On a Greedy 2-Matching Algorithm and Hamilton Cycles in Random Graphs with Minimum Degree at Least Three.

Alan Frieze[*]
Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh PA15217.

December 12, 2012

Abstract

We describe and analyse a simple greedy algorithm 2GREEDY that finds a good 2-matching M in the random graph $G = G^{\delta\geq 3}_{n,cn}$ when $c \geq 10$. A 2-matching is a spanning subgraph of maximum degree two and G is drawn uniformly from graphs with vertex set $[n]$, cn edges and minimum degree at least three. By good we mean that M has $O(\log n)$ components. We then use this 2-matching to build a Hamilton cycle in $O(n^{1.5+o(1)})$ time w.h.p..

1 Introduction

There have been many papers written on the existence of Hamilton cycles in random graphs. Komlós and Szemerédi [18], Bollobás [6], Ajtai, Komlós and Szemerédi [1] showed that the question is intimately related to the minimum degree. Loosely speaking, if we are considering random graphs with n vertices and minimum degree at least two then we need $\Omega(n \log n)$ edges in order that they are likely to be Hamiltonian.

For sparse random graphs with $O(n)$ random edges, one needs to have minimum degree at least three. This is to avoid having three vertices of degree two sharing a common neighbor. There are several models of a random graph in which minimum degree three is satisfied: Random regular graphs of degree at least three, Robinson and Wormald [22], [23] or the random graph G_{3-out}, Bohman and Frieze [5]. Bollobás, Cooper, Fenner and Frieze [8] considered the classical random graph $G_{n,m}$ with conditioning on the minimum degree k i.e. each graph with vertex set $[n]$ and m edges and minimum degree at least k is considered to be equally likely. Denote this model of a random graph by $G^{\delta\geq k}_{n,m}$. They showed that for every $k \geq 3$ there is a $c_k \leq (k+1)^3$ such that if $c \geq c_k$ then w.h.p. $G^{\delta\geq k}_{n,cn}$ has $(k-1)/2$ edge disjoint Hamilton cycles, where a perfect matching constitutes half a Hamilton cycle in the case where k is even. It is reasonable to conjecture that $c_k = k/2$. The results of this paper and a companion [15] reduce the known value of c_3 from 64 to below 10 (below 2 if you accept a "numerical proof"). It can be argued that replacing one incorrect upper bound

[*]Research supported in part by NSF Grant CCF1013110

by a smaller incorrect upper bound does not constitute significant progress. However, the main contribution of this paper is to introduce a new greedy algorithm for finding a large 2-matching in a random graph and to give a (partial) analysis of its performance and of course to apply it to the Hamilton cycle problem.

One is interested in the time taken to construct a Hamilton cycle in a random graph. Angluin and Valiant [2] and Bollobás, Fenner and Frieze [9] give polynomial time algorithms. The algorithm in [2] is very fast, $O(n\log^2 n)$ time, but requires $Kn\log n$ random edges for sufficiently large $K>0$. The algorithm in [9] is of order $n^{3+o(1)}$ but works w.h.p. at the exact threshold for Hamiltonicity. Frieze [13] gave an $O(n^{3+o(1)})$ time algorithm for finding large cycles in sparse random graphs and this could be adapted to find Hamilton cycles in $G_{n,cn}^{\delta\geq 3}$ in this time for sufficiently large c. Another aim of [15] and this paper is reduce this running time. The results of this paper and its companion [15] will reduce this to $n^{1.5+o(1)}$ for sufficiently large c, and perhaps in a later paper, we will further reduce the running time by borrowing ideas from a linear expected time algorithm for matchings due to Chebolu, Frieze and Melsted [11].

The idea of [11] is to begin the process of constructing a perfect matching by using the Karp-Sipser algorithm [17] to find a good matching and then build this up to a perfect matching by using alternating paths. The natural extension of this idea is to find a good 2-matching and then use extension-rotation arguments to transform it into a Hamilton cycle. A 2-matching M of G is a spanning subgraph of maximum degree 2. Each component of M is a cycle or a path (possibly an isolated vertex) and we let $\kappa(M)$ denote the number of components of M. The time taken to transform M into a Hamilton cycle depends heavily on $\kappa(M)$. The aim is to find a 2-matching M for which $\kappa(M)$ is small. The main result of this paper is the following:

Theorem 1. *There is an absolute constant $c_0>0$ and such that if $c\geq c_0$ then w.h.p.* 2GREEDY *finds a 2-matching M with $\kappa(M)=O(\log n)$. (This paper gives an analytic proof that $c_0\leq 10$. We have a numerical proof that $c_0\leq 1.6$).*

Given this theorem, we will show how we can use this and the result of [15] to show

Theorem 2. *If $c\geq c_0$ then w.h.p. a Hamilton cycle can be found in $O(n^{1.5+o(1)})$ time.*

Acknowledgement: I would like to thank my colleague Boris Pittel for his help with this paper. He ought to be a co-author, but he has declined to do so.

2 Outline of the paper

As already indicated, the idea is to use a greedy algorithm to find a good 2-matching and then transform it into a Hamilton cycle. We will first give an over-view of our greedy algorithm. As we proceed, we select edges to add to our 2-matching M. Thus M consists of paths and cycles (and isolated vertices). Vertices of the cycles and vertices interior to the paths get deleted from the current graph, which we denote by Γ. No more edges can be added incident to these interior vertices. Thus the paths can usefully be thought of as being contracted to the set of edges of a matching M^* on the remaining vertices of Γ. This matching is not part of Γ. We keep track of the vertices covered by M^* by using a 0/1 vector b so that for vertex v, $b(v)$ is the indicator that v is covered by M^*. Thus when v is still included in Γ and $b(v) = 1$, it will be the end-point of a path in the current 2-matching M.

The greedy algorithm first tries to cover vertices of degree at most two that are not covered by M or vertices of degree one that are covered by M. These choices are forced. When there are no such vertices, we choose an edge at random. We make sure that one of the end-points u, v of the chosen edge has b-value zero. The aim here is to try to quickly ensure that $b(v) = 1$ for all vertices of Γ. This will essentially reduce the problem to that of finding another (near) perfect matching in Γ. The first phase of the algorithm finishes when all of the vertices that remain have b-value one. This necessarily means that the contracted paths form a matching of the graph Γ that remains at this stage. Furthermore, we will see that Γ is distributed as $G_{\nu,\mu}^{\delta\geq 2}$ for some ν, μ and then we construct another (near) perfect matching M^{**} of Γ by using the linear expected time algorithm of [11]. We put M and M^{**} together to create a 2-matching along with the cycles that have been deleted. Note that some vertices may have become isolated during the construction of M and these will form single components of our 2-matching. The union of two random (near) perfect matchings is likely to have $O(\log n)$ components. Full details of this algorithm are given in Section 4.

