# On edge disjoint spanning trees in a randomly weighted complete graph 

Alan Frieze* Tony Johansson ${ }^{\dagger}$<br>Department of Mathematical Sciences<br>Carnegie Mellon University<br>Pittsburgh PA 15213<br>U.S.A.

June 13, 2018


#### Abstract

Assume that the edges of the complete graph $K_{n}$ are given independent uniform $[0,1]$ edges weights. We consider the expected minimum total weight $\mu_{k}$ of $k \geq 2$ edge disjoint spanning trees. When $k$ is large we show that $\mu_{k} \approx k^{2}$. Most of the paper is concerned with the case $k=2$. We show that $\mu_{2}$ tends to an explicitly defined constant and that $\mu_{2} \approx 4.1704288 \ldots$.


## 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 [12] 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 $k \rightarrow \infty$ and a rather difficult analysis for $k=2$.

Given a connected simple graph $G=(V, E)$ with edge lengths $\mathbf{x}=\left(x_{e}: e \in E\right)$ and a positive integer $k$, let $\mathrm{mst}_{k}(G, \mathbf{x})$ denote the minimum length of $k$ edge disjoint spanning trees of $G$. $\left(\operatorname{mst}_{k}(G)=\infty\right.$

[^0]if such trees do not exist.) When $\mathbf{X}=\left(X_{e}: e \in E\right)$ is a family of independent random variables, each uniformly distributed on the interval $[0,1]$, denote the expected value $\mathbf{E}\left[\operatorname{mst}_{k}(G, \mathbf{X})\right]$ by $\operatorname{mst}_{k}(G)$.

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

$$
\lim _{n \rightarrow \infty} \operatorname{mst}_{1}\left(K_{n}\right)=\zeta(3)=\sum_{k=1}^{\infty} \frac{1}{k^{3}} .
$$

Generalisations and refinements of this result were subsequently given in Steele [30], Frieze and McDiarmid [14], Janson [17], Penrose [28], Beveridge, Frieze and McDiarmid [4], Frieze, Ruszinko and Thoma [15] 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

$$
\mu_{k}^{*}=\liminf _{n \rightarrow \infty} \operatorname{mst}_{k}\left(K_{n}\right) \text { and } \mu_{k}^{* *}=\limsup _{n \rightarrow \infty} \operatorname{mst}_{k}\left(K_{n}\right) .
$$

Conjecture 1. $\mu_{k}^{*}=\mu_{k}^{* *}$ i.e. $\lim _{n \rightarrow \infty} m s t_{k}\left(K_{n}\right)$ exists.

## Theorem 1.

$$
\lim _{k \rightarrow \infty} \frac{\mu_{k}^{*}}{k^{2}}=\lim _{k \rightarrow \infty} \frac{\mu_{k}^{* *}}{k^{2}}=1 .
$$

Theorem 2. With $f_{k}$ and $c_{2}^{\prime} \approx 3.59$ and $\lambda_{2}^{\prime} \approx 2.688$ as defined in (1), (6), (20),

$$
\begin{array}{r}
\mu_{2}=2 c_{2}^{\prime}-\frac{\left(c_{2}^{\prime}\right)^{2}}{4}+\int_{\lambda=\lambda_{2}^{\prime}}^{\infty}\left(2-\frac{\lambda e^{\lambda}}{2 f_{2}(\lambda)}+\frac{\lambda f_{2}(\lambda)}{2 e^{\lambda}}-2 \frac{f_{3}(\lambda)}{e^{\lambda}}\right)\left(\frac{e^{\lambda}}{f_{2}(\lambda)}+\frac{\lambda e^{\lambda}}{f_{2}(\lambda)}-\frac{\lambda e^{\lambda} f_{1}(\lambda)}{f_{2}(\lambda)^{2}}\right) d \lambda \\
=4.17042881 \ldots
\end{array}
$$

There appears to be no clear connetion between $\mu_{2}$ and the $\zeta$ function.
Before proceeding to the proofs of Theorems 1 and 2 we note some properties of the $\kappa$-core of a random graph.

## 2 The $\kappa$-core

The functions

$$
\begin{equation*}
f_{i}(\lambda)=\sum_{j=i}^{\infty} \frac{\lambda^{j}}{j!}, \quad i=0,1,2, \ldots, \tag{1}
\end{equation*}
$$

figure prominently in our calculations. For $\lambda>0$ define

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

Properties of these functions are derived in Appendix B.
The $\kappa$-core $C_{\kappa}(G)$ of a graph $G$ is the largest set of vertices that induces a graph $H_{\kappa}$ such that the minimum degree $\delta\left(H_{\kappa}\right) \geq \kappa$. 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 $\kappa$-core and that if $c>c_{\kappa}$ then
w.h.p. $G_{n, p}$ has a $\kappa$-core of linear size. We list some facts about these cores that we will need in what follows.

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

$$
\pi_{r}(\lambda)=\operatorname{Pr}\{\operatorname{Po}(\lambda) \geq r\}=e^{-\lambda} f_{r}(\lambda) .
$$

Then

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

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

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

Then w.v.h.p. ${ }^{1}$ with $\lambda=\lambda_{\kappa}(c)$ we have that

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

Furthermore, when $\kappa$ is large,

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

Luczak [23] proved that $C_{\kappa}$ is $\kappa$-connected w.v.h.p. when $\kappa \geq 3$.
Next let $c_{\kappa}^{\prime}$ be the threshold for the $(\kappa+1)$-core having average degree $2 \kappa$. Here, see (3) and (4),

$$
\begin{equation*}
c_{\kappa}^{\prime}=\frac{\lambda e^{\lambda}}{f_{\kappa}(\lambda)} \text { where } \frac{\lambda f_{k}(\lambda)}{f_{k+1}(\lambda)}=2 \kappa \tag{6}
\end{equation*}
$$

We have $c_{2} \approx 3.35$ and $c_{2}^{\prime} \approx 3.59$.

## 3 Proof of Theorem 1: Large $k$.

We will prove Theorem 1 in this section. It is relatively straightforward. Theorem 2 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 $n \gg k$,

$$
\begin{equation*}
\operatorname{mst}_{k}\left(K_{n}\right) \geq \mathbf{E}\left[Z_{k}\right]=\sum_{\ell=1}^{k(n-1)} \frac{\ell}{\binom{n}{2}+1}=\frac{k(n-1)(k(n-1)+1)}{n(n-1)+2} \in\left[k^{2}\left(1-n^{-1}\right), k^{2}\right] . \tag{7}
\end{equation*}
$$

This gives us the lower bound in Theorem 1.
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 $\bar{E}_{H}$ of $H$ is at most $k_{0}^{2}$, see (7).

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

A theorem of Nash-Williams [25] states that a $2 k$-edge connected graph contains $k$ edge-disjoint spanning trees. Applying the result of Łuczak [23] we see that w.v.h.p. $C_{2 k}$ 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 $\left|C_{2 k}\right| \sim n \operatorname{Pr}\{\operatorname{Po}(\lambda) \geq 2 k\}$ w.v.h.p., we see that w.v.h.p. $D_{2 k}=[n] \backslash C_{2 k}$ satisfies $\left|D_{2 k}\right| \leq 2 n e^{-c_{1} k^{1 / 3}}$.

For each $v \in D_{2 k}$ we let $S_{v}$ be the $k$ shortest edges from $v$ to $C_{2 k}$. 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 $Y_{v}, v \in D_{2 k}$ ? We can bound this probabilistically by using the following lemma from Frieze and Grimmett [13]:

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 $k_{i}$ th minimum of $N$ independent uniform [0,1] random variables. If $\mu>1$ then

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

Let $\varepsilon=2 e^{-c_{1} k^{1 / 3}}$ and $\mu=10 \ln 1 / \varepsilon$ and let $M=k \varepsilon n, N=(1-\varepsilon) n, a=k^{2} \varepsilon n$. Let $\mathcal{B}_{0}$ be the event that there exists a set $S$ of size $\varepsilon n$ such that the sum over $v \in S$ of the lengths of the $k$ shortest edges from $v$ to $[n] \backslash S$ exceeds $\mu a /(N+1)$. Next let $\mathcal{B}$ be the event that the sum over $v \in S$ of the length of the $k$ th shortest edge from $v$ to $[n] \backslash S$ exceeds $\mu a /(k(N+1))$. We have $\mathcal{B}_{0} \subseteq \mathcal{B}$ and applying Lemma 1 we see that

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

It follows that

$$
\operatorname{mst}_{k}\left(K_{n}\right) \leq o(1)+k_{0}^{2}+\frac{\mu a}{N+1} \leq k^{2}+3 k^{5 / 3} .
$$

The $o(1)$ term is a bound $k n \times o\left(n^{-1}\right)$, to account for the cases that occur with probability $o\left(n^{-1}\right)$. Combining this with (7) we see that

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

which proves Theorem 1.

## 4 Proof of Theorem 2: $k=2$.

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 $0 \leq p \leq 1, \sum_{e \in F} 1_{\left(X_{e}>p\right)}$ is the number of edges of $F$ which are not in $G_{p}$, which equals $2 n-2-r_{2}\left(G_{p}\right)$. So,

$$
\operatorname{mst}_{2}\left(K_{n}, \mathbf{X}\right)=\sum_{e \in F} X_{e}=\sum_{e \in F} \int_{p=0}^{1} 1_{\left(X_{e}>p\right)} d p=\int_{p=0}^{1} \sum_{e \in F} 1_{\left(X_{e}>p\right)} d p .
$$

Hence, on taking expectations we obtain

$$
\begin{equation*}
\operatorname{mst}_{2}\left(K_{n}\right)=\int_{p=0}^{1}\left(2 n-2-\mathbf{E}\left[r_{2}\left(G_{p}\right)\right]\right) d p \tag{8}
\end{equation*}
$$

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

Lemma 2. Let $k \geq 2$. 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

$$
\begin{equation*}
r_{k}(G)=\left|E_{k}(G)\right|+r_{k}\left(C_{k+1}\right) . \tag{9}
\end{equation*}
$$

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 $F_{1}, \ldots, F_{k}$ are edge disjoint forests such that $r_{k}(G)=$ $\left|F_{1}\right|+\ldots+\left|F_{k}\right|$ then by removing $v$ we see, inductively, that $\left|E_{k}(G-v)\right|+r_{k}\left(C_{k+1}\right)=r_{k}(G-v) \geq$ $\left|F_{1}\right|+\ldots+\left|F_{k}\right|-d_{G}(v)=r_{k}(G)-d_{G}(v)$. On the other hand $G-v$ contains $k$ forests $F_{1}^{\prime}, \ldots, F_{k}^{\prime}$ such that $r_{k}(G-v)=\left|F_{1}^{\prime}\right|+\ldots+\left|F_{k}^{\prime}\right|=\left|E_{k}(G-v)\right|+r_{k}\left(C_{k+1}\right)$. We can then add $v$ as a vertex of degree one to $d_{G}(v)$ of the forests $F_{1}^{\prime}, \ldots, F_{k}^{\prime}$, implying that $r_{k}(G) \geq d_{G}(v)+\left|E_{k}(G-v)\right|+r_{k}\left(C_{k+1}\right)$. Thus, $r_{k}(G)=d_{G}(v)+\left|E_{k}(G-v)\right|+r_{k}\left(C_{k+1}\right)=\left|E_{k}(G)\right|+r_{k}\left(C_{k+1}\right)$.

Lemma 3. Let $k \geq 2$. If $c_{k}<c<c_{k}^{\prime}$, then w.h.p.

$$
\begin{equation*}
\left|E\left(G_{n, c / n}\right)\right|-o(n) \leq r_{k}\left(G_{n, c / n}\right) \leq\left|E\left(G_{n, c / n}\right)\right| . \tag{10}
\end{equation*}
$$

Proof. We will show that when $c<c_{k}^{\prime}$ we can find $k$ disjoint forests $F_{1}, F_{2}, \ldots, F_{k}$ contained in $C_{k+1}$ such that

$$
\begin{equation*}
\left|E\left(C_{k+1}\right)\right|-\sum_{i=1}^{k}\left|E\left(F_{i}\right)\right|=o(n) . \tag{11}
\end{equation*}
$$

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

Gao, Pérez-Giménez and Sato [16] show that when $c<c_{k}^{\prime}$, no subgraph of $G_{n, p}$ has average degree more than $2 k$, 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 $2 k-\varepsilon$, 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 $\varepsilon$.
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 $H_{i}=\left([n], \Phi_{i}\right)$ 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

$$
\mathbf{E}[Z] \leq \sum_{\ell=3}^{\omega_{0}} \ell!\binom{n}{\ell} \ell p^{\ell} \leq \omega_{0} c^{\omega_{0}} \leq n^{1 / 2}
$$

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 $k n / \omega_{0}=o(n)$ and this proves (11) and the lemma.

Lemma 4. If $c>c_{2}^{\prime}$, then w.h.p. the 3 -core of $G_{n, c / n}$ contains two edge-disjoint forests of total size $2\left|V\left(C_{3}\right)\right|-o(n)$. In particular, $r_{2}\left(C_{3}\left(G_{n, c / n}\right)\right)=2\left|V\left(C_{3}\right)\right|-o(n)$.

The proof of Lemma 4 is postponed to Section 6. We can now prove Theorem 2.

## 5 Proof of Theorem 2.

As noted in (8),

$$
\begin{equation*}
\operatorname{mst}_{2}\left(K_{n}\right)=\int_{p=0}^{1}\left(2 n-2-\mathbf{E}\left[r_{2}\left(G_{p}\right)\right]\right) d p \tag{12}
\end{equation*}
$$

A crude calculation shows that if $c$ is large then

$$
\begin{equation*}
p \geq \frac{c}{n} \text { implies that } \operatorname{Pr}\left\{r_{2}\left(G_{p}\right)<2 n-2-n A c^{6} e^{-c}\right\}=o(1) \tag{13}
\end{equation*}
$$

for some absolute constant $A>0$.
Indeed, we know that if $p=\frac{c}{n}$ and $c$ is suficently large, then $G_{p}$ contains a pair of edge disjoint cycles, each of length at least $n\left(1-c^{6} e^{-c}\right)$ with probability $1-\varepsilon_{1}$, where $\varepsilon_{1}=O\left(n^{-\alpha}\right)$, for some absolute constant $\alpha>0$, see Frieze [11]. If $p_{1}=\frac{c_{1}}{n}$ and $p_{2}=K p_{1}$ then $\operatorname{Pr}\left\{r_{2}\left(G_{p_{2}}\right)<2 n-2-n c^{6} e^{-c}\right\} \leq$ $\varepsilon_{1}^{p_{2} / p_{1}}=O\left(n^{-K \alpha}\right)$ since $G_{p_{2}}$ can be generated by adding edges to $p_{2} / p_{1}$ independent copies of $G_{p_{1}}$. This confirms (13).

