On edge disjoint spanning trees in a randomly weighted complete graph

Alan Frieze Tony Johansson
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA 15213
U.S.A.

May 13, 2015

1 Introduction

This paper can be considered to be a contribution to the following general problem. We are given a combinatorial optimization problem where the weights of variables are random. What can be said about the random variable equal to the minimum objective value in this model. The most studied examples of this problem are those of (i) Minimum Spanning Trees e.g. Frieze [10], (ii) Shortest Paths e.g. Janson [18], (iii) Minimum Cost Assignment e.g. Aldous [1], [2], Linusson and Wästlund [22] and Nair, Prabhakar and Sharma [24], Wästlund [31] and (iv) the Travelling Salesperson Problem e.g. Karp [20], Frieze [11] and Wästlund [32].

The minimum spanning tree problem is a special case of the problem of finding a minimum weight basis in an element weighted matroid. Extending the result of [10] has proved to be difficult for other matroids. We are aware of a general result due to Kordecki and Lyczkowska-Hanćkowiak [21] that expresses the expected minimum value of an integral using the Tutte Polynomial. The formulae obtained, although exact, are somewhat difficult to penetrate. In this paper we consider the union of k cycle matroids. We have a fairly simple analysis for $k \to \infty$ and a rather difficult analysis for $k = 2$.

Given a connected simple graph $G = (V, E)$ with edge lengths $\mathbf{x} = (x_e : e \in E)$ and a positive integer k, let $\mathrm{mst}_k(G, \mathbf{x})$ denote the minimum length of k edge disjoint spanning trees of G. ($\mathrm{mst}_k(G) = \infty$ if such trees do not exist.) When $\mathbf{X} = (X_e : e \in E)$ is a family of independent random variables, each uniformly distributed on the interval $[0, 1]$, denote the expected value $\mathbf{E}[\mathrm{mst}_k(G, \mathbf{X})]$ by $\mathrm{mst}_k(G)$.

As previously mentioned, the case $k = 1$ has been the subject of some attention. When G is the complete graph K_n, Frieze [10] proved that

$$\lim_{n\to\infty} \mathrm{mst}_1(K_n) = \zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3}.$$

Generalisations and refinements of this result were subsequently given in Steele [30], Frieze and McDiarmid [13], Janson [17], Penrose [28], Beveridge, Frieze and McDiarmid [4], Frieze, Ruszinko and Thoma [14] and most recently in Cooper, Frieze, Ince, Janson and Spencer [7].

In this paper we discuss the case $k \geq 2$ when $G = K_n$ and define

$$\mu_k^* = \liminf_{n\to\infty} \mathrm{mst}_k(K_n) \text{ and } \mu_k^{**} = \limsup_{n\to\infty} \mathrm{mst}_k(K_n).$$

Conjecture: $\mu_k^* = \mu_k^{**}$ i.e. $\lim_{n\to\infty} \mathrm{mst}_k(K_n)$ exists.

Theorem 1.

(a)

$$\lim_{k\to\infty} \frac{\mu_k^*}{k^2} = \lim_{k\to\infty} \frac{\mu_k^{**}}{k^2} = 1.$$

(b) *With f_k and $c_2' \approx 3.59$ and $\lambda_2' \approx 2.688$ as defined in* (1), (6), (17),

$$\mu_2$$
$$= 2c_2' - \frac{(c_2')^2}{4} + \int_{\lambda=\lambda_2'}^{\infty} \left(2 - \frac{\lambda e^\lambda}{2f_2(\lambda)} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right)\left(\frac{e^\lambda}{f_2(\lambda)} + \frac{\lambda e^\lambda}{f_2(\lambda)} - \frac{\lambda e^\lambda f_1(\lambda)}{f_2(\lambda)^2}\right) d\lambda$$
$$= 4.17042881\ldots$$

There appears to be no clear connetion between μ_2 and the ζ function.

Note also, in connection with Theorem 1(a), that if n is even and $k = (n-1)/2$ and we take a partition of the edge set of K_n into spanning trees then w.h.p. $\mu_k \approx \frac{n^2}{4} \approx k^2$.

Before proceeding to the proof of Theorems 1 we note some properties of the κ-core of a random graph.

2 The κ-core

The functions

$$f_i(\lambda) = \sum_{j=i}^{\infty} \frac{\lambda^j}{j!}, \quad i = 0, 1, 2, \ldots, \tag{1}$$

figure prominently in our calculations. We let

$$g_i(\lambda) = \frac{\lambda f_{2-i}(\lambda)}{f_{3-i}(\lambda)}, \quad i = 0, 1, 2.$$

Properties of these functions are derived in Appendix B.

The *κ-core* $C_\kappa(G)$ of a graph G is the largest set of vertices that induces a graph H_κ such that the minimum degree $\delta(H_\kappa) \geq \kappa$. Pittel, Spencer and Wormald [29] proved that there exist constants, $c_\kappa, \kappa \geq 3$ such that if $p = c/n$ and $c < c_\kappa$ then w.h.p. $G_{n,p}$ has no κ-core and that if $c > c_\kappa$ then w.h.p. $G_{n,p}$ has a κ-core of linear size. We list some facts about these cores that we will need in what follows.

Given λ let $\mathrm{Po}(\lambda)$ be the Poisson random variable with mean λ and let

$$\pi_r(\lambda) = \Pr\left\{\mathrm{Po}(\lambda) \geq r\right\} = e^{-\lambda} f_r(\lambda).$$

Then

$$c_\kappa = \inf\left(\frac{\lambda}{\pi_{\kappa-1}(\lambda)} : \lambda > 0\right). \tag{2}$$

When $c > c_\kappa$ define $\lambda_\kappa(c)$ by

$$\lambda_\kappa(c) \text{ is the larger of the two roots to the equation } c = \frac{\lambda}{\pi_{\kappa-1}(\lambda)} = \frac{\lambda e^\lambda}{f_{\kappa-1}(\lambda)}. \tag{3}$$

Then w.h.p.[1] with $\lambda = \lambda_\kappa(c)$ we have that

$$C_\kappa(G_{n,p}) \text{ has } \approx \pi_\kappa(\lambda)n = \frac{f_\kappa(\lambda)}{e^\lambda}n \text{ vertices and } \approx \frac{\lambda^2}{2c}n = \frac{\lambda f_{\kappa-1}(\lambda)}{2e^\lambda}n \text{ edges.} \tag{4}$$

Furthermore, when κ is large,

$$c_\kappa = \kappa + (\kappa \log \kappa)^{1/2} + O(\log \kappa). \tag{5}$$

Łuczak [23] proved that C_κ is κ-connected w.h.p. when $\kappa \geq 3$.

Next let c'_κ be the threshold for the $(\kappa+1)$-core having average degree 2κ. Here, see (3) and (4),

$$c'_\kappa = \frac{\lambda e^\lambda}{f_\kappa(\lambda)} \text{ where } \frac{\lambda f_k(\lambda)}{f_{k+1}(\lambda)} = 2\kappa. \tag{6}$$

We have $c_2 \approx 3.35$ and $c'_2 \approx 3.59$.

3 Proof of Theorem 1(a): Large k.

We will prove Part (a) of Theorem 1 in this section. It is relatively straightforward. Part (b) is more involved and occupies Section 4.

In this section we assume that $k = O(1)$ and large. Let Z_k denote the sum of the $k(n-1)$ shortest edge lengths in K_n. We have that for $n \gg k$,

$$\text{mst}_k(K_n) \geq \mathbf{E}[Z_k] = \sum_{\ell=1}^{k(n-1)} \frac{\ell}{\binom{n}{2}+1} = \frac{k(n-1)(k(n-1)+1)}{n(n-1)+2} \in [k^2(1-n^{-1}), k^2]. \tag{7}$$

This gives us the lower bound in Theorem 1(a).

For the upper bound let $k_0 = k + k^{2/3}$ and consider the random graph H generated by the $k_0(n-1)$ cheapest edges of K_n. The expected total edge weight $\overline{E}_H$ of H is at most k_0^2, see (7).

H is distributed as $G_{n,k_0 n}$. This is sufficiently close in distribution to $G_{n,p}, p = 2k_0/n$ that we can apply the results of Section 2 without further comment. It follows from (5) that $c_{2k} < 2k_0$. Putting $\lambda_0 = \lambda_{2k}(2k_0)$ we see from (4) that w.h.p. H has a $2k$-core C_{2k} with $\sim n\Pr\{\text{Po}(\lambda_0) \geq 2k\}$ vertices. It follows from (3) that $\lambda_0 = 2k_0 \pi_{2k-1}(2k_0) \leq 2k_0$ and since $\pi_{2k-1}(\lambda)$ increases with λ and $\pi_{2k-1}(2k + k^{2/3}) = \Pr\left\{\text{Po}(2k + k^{2/3}) \geq 2k - 1\right\} \geq 1 - e^{-c_1 k^{1/3}}$ for some constant $c_1 > 0$ we see that $\frac{2k+k^{2/3}}{\pi_{2k-1}(2k+k^{2/3})} \leq 2k_0$ and so $\lambda_0 \geq 2k + k^{2/3}$.

[1] For the purposes of this paper, a sequence of events $\mathcal{E}_n$ will be said to occur *with high probability* w.h.p. if $\Pr\{\mathcal{E}_n\} = 1 - o(n^{-1})$

A theorem of Nash-Williams [25] states that a $2k$-edge connected graph contains k edge-disjoint spanning trees. Applying the result of Łuczak [23] we see that w.h.p. C_{2k} contains k edge disjoint spanning trees $T_1, T_2, \ldots, T_k$. It remains to argue that we can cheaply augment these trees to spanning trees of K_n. Since $|C_{2k}| \sim n\Pr\{\mathrm{Po}(\lambda) \geq 2k\}$ w.h.p., we see that w.h.p. $D_{2k} = [n] \setminus C_{2k}$ satisfies $|D_{2k}| \leq 2ne^{-c_1 k^{1/3}}$.

For each $v \in D_{2k}$ we let S_v be the k shortest edges from v to C_{2k}. We can then add v as a leaf to each of the trees $T_1, T_2, \ldots, T_k$ by using one of these edges. What is the total weight of the edges Y_v, $v \in D_{2k}$? We can bound this probabilistically by using the following lemma from Frieze and Grimmett [12]:

Lemma 1. *Suppose that $k_1+k_2+\cdots+k_M \leq a$, and $Y_1, Y_2, \ldots, Y_M$ are independent random variables with Y_i distributed as the k_ith minimum of N independent uniform [0,1] random variables. If $\mu > 1$ then*

$$Pr\left\{Y_1 + \cdots + Y_M \geq \frac{\mu a}{N+1}\right\} \leq e^{a(1+\ln\mu-\mu)}.$$

Let $\varepsilon = 2e^{-c_1 k^{1/3}}$ and $\mu = 10\ln 1/\varepsilon$ and let $M = k\varepsilon n$, $N = (1-\varepsilon)n$, $a = \frac{k(k+1)}{2}\varepsilon n$. Let $\mathcal{B}$ be the event that there exists a set S of size εn such that the sum of the k shortest edges from each $v \in S$ to $[n] \setminus S$ exceeds $\mu a/(N+1)$. Applying Lemma 1 we see that

$$\Pr\{\mathcal{B}\} \leq \binom{n}{\varepsilon n} \exp\{k(k+1)\varepsilon n(1+\ln\mu-\mu)/2\} \leq \left(\frac{e}{\varepsilon} \cdot e^{-\mu k^2/3}\right)^{\varepsilon n} = o(n^{-1}).$$

It follows that

$$\mathrm{mst}_k(K_n) \leq o(1) + k_0^2 + \frac{\mu a}{N+1} \leq k^2 + 3k^{5/3}.$$

The $o(1)$ term is a bound $kn \times o(n^{-1})$, to account for the cases that occur with probability $o(n^{-1})$.

Combining this with (7) we see that

$$k^2 \leq \mu_k \leq k^2 + 3k^{5/3}$$

which proves Theorem 1(a).

4 Proof of Theorem 1(b): $k = 2$.

For this case we use the fact that for any graph $G = (V, E)$, the collection of subsets $I \subseteq E$ that can be partitioned into two edge disjoint forests form the independent sets in a matroid. This being the matroid which is the union of two copies of the cycle matroid of G. See for example Oxley [27] or Welsh [33]. Let r_2 denote the rank function of this matroid, when $G = K_n$. If G is a sub-graph of K_n then $r_2(G)$ is the rank of its edge-set.

We will follow the proof method in [3], [4] and [17]. Let F denote the random set of edges in the minimum weight pair of edge disjoint spanning trees. For any $0 \leq p \leq 1$ let G_p denote the graph induced by the edges e of K_n which satisfy $X_e \leq p$. Note that G_p is distributed as $G_{n,p}$.

For any $0 \leq p \leq 1$, $\sum_{e\in F} 1_{(X_e>p)}$ is the number of edges of F which are not in G_p, which equals $2n - 2 - r_2(G_p)$. So,

$$\mathrm{mst}_2(K_n, \mathbf{X}) = \sum_{e\in F} X_e = \sum_{e\in F} \int_{p=0}^{1} 1_{(X_e>p)} dp = \int_{p=0}^{1} \sum_{e\in F} 1_{(X_e>p)} dp.$$

Hence, on taking expectations we obtain

$$\mathrm{mst}_2(K_n) = \int_{p=0}^{1} (2n-2-\mathbf{E}\left[r_2(G_p)\right])dp. \tag{8}$$

It remains to estimate $\mathbf{E}\left[r_2(G_p)\right]$. The main contribution to the integral in (8) comes from $p = c/n$ where c is constant. Estimating $\mathbf{E}\left[r_2(G_p)\right]$ is easy enough for sufficiently small c, but it becomes more difficult for $c > c_2'$, see (6). When $p = \frac{c}{n}$ for $c > c_k$ we will need to be able to estimate $\mathbf{E}\left[r_k(C_{k+1}(G_{n,p}))\right]$. We give partial results for $k \geq 3$ and complete results for $k = 2$. We begin with a simple observation.

Lemma 2. *Let $C_{k+1} = C_{k+1}(G)$ denote the graph induced by the $(k+1)$-core of graph G (it may be an empty sub-graph). Let $E_k(G)$ denote the set of edges that are* **not** *contained in C_{k+1}. Then*

$$r_k(G) = |E_k(G)| + r_k(C_{k+1}). \tag{9}$$

Proof. By induction on $|V(G)|$. Trivial if $|V(G)| = 1$ and so assume that $|V(G)| > 1$. If $\delta(G) \geq k+1$ then $G = C_{k+1}$ and there is nothing to prove. Otherwise, G contains a vertex v of degree $d_G(v) \leq k$. Now $G - v$ has the same $(k+1)$-core as G. If $F_1, ..., F_k$ are edge disjoint forests such that $r_k(G) = |F_1| + ... + |F_k|$ then by removing v we see, inductively, that $|E_k(G-v)| + r_k(C_{k+1}) = r_k(G-v) \geq |F_1| + ... + |F_k| - d_G(v) = r_k(G) - d_G(v)$. On the other hand $G - v$ contains k forests $F_1', ..., F_k'$ such that $r_k(G-v) = |F_1'| + ... + |F_k'| = |E_k(G-v)| + r_k(C_{k+1})$. We can then add v as a vertex of degree one to $d_G(v)$ of the forests $F_1', ..., F_k'$, implying that $r_k(G) \geq d_G(v) + |E_k(G-v)| + r_k(C_{k+1})$. Thus, $r_k(G) = d_G(v) + |E_k(G-v)| + r_k(C_{k+1}) = |E_k(G)| + r_k(C_{k+1})$. □

Lemma 3. *Let $k \geq 2$. If $c_k < c < c_k'$, then w.h.p.*

$$|E(G_{n,c/n})| - o(n) \leq r_k(G_{n,c/n}) = |E(G_{n,c/n})|. \tag{10}$$

Proof. We will show that when $c < c_k'$ we can find k disjoint forests $F_1, F_2, \ldots, F_k$ contained in C_{k+1} such that

$$|E(C_{k+1})| - \sum_{i=1}^{k} |E(F_i)| = o(n). \tag{11}$$

This implies that $r_k(C_{k+1}) \geq |E(C_{k+1})| - o(n)$ and because $r_k(C_{k+1}) \leq |E(C_{k+1})|$ the lemma follows from this and Lemma 2.

Gao, Pérez-Giménez and Sato [15] show that when $c < c_k'$, no subgraph of $G_{n,p}$ has average degree more than $2k$, w.h.p. Fix $\varepsilon > 0$. Cain, Sanders and Wormald [6] proved that if the average degree of the $(k+1)$-core is at most $2k - \varepsilon$, then w.h.p. the edges of $G_{n,p}$ can be oriented so that no vertex has indegree more than k. It is clear from (4) that the edge density of the $(k+1)$-core increases smoothly w.h.p. and so we can apply the result of [6] for some value of ε.

