karp-Sipser on random graphs with a fixed degree sequence

Tom Bohman* and Alan Frieze†
Department of Mathematical Sciences,
Carnegie Mellon University
Pittsburgh, PA 15213

April 19, 2011

Abstract

Let $\Delta \geq 3$ be an integer. Given a fixed $\mathbf{z} \in \mathbb{R}_+^\Delta$ such that $z_\Delta > 0$, we consider a graph $G_\mathbf{z}$ drawn uniformly at random from the collection of graphs with $z_i n$ vertices of degree i for $i = 1, \ldots, \Delta$. We study the performance of the Karp-Sipser algorithm when applied to $G_\mathbf{z}$. If there is an index $\delta > 1$ such that $z_1 = \cdots = z_{\delta-1} = 0$ and $\delta z_\delta, \ldots, \Delta z_\Delta$ is a log-concave sequence of positive reals then with high probability the Karp-Sipser algorithm succeeds in finding a matching with $n\|\mathbf{z}\|_1/2 - o(n^{1-\epsilon})$ edges in $G_\mathbf{z}$ where $\epsilon = \epsilon(\Delta, \mathbf{z})$ is a constant.

1 Introduction.

The Karp-Sipser algorithm is the following randomized algorithm for finding a large matching in a graph. We begin with the empty matching and iteratively choose edges from the graph at random and add them to the matching. When an edge is added to the matching the vertices in that edge and all incident edges are removed from the graph. Care is taken in the choice of the random edge only when there are vertices of degree 1 in the graph. If vertices of degree 1 exist, one is chosen at random and the edge containing that vertex is added to the matching. Note that when the algorithm adds an edge that contains a vertex of degree 1 (sometimes called a pendant edge) no 'mistake' is made; to be precise, if G is a graph with a pendant edge e then G has a maximum matching that contains e.

There are polynomial time algorithms for finding maximum matchings in general graphs (see [6] [14]). However, these algorithms are somewhat complicated and require significantly more than linear time. The Karp-Sipser algorithm, on the other hand, is a simple linear time algorithm that finds a nearly optimal matching on the average case. Indeed, the performance of the Karp-Sipser algorithm on the sparse random graph $G_{n,cn/2}$ is well understood. If $c < e$

*Supported in part by NSF grant DMS-0401147. e-mail `tbohman@math.cmu.edu`
†Supported in part by NSF grant CCR-0502793. e-mail `alan@random.math.cmu.edu`

then **whp** the Karp-Sipser algorithm finds a maximum matching in $G_{n,cn/2}$. If $c > e$ then **whp** the Karp-Sipser algorithm finds a matching with cardinality within $n^{1/5+\epsilon}$ of the largest matching in $G_{n,cn/2}$ for any $\epsilon > 0$. These results were achieved by Aronson, Frieze and Pittel [1] in refinement of earlier results of Karp and Sipser [13]. While these results show that the Karp-Sipser algorithm succeeds in finding a nearly maximum matching **whp** [1] in almost all sparse graphs, we do not have an understanding of how it performs on more restrictive classes of sparse graphs; that is, when we impose conditions beyond simply insisting that the graph has certain a number of edges.

Consider the following model. We use $\mathbb{R}_+$ to denote the non-negative reals. Let Δ be a positive integer and let $\mathbf{z} \in \mathbb{R}_+^\Delta$ be a fixed vector with $z_\Delta > 0$. We define $G_\mathbf{z}$ to be a graph drawn uniformly at random from the collection of a graphs with $z_i n$ vertices of degree i for $i \geq 2$. (We systematically drop nearest integer notation and assume that the sum of the degrees in any degree sequence is even. These conventions have no impact on our results.) Note that we do not impose the condition $\sum_{i=1}^{\Delta} z_i = 1$. The number of vertices in $G_\mathbf{z}$ is $\|z\|_1 n$. ([5] and [15] treat similar models.) We consider the question:

> Which distributions $\mathbf{z}$ have the property that the Karp-Sipser algorithm is likely to produce a matching with $(1 - o(1))\|\mathbf{z}\|_1 n/2$ edges in $G_\mathbf{z}$?

We will refer to such matchings as **almost perfect** matchings. Of course, this question is closely related to the matching numbers of these graphs, a natural question from the perspective of the theory of random graphs that seems not to have been addressed. We return to a discussion of the matching numbers of these graphs in the Conclusion of this paper. Note that if the graph $G_\mathbf{z}$ has $z_1 > 0$ then **whp** there will be a linear number of pairs of vertices of degree 1 that have a common neighbor. At most one vertex from such a pair is saturated by a given matching. Therefore, if $z_1 > 0$ then **whp** $G_\mathbf{z}$ will not have an almost perfect matching. So, we henceforth consider only degree distributions $\mathbf{z}$ with $z_1 = 0$. (N.b. there are other degree distributions for which we can easily see that the Karp-Sipser algorithm does not produce an almost perfect matching. These are discussed in Section 2 below.)

Let δ be the minimum degree of the random graph $G_\mathbf{z}$ (i.e. the smallest i such that $z_i > 0$). Our results hold for vectors $\mathbf{z}$ such that

$$z_i > 0 \quad \text{and} \quad i^2 z_i^2 \geq (i-1)z_{i-1}(i+1)z_{i+1} \quad \text{for} \quad i = \delta+1, \ldots, \Delta-1. \tag{1}$$

Many natural degree sequences satisfy this condition. For example, if we have

$$z_i = \begin{cases} 1 & \text{if } \delta \leq i \leq \Delta \\ 0 & \text{if } 1 \leq i < \delta \end{cases}$$

then $\mathbf{z}$ satisfies (1). Regular graphs, for example, have degree distributions of this type. For a second example, consider the distribution given by

$$z_i = \begin{cases} \frac{\lambda^i}{i!} & \text{if } \delta \leq i < \Delta \\ 0 & \text{if } 1 \leq i < \delta \end{cases}$$

[1] a sequence of events $\mathcal{E}_1, \mathcal{E}_2, \ldots$ occur with high probability (**whp**) if $\lim_{n\to\infty} Pr(\mathcal{E}_n) = 1$.

(i.e. a truncated Poisson random variable). Then we have

$$i^2 z_i^2 = \frac{\lambda^i}{(i-1)!}\frac{\lambda^i}{(i-1)!} > \frac{\lambda^{i-1}}{(i-2)!}\frac{\lambda^{i+1}}{i!} = (i-1)z_{i-1}(i+1)z_{i+1}.$$

So the truncated Poisson distribution satisfies (1). We are now ready to state our main result.

Theorem 1. *If $\Delta \geq 3$ and $\mathbf{z} \in \mathbb{R}_+^\Delta$ such that $z_1 = 0$, $z_\Delta > 0$ and $\mathbf{z}$ satisfies (1) then* **whp** *the Karp-Sipser algorithm produces a matching with $(1 - o(n^{-\epsilon}))\,\|\mathbf{z}\|_1 n/2$ edges in $G_\mathbf{z}$ for some constant $\epsilon = \epsilon(\Delta, \mathbf{z})$.*

An analysis of the evolution of the Karp-Sipser algorithm on the graph $G_\mathbf{z}$ for a fixed $\mathbf{z}$ can be achieved by a (more or less) straightforward application of the so-called differential equations method for random graph processes (see Wormald [17] for an introduction to and discussion of this method). We write a system of ordinary differential equations with Δ variables (one corresponding to the set of vertices of degree i for each $1 \leq i \leq \Delta$) and show that **whp** the evolution of the algorithm is concentrated around the trajectory given by the solution of the system. Usually the solution of such a system of differential equations does not have a simple closed form. But one can get an approximation for the trajectory by solving the system numerically. Of course, this approach can only handle one distribution $\mathbf{z}$ at a time. Furthermore, numerically approximating this system all the way to the termination of the algorithm with good bounds on the errors in the approximation is a significant technical challenge because a certain Lipschitz constant goes to infinity as the number of remaining vertices goes to zero. (This issue is discussed in more detail in Section 2 below.) Thus, the main contribution of Theorem 1 is an *analytic* proof that the solution of the associated system of differential equations is well-behaved all the way to the termination of the algorithm for every distribution $\mathbf{z}$ that satisfies the given conditions.

The remainder of the paper is organized as follows. In the next section we introduce the system of differential equations that governs the evolution of the Karp-Sipser algorithm on $G_\mathbf{z}$ and, assuming that the evolution of Karp-Sipser follows the trajectory given by the solution of this system of differential equations, prove Theorem 1. In Section 3 we verify that the evolution of Karp-Sipser indeed follows the trajectory discussed in Section 2. The Conclusion contains an extension of Theorem 1 to random graphs $G_\mathbf{z}$ with $z_1 > 0$, a discussion of the matching number of the random graph $G_\mathbf{z}$ and a number of open questions.

2 Proof of Theorem 1

We analyze the Karp-Sipser algorithm applied to the random graph $G_\mathbf{z}$. We use the standard configuration model to generate $G_\mathbf{z}$ (for an introduction to and discussion of this model see Bollobás [2] or Section 9.1 in [11]). Each vertex v is associated with a set of **configuration points** where the number of points in this set is equal to the degree of v. A **configuration** is a uniform random matching on the set of configuration points. Each configuration corresponds to a (multi-)graph with the desired degree sequence: To get the graph that corresponds to a given configuration simply contract the set of configuration points associated with each vertex. We prove Theorem 1 in the configuration model; since we consider only graphs with bounded maximum degree, any property that holds **whp** in the configuration model also holds **whp** in $G_{\mathbf{z}}$.