So, for large $c$,

$$
\begin{equation*}
\operatorname{mst}_{2}\left(K_{n}\right)=\int_{p=0}^{\frac{c}{n}}\left(2 n-2-\mathbf{E}\left[r_{2}\left(G_{p}\right)\right]\right) d p+\varepsilon_{c} \tag{14}
\end{equation*}
$$

where

$$
0 \leq \varepsilon_{c} \leq A n \int_{p=\frac{c}{n}}^{1}(n p)^{6} e^{-n p} d p=A \int_{x=c}^{n} x^{6} e^{-x} d x \leq A \int_{x=c}^{\infty} x^{6} e^{-x} d x<c^{7} e^{-c}
$$

after changing variables to $x=p n$. Doing this once more we have,

$$
\begin{equation*}
\operatorname{mst}_{2}\left(K_{n}\right)=\int_{x=0}^{c}\left(2-2 n^{-1}-n^{-1} \mathbf{E}\left[r_{2}\left(G_{\frac{x}{n}}\right)\right]\right) d x+\varepsilon_{c} \tag{15}
\end{equation*}
$$

By Lemmas 2 and 3, for $x<c_{2}^{\prime}$ we have $n^{-1} \mathbf{E}\left[r_{2}\left(G_{\frac{x}{n}}\right)\right]=n^{-1} \mathbf{E}\left[\left|E\left(G_{\frac{x}{n}}\right)\right|\right]-\xi(x, n)=x / 2-\xi(x, n)$ where $\lim _{n \rightarrow \infty} \xi(x, n)=0$. Now $n^{-1} \mathbf{E}\left[r_{2}\left(G_{\frac{x}{n}}\right)\right], n=1,2, \ldots$, is a sequence of bounded monotone increasing continuous functions of $x$. This sequence converges pointwise to a continuous function $f$, and it actually converges uniformly to $f$. Thus we can bound $\max _{0 \leq x \leq c_{2}^{\prime}} \xi(x, n) \leq \eta(n)$ where $\lim _{n \rightarrow \infty} \eta(n)=0$. Clearly $f(x)=x / 2$ and so

$$
\int_{x=0}^{c_{2}^{\prime}} n^{-1} \mathbf{E}\left[r_{2}\left(G_{\frac{x}{n}}\right)\right] d x=\int_{x=0}^{c_{2}^{\prime}} \frac{x}{2} d x+o(1)
$$

By Lemma 4, for $x>c_{2}^{\prime}$ we have $\mathbf{E}\left[r_{2}\left(C_{3}\left(G_{\frac{x}{n}}\right)\right)\right]=\mathbf{E}\left[2\left|V\left(C_{3}\right)\right|\right]-o(n)$. So by Lemma 2 $\mathbf{E}\left[r_{2}\left(G_{\frac{x}{n}}\right)\right]=\mathbf{E}\left[\left|E\left(G_{\frac{x}{n}}\right)\right|-\left|E\left(C_{3}\right)\right|+2\left|V\left(C_{3}\right)\right|\right]-o(n)$, and
$\mu_{2}=\int_{x=0}^{c_{2}^{\prime}}\left(2-\frac{x}{2}\right) d x+\int_{x=c_{2}^{\prime}}^{c}\left(2-\frac{1}{n}\left(\frac{x n}{2}-\mathbf{E}\left[\left\lvert\, E\left(C_{3}\left(G_{\frac{x}{n}}\right)\right)\right.\right]+\mathbf{E}\left[2\left|V\left(C_{3}\left(G_{\frac{x}{n}}\right)\right)\right|\right]\right)\right) d x+\varepsilon_{c}+o(1)$

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

$$
\begin{aligned}
& \frac{1}{n} \mathbf{E}\left[\left|V\left(C_{3}\right)\right|\right]=\frac{f_{3}(\lambda)}{e^{\lambda}}+o(1) \\
& \frac{1}{n} \mathbf{E}\left[\left|E\left(C_{3}\right)\right|\right]=\frac{\lambda f_{2}(\lambda)}{2 e^{\lambda}}+o(1)
\end{aligned}
$$

where $\lambda$ is the largest solution to $\lambda e^{\lambda} / f_{2}(\lambda)=x$. Thus,

$$
\begin{equation*}
\mu_{2}=\lim _{n \rightarrow \infty} \operatorname{mst}_{2}\left(K_{n}\right)=\int_{x=0}^{c_{2}^{\prime}}\left(2-\frac{x}{2}\right) d x+\int_{x=c_{2}^{\prime}}^{c}\left(2-\frac{x}{2}+\frac{\lambda f_{2}(\lambda)}{2 e^{\lambda}}-2 \frac{f_{3}(\lambda)}{e^{\lambda}}\right) d x+\varepsilon_{c} . \tag{17}
\end{equation*}
$$

To calculate this, note that

$$
\begin{equation*}
\frac{d x}{d \lambda}=\frac{e^{\lambda}}{f_{2}(\lambda)}+\frac{\lambda e^{\lambda}}{f_{2}(\lambda)}-\frac{\lambda e^{\lambda} f_{1}(\lambda)}{f_{2}(\lambda)^{2}} \tag{18}
\end{equation*}
$$

so

$$
\begin{align*}
& \int_{x=c_{2}^{\prime}}^{c}\left(2-\frac{x}{2}+\frac{\lambda f_{2}(\lambda)}{2 e^{\lambda}}-2 \frac{f_{3}(\lambda)}{e^{\lambda}}\right) d x \\
= & \int_{\lambda\left(c_{2}^{\prime}\right)}^{\lambda(c)}\left(2-\frac{\lambda e^{\lambda}}{2 f_{2}(\lambda)}+\frac{\lambda f_{2}(\lambda)}{2 e^{\lambda}}-2 \frac{f_{3}(\lambda)}{e^{\lambda}}\right)\left(\frac{e^{\lambda}}{f_{2}(\lambda)}+\frac{\lambda e^{\lambda}}{f_{2}(\lambda)}-\frac{\lambda e^{\lambda} f_{1}(\lambda)}{f_{2}(\lambda)^{2}}\right) d \lambda+\varepsilon_{c} \tag{19}
\end{align*}
$$

where $\lambda(x)$ is the unique solution to $\lambda e^{\lambda} / f_{2}(\lambda)=x$.
Note that

$$
\begin{equation*}
\lambda\left(c_{2}^{\prime}\right) \approx 2.688 \text { and } \lambda(c)>\frac{c}{2} \text { for large } c . \tag{20}
\end{equation*}
$$

Now for large $\lambda$ we can bound

$$
\left(2-\frac{\lambda e^{\lambda}}{2 f_{2}(\lambda)}+\frac{\lambda f_{2}(\lambda)}{2 e^{\lambda}}-2 \frac{f_{3}(\lambda)}{e^{\lambda}}\right)\left(\frac{e^{\lambda}}{f_{2}(\lambda)}+\frac{\lambda e^{\lambda}}{f_{2}(\lambda)}-\frac{\lambda e^{\lambda} f_{1}(\lambda)}{f_{2}(\lambda)^{2}}\right)
$$

from above by $\lambda^{3} e^{-\lambda}$. So the range in the integral in (19) can be extended to $\infty$ at the cost of adding an amount $\delta_{c}$ where $0 \leq \delta_{c} \leq c^{4} e^{-c}$. Using the fact that we can make $\varepsilon_{c}, \delta_{c}$ arbitrarily close to zero by making $c$ abritrarily large, we obtain the expression for $\mu_{2}$ claimed in Theorem 2.

Attempts to transform the integral in the theorem into an explicit integral with explicit bounds have been unsuccesful. Numerical calculations give

$$
\begin{equation*}
\mu_{2} \approx 4.1704288 \ldots \tag{21}
\end{equation*}
$$

The Inverse Symbolic Calculator ${ }^{2}$ has yielded no symbolic representation of this number. An apparent connection to the $\zeta$ function lies in its representation as

$$
\begin{equation*}
\zeta(x)=\frac{1}{\Gamma(x)} \int_{\lambda=0}^{\infty} \frac{\lambda^{x-1}}{e^{\lambda}-1} d \lambda \tag{22}
\end{equation*}
$$

which is somewhat similar to terms of the form

$$
\begin{equation*}
\int_{\lambda=\lambda_{2}^{\prime}}^{\infty} \frac{\operatorname{poly}(\lambda)}{e^{\lambda}-1-\lambda} d \lambda \tag{23}
\end{equation*}
$$

appearing in $\mu_{2}$, but no real connection has been found.

## 6 Proof of Lemma 4.

### 6.1 More on the 3-core.

Suppose now that $c>c_{2}^{\prime}$ 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 $V\left(C_{3}\right)=[N]$.

The degree sequence $d(v), v \in[N]$ can be generated as follows: We independently choose for each $v \in V\left(C_{3}\right)$ a truncated Poisson random variable with parameter $\lambda$ satisfying $g_{0}(\lambda)=2 M / N$, conditioned on $d(v) \geq 3$. So for $v \in[N]$,

$$
\begin{equation*}
\operatorname{Pr}\{d(v)=k\}=\frac{\lambda^{k}}{k!f_{3}(\lambda)}, \quad k=3,4,5, \ldots, \quad \lambda=g_{0}^{-1}\left(\frac{2 M}{N}\right) \tag{24}
\end{equation*}
$$

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.

[^2]These independent variables are further conditioned so that the event

$$
\begin{equation*}
\mathcal{D}=\left\{\sum_{v \in[N]} d(v)=2 M\right\} \tag{25}
\end{equation*}
$$

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

$$
\begin{equation*}
\operatorname{Pr}\{\mathcal{E} \mid \mathcal{D}\} \leq O\left(n^{1 / 2}\right) \operatorname{Pr}\{\mathcal{E}\} \tag{26}
\end{equation*}
$$

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 $\left|W_{v}\right|=d(v)$, 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 $\Gamma_{F}$.

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 $\Gamma_{F}$ is simple will be bounded away from zero. This is because the degree sequence in (26) has exponential tails. Thus we only need to show that $\Gamma_{F}$ has certain properties w.h.p.

### 6.2 Setting up the main calculation.

Suppose now that $p=c / n$ where $c>c_{2}^{\prime}$. We will show that w.h.p., for any fixed $\varepsilon>0$,

$$
\begin{equation*}
i(S)=\left|\left\{e \in E\left(C_{3}\right): e \cap S \neq \emptyset\right\}\right| \geq(2-\varepsilon)|S| \text { for all } S \subseteq[N] . \tag{27}
\end{equation*}
$$

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 (27) 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 (27).

It follows from (27) that
$E\left(C_{3}\right)$ can be oriented so that at least $(1-\varepsilon) N$ vertices have indegree at least two.
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 $e \in E\left(C_{3}\right)$. The directed edges are (i) $(s, v), v \in[N]$ of capacity two; (ii) ( $u, e$ ), where $u \in e$ of infinite capacity; (iii) $(e, t), e \in E\left(C_{3}\right)$ 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 $E\left(C_{3}\right)$. The condition (27) 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 (28) 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 $\mathbf{E}\left[r_{2}\left(C_{3}\left(G_{n, c / n}\right)\right)\right] \geq 2(1-\varepsilon) N-o(N)$ and since $N=\Omega(n)$, we can have $\mathbf{E}\left[r_{2}\left(C_{3}\left(G_{n, c / n}\right)\right)\right]=2(1-\varepsilon) N-o(n)$. Because $\varepsilon$ is arbitrary, this implies $r_{2}\left(C_{3}\left(G_{n, c / n}\right)\right)=2 N-o(n)$ whenever $c>c_{2}^{\prime}$.

### 6.3 Proof of (27): Small $S$.

It will be fairly easy to show that (28) holds w.h.p. for all $|S| \leq s_{\varepsilon}$ where

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

We claim that w.h.p.

$$
\begin{equation*}
|S| \leq s_{\varepsilon} \text { implies } e(S)<(1+\varepsilon)|S| \text { in } G_{n, p} \tag{29}
\end{equation*}
$$

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

$$
\begin{aligned}
& \operatorname{Pr}\{\exists S \text { violating }(29)\} \leq \sum_{s=4}^{s_{\varepsilon}}\binom{n}{s}\binom{\binom{s}{2}}{(1+\varepsilon) s} p^{(1+\varepsilon) s} \leq \\
& \sum_{s=4}^{s_{\varepsilon}}\left(\frac{n e}{s}\right)^{s}\left(\frac{s e c}{2(1+\varepsilon) n}\right)^{(1+\varepsilon) s}=\sum_{s=4}^{s_{\varepsilon}}\left(\left(\frac{s}{n}\right)^{\varepsilon} \frac{e^{2+\varepsilon} c}{2(1+\varepsilon)}\right)^{s}=o(1)
\end{aligned}
$$

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 (29) holds and that $|S| \leq s_{\varepsilon}$ and $i(S) \leq(2-\varepsilon)|S|$. Then if $\bar{S}=[N] \backslash S$, we have

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

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

### 6.4 Proof of (27): 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 $T=[N] \backslash S$. For any $v \in S$, we have $i(S \backslash v) \geq(2-\varepsilon)|S \backslash v|$ while $i(S)<(2-\varepsilon)|S|$. This implies $d_{T}(v)=i(S)-i(S \backslash v)<2$.

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 $S^{\prime}=S \cup\{w\}$. We have $i\left(S^{\prime}\right) \leq i(S)+2<2\left|S^{\prime}\right|$. 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, i.e. $\left|S_{i}\right|=s_{i}$ and $|T|=t$.
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 $\mathbf{w}=\left(s_{0}, s_{1}, s_{2}, D_{0}, D_{1}, D_{2}, D_{3}, t, M\right)$. Here $D_{i}=d_{S}\left(S_{i}\right), i=0,1,2$ and $D_{3}=d_{T}(T)$.

### 6.4.1 Estimating the probability of w.

By construction, $D_{i} \geq(3-i) s_{i}$ for $i=0,1,2$ and $D_{3} \geq 3 t$. Define degree sequences $\left(d_{i}^{1}, \ldots, d_{i}^{s_{i}}\right)$ for $S_{i}, i=0,1,2$ and $\left(d_{3}^{1}, \ldots, d_{3}^{t}\right)$ 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 (25)), the probability that $S_{0}$ has degree sequence $\left(d_{0}^{1}, \ldots, d_{0}^{s_{0}}\right), d_{0}^{i} \geq 3$ for all $i$, is given by

$$
\begin{equation*}
\prod_{i=1}^{s_{0}} \frac{\lambda_{0}^{d_{0}^{i}}}{d_{0}^{i}!f_{3}(\lambda)} \tag{30}
\end{equation*}
$$

where $\lambda$ is the solution to

$$
g_{0}(\lambda)=\frac{2 M}{N} .
$$

Hence, letting $\left[x^{D}\right] f(x)$ denote the coefficient of $x^{D}$ in the power series $f(x)$, the probability $\pi_{0}\left(S_{0}, D_{0}\right)$ that $d\left(S_{0}\right)=D_{0}$ is bounded by

$$
\begin{align*}
\pi_{0}\left(S_{0}, D_{0}\right) \leq \sum_{\substack{d_{0}^{1}+\cdots+s_{0}^{s_{0}}=D_{0} \\
d_{0}^{d_{0} \geq 3}}} \prod_{i=1}^{s_{0}} \frac{\lambda^{d_{0}^{i}}}{d_{0}^{i}!f_{3}(\lambda)} & =\frac{\lambda^{D_{0}}}{f_{3}(\lambda)^{s_{0}}} \sum_{\substack{d_{0}^{1}+\cdots+d_{0}^{s_{0}}=D_{0} \\
d_{0}^{i} \geq 3}} \prod_{i=1}^{s_{0}} \frac{1}{d_{0}^{i}!} \\
& =\frac{\lambda^{D_{0}}}{f_{3}(\lambda)^{s_{0}}}\left[x^{D_{0}}\right]\left(\sum_{d_{0} \geq 3} \frac{x^{d_{0}}}{d_{0}!}\right)^{s_{0}} \\
& =\frac{\lambda^{D_{0}}}{f_{3}(\lambda)^{s_{0}}}\left[x^{D_{0}}\right] f_{3}(x)^{s_{0}} \\
& \leq \frac{\lambda^{D_{0}}}{f_{3}(\lambda)^{s_{0}}} \frac{f_{3}\left(\lambda_{0}\right)^{s_{0}}}{\lambda_{0}^{D_{0}}} \tag{31}
\end{align*}
$$

for all $\lambda_{0}$. Here we use the fact that for any function $f$ and any $y>0,\left[x^{D_{0}}\right] f(x) \leq f(y) / y^{D_{0}}$. To minimise (31) we choose $\lambda_{0}$ to be the unique solution to

$$
\begin{equation*}
g_{0}\left(\lambda_{0}\right)=\frac{D_{0}}{s_{0}} . \tag{32}
\end{equation*}
$$

If $D_{0}=3 s_{0}$ then $\lambda_{0}=0$ by Lemma 6, Appendix B. In this case, since $f_{3}\left(\lambda_{0}\right)=\frac{\lambda_{0}^{3}\left(1+O\left(\lambda_{0}\right)\right)}{6}$, we have

$$
\begin{equation*}
\pi_{0}\left(S_{0}, D_{0}\right) \leq\left(\frac{\lambda^{3}}{6 f_{3}(\lambda)}\right)^{s_{0}}, \quad \text { when } D_{0}=3 s_{0} \tag{33}
\end{equation*}
$$