It then follows that the edges of $G_{n,p}$ can be partitioned into k sets $\Phi_1, \Phi_2, \ldots, \Phi_k$ where each subgraph $H_i = ([n], \Phi_i)$ can be oriented so that each vertex has indegree at most one. We call such a graph a *Partial Functional Digraph* or PFD. Each component of a PFD is either a tree or contains exactly one cycle. We obtain $F_1, F_2, \ldots, F_k$ by removing one edge from each such cycle. We must show that w.h.p. we remove $o(n)$ vertices in total. Observe that if Z denotes the number of edges of $G_{n,p}$ that are on cycles of length at most $\omega_0 = \frac{1}{3} \log_c n$ then

$$\mathbf{E}\left[Z\right] \leq \sum_{\ell=3}^{\omega_0} \ell! \binom{n}{\ell} \ell p^{\ell} \leq \omega_0 c^{\omega_0} \leq n^{1/2}.$$

The Markov inequality implies that $Z \leq n^{2/3}$ w.h.p. The number of edges removed from the larger cycles to create $F_1, F_2, \ldots, F_k$ can be bounded by $kn/\omega_0 = o(n)$ and this proves (11) and the lemma. $\square$

Lemma 4. *If $c > c_2'$, then w.h.p. the 3-core of $G_{n,c/n}$ contains two edge-disjoint forests of total size $2|V(C_3)| - o(n)$. In particular, $r_2(C_3(G_{n,c/n})) = 2|V(C_3)| - o(n)$.*

The proof of Lemma 4 is postponed to Section 6. We can now prove Theorem 1 (b).

5 Proof of Theorem 1 (b).

As noted in (8),

$$\mathrm{mst}_2(K_n) = \int_{p=0}^{1} (2n - 2 - \mathbf{E}\left[r_2(G_p)\right])dp. \tag{12}$$

After changing variables to $x = pn$,

$$\mathrm{mst}_2(K_n) = \int_{x=0}^{n} (2 - 2n^{-1} - n^{-1}\mathbf{E}\left[r_2(G_{x/n})\right])dx \tag{13}$$

By Lemmas 2 and 3, for $x < c_2'$ we have $\mathbf{E}\left[r_2(G_{x/n})\right] = |E(G_{x/n})| - o(n) = xn/2 - o(n)$. By Lemma 4, for $x > c_2'$ we have $\mathbf{E}\left[r_2(C_3(G_{x/n}))\right] = 2|V(C_3)| - o(n)$. So by Lemma 2 $r_2(G_{x/n}) = |E(G_{x/n})| - |E(C_3)| + 2|V(C_3)| - o(n)$, and

$$\mu_2 = \int_{x=0}^{c_2'} \left(2 - \frac{x}{2}\right) dx + \int_{x=c_2'}^{n} \left(2 - \frac{1}{n}\left(\frac{xn}{2} - |E(C_3(G_{x/n})) + 2|V(C_3(G_{x/n}))|\right)\right) dx + o(1) \tag{14}$$

We have from (4) that for $p = x/n$ we have

$$\begin{aligned} \frac{1}{n}|V(C_3)| &= \frac{f_3(\lambda)}{e^\lambda} + o(1) \\ \frac{1}{n}|E(C_3)| &= \frac{\lambda f_2(\lambda)}{2e^\lambda} + o(1) \end{aligned}$$

where λ is the largest solution to $\lambda e^\lambda / f_2(\lambda) = x$. So

$$\mu_2 = \lim_{n\to\infty} \mathrm{mst}_2(K_n) = \int_{x=0}^{c_2'} \left(2 - \frac{x}{2}\right) dx + \int_{x=c_2'}^{\infty} \left(2 - \frac{x}{2} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right) dx \tag{15}$$

To calculate this, note that

$$\frac{dx}{d\lambda} = \frac{e^\lambda}{f_2(\lambda)} + \frac{\lambda e^\lambda}{f_2(\lambda)} - \frac{\lambda e^\lambda f_1(\lambda)}{f_2(\lambda)^2} \tag{16}$$

so

$$\begin{aligned} &\int_{x=c_2'}^{\infty} \left(2 - \frac{x}{2} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right) dx \\ &= \int_{\lambda_2'}^{\infty} \left(2 - \frac{\lambda e^\lambda}{2f_2(\lambda)} + \frac{\lambda f_2(\lambda)}{2e^\lambda} - 2\frac{f_3(\lambda)}{e^\lambda}\right)\left(\frac{e^\lambda}{f_2(\lambda)} + \frac{\lambda e^\lambda}{f_2(\lambda)} - \frac{\lambda e^\lambda f_1(\lambda)}{f_2(\lambda)^2}\right) d\lambda \end{aligned}$$

where, see (6),

$$\lambda_2' = g_0^{-1}(4) \approx 2.688 \tag{17}$$

is the unique solution to $\lambda f_2(\lambda)/f_3(\lambda) = 4$, see Appendix B. Attempts to transform this into an explicit integral with explicit bounds have been unsuccesful. Numerical calculations give

$$\mu_2 \approx 4.1704288\ldots \tag{18}$$

The Inverse Symbolic Calculator (https://isc.carma.newcastle.edu.au/) has yielded no symbolic representation of this number. An apparent connection to the ζ function lies in its representation as

$$\zeta(x) = \frac{1}{\Gamma(x)} \int_{\lambda=0}^{\infty} \frac{\lambda^{x-1}}{e^{\lambda}-1} d\lambda \tag{19}$$

which is somewhat similar to terms of the form

$$\int_{\lambda=\lambda_2'}^{\infty} \frac{\text{poly}(\lambda)}{e^{\lambda}-1-\lambda} d\lambda \tag{20}$$

appearing in μ_2, but no real connection has been found.

6 Proof of Lemma 4.

6.1 More on the 3-core.

Suppose now that $c > c_3'$ and that the 3-core C_3 of $G_{n,p}$ has $N = \Omega(n)$ vertices and M edges. It will be distributed as a random graph uniformly chosen from the set of graphs with vertex set $[N]$ and M edges and minimum degree at least three. This is an easy well known observation and follows from the fact that each such graph H can be extended in the same number of ways to a graph G with vertex set $[n]$ and m edges and such that H is the 3-core of G. We will for convenience now assume that $V(C_3) = [N]$.

The degree sequence $d(v), v \in [N]$ can be generated as follows: We independently choose for each $v \in V(C_3)$ a truncated Poisson random variable with parameter λ satisfying $g_0(\lambda) = 2M/N$, conditioned on $d(v) \geq 3$. So for $v \in [N]$,

$$\Pr\{d(v) = k\} = \frac{\lambda^k}{k! f_3(\lambda)}, \quad k = 3, 4, 5, \ldots, \quad \lambda = g_0^{-1}\left(\frac{2M}{N}\right) \tag{21}$$

Properties of the functions f_i, g_i are derived in Appendix B. In particular, the g_i are strictly increasing by Lemma 7, so g_0^{-1} is well defined.

These independent variables are further conditioned so that the event

$$\mathcal{D} = \left\{ \sum_{v \in [N]} d(v) = 2M \right\} \tag{22}$$

occurs. Now λ has been chosen so that $\mathbf{E}[d(v)] = 2M/N$ and then the local central limit theorem implies that $\Pr\{\mathcal{D}\} = \Omega(1/N^{1/2})$, see for example Durrett [8]. It follows that

$$\Pr\{\mathcal{E} \mid \mathcal{D}\} \leq O(n^{1/2}) \Pr\{\mathcal{E}\}, \tag{23}$$

for any event $\mathcal{E}$ that depends on the degree sequence of C_3.

In what follows we use the configuration model of Bollobás [5] to analyse C_3 after we have fixed its degree sequence. Thus, for each vertex v we define a set W_v of *points* such that $|W_v| = d(v)$, and write $W = \bigcup_v W_v$. A random configuration F is generated by selecting a random partition of W into M pairs. A pair $\{x, y\} \in F$ with $x \in W_u, y \in W_v$ yields an edge $\{u, v\}$ of the associated (multi-)graph Γ_F.

The key properties of F that we need are (i) conditional on F having no loops or multiple edges, it is equally likely to be any simple graph with the given degree sequence and (ii) for the degree sequences of interest, the probability that Γ_F is simple will be bounded away from zero. This is because the degree sequence in (23) has exponential tails. Thus we only need to show that Γ_F has certain properties w.h.p.

6.2 Setting up the main calculation.

Suppose now that $p = c/n$ where $c > c_2'$. We will show that w.h.p., for any fixed $\varepsilon > 0$,

$$i(S) = |\{e \in E(C_3) : e \cap S \neq \emptyset\}| \geq (2 - \varepsilon)|S| \text{ for all } S \subseteq [N]. \tag{24}$$

Proving this is the main computational task of the paper. In principle, it is just an application of the first moment method. We compute the expected number of S that violate (24) and show that tis expectation tends to zero. On the other hand, a moments glance at the expression $f(\mathbf{w})$ below shows that this is unlikely to be easy and it takes more than half of the paper to verify (24).

It follows from (24) that

$$E(C_3) \text{ can be oriented so that at least } (1 - \varepsilon)N \text{ vertices have indegree at least two.} \tag{25}$$

To see this consider the following network flow problem. We have a source s and a sink t plus a vertex for each $v \in [N]$ and a vertex for each edge $e \in E(C_3)$. The directed edges are (i) $(s, v), v \in [N]$ of capacity two; (ii) (u, e), where $u \in e$ of infinite capacity; (iii) $(e, t), e \in E(C_3)$ of capacity one. A $s - t$ flow decomposes into paths s, u, e, t corresponding to orienting the edge e into u. A flow thus corresponds to an orientation of $E(C_3)$. The condition (24) implies that the minimum cut in the network has capacity at least $(2 - \varepsilon)N$. This implies that there is a flow of value at least $(2 - \varepsilon)N$ and then the orientation claimed in (25) exists.

Thus w.h.p. C_3 contains two edge-disjoint PFD's, each containing $(1 - \varepsilon)N$ edges. Arguing as in the proof of Lemma 3, we see that we can w.h.p. remove $o(N)$ edges from the cycles of these PFD's and obtain forests. Thus w.h.p. C_3 contains two edge-disjoint forests of total size at least $2(1 - \varepsilon)N - o(N)$. This implies that $\mathbf{E}\left[r_2(C_3(G_{n,c/n}))\right] \geq 2(1 - \varepsilon)N - o(N)$ and since $N = \Omega(n)$, we can have $\mathbf{E}\left[r_2(C_3(G_{n,c/n}))\right] = 2(1 - \varepsilon)N - o(n)$. Because ε is arbitrary, this implies $r_2(C_3(G_{n,c/n})) = 2N - o(N)$ whenever $c > c_2'$.

6.3 Proof of (24): Small S.

It will be fairly easy to show that (25) holds w.h.p. for all $S| \leq s_0$ where

$$s_0 = \left(\frac{3(1 + \varepsilon)}{e^{2+\varepsilon}c}\right)^{1/\varepsilon} n.$$

We claim that w.h.p.

$$|S| \leq s_0 \text{ implies } e(S) < (1+\varepsilon)|S| \text{ in } G_{n,p}. \tag{26}$$

Here $e(S) = |\{e \in E(G_{n,p}) : e \subseteq S\}|$.

Indeed,

$$\Pr\{\exists S \text{ violating (26)}\} \leq \sum_{s=4}^{s_0} \binom{n}{s} \binom{\binom{s}{2}}{(1+\varepsilon)s} p^{(1+\varepsilon)s} \leq$$

$$\sum_{s=4}^{s_0} \left(\frac{ne}{s}\right)^s \left(\frac{sec}{2(1+\varepsilon)n}\right)^{(1+\varepsilon)s} = \sum_{s=4}^{s_0} \left(\left(\frac{s}{n}\right)^{\varepsilon} \frac{e^{2+\varepsilon}c}{2(1+\varepsilon)}\right)^s = o(1).$$

For sets A, B of vertices and $v \in A$ we will let $d_B(v)$ denote the number of neighbors of v in B. We then let $d_B(A) = \sum_{v \in A} d_B(v)$. We will drop the subscript B when $B = [N]$.

Suppose then that (26) holds and that $|S| \leq s_0$ and $i(S) \leq (2-\varepsilon)|S|$. Then if $\bar{S} = [N] \setminus S$, we have

$$e(S) + d_{\bar{S}}(S) \leq (2-\varepsilon)|S| \text{ and } d(S) = 2e(S) + d_{\bar{S}}(S) \geq 3|S|$$

which implies that $e(S) \geq (1+\varepsilon)|S|$, contradiction.

6.4 Proof of (24): Large S.

Suppose now that C_3 contains an S such that $i(S) < (2-\varepsilon)|S|$. Let such sets be *bad*. Let S be a minimal bad set, and write $T = [N] \setminus S$. For any $v \in S$, we have $i(S \setminus v) \geq (2-\varepsilon)|S \setminus v|$ while $i(S) < (2-\varepsilon)|S|$. This implies $d_T(v) = i(S) - i(S \setminus v) < 2$.

We will start with a minimal bad set and then carefully add more vertices. Consider a set S such that $i(S) < 2|S|$ and $d_T(v) \leq 2$ for all $v \in S$. If there is a $w \in T$ such that $d_T(w) \leq 2$, let $S' = S \cup \{w\}$. We have $i(S') \leq i(S) + 2 < 2|S'|$. This means we may add vertices to S in this fashion to aquire a partition $[N] = S \cup T$ where $d_T(v) \leq 2$ for all $v \in S$ and $d_T(v) \geq 3$ for all $v \in T$. We further partition $S = S_0 \cup S_1 \cup S_2$ so that $d_T(v) = i$ if and only if $v \in S_i$. Denote the size of any set by its lower case equivalent, e.g. $|S_0| = s_0$.

We now start to use the configuration model. Partition each point set into $W_v = W_v^S \cup W_v^T$, where a point is in W_v^S if and only if it is matched to a point in $\cup_{u \in S} W_u$. The sizes of W_v^S, W_v^T uniquely determine $\mathbf{w} = (s_0, s_1, s_2, D_0, D_1, D_2, D_3, t, M)$. Here $D_i = d_S(S_i), i = 0, 1, 2$ and $D_3 = d_T(T)$.

6.4.1 Estimating the probability of w.

We have $D_i \geq (3-i)s_i$ for $i = 0, 1, 2$ and $D_3 \geq 3t$. Define degree sequences $(d_i^1, \ldots, d_i^{s_i})$ for $S_i, i = 0, 1, 2$ and $(d_3^1, \ldots, d_3^t)$ for T. Furthermore, let $\widehat{d}_1^j = d_1^j - 1$, $\widehat{d}_2^j = d_2^j - 2$ and $\widehat{d}_3^j \geq 0$ be the S-degrees of vertices in S_1, S_2, T, respectively.

Dealing with S_0:
Ignoring for the moment, that we must condition on the event $\mathcal{D}$ (see (22)), the probability that S_0 has degree sequence $(d_0^1, \ldots, d_0^{s_0})$, $d_0^i \geq 3$ for all i, is given by

$$\prod_{i=1}^{s_0} \frac{\lambda^{d_0^i}}{d_0^i! f_3(\lambda)} \tag{27}$$

where λ is the solution to

$$g_0(\lambda) = \frac{2M}{N}.$$

Hence, letting $[x^D]f(x)$ denote the coefficient of x^D in the power series $f(x)$, the probability $\pi_0(S_0, D_0)$ that $d(S_0) = D_0$ is bounded by

$$\begin{aligned}
\pi_0(S_0, D_0) \le \sum_{\substack{d_0^1+\cdots+d_0^{s_0}=D_0 \\ d_0^i \ge 3}} \prod_{i=1}^{s_0} \frac{\lambda^{d_0^i}}{d_0^i! f_3(\lambda)} &= \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} \sum_{\substack{d_0^1+\cdots+d_0^{s_0}=D_0 \\ d_0^i \ge 3}} \prod_{i=1}^{s_0} \frac{1}{d_0^i!} \\
&= \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} [x^{D_0}] \left(\sum_{d_0 \ge 3} \frac{x^{d_0}}{d_0!} \right)^{s_0} \\
&= \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} [x^{D_0}] f_3(x)^{s_0} \\
&\le \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} \frac{f_3(\lambda_0)^{s_0}}{\lambda_0^{D_0}} \qquad (28)
\end{aligned}$$

for all λ_0. Here we use the fact that for any function f and any $y > 0$, $[x^{D_0}]f(x) \le f(y)/y^{D_0}$. To minimise (28) we choose λ_0 to be the unique solution to

$$g_0(\lambda_0) = \frac{D_0}{s_0}. \qquad (29)$$

If $D_0 = 3s_0$ then $\lambda_0 = 0$ by Lemma 6, Appendix B. In this case, since $f_3(\lambda_0) = \frac{\lambda_0^3(1+O(\lambda_0))}{6}$, we have

$$\pi_0(S_0, D_0) \le \left(\frac{\lambda^3}{6 f_3(\lambda)} \right)^{s_0}, \quad \text{when } D_0 = 3s_0. \qquad (30)$$