We reveal the configuration only as necessary as the algorithm proceeds. So, for example, when we choose a random edge we choose two configuration points at random and when we determine the neighbor of a particular configuration point it is chosen at random from the collection of remaining configuration points. Thus, our probability space is the uniform distribution over the set of all pairs consisting of a configuration and a sequence of bits (the sequence of bits determines the edge choices of the Karp-Sipser algorithm). Furthermore, if we condition on the evolution of the process up to a certain point then the remaining graph and the remaining bits form a probability space of the same form.

We follow the proof and notation of Theorem 5.1 in Wormald [17]. The vector $\mathbf{Y} \in \mathbb{Z}_+^\Delta$ corresponds to the configuration model with Y_i vertices of degree i for $i = 1, \dots, \Delta$. The vector $\mathbf{z}$ is our continuous approximation. Throughout this discussion we assume that we are working with $\mathbf{Y}$ that is close (in a way that will be made precise when we get to the actual proof) to $n\mathbf{z}$. Define

$$\mu(\mathbf{z}) = \sum_{i=1}^{\Delta} i z_i$$

for $\mathbf{z} \in \mathbb{R}^\Delta$. We have

$$\mu(\mathbf{Y}) = \sum_{i=1}^{\Delta} i Y_i.$$

Note that we assume $\mu(\mathbf{Y})$ to be approximately $n\mu(\mathbf{z})$ and that $\mu(\mathbf{Y})$ is the number of configuration points in the configuration model $\mathbf{Y}$.

There are two 'phases' of the Karp-Sipser algorithm applied to the graph $\mathbf{Y}$ that correspond to two different systems of differential equations governing the evolution of the algorithm: one for distributions $\mathbf{z}$ for which the algorithm spends *all* of its time dealing with pendant edges and another for distributions $\mathbf{z}$ where pendant edges do not significantly accumulate and the algorithm splits its time between dealing with pendant edges and choosing purely random edges. As we are only interested in those graphs $\mathbf{Y}$ for which the Karp-Sipser algorithm produces almost-perfect matchings, we restrict our attention to the latter phase. We restrict our attention to $\mathbf{z}$ such that $z_1 = 0$ and pendant edges don't quickly accumulate for $\mathbf{Y}$ in the vicinity of $\mathbf{z}$.

In order to characterize the set of distributions $\mathbf{z}$ with this property, we consider the expected change in the number of vertices of degree one in a step of the Karp-Sipser algorithm in which a pendant edge is added to the matching. Let v be a vertex of degree 1. We choose a configuration point uniformly at random to be in the edge with the configuration point corresponding to v. The probability that this configuration point corresponds to a vertex u of degree i is $iY_i/(\mu(\mathbf{Y}) - 1)$ for $i = 2, \dots, \Delta$. In this event we then delete the remaining $i - 1$ edges incident with the vertex u. If any of these edges is incident with a vertex (other than u) of degree 2 then a new pendant edge is created. The probability of this event is $2Y_2/(\mu(\mathbf{Y}) - 3)$ for each edge deleted (unless the vertex u has degree 2 in which case 2 configuration points corresponding to the vertices of degree 2 have already been removed).

Therefore, the expected number of vertices of degree one created is

$$\frac{2Y_2}{\mu(\mathbf{Y})-1}\cdot\frac{2Y_2-2}{\mu(\mathbf{Y})-3}+\sum_{i=3}^{\Delta}\frac{iY_i}{\mu(\mathbf{Y})-1}\cdot(i-1)\frac{2Y_2}{\mu(\mathbf{Y})-3}=$$
$$\frac{2Y_2}{\mu(\mathbf{Y})^2}\sum_{i=2}^{\Delta}Y_i i(i-1)+O\left(\frac{1}{\mu(\mathbf{Y})}\right),$$

Note further that if the Karp-Sipser algorithm adds a purely randomly edge from the graph $\mathbf{Y}$ to the matching then the expected number of vertices of degree one created is twice the final expression. Define

$$\varphi(\mathbf{z})=\frac{2z_2}{\mu(\mathbf{z})^2}\sum_{i=2}^{\Delta}z_i i(i-1)$$

for $\mathbf{z}\in\mathbb{R}_+^\Delta$. One can easily show that if $\varphi(\mathbf{z})>1$ and $\mathbf{Y}$ is close to $n\mathbf{z}$ then **whp** the Karp-Sipser algorithm quickly produces a linear number of vertices of degree 1 and therefore **whp** does not produce an almost-perfect matching in the graph $\mathbf{Y}$. So, we henceforth restrict our attention to the following region (i.e. we only consider $\mathbf{z}$ that lie in this region):

$$\mathcal{D}=\left\{\mathbf{x}\in\mathbb{R}_+^\Delta : x_1=0, x_\Delta>0 \text{ and } \varphi(\mathbf{x})<1\right\}.$$

By restricting $\mathbf{z}$ in this way we ensure that we only deal with the phase of the Karp-Sipser algorithm that splits time between adding pendant edges and adding purely random edges to the matching.

We are now ready to set-up our system of differential equations. We write the system of differential equations by making a heuristic argument that relates the derivatives of the continuous variables with the expected one-step changes in the corresponding random variables. The proof that the evolution of the Karp-Sipser algorithm actually follows the trajectory given by the solution of this differential equations is given below (in Section 3). We begin with the expected changes in the variables in one step of the process. Let $\mathbf{Y}^+$ be the degree distribution after we apply a single step of the Karp-Sipser algorithm to the graph $\mathbf{Y}$. If we add a pendant edge to our matching we have

$$E[Y_j^+-Y_j]=-\frac{jY_j}{\mu(\mathbf{Y})}+\sum_{i=2}^{\Delta}\frac{iY_i}{\mu(\mathbf{Y})}(i-1)\left(-\frac{jY_j}{\mu(\mathbf{Y})}+\frac{(j+1)Y_{j+1}}{\mu(\mathbf{Y})}\right)+O\left(\frac{1}{\mu(\mathbf{Y})}\right)$$

for $i=2,\ldots,\Delta$. This expression follows from reasoning that is similar to the calculation that gave the expected number of degree 1 vertices created given above: the first term accounts for change in Y_j from the event that the pendant edge added to the matching contains a vertex of degree j and the second term (i.e. the summation) accounts for the change in Y_j that results from the deletion of edges that are incident with the edge added to the matching. Note that if we add a purely random edge the expected change in Y_j is twice this expression. Since there are two types of steps in the process, we cannot write the differential equations in terms of the expected changes in a single step. Instead we write a system of differential equations given by the expected changes over many steps of the process. We do this by taking into account the proportion of steps used on each of the two types of steps in the Karp-Sipser algorithm. (Of course, this is not a 'standard' application of the differential equations method as we cannot directly apply Theorem 5.1 of [17], for example. The analysis due to Aronson, Frieze and Pittel [1] of the Karp-Sipser algorithm applied to random graph $G_{n,cn}$ also 'mixes' the two different types of Karp-Sipser steps into a single differential equation. In that case this is achieved, rather miraculously, by a differential equation in a single variable. A situation in which multiple types of moves are 'mixed' into a single system of differential equations also arises in Shi and Wormald's recent application of the differential equations method to a greedy coloring algorithm on random regular graphs [16].)

When the algorithm chooses a pendant edge for the matching, the expected number of pendant edges created is roughly $\varphi(\mathbf{z})$, and when the algorithm chooses a purely random edge for the matching, the expected number of pendant edges created is roughly $2\varphi(\mathbf{z})$. So, setting $\alpha(\mathbf{z})$ equal to the proportion of time that the algorithm spends choosing purely random edges for the matching (for distributions in the vicinity of $\mathbf{z}$) and noting that over many steps of the algorithm we expect nearly all of the pendant edges that appear to be added to the matching, we have

$$\alpha(\mathbf{z})(2\varphi(\mathbf{z})) + (1-\alpha(\mathbf{z}))\varphi(\mathbf{z}) = 1-\alpha(\mathbf{z}) \qquad \Rightarrow \qquad \alpha(\mathbf{z}) = \frac{1-\varphi(\mathbf{z})}{1+\varphi(\mathbf{z})}.$$