Dealing with $S_{1}$ :
For each $v \in S_{1}$, we have $W_{v}=W_{v}^{S} \cup W_{v}^{T}$ where $\left|W_{v}^{T}\right|=1$. Hence, the probability $\pi_{1}\left(S_{1}, D_{1}\right)$ that $d\left(S_{1}\right)=D_{1}+s_{1}$ is bounded by

$$
\begin{align*}
\pi_{1}\left(S_{1}, D_{1}\right) \leq \sum_{\substack{\hat{d}_{1}^{1}+\cdots+\widehat{d}_{1}^{s_{1}}=D_{1} \\
\widehat{d_{1}^{2}} \geq 2}} \prod_{i=1}^{s_{1}}\binom{\widehat{d_{1}^{2}}+1}{1} \frac{\lambda^{\widehat{d_{1}^{2}}+1}}{\left(\widehat{d_{1}^{i}}+1\right)!f_{3}(\lambda)} & =\frac{\lambda^{D_{1}+s_{1}}}{f_{3}(\lambda)^{s_{1}}} \sum_{\hat{d}_{1}^{1}+\cdots+\widehat{d}_{1}^{s_{1}}=D_{1}} \prod_{i=1}^{\widehat{d}_{1} \geq 2} \frac{1}{s_{1}^{\hat{d}_{1}^{u}}!} \\
& =\frac{\lambda^{D_{1}+s_{1}}}{f_{3}(\lambda)^{s_{1}}}\left[x^{D_{1}}\right] f_{2}(x)^{s_{1}} \\
& \leq \frac{\lambda^{D_{1}+s_{1}}}{f_{3}(\lambda)^{s_{1}}} \frac{f_{2}\left(\lambda_{1}\right)^{s_{1}}}{\lambda_{1}^{D_{1}}} . \tag{34}
\end{align*}
$$

We choose $\lambda_{1}$ to satisfy the equation

$$
\begin{equation*}
g_{1}\left(\lambda_{1}\right)=\frac{D_{1}}{s_{1}} . \tag{35}
\end{equation*}
$$

Similarly to what happens in (33) we have $\lambda_{1}=0$ when $D_{1}=2 s_{1}$ and $f_{2}\left(\lambda_{1}\right)=\frac{\lambda_{1}^{2}\left(1+O\left(\lambda_{1}\right)\right)}{2}$, so

$$
\begin{equation*}
\pi_{1}\left(S_{1}, D_{1}\right) \leq\left(\frac{\lambda^{3}}{2 f_{3}(\lambda)}\right)^{s_{1}}, \quad \text { when } D_{1}=2 s_{1} \tag{36}
\end{equation*}
$$

## 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 $\pi_{2}\left(S_{2}, D_{2}\right)$ that $d\left(S_{2}\right)=$ $D_{2}+2 s_{2}$ is bounded by

$$
\begin{equation*}
\pi_{2}\left(S_{2}, D_{2}\right) \leq \sum_{\substack{\hat{d}_{2}^{1}+\cdots+\widehat{d}_{2}^{s_{2}}=D_{2} \\ \widehat{d_{2}^{2}} \geq 1}} \prod_{i=1}^{s_{2}}\binom{\widehat{d_{2}^{l}}+2}{2} \frac{\lambda^{\widehat{d_{2}}+2}}{\left(\widehat{d_{2}^{l}}+2\right)!f_{3}(\lambda)} \leq \frac{\lambda^{D_{2}+2 s_{2}}}{f_{3}(\lambda)^{s_{2}}} \frac{f_{1}\left(\lambda_{2}\right)^{s_{2}}}{\lambda_{2}^{D_{2}}} 2^{-s_{2}} \tag{37}
\end{equation*}
$$

where we choose $\lambda_{2}$ to satisfy the equation

$$
\begin{equation*}
g_{2}\left(\lambda_{2}\right)=\frac{D_{2}}{s_{2}} . \tag{38}
\end{equation*}
$$

Similarly to what happens in (33) we have $\lambda_{2}=0$ when $D_{2}=s_{2}$ and $f_{1}\left(\lambda_{2}\right)=\lambda_{2}\left(1+O\left(\lambda_{2}\right)\right)$, so

$$
\begin{equation*}
\pi_{2}\left(S_{2}, D_{2}\right) \leq\left(\frac{\lambda^{3}}{2 f_{3}(\lambda)}\right)^{s_{2}}, \quad \text { when } D_{2}=s_{2} . \tag{39}
\end{equation*}
$$

## Dealing with $T$ :

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

$$
\sum_{i=1}^{t} \widehat{d}_{3}^{\widehat{i}}=d_{S}(T)=s_{1}+2 s_{2}
$$

by the definition of $S_{0}, S_{1}, S_{2}$. So the probability $\pi_{3}\left(T, D_{3}\right)$ that $d_{T}(T)=D_{3}$, given $s_{1}, s_{2}$ can be bounded by

$$
\begin{align*}
& \pi_{3}\left(T, D_{3}\right) \leq \sum_{\substack{\widehat{d}_{3}+\cdots+\widehat{d}_{3}^{d}=s_{1}+2 s_{2} \\
\widehat{d}_{3}^{\geq} \geq 0}} \sum_{\substack{\bar{d}_{3}^{1}+\cdots+\bar{d}_{3}^{t}=D_{3} \\
\bar{d}_{3}^{i} \geq 3}} \prod_{i=1}^{t}\binom{\widehat{d_{3}^{l}}+\bar{d}_{3}^{i}}{\widehat{d}_{3}^{i}} \frac{\lambda^{\widehat{d}_{3}}+\bar{d}_{3}^{i}}{\left(\widehat{d}_{3}^{l}+\bar{d}_{3}^{i}\right)!f_{3}(\lambda)} \\
& =\frac{\lambda^{D_{3}+s_{1}+2 s_{2}}}{f_{3}(\lambda)^{t}} \sum_{\substack{\widehat{d}_{3}^{1}+\cdots+\widehat{d}_{3}^{t}=s_{1}+2 s_{2} \\
\widehat{d}_{3}^{3} \geq 0}} \sum_{\substack{\bar{d}_{3}^{1}+\cdots+\bar{d}_{3}^{t}=D_{3} \\
\bar{d}_{3}^{i} \geq 3}} \prod_{i=1}^{t} \frac{1}{\widehat{d}_{3}^{i} \backslash \bar{d}{ }_{3}^{i}!} \\
& =\frac{\lambda^{D_{3}+s_{1}+2 s_{2}}}{f_{3}(\lambda)^{t}}\left(\left[x^{D_{3}}\right] f_{3}(x)^{t}\right)\left(\left[x^{s_{1}+2 s_{2}}\right] e^{x}\right) \\
& \leq \frac{\lambda^{D_{3}+s_{1}+2 s_{2}}}{f_{3}(\lambda)^{t}} \frac{f_{3}\left(\lambda_{3}\right)^{t}}{\lambda_{3}^{D_{3}}} \frac{t^{s_{1}+2 s_{2}}}{\left(s_{1}+2 s_{2}\right)!}, \tag{40}
\end{align*}
$$

where we choose $\lambda_{3}$ to satisfy the equation

$$
\begin{equation*}
g_{0}\left(\lambda_{3}\right)=\frac{D_{3}}{t} . \tag{41}
\end{equation*}
$$

Similarly to what happens in (33) we have $\lambda_{3}=0$ when $D_{3}=3 t$ and $f_{3}\left(\lambda_{3}\right)=\frac{\lambda_{3}^{3}\left(1+O\left(\lambda_{1}\right)\right)}{6}$, so

$$
\pi_{3}\left(T, D_{3}\right) \leq \frac{\lambda^{D_{3}+s_{1}+2 s_{2}}}{\left(6 f_{3}(\lambda)\right)^{t}} \frac{t^{s_{1}+2 s_{2}}}{\left(s_{1}+2 s_{2}\right)!}, \quad \text { when } D_{3}=3 t
$$

### 6.4.2 Putting the bounds together.

For a fixed w $=\left(s_{0}, s_{1}, s_{2}, D_{0}, D_{1}, D_{2}, D_{3}, t, M\right)$, 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 $W_{v}, v \in S \cup T$ into $W_{v}^{S} \cup W_{v}^{T}$. Note that our expressions (31), (34), (37), (40) account for these choices. Given the partitions of the $W_{v}$ 's, there are ( $D_{0}+D_{1}+$ $\left.D_{2}\right)!!D_{3}!!\left(s_{1}+2 s_{2}\right)$ ! configurations, where $(2 s)!!=(2 s-1) \times(2 s-3) \times \cdots \times 3 \times 1$ is the number of ways of partitioning a set of size $2 s$ into $s$ pairs. Here $\left(D_{0}+D_{1}+D_{2}\right)!$ ! 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 $\left(s_{1}+2 s_{2}\right)$ ! is the number of ways of pairing points associated with $S$ to points associated with $T$. Each configuration has probability $1 /(2 M)!!$. So, the total probability of all configurations whose vertex partition and degrees are described by w can be bounded by

$$
\begin{aligned}
& \binom{t+s}{s_{0}, s_{1}, s_{2}, t} \frac{\lambda^{D_{0}}}{f_{3}(\lambda)^{s_{0}}} \frac{f_{3}\left(\lambda_{0}\right)^{s_{0}}}{\lambda_{0}^{D_{0}}} \frac{\lambda^{D_{1}+s_{1}}}{f_{3}(\lambda)^{s_{1}}} \frac{f_{2}\left(\lambda_{1}\right)^{s_{1}}}{\lambda_{1}^{D_{1}}} \frac{\lambda^{D_{2}+2 s_{2}}}{f_{3}(\lambda)^{s_{2}}} \frac{f_{1}\left(\lambda_{2}\right)^{s_{2}}}{\lambda_{2}^{D_{2}}} 2^{-s_{2}} \\
& \times \frac{\lambda^{D_{3}+s_{1}+2 s_{2}}}{f_{3}(\lambda)^{t}} \frac{f_{3}\left(\lambda_{3}\right)^{t}}{\lambda_{3}^{D_{3}}} \frac{t^{s_{1}+2 s_{2}}}{\left(s_{1}+2 s_{2}\right)!} \frac{\left(D_{0}+D_{1}+D_{2}\right)!!D_{3}!!\left(s_{1}+2 s_{2}\right)!}{(2 M)!!} \\
= & \binom{t+s}{s_{0}, s_{1}, s_{2}, t} \frac{\lambda^{2 M}}{f_{3}(\lambda)^{N}} \frac{f_{3}\left(\lambda_{0}\right)^{s_{0}}}{\lambda_{0}^{D_{0}}} \frac{f_{2}\left(\lambda_{1}\right)^{s_{1}}}{\lambda_{1}^{D_{1}}} \frac{f_{1}\left(\lambda_{2}\right)^{s_{2}}}{\lambda_{2}^{D_{2}}} 2^{-s_{2}} \frac{f_{3}\left(\lambda_{3}\right)^{t}}{\lambda_{3}^{D_{3}}} \frac{t^{s_{1}+2 s_{2}}}{\left(s_{1}+2 s_{2}\right)!} \\
& \times \frac{\left(D_{0}+D_{1}+D_{2}\right)!D_{3}!!\left(s_{1}+2 s_{2}\right)!}{(2 M)!!}
\end{aligned}
$$

Write $D_{i}=\Delta_{i} s,\left|S_{i}\right|=\sigma_{i} s, t=\tau s, M=\mu s$ and $N=\nu s$. We have $k!!\sim \sqrt{2}(k / e)^{k / 2}$ as $k \rightarrow \infty$ by Stirling's formula, so the expression above, modulo an $e^{o(s)}$ factor, can be written as

$$
\begin{align*}
& f(\mathbf{w})^{s}=\left(\frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}} \sigma_{1}^{\sigma_{1}}\left(1-\sigma_{0}-\sigma_{1}\right)^{1-\sigma_{0}-\sigma_{1}} \tau^{\tau}} \frac{\lambda^{2 \mu}}{f_{3}(\lambda)^{\nu}} \frac{f_{3}\left(\lambda_{0}\right)^{\sigma_{0}}}{\lambda_{0}^{\Delta_{0}}} \frac{f_{2}\left(\lambda_{1}\right)^{\sigma_{1}}}{\lambda_{1}^{\Delta_{1}}} \frac{f_{1}\left(\lambda_{2}\right)^{\sigma_{2}}}{\lambda_{2}^{\Delta_{2}}} \frac{f_{3}\left(\lambda_{3}\right)^{\tau}}{\lambda_{3}^{\Delta_{3}}} \frac{(\tau e)^{\sigma_{1}+2 \sigma_{2}}}{2^{\sigma_{2}}}\right. \\
&\left.\frac{\left(\Delta_{0}+\Delta_{1}+\Delta_{2}\right)^{\left(\Delta_{0}+\Delta_{1}+\Delta_{2}\right) / 2} \Delta_{3}^{\Delta_{3} / 2}}{(2 \mu)^{\mu}}\right)^{s} \tag{42}
\end{align*}
$$

We note that

$$
\begin{align*}
\sigma_{2} & =1-\sigma_{0}-\sigma_{1},  \tag{43}\\
\Delta_{3} & =2 \mu-\Delta_{0}-\Delta_{1}-\Delta_{2}-2 \sigma_{1}-4 \sigma_{2} \\
& =2 \mu-4-\Delta_{0}-\Delta_{1}-\Delta_{2}+4 \sigma_{0}+2 \sigma_{1}  \tag{44}\\
\nu & =1+\tau . \tag{45}
\end{align*}
$$

Hence $\sigma_{2}, \Delta_{3}, \nu$ may be eliminated, and we can consider w to be ( $\sigma_{0}, \sigma_{1}, \Delta_{0}, \Delta_{1}, \Delta_{2}, \tau, \mu$ ). When convenient, $\Delta_{3}$ 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 w satisfying

$$
\begin{aligned}
& \Delta_{0} \geq 3 \sigma_{0}, \Delta_{1} \geq 2 \sigma_{1}, \Delta_{2} \geq 1-\sigma_{0}-\sigma_{1}, \Delta_{3} \geq 3 \tau . \\
& \frac{\Delta_{0}+\Delta_{1}+\Delta_{2}}{2}+\sigma_{1}+2\left(1-\sigma_{0}-\sigma_{1}\right)<2-\varepsilon \quad \text { since } i(S)<(2-\varepsilon)|S|, \quad \text { see }(27) . \\
& \sigma_{0}, \sigma_{1} \geq 0, \sigma_{0}+\sigma_{1} \leq 1 . \\
& 0 \leq \tau \leq(1-\varepsilon) / \varepsilon \text { since }|S| \geq \varepsilon N . \\
& \mu \geq(2+\varepsilon)(1+\tau) \text { since } M \geq(2+\varepsilon) N . \\
& \sigma_{0}<1, \quad \text { otherwise } C_{3} \text { is not connected. }
\end{aligned}
$$

Here $\varepsilon$ is a sufficiently small positive constant such that we can (i) exclude the case of small $S$, (ii) satisfy condition (27) and (iii) have $M \geq(2+\varepsilon) N$ since $c>c_{2}^{\prime}$.

For a given $s$, there are $O(\operatorname{poly}(s))$ 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

$$
\begin{equation*}
\sum_{s \geq \varepsilon N} \sum_{\mathbf{w}} O(\operatorname{poly}(s)) f(\mathbf{w})^{s} \leq \sum_{s}\left(e^{o(1)} \max _{F} f(\mathbf{w})\right)^{s} \leq N\left(e^{o(1)} \max _{F} f(\mathbf{w})\right)^{\varepsilon N} \tag{46}
\end{equation*}
$$

As $N \rightarrow \infty$, 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 (25). Inequality (26) implies that it is sufficient to inflate the RHS of (46) by $O\left(n^{1 / 2}\right)$ to obtain our result.