Once we have described the algorithm, we can begin its analysis. We first describe the random graph model that we will use. We call it the *Random Sequence model*. It was first used in Bollobás and Frieze [10] and independently in Chvatál [12]. We used it in [3] for our analysis of the Karp-Sipser algorithm. We prove the truncated Poisson nature of the degree sequence of the graph Γ that remains at each stage in Section 3. We then, in Section 4, give a detailed description of 2Greedy. In Section 5 we show that the distribution of the evolving graph Γ can be succinctly described by a 6-component vector $\mathbf{v} = (y_1, y_2, z_1, y, z, \mu)$ that evolves as a Markov chain. Here $y_j, j = 1, 2$ denotes the number of vertices of degree j that are not incident with M and z_1 denotes the number of vertices of degree one that are incident with M. y denotes the number of vertices of degree at least three that are not incident with M and z denotes the number of vertices of degree at least two that are incident with M. μ denotes the number of edges. It is important to keep $\zeta = y_1 + 2y_2 + z_1$ small and 2Greedy will deal immediately with low degree vertices when $\zeta > 0$. In this way we keep ζ small w.h.p. throughout the algorithm and this will mean that the final 2-matching produced will have few components. Section 6 first describes the (approximate) transition probabilities of this chain. There are four types of step in 2Greedy that depend on which if any of y_1, y_2, z_1 are positive. Thus there are four sets of transition probabilities. Given the expected changes in $\mathbf{v}$, we first show that in all cases the expected change in ζ is negative, when ζ is positive. This indicates that ζ will not get large and a high probability polylog bound is proven.

We are using the differential equation method and Section 7 describes the sets of differential equations that can be used to track the progress of the algorithm w.h.p.. The parameters for these equations will be $\hat{\mathbf{v}} = (\hat{y}_1, \hat{y}_2, \hat{z}, \hat{y}, \hat{z}, \hat{\mu})$. There are four sets of equations corresponding to the four types of step in 2Greedy. It is important to know the proportion of each type of step over a small interval. We thus consider a *sliding trajectory* i.e. a weighted sum of these four sets of equations. The weights are chosen so that in the weighted set of equations we have $\hat{y}_1' = \hat{y}_2' = \hat{z}_1' = 0$. This is in line with the fact that $\hat{y}, \hat{z}, \hat{\mu} \gg \zeta$ for most of the algorithm. We verify that the expressions for the weights are non-negative. We then verify that w.h.p. the sliding trajectory and the process parameters remain close.

Our next aim is to show that w.h.p. there is a time T such that $y(T) = 0, z(T) = \Omega(n)$. It would therefore be most natural to show that for the sliding trajectory, there is a time $\hat{T}$ such that $\hat{y}(\hat{T}) = 0, \hat{z}(\hat{T}) = \Omega(n)$. The equations for the sliding trajectory are complicated and we have not been able to do this directly. Instead, we have set up an approximate system of equations (in parameters $\tilde{y}, \tilde{z}, \tilde{\mu}$) that are close when $c \geq 10$. We can prove these parameters stay close to $\hat{y}, \hat{z}, \hat{\mu}$ and that there is a time $\tilde{T}$ such that $\tilde{y}(\tilde{T}) = 0, \tilde{z}(\tilde{T}) = \Omega(n)$. The existence of $\hat{T}$ is deduced from this and then we can deduce the existence of T. We then in Section 9 show that w.h.p. 2greedy creates a matching with $O(\log n)$ components, completing the proof of Theorem 1.

Section 10 shows how to use an extension-rotation procedure on our graph G to find a Hamilton cycle within the claimed time bounds. This procedure works by extending paths one edge at a time and using an operation called a rotation to increase the number of chances of extending a path. It is not guaranteed to extend a path, even if it is possible some other way. There is the notion of a *booster*. This is a non-edge whose addition will allow progress in the extension-rotation algorithm. The companion paper [15] shows that for $c \geq 2.67$ there will w.h.p. always be many boosters. To get the non-edges we first randomly choose $s = n^{1/2} \log^{-2} n$ random edges X of G, none of which are incident with a vertex of degree three. We then write $G = G' + X$ and argue in Section 10.1 that the pair (G', X) can be replaced by (H, Y) where $H = G_{n,cn-s}^{\delta \geq 3}$ and Y is an independent random set of edges disjoint from $E(H)$. We then argue in Section 10.3 that w.h.p. Y contains enough boosters to create a Hamilton cycle within the claimed time bound.

Section 11 contains some concluding remarks.

3 Random Sequence Model

A small change of model will simplify the analysis. Given a sequence $\mathbf{x} = (x_1, x_2, \ldots, x_{2M}) \in [n]^{2M}$ of $2M$ integers between 1 and N we can define a (multi)-graph $G_{\mathbf{x}} = G_{\mathbf{x}}(N, M)$ with vertex set $[N]$ and edge set $\{(x_{2i-1}, x_{2i}) : 1 \leq i \leq M\}$. The degree $d_{\mathbf{x}}(v)$ of $v \in [N]$ is given by

$$d_{\mathbf{x}}(v) = |\{j \in [2M] : x_j = v\}|.$$

If $\mathbf{x}$ is chosen randomly from $[N]^{2M}$ then $G_{\mathbf{x}}$ is close in distribution to $G_{N,M}$. Indeed, conditional on being simple, $G_{\mathbf{x}}$ is distributed as $G_{N,M}$. To see this, note that if $G_{\mathbf{x}}$ is simple then it has vertex set $[N]$ and M edges. Also, there are $M!2^M$ distinct equally likely values of $\mathbf{x}$ which yield the same graph.

Our situation is complicated by there being lower bounds of 2, 3 respectively on the minimum degree in two disjoint sets $J_2, J_3 \subseteq [N]$. Initially $J_2 = J_3 = \emptyset$ but we will have to consider instances where they are non-empty, as our 2-matching algorithm progresses. The vertices in $J_0 = [N] \setminus (J_2 \cup J_3)$ are of fixed bounded degree and the sum of their degrees is $D = o(N)$. So we let

$$[N]_{J_2,J_3;D}^{2M} = \{\mathbf{x} \in [N]^{2M} : d_{\mathbf{x}}(j) \geq i \text{ for } j \in J_i,\ i = 2, 3 \text{ and } \sum_{j \in J_0} d_{\mathbf{x}}(j) = D\}.$$

Let $G = G(N, M, J_2, J_3; D)$ be the multi-graph $G_{\mathbf{x}}$ for $\mathbf{x}$ chosen uniformly from $[N]_{J_2,J_3;D}^{2M}$. It is clear then that conditional on being simple, $G(n, m, \emptyset, [n]; 0)$ has the same distribution as $G_{n,m}^{\delta \geq 3}$. It is important therefore to estimate the probability that this graph is simple. For this and other reasons, we need to have an understanding of the degree sequence $d_{\mathbf{x}}$ when $\mathbf{x}$ is drawn uniformly from $[N]^{2M}_{J_2,J_3;D}$. Let

$$f_k(\lambda) = e^{\lambda} - \sum_{i=0}^{k-1} \frac{\lambda^i}{i!}$$

for $k \geq 0$.