Dealing with S_1:
For each $v \in S_1$, we have $W_v = W_v^S \cup W_v^T$ where $|W_v^T| = 1$. Hence, the probability $\pi_1(S_1, D_1)$ that $d(S_1) = D_1 + s_1$ is bounded by

$$\begin{aligned}
\pi_1(S_1, D_1) \le \sum_{\substack{\widehat{d}_1^1+\cdots+\widehat{d}_1^{s_1}=D_1 \\ \widehat{d}_1^i \ge 2}} \prod_{i=1}^{s_1} \binom{\widehat{d}_1^i + 1}{1} \frac{\lambda^{\widehat{d}_1^i+1}}{(\widehat{d}_1^i + 1)! f_3(\lambda)} &= \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} \sum_{\substack{\widehat{d}_1^1+\cdots+\widehat{d}_1^{s_1}=D_1 \\ \widehat{d}_1^i \ge 2}} \prod_{i=1}^{s_1} \frac{1}{\widehat{d}_1^i!} \\
&= \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} [x^{D_1}] f_2(x)^{s_1} \\
&\le \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} \frac{f_2(\lambda_1)^{s_1}}{\lambda_1^{D_1}}. \qquad (31)
\end{aligned}$$

We choose λ_1 to satisfy the equation

$$g_1(\lambda_1) = \frac{D_1}{s_1}. \qquad (32)$$

Similarly to what happens in (30) we have $\lambda_1 = 0$ when $D_1 = 2s_1$ and we have $f_2(\lambda_1) = \frac{\lambda_1^2(1+O(\lambda_1)}{2}$ and then we have

$$\pi_1(S_1, D_1) \le \left(\frac{\lambda^3}{2 f_3(\lambda)} \right)^{s_1}, \quad \text{when } D_1 = 2s_1. \qquad (33)$$

Dealing with S_2:
For $v \in S_2$, we choose 2 points from W_v to be in W_v^T , so the probability $\pi_2(S_2, D_2)$ that $d(S_2) = D_2 + 2s_2$ is bounded by

$$\pi_2(S_2, D_2) \leq \sum_{\substack{\widehat{d}_2^1+\cdots+\widehat{d}_2^{s_2}=D_2 \\ \widehat{d}_2^i \geq 1}} \prod_{i=1}^{s_2} \binom{\widehat{d}_2^i + 2}{2} \frac{\lambda^{\widehat{d}_2^i+1}}{(\widehat{d}_2^i + 2)!f_3(\lambda)} \leq \frac{\lambda^{D_2+2s_2}}{f_3(\lambda)^{s_2}} \frac{f_1(\lambda_2)^{s_2}}{\lambda_2^{D_2}} 2^{-s_2} \tag{34}$$

where we choose λ_2 to satisfy the equation

$$g_2(\lambda_2) = \frac{D_2}{s_2}. \tag{35}$$

Similarly to what happens in (30) we have $\lambda_2 = 0$ when $D_2 = s_2$ and we have $f_1(\lambda_2) = \lambda_2(1+O(\lambda_2))$ and then we have

$$\pi_2(S_2, D_2) \leq \left(\frac{\lambda^3}{2f_3(\lambda)}\right)^{s_2}, \quad \text{when } D_2 = s_2. \tag{36}$$

Dealing with T:
Finally, the degree of vertex i in T can be written as $d_3^i = \widehat{d}_3^i + \overline{d}_3^i$ where $\widehat{d}_3^i \geq 0$ is the S-degree and $\overline{d}_3^i \geq 3$ is the T-degree. Here, with $t = |T|$, we have

$$\sum_{i=1}^{t} \widehat{d}_3^i = d_S(T) = s_1 + 2s_2$$

by the definition of S_0, S_1, S_2. So the probability $\pi_3(T, D_3)$ that $d_T(T) = D_3$, given s_1, s_2 can be bounded by

$$\begin{aligned}
\pi_3(T, D_3) &\leq \sum_{\substack{\widehat{d}_3^1+\cdots+\widehat{d}_3^t=s_1+2s_2 \\ \widehat{d}_3^i \geq 0}} \sum_{\substack{\overline{d}_3^1+\cdots+\overline{d}_3^t=D_3 \\ \overline{d}_3^i \geq 3}} \prod_{i=1}^{t} \binom{\widehat{d}_3^i + \overline{d}_3^i}{\widehat{d}_3^i} \frac{\lambda^{\widehat{d}_3^i+\overline{d}_3^i}}{(\widehat{d}_3^i + \overline{d}_3^i)!f_3(\lambda)} \\
&= \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \sum_{\substack{\widehat{d}_3^1+\cdots+\widehat{d}_3^t=s_1+2s_2 \\ \widehat{d}_3^i \geq 0}} \sum_{\substack{\overline{d}_3^1+\cdots+\overline{d}_3^t=D_3 \\ \overline{d}_3^i \geq 3}} \prod_{i=1}^{t} \frac{1}{\widehat{d}_3^i!\overline{d}_3^i!} \\
&= \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \left([x^{D_3}]f_3(x)^t\right) \left([x^{s_1+2s_2}]e^x\right) \\
&\leq \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \frac{f_3(\lambda_3)^t}{\lambda_3^{D_3}} \frac{t^{s_1+2s_2}}{(s_1 + 2s_2)!},
\end{aligned} \tag{37}$$

where we choose λ_3 to satisfy the equation

$$g_0(\lambda_3) = \frac{D_3}{t}. \tag{38}$$

Similarly to what happens in (30) we have $\lambda_3 = 0$ when $D_3 = 3t$ and we have $f_3(\lambda_3) = \frac{\lambda_3^3(1+O(\lambda_1))}{6}$ and then we have

$$\pi_3(T, D_3) \leq \frac{\lambda^{D_3+s_1+2s_2}}{(6f_3(\lambda))^t} \frac{t^{s_1+2s_2}}{(s_1 + 2s_2)!}, \quad \text{when } D_3 = 3t.$$

6.4.2 Putting the bounds together.

For a fixed $\mathbf{w} = (s_0, s_1, s_2, D_0, D_1, D_2, D_3, t, M)$, there are $\binom{t+s}{s_0,s_1,s_2,t}$ choices for S_0, S_1, S_2, T. Having chosen these sets we partition the $W_v, v \in S$ into $W_v^S \cup W_v^T$. Note that our expressions (28), (31), (34), (37) account for these choices. Given the partitions of the W_v's, there are $(D_0 + D_1 + D_2)!!D_3!!(s_1 + 2s_2)!$ configurations, where $(2s)!! = (2s-1) \times (2s-3) \times \cdots \times 3 \times 1$ is the number of ways of partitioning a set of size $2s$ into s pairs. Here $(D_0 + D_1 + D_2)!!$ is the number of ways of pairing up $\bigcup_{v\in S} W_v^S$, $D_3!!$ is the number of ways of pairing up $\bigcup_{v\in T} W_v^T$ and $(s_1 + 2s_2)!$ is the number of ways of pairing points associated with S to points associated with T. Each configuration has probability $1/(2M)!!$. So, the total probability of all configurations whose vertex partition and degrees are described by $\mathbf{w}$ can be bounded by

$$
\begin{aligned}
&\binom{t+s}{s_0, s_1, s_2, t} \frac{\lambda^{D_0}}{f_3(\lambda)^{s_0}} \frac{f_3(\lambda_0)^{s_0}}{\lambda_0^{D_0}} \frac{\lambda^{D_1+s_1}}{f_3(\lambda)^{s_1}} \frac{f_2(\lambda_1)^{s_1}}{\lambda_1^{D_1}} \frac{\lambda^{D_2+2s_2}}{f_3(\lambda)^{s_2}} \frac{f_1(\lambda_2)^{s_2}}{\lambda_2^{D_2}} 2^{-s_2}\\
&\times \frac{\lambda^{D_3+s_1+2s_2}}{f_3(\lambda)^t} \frac{f_3(\lambda_3)^t}{\lambda_3^{D_3}} \frac{t^{s_1+2s_2}}{(s_1+2s_2)!} \frac{(D_0 + D_1 + D_2)!!D_3!!(s_1 + 2s_2)!}{(2M)!!}\\
&= \binom{t+s}{s_0, s_1, s_2, t} \frac{\lambda^{2M}}{f_3(\lambda)^N} \frac{f_3(\lambda_0)^{s_0}}{\lambda_0^{D_0}} \frac{f_2(\lambda_1)^{s_1}}{\lambda_1^{D_1}} \frac{f_1(\lambda_2)^{s_2}}{\lambda_2^{D_2}} 2^{-s_2} \frac{f_3(\lambda_3)^t}{\lambda_3^{D_3}} \frac{t^{s_1+2s_2}}{(s_1+2s_2)!}\\
&\times \frac{(D_0 + D_1 + D_2)!!D_3!!(s_1 + 2s_2)!}{(2M)!!}
\end{aligned}
$$

Write $D_i = \Delta_i s$, $|S_i| = \sigma_i s$, $t = \tau s$, $M = \mu s$ and $N = \nu s$. We have $k!! \sim \sqrt{2}(k/e)^{k/2}$ as $k \to \infty$ by Stirling's formula, so the expression above, modulo an $e^{o(s)}$ factor, can be written as

$$
\begin{aligned}
f(\mathbf{w})^s = \Bigg(&\frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau} \frac{\lambda^{2\mu}}{f_3(\lambda)^\nu} \frac{f_3(\lambda_0)^{\sigma_0}}{\lambda_0^{\Delta_0}} \frac{f_2(\lambda_1)^{\sigma_1}}{\lambda_1^{\Delta_1}} \frac{f_1(\lambda_2)^{\sigma_2}}{\lambda_2^{\Delta_2}} \frac{f_3(\lambda_3)^\tau}{\lambda_3^{\Delta_3}} \frac{(\tau e)^{\sigma_1+2\sigma_2}}{2^{\sigma_2}}\\
&\frac{(\Delta_0+\Delta_1+\Delta_2)^{(\Delta_0+\Delta_1+\Delta_2)/2}\Delta_3^{\Delta_3/2}}{(2\mu)^\mu}\Bigg)^s
\end{aligned}
\tag{39}
$$

We note that

$$
\sigma_2 = 1 - \sigma_0 - \sigma_1, \tag{40}
$$

$$
\begin{aligned}
\Delta_3 &= 2\mu - \Delta_0 - \Delta_1 - \Delta_2 - 2\sigma_1 - 4\sigma_2\\
&= 2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1
\end{aligned}
\tag{41}
$$

$$
\nu = 1 + \tau. \tag{42}
$$

Hence σ_2, Δ_3, ν may be eliminated, and we can consider $\mathbf{w}$ to be $(\sigma_0, \sigma_1, \Delta_0, \Delta_1, \Delta_2, \tau, \mu)$. When convenient, Δ_3 may be used to denote $2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1$. Define the constraint set F to be all $\mathbf{w}$ satisfying

$$\Delta_0 \geq 3\sigma_0, \Delta_1 \geq 2\sigma_1, \Delta_2 \geq 1-\sigma_0-\sigma_1, \Delta_3 \geq 3\tau. \tag{43a}$$

$$\frac{\Delta_0+\Delta_1+\Delta_2}{2}+\sigma_1+2(1-\sigma_0-\sigma_1) < 2-\varepsilon \qquad \text{since } i(S) < (2-\varepsilon)|S|, \qquad \text{see (24).} \tag{43b}$$

$$\sigma_0, \sigma_1 \geq 0, \sigma_0+\sigma_1 \leq 1. \tag{43c}$$

$$0 \leq \tau \leq (1-\varepsilon)/\varepsilon \text{ since } |S| \geq \varepsilon N. \tag{43d}$$

$$\mu \geq (2+\varepsilon)(1+\tau) \text{ since } M \geq (2+\varepsilon)N. \tag{43e}$$

$$\sigma_0 < 1, \qquad \text{otherwise } C_3 \text{ is not connected.} \tag{43f}$$

Here ε is a sufficiently small positive constant such that (i) we can exclude the case of small S, (ii) satisfy condition (24) and (iii) have $M \geq (2+\varepsilon)N$ since $c > c_2'$.

For a given s, there are $O(\text{poly}(s))$ choices of $\mathbf{w} \in F$, and the probability that the randomly chosen configuration corresponds to a $\mathbf{w} \in F$ can be bounded by

$$\sum_{s \geq \varepsilon N} \sum_{\mathbf{w}} O(\text{poly}(s)) f(\mathbf{w})^s \leq \sum_s (e^{o(1)} \max_F f(\mathbf{w}))^s \leq N(e^{o(1)} \max_F f(\mathbf{w}))^{\varepsilon N}. \tag{44}$$

As $N \to \infty$, it remains to show that $f(\mathbf{w}) \leq 1-\delta$ for all $\mathbf{w} \in F$, for some $\delta = \delta(\varepsilon) > 0$. At this point we remind the reader that we have so far ignored conditioning on the event $\mathcal{D}$ defined in (22). Inequality (23) implies that it is sufficient to inflate the RHS of (44) by $O(n^{1/2})$ to obtain our result.

So, let

$$\begin{aligned} & f(\Delta_0, \Delta_1, \Delta_2, \sigma_0, \sigma_1, \tau, \mu) = \\ & \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau} \frac{\lambda^{2\mu}}{f_3(\lambda)^{\tau+1}} \frac{f_3(\lambda_0)^{\sigma_0}}{\lambda_0^{\Delta_0}} \frac{f_2(\lambda_1)^{\sigma_1}}{\lambda_1^{\Delta_1}} \frac{f_1(\lambda_2)^{1-\sigma_0-\sigma_1}}{\lambda_2^{\Delta_2}} \frac{f_3(\lambda_3)^\tau}{\lambda_3^{\Delta_3}} \\ \times \quad & \frac{(e\tau)^{2-2\sigma_0-\sigma_1}}{2^{1-\sigma_0-\sigma_1}} \frac{(\Delta_0+\Delta_1+\Delta_2)^{(\Delta_0+\Delta_1+\Delta_2)/2}\Delta_3^{\Delta_3/2}}{(2\mu)^\mu} \end{aligned}$$

We complete the proof of Theorem 1(b) by showing that

$$f(\mathbf{w}) \leq \exp\left\{-\frac{\varepsilon^2}{3}\right\} \text{ for all } \mathbf{w} \in F. \tag{45}$$

6.4.3 Eliminating μ

We begin by showing that it is enough to consider $\mu = (2+\varepsilon)(1+\tau)$. We collect all terms involving μ, including Δ_3, λ and λ_3 whose values are determined in part by μ. It is enough to consider the logarithm of f. We have

$$\begin{aligned} \frac{\partial \log f}{\partial \mu} = 2\log\lambda + \frac{\partial \lambda}{\partial \mu}\left(\frac{2\mu}{\lambda} - \nu\frac{f_2(\lambda)}{f_3(\lambda)}\right) + \frac{\partial \lambda_3}{\partial \mu}\left(\tau\frac{f_2(\lambda_3)}{f_3(\lambda_3)} - \frac{\Delta_3}{\lambda_3}\right) \\ -2\log\lambda_3 + \log\Delta_3 + 1 - \log 2\mu - 1 \end{aligned}$$

by definition of λ, λ_3, we have

$$\frac{2\mu}{\lambda} - \nu\frac{f_2(\lambda)}{f_3(\lambda)} = 0 \text{ and } \frac{\Delta_3}{\lambda_3} - \tau\frac{f_2(\lambda_3)}{f_3(\lambda_3)} = 0,$$

and so

$$\frac{\partial \log f}{\partial \mu} = 2\log\left(\frac{\lambda}{\lambda_3}\right) + \log\left(\frac{\Delta_3}{2\mu}\right) \tag{46}$$

We have $\Delta_3 \leq 2\mu$ and furthermore, $\lambda \leq \lambda_3$ since g_0 is an increasing function. Indeed, writing $\iota = i(S)/s \leq 2$, we have $\Delta_3 + 2\iota = 2\mu \geq 4(\tau+1)$, so

$$g_0(\lambda_3) - g_0(\lambda) = \frac{\Delta_3}{\tau} - \frac{2\mu}{\nu} = \frac{2\mu - 2\iota}{\tau} - \frac{2\mu}{\tau+1} = \frac{2\mu - 2\iota(\tau+1)}{\tau(\tau+1)} \geq \frac{4-2\iota}{\tau} \geq 0. \tag{47}$$

This shows that $\log f$ is decreasing with respect to μ, and in discussing the maximum value of f for $\mu \geq (2+\varepsilon)(1+\tau)$ we may assume that $\mu = (2+\varepsilon)(1+\tau)$.