So, let

$$
\begin{aligned}
& \quad f\left(\Delta_{0}, \Delta_{1}, \Delta_{2}, \sigma_{0}, \sigma_{1}, \tau, \mu\right)= \\
& \quad \frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}} \sigma_{1}^{\sigma_{1}}\left(1-\sigma_{0}-\sigma_{1}\right)^{1-\sigma_{0}-\sigma_{1}} \tau^{\tau}} \frac{\lambda^{2 \mu}}{f_{3}(\lambda)^{\tau+1}} \frac{f_{3}\left(\lambda_{0}\right)^{\sigma_{0}}}{\lambda_{0}^{\Delta_{0}}} \frac{f_{2}\left(\lambda_{1}\right)^{\sigma_{1}}}{\lambda_{1}^{\Delta_{1}}} \frac{f_{1}\left(\lambda_{2}\right)^{1-\sigma_{0}-\sigma_{1}}}{\lambda_{2}^{\Delta_{2}}} \frac{f_{3}\left(\lambda_{3}\right)^{\tau}}{\lambda_{3}^{\Delta_{3}}} \\
& \times \\
& \frac{(e \tau)^{2-2 \sigma_{0}-\sigma_{1}}}{2^{1-\sigma_{0}-\sigma_{1}}} \frac{\left(\Delta_{0}+\Delta_{1}+\Delta_{2}\right)^{\left(\Delta_{0}+\Delta_{1}+\Delta_{2}\right) / 2} \Delta_{3}^{\Delta_{3} / 2}}{(2 \mu)^{\mu}}
\end{aligned}
$$

We complete the proof of Theorem 2 by showing that

$$
\begin{equation*}
f(\mathbf{w}) \leq \exp \left\{-\frac{\varepsilon^{2}}{3}\right\} \text { for all } \mathbf{w} \in F \tag{47}
\end{equation*}
$$

The proof of (47) is a very long careful calculation and we have placed it in Section A of the appendix.

## 7 Final Remarks

There are a number of loose ends to be taken care of. Is Conjecture 1 true? Is there a simpler expression for $\mu_{2}$ of Theorem 2? Is it possible to get an exact expression for $\mu_{3}$ ? On another tack, what are the expected running times of algorithms for computing these edge disjoint trees? They are polynomial time solvable problems, in the worst-case, but maybe their average complexity is significantly better than worst-case.

## 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 large matchings and cycles in sparse random graphs, Discrete Mathematics 59 (1986) 243-256.
[12] A.M. Frieze, On random symmetric travelling salesman problems, Mathematics of Operations Research 29, 878-890.
[13] A.M. Frieze and G.R. Grimmett, The shortest path problem for graphs with random arclengths, Discrete Applied Mathematics 10 (1985) 57-77.
[14] A.M. Frieze and C.J.H. McDiarmid, On random minimum length spanning trees, Combinatorica 9 (1989) 363-374.
[15] A.M. Frieze, M. Ruszinko and L. Thoma, A note on random minimum length spanning trees, The electronic journal of combinatorics 7 (2000).
[16] 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)
[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.

## A Proof of (47)

We remind the reader that the aim of this section is to show that $f(\mathbf{w})<1$ for all $\mathbf{w}=\left(\sigma_{0}, \sigma_{1}, \Delta_{0}, \Delta_{1}, \Delta_{2}, \tau, \mu\right) \in F$, where

$$
\begin{gathered}
f(\mathbf{w})=\frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}} \sigma_{1}^{\sigma_{1}}\left(1-\sigma_{0}-\sigma_{1}\right)^{1-\sigma_{0}-\sigma_{1}} \tau^{\tau}} \frac{\lambda^{2 \mu}}{f_{3}(\lambda)^{\tau+1}} \frac{f_{3}\left(\lambda_{0}\right)^{\sigma_{0}}}{\lambda_{0}^{\Delta_{0}}} \frac{f_{2}\left(\lambda_{1}\right)^{\sigma_{1}}}{\lambda_{1}^{\Delta_{1}}} \frac{f_{1}\left(\lambda_{2}\right)^{1-\sigma_{0}-\sigma_{1}}}{\lambda_{2}^{\Delta_{2}}} \frac{f_{3}\left(\lambda_{3}\right)^{\tau}}{\lambda_{3}^{\Delta_{3}}} \\
\times \frac{(e \tau)^{2-2 \sigma_{0}-\sigma_{1}}}{2^{1-\sigma_{0}-\sigma_{1}}} \frac{\left(\Delta_{0}+\Delta_{1}+\Delta_{2}\right)^{\left(\Delta_{0}+\Delta_{1}+\Delta_{2}\right) / 2} \Delta_{3}^{\Delta_{3} / 2}}{(2 \mu)^{\mu}}
\end{gathered}
$$

and $F$ is the set of solutions to

$$
\begin{align*}
& \Delta_{0} \geq 3 \sigma_{0}, \Delta_{1} \geq 2 \sigma_{1}, \Delta_{2} \geq 1-\sigma_{0}-\sigma_{1}, \Delta_{3} \geq 3 \tau  \tag{48a}\\
& \frac{\Delta_{0}+\Delta_{1}+\Delta_{2}}{2}+\sigma_{1}+2\left(1-\sigma_{0}-\sigma_{1}\right)<2-\varepsilon \quad \text { since } i(S)<(2-\varepsilon)|S|, \quad \text { see }(27) .  \tag{48b}\\
& \sigma_{0}, \sigma_{1} \geq 0, \sigma_{0}+\sigma_{1} \leq 1 .  \tag{48c}\\
& 0 \leq \tau \leq(1-\varepsilon) / \varepsilon \text { since }|S| \geq \varepsilon N .  \tag{48d}\\
& \mu \geq(2+\varepsilon)(1+\tau) \text { since } M \geq(2+\varepsilon) N .  \tag{48e}\\
& \sigma_{0}<1, \quad \text { otherwise } C_{3} \text { is not connected. } \tag{48f}
\end{align*}
$$

## A.0. 1 Eliminating $\mu$

We begin by showing that it is enough to consider $\mu=(2+\varepsilon)(1+\tau)$. We collect all terms involving $\mu$, including $\Delta_{3}, \lambda$ and $\lambda_{3}$ whose values are determined in part by $\mu$. It is enough to consider the logarithm of $f$. We have

$$
\begin{aligned}
\frac{\partial \log f}{\partial \mu}=2 \log \lambda+\frac{\partial \lambda}{\partial \mu}\left(\frac{2 \mu}{\lambda}-\nu \frac{f_{2}(\lambda)}{f_{3}(\lambda)}\right)+\frac{\partial \lambda_{3}}{\partial \mu}\left(\tau \frac{f_{2}\left(\lambda_{3}\right)}{f_{3}\left(\lambda_{3}\right)}-\right. & \left.\frac{\Delta_{3}}{\lambda_{3}}\right) \\
& -2 \log \lambda_{3}+\log \Delta_{3}+1-\log 2 \mu-1
\end{aligned}
$$

by definition of $\lambda, \lambda_{3}$, we have

$$
\frac{2 \mu}{\lambda}-\nu \frac{f_{2}(\lambda)}{f_{3}(\lambda)}=0 \text { and } \frac{\Delta_{3}}{\lambda_{3}}-\tau \frac{f_{2}\left(\lambda_{3}\right)}{f_{3}\left(\lambda_{3}\right)}=0,
$$

and so

$$
\begin{equation*}
\frac{\partial \log f}{\partial \mu}=2 \log \left(\frac{\lambda}{\lambda_{3}}\right)+\log \left(\frac{\Delta_{3}}{2 \mu}\right) \tag{49}
\end{equation*}
$$

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

$$
\begin{equation*}
g_{0}\left(\lambda_{3}\right)-g_{0}(\lambda)=\frac{\Delta_{3}}{\tau}-\frac{2 \mu}{\nu}=\frac{2 \mu-2 \iota}{\tau}-\frac{2 \mu}{\tau+1}=\frac{2 \mu-2 \iota(\tau+1)}{\tau(\tau+1)} \geq \frac{4-2 \iota}{\tau} \geq 0 . \tag{50}
\end{equation*}
$$

This shows that $\log f$ is decreasing with respect to $\mu$, 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 $f \leq \exp \left\{-\varepsilon^{2} / 3\right\}$ 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 (44) and (48a)

$$
\begin{aligned}
\Delta_{3} & =2 \mu-4-\Delta_{0}-\Delta_{1}-\Delta_{2}+4 \sigma_{0}+2 \sigma_{1} \\
& \leq 2 \mu-4-3 \sigma_{0}-2 \sigma_{1}-\left(1-\sigma_{0}-\sigma_{1}\right)+4 \sigma_{0}+2 \sigma_{1} \\
& =2 \mu-5+2 \sigma_{0}+\sigma_{1} \\
& \leq 2 \mu-2
\end{aligned}
$$

and since $\tau \leq 1 / \varepsilon-1$ by (48d), $\mu \leq(2+\varepsilon)(1+\tau)$ implies $\mu \leq 2 / \varepsilon+1<3 / \varepsilon$. So,

$$
\begin{equation*}
\frac{\partial \log f}{\partial \mu} \leq 2 \log \left(\frac{\lambda}{\lambda_{3}}\right)+\log \left(\frac{2 \mu-2}{2 \mu}\right) \leq \log \left(1-\frac{\varepsilon}{3}\right) \tag{51}
\end{equation*}
$$

So, fixing $\mathbf{w}^{\prime}=\left(\sigma_{0}, \sigma_{1}, \Delta_{0}, \Delta_{1}, \Delta_{2}, \tau\right)$, let $\mu=2(1+\tau)$ and $\mu^{\prime}=(2+\varepsilon)(1+\tau)$. If $f\left(\mathbf{w}^{\prime}, \mu\right) \leq 1$, then

$$
\begin{equation*}
\log f\left(\mathbf{w}^{\prime}, \mu^{\prime}\right) \leq \log f\left(\mathbf{w}^{\prime}, \mu\right)+\varepsilon(1+\tau) \log \left(1-\frac{\varepsilon}{3}\right) \leq-\frac{\varepsilon^{2}}{3} \tag{52}
\end{equation*}
$$

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

$$
\begin{align*}
& \Delta_{0} \geq 3 \sigma_{0}, \Delta_{1} \geq 2 \sigma_{1}, \Delta_{2} \geq 1-\sigma_{0}-\sigma_{1}, \Delta_{3} \geq 3 \tau  \tag{53a}\\
& \Delta_{0}+\Delta_{1}+\Delta_{2} \leq 4 \sigma_{0}+2 \sigma_{1}  \tag{53b}\\
& \sigma_{0}, \sigma_{1} \geq 0, \sigma_{0}+\sigma_{1} \leq 1  \tag{53c}\\
& 0 \leq \tau<\infty  \tag{53~d}\\
& \mu=2(1+\tau) \tag{53e}
\end{align*}
$$

We have relaxed equation (48b) to give (53b) in order to simplify later calculations. In $F^{\prime}$, $\lambda$ 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 $3 \sigma_{0}+2 \sigma_{1}+\left(1-\sigma_{0}-\sigma_{1}\right)=$ $2 \sigma_{0}+\sigma_{1}+1$, so by (53a), (53b),

$$
\begin{equation*}
2 \sigma_{0}+\sigma_{1}+1 \leq \Delta \leq 4 \sigma_{0}+2 \sigma_{1} \tag{54}
\end{equation*}
$$

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

$$
\begin{equation*}
\Delta_{3}=2 \mu-4-\Delta_{0}-\Delta_{1}-\Delta_{2}+4 \sigma_{0}+2 \sigma_{1}=4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta \tag{55}
\end{equation*}
$$

The quantity $2 \sigma_{0}+\sigma_{1}$ will appear frequently. We note that (54) and $\sigma_{0}+\sigma_{1} \leq 1$ imply

$$
\begin{equation*}
1 \leq 2 \sigma_{0}+\sigma_{1} \leq 2 \tag{56}
\end{equation*}
$$

## A.0.2 Eliminating $\tau$

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

$$
\begin{align*}
f\left(\sigma_{0}, \sigma_{1}, \Delta_{0}, \Delta_{1}, \Delta_{2}, \tau\right)= & \frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}} \sigma_{1}^{\sigma_{1}}\left(1-\sigma_{0}-\sigma_{1}\right)^{1-\sigma_{0}-\sigma_{1} \tau^{\tau}}}\left(\frac{\lambda^{4}}{f_{3}(\lambda)}\right)^{\tau+1} \frac{f_{3}\left(\lambda_{0}\right)^{\sigma_{0}}}{\lambda_{0}^{\Delta_{0}}} \frac{f_{2}\left(\lambda_{1}\right)^{\sigma_{1}}}{\lambda_{1}^{\Delta_{1}}} \\
& \times \frac{f_{1}\left(\lambda_{2}\right)^{1-\sigma_{0}-\sigma_{1}}}{\lambda_{2}^{\Delta_{2}}} \frac{f_{3}\left(\lambda_{3}\right)^{\tau}}{\lambda_{3}^{\Delta_{3}}} \frac{(e \tau)^{2-2 \sigma_{0}-\sigma_{1}}}{2^{1-\sigma_{0}-\sigma_{1}}} \times \frac{\Delta^{\Delta / 2} \Delta_{3}^{\Delta_{3} / 2}}{(4+4 \tau)^{2+2 \tau}} . \tag{57}
\end{align*}
$$

Here $\lambda_{0}=\lambda_{0}\left(\Delta_{0}, \sigma_{0}\right), \lambda_{1}=\lambda_{1}\left(\Delta_{1}, \sigma_{1}\right), \lambda_{2}=\lambda_{2}\left(\Delta_{2}, \sigma_{0}, \sigma_{1}\right), \lambda_{3}=\lambda_{3}\left(\sigma_{0}, \sigma_{1}, \Delta_{0}, \Delta_{1}, \Delta_{2}, \tau\right)$ as defined in (32), (35), (38), (41). Since $\tau f_{2}\left(\lambda_{3}\right) / f_{3}\left(\lambda_{3}\right)-\Delta_{3} / \lambda_{3}=0$ by the definition of $\lambda_{3}$, the partial derivative of $\log f$ with respect to $\tau$ is given by

$$
\begin{aligned}
\frac{\partial}{\partial \tau} \log f\left(\sigma_{0}, \sigma_{1}, \Delta_{0}, \Delta_{1}, \Delta_{2}, \tau\right)= & \log (\tau+1)+1-\log \tau-1+\log \left(\frac{\lambda^{4}}{f_{3}(\lambda)}\right) \\
& +\frac{\partial \lambda_{3}}{\partial \tau}\left(\tau \frac{f_{2}\left(\lambda_{3}\right)}{f_{3}\left(\lambda_{3}\right)}-\frac{\Delta_{3}}{\lambda_{3}}\right)+\log \left(f_{3}\left(\lambda_{3}\right)\right)-4 \log \lambda_{3} \\
& +\frac{2-2 \sigma_{0}-\sigma_{1}}{\tau}+2\left(1+\log \Delta_{3}\right)-2 \log (4+4 \tau)-2 \\
= & \log (\tau+1)-\log \tau+\log \left(\frac{\lambda^{4}}{\lambda_{3}^{4}} \frac{f_{3}\left(\lambda_{3}\right)}{f_{3}(\lambda)}\right)+\frac{2-2 \sigma_{0}-\sigma_{1}}{\tau} \\
& +2 \log \Delta_{3}-2 \log (4+4 \tau)
\end{aligned}
$$

This is positive for $\tau$ close to zero. This is clear as long as $2 \sigma_{0}+\sigma_{1}<2$. But if $2 \sigma_{0}+\sigma_{1}=2$ then $\sigma_{0}+\sigma_{1} \leq 1$ 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 (48f). On the other hand, $\frac{\partial}{\partial \tau} \log f$ vanishes if

$$
\begin{equation*}
2-2 \sigma_{0}-\sigma_{1}-\tau\left[\log \left(1+\frac{1}{\tau}\right)-2 \log \left(\frac{\Delta_{3}}{4 \tau}\right)-\log \left(\frac{\lambda^{4} f_{3}\left(\lambda_{3}\right)}{\lambda_{3}^{4} f_{3}(\lambda)}\right)\right]=0 \tag{58}
\end{equation*}
$$

So any local maximum of $f$ must satisfy this equation. If no solution exists, then it is optimal to let $\tau \rightarrow \infty$. We will see below how to choose $\tau$ to guarantee maximality. For now, we only assume $\tau$ satisfies (58).