Lemma 3.1. *Let* $\mathbf{x}$ *be chosen randomly from* $[N]^{2M}_{J_2,J_3;D}$. *For* $i = 2, 3$ *let* Z_j $(j \in [J_i])$ *be independent copies of a* truncated Poisson *random variable* $\mathcal{P}_i$, *where*

$$\mathbb{P}(\mathcal{P}_i = t) = \frac{\lambda^t}{t! f_i(\lambda)}, \qquad t = i, i+1, \ldots .$$

Here λ *satisfies*

$$\sum_{i=2}^{3} \frac{\lambda f_{i-1}(\lambda)}{f_i(\lambda)} |J_i| = 2M - D. \tag{1}$$

For $j \in J_0$, $Z_j = d_j$ *is a constant and* $\sum_{j \in J_0} d_j = D$. *Then* $\{d_{\mathbf{x}}(j)\}_{j \in [N]}$ *is distributed as* $\{Z_j\}_{j \in [N]}$ *conditional on* $Z = \sum_{j \in [n]} Z_j = 2M$.

Proof Note first that the value of λ in (1) is chosen so that

$$\mathbb{E}(Z) = 2M.$$

Fix J_0, J_2, J_3 and $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_N)$ such that $\xi_j = d_j$ for $j \in J_0$ and $\xi_j \geq k$ for $k = 2, 3$ and $j \in J_k$. Then,

$$\mathbb{P}(d_{\mathbf{x}} = \boldsymbol{\xi}) = \left(\frac{(2M)!}{\xi_1! \xi_2! \ldots \xi_N!} \right) \Bigg/ \left(\sum_{\mathbf{x} \in [N]^{2M}_{J_2,J_3;D}} \frac{(2M)!}{x_1! x_2! \ldots x_N!} \right).$$

On the other hand,

$$\mathbb{P}\left((Z_1, Z_2, \ldots, Z_N) = \boldsymbol{\xi} \,\Bigg|\, \sum_{j \in [N]} Z_j = 2M \right) =$$

$$\left((2M)! \prod_{j \in J_0} \frac{1}{d_j!} \prod_{i=2}^{3} \prod_{j \in J_i} \frac{\lambda^{\xi_j}}{f_i(\lambda) \xi_j!} \right) \Bigg/ \left(\sum_{\mathbf{x} \in [N]^{2M}_{J_2,J_3;D}} (2M)! \prod_{j \in J_0} \frac{1}{d_j!} \prod_{i=2}^{3} \prod_{j \in J_i} \frac{\lambda^{x_j}}{f_i(\lambda) x_j!} \right)$$

$$= \left(\frac{\prod_{i=2}^{3} f_i(\lambda)^{-|J_i|} \lambda^{2M-D}}{\xi_1! \xi_2! \ldots \xi_N!} \right) \Bigg/ \left(\sum_{\mathbf{x} \in [N]^{2M}_{J_2,J_3;D}} \frac{\prod_{i=2}^{3} f_i(\lambda)^{-|J_i|} \lambda^{2M-D}}{x_1! x_2! \ldots x_N!} \right)$$

$$= \mathbb{P}(d_{\mathbf{x}} = \boldsymbol{\xi}).$$

□

To use Lemma 3.1 for the approximation of vertex degrees distributions we need to have sharp estimates of the probability that Z is close to its mean $2M$. In particular we need sharp estimates of $\mathbb{P}(Z = 2M)$ and $\mathbb{P}(Z - Z_1 = 2M - k)$, for $k = o(N)$. These estimates are possible precisely because $\mathbb{E}(Z) = 2M$. Using the special properties of Z, we can refine a standard argument to show (Appendix 1) that where $N_\ell = |J_\ell|$ and $N^* = N_2 + N_3$ and the variances are

$$\sigma_\ell^2 = \frac{f_\ell(\lambda)(\lambda^2 f_{\ell-2}(\lambda) + \lambda f_{\ell-1}(\lambda)) - \lambda^2 f_{\ell-1}(\lambda)^2}{f_\ell(\lambda)^2} \text{ and } \sigma^2 = \frac{1}{N^*}\sum_{\ell=2}^{3} N_\ell \sigma_\ell^2, \tag{2}$$

that if $N^*\sigma^2 \to \infty$ and $k = O(\sqrt{N^*}\sigma)$ then

$$\mathbb{P}(Z = 2M - k) = \frac{1}{\sigma\sqrt{2\pi N^*}}\left(1 + O\left(\frac{k^2+1}{N^*\sigma^2}\right)\right). \tag{3}$$

A proof for $J_2 = [N]$ was given in the appendix of [3]. We need to modify the proof in a trivial way. Given (3) and

$$\sigma_\ell^2 = O(\lambda), \qquad \ell = 2, 3,$$

we obtain

Lemma 3.2. *Let* $\mathbf{x}$ *be chosen randomly from* $[N]^{2M}_{J_2,J_3;D}$.

(a) *Assume that* $\log N^* = O((N^*\lambda)^{1/2})$. *For every* $j \in J_\ell$ *and* $\ell \le k \le \log N^*$,

$$\mathbb{P}(d_{\mathbf{x}}(j) = k) = \frac{\lambda^k}{k! f_\ell(\lambda)}\left(1 + O\left(\frac{k^2+1}{N^*\lambda}\right)\right). \tag{4}$$

Furthermore, for all $\ell_1, \ell_2 \in \{2,3\}$ *and* $j_1 \in J_{\ell_1}, j_2 \in J_{\ell_2}$, $j_1 \neq j_2$, *and* $\ell_i \le k_i \le \log N^*$,

$$\mathbb{P}(d_{\mathbf{x}}(j_1) = k_1, d_{\mathbf{x}}(j_2) = k_2) = \frac{\lambda^{k_1}}{k_1! f_{\ell_1}(\lambda)}\frac{\lambda^{k_2}}{k_2! f_{\ell_2}(\lambda)}\left(1 + O\left(\frac{\log^2 N^*}{N^*\lambda}\right)\right). \tag{5}$$

(b)

$$d_{\mathbf{x}}(j) \le \frac{\log N}{(\log\log N)^{1/2}} \quad q.s.^1 \tag{6}$$

for all $j \in J_2 \cup J_3$.

Proof Assume that $j = 1 \notin J_0$. Then

$$\begin{aligned}\mathbb{P}(d_{\mathbf{x}}(1) = k) &= \frac{\mathbb{P}\left(Z_1 = k \text{ and } \sum_{i=1}^N Z_i = 2M\right)}{\mathbb{P}\left(\sum_{i=1}^N Z_i = 2M\right)} \\ &= \frac{\lambda^k}{k! f_\ell(\lambda)}\frac{\mathbb{P}\left(\sum_{i=2}^N Z_i = 2M - k\right)}{\mathbb{P}\left(\sum_{i=1}^N Z_i = 2M\right)}.\end{aligned}$$

Likewise, with $j_1 = 1, j_2 = 2$,

$$\mathbb{P}(d_{\mathbf{x}}(1) = k_1, d_{\mathbf{x}}(2) = k_2) = \frac{\lambda^{k_1}}{k_1! f_{\ell_1}(\lambda)}\frac{\lambda^{k_2}}{k_2! f_{\ell_2}(\lambda)}\frac{\mathbb{P}\left(\sum_{i=3}^N Z_i = 2M - k_1 - k_2\right)}{\mathbb{P}\left(\sum_{i=1}^N Z_i = 2M\right)}.$$