Let $\mathbf{Y}(\ell)$ be the degree distribution of the remaining graph after ℓ steps of the algorithm (each edge added to the matching corresponds to a single step). We introduce a continuous parameter t defined by $t = \ell/n$, and we suppose $\mathbf{Y}(\ell) \approx \mathbf{z}(t)n$. (We generally write $\mathbf{z}$ in the place of $\mathbf{z}(t)$ for ease of notation). Taking into account the two types of steps in the algorithm, we arrive at the following system of differential equations:

$$\begin{aligned}\frac{dz_i}{dt} = f_i(\mathbf{z}) = \frac{2\varphi(\mathbf{z})}{1+\varphi(\mathbf{z})}&\left[-\frac{iz_i}{\mu(\mathbf{z})} + \left(\sum_{j=2}^{\Delta}\frac{jz_j}{\mu(\mathbf{z})}(j-1)\right)\left(-\frac{iz_i}{\mu(\mathbf{z})} + \frac{(i+1)z_{i+1}}{\mu(\mathbf{z})}\right)\right]\\ &+\frac{1-\varphi(\mathbf{z})}{1+\varphi(\mathbf{z})}\left[-\frac{2iz_i}{\mu(\mathbf{z})} + \left(\sum_{j=2}^{\Delta}2\frac{jz_j}{\mu(\mathbf{z})}(j-1)\right)\left(-\frac{iz_i}{\mu(\mathbf{z})} + \frac{(i+1)z_{i+1}}{\mu(\mathbf{z})}\right)\right] \\ =\frac{2}{1+\varphi(\mathbf{z})}&\left[-\frac{iz_i}{\mu(\mathbf{z})} + \left(\sum_{j=2}^{\Delta}\frac{jz_j}{\mu(\mathbf{z})}(j-1)\right)\left(-\frac{iz_i}{\mu(\mathbf{z})} + \frac{(i+1)z_{i+1}}{\mu(\mathbf{z})}\right)\right]\end{aligned} \qquad (2)$$

for $2 \le i \le \Delta$. We emphasize here that we have not yet established that Karp-Sipser actually follows the solution of this autonomous system of differential equations. This is done precisely below (see Theorem 2).

Before proceeding with the proof we make a number of observations regarding the system of differential equations given in (2). These are purely analytic observations; they do not make use of the correspondence with the Karp-Sipser algorithm. First note that this system of differential equations does not have a Lipschitz constant over the entire region $\mathcal{D}$; in particular, the partial derivatives of f_i are unbounded as $\mu \to 0$. However, we do have

$$\frac{\partial f_i}{\partial z_j} = O\left(\frac{1}{\mu(\mathbf{z})}\right). \qquad (3)$$

This follows from

$$f_i(\mathbf{z}) = \frac{2\left(-iz_i\mu(\mathbf{z}) + \sum_{j=2}^{\Delta} jz_j(j-1)(-iz_i + (i+1)z_{i+1})\right)}{\mu^2(\mathbf{z}) + 2z_2\sum_{j=2}^{\Delta} jz_j(j-1)},$$

the fact that the partial derivatives of both the numerator and denominator are linear in $z_2, \dots, z_\Delta$ and the simple observation $jz_j \le \mu(\mathbf{z})$. It follows that f_i is Lipschitz on regions of the form

$$\mathcal{D}_\epsilon = \left\{\mathbf{x} \in \mathcal{D} : \sum_{i=2}^{\Delta} ix_i > \epsilon\right\}.$$

Therefore, for any initial condition $\mathbf{z}(0) \in \mathcal{D}$ there is a unique solution $\mathbf{z}(t)$ to the system of differential equations (2) that extends to points arbitrarily close to the boundary of $\mathcal{D}$. Note next that when we compute $d\mu(\mathbf{z})/dt$ many of the terms cancel:

$$\begin{aligned}\frac{d\mu(\mathbf{z})}{dt} = \sum_{i=2}^{\Delta} if_i(\mathbf{z}) &= -\frac{2}{1+\varphi(\mathbf{z})}\sum_{i=2}^{\Delta}\frac{iz_i}{\mu(\mathbf{z})}\left[2i-1+\frac{2z_2}{\mu(\mathbf{z})}(i-1)\right] \\ &= -\left(\frac{2\varphi(\mathbf{z})\mu(\mathbf{z})}{z_2(1+\varphi(\mathbf{z}))}+2\right) \le -4. \qquad (4)\end{aligned}$$

Similarly, we can bound dz_i/dt; that is, we have

$$-2\Delta \le -\frac{2\Delta iz_i}{\mu(\mathbf{z})} \le f_i(\mathbf{z}) \le 2\Delta. \qquad (5)$$

We note next that if $\mathbf{z}(0)$ is on the boundary of $\mathcal{D}$ (i.e. because some $z_i(0)$ equals 0) then there is some $\delta > 0$ such that $\mathbf{z}(t)$ is in the interior of $\mathcal{D}$ for $t \in (0, \delta]$. To see this, consider an index $i \ge 2$ such that $z_i(0) = 0$ and let $k \le \Delta$ be the smallest index larger than i such that $z_k(0) > 0$. We claim that for $i \le l \le k-1$ we have

$$\frac{d^j z_l}{dt^j}(0) = \frac{d^{j-1}f_l}{dt^{j-1}}(0)\begin{cases} = 0 & \text{if } j < k-l \\ > 0 & \text{if } j = k-l, \end{cases} \qquad (6)$$

which follows from the fact that we can write $f_l = z_l g_1 + z_{l+1}g_2$ where $g_1 < 0$ and $g_2 > 0$. We verify (6) by induction on j. As a base case observe that

$$\frac{dz_l}{dt} = f_l = z_l g_1 + z_{l+1}g_2\begin{cases} = 0 & \text{if } l < k-1 \\ > 0 & \text{if } l = k-1. \end{cases}$$

Now suppose that (6) is true for $j \le m-1$. Consider $i \le l \le k-m$ and write

$$\frac{d^m z_l}{dt^m} = \frac{d^{m-1}f_l}{dt^{m-1}} = \sum_{r=0}^{m-1}\binom{m-1}{r}\left(\frac{d^r z_l}{dt^r}\frac{d^{m-1-r}g_1}{dt^{m-1-r}} + \frac{d^r z_{l+1}}{dt^r}\frac{d^{m-1-r}g_2}{dt^{m-1-r}}\right). \qquad (7)$$

It follows from our inductive assumption that $d^r z_l/dt^r = 0$ for $r + l < k$ and $r \le m-1$. Therefore, every term in (7) is zero if $l < k-m$, and in the $l = k-m$ case we have

$$\frac{d^m z_{k-m}}{dt^m} = g_2\frac{d^{m-1}z_{k-m+1}}{dt^{m-1}} > 0,$$

again by induction.

Having established that the solution to (2) begins by moving into the interior of $\mathcal{D}$, we now characterize the ways in which the solution $\mathbf{z}(t)$ of the system (2) with initial condition $\mathbf{z}(0) \in \mathcal{D}$ can leave the region $\mathcal{D}$. It follows from (5) that

$$z_i(t) \geq z_i(\delta) \exp\left(-2i\Delta \int_\delta^t \frac{d\tau}{\mu(\mathbf{z}(\tau))}\right) \quad \text{for} \quad t \geq \delta \quad \text{and} \quad i = 2, \ldots, \Delta$$

(this is an application of Gronwall's inequality). Thus

$$z_i(t) > 0 \quad \text{for} \quad i = 2, \ldots, \Delta \quad \text{and} \quad t > 0 \tag{8}$$

and

$$\lim_{t \to s^-} z_i(t) = 0 \quad \Rightarrow \quad \lim_{t \to s^-} \mu(\mathbf{z}(t)) = 0. \tag{9}$$

On the other hand, it follows from (4) that, as long as $\mathbf{z}(t)$ remains in the interior of $\mathcal{D}$, $\mu(\mathbf{z}(t))$ decreases at rate 4, at least. So there exists a finite s such that $\mathbf{z}(t)$ is defined for $t \in (0, s)$ and either

$$\lim_{t \to s^-} \mathbf{z}(t) = 0 \qquad \text{or} \qquad \lim_{t \to s^-} \varphi(\mathbf{z}(t)) = 1.$$

So, the point where the solution to the system of the differential equations (2) leaves the interior of $\mathcal{D}$ is either at the origin or at a point $\mathbf{z}$ with $\varphi(\mathbf{z}) = 1$. Note that these two outcomes correspond to success and failure, respectively, of the Karp-Sipser algorithm applied to the graph $\mathbf{Y}(0)$ (assuming, of course, that the evolution of the Karp-Sipser algorithm follows the solution of this system of differential equations): If we arrive at a distribution in the vicinity of the origin then nearly all vertices are saturated by the matching produced by the Karp-Sipser algorithm and if we arrive at a distribution with $\varphi > 1$ then pendant edges quickly accumulate and Karp-Sipser does not produce an almost perfect matching. This concludes our discussion of the analytic properties of the solution of system of differential equations (2).

Now we return to our discussion of the Karp-Sipser algorithm itself. We first note that we can apply the differential equations method to conclude that the evolution of the algorithm indeed follows the trajectory given by the solution of (2). Note that in order to prove Theorem 1 we also need to show that **whp** very few vertices become isolated (i.e. have degree drop to zero before being saturated by the matching) in the course of this evolution. Let $\mathbf{Y}(\ell)$ be the degree distribution after ℓ steps of the Karp-Sipser algorithm have been applied to the graph $\mathbf{Y}(0)$.