## A.0.3 Eliminating $\Delta_{0}, \Delta_{1}, \Delta_{2}$.

We now eliminate $\Delta_{0}, \Delta_{1}, \Delta_{2}$. Fix $\sigma_{0}, \sigma_{1}$. For $\Delta_{i}>(3-i) \sigma_{i}$ such that $\Delta_{0}+\Delta_{1}+\Delta_{2}<4 \sigma_{0}+2 \sigma_{1}$,

$$
\begin{align*}
\frac{\partial}{\partial \Delta_{i}} \log f= & \frac{\partial \lambda_{i}}{\partial \Delta_{i}}\left(\sigma_{i} \frac{f_{2-i}\left(\lambda_{i}\right)}{f_{3-i}\left(\lambda_{i}\right)}-\frac{\Delta_{i}}{\lambda_{i}}\right)-\log \lambda_{i}+\log \lambda_{3} \\
& +\frac{\partial}{\partial \tau} \log f \frac{\partial \tau}{\partial \Delta_{i}}+\frac{1}{2} \log \Delta+\frac{1}{2}-\frac{1}{2} \log \Delta_{3}-\frac{1}{2}  \tag{59}\\
= & -\log \lambda_{i}+\log \left(\lambda_{3} \sqrt{\frac{\Delta}{\Delta_{3}}}\right)
\end{align*}
$$

since $g_{i}\left(\lambda_{i}\right)=\Delta_{i} / \sigma_{i}$ by definition of $\lambda_{i}$, and the term $\frac{\partial}{\partial \tau} \log f \partial \tau / \partial \Delta_{i}$ vanishes because (58) is assumed to hold. We note that $\lambda_{i}>0$ when $\Delta_{i}>(3-i) \sigma_{i}$ (Appendix A), allowing division by $\lambda_{i}$.

As $\Delta_{i}$ tends to its lower bound $(3-i) \sigma_{i}$, we have $\log \lambda_{i} \rightarrow-\infty$ while the other terms remain bounded, so the derivative is positive at the lower bound of $\Delta_{i}$. Any stationary point must satisfy $\lambda_{0}=\lambda_{1}=\lambda_{2}=\lambda_{3} \sqrt{\Delta / \Delta_{3}}=: \widehat{\lambda}$. This can only happen if

$$
\begin{equation*}
\sigma_{0} g_{0}(\widehat{\lambda})+\sigma_{1} g_{1}(\widehat{\lambda})+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(\widehat{\lambda})=\sigma_{0} \frac{\Delta_{0}}{\sigma_{0}}+\sigma_{1} \frac{\Delta_{1}}{\sigma_{1}}+\left(1-\sigma_{0}-\sigma_{1}\right) \frac{\Delta_{2}}{1-\sigma_{0}-\sigma_{1}}=\Delta . \tag{60}
\end{equation*}
$$

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

$$
\begin{align*}
\widehat{\lambda} & =\lambda_{3} \sqrt{\frac{\Delta}{\Delta_{3}}} \\
\Delta & =\sigma_{0} g_{0}(\hat{\lambda})+\sigma_{1} g_{1}(\widehat{\lambda})+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(\widehat{\lambda})  \tag{61}\\
2-2 \sigma_{0}-\sigma_{1} & =\tau\left[\log \left(1+\frac{1}{\tau}\right)-2 \log \left(\frac{\Delta_{3}}{4 \tau}\right)-\log \left(\frac{\lambda^{4} f_{3}\left(\lambda_{3}\right)}{\lambda_{3}^{4} f_{3}(\lambda)}\right)\right]
\end{align*}
$$

In Appendix B we show that this system has no solution such that $2 \sigma_{0}+\sigma_{1}+1 \leq \Delta \leq 4 \sigma_{0}+2 \sigma_{1}$ (see (54)). 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

$$
\begin{equation*}
\Delta_{0}+\Delta_{1}+\Delta_{2}=4 \sigma_{0}+2 \sigma_{1} \text { which implies that } \Delta_{3}=4 \tau, \text { see }(55) . \tag{62}
\end{equation*}
$$

This eliminates one degree of freedom. We now set

$$
\Delta_{2}=4 \sigma_{0}+2 \sigma_{1}-\Delta_{0}-\Delta_{1} .
$$

Then for $\Delta_{0}, \Delta_{1}$, we have

$$
\begin{equation*}
\frac{\partial}{\partial \Delta_{i}} \log f=-\log \lambda_{i}+\log \lambda_{2}, \quad i=0,1 \tag{63}
\end{equation*}
$$

To see this note that (59) 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 $\lambda_{0}=\lambda_{1}=\lambda_{2}=\bar{\lambda}$, defined by

$$
\begin{equation*}
\sigma_{0} g_{0}(\bar{\lambda})+\sigma_{1} g_{1}(\bar{\lambda})+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(\bar{\lambda})=4 \sigma_{0}+2 \sigma_{1} \tag{64}
\end{equation*}
$$

This has a unique solution $\bar{\lambda} \geq 0$ whenever $2 \sigma_{0}+\sigma_{1} \geq 1$, since for fixed $\sigma_{0}, \sigma_{1}$, the left-hand side is a convex combination of increasing functions, by Lemma 7, Appendix C. This defines $\Delta_{i}=\Delta_{i}\left(\sigma_{0}, \sigma_{1}\right)$ by

$$
\begin{equation*}
\Delta_{0}=g_{0}(\bar{\lambda}) \sigma_{0}, \quad \Delta_{1}=g_{1}(\bar{\lambda}) \sigma_{1}, \quad \Delta_{2}=g_{2}(\bar{\lambda})\left(1-\sigma_{0}-\sigma_{1}\right) \tag{65}
\end{equation*}
$$

We note at this point that $\bar{\lambda} \leq \lambda$. Indeed, by (62) and (48a),

$$
\Delta_{0}=4 \sigma_{0}+2 \sigma_{1}-\Delta_{1}-\Delta_{2} \leq 4 \sigma_{0}+2 \sigma_{1}-2 \sigma_{1}-\left(1-\sigma_{0}-\sigma_{1}\right) \leq 4 \sigma_{0}
$$

so

$$
\begin{equation*}
g_{0}(\bar{\lambda})=\frac{\Delta_{0}}{\sigma_{0}} \leq 4=g_{0}(\lambda) \tag{66}
\end{equation*}
$$

implying that $\bar{\lambda} \leq \lambda$, since $g_{0}$ is increasing.
This choice (65) 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 (62), and so

$$
\begin{equation*}
\lambda_{3}=g_{0}^{-1}\left(\frac{4 \tau}{\tau}\right)=\lambda \tag{67}
\end{equation*}
$$

is fixed. In particular, the relation (58) for $\tau$ simplifies to

$$
\begin{equation*}
2-2 \sigma_{0}-\sigma_{1}=\tau \log \left(1+\frac{1}{\tau}\right) \tag{68}
\end{equation*}
$$

Let $\phi(\tau)=\tau \log (1+1 / \tau)$. Then $\phi^{\prime \prime}(\tau)=-\tau^{-1}(\tau+1)^{-2}$, so $\phi$ is concave and then $\lim _{\tau \rightarrow 0} \phi(\tau)=0$, $\lim _{\tau \rightarrow \infty} \phi(\tau)=1$ implies that $\phi$ is strictly increasing and takes values in $[0,1)$ for $\tau \geq 0$. This means that (68) 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 $\tau$, and we treat this case now.

If $2 \sigma_{0}+\sigma_{1}=1$, then (54) implies that $\Delta=2$. Furthermore, $\Delta_{3}=4 \tau$ (see (55)) and $\lambda_{3}=\lambda$ (see (67)) and $g_{i}(0)=3-i$ implies that

$$
\sigma_{0} g_{0}(0)+\sigma_{1} g_{1}(0)+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(0)=2 \sigma_{0}+\sigma_{1}+1=4 \sigma_{0}+2 \sigma_{1},
$$

so $\bar{\lambda}=0$ is the unique solution to (64). Then since $\Delta_{i} / \sigma_{i}=g_{i}(0)=3-i$ (Lemma 6, Appendix C), we have $\Delta_{i}=(3-i) \sigma_{i}, i=0,1,2$, and as in (33), (36), (39),

$$
\begin{equation*}
\frac{f_{3}(\bar{\lambda})^{\sigma_{0}} f_{2}(\bar{\lambda})^{\sigma_{1}} f_{1}(\bar{\lambda})^{1-\sigma_{0}-\sigma_{1}}}{\bar{\lambda}^{\Delta}}=\left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{3}}\right)^{\sigma_{0}}\left(\frac{f_{2}(\bar{\lambda})}{\bar{\lambda}^{2}}\right)^{\sigma_{1}}\left(\frac{f_{1}(\bar{\lambda})}{\bar{\lambda}}\right)^{1-\sigma_{0}-\sigma_{1}}=\frac{1}{6^{\sigma_{0}}} \frac{1}{2^{\sigma_{1}}} \tag{69}
\end{equation*}
$$

so when $2 \sigma_{0}+\sigma_{1}=1$, (57) becomes

$$
\begin{equation*}
f\left(\sigma_{0}, \sigma_{1}, \tau\right)=\frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}} \sigma_{1}^{\sigma_{1}}\left(1-\sigma_{0}-\sigma_{1}\right)^{1-\sigma_{0}-\sigma_{1}} \tau^{\tau}} \frac{\lambda^{4}}{f_{3}(\lambda)} \frac{1}{6^{\sigma_{0}}} \frac{1}{2^{\sigma_{1}}} \frac{e \tau}{2^{1-\sigma_{0}-\sigma_{1}}} \frac{2^{2 / 2}(4 \tau)^{2 \tau}}{(4+4 \tau)^{2+2 \tau}} . \tag{70}
\end{equation*}
$$

In this computation we also used the fact that $\lambda=\lambda_{3}$ (see (67)) and $\Delta_{3}=4 \tau$ (see (55)) to find that

$$
\left(\frac{\lambda^{4}}{f_{3}(\lambda)}\right)^{\tau+1} \frac{f_{3}\left(\lambda_{3}\right)^{\tau}}{\lambda_{3}^{\Delta_{3}}}=\frac{\lambda^{4}}{f_{3}(\lambda)} .
$$

Here $\lambda^{4} / f_{3}(\lambda) \approx 7.05$ is fixed. As noted in the discussion after (68), the partial derivative in $\tau$ is positive for all $\tau$, so we let $\tau \rightarrow \infty$. Substituting $\sigma_{1}=1-2 \sigma_{0}$ we are reduced to

$$
\begin{aligned}
f\left(\sigma_{0}\right) & =\lim _{\tau \rightarrow \infty} \frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}}\left(1-2 \sigma_{0}\right)^{\left(1-2 \sigma_{0}\right)} \sigma_{0}^{\sigma_{0}} \tau^{\tau}} \frac{\lambda^{4}}{f_{3}(\lambda)} \frac{1}{6^{\sigma_{0}}} \frac{1}{2^{1-2 \sigma_{0}}} \frac{e \tau}{2^{\sigma_{0}}} \frac{2(4 \tau)^{2 \tau}}{(4+4 \tau)^{2+2 \tau}} \\
& =\frac{\lambda^{4}}{16 f_{3}(\lambda)} \frac{1}{\sigma_{0}^{2 \sigma_{0}}\left(1-2 \sigma_{0}\right)^{1-2 \sigma_{0} 3^{\sigma_{0}}}}
\end{aligned}
$$

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 $\sigma_{0}$.

## A.0.4 Dealing with $\sigma_{0}, \sigma_{1}$

With this, we have reduced our analysis to the variables $\sigma_{0}, \sigma_{1}$ in the domain

$$
E=\left\{\left(\sigma_{0}, \sigma_{1}\right): \sigma_{0}, \sigma_{1} \geq 0, \sigma_{0}+\sigma_{1} \leq 1,2 \sigma_{0}+\sigma_{1} \geq 1\right\}
$$

We just showed that $f \leq 1$ in

$$
E_{0}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 2 \sigma_{0}+\sigma_{1}=1\right\} .
$$

Further define

$$
\begin{aligned}
& E_{1}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 0.01 \leq \sigma_{1} \leq 0.99\right\}, \\
& E_{2}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 0 \leq \sigma_{1}<0.01\right\}, \\
& E_{3}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 0.99<\sigma_{1} \leq 1\right\} .
\end{aligned}
$$

We will show that $f \leq 1$ 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 mentioned above, $\Delta=4 \sigma_{0}+2 \sigma_{1}$ simplifies $f$. Specifically, if $2 \sigma_{0}+\sigma_{1}>1$ then (57) becomes, after using (62) and (67),

$$
\begin{align*}
f\left(\sigma_{0}, \sigma_{1}\right)= & \frac{(\tau+1)^{\tau+1}}{\sigma_{0}^{\sigma_{0}} \sigma_{1}^{\sigma_{1}}\left(1-\sigma_{0}-\sigma_{1}\right)^{1-\sigma_{0}-\sigma_{1}} \tau^{\tau}} \frac{\lambda^{4}}{f_{3}(\lambda)} \frac{f_{3}(\bar{\lambda})^{\sigma_{0}} f_{2}(\bar{\lambda})^{\sigma_{1}} f_{1}(\bar{\lambda})^{1-\sigma_{0}-\sigma_{1}}}{\bar{\lambda}^{4 \sigma_{0}+2 \sigma_{1}}} \\
& \times \frac{(e \tau)^{2-2 \sigma_{0}-\sigma_{1}}}{2^{1-\sigma_{0}-\sigma_{1}}} \frac{\left(4 \sigma_{0}+2 \sigma_{1}\right)^{2 \sigma_{0}+\sigma_{1}}(4 \tau)^{2 \tau}}{(4+4 \tau)^{2+2 \tau}} \tag{71}
\end{align*}
$$

In (68), (64) respectively, $\tau$ and $\bar{\lambda}$ are given as functions of $\sigma_{0}, \sigma_{1}$. Recall that $\lambda=g_{0}^{-1}(4)$ is constant. So

$$
\begin{align*}
& \partial_{0} \log f\left(\sigma_{0}, \sigma_{1}\right)= \\
& -\log \sigma_{0}-1+\log \left(1-\sigma_{0}-\sigma_{1}\right)+1+\log f_{3}(\bar{\lambda})-\log f_{1}(\bar{\lambda}) \\
& -4 \log \bar{\lambda}-2 \log (e \tau)+\log 2+2 \log \left(4 \sigma_{0}+2 \sigma_{1}\right)+2 \\
& +\frac{\partial \bar{\lambda}}{\partial \sigma_{0}}\left(\sigma_{0} \frac{f_{2}(\bar{\lambda})}{f_{3}(\bar{\lambda})}+\sigma_{1} \frac{f_{1}(\bar{\lambda})}{f_{2}(\bar{\lambda})}+\left(1-\sigma_{0}-\sigma_{1}\right) \frac{f_{0}(\bar{\lambda})}{f_{1}(\bar{\lambda})}-\frac{4 \sigma_{0}+2 \sigma_{1}}{\bar{\lambda}}\right) \\
& +\frac{\partial \tau}{\partial \sigma_{0}}\left(\log (\tau+1)+1-\log \tau-1+\frac{2-2 \sigma_{0}-\sigma_{1}}{\tau}+2 \log 4 \tau+2-2 \log (4+4 \tau)-2\right) \\
& =\log \left(\frac{1-\sigma_{0}-\sigma_{1}}{\sigma_{0}}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{4} f_{1}(\bar{\lambda})}\right)-2 \log \tau+\log 2+2 \log \left(4 \sigma_{0}+2 \sigma_{1}\right) \tag{72}
\end{align*}
$$

where, as expected, the terms involving $\partial_{0} \tau$ and $\partial_{0} \bar{\lambda}$ vanish since $\tau, \bar{\lambda}$ were chosen to maximize $\log f$. (See (68) and (64) respectively).