Statement (a) follows immediately from (3) and (b) follows from simple estimations. □

Let $\nu_{\mathbf{x}}^\ell(s)$ denote the number of vertices in $J_\ell, \ell = 2, 3$ of degree s in $G_{\mathbf{x}}$. Equation (3) and a standard tail estimate for the binomial distribution shows the following:

[1] An event $\mathcal{E} = \mathcal{E}(N^*)$ occurs quite surely (q.s., in short) if $\mathbb{P}(\mathcal{E}) = 1 - O(N^{-a})$ for any constant $a > 0$

Lemma 3.3. *Suppose that* $\log N^* = O((N^*\lambda)^{1/2})$ *and* $N_\ell \to \infty$ *with* N. *Let* $\mathbf{x}$ *be chosen randomly from* $[N]^{2M}_{J_2,J_3;D}$. *Then q.s.,*

$$\mathcal{D}(\mathbf{x}) = \left\{ \left| \nu^\ell_{\mathbf{x}}(j) - \frac{N_\ell \lambda^j}{j! f(\lambda)} \right| \le \left(1 + \left(\frac{N_\ell \lambda^j}{j! f(\lambda)} \right)^{1/2} \right) \log^2 N,\ k \le j \le \log N \right\}. \tag{7}$$

□

We can now show $G_{\mathbf{x}}$, $\mathbf{x} \in [n]^{2m}_{\emptyset,[n];0}$ is a good model for $G^{\delta\geq 3}_{n,m}$. For this we only need to show now that

$$\mathbb{P}(G_{\mathbf{x}} \text{ is simple}) = \Omega(1). \tag{8}$$

For this we can use a result of McKay [20]. If we fix the degree sequence of $\mathbf{x}$ then $\mathbf{x}$ itself is just a random permutation of the multi-graph in which each $j \in [n]$ appears $d_{\mathbf{x}}(j)$ times. This in fact is another way of looking at the configuration model of Bollobás [7]. The reference [20] shows that the probability $G_{\mathbf{x}}$ is simple is asymptotically equal to $e^{-(1+o(1))\rho(\rho+1)}$ where $\rho = m_2/m$ and $m_2 = \sum_{j\in[n]} d_{\mathbf{x}}(j)(d_{\mathbf{x}}(j) - 1)$. One consequence of the exponential tails in Lemma 3.3 is that $m_2 = O(m)$. This implies that $\rho = O(1)$ and hence that (8) holds. We can thus use the Random Sequence Model to prove the occurrence of high probability events in $G^{\delta\geq 3}_{n,m}$.

With this in hand, we can now proceed to describe our 2-matching algorithm.

4 Greedy Algorithm

Our algorithm will be applied to the random graph $G = G^{\delta\geq 3}_{n,m}$ and analysed in the context of $G_{\mathbf{x}}$. As the algorithm progresses, it makes changes to G and we let Γ denote the current state of G. The algorithm grows a 2-matching M and for $v \in [n]$ we let $b(v)$ be the 0/1 indicator for vertex v being incident to an edge of M. We let

- μ be the number of edges in Γ,
- $Y_k = \{v \in [n] : d_\Gamma(v) = k \text{ and } b(v) = 0\}$, $k = 0, 1, 2$,
- $Z_k = \{v \in [n] : d_\Gamma(v) = k \text{ and } b(v) = 1\}$, $k = 0, 1$.
- $Y = \{v \in [n] : d_\Gamma(v) \geq 3 \text{ and } b(v) = 0\}$, This is J_3 of Section 3.
- $Z = \{v \in [n] : d_\Gamma(v) \geq 2 \text{ and } b(v) = 1\}$, This is J_2 of Section 3.
- M is the set of edges in the current 2-matching.
- M^* is the matching induced by the path components of M i.e. if $P \subseteq M$ is a path from x to y then (x, y) will be an edge of M^* and the internal edges of P will have been deleted from Γ.

Observe that the sequence $\mathbf{b} = (b(v))$ is determined by $Y_0, Y_1, Y_2, Z_0, Z_1, Y, Z$. Note that Y_0, Z_0 play no active role inn the algorithm. They are the vertices that have been removed from Γ.

If $Y_1 \neq \emptyset$ then we choose $v \in Y_1$ and add the edge incident to v to M, because doing so is not a mistake i.e. there is a maximum size 2-matching of Γ that contains this edge. If $Y_1 = \emptyset$ and $Y_2 \neq \emptyset$ then we choose $v \in Y_2$ and add one of the two edges incident to v to M, because doing so is also not a mistake i.e. there is a maximum size 2-matching of Γ that contains this edge. We could immediately take both edges, but curiously enough the differential equations are simpler when we only take one. The other edge will eventually be picked up by the next case. We move v into Z_1. Similarly, if $Y_2 = \emptyset$ and $Z_1 \neq \emptyset$ we choose $v \in Z_1$ and add the unique edge of Γ incident to v to M. When we add an edge to M it can cause vertices of Γ to become internal vertices of paths of M and be deleted from Γ. In particular, this happens to the neighbor of $v \in Z_1$ in the case just described. When $Y_1 = Y_2 = Z_1 = \emptyset \neq Y$ we choose a random edge incident to a vertex of Y. In this way we hope to end up in a situation where $Y_2 = Z_1 = Y = \emptyset$ and $|Z| = \Omega(n)$. This has advantages that will be explained later in Section 9 and we have only managed to prove that this happens w.h.p. when $c \geq 10$. When $Y_1 = Y_2 = Z_1 = Y = \emptyset$ we are looking for a maximum matching in the graph Γ that remains and we can use the results of [11].

We now give details of the steps of

Algorithm 2GREEDY**:**

Step 1(a) $Y_1 \neq \emptyset$

Choose a random vertex v from Y_1. Suppose that its neighbor in Γ is w. We add (v, w) to M and move v to Z_0.

(i) If $b(w) = 0$ then we add (v, w) to M^*. If w is currently in Y then move it to Z. If it is currently in Y_1 then move it to Z_0. If it is currently in Y_2 then move it to Z_1. Call this re-assigning w.

(ii) If $b(w) = 1$ let u be the other end point of the path P of M that contains w. We remove (w, u) from M^* and replace it with (v, u). We move w to Z_0 and make the requisite changes due to the loss of other edges incident with w. Call this *tidying up*.

Step 1(b): $Y_1 = \emptyset$ **and** $Y_2 \neq \emptyset$

Choose a random vertex v from Y_2. Suppose that its neighbors in Γ are w_1, w_2.

If $w_1 = w_2 = v$ then we simply delete v from Γ. (We are dealing with loops because we are analysing the algorithm within the context of $G_{\mathbf{x}}$. This case is of course unnecessary when the input is simple i.e. for $G_{n,m}^{\delta \geq k}$).

Continuing with the most likely case, we choose one of the neighbors at random, say w_1. We move v to Z_1. We delete the edge (v, w_1) from Γ and place it into M. In addition,

(i) If $b(w_1) = 0$ then put $b(w_1) = 1$ and add the edge (v, w_1) to M^*. Re-assign w_1.