We now argue that to show that $f \leq \exp\{-\varepsilon^2/3\}$ when $\mu = (2+\varepsilon)(1+\tau)$, it is enough to show that $f \leq 1$ when $\mu = 2(1+\tau)$. Let $2(1+\tau) < \mu < (2+\varepsilon)(1+\tau)$. Then by (41) and (43a)

$$\begin{aligned}
\Delta_3 &= 2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1 \\
&\leq 2\mu - 4 - 3\sigma_0 - 2\sigma_1 - (1 - \sigma_0 - \sigma_1) + 4\sigma_0 + 2\sigma_1 \\
&= 2\mu - 5 + 2\sigma_0 + \sigma_1 \\
&\leq 2\mu - 2
\end{aligned}$$

and since $\tau \leq 1/\varepsilon - 1$, $\mu \leq (2+\varepsilon)(1+\tau)$ implies $\mu \leq 2/\varepsilon + 1 < 3/\varepsilon$. So,

$$\frac{\partial \log f}{\partial \mu} \leq 2\log\left(\frac{\lambda}{\lambda_3}\right) + \log\left(\frac{2\mu-2}{2\mu}\right) \leq \log\left(1 - \frac{\varepsilon}{3}\right) \tag{48}$$

So, fixing $\mathbf{w}' = (\sigma_0, \sigma_1, \Delta_0, \Delta_1, \Delta_2, \tau)$, let $\mu = 2(1+\tau)$ and $\mu' = (2+\varepsilon)(1+\tau)$. If $f(\mathbf{w}', \mu) \leq 1$, then

$$\log f(\mathbf{w}', \mu') \leq \log f(\mathbf{w}', \mu) + \varepsilon(1+\tau)\log\left(1 - \frac{\varepsilon}{3}\right) \leq -\frac{\varepsilon^2}{3}. \tag{49}$$

This shows that it is enough to prove that $f(\mathbf{w}) \leq 1$ for $\mathbf{w} \in F'$, defined by

$$\Delta_0 \geq 3\sigma_0, \Delta_1 \geq 2\sigma_1, \Delta_2 \geq 1 - \sigma_0 - \sigma_1, \Delta_3 \geq 3\tau \tag{50a}$$

$$\Delta_0 + \Delta_1 + \Delta_2 \leq 4\sigma_0 + 2\sigma_1 \tag{50b}$$

$$\sigma_0, \sigma_1 \geq 0, \sigma_0 + \sigma_1 \leq 1 \tag{50c}$$

$$0 \leq \tau < \infty \tag{50d}$$

$$\mu = 2(1+\tau). \tag{50e}$$

We have relaxed equation (43b) to give (50b) in order to simplify later calculations. In F', λ is defined by

$$g_0(\lambda) = \frac{2\mu}{\nu} = \frac{4(1+\tau)}{1+\tau} = 4,$$

so in the remainder of the proof

$$\lambda = g_0^{-1}(4) \approx 2.688 \text{ is fixed.}$$

It will be convenient at times to write $\Delta = \Delta_0 + \Delta_1 + \Delta_2$. We observe that $3\sigma_0 + 2\sigma_1 + (1 - \sigma_0 - \sigma_1) = 2\sigma_0 + \sigma_1 + 1$, so by (50a), (50b),

$$2\sigma_0 + \sigma_1 + 1 \leq \Delta \leq 4\sigma_0 + 2\sigma_1. \tag{51}$$

Note also that $\mu = 2(1+\tau)$ implies

$$\Delta_3 = 2\mu - 4 - \Delta_0 - \Delta_1 - \Delta_2 + 4\sigma_0 + 2\sigma_1 = 4\tau + 4\sigma_0 + 2\sigma_1 - \Delta. \tag{52}$$

The quantity $2\sigma_0 + \sigma_1$ will appear frequently. We note that (51) and $\sigma_0 + \sigma_1 \le 1$ imply

$$1 \le 2\sigma_0 + \sigma_1 \le 2. \tag{53}$$

6.4.4 Eliminating τ

We now turn to choosing the optimal τ. With $\mu = 2(1+\tau)$,

$$\begin{aligned} f(\sigma_0,\sigma_1,\Delta_0,\Delta_1,\Delta_2,\tau) \quad = \quad & \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau}\left(\frac{\lambda^4}{f_3(\lambda)}\right)^{\tau+1}\frac{f_3(\lambda_0)^{\sigma_0}}{\lambda_0^{\Delta_0}}\frac{f_2(\lambda_1)^{\sigma_1}}{\lambda_1^{\Delta_1}} \\ & \times\frac{f_1(\lambda_2)^{1-\sigma_0-\sigma_1}}{\lambda_2^{\Delta_2}}\frac{f_3(\lambda_3)^\tau}{\lambda_3^{\Delta_3}}\frac{(e\tau)^{2-2\sigma_0-\sigma_1}}{2^{1-\sigma_0-\sigma_1}}\times\frac{\Delta^{\Delta/2}\Delta_3^{\Delta_3/2}}{(4+4\tau)^{2+2\tau}}. \end{aligned} \tag{54}$$

Here $\lambda_0 = \lambda_0(\Delta_0,\sigma_0)$, $\lambda_1 = \lambda_1(\Delta_1,\sigma_1)$, $\lambda_2 = \lambda_2(\Delta_2,\sigma_0,\sigma_1)$, $\lambda_3 = \lambda_3(\sigma_0,\sigma_1,\Delta_0,\Delta_1,\Delta_2,\tau)$ as defined in (29), (32), (35), (38). Since $\tau f_2(\lambda_3)/f_3(\lambda_3) - \Delta_3/\lambda_3 = 0$ by the definition of λ_3, the partial derivative of $\log f$ with respect to τ is given by

$$\begin{aligned} \frac{\partial}{\partial\tau}\log f(\sigma_0,\sigma_1,\Delta_0,\Delta_1,\Delta_2,\tau) \quad = \quad & \log(\tau+1) + 1 - \log\tau - 1 + \log\left(\frac{\lambda^4}{f_3(\lambda)}\right) \\ & + \frac{\partial\lambda_3}{\partial\tau}\left(\tau\frac{f_2(\lambda_3)}{f_3(\lambda_3)} - \frac{\Delta_3}{\lambda_3}\right) + \log(f_3(\lambda_3)) - 4\log\lambda_3 \\ & + \frac{2-2\sigma_0-\sigma_1}{\tau} + 2\left(1+\log\Delta_3\right) - 2\log(4+4\tau) - 2 \\ = \quad & \log(\tau+1) - \log\tau + \log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right) + \frac{2-2\sigma_0-\sigma_1}{\tau} \\ & + 2\log\Delta_3 - 2\log(4+4\tau) \end{aligned}$$

This is positive for τ close to zero. This is clear as long as $2\sigma_0 + \sigma_1 < 2$. But if $2\sigma_0 + \sigma_1 = 2$ then $\sigma_0 + \sigma_1 \le 1$ implies that $\sigma_0 = 1, \sigma_1 = 0$. But then if $\tau > 0$ we have that C_3 is not connected and that if $\tau = 0$, $S = [N]$ which violates (43f). On the other hand, $\frac{\partial}{\partial\tau}\log f$ vanishes if

$$2 - 2\sigma_0 - \sigma_1 - \tau\left[\log\left(1+\frac{1}{\tau}\right) - 2\log\left(\frac{\Delta_3}{4\tau}\right) - \log\left(\frac{\lambda^4 f_3(\lambda_3)}{\lambda_3^4 f_3(\lambda)}\right)\right] = 0. \tag{55}$$

So any local maximum of f must satisfy this equation. If no solution exists, then it is optimal to let $\tau \to \infty$. We will see below how to choose τ to guarantee maximality. For now, we only assume τ satisfies (55).

6.4.5 Eliminating $\Delta_0, \Delta_1, \Delta_2$.

We now eliminate $\Delta_0, \Delta_1, \Delta_2$. Fix σ_0, σ_1. For $\Delta_i > (3-i)\sigma_i$ such that $\Delta_0 + \Delta_1 + \Delta_2 < 4\sigma_0 + 2\sigma_1$,

$$\begin{aligned} \frac{\partial}{\partial \Delta_i} \log f &= \frac{\partial \lambda_i}{\partial \Delta_i} \left(\sigma_i \frac{f_{2-i}(\lambda_i)}{f_{3-i}(\lambda_i)} - \frac{\Delta_i}{\lambda_i} \right) - \log \lambda_i + \log \lambda_3 \\ &\quad + \frac{\partial}{\partial \tau} \log f \;\; \frac{\partial \tau}{\partial \Delta_i} + \frac{1}{2} \log \Delta + \frac{1}{2} - \frac{1}{2} \log \Delta_3 - \frac{1}{2} \\ &= -\log \lambda_i + \log \left(\lambda_3 \sqrt{\frac{\Delta}{\Delta_3}} \right), \end{aligned} \tag{56}$$

since $g_i(\lambda_i) = \Delta_i/\sigma_i$ by definition of λ_i, and the term $\frac{\partial}{\partial \tau} \log f \;\; \partial\tau/\partial\Delta_i$ vanishes because (55) is assumed to hold. We note that $\lambda_i > 0$ when $\Delta_i > (3-i)\sigma_i$ (Appendix B), allowing division by λ_i.

As Δ_i tends to its lower bound $(3-i)\sigma_i$, we have $\log \lambda_i \to -\infty$ while the other terms remain bounded, so the derivative is positive at the lower bound of Δ_i. Any stationary point must satisfy $\lambda_0 = \lambda_1 = \lambda_2 = \lambda_3 \sqrt{\Delta/\Delta_3} =: \widehat{\lambda}$. This can only happen if

$$\sigma_0 g_0(\widehat{\lambda}) + \sigma_1 g_1(\widehat{\lambda}) + (1-\sigma_0-\sigma_1) g_2(\widehat{\lambda}) = \sigma_0 \frac{\Delta_0}{\sigma_0} + \sigma_1 \frac{\Delta_1}{\sigma_1} + (1-\sigma_0-\sigma_1) \frac{\Delta_2}{1-\sigma_0-\sigma_1} = \Delta. \tag{57}$$

So we choose $\widehat{\lambda}$, Δ, τ to solve the system of equations

$$\begin{aligned} \widehat{\lambda} &= \lambda_3 \sqrt{\frac{\Delta}{\Delta_3}} \\ \Delta &= \sigma_0 g_0(\widehat{\lambda}) + \sigma_1 g_1(\widehat{\lambda}) + (1-\sigma_0-\sigma_1) g_2(\widehat{\lambda}) \\ 2 - 2\sigma_0 - \sigma_1 &= \tau \left[\log\left(1 + \frac{1}{\tau}\right) - 2 \log\left(\frac{\Delta_3}{4\tau}\right) - \log\left(\frac{\lambda^4 f_3(\lambda_3)}{\lambda_3^4 f_3(\lambda)}\right) \right] \end{aligned} \tag{58}$$

In Appendix A we show that this system has no solution such that $2\sigma_0 + \sigma_1 + 1 \le \Delta \le 4\sigma_0 + 2\sigma_1$ (see (51)). This means that no stationary point exists, and $\log f$ is increasing in each of $\Delta_0, \Delta_1, \Delta_2$. In particular, it is optimal to set

$$\Delta_0 + \Delta_1 + \Delta_2 = 4\sigma_0 + 2\sigma_1 \text{ which implies that } \Delta_3 = 4\tau, \text{ see (52).} \tag{59}$$

This eliminates one degree of freedom. We now set

$$\Delta_2 = 4\sigma_0 + 2\sigma_1 - \Delta_0 - \Delta_1.$$

Then for Δ_0, Δ_1, we have

$$\frac{\partial}{\partial \Delta_i} \log f = -\log \lambda_i + \log \lambda_2, \quad i = 0, 1. \tag{60}$$

To see this note that (56) has to be modified via the addition of $\frac{\partial}{\partial \Delta_2} \log f \times \frac{\partial \Delta_2}{\partial \Delta_i}$, for $i = 0, 1$.

So it is optimal to let $\lambda_0 = \lambda_1 = \lambda_2 = \overline{\lambda}$, defined by

$$\sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1-\sigma_0-\sigma_1) g_2(\overline{\lambda}) = 4\sigma_0 + 2\sigma_1 \tag{61}$$

This has a unique solution $\overline{\lambda} \geq 0$ whenever $2\sigma_0 + \sigma_1 \geq 1$, since for fixed σ_0, σ_1, the left-hand side is a convex combination of increasing functions, by Lemma 7, Appendix B. This defines $\Delta_i = \Delta_i(\sigma_0, \sigma_1)$ by

$$\Delta_0 = g_0(\overline{\lambda})\sigma_0, \quad \Delta_1 = g_1(\overline{\lambda})\sigma_1, \quad \Delta_2 = g_2(\overline{\lambda})(1 - \sigma_0 - \sigma_1) \tag{62}$$

We note at this point that $\overline{\lambda} \leq \lambda$. Indeed, by (59) and (43a),

$$\Delta_0 = 4\sigma_0 + 2\sigma_1 - \Delta_1 - \Delta_2 \leq 4\sigma_0 + 2\sigma_1 - 2\sigma_1 - (1 - \sigma_0 - \sigma_1) \leq 4\sigma_0,$$

so

$$g_0(\overline{\lambda}) = \frac{\Delta_0}{\sigma_0} \leq 4 = g_0(\lambda) \tag{63}$$

implying that $\overline{\lambda} \leq \lambda$, since g_0 is increasing.

This choice (62) of $\Delta_0, \Delta_1, \Delta_2$ simplifies f significantly. With $\Delta = 4\sigma_0 + 2\sigma_1$ we have $\Delta_3 = 4\tau$, see (59), and so

$$\lambda_3 = g_0^{-1}\left(\frac{4\tau}{\tau}\right) = \lambda \tag{64}$$

is fixed. In particular, the relation (55) for τ simplifies to

$$2 - 2\sigma_0 - \sigma_1 = \tau \log\left(1 + \frac{1}{\tau}\right) \tag{65}$$

Let $\phi(\tau) = \tau \log(1 + 1/\tau)$. Then $\phi''(\tau) = -\tau^{-1}(\tau+1)^{-2}$, so ϕ is concave and then $\lim_{\tau \to 0} \phi(\tau) = 0$, $\lim_{\tau \to \infty} \phi(\tau) = 1$ implies that ϕ is strictly increasing and takes values in $[0, 1)$ for $\tau \geq 0$. This means that (65) has a unique solution if and only if $2\sigma_0 + \sigma_1 > 1$. When $2\sigma_0 + \sigma_1 = 1$, f is increasing with respect to τ, and we treat this case now.

If $2\sigma_0 + \sigma_1 = 1$, then (51) implies that $\Delta = 2$. Furthermore, $\Delta_3 = 4\tau$ (see (52)) and $\lambda_3 = \lambda$ (see (64)) and $g_i(0) = 3 - i$ implies that

$$\sigma_0 g_0(0) + \sigma_1 g_1(0) + (1 - \sigma_0 - \sigma_1) g_2(0) = 2\sigma_0 + \sigma_1 + 1 = 4\sigma_0 + 2\sigma_1,$$

so $\overline{\lambda} = 0$ is the unique solution to (61). Then since $\Delta_i/\sigma_i = g_i(0) = 3 - i$ (Lemma 6, Appendix B), we have $\Delta_i = (3 - i)\sigma_i$, $i = 0, 1, 2$, and as in (30), (33), (36),

$$\frac{f_3(\overline{\lambda})^{\sigma_0} f_2(\overline{\lambda})^{\sigma_1} f_1(\overline{\lambda})^{1-\sigma_0-\sigma_1}}{\overline{\lambda}^{\Delta}} = \left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^3}\right)^{\sigma_0} \left(\frac{f_2(\overline{\lambda})}{\overline{\lambda}^2}\right)^{\sigma_1} \left(\frac{f_1(\overline{\lambda})}{\overline{\lambda}}\right)^{1-\sigma_0-\sigma_1} = \frac{1}{6^{\sigma_0}} \frac{1}{2^{\sigma_1}} \tag{66}$$

so when $2\sigma_0 + \sigma_1 = 1$, (54) becomes

$$f(\sigma_0, \sigma_1, \tau) = \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0} \sigma_1^{\sigma_1} (1 - \sigma_0 - \sigma_1)^{1-\sigma_0-\sigma_1} \tau^{\tau}} \frac{\lambda^4}{f_3(\lambda)} \frac{1}{6^{\sigma_0}} \frac{1}{2^{\sigma_1}} \frac{e\tau}{2^{1-\sigma_0-\sigma_1}} \frac{2^{2/2}(4\tau)^{2\tau}}{(4+4\tau)^{2+2\tau}}. \tag{67}$$

In this computation we also used the fact that $\lambda = \lambda_3$ (see (64)) and $\Delta_3 = 4\tau$ (see (52)) to find that

$$\left(\frac{\lambda^4}{f_3(\lambda)}\right)^{\tau+1} \frac{f_3(\lambda_3)^{\tau}}{\lambda_3^{\Delta_3}} = \frac{\lambda^4}{f_3(\lambda)}.$$

Here $\lambda^4/f_3(\lambda) \approx 7.05$ is fixed. We show in Appendix A that in this case, the partial derivative in τ is positive for all τ, so we let $\tau \to \infty$. Substituting $\sigma_1 = 1 - 2\sigma_0$ we are reduced to

$$\begin{aligned} f(\sigma_0) &= \lim_{\tau\to\infty} \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}(1-2\sigma_0)^{(1-2\sigma_0)}\sigma_0^{\sigma_0}\tau^\tau} \frac{\lambda^4}{f_3(\lambda)} \frac{1}{6^{\sigma_0}} \frac{1}{2^{1-2\sigma_0}} \frac{e\tau}{2^{\sigma_0}} \frac{2(4\tau)^{2\tau}}{(4+4\tau)^{2+2\tau}} \\ &= \frac{\lambda^4}{16 f_3(\lambda)} \frac{1}{\sigma_0^{2\sigma_0}(1-2\sigma_0)^{1-2\sigma_0}3^{\sigma_0}} \end{aligned}$$

This has the stationary point $\sigma_0 = 2 - \sqrt{3}$, and $f(2-\sqrt{3}) \approx 0.95$. We also have $f(0) \approx 0.44$ and $f(1/2) \approx 0.51$ at the lower and upper bounds for σ_0.