Similarly,

$$
\begin{align*}
& \partial_{1} \log f\left(\sigma_{0}, \sigma_{1}\right)= \\
& -\log \sigma_{1}-1+\log \left(1-\sigma_{0}-\sigma_{1}\right)+1+\log f_{2}(\bar{\lambda})-\log f_{1}(\bar{\lambda}) \\
& -2 \log \bar{\lambda}-\log (e \tau)+\log 2+\log \left(4 \sigma_{0}+2 \sigma_{1}\right)+1 \\
& +\frac{\partial \bar{\lambda}}{\partial \sigma_{1}}\left(\sigma_{0} \frac{f_{2}(\bar{\lambda})}{f_{3}(\bar{\lambda})}+\sigma_{1} \frac{f_{1}(\bar{\lambda})}{f_{2}(\bar{\lambda})}+\left(1-\sigma_{0}-\sigma_{1}\right) \frac{f_{0}(\bar{\lambda})}{f_{1}(\bar{\lambda})}-\frac{4 \sigma_{0}+2 \sigma_{1}}{\bar{\lambda}}\right) \\
& +\frac{\partial \tau}{\partial \sigma_{1}}\left(\log (\tau+1)+1-\log \tau-1+\frac{2-2 \sigma_{0}-\sigma_{1}}{\tau}+2 \log 4 \tau+2-2 \log (4+4 \tau)-2\right) \\
& =\log \left(\frac{1-\sigma_{0}-\sigma_{1}}{\sigma_{1}}\right)+\log \left(\frac{f_{2}(\bar{\lambda})}{\bar{\lambda}^{2} f_{1}(\bar{\lambda})}\right)-\log \tau+\log 2+\log \left(4 \sigma_{0}+2 \sigma_{1}\right) . \tag{73}
\end{align*}
$$

Any stationary point must satisfy

$$
\begin{equation*}
\left(\partial_{0}-2 \partial_{1}\right) \log f=\log \left(\frac{\sigma_{1}^{2}}{\sigma_{0}\left(1-\sigma_{0}-\sigma_{1}\right)}\right)+\log \left(\frac{f_{1}(\bar{\lambda}) f_{3}(\bar{\lambda})}{f_{2}(\bar{\lambda})^{2}}\right)-\log 2=0 \tag{74}
\end{equation*}
$$

It is shown in Lemma 8, Appendix $C$ that

$$
1 \leq \frac{f_{2}(\bar{\lambda})^{2}}{f_{1}(\bar{\lambda}) f_{3}(\bar{\lambda})} \leq 2
$$

This means from (74) that if $\left(\partial_{0}-2 \partial_{1}\right) \log f=0$ then

$$
2 \leq \frac{\sigma_{1}^{2}}{\sigma_{0}\left(1-\sigma_{0}-\sigma_{1}\right)} \leq 4 .
$$

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

$$
\left(\partial_{0}-2 \partial_{1}\right) \log f=0 \Longrightarrow\left\{\begin{array}{l}
\sigma_{0} \geq\left(1-\sigma_{1}\right) / 2+\sqrt{1-2 \sigma_{1}-\sigma_{1}^{2}} / 2  \tag{75}\\
\sigma_{1}<1 / 2
\end{array}\right.
$$

Case One. $E_{1}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 0.01 \leq \sigma_{1} \leq 0.99\right\}$
We need a lower bound for $\bar{\lambda} \tau$. We first note that $g_{i}(\bar{\lambda}) \leq 3-i+\bar{\lambda}$ (Lemma 6, Appendix C) implies

$$
\begin{equation*}
4 \sigma_{0}+2 \sigma_{1}=\sigma_{0} g_{0}(\bar{\lambda})+\sigma_{1} g_{1}(\bar{\lambda})+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(\bar{\lambda}) \leq 2 \sigma_{0}+\sigma_{1}+1+\bar{\lambda} \tag{76}
\end{equation*}
$$

so

$$
\bar{\lambda} \geq 2 \sigma_{0}+\sigma_{1}-1=1-\tau \log (1+1 / \tau)
$$

Here we have used (68).
For $\tau$, 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 $\tau$ by the discussion after (68). This implies

$$
\begin{equation*}
\tau>10^{-3} \tag{77}
\end{equation*}
$$

since $0.001 \log (1001)<0.01$.
If $\tau \leq 1.1$,

$$
\bar{\lambda} \geq 1-1.1 \log \left(1+\frac{1}{1.1}\right)>0.1
$$

So, if $\tau \leq 1.1$,

$$
\bar{\lambda} \tau \geq 10^{-4}
$$

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

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

So, in $E_{1}$, we have

$$
\begin{equation*}
\bar{\lambda}_{\tau} \geq 10^{-4} \tag{78}
\end{equation*}
$$

By definition of $E_{1}, \sigma_{0} \geq 0.01$ and $\sigma_{1} \geq 0.01$. By ( 66 ), $0 \leq \bar{\lambda} \leq \lambda$. This implies $f_{3}(\bar{\lambda}) / \lambda^{2} f_{1}(\bar{\lambda}) \leq 1 / 6$ and $f_{2}(\bar{\lambda}) / \bar{\lambda} f_{1}(\bar{\lambda}) \leq 1 / 3$ (Lemma 8, Appendix C). So after rewriting (72) slightly,

$$
\begin{align*}
\partial_{0} \log f\left(\sigma_{0}, \sigma_{1}\right) & =\log \left(\frac{1-\sigma_{0}-\sigma_{1}}{\sigma_{0}}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{2} f_{1}(\bar{\lambda})}\right)-2 \log \bar{\lambda} \tau+\log 2+2 \log \left(4 \sigma_{0}+2 \sigma_{1}\right) \\
& \leq \log \frac{1}{0.01}+\log \frac{1}{6}-2 \log 10^{-4}+\log 2+2 \log 4  \tag{79}\\
& \leq 25
\end{align*}
$$

Similarly, (73) is bounded by

$$
\partial_{1} \log f\left(\sigma_{0}, \sigma_{1}\right) \leq \log \frac{1}{0.01}+\log \frac{1}{3}-\log 10^{-4}+\log 2+\log 4 \leq 15 .
$$

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 $\sigma_{0}, \sigma_{1}$, given by $\sigma_{0} \geq\left(1-\sigma_{1}\right) / 2$ and $\sigma_{1} \geq 0.01$. Note that if $\sigma_{0}=\left(1-\sigma_{1}\right) / 2$, then $\left(\sigma_{0}, \sigma_{1}\right) \in E_{0}$ and it was shown above that $\log f\left(\sigma_{0}, \sigma_{1}\right) \leq \log 0.96 \leq-0.01$. Define a finite grid $P \subseteq E_{1}$ such that for any $\left(\sigma_{0}, \sigma_{1}\right) \in E_{1}$, there exists $\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right) \in P \cup E_{0}$ where $0 \leq \sigma_{0}-\bar{\sigma}_{0} \leq \delta$ and $0 \leq \sigma_{1}-\bar{\sigma}_{1} \leq \delta$. Here $\delta=1 / 4000$. Numerical calculations will show that $\log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right) \leq-0.01$ for all $\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right) \in P$. This implies that for all $\sigma_{0}, \sigma_{1} \in E_{1}$,

$$
\log f\left(\sigma_{0}, \sigma_{1}\right) \leq \max _{\bar{\sigma}_{0}, \bar{\sigma}_{1} \in P \cup E_{0}} \log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right)+25 \delta+15 \delta \leq-0.01+40 \delta \leq 0 .
$$

When calculating $\log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right)$, approximations $\bar{\lambda}_{\text {num }}, \tau_{n u m}$ of $\bar{\lambda}\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right), \tau\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right)$ must be calculated with sufficient precision. By definition of $\bar{\lambda}, \partial \log f / \partial \bar{\lambda}=0$, while

$$
\begin{aligned}
& \begin{aligned}
& \begin{aligned}
\partial^{2} \log f \\
\partial \bar{\lambda}^{2}
\end{aligned} \\
&=\left\lvert\, \sigma_{0}\left(\frac{f_{1}(\bar{\lambda})}{f_{3}(\bar{\lambda})}-\right.\right.\left.\frac{f_{2}(\bar{\lambda})^{2}}{f_{3}(\bar{\lambda})^{2}}\right)+\sigma_{1}\left(\frac{f_{0}(\bar{\lambda})}{f_{2}(\bar{\lambda})}-\frac{f_{1}(\bar{\lambda})^{2}}{f_{2}(\bar{\lambda})^{2}}\right) \\
& \left.\quad+\left(1-\sigma_{0}-\sigma_{1}\right)\left(\frac{f_{0}(\bar{\lambda})}{f_{1}(\bar{\lambda})}-\frac{f_{0}(\bar{\lambda})^{2}}{f_{1}(\bar{\lambda})^{2}}\right)+\frac{4 \sigma_{0}+2 \sigma_{1}}{\bar{\lambda}^{2}} \right\rvert\,
\end{aligned} \\
& =\frac{1}{\bar{\lambda}^{2}} \left\lvert\, \sigma_{0}\left(\bar{\lambda}^{2} \frac{f_{1}(\bar{\lambda})}{f_{3}(\bar{\lambda})}-\frac{\bar{\lambda}^{2} f_{2}(\bar{\lambda})^{2}}{f_{3}(\bar{\lambda})^{2}}\right)+\sigma_{1}\left(\bar{\lambda}^{2} \frac{f_{0}(\bar{\lambda})}{f_{2}(\bar{\lambda})}-\frac{\bar{\lambda}^{2} f_{1}(\bar{\lambda})^{2}}{f_{2}(\bar{\lambda})^{2}}\right)\right.
\end{aligned} \quad \begin{aligned}
& \left.+\left(1-\sigma_{0}-\sigma_{1}\right)\left(\bar{\lambda}^{2} \frac{f_{0}(\bar{\lambda})}{f_{1}(\bar{\lambda})}-\frac{\bar{\lambda}^{2} f_{0}(\bar{\lambda})^{2}}{f_{1}(\bar{\lambda})^{2}}\right)+\sigma_{0} g_{0}(\bar{\lambda})+\sigma_{1} g_{1}(\bar{\lambda})+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(\bar{\lambda}) \right\rvert\, \\
& \left.=\frac{1}{\bar{\lambda}^{2}} \right\rvert\, \sigma_{0} g_{0}(\bar{\lambda})\left(g_{1}(\bar{\lambda})-g_{0}(\bar{\lambda})+1\right)+\sigma_{1} g_{1}(\bar{\lambda})\left(g_{2}(\bar{\lambda})-g_{1}(\bar{\lambda})+1\right) \\
& \left.\left.\leq \frac{9}{\bar{\lambda}^{2}} \right\rvert\, \sigma_{0} g_{0}(\bar{\lambda})+\sigma_{1}\right) g_{2}(\bar{\lambda})\left(\bar{\lambda}-g_{2}(\bar{\lambda})+1\right) \mid \\
& =\frac{9}{\bar{\lambda}^{2}}\left|4 \sigma_{0}+2 \sigma-1\right|, \quad \quad \text { by }(64) \\
& \leq \frac{36}{\bar{\lambda}^{2}} .
\end{aligned}
$$

Here we use the fact that $g_{i}(\bar{\lambda}) \leq 4$ for $0 \leq \bar{\lambda} \leq \lambda, i=0,1,2$ to conclude that $\left|g_{1}-g_{0}+1\right|, \mid g_{2}-$ $g_{1}+1\left|,\left|\bar{\lambda}-g_{2}+1\right| \leq 9\right.$, and the final step uses $4 \sigma_{0}+2 \sigma_{1} \leq 4$. So the error contributed by $\bar{\lambda}_{\text {num }}$ is

$$
\begin{equation*}
\left|\log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1} ; \bar{\lambda}_{n u m}\right)-\log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1} ; \bar{\lambda}\right)\right| \leq\left(\bar{\lambda}_{n u m}-\bar{\lambda}\right)^{2} \frac{36}{\bar{\lambda}^{2}} \tag{80}
\end{equation*}
$$

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

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

$$
\left|\frac{\partial^{2} \log f}{\partial \tau^{2}}\right|=\left|\frac{1}{\tau(\tau+1)}\right| \leq 10^{3}, \text { by (77). }
$$

Thus to achieve a numerical error of at most $10^{-4}$, it suffices to have $\left|\tau_{\text {num }} / \tau-1\right| \leq 10^{-2}$.
With the above precision, it is found that over all $\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right) \in P \cup E_{0}, \log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right) \leq-0.0105$ numerically. With an error tolerance of $10^{-4}$, this shows that $\log f\left(\bar{\sigma}_{0}, \bar{\sigma}_{1}\right) \leq-0.01$.

Case Two. $E_{2}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 0 \leq \sigma_{1}<0.01\right\}$
We divide $E_{2}$ into three subregions,

$$
\begin{aligned}
& E_{2,1}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E_{2}: \sigma_{1}=0\right\}, \\
& E_{2,2}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E_{2}: \sigma_{0}+\sigma_{1}=1\right\}, \\
& E_{2,3}=E_{2} \backslash\left(E_{2,1} \cup E_{2,2}\right) .
\end{aligned}
$$

We begin by considering the point $\left(\sigma_{0}, \sigma_{1}\right)=(1,0)$. Here $4 \sigma_{0}+2 \sigma_{1}=4$, and from (64) $\bar{\lambda}$ is defined by $g_{0}(\bar{\lambda})=4$. So $\bar{\lambda}=g_{0}^{-1}(4)=\lambda$. We also have $2-2 \sigma_{0}-\sigma_{1}=0$, and from the definition (68) of $\tau$ we have $\tau=0$. Plugging this into the definition of $f(71)$ 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

$$
\partial_{0} \log f\left(\sigma_{0}, 0\right)=\log \left(\frac{1-\sigma_{0}}{\sigma_{0}}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{2} f_{1}(\bar{\lambda})}\right)-2 \log \bar{\lambda} \tau+\log 2+2 \log \left(4 \sigma_{0}\right)
$$

Within $E_{2,1}$, we consider two cases. First suppose $\sigma_{0} \leq 0.99$. As noted in (78), $\sigma_{0} \leq 0.99$ implies $\bar{\lambda} \tau \geq 10^{-4}$. Applying the same bounds as in (79),

$$
\begin{equation*}
\partial_{0} \log f\left(\sigma_{0}, 0\right) \leq \log \frac{1}{6}-2 \log 10^{-4}+\log 2+2 \log 4 \leq 21 \tag{81}
\end{equation*}
$$

and we show numerically that $f \leq 1$. The numerical 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 $\bar{\lambda} \leq \lambda$ (see (66)) implies $f_{3}(\bar{\lambda}) / \bar{\lambda}^{4} f_{1}(\bar{\lambda}) \geq 0.01$ by Lemma 8, Appendix C. 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 (68)), it follows from a numerical calculation that $\tau \leq 0.004$. This implies

$$
\begin{equation*}
\frac{1-\sigma_{0}}{\tau^{2}}=\frac{\log \left(1+\frac{1}{\tau}\right)}{2 \tau} \geq 125 \log 250 \tag{82}
\end{equation*}
$$

and

$$
\begin{aligned}
\partial_{0} \log f\left(\sigma_{0}, 0\right) & =\log \left(\frac{1-\sigma_{0}}{\sigma_{0}}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{4} f_{1}(\bar{\lambda})}\right)-2 \log \tau+\log 2+4 \log \left(4 \sigma_{0}\right) \\
& =\log \left(\frac{1-\sigma_{0}}{\tau^{2}}\right)-\log \sigma_{0}+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{4} f_{1}(\bar{\lambda})}\right)+\log 2+4 \log \left(4 \sigma_{0}\right) \\
& \geq \log (125 \log 250)+\log 0.01+\log 2+2 \log 3.96>0
\end{aligned}
$$

which implies $f\left(\sigma_{0}, 0\right)<f(1,0)=1$ 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

$$
\begin{equation*}
\partial_{0} \log f\left(\sigma_{0}, 1-\sigma_{0}\right)=\log \left(\frac{1-\sigma_{0}}{\sigma_{0}}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{2} f_{2}(\bar{\lambda})}\right)-\log \tau+\log \left(2+2 \sigma_{0}\right) \tag{83}
\end{equation*}
$$

By Lemma 8, Appendix C, $\bar{\lambda} \leq \lambda$ implies

$$
\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{2} f_{2}(\bar{\lambda})}>0.09
$$

As $\sigma_{1}=1-\sigma_{0}, \tau$ 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 so

$$
\begin{equation*}
\frac{1-\sigma_{0}}{\tau}=\frac{\sigma_{1}}{\tau}=\log \left(1+\frac{1}{\tau}\right)>\log 333 . \tag{84}
\end{equation*}
$$

So,

$$
\begin{aligned}
\partial_{0} \log f\left(\sigma_{0}, 1-\sigma_{0}\right) & =\log \left(\frac{1-\sigma_{0}}{\tau}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{2} f_{2}(\bar{\lambda})}\right)-\log \sigma_{0}+\log \left(2+2 \sigma_{0}\right) \\
& \geq \log \log 333+\log 0.09+\log 3.98 \\
& >0
\end{aligned}
$$

and for all $0.99 \leq \sigma_{0}<1, f\left(\sigma_{0}, 1-\sigma_{0}\right)<f(1,0)=1$.