(ii) If $b(w_1) = 1$ let u_1 be the other end point of the paths P_1 of M that contains w_1 respectively. Adding the edge (v, w_1) creates a path (u_1, P_1', w_1, v) to M, where P_1' is the reversal of P_1. We delete the edge (w_1, u_1) from M^* and add (u_1, v) in its place. Vertex w_1 is deleted from Γ. Tidy up.

Step 1(c): $Y_2 = \emptyset$ **and** $Z_1 \neq \emptyset$

Choose a random vertex v from Z_1. Let u be the other endpoint of the path P of M that contains v. Let w be the unique neighbor of v in Γ. We delete v from Γ and add the edge (v,w) to M. In addition there are two cases.

(1) If $b(w)=0$ then we delete (v,u) from M^* and replace it with (w,u) and put $b(w)=1$ and re-assign w.

(2) If $b(w)=1$ then let u' be the other end-point of the path containing w in M. If $u'\neq u$ then we delete vertex w and the edge (u',w) from M^* and replace it with (u,u'). Tidy up. If $u'=u$ then we have created a cycle C and we delete it from Γ.

Step 2: $Y_1=Y_2=Z_1=\emptyset$ **and** $Y\neq\emptyset$
Choose a random edge (v,w) incident with a vertex $v\in Y$. We delete the edge (v,w) from Γ and add it to M. We put $b(v)=1$ and move it from Y to Z. There are two cases.

(i) If $b(w)=0$ then put $b(w)=1$ and move it from Y to Z. We add the edge (v,w) to M^*.

(ii) If $b(w)=1$ let u be the other end point of the path in M containing w. We delete vertex w and the edge (u,w) from M^* and replace it with (u,v). Tidy up.

Step 3: $Y_1=Y_2=Z_1=Y=\emptyset$
At this point Γ will be seen to be distributed as $G_{\nu,\mu}^{\delta\geq 2}$ for some ν,μ where $\mu=O(\nu)$. As such, it contains a (near) perfect matching M^{**} [15] and it can be found in $O(\nu)$ expected time [11].

The output of 2GREEDY is set of edges in $M\cup M^{**}$.

No explicit mention has been made of vertices contributing to Y_0. When we we tidy up after removing a vertex w, any vertex whose sole neighbor is w will be placed in Y_0.

5 Uniformity

In the previous section, we described the action of the algorithm as applied to Γ. In order to prove a uniformity property, it is as well to consider the changes induced by the algorithm in terms of $\mathbf{x}$.

When an edge is removed we will replace it in $\mathbf{x}$ by a pair of $\star$'s. This goes for all of the edges removed at an iteration, not just the edges of the 2-matching M. Thus at the end of this and subsequent iterations we will have a sequence in $\Lambda=([n]\cup\{\star\})^{2m}$ where for all i, $x_{2i-1}=\star$ if and only if $x_{2i}=\star$. We call such sequences *proper*.

We use the same notation as in Section 3. Let $S=S(\mathbf{x})=\{i: x_{2i-1}=x_{2i}=\star\}$. Note that the number of edges μ in $G_{\mathbf{x}}$ is given by

$$\mu=m-|S|.$$

For a tuple $\mathbf{v}=(Y_0,Y_1,Y_2,Z_0,Z_1,Y,Z,S)$ we let $\Lambda_{\mathbf{v}}$ denote the set of pairs $(\mathbf{x},\mathbf{b})$ where $\mathbf{x}\in\Lambda$ is proper and

- $Y_k=\{v\in[n]: d_{\mathbf{x}}(v)=k \text{ and } b(v)=0\}$, $k=0,1,2$,
- $Z_k=\{v\in[n]: d_{\mathbf{x}}(v)=k \text{ and } b(v)=1\}$, $k=0,1$.

- $Y = \{v \in [n] : d_{\mathbf{x}}(v) \geq 3 \text{ and } b(v) = 0\}$,
- $Z = \{v \in [n] : d_{\mathbf{x}}(v) \geq 2 \text{ and } b(v) = 1\}$.
- $S = S(\mathbf{x})$.

(Recall that $\mathbf{b}$ is determined by $\mathbf{v}$).

For vectors $\mathbf{x}, \mathbf{b}$ we define $\mathbf{v}(\mathbf{x}, \mathbf{b})$ by $(\mathbf{x}, \mathbf{b}) \in \Lambda_{\mathbf{v}(\mathbf{x},\mathbf{b})}$. We also use the notation $\mathbf{x} \in \Lambda_{\mathbf{v}(\mathbf{x})}$ when the second component $\mathbf{b}$ is assumed.

Given two sequences $\mathbf{x}, \mathbf{x}' \in \Lambda$, we say that $\mathbf{x}' \subseteq \mathbf{x}$ if $x_j = \star$ implies $x'_j = \star$. In which case we define $\mathbf{y} = \mathbf{x} - \mathbf{x}'$ by

$$y_j = \begin{cases} x_j & \text{If } x_j \neq \star = x'_j \\ \star & \text{Otherwise} \end{cases}$$

Thus $\mathbf{y}$ records the changes in going from $\mathbf{x}$ to $\mathbf{x}'$.

Given two sequences $\mathbf{x}, \mathbf{x}' \in \Lambda$ we say that $\mathbf{x}, \mathbf{x}'$ are *disjoint* if $x_j \neq \star$ implies that $x'_j = \star$. In which case we define $\mathbf{y} = \mathbf{x} + \mathbf{x}'$ by

$$y_j = \begin{cases} x_j & \text{If } x_j \neq \star \\ x'_j & \text{If } x'_j \neq \star \\ \star & \text{Otherwise} \end{cases}$$

Thus,

$$\text{if } \mathbf{x}' \subseteq \mathbf{x} \text{ then } \mathbf{x}' \text{ and } \mathbf{x} - \mathbf{x}' \text{ are disjoint and } \mathbf{x} = \mathbf{x}' + (\mathbf{x} - \mathbf{x}'). \tag{9}$$

Suppose now that $(\mathbf{x}(0), \mathbf{b}(0)), (\mathbf{x}(1), \mathbf{b}(1)), \ldots, (\mathbf{x}(t), \mathbf{b}(t))$ is the sequence of pairs representing the graphs constructed by the algorithm 2GREEDY. Here $\mathbf{x}(i-1) \supseteq \mathbf{x}(i)$ for $i \geq 1$ and so we can define $\mathbf{y}(i) = \mathbf{x}(i-1) - \mathbf{x}(i)$. Suppose that $\mathbf{v}(i) = \mathbf{v}(\mathbf{x}(i))$ for $1 \leq i \leq t$ where $\mathbf{v}(0) = (\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, [n], \emptyset, \emptyset)$ and $\mathbf{b}(0) = 0$.