Theorem 2. *Let $\mathbf{a} \in \mathcal{D}$. Let $\mathbf{z}(t)$ be the solution of the system of differential equations (2) with initial condition $\mathbf{z}(0) = \mathbf{a}$. Let $\mathbf{Y}(\mathbf{0}) = G_{\mathbf{a}}$. There exists a constant $\epsilon = \epsilon(\Delta) > 0$ such that the following holds. If $\sigma = \sigma(n)$ satisfies the condition*

$$\varphi(\mathbf{z}(t)) < 1 - n^{-1/6} \quad \textit{and} \quad \mu(\mathbf{z}(t)) > n^{-\epsilon} \quad \textit{for} \quad 0 \leq t \leq \sigma \tag{10}$$

then **whp** *we have*

$$|Y_i(\ell) - nz_i(\ell/n)| < n^{9/10} \quad \textit{for} \quad i = 2, \ldots, \Delta \quad \textit{and} \quad 0 \leq \ell \leq \sigma n, \tag{11}$$

and the number of vertices stranded (i.e. vertices whose degrees drops to zero before being saturated by the matching) in the first σn rounds is at most $n^{3/4}$.

The proof of Theorem 2 is given in Section 3.

With this understanding of the evolution of the Karp-Sipser algorithm within the region $\mathcal{D}$ in hand, we are ready to consider log-concave degree-distributions. Define

$$\mathcal{C} = \left\{ \mathbf{z} \in \mathbb{R}_+^\Delta : z_1 = 0, z_\Delta > 0 \text{ and } \mathbf{z} \text{ satisfies (1)} \right\}.$$

We establish a relationship between the region $\mathcal{C}$ and the region $\mathcal{D}$ as well as a relationship between the region $\mathcal{C}$ and the system of differential equations (2). In order to do so, we replace (1) with the following equivalent condition. We say that $\mathbf{z}$ is **fully log-concave** if

$$x + y = a + b \text{ and } x < a \le b < y \qquad \Rightarrow \qquad xz_x yz_y \le az_a bz_b. \tag{12}$$

It is easy to see that if $\mathbf{z}$ is fully log-concave then the sequence iz_i is log-concave. The converse also holds. Indeed, if $\delta z_\delta, \ldots, \Delta z_\Delta > 0$ and $\delta z_\delta, \ldots, \Delta z_\Delta$ is log-concave and $\delta \le x \le y \le \Delta$ then we have

$$\prod_{i=x+1}^{y-1} i^2 z_i^2 \ge \prod_{i=x+1}^{y-1} (i-1)z_{i-1}(i+1)z_{i+1} = xz_x(x+1)z_{x+1} \left[\prod_{i=x+2}^{y-2} i^2 z_i^2 \right] (y-1)z_{y-1} yz_y.$$

This inequality implies $(x+1)z_{x+1}(y-1)z_{y-1} \le xz_x yz_y$.

Theorem 3. *The region $\mathcal{C}$ is a subset of the region $\mathcal{D}$, and*

$$\mathbf{z} \in \overline{\mathcal{C}} \quad \text{and} \quad \varphi(\mathbf{z}) = 1 \qquad \Rightarrow \qquad z_1 = z_3 = \cdots = z_\Delta = 0. \tag{13}$$

Proof. If $\mathbf{z} \in \mathbb{R}_+^\Delta$ satisfies (12), $z_\Delta > 0$ and $\Delta \ge 3$ then we have

$$\begin{aligned} \left(\sum_{i=2}^{\Delta} i(i-1)z_i \right) 2z_2 &= \sum_{i=2}^{\Delta} \sum_{j=2}^{i} iz_i 2z_2 \\ &\le \sum_{i=2}^{\Delta} \sum_{j=2}^{i} jz_j (i+2-j) z_{i+2-j} \\ &= \sum_{k=2}^{\Delta} \sum_{l=2}^{\Delta+2-k} kz_k \cdot lz_l \\ &\le \left(\sum_{k=2}^{\Delta} kz_k \right)^2 - \Delta^2 z_\Delta^2. \end{aligned} \tag{14}$$

Note that we use the condition $\Delta \ge 3$ in the last line. So $\varphi(\mathbf{z}) < 1$ so long as $\mathbf{z} > \mathbf{0}$ is log-concave *and* z_i is non-zero for some $i \ge 3$. This gives $\mathcal{C} \subseteq \overline{\mathcal{D}}$. To prove the boundary condition (13), consider $\mathbf{z} \in \overline{\mathcal{C}}$ and let k be the largest index such that $z_k > 0$. It follows from (14), replacing Δ in the upper limit of the summation with k, that if $k \ge 3$ then $\varphi(\mathbf{z}) < 1$. $\square$

Next we show that log-concavity is preserved along the trajectory $\mathbf{z}(t)$ as long as $\mathbf{z}(t)$ is defined. This observation is the key new idea in this paper.

Theorem 4. *The region $\mathcal{C}$ is invariant with respect to the system of differential equations (2). To be precise, if $\mathbf{z}(0) \in \mathcal{C}$ where $\mathbf{z}(t)$ is the solution to (2) then $\mathbf{z}(t) \in \mathcal{C}$ for all $t > 0$ (for which $\mathbf{z}(t) \in \mathcal{D}$ is defined).*

Proof. It suffices to show that if $\mathbf{z} \in \mathcal{D}$, $x + y = a + b$ and $x < a \leq b < y$ and

$$rz_r s z_s \geq u z_u v z_v \text{ for all } u < r \leq s < v \text{ such that } r + s = u + v > x + y = a + b \tag{15}$$

then

$$\frac{d}{dt}\left(az_a bz_b - xz_x yz_y\right) \geq -(a+b)\left(\frac{2\Delta}{\mu(\mathbf{z})}\right)\left(az_a bz_b - xz_x yz_y\right). \tag{16}$$

To see this, we suppose (15) implies (16) and prove that the inequalities in (12) remain valid along the trajectory $\mathbf{z}(t)$ by (reverse) induction on $x + y = a + b$. Our base case is $x = \Delta - 2, a = b = \Delta - 1, y = \Delta$, where (15) is vacuously true. Here we have

$$\begin{aligned} &az_a(t)bz_b(t) - xz_x(t)yz_y(t) \geq \\ &\qquad\qquad [az_a(0)bz_b(0) - xz_x(0)yz_y(0)] \exp\left\{-(a+b)\int_0^t \frac{2\Delta}{\mu(\mathbf{z}(s))}ds\right\} \geq 0 \end{aligned}$$

as long as $\mathbf{z}(t)$ is defined by Gronwall's inequality. As each step in the induction provides condition (15) for the next step, the remainder of the induction follows similarly.

It remains to prove that (15) implies (16). We have

$$\begin{aligned} \frac{d}{dt}az_a bz_b &= -(a+b)\frac{az_a bz_b}{\mu(\mathbf{z})}\left(\frac{2}{1+\varphi(\mathbf{z})}\right)\left[1 + \sum_{j=2}^{\Delta}\frac{jz_j}{\mu(\mathbf{z})}(j-1)\right] \\ &\qquad + \frac{(a(a+1)z_{a+1}bz_b + baz_a(b+1)z_{b+1}}{\mu(\mathbf{z})}\left(\frac{2}{1+\varphi(\mathbf{z})}\sum_{j=2}^{\Delta}\frac{jz_j}{\mu(\mathbf{z})}(j-1)\right) \\ &= C_1(\mathbf{z})\left[-(a+b)az_a bz_b\right] + C_2(\mathbf{z})\left[a(a+1)z_{a+1}bz_b + baz_a(b+1)z_{b+1}\right] \end{aligned}$$

where $0 < C_1(\mathbf{z}) \leq 2\Delta/\mu(\mathbf{z})$ and $0 < C_2(\mathbf{z})$ and $C_1(\mathbf{z}), C_2(\mathbf{z})$ do not depend on a, b. Since $a + b = x + y$ we have

$$\begin{aligned} &\frac{d}{dt}\left(az_a bz_b - xz_x yz_y\right) = -C_1(\mathbf{z})(a+b)\left(az_a bz_b - xz_x yz_y\right) \\ &\qquad + C_2(\mathbf{z})\left[a(a+1)z_{a+1}bz_b + baz_a(b+1)z_{b+1} - x(x+1)z_{x+1}yz_y - yxz_x(y+1)z_{y+1}\right]. \end{aligned}$$

Noting that (15) implies

$$\begin{aligned} a(a+1)z_{a+1}bz_b + baz_a(b+1)z_{b+1} &\geq (a+b)az_a(b+1)z_{b+1} \\ &\geq (x+y)(x+1)z_{x+1}yz_y \\ &\geq x(x+1)z_{x+1}yz_y + yxz_x(y+1)z_{y+1}, \end{aligned}$$

we see that (16) follows. □

We are now ready to prove Theorem 1. Let $\mathbf{a} \in \mathcal{C}$ and let $\mathbf{z}(t)$ be the solution of the system of differential equations (2) with initial condition $\mathbf{z}(0) = \mathbf{a}$. It follows from Theorems 3 and 4 that the point where $\mathbf{z}(t)$ leaves the region $\mathcal{D}$ is along the 'edge'

$$\left\{\mathbf{x} \in \mathbb{R}_+^\Delta : x_1 = x_3 = \cdots = x_\Delta = 0\right\}.$$

Since, by (9), no variable goes to zero until the sum of the variables goes to zero, it follows that $\mathbf{z}(t)$ leaves $\mathcal{D}$ at the origin.