6.4.6 Dealing with σ_0, σ_1

With this, we have reduced our analysis to the variables σ_0, σ_1 in the domain

$$E = \{(\sigma_0, \sigma_1) : \sigma_0, \sigma_1 \ge 0, \sigma_0 + \sigma_1 \le 1, 2\sigma_0 + \sigma_1 \ge 1\}.$$

We just showed that $f \le 1$ in

$$E_0 = \{(\sigma_0, \sigma_1) \in E : 2\sigma_0 + \sigma_1 = 1\}.$$

Further define

$$\begin{aligned} E_1 &= \{(\sigma_0, \sigma_1) \in E : 0.01 \le \sigma_1 \le 0.99\}, \\ E_2 &= \{(\sigma_0, \sigma_1) \in E : 0 \le \sigma_1 < 0.01\}, \\ E_3 &= \{(\sigma_0, \sigma_1) \in E : 0.99 < \sigma_1 \le 1\}. \end{aligned}$$

We will show that $f \le 1$ in each of these sets, whose union covers E.

From this point on, let $\partial_i = \frac{\partial}{\partial\sigma_i}, i = 0, 1$. As noted above, $\Delta = 4\sigma_0 + 2\sigma_1$ simplifies f. Specifically, if $2\sigma_0 + \sigma_1 > 1$ then (54) becomes, after using (59) and (64),

$$\begin{aligned} f(\sigma_0, \sigma_1) &= \frac{(\tau+1)^{\tau+1}}{\sigma_0^{\sigma_0}\sigma_1^{\sigma_1}(1-\sigma_0-\sigma_1)^{1-\sigma_0-\sigma_1}\tau^\tau} \frac{\lambda^4}{f_3(\lambda)} \frac{f_3(\overline{\lambda})^{\sigma_0} f_2(\overline{\lambda})^{\sigma_1} f_1(\overline{\lambda})^{1-\sigma_0-\sigma_1}}{\overline{\lambda}^{4\sigma_0+2\sigma_1}} \\ &\quad \times \frac{(e\tau)^{2-2\sigma_0-\sigma_1}}{2^{1-\sigma_0-\sigma_1}} \frac{(4\sigma_0+2\sigma_1)^{2\sigma_0+\sigma_1}(4\tau)^{2\tau}}{(4+4\tau)^{2+2\tau}} \end{aligned} \tag{68}$$

In (65), (61) respectively, τ and $\overline{\lambda}$ are given as functions of σ_0, σ_1. Recall that $\lambda = g_0^{-1}(4)$ is constant. So

$$\begin{aligned} &\partial_0 \log f(\sigma_0, \sigma_1) = \\ &- \log\sigma_0 - 1 + \log(1-\sigma_0-\sigma_1) + 1 + \log f_3(\overline{\lambda}) - \log f_1(\overline{\lambda}) \\ &- 4\log\overline{\lambda} - 2\log(e\tau) + \log 2 + 2\log(4\sigma_0+2\sigma_1) + 2 \\ &+ \frac{\partial\overline{\lambda}}{\partial\sigma_0}\left(\sigma_0\frac{f_2(\overline{\lambda})}{f_3(\overline{\lambda})} + \sigma_1\frac{f_1(\overline{\lambda})}{f_2(\overline{\lambda})} + (1-\sigma_0-\sigma_1)\frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{4\sigma_0+2\sigma_1}{\overline{\lambda}}\right) \\ &+ \frac{\partial\tau}{\partial\sigma_0}\left(\log(\tau+1) + 1 - \log\tau - 1 + \frac{2-2\sigma_0-\sigma_1}{\tau} + 2\log 4\tau + 2 - 2\log(4+4\tau) - 2\right) \\ &= \log\left(\frac{1-\sigma_0-\sigma_1}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^4 f_1(\overline{\lambda})}\right) - 2\log\tau + \log 2 + 2\log(4\sigma_0+2\sigma_1) \end{aligned} \tag{69}$$

where, as expected, the terms involving $\partial_0\tau$ and $\partial_0\overline{\lambda}$ vanish since $\tau, \overline{\lambda}$ were chosen to maximize $\log f$. (See (65) and (61) respectively).

Similarly,

$$
\begin{aligned}
&\partial_1 \log f(\sigma_0, \sigma_1) = \\
&- \log \sigma_1 - 1 + \log(1 - \sigma_0 - \sigma_1) + 1 + \log f_2(\overline{\lambda}) - \log f_1(\overline{\lambda}) \\
&- 2\log \overline{\lambda} - \log(e\tau) + \log 2 + \log(4\sigma_0 + 2\sigma_1) + 1 \\
&+ \frac{\partial\overline{\lambda}}{\partial\sigma_1}\left(\sigma_0 \frac{f_2(\overline{\lambda})}{f_3(\overline{\lambda})} + \sigma_1 \frac{f_1(\overline{\lambda})}{f_2(\overline{\lambda})} + (1 - \sigma_0 - \sigma_1)\frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{4\sigma_0 + 2\sigma_1}{\overline{\lambda}}\right) \\
&+ \frac{\partial\tau}{\partial\sigma_1}\left(\log(\tau + 1) + 1 - \log\tau - 1 + \frac{2 - 2\sigma_0 - \sigma_1}{\tau} + 2\log 4\tau + 2 - 2\log(4 + 4\tau) - 2\right) \\
&= \log\left(\frac{1 - \sigma_0 - \sigma_1}{\sigma_1}\right) + \log\left(\frac{f_2(\overline{\lambda})}{\overline{\lambda}^2 f_1(\overline{\lambda})}\right) - \log\tau + \log 2 + \log(4\sigma_0 + 2\sigma_1).
\end{aligned}
\tag{70}
$$

Any stationary point must satisfy

$$
(\partial_0 - 2\partial_1)\log f = \log\left(\frac{\sigma_1^2}{\sigma_0(1 - \sigma_0 - \sigma_1)}\right) + \log\left(\frac{f_1(\overline{\lambda})f_3(\overline{\lambda})}{f_2(\overline{\lambda})^2}\right) - \log 2 = 0. \tag{71}
$$

Now we show in Lemma 8, Appendix B that

$$
1 \le \frac{f_2(\overline{\lambda})^2}{f_1(\overline{\lambda})f_3(\overline{\lambda})} \le 2.
$$

This means from (71) that if $(\partial_0 - 2\partial_1)\log f = 0$ then

$$
2 \le \frac{\sigma_1^2}{\sigma_0(1 - \sigma_0 - \sigma_1)} \le 4.
$$

In particular, the lower bound implies $\sigma_0 \ge (1 - \sigma_1)/2 + \sqrt{1 - 2\sigma_1 - \sigma_1^2}/2$ and the upper bound implies $\sigma_1 \le -2\sigma_0 + \sqrt{4\sigma_0 - 4\sigma_0^2}$. The latter bound is used only to conclude that $\sigma_1 < 1/2$, by noting that $-2\sigma_0 + \sqrt{4\sigma_0 - 4\sigma_0^2} \le (5^{1/2} - 1)/3 < 1/2$ for $0 \le \sigma_0 \le 1$. In conclusion,

$$
(\partial_0 - 2\partial_1)\log f = 0 \Longrightarrow \begin{cases} \sigma_0 \ge (1 - \sigma_1)/2 + \sqrt{1 - 2\sigma_1 - \sigma_1^2}/2. \\ \sigma_1 < 1/2. \end{cases} \tag{72}
$$

Case One. $E_1 = \{(\sigma_0, \sigma_1) \in E : 0.01 \le \sigma_1 \le 0.99\}$

When $\sigma_0 < 0.99$, we need a lower bound for $\overline{\lambda}\tau$. We first note that $g_i(\overline{\lambda}) \le 3 - i + \overline{\lambda}$ (Lemma 6, Appendix B) implies

$$
4\sigma_0 + 2\sigma_1 = \sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1 - \sigma_0 - \sigma_1)g_2(\overline{\lambda}) \le 2\sigma_0 + \sigma_1 + 1 + \overline{\lambda} \tag{73}
$$

so

$$
\overline{\lambda} \ge 2\sigma_0 + \sigma_1 - 1 = 1 - \tau\log(1 + 1/\tau).
$$

Here we have used (65).

For τ, note that $\sigma_0 < 0.99$ and $\sigma_0 + \sigma_1 \leq 1$ implies $\tau \log(1 + 1/\tau) = 2 - 2\sigma_0 - \sigma_1 \geq 1 - \sigma_0 > 0.01$. The function $\tau \log(1 + 1/\tau)$ is increasing in τ by the discussion after (65). This implies

$$\tau > 10^{-3}, \tag{74}$$

since $0.001 \log(1001) < 0.01$.

If $\tau \leq 1.1$,

$$\overline{\lambda} \geq 1 - 1.1 \log 2 > 0.1.$$

So, if $\tau \leq 1.1$,

$$\overline{\lambda}\tau \geq 10^{-4}.$$

If $1.1 < \tau$ then we use $\log(1 + x) \leq x - x^2/2 + x^3/3$ for $|x| \leq 1$ to write

$$\overline{\lambda}\tau \geq \tau - \tau^2 \log(1 + 1/\tau) \geq \frac{1}{2} - \frac{1}{3\tau} \geq \frac{1}{6}.$$

So, in E_1, we have

$$\overline{\lambda}\tau \geq 10^{-4}. \tag{75}$$

By definition of E_1, $\sigma_0 \geq 0.01$ and $\sigma_1 \geq 0.01$. By (63), $0 \leq \overline{\lambda} \leq \lambda$. This implies $f_3(\overline{\lambda})/\overline{\lambda}^2 f_1(\overline{\lambda}) \leq 1/6$ and $f_2(\overline{\lambda})/\overline{\lambda} f_1(\overline{\lambda}) \leq 1/3$ (Lemma 8, Appendix B). So after rewriting (69) slightly,

$$\begin{aligned}
\partial_0 \log f(\sigma_0, \sigma_1) &= \log\left(\frac{1 - \sigma_0 - \sigma_1}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_1(\overline{\lambda})}\right) - 2\log\overline{\lambda}\tau + \log 2 + 2\log(4\sigma_0 + 2\sigma_1) \\
&\leq \log\frac{1}{0.01} + \log\frac{1}{6} - 2\log 10^{-4} + \log 2 + 2\log 4 \qquad (76) \\
&\leq 25.
\end{aligned}$$

Similarly, (70) is bounded by

$$\partial_1 \log f(\sigma_0, \sigma_1) \quad \leq \quad \log\frac{1}{0.01} + \log\frac{1}{3} - \log 10^{-4} + \log 2 + \log 4 \leq 15.$$

We now show numerically that $\log f \leq 0$ in E_1.

Numerics of Case One:
Since $\partial_i \log f$ is only bounded from above, $i = 0, 1$, this requires some care at the lower bounds of σ_0, σ_1, given by $\sigma_0 \geq (1 - \sigma_1)/2$ and $\sigma_1 \geq 0.01$. Note that if $\sigma_0 = (1 - \sigma_1)/2$, then $(\sigma_0, \sigma_1) \in E_0$ and it was shown above that $\log f(\sigma_0, \sigma_1) \leq \log 0.95 \leq -0.01$. Define a finite grid $P \subseteq E_1$ such that for any $(\sigma_0, \sigma_1) \in E_1$, there exists $(\overline{\sigma}_0, \overline{\sigma}_1) \in P \cup E_0$ where $0 \leq \sigma_0 - \overline{\sigma}_0 \leq \delta$ and $0 \leq \sigma_1 - \overline{\sigma}_1 \leq \delta$. Here $\delta = 1/4000$. Numerical calculations will show that $\log f(\overline{\sigma}_0, \overline{\sigma}_1) \leq -0.01$ for all $(\overline{\sigma}_0, \overline{\sigma}_1) \in P$. This implies that for all $\sigma_0, \sigma_1 \in E_1$,

$$\log f(\sigma_0, \sigma_1) \leq \max_{\overline{\sigma}_0, \overline{\sigma}_1 \in P \cup E_0} \log f(\overline{\sigma}_0, \overline{\sigma}_1) + 25\delta + 15\delta \leq -0.01 + 40\delta \leq 0.$$

When calculating $\log f(\overline{\sigma}_0, \overline{\sigma}_1)$, approximations $\overline{\lambda}_{num}, \tau_{num}$ of $\overline{\lambda}(\overline{\sigma}_0, \overline{\sigma}_1), \tau(\overline{\sigma}_0, \overline{\sigma}_1)$ must be calculated with sufficient precision. By definition of $\overline{\lambda}$, $\partial \log f/\partial\overline{\lambda} = 0$, while

$$\begin{aligned}
&\left|\frac{\partial^2 \log f}{\partial \overline{\lambda}^2}\right| \\
&= \left|\sigma_0\left(\frac{f_1(\overline{\lambda})}{f_3(\overline{\lambda})} - \frac{f_2(\overline{\lambda})^2}{f_3(\overline{\lambda})^2}\right) + \sigma_1\left(\frac{f_0(\overline{\lambda})}{f_2(\overline{\lambda})} - \frac{f_1(\overline{\lambda})^2}{f_2(\overline{\lambda})^2}\right)\right. \\
&\qquad\qquad\left. + (1-\sigma_0-\sigma_1)\left(\frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{f_0(\overline{\lambda})^2}{f_1(\overline{\lambda})^2}\right) + \frac{4\sigma_0+2\sigma_1}{\overline{\lambda}^2}\right| \\
&= \frac{1}{\overline{\lambda}^2}\left|\sigma_0\left(\overline{\lambda}^2\frac{f_1(\overline{\lambda})}{f_3(\overline{\lambda})} - \frac{\overline{\lambda}^2 f_2(\overline{\lambda})^2}{f_3(\overline{\lambda})^2}\right) + \sigma_1\left(\overline{\lambda}^2\frac{f_0(\overline{\lambda})}{f_2(\overline{\lambda})} - \frac{\overline{\lambda}^2 f_1(\overline{\lambda})^2}{f_2(\overline{\lambda})^2}\right)\right. \\
&\left. + (1-\sigma_0-\sigma_1)\left(\overline{\lambda}^2\frac{f_0(\overline{\lambda})}{f_1(\overline{\lambda})} - \frac{\overline{\lambda}^2 f_0(\overline{\lambda})^2}{f_1(\overline{\lambda})^2}\right) + \sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1-\sigma_0-\sigma_1)g_2(\overline{\lambda})\right| \\
&= \frac{1}{\overline{\lambda}^2}|\sigma_0 g_0(\overline{\lambda})(g_1(\overline{\lambda}) - g_0(\overline{\lambda}) + 1) + \sigma_1 g_1(\overline{\lambda})(g_2(\overline{\lambda}) - g_1(\overline{\lambda}) + 1) \\
&\qquad\qquad\qquad + (1-\sigma_0-\sigma_1)g_2(\overline{\lambda})(\overline{\lambda} - g_2(\overline{\lambda}) + 1)| \\
&\leq \frac{9}{\overline{\lambda}^2}|\sigma_0 g_0(\overline{\lambda}) + \sigma_1 g_1(\overline{\lambda}) + (1-\sigma_0-\sigma_1)g_2(\overline{\lambda})| \\
&= \frac{9}{\overline{\lambda}^2}|4\sigma_0 + 2\sigma - 1|, \qquad \text{by (61)} \\
&\leq \frac{36}{\overline{\lambda}^2}.
\end{aligned}$$

Here we use the fact that $g_i(\overline{\lambda}) \leq 4$ for $0 \leq \overline{\lambda} \leq \lambda$, $i = 0, 1, 2$ to conclude that $|g_1 - g_0 + 1|, |g_2 - g_1 + 1|, |\overline{\lambda} - g_2 + 1| \leq 9$, and the final step uses $4\sigma_0 + 2\sigma_1 \leq 4$. So the error contributed by $\overline{\lambda}_{num}$ is

$$|\log f(\overline{\sigma}_0, \overline{\sigma}_1; \overline{\lambda}_{num}) - \log f(\overline{\sigma}_0, \overline{\sigma}_1; \overline{\lambda})| \leq (\overline{\lambda}_{num} - \overline{\lambda})^2 \frac{36}{\overline{\lambda}^2} \tag{77}$$

and to achieve a numerical error of at most 10^{-4}, we require that $|\overline{\lambda}_{num}/\overline{\lambda} - 1| \leq 10^{-2}/6$.

Similarly by definition of τ, $\partial \log f/\partial\tau = 0$, while

$$\left|\frac{\partial^2 \log f}{\partial \tau^2}\right| = \left|\frac{1}{\tau(\tau+1)}\right| \leq 10^3, \text{ by (74).}$$

Thus to achieve a numerical error of at most 10^{-4}, it suffices to have $|\tau_{num}/\tau - 1| \leq 10^{-2}$.

With the above precision, it is found that over all $(\overline{\sigma}_0, \overline{\sigma}_1) \in P \cup E_0$, $\log f(\overline{\sigma}_0, \overline{\sigma}_1) \leq -0.0105$ numerically. With an error tolerance of 10^{-4}, this shows that $\log f(\overline{\sigma}_0, \overline{\sigma}_1) \leq -0.01$.