Let

$$\Lambda_{\mathbf{v}|\mathbf{b}} = \{\mathbf{x} : (\mathbf{x}, \mathbf{b}) \in \Lambda_{\mathbf{v}}\}.$$

Lemma 5.1. *Suppose that $\mathbf{x}(0)$ is a random member of $\Lambda_{\mathbf{v}(0)|\mathbf{b}(0)}$. Then given $\mathbf{v}(0), \mathbf{v}(1), \ldots, \mathbf{v}(t)$, the vector $\mathbf{x}(t)$ is a random member of $\Lambda_{\mathbf{v}(t)|\mathbf{b}(t)}$ for all $t \geq 0$, that is, the distribution of $\mathbf{x}(t)$ is uniform, conditional on the edges deleted in the first t steps. (Note that $\mathbf{b}(t)$ is fixed by $\mathbf{v}(t)$ here).*

Proof We prove this by induction on t. It is trivially true for $t = 0$. Fix $t \geq 0, \mathbf{x}(t), \mathbf{b}(t), \mathbf{x}(t+1), \mathbf{b}(t+1)$. We define a sequence $\mathbf{x}(t) = \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_s = \mathbf{x}(t+1)$ where $\mathbf{z}_{i+1}$ is obtained from $\mathbf{z}_i$ by a *basic step*

Basic Step: Given $\mathbf{x}, \mathbf{b}$ and $\mathbf{v} = \mathbf{v}(\mathbf{x}, \mathbf{b})$ we create new sequences $\mathbf{x}' = A_j(\mathbf{x}), \mathbf{b}' = B_j(\mathbf{b})$ and $\mathbf{v}' = \mathbf{v}(\mathbf{x}', \mathbf{b}')$. Let $\mathbf{w} = \mathbf{x} - \mathbf{x}'$. A basic step corresponds to replacing the edge (w_{2j-1}, w_{2j}) by an edge of the matching M, for some index j. Let $u = w_{2j-1}, v = w_{2j}$.

Case 1: Here we assume $b(u) = b(v) = 0$.

Replace x_{2j-1}, x_{2j} by $\star$'s and put $b(u) = b(v) = 1$.

Case 2: Here we assume $b(u) = 0$, $b(v) = 1$.

Replace x_{2k-1}, x_{2k} by $\star$'s for every k such that $v \in \{x_{2k-1}, x_{2k}\}$ and put $b(u) = 1$.

Case 3: Here we assume $b(u) = b(v) = 1$.

Replace w_{2k-1}, w_{2k} by $\star$'s for every k such that $\{u, v\} \cap \{w_{2k-1}, w_{2k}\} \neq \emptyset$.

Claim 2.1. *Suppose that* $\mathbf{x}' = A_j(\mathbf{x})$ *and* $\mathbf{y} = \mathbf{x} - \mathbf{x}'$ *and* $\mathbf{b}' = B_j(\mathbf{b})$. *Then the map* $\phi : \mathbf{z} \in \Lambda^{\mathbf{y}}_{\mathbf{v}(\mathbf{x},\mathbf{b})} \longmapsto (\mathbf{z} - \mathbf{y}, \mathbf{b}')$ *is 1-1 and each* $(\mathbf{z}', \mathbf{b}') \in \Lambda_{\mathbf{v}(\mathbf{x}',\mathbf{b}')}$ *is the image under* ϕ *of a unique member of* $\Lambda^{\mathbf{y}}_{\mathbf{v}(\mathbf{x},\mathbf{b})}$*, where* $\Lambda^{\mathbf{y}}_{\mathbf{v}(\mathbf{x},\mathbf{b})} = \left\{(\mathbf{z}, \mathbf{b}) \in \Lambda_{\mathbf{v}(\mathbf{x},\mathbf{b})} : \mathbf{z} \supseteq \mathbf{y}\right\}$.

Proof of Claim 2.1. Equation (9) implies that ϕ is 1-1 i.e. $\mathbf{z}^* - \mathbf{y} = \mathbf{z} - \mathbf{y}$ implies $\mathbf{z}^* = \mathbf{z}$. Let $\mathbf{v} = \mathbf{v}(\mathbf{x}, \mathbf{b})$ and $\mathbf{v}' = \mathbf{v}(\mathbf{x}', \mathbf{b}')$. Choose $(\mathbf{w}, \mathbf{b}') \in \Lambda_{\mathbf{v}'}$. Because S' is determined by $\mathbf{v}'$, we see that $\mathbf{y}$ and $\mathbf{w}$ are necessarily disjoint and we simply have to check that if $\mathbf{x}^* = \mathbf{w} + \mathbf{y}$ then $(\mathbf{x}^*, \mathbf{b}) \in \Lambda_{\mathbf{v}}$. But in all cases, $\mathbf{v}(\mathbf{x}^*, \mathbf{b})$ is determined by $\mathbf{v}'$ and $\mathbf{y}$ and this implies that $\mathbf{v}(\mathbf{x}^*, \mathbf{b}) = \mathbf{v}(\mathbf{x}, \mathbf{b})$.

This statement is the crux of the proof and we should perhaps justify it a little more. Suppose then that we are given $\mathbf{v}'$ (and hence $\mathbf{b}'$) and $\mathbf{y}$ and $\mathbf{b}$. Observe that this determines $d_{\mathbf{x}^*}(v)$ for all $v \in Y_0' \cup Y_1' \cup Y_2' \cup Z_0' \cup Z_1'$. Together with $b(v)$ this determines the place of v in the partition defined by $\mathbf{v}$. Now $Y' \subseteq Y$ and it only remains to deal with $v \in Z'$. If $d_{\mathbf{y}}(v) > 0$ then $v \in Y \cup Z$ and $b(v)$ determines which of the sets v is in. If $d_{\mathbf{y}}(v) = 0$ and $b(v) = 1$ then $v \in Z$. If $d_{\mathbf{y}}(v) = 0$ and $b(v) = 0$ then $v \in Y$. This is because $b(v) = 0$ and $b'(v) = 1$ implies that we have put one of the edges incident with v into M.

End of proof of Claim 2.1

The claim implies (inductively) that if $\mathbf{x}$ is a uniform random member of $\Lambda_{\mathbf{v}|\mathbf{b}}$ and we do a sequence of basic steps involving the "deletion" of $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_s$ where $\mathbf{y}_{i+1} \subseteq \mathbf{x} - \mathbf{y}_1 - \cdots \mathbf{y}_i$, then $\mathbf{x}' = \mathbf{x} - \mathbf{y}_1 - \cdots - \mathbf{y}_s$ is a uniform random member of $\Lambda_{\mathbf{v}'|\mathbf{b}'}$, where $\mathbf{v}' = \mathbf{v}(\mathbf{x}', \mathbf{b}')$ for some $\mathbf{b}'$. This will imply Lemma 5.1 once we check that a step of 2GREEDY can be broken into basic steps.

First consider Step 1(a). First we choose a vertex in $x \in Y_1$. Then we apply Case 1 or 2 with probabilities determined by $\mathbf{v}$.

Now consider Step 1(b). First we choose a vertex in $x \in Y_2$. We can then replace one of the edges incident with x by a matching edge. We apply Case 1 or Case 2 with probabilities determined by $\mathbf{v}$.

For Step 1(c) we apply one of Case 2 or Case 3 with probabilities determined by $\mathbf{v}$.