## 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 $\left(\partial_{0}-2 \partial_{1}\right) \log f=0$. By (75) we must have $\sigma_{0} \geq\left(1-\sigma_{1}\right) / 2+\sqrt{1-2 \sigma_{1}-\sigma_{1}^{2}} / 2$. 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 $1-\sigma_{0} \geq\left(2-2 \sigma_{0}-\sigma_{1}\right) / 2$, so

$$
\begin{equation*}
\frac{\sigma_{1}}{\tau} \geq \frac{1}{1.1} \frac{1-\sigma_{0}}{\tau} \geq \frac{1}{2.2} \frac{2-2 \sigma_{0}-\sigma_{1}}{\tau}=\frac{1}{2.2} \log \left(1+\frac{1}{\tau}\right) \tag{85}
\end{equation*}
$$

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 $\left(\partial_{0}-2 \partial_{1}\right) \log f=0$ and $\sigma_{1} \leq 0.01$, then $\left(\partial_{0}-\partial_{1}\right) \log f \neq 0$. Noting that $4 \sigma_{0}+2 \sigma_{1} \geq 4\left(1-1.1 \sigma_{1}\right)+2 \sigma_{1} \geq 3.976$,

$$
\begin{aligned}
\left(\partial_{0}-\partial_{1}\right) \log f & =\log \left(\frac{\sigma_{1}}{\tau}\right)+\log \left(\frac{f_{3}(\bar{\lambda})}{\bar{\lambda}^{2} f_{2}(\bar{\lambda})}\right)-\log \sigma_{0}+\log \left(4 \sigma_{0}+2 \sigma_{1}\right) \\
& \geq \log (\log (500) / 2.2)+\log 0.09+\log 3.976 \\
& =1.038445 \ldots-2.407945 \ldots+1.380276 \ldots \\
& >0
\end{aligned}
$$

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: $E_{3}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E: 0.99<\sigma_{1} \leq 1\right\}$.
Further divide $E_{3}$ into

$$
\begin{aligned}
& E_{3,1}=\left\{\left(\sigma_{0}, \sigma_{1}\right) \in E_{3}: \sigma_{0}+\sigma_{1}=1\right\}, \\
& E_{3,2}=E_{3} \backslash E_{3,1} .
\end{aligned}
$$

## 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 (83),

$$
\begin{equation*}
\partial_{0} \log f\left(\sigma_{0}, 1-\sigma_{0}\right)=\log \left(\frac{1-\sigma_{0}}{\sigma_{0}}\right)+\log \left(\frac{1}{g_{0}(\bar{\lambda})}\right)-\log \bar{\lambda} \tau+\log \left(2+2 \sigma_{0}\right) \tag{86}
\end{equation*}
$$

To show that this is positive, we bound $\bar{\lambda} \tau$ from above. From (98) (Appendix B) with $\Delta=4 \sigma_{0}+2 \sigma_{1}$ we have $\tau \leq 1 /\left(4 \sigma_{0}+2 \sigma_{1}-2\right)$. For $\bar{\lambda}$, we use the bound derived in Appendix B (99). Note that if $\Delta=4 \sigma_{0}+2 \sigma_{1}$ then $L_{2}=\bar{\lambda}$ in (99). So,

$$
\begin{equation*}
\bar{\lambda} \leq \frac{12\left(4 \sigma_{0}+2 \sigma_{1}-2 \sigma_{0}-\sigma_{1}-1\right)}{6-3 \sigma_{0}-2 \sigma_{1}} \leq 12\left(2 \sigma_{0}+\sigma_{1}-1\right) \leq 12 . \tag{87}
\end{equation*}
$$

These two bounds together imply $\bar{\lambda}_{\tau} \leq 6$. For all $0 \leq \bar{\lambda} \leq \lambda$ we have $3 \leq g_{0}(\bar{\lambda}) \leq 4$ since $3 \leq \Delta_{0} / \sigma_{0} \leq 4$ (see the discussion before (66)).

We conclude that

$$
\begin{equation*}
\partial_{0} \log f\left(\sigma_{0}, 1-\sigma_{0}\right) \geq \log \frac{0.99}{0.01}+\log \frac{1}{4}-\log 6+\log 2>0 \tag{88}
\end{equation*}
$$

This implies that for all $\left(\sigma_{0}, \sigma_{1}\right) \in E_{3,1}, f\left(\sigma_{0}, \sigma_{1}\right) \leq f(0.01,0.99) \leq 1$, since $(0.01,0.99) \in E_{1}$.

## Sub-Case 3.2:

Now consider $E_{3,2}$. As noted in (75), 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 2.

## Appendix B

This section is concerned with showing that the system of equations (61) under certain conditions has no solution. Throughout the section, assume $\tau$ satisfies (58): Recall that $\Delta_{3}=4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta$,

$$
\begin{equation*}
\tau\left(\log \left(1+\frac{1}{\tau}\right)-2 \log \left(\frac{4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta}{4 \tau}\right)-\log \left(\frac{\lambda^{4}}{\lambda_{3}^{4}} \frac{f_{3}\left(\lambda_{3}\right)}{f_{3}(\lambda)}\right)\right)=2-2 \sigma_{0}-\sigma_{1} . \tag{89}
\end{equation*}
$$

Here $\lambda=g_{0}^{-1}(4) \approx 2.688$ is fixed, and $\lambda_{3}$ is defined by $\lambda_{3}=g_{0}^{-1}\left(\Delta_{3} / \tau\right)$.
Define for $2 \sigma_{0}+\sigma_{1}+1 \leq \Delta \leq 4 \sigma_{0}+2 \sigma_{1}$

$$
\begin{equation*}
L_{1}\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right)=\lambda_{3} \sqrt{\frac{\Delta}{4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta}} \tag{90}
\end{equation*}
$$

and define $L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)$ as the unique solution to $G\left(\sigma_{0}, \sigma_{1}, L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)\right)=\Delta$, where $G$ is defined by

$$
\begin{equation*}
G\left(\sigma_{0}, \sigma_{1}, x\right)=\sigma_{0} g_{0}(x)+\sigma_{1} g_{1}(x)+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}(x) \tag{91}
\end{equation*}
$$

This is well defined because each $g_{i}$ is strictly increasing, and for fixed $\sigma_{0}, \sigma_{1}$ we have $G\left(\sigma_{0}, \sigma_{1}, 0\right)=$ $2 \sigma_{0}+\sigma_{1}+1 \leq \Delta$ and $\lim _{x \rightarrow \infty} G\left(\sigma_{0}, \sigma_{1}, x\right)=\infty$ (see Appendix B). Define
$R=\left\{\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in \mathbb{R}_{+}^{4}: \sigma_{0}+\sigma_{1} \leq 1 ; 2 \sigma_{0}+\sigma_{1} \geq 1 ; 2 \sigma_{0}+\sigma_{1}+1 \leq \Delta \leq 4 \sigma_{0}+2 \sigma_{1} ;\right.$ (89) holds. $\}$
We prove that the system (61) is inconsistent by proving
Lemma 5. Let $\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in R$. Then $L_{1}\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right)>L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)$
Proof. Define $L\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right)=L_{1}\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right)-L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)$. We will bound $|\nabla L|$ in $R$ in order to show numerically that $L>0$. However, $\nabla L$ is unbounded for $\Delta$ close to 4 and $2 \sigma_{0}+\sigma_{1}$ close to 1. For this reason, define

$$
\begin{aligned}
& R_{1}=\left\{\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in R: \Delta \geq 3.6\right\}, \\
& R_{2}=\left\{\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in R: 2 \sigma_{0}+\sigma_{1} \leq 1.1\right\}, \\
& R_{3}=R \backslash\left(R_{1} \cup R_{2}\right)
\end{aligned}
$$

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 $\sigma_{0}, \sigma_{1}$ we have $L_{2}\left(\sigma_{0}, \sigma_{1}, 2 \sigma_{0}+\sigma_{1}+1\right)=0$, since $G\left(\sigma_{0}, \sigma_{1}, 0\right)=2 \sigma_{0}+\sigma_{1}+1$, see (91). Here we use $g_{i}(0)=3-i, i=0,1,2$. This implies that $L_{1}\left(\sigma_{0}, \sigma_{1}, 2 \sigma_{0}+\sigma_{1}+1, \tau\right)>0=$ $L_{2}\left(\sigma_{0}, \sigma_{1}, 2 \sigma_{0}+\sigma_{1}+1\right)$, and we may therefore assume $\Delta>2 \sigma_{0}+\sigma_{1}+1$.
We proceed by finding an upper bound for $\tau$, given that it satisfies (89). Fix $\sigma_{0}, \sigma_{1}, \Delta$ and define

$$
\begin{equation*}
r(\zeta)=\zeta\left(\log \left(1+\frac{1}{\zeta}\right)-2 \log \left(\frac{4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta}{4 \tau}\right)-\log \left(\frac{\lambda^{4}}{\lambda_{3}^{4}} \frac{f_{3}\left(\lambda_{3}\right)}{f_{3}(\lambda)}\right)\right) \tag{92}
\end{equation*}
$$

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 $\zeta$,

$$
\begin{equation*}
2 \zeta \log \left(1+\frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{4 \zeta}\right) \leq 2 \zeta \frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{4 \zeta}=\frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{2} \tag{93}
\end{equation*}
$$

Let $h(x)=\log f_{3}(x)-4 \log x$. Then $h^{\prime}(x)=f_{2}(x) / f_{3}(x)-4 / x$, and we note that $h^{\prime}(\lambda)=0$, by definition of $\lambda$. The second derivative is $h^{\prime \prime}(x)=f_{1}(x) / f_{3}(x)-f_{2}(x)^{2} / f_{3}(x)^{2}+4 / x^{2}$. Substituting $f_{1}(x)=f_{3}(x)+x+x^{2} / 2$ and $f_{2}(x)=f_{3}(x)+x^{2} / 2$, for all $x \geq \lambda$

$$
\begin{aligned}
h^{\prime \prime}(x) & =\frac{4}{x^{2}}+1+\frac{x+x^{2} / 2}{f_{3}(x)}-1-\frac{x^{2}}{f_{3}(x)}-\frac{x^{4}}{4 f_{3}(x)^{2}} \\
& =\frac{4}{x^{2}}-\frac{x^{2}-2 x}{2 f_{3}(x)}-\frac{x^{4}}{4 f_{3}(x)^{2}}
\end{aligned}
$$

Since $x \geq \lambda>2$ we have $x^{2}-2 x>0$, and $f_{3}(x)<e^{x}$ implies

$$
\begin{aligned}
h^{\prime \prime}(x) & =\frac{4}{x^{2}}-\frac{x^{2}-2 x}{2 f_{3}(x)}-\frac{x^{4}}{4 f_{3}(x)^{2}} \\
& \leq \frac{4}{x^{2}}-\frac{x^{2}-2 x}{2 e^{x}} \\
& \leq \frac{4}{x^{2}}+\frac{2 x}{2 e^{x}} \\
& \leq \frac{4}{x^{2}}+x^{1-\lambda}
\end{aligned}
$$

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

By Taylor's theorem, for some $x \in\left[\lambda, \lambda_{3}\right]$

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

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 $g_{0}^{\prime}(x) \geq g_{0}^{\prime}(\lambda) \geq 1 / 2$ for $x \geq \lambda$, so $d g_{0}^{-1}(y) / d y \leq 2$ for $y \geq 4$, and for some $y \geq 4$

$$
\begin{equation*}
\lambda_{3}=\lambda+\frac{d g_{0}^{-1}(y)}{d y}\left(\frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{\tau}\right) \leq \lambda+2 \frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{\tau} \tag{94}
\end{equation*}
$$

and so

$$
\begin{equation*}
\log \left(\frac{\lambda^{4}}{\lambda_{3}^{4}} \frac{f_{3}\left(\lambda_{3}\right)}{f_{3}(\lambda)}\right) \leq \frac{3}{8}\left(\lambda_{3}-\lambda\right)^{2} \leq \frac{3}{2}\left(\frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{\tau}\right)^{2} \tag{95}
\end{equation*}
$$

Define $\tau_{1}$ as the unique solution $\zeta$ 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 by the discussion after (68). So, since $r_{1}\left(\tau_{1}\right)=$ $r(\tau)=2-2 \sigma_{0}-\sigma_{1}$, it follows that $\tau \leq \tau_{1}$.

## Case of $R_{1}$ :

Now fix $\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in R_{1}$, i.e. suppose $\Delta \geq 3.6$. Then

$$
\begin{aligned}
r_{1}\left(\frac{3}{4}\right) & =\frac{3}{4} \log \left(1+\frac{4}{3}\right)-\frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{2}-2\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2} \\
& =\frac{3}{4} \log \frac{7}{3}-2 \sigma_{0}-\sigma_{1}+\frac{\Delta}{2}-2\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2} \\
& \geq \frac{3}{4} \log \frac{7}{3}-2 \sigma_{0}-\sigma_{1}+\frac{3.6}{2}-2(4-3.6)^{2} \\
& >2-2 \sigma_{0}-\sigma_{1}
\end{aligned}
$$

We have $\lim _{\zeta \rightarrow 0} r_{1}(\zeta) \leq 0$, and $r_{1}$ is continous and increasing, so $\tau \leq \tau_{1}<3 / 4$. Since $\Delta \geq 3.6$ and $2 \sigma_{0}+\sigma_{1} \leq 2$,

$$
\begin{equation*}
\Delta-\left(4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta\right) \geq 2 \Delta-3-4 \sigma_{0}-2 \sigma_{1} \geq 7.2-7>0 \tag{96}
\end{equation*}
$$

This implies that

$$
\begin{equation*}
L_{1}\left(\sigma_{0}, \sigma_{1}, \Delta\right)=\lambda_{3} \sqrt{\frac{\Delta}{4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta}}>\lambda_{3} \tag{97}
\end{equation*}
$$

Note that

$$
G\left(\sigma_{0}, \sigma_{1}, \lambda\right) \geq G\left(\sigma_{0}, \sigma_{1}, \bar{\lambda}\right)=4 \sigma_{0}+2 \sigma_{1} \geq \Delta
$$

implies that

$$
L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right) \leq \lambda=g_{0}^{-1}(4)
$$

Also note that by (41) and (55) 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

$$
L_{1}\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right)>\lambda_{3} \geq \lambda \geq L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)
$$

for $\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in R_{1}$.
Case of $R_{2}, R_{3}$ :
For $R_{2}, R_{3}$ we will need a new bound on $\tau$. 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 $\tau_{2}$ be defined by $r_{2}\left(\tau_{2}\right)=2-2 \sigma_{0}-\sigma_{1}$, which can be solved for $\tau_{2}$;

$$
\tau_{2}=\frac{1+3\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2}}{\Delta-2}
$$

It follows from $r(\tau) \geq r_{2}(\tau)$ and the fact that $r_{2}$ is increasing that

$$
\begin{equation*}
\tau \leq \frac{1+3\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2}}{\Delta-2} \tag{98}
\end{equation*}
$$