Case Two. $E_2 = \{(\sigma_0, \sigma_1) \in E : 0 \leq \sigma_1 < 0.01\}$

We divide E_2 into three subregions,

$$\begin{aligned}
E_{2,1} &= \{(\sigma_0, \sigma_1) \in E_2 : \sigma_1 = 0\}, \\
E_{2,2} &= \{(\sigma_0, \sigma_1) \in E_2 : \sigma_0 + \sigma_1 = 1\}, \\
E_{2,3} &= E_2 \setminus (E_{2,1} \cup E_{2,2}).
\end{aligned}$$

We begin by considering the point $(\sigma_0, \sigma_1) = (1, 0)$. Here $4\sigma_0 + 2\sigma_1 = 4$, and from (61) $\overline{\lambda}$ is defined by $g_0(\overline{\lambda}) = 4$. So $\overline{\lambda} = g_0^{-1}(4) = \lambda$. We also have $2 - 2\sigma_0 - \sigma_1 = 0$, and from the definition (65) of τ we have $\tau = 0$. Plugging this into the definition of f (68) gives $f(1, 0) = 1$.

Sub-Case 2.1a:
Now consider $E_{2,1}$, where $\sigma_1 = 0$. Here $\sigma_0 \geq 1/2$, from the definition of E and

$$\partial_0 \log f(\sigma_0, 0) = \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_1(\overline{\lambda})}\right) - 2\log\overline{\lambda}\tau + \log 2 + 2\log(4\sigma_0)$$

Within $E_{2,1}$, we consider two cases. First suppose $\sigma_0 \leq 0.99$. As noted in (75), $\sigma_0 \leq 0.99$ implies $\overline{\lambda}\tau \geq 10^{-4}$. Applying the same bounds as in (76),

$$\partial_0 \log f(\sigma_0, 0) \leq \log\frac{1}{6} - 2\log 10^{-4} + \log 2 + 2\log 4 \leq 21 \tag{78}$$

and we show numerically that $f \leq 1$. The umerical calculations for this case now follow the same outline as above. The precision requirements given there will suffice in this case.

Sub-Case 2.1b:
Now suppose $\sigma_0 \geq 0.99$, still assuming $\sigma_1 = 0$. Here $\overline{\lambda} \leq \lambda$ (see (63)) implies $f_3(\overline{\lambda})/\overline{\lambda}^4 f_1(\overline{\lambda}) \geq 0.01$ by Lemma 8, Appendix B. We have $\tau\log(1 + 1/\tau) = 2 - 2\sigma_0 - \sigma_1 = 2 - 2\sigma_0 \leq 0.02$ and since $\tau\log(1 + 1/\tau)$ is increasing (see (65)), it follows from a numerical calculation that $\tau \leq 0.004$. This implies

$$\frac{1-\sigma_0}{\tau^2} = \frac{\log\left(1+\frac{1}{\tau}\right)}{2\tau} \geq 125\log 250 \tag{79}$$

and

$$\begin{aligned}
\partial_0 \log f(\sigma_0, 0) &= \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^4 f_1(\overline{\lambda})}\right) - 2\log\tau + \log 2 + 4\log(4\sigma_0) \\
&= \log\left(\frac{1-\sigma_0}{\tau^2}\right) - \log\sigma_0 + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^4 f_1(\overline{\lambda})}\right) + \log 2 + 4\log(4\sigma_0) \\
&\geq \log(125\log 250) + \log 0.01 + \log 2 + 2\log 3.96 > 0
\end{aligned}$$

which implies $f(\sigma_0, 0) < f(1, 0) = 1$ for $\sigma_0 \geq 0.99$.

Sub-Case 2.2:
Now consider $E_{2,2}$, i.e. suppose $\sigma_0 + \sigma_1 = 1$ and $\sigma_1 < 0.01$. Then

$$\partial_0 \log f(\sigma_0, 1-\sigma_0) = \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})}\right) - \log\tau + \log(2 + 2\sigma_0) \tag{80}$$

By Lemma 8, Appendix B, $\overline{\lambda} \leq \lambda$ implies

$$\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})} > 0.09.$$

As $\sigma_1 = 1 - \sigma_0$, τ is defined by $\tau\log(1 + 1/\tau) = 2 - 2\sigma_0 - \sigma_1 = \sigma_1$. So $\tau\log(1 + 1/\tau) \leq 0.01$, implying $\tau \leq 0.003$ since $\tau\log(1 + 1/\tau)$ is increasing, and

$$\frac{1-\sigma_0}{\tau} = \frac{\sigma_1}{\tau} = \log\left(1 + \frac{1}{\tau}\right) > \log 333. \tag{81}$$

So,

$$\begin{aligned}
\partial_0 \log f(\sigma_0, 1-\sigma_0) &= \log\left(\frac{1-\sigma_0}{\tau}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})}\right) - \log\sigma_0 + \log(2+2\sigma_0) \\
&\geq \log\log 333 + \log 0.09 + \log 3.98 \\
&> 0
\end{aligned}$$

and for all $0.99 \leq \sigma_0 < 1$, $f(\sigma_0, 1-\sigma_0) < f(1,0) = 1$.

Sub-Case 2.3:
Now consider $E_{2,3}$, i.e. suppose $0 < \sigma_1 < 1-\sigma_0$ and $\sigma_1 < 0.01$. We show that the gradient $\nabla \log f \neq 0$. Assume $(\partial_0 - 2\partial_1)\log f = 0$. By (72) we must have $\sigma_0 \geq (1-\sigma_1)/2 + \sqrt{1-2\sigma_1-\sigma_1^2}/2$. Since $\sigma_1 \leq 0.01$, we can replace this by the weaker bound $\sigma_0 \geq 1 - 1.1\sigma_1$. We trivially have $1-\sigma_0 \geq (2-2\sigma_0-\sigma_1)/2$, so

$$\frac{\sigma_1}{\tau} \geq \frac{1}{1.1}\frac{1-\sigma_0}{\tau} \geq \frac{1}{2.2}\frac{2-2\sigma_0-\sigma_1}{\tau} = \frac{1}{2.2}\log\left(1+\frac{1}{\tau}\right) \tag{82}$$

Since $\tau\log(1+1/\tau) = 2-2\sigma_0-\sigma_1 \leq 1.2\sigma_1 \leq 0.012$, we have $\tau < 0.002$. So $\sigma_1/\tau \geq \log(500)/2.2$.

This allows us to show that if $(\partial_0 - 2\partial_1)\log f = 0$ and $\sigma_1 \leq 0.01$, then $(\partial_0 - \partial_1)\log f \neq 0$. Noting that $4\sigma_0 + 2\sigma_1 \geq 4(1-1.1\sigma_1) + 2\sigma_1 \geq 3.976$,

$$\begin{aligned}
(\partial_0 - \partial_1)\log f &= \log\left(\frac{\sigma_1}{\tau}\right) + \log\left(\frac{f_3(\overline{\lambda})}{\overline{\lambda}^2 f_2(\overline{\lambda})}\right) - \log\sigma_0 + \log(4\sigma_0 + 2\sigma_1) \\
&\geq \log(\log(500)/2.2) + \log 0.09 + \log 3.976 \\
&= 1.038445... - 2.407945... + 1.380276... \\
&> 0
\end{aligned}$$

This shows that $\nabla\log f \neq 0$ in $E_{2,3}$. The boundary of $E_{2,3}$ is contained in $E_0 \cup E_{2,1} \cup E_{2,2} \cup E_1$. Since $f \leq 1$ on the boundary of $E_{2,3}$ and $\nabla \log f \neq 0$ in $E_{2,3}$, it follows that $f \leq 1$ in $E_{2,3}$.

Case Three: $E_3 = \{(\sigma_0, \sigma_1) \in E : 0.99 < \sigma_1 \leq 1\}$.

Further divide E_3 into

$$\begin{aligned}
&E_{3,1} = \{(\sigma_0, \sigma_1) \in E_3 : \sigma_0 + \sigma_1 = 1\}, \\
&E_{3,2} = E_3 \setminus E_{3,1}.
\end{aligned}$$

Sub-Case 3.1:
Consider $E_{3,1}$, i.e. suppose $\sigma_0 + \sigma_1 = 1$ and $\sigma_0 < 0.01$. Then we write, see (80),

$$\partial_0 \log f(\sigma_0, 1-\sigma_0) = \log\left(\frac{1-\sigma_0}{\sigma_0}\right) + \log\left(\frac{1}{g_0(\overline{\lambda})}\right) - \log\overline{\lambda}\tau + \log(2+2\sigma_0) \tag{83}$$

To show that this is positive, we bound $\overline{\lambda}\tau$ from above. From (95) (Appendix A) with $\Delta = 4\sigma_0 + 2\sigma_1$ we have $\tau \leq 1/(4\sigma_0 + 2\sigma_1 - 2)$. For $\overline{\lambda}$, we use the bound derived in Appendix A (96). Note that if $\Delta = 4\sigma_0 + 2\sigma_1$ then $L_2 = \overline{\lambda}$ in (96). So,

$$\overline{\lambda} \leq \frac{12(4\sigma_0 + 2\sigma_1 - 2\sigma_0 - \sigma_1 - 1)}{6 - 3\sigma_0 - 2\sigma_1} \leq 12(2\sigma_0 + \sigma_1 - 1) \leq 12. \tag{84}$$

These two bounds together imply $\overline{\lambda}\tau \leq 6$. For all $0 \leq \overline{\lambda} \leq \lambda$ we have $3 \leq g_0(\overline{\lambda}) \leq 4$ since $3 \leq \Delta_0/\sigma_0 \leq 4$ (see the discussion before (63)).

We conclude that

$$\partial_0 \log f(\sigma_0, 1-\sigma_0) \geq \log \frac{0.99}{0.01} + \log \frac{1}{4} - \log 6 + \log 2 > 0 \tag{85}$$

This implies that for all $(\sigma_0, \sigma_1) \in E_{3,1}$, $f(\sigma_0, \sigma_1) \leq f(0.01, 0.99) \leq 1$, since $(0.01, 0.99) \in E_1$.

Sub-Case 3.2:
Now consider $E_{3,2}$. As noted in (72), any stationary point of $\log f$ must satisfy $\sigma_1 < 1/2$, so $E_{3,2}$ contains no stationary point. The boundary of $E_{3,2}$ is contained in $E_0 \cup E_1 \cup E_{3,1}$, and it has been shown that $f \leq 1$ in each of $E_0, E_1, E_{3,1}$. It follows that $f \leq 1$ in $E_{3,2}$.

This completes the proof of Lemma 4 and Theorem 1.

7 Final Remarks

References

[1] D. Aldous, Asymptotics in the random assignment problem, *Probability Theory and Related Fields* 93 (1992) 507-534.

[2] D. Aldous, The $\zeta(2)$ limit in the random assignment problem, *Random Structures and Algorithms* 4 (2001) 381-418.

[3] F. Avram and D. Bertsimas, The minimum spanning tree constant in geometrical probability and under the independent model: a unified approach, *Annals of Applied Probability* **2** (1992) 113 - 130.

[4] A. Beveridge, A. M. Frieze and C. J. H. McDiarmid, Minimum length spanning trees in regular graphs, *Combinatorica* **18** (1998) 311-333.

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

[6] J.A. Cain, P. Sanders and Nick Wormald, The random graph threshold for k-orientability and a fast algorithm for optimal multiple-choice allocation, *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (2007) 469-476.

[7] C. Cooper, A.M. Frieze, N. Ince, S. Janson and J. Spencer, On the length of a random minimum spanning tree, to appear.

[8] R.Durrett, Probability: Theory and examples, Wadsworth and Brooks/Cole, 1991.

[9] T.I. Fenner and A.M. Frieze, On the connectivity of random m-orientable graphs and digraphs, *Combinatorica* 2 (1982) 347-359.

[10] A.M. Frieze, On the value of a random minimum spanning tree problem, *Discrete Applied Mathematics* **10** (1985) 47-56.

[11] A.M. Frieze, On random symmetric travelling salesman problems, *Mathematics of Operations Research* 29, 878-890.

[12] A.M. Frieze and G.R. Grimmett, The shortest path problem for graphs with random arc-lengths, *Discrete Applied Mathematics* **10** (1985) 57-77.

[13] A.M. Frieze and C.J.H. McDiarmid, On random minimum length spanning trees, *Combinatorica* **9** (1989) 363-374.

[14] A.M. Frieze, M. Ruszinko and L. Thoma, A note on random minimum length spanning trees, *The electronic journal of combinatorics* 7 (2000).

[15] P. Gao, X. Pérez-Giménez and C. M. Sato, Arboricity and spanning-tree packing in random graphs with an application to load balancing, arXiv:1303.3881, extended abstract published in Proc. SODA 2014, Portland, Oregon, 317-326 (2014)

[16] S. L. Hakimi, On the degrees of the vertices of a directed graph, *J. Franklin Inst.* **279**, 290-308 (1965).

[17] S. Janson, The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph, *Random Structures and Algorithms* **7** (1995) 337 - 355.

[18] S. Janson, One, two and three times $\log n/n$ for paths in a complete graph with random weights, *Combinatorics, Probability and Computing* 8 (1999) 347-361.

[19] S. Janson and M. J. Łuczak, A simple solution to the k-core problem, *Random Structures and Algorithms* **30** (2007) 50 - 62.

[20] R.M. Karp, *A patching algorithm for the non-symmetric traveling salesman problem*, SIAM Journal on Computing 8 (1979) 561–573.

[21] W. Kordecki and A. Lyczkowska-Hanćkowiak, Exact Expectation and Variance of Minimal Basis of Random Matroids, *Discussiones Mathematicae Graph Theory* 33 (2013) 277-288.

[22] S. Linusson and J. Wästlund, A proof of Parisi's conjecture on the random assignment problem, *Probability Theory and Related Fields* 128 (2004) 419-440.

[23] T. Łuczak, Size and connectivity of the k-core of a random graph, *Discrete Mathematics* **91** (1991) 61-68.

[24] C. Nair, B. Prabhakar and M. Sharma, Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures, *Random Structures and Algorithms* 27 (2005) 413-444.

[25] C. St. J. A. Nash-Williams, Edge-disjoint spanning trees of finite graphs. *Journal of the London Mathematical Society* **36** (1961) 445-450.

[26] C. St. J. A. Nash-Williams, Decomposition of finite graphs into forests, *Journal of the London Mathematical Society* **39** (1964) 12.

[27] J. Oxley, Matroid Theory, Oxford University Press, Oxford, 1992.

[28] M. Penrose, Random minimum spanning tree and percolation on the n-cube, *Random Structures and Algorithms* **12** (1998) 63 - 82.

[29] B. Pittel, J. Spencer and N. Wormald, Sudden emergence of a giant k-core in a random graph, *Journal of Combinatorial Theory (B)* **67** (1996) 111–151.

[30] J.M. Steele, On Frieze's Zeta(3) limit for lengths of minimal spanning trees, *Discrete Applied Mathematics* **18** (1987) 99–103.

[31] J. Wástlund, An easy proof of the $\zeta(2)$ limit in the random assignment problem, *Electronic Communications in Probability* 14 (2009) 261-269.

[32] J. Wástlund, The mean field traveling salesman and related problems, *Acta Mathematica* 204 (2010) 91-150.

[33] D.J.A. Welsh, Matroid Theory, Academic Press, London, 1976.

Appendix A

This section is concerned with showing that the system of equations (58) under certain conditions has no solution. Throughout the section, assume τ satisfies (55): Recall that $\Delta_3 = 4\tau+4\sigma_0+2\sigma_1-\Delta$,

$$\tau\left(\log\left(1+\frac{1}{\tau}\right)-2\log\left(\frac{4\tau+4\sigma_0+2\sigma_1-\Delta}{4\tau}\right)-\log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right)\right)=2-2\sigma_0-\sigma_1. \tag{86}$$

Here $\lambda = g_0^{-1}(4) \approx 2.688$ is fixed.

Define for $2\sigma_0+\sigma_1+1 \le \Delta \le 4\sigma_0+2\sigma_1$

$$L_1(\sigma_0,\sigma_1,\Delta,\tau) = \lambda_3\sqrt{\frac{\Delta}{4\tau+4\sigma_0+2\sigma_1-\Delta}} \tag{87}$$

and define $L_2(\sigma_0,\sigma_1,\Delta)$ as the unique solution to $G(\sigma_0,\sigma_1,L_2(\sigma_0,\sigma_1,\Delta)) = \Delta$, where G is defined by

$$G(\sigma_0,\sigma_1,x) = \sigma_0 g_0(x) + \sigma_1 g_1(x) + (1-\sigma_0-\sigma_1)g_2(x). \tag{88}$$

This is well defined because each g_i is strictly increasing, and for fixed σ_0,σ_1 we have $G(\sigma_0,\sigma_1,0) = 2\sigma_0+\sigma_1+1 \le \Delta$ and $\lim_{x\to\infty} G(\sigma_0,\sigma_1,x) = \infty$ (see Appendix B). Define

$$R = \left\{(\sigma_0,\sigma_1,\Delta,\tau) \in \mathbb{R}_+^4 : \sigma_0+\sigma_1 \le 1; 2\sigma_0+\sigma_1 \ge 1;\ 2\sigma_0+\sigma_1+1 \le \Delta \le 4\sigma_0+2\sigma_1;\ (86) \text{ holds.}\right\}$$

We prove that the system (58) is inconsistent by proving

Lemma 5. *Let $(\sigma_0,\sigma_1,\Delta,\tau) \in R$. Then $L_1(\sigma_0,\sigma_1,\Delta,\tau) > L_2(\sigma_0,\sigma_1,\Delta)$*

Proof. Define $L(\sigma_0,\sigma_1,\Delta,\tau) = L_1(\sigma_0,\sigma_1,\Delta,\tau) - L_2(\sigma_0,\sigma_1,\Delta)$. We will bound $|\nabla L|$ in R in order to show numerically that $L \ge 0$. However, ∇L is unbounded for Δ close to 4 and $2\sigma_0+\sigma_1$ close to 1. For this reason, define

$$\begin{aligned}
R_1 &= \{(\sigma_0,\sigma_1,\Delta,\tau) \in R : \Delta \ge 3.6\},\\
R_2 &= \{(\sigma_0,\sigma_1,\Delta,\tau) \in R : 2\sigma_0+\sigma_1 \le 1.1\},\\
R_3 &= R \setminus (R_1 \cup R_2).
\end{aligned}$$

Analytical proofs will be provided for R_1, R_2, and a numerical calculation will have to suffice for R_3.