For Step 2, we apply one of Case 1 or Case 2 with probabilities determined by $\mathbf{v}$.

This completes the proof of Lemma 5.1. □

As a consequence

Lemma 5.2. *The random sequence* $(\mathbf{v}(t),\ t = 0, 1, 2, \ldots,)$ *is a Markov chain.*

Proof Slightly abusing notation,

$$\begin{aligned}
&\mathbb{P}(\mathbf{v}(t+1) \mid \mathbf{v}(0), \ldots, \mathbf{v}(t)) \\
&= \sum_{\mathbf{w}' \in \Lambda_{\mathbf{v}(t+1)}} \mathbb{P}(\mathbf{w}' \mid \mathbf{v}(0), \ldots, \mathbf{v}(t))
\end{aligned}$$

$$= \sum_{\mathbf{w}'\in\Lambda_{\mathbf{v}(t+1)}} \sum_{\mathbf{w}\in\Lambda_{\mathbf{v}(t)}} \mathbb{P}(\mathbf{w}', \mathbf{w} \mid \mathbf{v}(0), \ldots, \mathbf{v}(t))$$
$$= \sum_{\mathbf{x}'\in\Lambda_{\mathbf{v}(t+1)}} \sum_{\mathbf{w}\in\Lambda_{\mathbf{v}(t)}} \mathbb{P}(\mathbf{w}' \mid \mathbf{v}(0), \ldots, \mathbf{v}(t-1), \mathbf{w})\mathbb{P}(\mathbf{w} \mid \mathbf{v}(0), \ldots, \mathbf{v}(t))$$
$$= \sum_{\mathbf{w}'\in\Lambda_{\mathbf{v}(t+1)}} \sum_{\mathbf{w}\in\Lambda_{\mathbf{v}(t)}} \mathbb{P}(\mathbf{w}' \mid \mathbf{w})|\Lambda_{\mathbf{v}(t)}|^{-1}, \qquad \textit{using Lemma } 5.1.$$

which depends only on $\mathbf{v}(t), \mathbf{v}(t+1)$. □

We now let

$$|\mathbf{v}| = \{|V_{0,0}|, |V_{0,1}|, |Y_1|, |Y_2|, |Z_1|, |Y|, |Z|, |S|\}.$$

Then we let $\Lambda_{|\mathbf{v}|}$ denote the set of $(\mathbf{x}, \mathbf{b}) \in \Lambda$ with $|\mathbf{v}(\mathbf{x}, \mathbf{b})| = |\mathbf{v}|$ and we let $\Lambda_{|\mathbf{v}|\,|\mathbf{b}} = \left\{\mathbf{x} : (\mathbf{x}, \mathbf{b}) \in \Lambda_{|\mathbf{v}|}\right\}$.

It then follows from Lemma 5.2 that by symmetry,

Lemma 5.3. *The random sequence* $|\mathbf{v}(t)|$, $t = 0, 1, 2, \ldots,$ *is a Markov chain.*

A component of a graph is *trivial* if it consists of a single isolated vertex.

Lemma 5.4. Whp *the number of non-trivial components of the graph induced by* $M \cup M^{**}$ *is* $O(\log n)$.

Proof Lemma 3 of Frieze and Łuczak [14] proves that w.h.p. the union of two random (near) perfect matchings of $[n]$ has at most $3 \log n$ components. Lemma 5.1 implies that at the end of Phase 1, Γ is a copy of $G_{\nu,\mu}^{\delta\geq 2}$, independent of M^*. In which case the (near) perfect matching of Γ is independent of M^* and we can apply [14]. □

6 Conditional expected changes

We now set up a system of differential equations that closely describe the path taken by the parameters of Algorithm 2GREEDY, as applied to $G_{\mathbf{x}}$ where $\mathbf{x}$ is chosen randomly from $[n]^{2m}_{\emptyset,[n];0}$. We introduce the following notation: At some point in the algorithm, the state of Γ is described by $\mathbf{x} \in [n]^{2M}_{J_2,J_3;D}$, together with an indicator vector $\mathbf{b}$. We let $y_i = |\{v : d_{\mathbf{x}}(v) = i \text{ and } b(v) = 0\}|$ and let $z_i = |\{v : d_{\mathbf{x}}(v) = i \text{ and } b(v) = 1\}|$ for $i \geq 0$. We let $y = \sum_{i\geq 3} y_i$ and $z = \sum_{i\geq 2} z_i$ and let $2\mu = \sum_{i\geq 0} i(y_i + z_i)$ be the total degree. Thus in the notation of Section 4 we have $y_i = |Y_i|, i = 1, 2,$ $J_3 = Y, N_3 = y,$ $z_1 = |Z_1|,$ $J_2 = Z, N_2 = z, D = y_1 + 2y_2 + z_1, M = \mu$ and $N = y + z$. The definition of N is a small departure from the notation of Section 3. Then it follows from Lemma 3.3, that as long as $(y + z)\lambda = \Omega(\log^2 n)$, we have q.s.,

$$y_k \approx \frac{\lambda^k}{k! f_3(\lambda)} y, \ (k \geq 3); \quad z_k \approx \frac{\lambda^k}{k! f_2(\lambda)} z, \ (k \geq 2). \tag{10}$$

Here λ is the root of

$$y \frac{\lambda f_2(\lambda)}{f_3(\lambda)} + z \frac{\lambda f_1(\lambda)}{f_2(\lambda)} = 2\mu - y_1 - 2y_2 - z_1. \tag{11}$$

□

Notational Convention: There are a large number of parameters that change as 2GREEDY progresses. Our convention will be that if we write a parameter ξ then by default it means $\xi(t)$, the value of ξ after t steps of the algorithm. Thus the initial value of ξ will be $\xi(0)$. When ξ is evaluated at a different point, we make this explicit.

We now keep track of the expected changes in $\mathbf{v} = (y_1, y_2, y, z_1, z_2, \mu)$ due to one step of 2GREEDY. These expectations are conditional on the current values of $\mathbf{b}$ and the degree sequence $\mathbf{d}$. In the following sequence of equations, $\xi' = \xi(t+1)$ represents the value of parameter ξ after the corresponding step of 2GREEDY.

Lemma 6.1. *The following are the expected one step changes in the parameters $(y_1, y_2, y, z_1, z, \mu)$. We will compute them conditional on the degree sequence $\mathbf{d}$ and on $|\mathbf{v}|$. We give both, because the first are more transparent and the second are what is needed. The error terms $\varepsilon_?$ are the consequence of multi-edges and we will argue that they are small.*

Step 1. $y_1 + y_2 + z_1 > 0$.