It only remains for us to deduce Theorem 1 from this understanding of the trajectory $\mathbf{z}(t)$. Let $\mathbf{Y}(0) = n\mathbf{a}$ and let $\mathbf{Y}(\ell)$ be the degree sequence of the remaining graph after ℓ steps of the Karp-Sipser algorithm. We apply Theorem 2. Let $\sigma = \sigma(n)$ be defined by

$$\mu(\mathbf{z}(\sigma)) = n^{-\delta}$$

where $\delta < 1/(6\Delta^2)$ and δ is less than the ϵ of Theorem 2. To get the condition $\varphi(\mathbf{z}(t)) < 1 - n^{-1/6}$ for $0 < t < \sigma$ we use the fact that $f_\Delta(\mathbf{z}) \geq -2\Delta^2 z_\Delta/\mu(\mathbf{z})$ (which follows from (5)) and (4) to write

$$\frac{dz_\Delta}{dt} \geq -\frac{2\Delta^2 z_\Delta}{\mu(\mathbf{z}(\sigma)) + 4(\sigma - t)} \qquad \text{for} \quad 0 \leq t \leq \sigma.$$

So, for $t \in [0, \sigma]$,

$$z_\Delta(t) \geq z_\Delta(0)\left(\frac{\mu(\mathbf{z}(\sigma)) + 4\sigma - 4t}{\mu(\mathbf{z}(\sigma)) + 4\sigma}\right)^{\Delta^2/2}$$

or

$$z_\Delta(t) = \Omega\left(\mu(\mathbf{z}(\sigma))^{\Delta^2/2}\right),$$

since

$$\mu(\mathbf{z}(\sigma)) + 4\sigma \leq \mu(\mathbf{z}(0)).$$

Since (14) implies $\varphi(\mathbf{z}) \leq 1 - (\Delta z_\Delta/\mu(\mathbf{z}))^2$ for $\mathbf{z} \in \mathcal{C}$, we have the bound

$$\varphi(\mathbf{z}(t)) = 1 - \Omega\left(\mu(\mathbf{z}(\sigma))^{\Delta^2 - 2}\right) = 1 - \Omega\left(n^{-\delta(\Delta^2-2)}\right) = 1 - \Omega\left(n^{-1/6}\right)$$

for $0 \leq t \leq \sigma$. Theorem 1 now follows directly from Theorem 2.

3 Proof of Theorem 2

We follow the proof of Theorem 5.1 in [17]. There are two technical difficulties that force some modification of the proof: (i) the system of differential equations (2) does not model the expected change in a single step of the process and (ii) there is a singularity in $\frac{\partial f_i}{\partial z_j}$ on the boundary of $\mathcal{D}$ (i.e. at the origin).

We begin with a Lemma that we apply inductively to prove Theorem 2. Set $\omega = n^{2/3}$. (This 2/3 in the exponent is an arbitrary but convenient choice; our argument can be made to work for other values of ω.)

Lemma 5. *Suppose the degree sequence $\mathbf{Y} \in \mathbb{N}^\Delta$ has $\varphi(\mathbf{Y}) < 1 - n^{-1/5}$, $\mu(\mathbf{Y}) > n^{26/27} = \omega^{13/9}$ and $Y_1 < (\log n)\max\{\log^9 n, (1 - \varphi(\mathbf{Y}))^{-9/4}\}$. Let $\mathbf{Y}(t)$ be the degree sequence after we apply t steps of the Karp-Sipser algorithm to the graph $\mathbf{Y}$. With probability $1 - O(1/n)$ we have*

1. $|Y_i(\omega) - Y_i - \omega f_i(\mathbf{Y}/n)| \le \omega^{3/4}$ *for* $i = 2, \dots, \Delta$,

2. $Y_1(t) \le 2(\log n) \max\left\{\log^9 n, (1-\varphi(\mathbf{Y}))^{-9/4}\right\}$ *for* $t = 1, \dots, \omega$ *and*

3. $Y_1(\omega) < (\log n) \max\left\{\log^9 n, (1-\varphi(\mathbf{Y}(\omega)))^{-9/4}\right\}$.

Note that we expect the algorithm to spend more time on pendant edges as φ gets closer to 1. So our control on Y_1 should deteriorate as φ grows closer to one, and this is reflected in Lemma 5. The proof of Lemma 5 is given in Section 3.1.

Let $0 = k_1 < \cdots < k_N = \sigma n$ be a sequence of integers such that $k_{\ell+1} - k_\ell = \omega$ for $\ell = 1, \dots, N-2$ and $k_N - k_{N-1} < \omega$. We bound the difference between $\mathbf{z}$ and $\mathbf{Y}$ (appropriately scaled) at each of these time steps by induction on ℓ by iterative application of Lemma 5. The application of Lemma 5 to the random graph with degree sequence $\mathbf{Y}(k_\ell)$ requires 3 conditions: a condition on $\varphi(\mathbf{Y}(k_\ell))$, a condition on $\mu(\mathbf{Y}(k_\ell))$ and a condition on $Y_1(k_\ell)$. It is easy to see that $\mathbf{Y}(k_1)$ satisfies these three conditions. The condition on Y_1 needed for subsequent application of the Lemma 5 is guaranteed by conclusion 3 of Lemma 5. We will see below that the conditions on $\varphi(\mathbf{Y}(k_\ell))$ and $\mu(\mathbf{Y}(k_\ell))$ follow from the control on $n\mathbf{Y}(k_\ell) - \mathbf{z}(k_\ell/n)$ we are able to maintain and the condition on $\mathbf{z}(t)$ for $t \in [0, \sigma]$ assumed in Theorem 2 . So, for the moment, we assume that these conditions are satisfied and calculate the control on $\mathbf{Y}$ that follows from Lemma 5. We then justify the inductive application of said assumption from the control $\mathbf{Y}$ that we prove. Also note that, as the number steps in our induction is $o(n)$ and the probability of an error in any step is $O(1/n)$, we can conclude that this description of the evolution indeed holds **whp** (in fact, the argument given here can – with some minor adjustments – be used to show that the probability that the description of the evolution given in Theorem 2 does not hold is at most n^{-c} for any constant c.)

Define

$$e_i(k_\ell) = |Y_i(k_\ell) - nz_i(k_\ell/n)| \qquad \text{for } i = 2, \dots, \Delta$$

We have

$$e_i(k_{\ell+1}) = |Y_i(k_{\ell+1}) - nz_i(k_{\ell+1}/n)| \le A_1 + A_2 + A_3 + A_4$$

where

$$\begin{aligned}
A_1 &= |Y_i(k_\ell) - nz_i(k_\ell/n)| = e_i(k_\ell) \\
A_2 &= |Y_i(k_{\ell+1}) - Y_i(k_\ell) - \omega f_i(\mathbf{Y}(k_\ell)/n)| \\
A_3 &= |\omega z_i'(k_\ell/n) - n\left(z_i(k_{\ell+1}/n) - z_i(k_\ell/n)\right)| \\
A_4 &= |\omega f_i(\mathbf{Y}(k_\ell)/n) - \omega z_i'(k_\ell/n)|\,.
\end{aligned}$$

Note that we can bound $A_2 \le \omega^{3/4}$ with an application of Lemma 5. We use the Lipschitz constant, as given by (3), to bound A_3 and A_4. We have

$$A_3 = \left|\frac{\omega^2}{2n} z_i''(s)\right|$$

for some value s such that $k_\ell/n \le s \le k_{\ell+1}/n$. Therefore, since $d\mu(\mathbf{z}(t))/dt \ge -4\Delta$ (from (4)), we have

$$A_3 \le \frac{\omega^2}{n} \frac{M}{\mu(\mathbf{z}(k_\ell/n)) - 4\Delta\omega/n}$$

for some absolute constant M. (Note that $\mu(\mathbf{z}(k_\ell/n)) > n^{-\epsilon} \gg n^{-1/3}$ so long as $\epsilon < 1/3$.) Furthermore, it follows from (3) that

$$\begin{aligned} A_4 &= |\omega\left(f_i(\mathbf{Y}(k_\ell)/n) - f_i(\mathbf{z}(k_\ell/n))\right)| \\ &= \omega\left|\sum_{j=2}^{\Delta}\frac{\partial f_i}{\partial z_j}(\xi_j)(\mathbf{Y}_j(k_\ell)/n - \mathbf{z}_j(k_\ell/n))\right| \\ &\leq \omega\left(\frac{L}{\mu(\mathbf{z}(k_\ell/n)) - \mu(\mathbf{e}(k_\ell))/n}\right)\frac{\|\mathbf{e}(k_\ell)\|_\infty}{n} \end{aligned}$$

where $\xi_j \in [\mathbf{Y}_j(k_\ell)/n, \mathbf{z}_j(k_\ell/n)]$ and L is an absolute constant.