First note that for any σ_0, σ_1 we have $L_2(\sigma_0, \sigma_1, 2\sigma_0+\sigma_1+1) = 0$, since $G(\sigma_0, \sigma_1, 0) = 2\sigma_0+\sigma_1+1$, see (88). Here we use the fact that $g_i(0) = 3 - i$, $i = 0, 1, 2$ by Lemma 6, Appendix B. This implies that $L_1(\sigma_0, \sigma_1, 2\sigma_0+\sigma_1+1, \tau) \geq 0 = L_2(\sigma_0, \sigma_1, 2\sigma_0+\sigma_1+1)$, and we may therefore assume $\Delta > 2\sigma_0 + \sigma_1 + 1$.

We proceed by finding an upper bound for τ, given that it satisfies (86). Fix $\sigma_0, \sigma_1, \Delta$ and define

$$r(\zeta) = \zeta\left(\log\left(1+\frac{1}{\zeta}\right) - 2\log\left(\frac{4\tau+4\sigma_0+2\sigma_1-\Delta}{4\tau}\right) - \log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right)\right) \tag{89}$$

We first derive a lower bound $r_1(\zeta) \leq r(\zeta)$.

For $x \geq 0$ we have $x - x^2/2 \leq \log(1+x) \leq x$. This implies, that for all ζ,

$$2\zeta\log\left(1+\frac{4\sigma_0+2\sigma_1-\Delta}{4\zeta}\right) \leq 2\zeta\frac{4\sigma_0+2\sigma_1-\Delta}{4\zeta} = \frac{4\sigma_0+2\sigma_1-\Delta}{2} \tag{90}$$

Let $h(x) = \log f_3(x) - 4\log x$. Then $h'(x) = f_2(x)/f_3(x) - 4/x$, and we note that $h'(\lambda) = 0$, by definition of λ. The second derivative is $h''(x) = f_1(x)/f_3(x) - f_2(x)^2/f_3(x)^2 + 4/x^2$. Substituting $f_1(x) = f_3(x) + x + x^2/2$ and $f_2(x) = f_3(x) + x^2/2$, for all $x \geq \lambda$

$$\begin{aligned} h''(x) &= \frac{4}{x^2} + 1 + \frac{x+x^2/2}{f_3(x)} - 1 - \frac{x^2}{f_3(x)} - \frac{x^4}{4f_3(x)^2} \\ &= \frac{4}{x^2} - \frac{x^2-2x}{2f_3(x)} - \frac{x^4}{4f_3(x)^2} \end{aligned}$$

Since $x \geq \lambda > 2$ we have $x^2 - 2x > 0$, and $f_3(x) < e^x$ implies

$$\begin{aligned} h''(x) &= \frac{4}{x^2} - \frac{x^2-2x}{2f_3(x)} - \frac{x^4}{4f_3(x)^2} \\ &\leq \frac{4}{x^2} - \frac{x^2-2x}{2e^x} \\ &\leq \frac{4}{x^2} + \frac{2x}{2e^x} \\ &\leq \frac{4}{x^2} + x^{1-\lambda} \end{aligned}$$

Here we use the fact that $e^x \geq x^\lambda$ for $x \geq \lambda$, since $\lambda < e$. Since $4x^{-2} + x^{1-\lambda}$ is decreasing, we have $h''(x) \leq 4\lambda^{-2} + \lambda^{1-\lambda} < 3/4$ for all $x \geq \lambda$.

By Taylor's theorem, for some $x \in [\lambda, \lambda_3]$

$$\begin{aligned} \log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right) &= h(\lambda_3) - h(\lambda) \\ &= h(\lambda) + h'(\lambda)(\lambda_3-\lambda) + \frac{1}{2}h''(x)(\lambda_3-\lambda)^2 - h(\lambda) \\ &\leq \frac{3}{8}(\lambda_3-\lambda)^2 \end{aligned}$$

Another application of Taylor's theorem lets us bound

$$\lambda_3 - \lambda = g_0^{-1}\left(4 + \frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right) - g_0^{-1}(4).$$

By Lemma 7, Appendix B, we have $g_0'(x) \geq g_0'(\lambda) \geq 1/2$ for $x \geq \lambda$, so $dg_0^{-1}(y)/dy \leq 2$ for $y \geq 4$, and for some $y \geq 4$

$$\lambda_3 = \lambda + \frac{dg_0^{-1}(y)}{dy}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right) \leq \lambda + 2\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau} \tag{91}$$

and so

$$\log\left(\frac{\lambda^4}{\lambda_3^4}\frac{f_3(\lambda_3)}{f_3(\lambda)}\right) \leq \frac{3}{8}(\lambda_3 - \lambda)^2 \leq \frac{3}{2}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right)^2 \tag{92}$$

Define τ_1 as the unique solution ζ to

$$2 - 2\sigma_0 - \sigma_1 = r_1(\zeta)$$

where

$$r_1(\zeta) = \zeta\left(\log\left(1 + \frac{1}{\zeta}\right) - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2\zeta} - \frac{3}{2}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\zeta}\right)^2\right).$$

Then $r_1(\zeta) \leq r(\zeta)$, and $r_1(\zeta)$ is strictly increasing. So, since $r_1(\tau_1) = r(\tau) = 2 - 2\sigma_0 - \sigma_1$, it follows that $\tau \leq \tau_1$.

Case of R_1:
Now fix $(\sigma_0, \sigma_1, \Delta, \tau) \in R_1$, i.e. suppose $\Delta \geq 3.6$. Then

$$\begin{aligned}
r_1\left(\frac{3}{4}\right) &= \frac{3}{4}\log\left(1 + \frac{4}{3}\right) - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2} - 2(4\sigma_0 + 2\sigma_1 - \Delta)^2 \\
&= \frac{3}{4}\log\frac{7}{3} - 2\sigma_0 - \sigma_1 + \frac{\Delta}{2} - 2(4\sigma_0 + 2\sigma_1 - \Delta)^2 \\
&\geq \frac{3}{4}\log\frac{7}{3} - 2\sigma_0 - \sigma_1 + \frac{3.6}{2} - 2(4 - 3.6)^2 \\
&> 2 - 2\sigma_0 - \sigma_1
\end{aligned}$$

We have $\lim_{\zeta\to 0} r_1(\zeta) \leq 0$, and r_1 is continous and increasing, so $\tau \leq \tau_1 < 3/4$. Since $\Delta \geq 3.6$ and $2\sigma_0 + \sigma_1 \leq 2$,

$$\Delta - (4\tau + 4\sigma_0 + 2\sigma_1 - \Delta) \geq 2\Delta - 3 - 4\sigma_0 - 2\sigma_1 \geq 7.2 - 7 > 0 \tag{93}$$

This implies that

$$L_1(\sigma_0, \sigma_1, \Delta) = \lambda_3\sqrt{\frac{\Delta}{4\tau + 4\sigma_0 + 2\sigma_1 - \Delta}} > \lambda_3 \tag{94}$$

Note that

$$G(\sigma_0, \sigma_1, \lambda) \geq G(\sigma_0, \sigma_1, \overline{\lambda}) = 4\sigma_0 + 2\sigma_1 \geq \Delta$$

implies that

$$L_2(\sigma_0, \sigma_1, \Delta) \leq \lambda = g_0^{-1}(4).$$

Also note that by (38) and (52) we have

$$\lambda_3 = g_0^{-1}\left(\frac{\Delta_3}{\tau}\right) = g_0^{-1}\left(4 + \frac{4\sigma_0 + 2\sigma_1 - \Delta}{\tau}\right) \geq g_0^{-1}(4) = \lambda,$$

since g_0^{-1} is increasing (Lemma 7, Appendix B). So

$$L_1(\sigma_0, \sigma_1, \Delta, \tau) > \lambda_3 \geq \lambda \geq L_2(\sigma_0, \sigma_1, \Delta)$$

for $(\sigma_0, \sigma_1, \Delta, \tau) \in R_1$.

Case of R_2, R_3:
For R_2, R_3 we will need a new bound on τ. Since $x - x^2/2 \leq \log(1+x)$ for all $x \geq 0$,

$$r_1(\zeta) \geq r_2(\zeta) = \zeta\left(\frac{1}{\zeta} - \frac{1}{2\zeta^2} - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{2\zeta} - \frac{3}{2}\left(\frac{4\sigma_0 + 2\sigma_1 - \Delta}{\zeta}\right)^2\right).$$

Let τ_2 be defined by $r_2(\tau_2) = 2 - 2\sigma_0 - \sigma_1$, which can be solved for τ_2;

$$\tau_2 = \frac{1 + 3(4\sigma_0 + 2\sigma_1 - \Delta)^2}{\Delta - 2}.$$

It follows from $r(\tau) \geq r_2(\tau)$ and the fact that r_2 is increasing that

$$\tau \leq \frac{1 + 3(4\sigma_0 + 2\sigma_1 - \Delta)^2}{\Delta - 2}. \tag{95}$$

An upper bound for $L_2(\sigma_0, \sigma_1, \Delta)$ will follow from bounding the partial derivative of $G(\sigma_0, \sigma_1, x)$ with respect to x. We have $g_0' \geq 1/4$, $g_1' \geq 1/3$ and $g_2' \geq 1/2$ by Lemma 7 (Appendix B), so

$$\begin{aligned}
\frac{\partial}{\partial x} G(\sigma_0, \sigma_1, x) &= \sigma_0 g_0'(x) + \sigma_1 g_1'(x) + (1 - \sigma_0 - \sigma_1) g_2'(x) \\
&\geq \frac{\sigma_0}{4} + \frac{\sigma_1}{3} + \frac{1 - \sigma_0 - \sigma_1}{2} \\
&= \frac{6 - 3\sigma_0 - 2\sigma_1}{12}
\end{aligned}$$

and $G(\sigma_0, \sigma_1, 0) = 2\sigma_0 + \sigma_1 + 1$ implies

$$\begin{aligned}
\Delta &= G(\sigma_0, \sigma_1, L_2(\Delta)) \\
&\geq G(\sigma_0, \sigma_1, 0) + \min_x \frac{\partial}{\partial x} G(\sigma_0, \sigma_1, x) L_2(\Delta) \\
&\geq 2\sigma_0 + \sigma_1 + 1 + \frac{6 - 3\sigma_0 - 2\sigma_1}{12} L_2(\Delta)
\end{aligned}$$

So

$$L_2(\Delta) \leq \frac{12(\Delta - 2\sigma_0 - \sigma_1 - 1)}{6 - 3\sigma_0 - 2\sigma_1}. \tag{96}$$

So, to show $L_1(\sigma_0, \sigma_1, \Delta, \tau) \geq L_2(\sigma_0, \sigma_1, \Delta)$, it is enough to show that

$$\lambda_3 \sqrt{\frac{\Delta}{4\tau + 4\sigma_0 + 2\sigma_1 - \Delta}} > \frac{12(\Delta - 2\sigma_0 - \sigma_1 - 1)}{6 - 3\sigma_0 - 2\sigma_1} \tag{97}$$

Solving for τ, this is equivalent to showing

$$\tau < \Delta \left[\frac{\lambda_3(6 - 3\sigma_0 - 2\sigma_1)}{24(\Delta - 2\sigma_0 - \sigma_1 - 1)} \right]^2 - \frac{4\sigma_0 + 2\sigma_1 - \Delta}{4} \tag{98}$$

and by (95), and $\lambda_3 \geq \lambda$, it is enough to show

$$\frac{1 + 3(4\sigma_0 + 2\sigma_1 - \Delta)^2}{\Delta - 2} < \Delta \left[\frac{\lambda(6 - 3\sigma_0 - 2\sigma_1)}{24(\Delta - 2\sigma_0 - \sigma_1 - 1)} \right]^2 - \frac{(4\sigma_0 + 2\sigma_1 - \Delta)}{4} \tag{99}$$

for $(\sigma_0, \sigma_1, \Delta, \tau) \in R_2 \cup R_3$.

Case of R_2:
Consider R_2, i.e. suppose $2\sigma_0 + \sigma_1 \leq 1.1$. Then $4\sigma_0 + 2\sigma_1 - \Delta \leq 2\sigma_0 + \sigma_1 - 1 \leq 0.1$ since $\Delta \geq 2\sigma_0 + \sigma_1 + 1$. This implies

$$\frac{1 + 3(4\sigma_0 + 2\sigma_1 - \Delta)^2}{\Delta - 2} \leq \frac{1.03}{\Delta - 2} \tag{100}$$

Furthermore, $6 - 3\sigma_0 - 2\sigma_1 \geq 4.9 - \sigma_0 - \sigma_1 \geq 3.9$, while $2\sigma_0 + \sigma_1 \geq 1$ implies $\Delta - 2\sigma_0 - \sigma_1 - 1 \leq \Delta - 2$. We have $\lambda_3 \geq \lambda = g_0^{-1}(4) > 2.5$. So it holds that

$$\Delta \left[\frac{\lambda_3(6 - 3\sigma_0 - 2\sigma_1)}{24(\Delta - 2\sigma_0 - \sigma_1 - 1)} \right]^2 - \frac{(4\sigma_0 + 2\sigma_1 - \Delta)}{4} > \Delta \left[\frac{2.5 \times 3.9}{24(\Delta - 2)} \right]^2 - 0.025 \tag{101}$$

and it is enough to show that

$$\frac{1.03}{\Delta - 2} \leq \Delta \left[\frac{2.5 \times 3.9}{24(\Delta - 2)} \right]^2 - 0.025 \tag{102}$$

We have $\Delta \geq 2\sigma_0 + \sigma_1 + 1 > 2$, so multipling both sides by $\Delta - 2 > 0$, this amounts to solving a second-degree polynomial inequality. Numerically, the zeros of the resulting second-degree polynomial are $\Delta \approx -33$ and $\Delta \approx 2.37$. The inequality holds at $\Delta = 2.3$, and so it holds for all $2 < \Delta \leq 2.37$. In particular, it holds for $2\sigma_0 + \sigma_1 + 1 < \Delta \leq 4\sigma_0 + 2\sigma_1$ when $1 \leq 2\sigma_0 + \sigma_1 \leq 1.1$.

Case of R_3:
Lastly, consider R_3. Here more extensive numerical methods will be used, and we begin by reducing the analysis from three variables to two. Divide R_3 into four subregions,

$$\begin{aligned}
R_{3,1} &= \{(\sigma_0, \sigma_1, \Delta) \in R_3 : 1/2 \leq \sigma_1 \leq 1\}, \\
R_{3,2} &= \{(\sigma_0, \sigma_1, \Delta) \in R_3 : 1/4 \leq \sigma_1 < 1/2\}, \\
R_{3,3} &= \{(\sigma_0, \sigma_1, \Delta) \in R_3 : 1/8 \leq \sigma_1 < 1/4\}, \\
R_{3,4} &= \{(\sigma_0, \sigma_1, \Delta) \in R_3 : 0 \leq \sigma_1 < 1/8\}.
\end{aligned}$$

Define

$$u_1 = 5.5, \quad u_2 = 5.75, \quad u_3 = 5.875, \quad u_4 = 5.9375.$$

Then

$$6 - 3\sigma_0 - 2\sigma_1 = \left(6 - \frac{\sigma_1}{2}\right) - 3\sigma_0 - \frac{3\sigma_1}{2} \geq u_i - \frac{3(2\sigma_0 + \sigma_1)}{2}$$

in $R_{3,i}$, $i = 1, 2, 3, 4$.

Fixing i, (99) will hold in $R_{3,i}$ if we can show that

$$\frac{1+3(4\sigma_0+2\sigma_1-\Delta)^2}{\Delta-2} \leq \Delta \left[\frac{\lambda(u_i-3(2\sigma_0+\sigma_1)/2)}{24(\Delta-2\sigma_0-\sigma_1-1)}\right]^2 - \frac{(4\sigma_0+2\sigma_1-\Delta)}{4} \tag{103}$$

Note that σ_0, σ_1 only appear as $\Sigma = 2\sigma_0+\sigma_1$ in (103). For this reason we clear denominators in (103) and define for $i = 1,2,3,4$,

$$\begin{aligned} \varphi_i(\Sigma,\Delta) &= \lambda^2\Delta(\Delta-2)(u_i-3\Sigma/2)^2 - 144(\Delta-2)(\Delta-\Sigma-1)^2(2\Sigma-\Delta) \\ &\quad -576(\Delta-\Sigma-1)^2 - 1728(\Delta-\Sigma-1)^2(2\Sigma-\Delta)^2. \end{aligned}$$

In which case, (103) is equivalent to $\varphi(\Sigma,\Delta) \geq 0$.