An upper bound for $L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)$ will follow from bounding the partial derivative of $G\left(\sigma_{0}, \sigma_{1}, x\right)$ with respect to $x$. We have $g_{0}^{\prime} \geq 1 / 4, g_{1}^{\prime} \geq 1 / 3$ and $g_{2}^{\prime} \geq 1 / 2$ by Lemma 7 (Appendix B), so

$$
\begin{aligned}
\frac{\partial}{\partial x} G\left(\sigma_{0}, \sigma_{1}, x\right) & =\sigma_{0} g_{0}^{\prime}(x)+\sigma_{1} g_{1}^{\prime}(x)+\left(1-\sigma_{0}-\sigma_{1}\right) g_{2}^{\prime}(x) \\
& \geq \frac{\sigma_{0}}{4}+\frac{\sigma_{1}}{3}+\frac{1-\sigma_{0}-\sigma_{1}}{2} \\
& =\frac{6-3 \sigma_{0}-2 \sigma_{1}}{12}
\end{aligned}
$$

and $G\left(\sigma_{0}, \sigma_{1}, 0\right)=2 \sigma_{0}+\sigma_{1}+1$ implies

$$
\begin{aligned}
\Delta & =G\left(\sigma_{0}, \sigma_{1}, L_{2}(\Delta)\right) \\
& \geq G\left(\sigma_{0}, \sigma_{1}, 0\right)+\min _{x} \frac{\partial}{\partial x} G\left(\sigma_{0}, \sigma_{1}, x\right) L_{2}(\Delta) \\
& \geq 2 \sigma_{0}+\sigma_{1}+1+\frac{6-3 \sigma_{0}-2 \sigma_{1}}{12} L_{2}(\Delta)
\end{aligned}
$$

So

$$
\begin{equation*}
L_{2}(\Delta) \leq \frac{12\left(\Delta-2 \sigma_{0}-\sigma_{1}-1\right)}{6-3 \sigma_{0}-2 \sigma_{1}} . \tag{99}
\end{equation*}
$$

So, to show $L_{1}\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \geq L_{2}\left(\sigma_{0}, \sigma_{1}, \Delta\right)$, it is enough to show that

$$
\begin{equation*}
\lambda_{3} \sqrt{\frac{\Delta}{4 \tau+4 \sigma_{0}+2 \sigma_{1}-\Delta}}>\frac{12\left(\Delta-2 \sigma_{0}-\sigma_{1}-1\right)}{6-3 \sigma_{0}-2 \sigma_{1}} \tag{100}
\end{equation*}
$$

Solving for $\tau$, this is equivalent to showing

$$
\begin{equation*}
\tau<\Delta\left[\frac{\lambda_{3}\left(6-3 \sigma_{0}-2 \sigma_{1}\right)}{24\left(\Delta-2 \sigma_{0}-\sigma_{1}-1\right)}\right]^{2}-\frac{4 \sigma_{0}+2 \sigma_{1}-\Delta}{4} \tag{101}
\end{equation*}
$$

and by (98), and $\lambda_{3} \geq \lambda$, it is enough to show

$$
\begin{equation*}
\frac{1+3\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2}}{\Delta-2}<\Delta\left[\frac{\lambda\left(6-3 \sigma_{0}-2 \sigma_{1}\right)}{24\left(\Delta-2 \sigma_{0}-\sigma_{1}-1\right)}\right]^{2}-\frac{\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)}{4} \tag{102}
\end{equation*}
$$

for $\left(\sigma_{0}, \sigma_{1}, \Delta, \tau\right) \in R_{2} \cup R_{3}$.

## 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

$$
\begin{equation*}
\frac{1+3\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2}}{\Delta-2} \leq \frac{1.03}{\Delta-2} \tag{103}
\end{equation*}
$$

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 $\lambda>2.5$, so it holds that

$$
\begin{equation*}
\Delta\left[\frac{\lambda\left(6-3 \sigma_{0}-2 \sigma_{1}\right)}{24\left(\Delta-2 \sigma_{0}-\sigma_{1}-1\right)}\right]^{2}-\frac{\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)}{4}>\Delta\left[\frac{2.5 \times 3.9}{24(\Delta-2)}\right]^{2}-0.025 \tag{104}
\end{equation*}
$$

and it is enough to show that

$$
\begin{equation*}
\frac{1.03}{\Delta-2} \leq \Delta\left[\frac{2.5 \times 3.9}{24(\Delta-2)}\right]^{2}-0.025 \tag{105}
\end{equation*}
$$

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,

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

Define

$$
u_{1}=5.5, \quad u_{2}=5.75, \quad u_{3}=5.875, \quad u_{4}=5.9375 .
$$

Then

$$
6-3 \sigma_{0}-2 \sigma_{1}=\left(6-\frac{\sigma_{1}}{2}\right)-3 \sigma_{0}-\frac{3 \sigma_{1}}{2} \geq u_{i}-\frac{3\left(2 \sigma_{0}+\sigma_{1}\right)}{2}
$$

in $R_{3, i}, i=1,2,3,4$.
Fixing $i,(102)$ will hold in $R_{3, i}$ if we can show that

$$
\begin{equation*}
\frac{1+3\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)^{2}}{\Delta-2} \leq \Delta\left[\frac{\lambda\left(u_{i}-3\left(2 \sigma_{0}+\sigma_{1}\right) / 2\right)}{24\left(\Delta-2 \sigma_{0}-\sigma_{1}-1\right)}\right]^{2}-\frac{\left(4 \sigma_{0}+2 \sigma_{1}-\Delta\right)}{4} \tag{106}
\end{equation*}
$$

Note that $\sigma_{0}, \sigma_{1}$ only appear as $\Sigma=2 \sigma_{0}+\sigma_{1}$ in (106). For this reason we clear denominators in (106) and define for $i=1,2,3,4$,

$$
\begin{aligned}
\varphi_{i}(\Sigma, \Delta)= & \lambda^{2} \Delta(\Delta-2)\left(u_{i}-3 \Sigma / 2\right)^{2}-144(\Delta-2)(\Delta-\Sigma-1)^{2}(2 \Sigma-\Delta) \\
& -576(\Delta-\Sigma-1)^{2}-1728(\Delta-\Sigma-1)^{2}(2 \Sigma-\Delta)^{2} .
\end{aligned}
$$

In which case, (106) is equivalent to $\varphi_{i}(\Sigma, \Delta) \geq 0$.
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 (54).
Equation (106) will follow from showing that $\varphi_{i}(\Sigma, \Delta) \geq 0$ in $\widetilde{R}_{3, i}, i=1,2,3,4$.
The $\varphi_{i}$ are degree four polynomials, and bounds on $\left|\nabla \varphi_{i}\right|$ are found by applying the triangle inequality to the partial derivatives of $\varphi_{i}$. 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

$$
\begin{aligned}
\left|\frac{\partial \varphi_{i}}{\partial \Sigma}\right|= & \mid-3 \lambda^{2} \Delta(\Delta-2)\left(u_{i}-3 \Sigma / 2\right)+288(\Delta-2)(\Delta-\Sigma-1)(2 \Sigma-\Delta) \\
& -288(\Delta-2)(\Delta-\Sigma-1)^{2}+1152(\Delta-\Sigma-1) \\
& +3456(\Delta-\Sigma-1)(2 \Sigma-\Delta)^{2}-6912(\Delta-\Sigma-1)^{2}(2 \Sigma-\Delta) \mid \\
\leq & 3 \lambda^{2} \Delta(\Delta-2)\left(u_{i}-3 \Sigma / 2\right)+288(\Delta-2)(\Delta-\Sigma-1)|3 \Sigma-2 \Delta+1| \\
& +1152(\Delta-\Sigma-1)+3456(\Delta-\Sigma-1)(2 \Sigma-\Delta)|4 \Sigma-3 \Delta+2| \\
\leq & 27 \cdot 4 \cdot 2 \cdot 9 / 2+288 \cdot 2 \cdot 2 \cdot 1+1152 \cdot 2+3456 \cdot 3 / 4 \cdot 2 \\
= & 9612
\end{aligned}
$$

For $\Delta$,

$$
\begin{aligned}
\left|\frac{\partial \varphi_{i}}{\partial \Delta}\right|= & \mid \lambda^{2} \Delta\left(u_{i}-3 \Sigma / 2\right)^{2}+\lambda^{2}(\Delta-2)\left(u_{i}-3 \Sigma / 2\right)^{2}-144(\Delta-\Sigma-1)^{2}(2 \Sigma-\Delta) \\
& -288(\Delta-2)(\Delta-\Sigma-1)(2 \Sigma-\Delta)+144(\Delta-2)(\Delta-\Sigma-1)^{2} \\
& -1152(\Delta-\Sigma-1)-3456(\Delta-\Sigma-1)(2 \Sigma-\Delta)^{2}+3456(\Delta-\Sigma-1)^{2}(2 \Sigma-\Delta) \mid \\
\leq & \lambda^{2}(2 \Delta-2)\left(u_{i}-3 \Sigma / 2\right)^{2}+144(\Delta-\Sigma-1)^{2}(2 \Sigma-\Delta)+288(\Delta-2)(\Delta-\Sigma-1)(2 \Sigma-\Delta) \\
& +144(\Delta-2)(\Delta-\Sigma-1)^{2}+1152(\Delta-\Sigma-1)+3456(\Delta-\Sigma-1)(2 \Sigma-\Delta)|2 \Delta-3 \Sigma-1| \\
\leq & 9 \cdot 6 \cdot(9 / 2)^{2}+144 \cdot 2^{2} \cdot 1+288 \cdot 2 \cdot 2 \cdot 1+144 \cdot 2 \cdot 2^{2}+1152 \cdot 2+3456 \cdot 3 / 4 \cdot 1 \\
= & 8383.5
\end{aligned}
$$

so $\left|\nabla \varphi_{i}\right| \leq 12755$ 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 $\left|x-x_{0}\right| \leq 0.001$. On this grid, $\varphi_{i}$ is calculated numerically, and it is found that

$$
\min _{x_{0} \in \mathcal{P}_{i}} \varphi_{i}\left(x_{0}\right)= \begin{cases}22.49, & i=1  \tag{107}\\ 25.50, & i=2 \\ 27.08, & i=3 \\ 19.04, & i=4\end{cases}
$$

So for any $i$ and any $x \in \widetilde{R}_{3, i}$, there exists an $x_{0}$ such that $\left|\varphi_{i}(x)-\varphi_{i}\left(x_{0}\right)\right| \leq\left|\nabla \varphi_{i}\right|\left|x-x_{0}\right| \leq$ $12755 \cdot 0.001<13$, which implies $\varphi_{i}(x)>\varphi_{i}\left(x_{0}\right)-13>0$. This proves (102) for $\sigma_{0}, \sigma_{1}, \Delta \in R_{3}$.

## B Appendix C

This section is concerned with the functions

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

and the related functions

$$
\begin{equation*}
g_{0}(x)=\frac{x f_{2}(x)}{f_{3}(x)}, \quad g_{1}(x)=\frac{x f_{1}(x)}{f_{2}(x)}, \quad g_{2}(x)=\frac{x f_{0}(x)}{f_{1}(x)} . \tag{108}
\end{equation*}
$$

Since $f_{k}(0)=0$ for $k \geq 1$, we define $g_{i}(0)=\lim _{x \rightarrow 0} g_{i}(x)=3-i$. Note that

$$
\begin{equation*}
\frac{d}{d x} f_{k}(x)=f_{k-1}(x), \quad k \geq 1 \tag{109}
\end{equation*}
$$

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

$$
\begin{equation*}
x<g_{i}(x) \leq 3-i+x \tag{110}
\end{equation*}
$$

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

$$
\begin{equation*}
g_{i}(x)-x=\frac{x f_{2-i}(x)}{f_{3-i}(x)}-x=\frac{x\left(f_{2-i}(x)-f_{3-i}(x)\right)}{f_{3-i}(x)}=\frac{x^{3-i}}{(2-i)!f_{3-i}(x)} . \tag{111}
\end{equation*}
$$

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

$$
\begin{equation*}
(3-i)(2-i)!f_{3-i}(x)-x^{3-i}=(3-i)!\sum_{k \geq 3-i} \frac{x^{k}}{k!}-x^{3-i}=(3-i)!\sum_{k \geq 4-i} \frac{x^{k}}{k!}>0 \tag{112}
\end{equation*}
$$

for $x>0$, implying $g_{i}(x)-x<3-i$.
Lemma 7. The functions $g_{0}, g_{1}, g_{2}$ are convex, and $g_{i}^{\prime}(x) \geq 1 /(4-i)$ 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

$$
\begin{equation*}
g_{0}(x)=\frac{x f_{2}(x)}{f_{3}(x)}=x+\frac{x^{3}}{2 f_{3}(x)} \tag{113}
\end{equation*}
$$

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

$$
\begin{equation*}
g_{0}^{\prime}(x)=1-\frac{q^{\prime}(x)}{2 q(x)^{2}}, \quad g_{0}^{\prime \prime}(x)=\frac{2 q^{\prime}(x)^{2}-q(x) q^{\prime \prime}(x)}{2 q(x)^{3}} \tag{114}
\end{equation*}
$$

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

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

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

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

(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 $2 q^{\prime}(x)^{2}-q(x) q^{\prime \prime}(x) \geq 0$, and it follows that $g_{0}$ is convex. Similar arguments show that $g_{1}, g_{2}$ are convex.

For $i=0,1$,

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

Now

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

And

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

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

$$
g_{i}^{\prime}(x)=\frac{1}{4-i}+O(x)
$$

For $i=2$ we have

$$
g_{2}^{\prime}(x)=\frac{e^{x}}{f_{1}(x)}+\frac{x e^{x}}{f_{1}(x)}-\frac{x e^{2 x}}{f_{1}(x)^{2}}=e^{x}\left(\frac{f_{1}(x)(1+x)-x e^{x}}{f_{1}(x)^{2}}\right)=e^{x}\left(\frac{\frac{x^{2}}{2}+O\left(x^{3}\right)}{x^{2}+O\left(x^{3}\right)}\right)=\frac{1}{2}+O(x) .
$$

And by the convexity of $g_{i}$ we have $g_{i}^{\prime}(x) \geq 1 /(4-i)$ for all $x \geq 0$.

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.

$$
\begin{array}{ll}
\text { (i) } & 1 \leq \frac{f_{2}(x)^{2}}{f_{1}(x) f_{3}(x)} \leq 2 \\
\text { (ii) } & 0.09<\frac{f_{3}(x)}{x^{2} f_{1}(x)} \leq \frac{1}{6} \\
\text { (iii) } & \frac{f_{2}(x)}{x f_{1}(x)} \leq \frac{1}{3} \\
\text { (iv) } & 0.01<\frac{f_{3}(x)}{x^{4} f_{1}(x)} \\
\text { (v) } & 0.09<\frac{f_{3}(x)}{x^{2} f_{2}(x)}
\end{array}
$$

Proof. Consider (i). 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

$$
\begin{equation*}
\left(e^{x}-1-x\right)^{2}=\left(e^{x}-1\right)\left(e^{x}-1-x-\frac{x^{2}}{2}\right) \tag{115}
\end{equation*}
$$

Expanding and reordering terms, we have

$$
\begin{equation*}
e^{x}\left(x+\frac{x^{2}}{2}\right)=x+\frac{x^{2}}{2} \tag{116}
\end{equation*}
$$

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 $f_{2}(x)^{2}=2 f_{1}(x) f_{3}(x)$. This simplifies to

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

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;

$$
\begin{equation*}
\frac{f_{3}(x)}{x^{2} f_{1}(x)}=\frac{1}{g_{0}(x) g_{1}(x)}, \quad \frac{f_{2}(x)}{x f_{1}(x)}=\frac{1}{g_{1}(x)}, \quad \frac{f_{3}(x)}{x^{4} f_{1}(x)}=\frac{1}{x^{2} g_{0}(x) g_{1}(x)}, \quad \frac{f_{3}(x)}{x^{2} f_{2}(x)}=\frac{1}{x g_{0}(x)} \tag{117}
\end{equation*}
$$

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 $x=2.688>\lambda$.


[^0]:    *Research supported in part by NSF Grant DMS1362785. Email: alan@random.math.cmu.edu
    ${ }^{\dagger}$ Research supported in part by NSF Grant DMS1362785. Email: alan@random.math.cmu.edu

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

[^2]:    ${ }^{2}$ https://isc.carma.newcastle.edu.au/