Now we combine the bounds on A_1, A_2, A_3 and A_4 to bound $e_i(k_{\ell+1})$ itself. Of course, $\mathbf{e}(k_1) = \mathbf{0}$. We add the inductive assumption (which we also justify below) that

$$\frac{\mu(\mathbf{e}(k_\ell))}{n}, \frac{4\Delta\omega}{n} < \frac{1}{2}\mu(\mathbf{z}(k_\ell/n)). \tag{17}$$

We have for $\ell \geq 1$

$$\|\mathbf{e}(k_{\ell+1})\|_\infty \leq \|\mathbf{e}(k_\ell)\|_\infty\left(1 + \frac{\omega}{n}\cdot\frac{2L}{\mu(\mathbf{z}(k_\ell/n))}\right) + \left(\omega^{3/4} + \frac{\omega^2}{n}\frac{2M}{\mu(\mathbf{z}(k_\ell/n))}\right). \tag{18}$$

Let

$$\alpha_j = 2L/\mu(\mathbf{z}(k_j/n)) \quad \text{and} \quad \beta = 2M/\mu(\mathbf{z}(k_N/n)).$$

By iterative application of (18) we have

$$\|\mathbf{e}(k_r)\|_\infty \leq \left(\omega^{3/4} + \frac{\omega^2}{n}\beta\right) + \sum_{i=1}^{r-1}\left[\prod_{j=r-i}^{r-1}\left(1 + \frac{\omega}{n}\alpha_j\right)\right]\left(\omega^{3/4} + \frac{\omega^2}{n}\beta\right) \tag{19}$$

for $r = 2, \ldots N$. Now, it follows from (4) that $\mu(\mathbf{z}(t))$ is decreasing with derivative bounded above by -4. Therefore, letting $\mu_{\text{final}} = \mu(\mathbf{z}(k_N/n)) = \mu(\mathbf{z}(\sigma))$, we have $\mu(\mathbf{z}(t)) \geq \mu_{\text{final}} + 4(\sigma - t)$. So, we have the bound

$$\begin{aligned} \prod_{j=r-i}^{r-1}\left(1 + \frac{\omega}{n}\alpha_j\right) &\leq \prod_{j=r-i}^{r-1}\left(1 + \frac{2L\omega}{n}\cdot\frac{1}{\mu_{\text{final}} + 4(\sigma - k_j/n)}\right) \\ &\leq \exp\left\{\frac{2L\omega}{n}\sum_{j=1}^{N-1}\frac{1}{\mu_{\text{final}} + 4(\sigma - k_j/n)}\right\} \\ &\leq \exp\left\{\frac{2L\omega}{n}\cdot\frac{n}{4\omega}\int_{\mu_{\text{final}}}^{\mu_{\text{final}}+4\sigma}\frac{dx}{x}\right\} \\ &= O\left(\exp\left\{-L\ln\left(\mu_{\text{final}}\right)\right\}\right) \\ &= O\left(\mu_{\text{final}}^{-L}\right). \end{aligned}$$

Plugging this estimate into (19) and recalling that $r \leq \sigma n/\omega$ and $\mu_{\text{final}} \geq n^{-\epsilon}$ we have

$$\begin{aligned} \|\mathbf{e}(k_r)\|_\infty &\leq \left(\omega^{3/4} + \frac{\omega^2}{n}\beta\right)\left(1 + \frac{\sigma n}{\omega}O\left(\mu_{\text{final}}^{-L}\right)\right) \\ &= O\left(\sigma n\omega^{-1/4}\mu_{\text{final}}^{-L}\right) + O\left(2M\omega\sigma\mu_{\text{final}}^{-L-1}\right) = O\left(n^{5/6+(L+1)\epsilon}\right). \end{aligned} \tag{20}$$

We assume that

$$\epsilon < 1/(30(L+2)).$$

Given this choice of ϵ, we see that (20) implies (11).

It remains to justify the assumption (17) and the assumption that $\varphi(\mathbf{Y})$ and $\mu(\mathbf{Y})$ satisfy the conditions required for the application of Lemma 5. Note that, since $\mu(\mathbf{x}) \le \Delta^2 \|\mathbf{x}\|_\infty$ and $\mu(\mathbf{z}(k_r/n)) > n^{-\epsilon}$ (by assumption (10)), (20) inductively justifies the assumption (17) above as well as the conditions required to apply Lemma 5 at each step. Indeed, we have

$$\mu(\mathbf{Y}(k_r)) \ge n\mu(\mathbf{z}(k_r/n)) - \Delta \|\mathbf{e}(k_r)\|_\infty \ge n\mu(\mathbf{z}(k_r/n)) - O\left(n^{5/6+(L+1)\epsilon}\right) \tag{21}$$

and

$$\varphi(\mathbf{Y}(k_r)) \le \varphi(\mathbf{z}(k_r/n)) + O\left(\frac{\|\mathbf{e}(k_r)\|_\infty}{\mu(\mathbf{Y}(k_r))}\right) \le 1 - n^{-1/6} + O\left(n^{-1/6+(L+2)\epsilon}\right). \tag{22}$$

Finally, note that when we apply a step of the Karp-Sipser algorithm to the random graph given by $\mathbf{Y}$ the probability that we isolate a vertex is at most $2\Delta Y_1/\mu(\mathbf{Y})$. Thus, if Y_1 is bounded by $n^{1/2}$ (as given by Lemma 5) then the expected number of isolated vertices is bounded above by

$$\sum_{i=1}^{\sigma n} \frac{\Delta n^{1/2}}{n^{1-\epsilon} + 4i} \le \Delta n^{1/2} \sum_{j=n^{1-\epsilon}}^{5\sigma n} \frac{1}{j} \le \Delta (\log n) n^{1/2}.$$

An application of Markov's inequality then gives the desired result.

3.1 Proof of Lemma 5

The idea here is to break the ω steps of the algorithm in question into a collection of simple moves, apply Chernoff bounds to sums of simple moves, and use the fact that the outcomes of the sums are close to their expected values to conclude that the aggregate behavior is close to the expected value. To this end we define some auxiliary random variables.

We view the Karp-Sipser algorithm running on a configuration model as a sequence of two types of moves: selections and deletions. A **selection move** is simply choosing a configuration point (a.k.a. half edge) for an edge in the matching. When we choose a random pendant edge we make one selection move. When we choose a purely random edge we make two selection moves. After we make a selection move we delete some number of randomly chosen configuration points for the remaining edges incident to the selected vertex. These are **deletion moves**.

Set $\mu = \mu(\mathbf{Y})$. Let $S_i(s)$ be the indicator variable for the event that the s^{th} selection is a configuration point corresponding to a vertex of degree i. Similarly, let $D_i(s)$ the indicator variable for the event that the s^{th} deletion is a configuration point corresponding to a vertex of degree i. Define

$$p_1 = \frac{iY_i - 2\Delta\omega}{\mu - 2\omega} \qquad p_2 = \frac{iY_i + 2\Delta\omega}{\mu - 2\Delta\omega}.$$

Note that in a rounds of the Karp-Sipser algorithm at most $2\Delta a$ and at least $2a$ configuration points are removed. It follows that $S_i(s)$ and $D_i(s)$, conditioned on the configuration given by the previous moves, are Bernoulli random variables with success probability bounded between p_1 and p_2. Therefore, the probability that there exists a degree $i \geq 1$ and moves $0 < s_1 < s_2$ such that $s_2 - s_1 > \log^{10} n$ and

$$\left| \sum_{r=s_1}^{s_2} S_i(r) - (s_2 - s_1)\frac{iY_i}{\mu} \right| \geq (s_2 - s_1)^{5/9} + (s_2 - s_1)(p_2 - p_1)$$

is bounded above (by way of an application of the Chernoff bounds

$$\Pr(Bi(N,p) - Np \leq -a) \leq e^{-a^2/(2Np)} \qquad for\ a \geq 0. \tag{23}$$

$$\Pr(Bi(N,p) - Np \geq a) \leq e^{-2a^2/(Np)} \qquad for\ a \geq 0. \tag{24}$$

to $Bi(s_2 - s_1, p_1)$ and $Bi(s_2 - s_1, p_2)$) by

$$2\Delta^2\omega \sum_{k=\log^{10} n}^{2\Delta\omega} \left[\exp\{-k^{10/9}/(2kp_1)\} + \exp\{-2k^{10/9}/(kp_2)\}\right] = O(1/n).$$

The same statement holds for the D_i's. For the remainder of the proof of this Lemma we assume that these small probability events do not occur. Note that we have $p_2 - p_1 = O(\omega/\mu)$ and the condition on $\mu(\mathbf{Y})$ ensures $(s_2-s_1)\omega/\mu < (s_2-s_1)^{5/9}$. Therefore, we may henceforth assume

$$\left| \sum_{r=s_1}^{s_2} S_i(r) - (s_2 - s_1)\frac{iY_i}{\mu} \right| = O\left((s_2 - s_1)^{5/9}\right), \tag{25}$$

and by a similar argument

$$\left| \sum_{r=s_1}^{s_2} D_i(r) - (s_2 - s_1)\frac{iY_i}{\mu} \right| = O\left((s_2 - s_1)^{5/9}\right). \tag{26}$$

for all $0 < s_1 < s_2$ such that $s_2 - s_1 > \log^{10} n$.