In $R_{3,1}$ we have $1.1 \leq \Sigma \leq 1.5$ since $2\sigma_0+\sigma_1 \geq 1.1$ is assumed, and $\sigma_1 \geq 1/2$ and $\sigma_0+\sigma_1 \leq 1$ imply $2\sigma_0+\sigma_1 \leq 2-\sigma_1 \leq 1.5$. For this reason define

$$\begin{aligned} &\widetilde{R}_{3,1} = \{(\Sigma,\Delta) : 1.1 \leq \Sigma \leq 1.5, \Sigma+1 \leq \Delta \leq 2\Sigma\} \\ &\widetilde{R}_{3,2} = \{(\Sigma,\Delta) : 1.5 \leq \Sigma \leq 1.75, \Sigma+1 \leq \Delta \leq 2\Sigma\}, \\ &\widetilde{R}_{3,3} = \{(\Sigma,\Delta) : 1.75 \leq \Sigma \leq 1.875, \Sigma+1 \leq \Delta \leq \min\{2\Sigma, 3.6\}\}, \\ &\widetilde{R}_{3,4} = \{(\Sigma,\Delta) : 1.875 \leq \Sigma \leq 2, \Sigma+1 \leq \Delta \leq \min\{2\Sigma, 3.6\}\}. \end{aligned}$$

Here $\Sigma+1 \leq \Delta \leq 2\Sigma$ is (51).

Equation (103) will follow from showing that $\varphi_i(\Sigma,\Delta) \geq 0$ in $\widetilde{R}_{3,i}$, $i = 1,2,3,4$.

The φ_i are degree four polynomials, and bounds on $|\nabla\varphi_i|$ are found by applying the triangle inequality to the partial derivatives of φ_i. The same bound will be applied to $\nabla\varphi_i$ for all i. using,

$$2 \leq \Sigma+1 \leq \Delta \leq 2\Sigma \leq 4, \quad u_i \leq 6, \quad \lambda < 3$$

from which we obtain

$$\begin{aligned} &u_i - \frac{3\Sigma}{2} \leq \frac{9}{2}, \; -1 \leq 3\Sigma-2\Delta+1 \leq 1, -2 \leq 4\Sigma-3\Delta+2 \leq 1, \\ &(\Delta-\Sigma-1)(2\Sigma-\Delta) \leq \frac{(\Sigma-1)^2}{4} \leq \frac{1}{4}. \end{aligned}$$

we have

$$\begin{aligned} \left|\frac{\partial\varphi_i}{\partial\Sigma}\right| &= |-3\lambda^2\Delta(\Delta-2)(u_i-3\Sigma/2) + 288(\Delta-2)(\Delta-\Sigma-1)(2\Sigma-\Delta) \\ &\quad -288(\Delta-2)(\Delta-\Sigma-1)^2 + 1152(\Delta-\Sigma-1) \\ &\quad +3456(\Delta-\Sigma-1)(2\Sigma-\Delta)^2 - 6912(\Delta-\Sigma-1)^2(2\Sigma-\Delta)| \\ &\leq 3\lambda^2\Delta(\Delta-2)(u_i-3\Sigma/2) + 288(\Delta-2)(\Delta-\Sigma-1)|3\Sigma-2\Delta+1| \\ &\quad +1152(\Delta-\Sigma-1) + 3456(\Delta-\Sigma-1)(2\Sigma-\Delta)|4\Sigma-3\Delta+2| \\ &\leq 27\cdot4\cdot2\cdot9/2 + 288\cdot2\cdot2\cdot1 + 1152\cdot2 + 3456\cdot3/4\cdot2 \\ &= 9612 \end{aligned}$$

For Δ,

$$
\begin{aligned}
\left|\frac{\partial \varphi_i}{\partial \Delta}\right| &= |\lambda^2\Delta(u_i-3\Sigma/2)^2+\lambda^2(\Delta-2)(u_i-3\Sigma/2)^2-144(\Delta-\Sigma-1)^2(2\Sigma-\Delta)\\
&\quad -288(\Delta-2)(\Delta-\Sigma-1)(2\Sigma-\Delta)+144(\Delta-2)(\Delta-\Sigma-1)^2\\
&\quad -1152(\Delta-\Sigma-1)-3456(\Delta-\Sigma-1)(2\Sigma-\Delta)^2+3456(\Delta-\Sigma-1)^2(2\Sigma-\Delta)|\\
&\le \lambda^2(2\Delta-2)(u_i-3\Sigma/2)^2+144(\Delta-\Sigma-1)^2(2\Sigma-\Delta)+288(\Delta-2)(\Delta-\Sigma-1)(2\Sigma-\Delta)\\
&\quad +144(\Delta-2)(\Delta-\Sigma-1)^2+1152(\Delta-\Sigma-1)+3456(\Delta-\Sigma-1)(2\Sigma-\Delta)|2\Delta-3\Sigma-1|\\
&\le 9\cdot 6\cdot(9/2)^2+144\cdot 2^2\cdot 1+288\cdot 2\cdot 2\cdot 1+144\cdot 2\cdot 2^2+1152\cdot 2+3456\cdot 3/4\cdot 1\\
&= 8383.5
\end{aligned}
$$

so $|\nabla\varphi_i| \le 12755$ for $i=1,2,3,4$.

For each i, a grid $\mathcal{P}_i \subseteq \widetilde{R}_{3,i}$ of $4\cdot 10^6$ points is generated such that for each $x \in \widetilde{R}_{3,i}$, there exists an $x_0 \in \mathcal{P}_i$ for which $|x-x_0| \le 0.001$. On this grid, φ_i is calculated numerically, and it is found that

$$
\min_{x_0\in\mathcal{P}_i} \varphi_i(x_0) = \begin{cases} 22.49, & i=1\\ 25.50, & i=2\\ 27.08, & i=3\\ 19.04, & i=4 \end{cases} \tag{104}
$$

So for any i and any $x \in \widetilde{R}_{3,i}$, there exists an x_0 such that $|\varphi_i(x)-\varphi_i(x_0)| \le |\nabla\varphi_i||x-x_0| \le 12755\cdot 0.001 < 13$, which implies $\varphi_i(x) > \varphi_i(x_0)-13 > 0$. This proves (99) for $\sigma_0,\sigma_1,\Delta \in R_3$.

□

Appendix B

This section is concerned with the functions

$$
f_0(x)=e^x \text{ and } f_k(x)=e^x-\sum_{j=0}^{k-1}\frac{x^j}{j!},\ x\ge 0,\ k=1,2,3,
$$

and the related functions

$$
g_0(x)=\frac{xf_2(x)}{f_3(x)},\quad g_1(x)=\frac{xf_1(x)}{f_2(x)},\quad g_2(x)=\frac{xf_0(x)}{f_1(x)}. \tag{105}
$$

Since $f_k(0)=0$ for $k\ge 1$, we define $g_i(0)=\lim_{x\to 0} g_i(x)=3-i$. Note that

$$
\frac{d}{dx}f_k(x)=f_{k-1}(x),\quad k\ge 1 \tag{106}
$$

Lemma 6. *For all $x \ge 0$ and $i=0,1,2$,*

$$
x < g_i(x) \le 3-i+x \tag{107}
$$

with equality in the upper bound if and only if $x=0$.

Proof. Fix i. By definition, $g_i(0) = 3 - i$. For $x > 0$ consider

$$g_i(x) - x = \frac{x f_{2-i}(x)}{f_{3-i}(x)} - x = \frac{x(f_{2-i}(x) - f_{3-i}(x))}{f_{3-i}(x)} = \frac{x^{3-i}}{(2-i)! f_{3-i}(x)}. \tag{108}$$

Since $f_{3-i}(x) > 0$ we have $g_i(x) - x > 0$. Now

$$(3-i)(2-i)! f_{3-i}(x) - x^{3-i} = (3-i)! \sum_{k \geq 3-i} \frac{x^k}{k!} - x^{3-i} = (3-i)! \sum_{k \geq 4-i} \frac{x^k}{k!} > 0 \tag{109}$$

for $x > 0$, implying $g_i(x) - x < 3 - i$. □

Lemma 7. *The functions g_0, g_1, g_2 are convex, and $g_i'(x) \geq 1/(4-i)$ for $x \geq 0$, $i = 0, 1, 2$.*

Proof. Consider g_0. Since $f_2(x) = f_3(x) + x^2/2$, g_0 can be written as

$$g_0(x) = \frac{x f_2(x)}{f_3(x)} = x + \frac{x^3}{2 f_3(x)} \tag{110}$$

Let $q(x) = f_3(x)/x^3 = \sum_{j \geq 0} x^j/(j+3)!$. Then $g_0(x) = x + 1/2q(x)$, and

$$g_0'(x) = 1 - \frac{q'(x)}{2q(x)^2}, \quad g_0''(x) = \frac{2q'(x)^2 - q(x)q''(x)}{2q(x)^3} \tag{111}$$

and we show that $2q'(x)^2 - q(x)q''(x) \geq 0$. We have $q'(x) = \sum_{j \geq 0} (j+1)x^j/(j+4)!$ and $q''(x) = \sum_{j \geq 0} (j+1)(j+2)x^j/(j+5)!$, so the jth Taylor coefficient of $2q'(x)^2 - q(x)q''(x)$ is given by

$$\begin{aligned}
[x^j][2q'(x)^2 - q(x)q''(x)] &= \sum_{\substack{j_1, j_2 \geq 0 \\ j_1 + j_2 = j}} 2\frac{(j_1+1)}{(j_1+4)!}\frac{(j_2+1)}{(j_2+4)!} - \frac{1}{(j_1+3)!}\frac{(j_2+1)(j_2+2)}{(j_2+5)!} \\
&= \sum_{j_1, j_2} \frac{2(j_1+1)(j_2+1)(j_2+5) - (j_1+4)(j_2+1)(j_2+2)}{(j_1+4)!(j_2+5)!} \\
&= \sum_{j_1, j_2} \frac{(j_2+1)(2(j_1+1)(j_2+5) - (j_1+4)(j_2+2))}{(j_1+4)!(j_2+5)!} \\
&= \sum_{j_1, j_2} \frac{(j_2+1)(j_1 j_2 + 8j_1 - 2j_2 + 2)}{(j_1+4)!(j_2+5)!}
\end{aligned}$$

It is seen that this is positive for $j = 0, 1, 2$. Let $Q(j_1, j_2)$ denote the summand. If $j \geq 3$ then since $Q(j_1, j_2) \geq 0$ whenever $j_1 \geq 2$.

$$\begin{aligned}
\sum_{\substack{j_1, j_2 \geq 0 \\ j_1 + j_2 = j}} Q(j_1, j_2) &\geq Q\left(\left\lfloor \frac{j}{2} \right\rfloor, \left\lceil \frac{j}{2} \right\rceil\right) + Q(0, j) + Q(1, j-1) \\
&= \frac{(\lfloor j/2 \rfloor + 1)\left(\lceil j/2 \rceil \lfloor j/2 \rfloor + 8\lceil j/2 \rceil - 2\lfloor j/2 \rfloor + 2\right)}{(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{2(j^2-1)}{24(j+5)!} - \frac{j^2 - 11j}{120(j+4)!} \\
&\geq \frac{j^3}{8(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{j^2}{12(j+5)!} - \frac{j^2 - 11j}{120(j+4)!} \\
&= \frac{j^3}{8(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{10j^2 + (j^2 - 11j)(j+5)}{120(j+5)!} \\
&\geq \frac{j^3}{8}\left(\frac{1}{(\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!} - \frac{1}{15(j+5)!}\right).
\end{aligned}$$

(To get the final inequality, consider $j \leq 11$ and $j > 11$ seperately).

It remains to show that $a_j = (\lceil j/2 \rceil + 4)!(\lfloor j/2 \rfloor + 5)!$ is smaller than $b_j = 15(j+5)!$ for $j \geq 3$. For $j = 3$, $a_3 = 6! \cdot 6! < 15 \cdot 8! = b_3$. For the induction step, $a_{j+1}/a_j \leq j/2 + 6$ while $b_{j+1}/b_j = j + 6$, so $a_3 < b_3$ implies $a_j < b_j$ for all $j \geq 3$. So $2q'(x)^2 - q(x)q''(x) \geq 0$, and it follows that g_0 is convex. Similar arguments show that g_1, g_2 are convex.

For $i = 0, 1$,

$$\begin{aligned} g_i'(x) &= \frac{f_{2-i}(x)}{f_{3-i}(x)} + \frac{xf_{1-i}(x)}{f_{3-i}(x)} - \frac{xf_{2-i}(x)^2}{f_{3-i}(x)^2} \\ &= \frac{f_{2-i}(x)f_{3-i}(x) + xf_{1-i}(x)f_{3-i}(x) - xf_{3-i}(x)^2}{f_{3-i}(x)^2}. \end{aligned}$$

Now

$$\begin{aligned} &f_{2-i}(x)f_{3-i}(x) + xf_{1-i}(x)f_{3-i}(x) - xf_{3-i}(x)^2 = \\ &x^{6-2i}\left(\frac{1}{(2-i)!(4-i)!} + \frac{1}{(3-i)!^2} + \frac{1}{(1-i)!(4-i)!} + \frac{1}{(2-i)!(3-i)!} - \frac{2}{(2-i)!(3-i)!} + O(x)\right) \\ &= x^{6-2i}\left(\frac{1}{(3-i)!(4-i)!} + O(x)\right). \end{aligned}$$

And

$$f_{3-i}(x)^2 = x^{6-2i}\left(\frac{1}{(3-i)!^2} + O(x)\right).$$

So, for $i = 0, 1$ we have

$$g_i'(x) = \frac{1}{4-i} + O(x).$$

For $i = 2$ we have

$$g_2'(x) = \frac{e^x}{f_1(x)} + \frac{xe^x}{f_1(x)} - \frac{xe^{2x}}{f_1(x)^2} = e^x\left(\frac{f_1(x)(1+x) - xe^x}{f_1(x)^2}\right) = e^x\left(\frac{\frac{x^2}{2} + O(x^3)}{x^2 + O(x^3)}\right) = \frac{1}{2} + O(x).$$

And by the convexity of g_i we have $g_i'(x) \geq 1/(4-i)$ for all $x \geq 0$. $\square$

Lemma 7 allows us to define inverses g_i^{-1}, $i = 0, 1, 2$.

Lemma 8. *For $0 \leq x \leq \lambda = g_0^{-1}(4)$, the following inequalities hold.*

$$\begin{aligned} &(i) && 1 \leq \frac{f_2(x)^2}{f_1(x)f_3(x)} \leq 2 \\ &(ii) && 0.09 < \frac{f_3(x)}{x^2 f_1(x)} \leq \frac{1}{6} \\ &(iii) && \frac{f_2(x)}{xf_1(x)} \leq \frac{1}{3} \\ &(iv) && 0.01 < \frac{f_3(x)}{x^4 f_1(x)} \\ &(v) && 0.09 < \frac{f_3(x)}{x^2 f_2(x)} \end{aligned}$$

Proof. For the lower bound, let $x > 0$ and consider the equation $f_2(x)^2 = f_1(x)f_3(x)$. By definition of f_i, this equation can be written as

$$(e^x - 1 - x)^2 = (e^x - 1)\left(e^x - 1 - x - \frac{x^2}{2}\right) \tag{112}$$

Expanding and reordering terms, we have

$$e^x\left(x + \frac{x^2}{2}\right) = x + \frac{x^2}{2} \tag{113}$$

which clearly has no positive solution. Since $f_2(0)^2/f_1(0)f_3(0) = 3/2 > 1$, this implies that $f_2(x)^2/f_1(x)f_3(x) > 1$ for all $x \geq 0$.

For the upper bound we consider the equation $f_2(x)^2 = 2f_1(x)f_3(x)$. This simplifies to

$$(e^x - 1)^2 = x^2e^x \text{ or } e^x = 1 + xe^{x/2}$$

which has no positive solution.

Since g_0, g_1 are increasing by Lemma 7 and positive, the expressions in (ii) – (v) are all decreasing;

$$\frac{f_3(x)}{x^2f_1(x)} = \frac{1}{g_0(x)g_1(x)}, \quad \frac{f_2(x)}{xf_1(x)} = \frac{1}{g_1(x)}, \quad \frac{f_3(x)}{x^4f_1(x)} = \frac{1}{x^2g_0(x)g_1(x)}, \quad \frac{f_3(x)}{x^2f_2(x)} = \frac{1}{xg_0(x)} \tag{114}$$

The upper bounds are obtained by noting that $g_i(0) = 3 - i$ by Lemma 6, while the lower bounds are obtained numerically by letting $x = 2.688 > \lambda$. $\square$