We now consider the evolution of the number of vertices of degree 1 during the ω steps of Karp-Sipser we consider here. Define

$$K = \log n \cdot \max\left\{\log^9 n, (1 - \varphi(\mathbf{Y}))^{-9/4}\right\}.$$

Note that the condition on $\varphi(\mathbf{Y})$ implies $K < (\log n)n^{9/20} \ll \omega$. Suppose there is a sequence of $k \geq K$ consecutive steps of Karp-Sipser such that the algorithm chooses a pendant edge on every step in this run. Note that during this span of steps of Karp-Sipser there is exactly one selection move per step of the algorithm. During these steps we perform $(i-1)$ deletion moves each time we select a vertex of degree i. Therefore, the number of deletion moves performed is, by (25),

$$k\sum_{i=2}^{\Delta} \frac{iY_i(i-1)}{\mu} + O(k^{5/9}).$$

Each of these deletion moves creates a new pendant edge when it deletes a configuration point corresponding to a vertex of degree 2. So, by an application of (26), the number of pendant edges created during these k steps is at most

$$\left(k\sum_{i=2}^{\Delta} \frac{i(i-1)Y_i}{\mu} + O(k^{5/9})\right)\frac{2Y_2}{\mu} + O(k^{5/9}) = k\varphi(\mathbf{Y}) + O(k^{5/9}). \tag{27}$$

We first apply this observation to show that the number of vertices of degree 1 drops to zero during the ω steps of the algorithm we consider here. Under the assumption that in all ω moves the algorithm adds a pendant edge to the matching, we set $k = \omega$ and see that the number of vertices of degree 1 produced is at most

$$\omega\varphi(\mathbf{Y}) + O(\omega^{5/9}).$$

This number is less than $\omega - K$ (for n sufficiently large) by the assumption in the Lemma that $\varphi(\mathbf{Y}) < 1 - n^{-1/5}$. In other words, the number of vertices of degree 1 at the beginning (i.e. Y_1) plus the number of vertices of degree 1 produced in the ω rounds is less than ω. But this is a contradiction as the algorithm processes a pendant edge in each step. Next we show that $Y_1(t)$ is never too large. Note that the number of degree 1 vertices produced in K consecutive steps in which Karp-Sipser chooses pendant edges is at most

$$K\varphi(\mathbf{Y}) + O(K^{5/9}) < K.$$

We conclude that the number of degree 1 vertices at the end of such a run is at most the number of degree 1 vertices at the beginning. This implies that the number of vertices of degree 1 never exceeds $2K$, say. (Note that we have established condition 2 of Lemma 5.) But we can say a little more. Suppose we arrive at a step at which there are zero vertices of degree 1. And suppose that in the ensuing K rounds we see the worst-case K new degree 1 vertices. Note that the maximum number of new degree 1 vertices created in a single round in which Karp-Sipser adds a pendant edge is $\Delta - 1$. So it must take at least $K/(\Delta - 1)$ steps for the new degree 1 vertices to appear. Thus, as Karp-Sipser is also processing pendant edges during these rounds, we see that the number of vertices of degree 1 does not exceed

$$\frac{K}{\Delta - 1}(\Delta - 2) = K\left(1 - \frac{1}{\Delta - 1}\right)$$

during these rounds. In particular, since we have already shown that the number of vertices of degree 1 drops to zero during our ω rounds, the number of degree 1 vertices at the end of the ω rounds is at most

$$K\left(1 - \frac{1}{\Delta - 1}\right).$$

As

$$1 - \varphi(\mathbf{Y}(\omega)) = 1 - \varphi(\mathbf{Y}) + O\left(\frac{\omega}{\mu(\mathbf{Y})}\right)$$
$$\leq 1 - \varphi(\mathbf{Y}) + O\left(\frac{n^{2/3}}{n^{26/27}}\right) = (1 - \varphi(\mathbf{Y}))(1 + o(1)),$$

we have established Condition 3 in Lemma 5.

It remains to prove Condition 1. We begin by studying the proportion of each type of step Karp-Sipser makes. To this end, we first note – applying the assumption that the low probability events discussed above do not occur – that the algorithm selects or deletes a configuration point corresponding to a vertex of degree 1 at most K times among the at most $2\Delta\omega$ selection/deletion moves in the ω steps of Karp-Sipser we consider here. Let $P(t)$

be the number of vertices of degree 1 created through the first t edges added to the matching by the Karp-Sipser algorithm. Let $R(t)$ be the number of deletion moves made through the first t steps of Karp-Sipser. Since new vertices of degree 1 are only produced by deletion moves, we have

$$P(t) = \sum_{s=1}^{R(t)} D_2(s).$$

If $t > K$ then (26) implies

$$P(t) = R(t)\frac{2Y_2}{\mu} + O(R(t)^{5/9}). \tag{28}$$

Note that if $Y_1 + P(t) \leq t$ then Karp-Sipser adds at least one purely random edge (rather than pendant edge) during the first step $t+1$ steps. Indeed if there exists a step $r < t$ such that $Y_1 + P(r) \leq r$ then – applying the observations from the previous paragraph – $Y_1(t) < K$ and the number of pendant edges added during the first t steps is between

$$Y_1 + P(t) - 2K \qquad \text{and} \qquad Y_1 + P(t).$$

(We subtract $2K$ in the lower bound to account for vertices of degree 1 that might exist at step t and vertices of degree 1 that the algorithm might have selected or deleted in the first t steps of Karp-Sipser.) Let $T(t)$ be the number of selection moves through t steps of the process. Since we make 2 selection moves when Karp-Sipser adds a purely random edge and only 1 selection move when Karp-Sipser adds a pendant edge, it follows that if t is sufficiently large then

$$2t - (Y_1 + P(t)) \leq T(t) \leq 2t - (Y_1 + P(t) - 2K) \tag{29}$$

and

$$\sum_{s=1}^{2t-P(t)-K} \sum_{i=2}^{\Delta} (i-1)S_i(s) \leq R(t) \leq \sum_{s=1}^{2t-P(t)+2K} \sum_{i=2}^{\Delta} (i-1)S_i(s).$$

We see that, by (25), for t sufficiently large we have

$$R(t) = (2t - P(t)) \sum_{i=2}^{\Delta} \frac{i(i-1)Y_i}{\mu} + O\left((2t-P(t))^{5/9}\right) + O(K). \tag{30}$$

Putting (28) and (30) together (with $t = \omega$ and $K \leq n^{9/20}\log n \leq n^{19/40}$) yields

$$\begin{aligned} P(\omega) &= R(\omega)\frac{2Y_2}{\mu} + O(\omega^{5/9}) \\ &= \left[(2\omega - P(\omega))\sum_{i=2}^{\Delta}\frac{i(i-1)Y_i}{\mu} + O\left(\omega^{5/9}\right) + O(K)\right]\frac{2Y_2}{\mu} + O(\omega^{5/9}) \\ &= (2\omega - P(\omega))\varphi(\mathbf{Y}) + O(n^{19/40}). \end{aligned}$$

and therefore,

$$P(\omega) = \omega \cdot \frac{2\varphi(\mathbf{Y})}{1+\varphi(\mathbf{Y})} + O\left(n^{19/40}\right).$$

It now follows from (29) that we have

$$T(\omega) = 2\omega - P(\omega) + O(K) = \omega \cdot \frac{2}{1+\varphi(\mathbf{Y})} + O(n^{19/40}).$$

(It should come as no surprise that our calculations now closely parallel the heuristic discussion that lead to the system of equations (2).) The number of deletion moves in these ω steps is

$$R(\omega) = (2\omega - P(\omega))\sum_{i=2}^{\Delta} \frac{i(i-1)Y_i}{\mu} + O(n^{19/40}) = \omega \cdot \frac{2}{1+\varphi(\mathbf{Y})} \cdot \sum_{i=2}^{\Delta} \frac{i(i-1)Y_i}{\mu} + O(n^{19/40})$$

As

$$Y_i(\omega) = Y_i - \sum_{s=1}^{T(\omega)} S_i(s) - \sum_{s=1}^{R(\omega)} D_i(s) + \sum_{s=1}^{R(\omega)} D_{i+1}(s),$$

we have the desired bound from an application of (25) and (26).

4 Conclusion

Theorem 1 can be extended to apply to degree distributions $\mathbf{z}$ with $z_1 > 0$.

Theorem 6. *Let $\mathbf{z} \in \mathbb{R}_+^{\Delta}$ be a degree distribution with $z_1 > 0$ and $z_2, \ldots, z_\Delta$ strongly log-concave.* **Whp** *the Karp-Sipser algorithm applied to $G_{\mathbf{z}}$ produces a matching with cardinality within $n^{1-\epsilon}$ of the maximum matching in $G_{\mathbf{z}}$ for some constant $\epsilon = \epsilon(\mathbf{z}, \Delta)$.*

The proof follows from noting that the region

$$\mathcal{C}' = \left\{\mathbf{z} \in \mathbb{R}_+^{\Delta} : z_2, \ldots, z_\Delta \text{ is strongly log-concave}\right\}$$

is invariant with respect to the system of differential equations that governs the phase of the Karp-Sipser algorithm that spends *all* of its time adding pendant edges to the matching (the proof of this fact is very similar to the proof of Theorem 4). Using the fact that no 'mistakes' are made during this phase, we see that once the Karp-Sipser algorithm on $G_{\mathbf{z}}$ gets to a distribution where there are no vertices of degree 1, we can apply Theorem 1 directly to the remaining graph. Note that Theorem 6 is a kind of generalization of the analysis of the Karp-Sipser algorithm on the random graph $G_{n,cn}$ given by Karp and Sipser and Aronson, Frieze and Pittel. The Karp-Sipser algorithm on $G_{n,cn}$ has two distinct phases: in the first phase pendant edges are removed, and if anything remains after the first phase then the remainder of evolution of the Karp-Sipser algorithm is a phase in which pendant edges do not significantly accumulate.

Of course, Theorem 1 begs the question: which distributions $\mathbf{z}$ have the property that **whp** $G_{\mathbf{z}}$ has a perfect (or almost perfect) matching? This is a natural question from the perspective of the theory of random graphs. The existence of a perfect matching is a graph property for which there is a so-called minimum degree phenomenon: If p is large enough so that **whp** the random graph $G_{n,p}$ has minimum degree 1 then **whp** $G_{n,p}$ has a perfect matching (proved by Erdős and Rényi [7]). Furthermore, long before a perfect matching actually appears **whp** $G_{n,p}$ has a matching that saturates all non-isolated vertices. So, it is natural to ask for the number of edges needed for a perfect matchings in random graphs when we impose the condition that the minimum degree is at least 1. This problem was solved for $G_{n,p}$ by Bollobás and Frieze [4]: a super-linear number of edges are still needed as the barrier to the existence of a perfect matching is now a pair of vertices of degree 1 that have a common neighbor. Since a random graph with a significant number of vertices of degree 1 will likely have such a pair, if we are to find perfect matchings in random graphs with a linear number of edges we must ensure that there are no (or very few) vertices of degree 1. So, for sparse random graphs it is natural to condition on minimum degree 2. The matching question has been resolved for a handful of models of sparse random graphs with such a minimum degree condition: the random regular graph, $G_{n,cn}$ where we condition on minimum degree at least 2, the model $G_{k-\text{out}}$ with $k \geq 2$ (in $G_{k-\text{out}}$ every vertex independently chooses k neighbors at random, producing a graph with average degree $2k$ and minimum degree k), and certain natural variations on $G_{k-\text{out}}$ [3] [8] [9] [10] [12] [18].

It seems natural to suspect that if the Karp-Sipser algorithm **whp** produces an almost-perfect matching in $G_{\mathbf{z}}$ then **whp** $G_{\mathbf{z}}$ has a perfect matching, up to the odd cycle components which will appear in $G_{\mathbf{z}}$ if $z_2 > 0$. We conjecture that this is indeed the case for the distributions considered in Theorem 1.

Conjecture 7. *If $\Delta \geq 3$ and $\mathbf{z} \in \mathbb{R}_+^{\Delta}$ such that $z_1 = 0$, $z_\Delta > 0$ and $\mathbf{z}$ satisfies (1) then* **whp** *$G_{\mathbf{z}}$ has a matching with defect equal to the number of odd cycle components in $G_{\mathbf{z}}$.*

In a companion paper the authors prove Conjecture 7 for a significant subset of the degree distributions that satisfy (1). Theorem 1 is a key part of that proof. The full Conjecture 7 remains open.

At first glance, it is tempting to think that $G_{\mathbf{z}}$ has a matching with defect equal to the number of odd cycle components **whp** whenever $z_1 = 0$. We now illustrate by example that this is not the case. Consider the distribution $\mathbf{z} \in \mathbb{R}^6$ defined by

$$z_i = \begin{cases} m & \text{if } i = 2 \\ 1 & \text{if } i = 6 \\ 0 & \text{if } i \notin \{2, 6\} \end{cases}$$

where m is some large constant. We can view $G_{\mathbf{z}}$ as a graph consisting of n vertices of degree 6 connected by a collection $\mathcal{P}$ of $3n$ paths of varying lengths where all internal vertices of the paths have degree 2. Note that the lengths of the paths in $\mathcal{P}$ are **whp** approximately distributed as a geometric random variable with mean $m+1$. Let $\mathcal{P}_{\text{even}}$ be the collection of paths $P \in \mathcal{P}$ that have an even number of edges. Note that if m is sufficiently large then **whp** we have $|\mathcal{P}_{\text{even}}| > 5n/4$ as the path lengths are smoothly distributed. (Indeed, the number of paths in $\mathcal{P}$ with $2i$ edges is roughly $m/(m+1)$ times the number of paths in $\mathcal{P}$ with $2i-1$ edges.) Assume for the sake of contradiction that $|\mathcal{P}_{\text{even}}| > 5n/4$ and there is an almost perfect matching M in $G_{\mathbf{z}}$. Consider a path $P \in \mathcal{P}_{\text{even}}$ that connects vertices (of degree 6) u and v. If every internal vertex of P is saturated by M then there must be edge in $M \cap P$ that contains either u or v. Now, since M is almost perfect, nearly all of the paths in $\mathcal{P}_{\text{even}}$ have the property that every internal vertex is saturated. Therefore there are at least $(1-o(1))|\mathcal{P}_{\text{even}}| > 9n/8$ edges in M that contain vertices of degree 6. As there are n vertices of degree 6, this is a contradiction. Mike Picollelli and Mike Molloy independently point out that this example can be extended to show that there exists $G_{\mathbf{z}}$ with no perfect matching **whp** and arbitrarily large minimum degree. To see this, consider $\mathbf{z}$ defined by

$$z_i = \begin{cases} d & \text{if } i = d \\ 1 & \text{if } i = d^3 \\ 0 & \text{if } i \notin \{d, d^3\} \end{cases}$$

for some d large. Then there will be roughly nd/e vertices of degree d whose neighborhoods are contained in the set of n vertices of degree d^3. Clearly, such a graph does not have a perfect matching.

Theorem 1 and Conjecture 7 lead naturally to a number of interesting questions regarding $G_{\mathbf{z}}$ for $\mathbf{z} \in \mathcal{D} \setminus \mathcal{C}$. We mention two of these here.

1. Can we analyze the performance of the Karp-Sipser algorithm on random graphs $G_{\mathbf{z}}$ where $\mathbf{z} \in \mathcal{D} \setminus \mathcal{C}$? Are there distributions $\mathbf{z}$ such that the Karp-Sipser algorithm applied to $G_{\mathbf{z}}$ passes between $\overline{\mathcal{D}}$ and $\mathcal{D}$ an arbitrary number of times?

2. Which distributions $\mathbf{z} \in \mathcal{D}$ have the property that **whp** $G_{\mathbf{z}}$ has a matching with defect equal to the number of odd cycle components?

Acknowledgment. The authors are deeply indebted to the anonymous referees who made a number of observations and suggestions that greatly improved the presentation of this paper.

References

[1] J. Aronson, A. Frieze and B. Pittel, Maximum matchings in random graphs: Karp-Sipser revisited, *Random Structures and Algorithms* **12** (1998), 111-178.

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

[3] B. Bollobás, Random graphs, in *Combinatorics* (H. Temperley editor) London mathematical society lecture note series 52, Cambridge university press (1981) 80-102.

[4] B. Bollobás and A. Frieze, On matchings and hamiltonian cycles in random graphs, *Annals of Discrete Mathematics* **28** (1985), 23-46.

[5] F. Chung, L. Lu, and V. Vu, The spectra of random graphs with given expected degrees, *Proceedings of National Academy of Science* **100** (2003), 6313-6318.

[6] J. Edmonds, Paths, trees and flowers, *Canadian Journal of Mathematics* **17** (1965), 449-467.

[7] P. Erdős and A. Rényi, On the existence of a factor of degree one of a connected random graph, *Acta. Math. Acad. Sci. Hungar.* **17** (1966), 359-368.

[8] A. Frieze, Maximum matchings in a class of random graphs, *Journal of Combinatorial Theory Series B* **40** (1986), 196-212.

[9] A. Frieze, Perfect matchings in random bipartite graphs with minimum degree 2, *Random Structures and Algorithms* **26** (2005), 319-358.

[10] A. Frieze and B. Pittel, Perfect matchings in random graphs with prescribed minimal degree, in *Trends in Mathematics*, Birkhauser Verlag, Basel 95-132.

[11] S. Janson, T. Luckaz and A. Rucinski, *Random Graphs*. John Wiley and Sons, 2000.

[12] M. Karonski and B. Pittel, Random proposals with a second chance for nerds, manuscript.

[13] R. Karp and M. Sipser, Maximum matchings in sparse random graphs, *Proceedings of the 22nd IEEE Symposium on the Foundations of Computing* (1981) 364-375.

[14] S. Micali and V. Vazirani, An $O(|V|^{1/2}|E|)$ algorithm for finding maximum matchings in general graphs. *Proceedings of the 21st IEEE Symposium on the Foundations of Computing* (1980).

[15] M. Molloy and B. Reed, A Critical Point for Random Graphs with a Given Degree Sequence, *Random Structures and Algorithms* **6** (1995), 161-180.

[16] L. Shi and N. Womrald, Colouring random 4-regular graphs, manuscript.

[17] N. Wormald, The Differential Equation Method For Random Graph Processes and Greedy Algorithms, in *Lectures on Approximation and Randomized Algorithms*, M. Karonski and H.J. Promel, editors, 1999, pp. 73–155.

[18] D. Walkup, Matchings in random regular bipartite graphs, *Discrete Mathematics* **31** (1980) 59-64.