Step 1(a). $y_1 > 0$.

$$\mathbb{E}[y_1' - y_1 \mid \mathbf{b}, \mathbf{d}] = -1 - \left(\frac{y_1}{2\mu} + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{y_1}{2\mu} \right) + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{2y_2}{2\mu} + \varepsilon_{12}. \quad (12)$$

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = -1 - \frac{y_1}{2\mu} - \frac{y_1 z}{4\mu^2} \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{y_2 z}{2\mu^2} \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N} \right) \quad (13)$$

$$\mathbb{E}[y_2' - y_2 \mid \mathbf{b}, \mathbf{d}] = -\left(\frac{2y_2}{2\mu} + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{2y_2}{2\mu} \right) + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{3y_3}{2\mu} + \varepsilon_{14}. \quad (14)$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = -\frac{y_2}{\mu} - \frac{y_2 z}{2\mu^2} \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{yz}{8\mu^2} \frac{\lambda^3}{f_3(\lambda)} \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N} \right) \quad (15)$$

$$\mathbb{E}[z_1' - z_1 \mid \mathbf{b}, \mathbf{d}] = -\left(\frac{z_1}{2\mu} + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{z_1}{2\mu} \right) + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{2z_2}{2\mu} + \varepsilon_{16}. \quad (16)$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = -\frac{z_1}{2\mu} - \frac{z_1 z}{4\mu^2} \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{z^2}{4\mu^2} \frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N} \right). \quad (17)$$

$$\mathbb{E}[y' - y \mid \mathbf{b}, \mathbf{d}] = -\left(\sum_{k \geq 3} \frac{k y_k}{2\mu} + \sum_{k \geq 2} \frac{k z_k}{2\mu} (k-1) \frac{3y_3}{2\mu} \right) + \varepsilon_{18}. \quad (18)$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -\frac{y}{2\mu} \frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{yz}{8\mu^2} \frac{\lambda^3}{f_3(\lambda)} \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N} \right). \quad (19)$$

$$\mathbb{E}[z' - z \mid \mathbf{b},\mathbf{d}] = \sum_{k\geq 3}\frac{ky_k}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \varepsilon_{20}. \tag{20}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = \frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{z}{2\mu}\frac{\lambda f_1(\lambda)}{f_2(\lambda)} - \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{21}$$

$$\mathbb{E}[\mu' - \mu \mid \mathbf{b},\mathbf{d}] = -1 - \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1) + \varepsilon_{22}. \tag{22}$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - \frac{z}{2\mu}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{23}$$

Step 1(b). $y_1 = 0, y_2 > 0$.

$$\mathbb{E}[y_1' - y_1 \mid \mathbf{b},\mathbf{d}] = \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{2y_2}{2\mu} + \varepsilon_{24}. \tag{24}$$

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = \frac{y_2 z}{2\mu^2}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{25}$$

$$\mathbb{E}[y_2' - y_2 \mid \mathbf{b},\mathbf{d}] = -1 - \frac{2y_2}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{2y_2}{2\mu} + \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{3y_3}{2\mu} + \varepsilon_{26}. \tag{26}$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = -1 - \frac{y_2}{\mu} - \frac{y_2 z}{2\mu^2}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{yz}{8\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{27}$$

$$\mathbb{E}[z_1' - z_1 \mid \mathbf{b},\mathbf{d}] = 1 - \frac{z_1}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{z_1}{2\mu} + \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \varepsilon_{28}. \tag{28}$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = 1 - \frac{z_1}{2\mu} - \frac{z_1 z}{4\mu^2}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{29}$$

$$\mathbb{E}[y' - y \mid \mathbf{b},\mathbf{d}] = -\sum_{k\geq 3}\frac{ky_k}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{3y_3}{2\mu} + \varepsilon_{30}. \tag{30}$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -\frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{yz}{8\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{31}$$

$$\mathbb{E}[z' - z \mid \mathbf{b},\mathbf{d}] = \sum_{k\geq 3}\frac{ky_k}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu} - \sum_{k\geq 2}\frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \varepsilon_{32}. \tag{32}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = \frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{z}{2\mu}\frac{\lambda f_1(\lambda)}{f_2(\lambda)} - \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{33}$$

$$\mathbb{E}[\mu' - \mu \mid \mathbf{b}, \mathbf{d}] = -1 - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1) + \varepsilon_{34}. \quad (34)$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - \frac{z}{2\mu}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \quad (35)$$

$$(36)$$

Step 1(c). $y_1 = y_2 = 0, z_1 > 0$.

$$\mathbb{E}[y_1' - y_1 \mid \mathbf{b}, \mathbf{d}] = O\left(\frac{1}{N}\right). \quad (37)$$

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = O\left(\frac{1}{N}\right). \quad (38)$$

$$\mathbb{E}[y_2' - y_2 \mid \mathbf{b}, \mathbf{d}] = \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{3y_3}{2\mu} + \varepsilon_{39}. \quad (39)$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = \frac{yz}{8\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \quad (40)$$

$$\mathbb{E}[z_1' - z_1 \mid \mathbf{b}, \mathbf{d}] = -1 - \frac{z_1}{2\mu} - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{z_1}{2\mu} + \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \varepsilon_{41}. \quad (41)$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = -1 - \frac{z_1}{2\mu} - \frac{z_1 z}{4\mu^2}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N}\right). \quad (42)$$

$$\mathbb{E}[y' - y \mid \mathbf{b}, \mathbf{d}] = -\sum_{k\geq 3} \frac{ky_k}{2\mu} - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{3y_3}{2\mu} + \varepsilon_{43}. \quad (43)$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -\frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{yz}{8\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \quad (44)$$

$$\mathbb{E}[z' - z \mid \mathbf{b}, \mathbf{d}] = \sum_{k\geq 3} \frac{ky_k}{2\mu} - \sum_{k\geq 2} \frac{kz_k}{2\mu} - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \varepsilon_{45}. \quad (45)$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = \frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{z}{2\mu}\frac{\lambda f_1(\lambda)}{f_2(\lambda)} - \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N}\right). \quad (46)$$

$$\mathbb{E}[\mu' - \mu \mid \mathbf{b}, \mathbf{d}] = -1 - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1) + \varepsilon_{47}. \quad (47)$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - \frac{z}{2\mu}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \quad (48)$$

Step 2. $y_1 = y_2 = z_1 = 0$.

$$\mathbb{E}[y_1' - y_1 \mid \mathbf{b}, \mathbf{d}] = O\left(\frac{1}{N}\right). \tag{49}$$

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = O\left(\frac{1}{N}\right). \tag{50}$$

$$\mathbb{E}[y_2' - y_2 \mid \mathbf{b}, \mathbf{d}] = \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{3y_3}{2\mu} + \varepsilon_{51}. \tag{51}$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = \frac{yz}{8\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{52}$$

$$\mathbb{E}[z_1' - z_1 \mid \mathbf{b}, \mathbf{d}] = \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \varepsilon_{53}. \tag{53}$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{54}$$

$$\mathbb{E}[y' - y \mid \mathbf{b}, \mathbf{d}] = -1 - \sum_{k\geq 3} \frac{ky_k}{2\mu} - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{3y_3}{2\mu} + \varepsilon_{55}. \tag{55}$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -1 - \frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} - \frac{yz}{8\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{56}$$

$$\mathbb{E}[z' - z \mid \mathbf{b}, \mathbf{d}] = 1 - \sum_{k\geq 2} \frac{kz_k}{2\mu} - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1)\frac{2z_2}{2\mu} + \sum_{k\geq 3} \frac{ky_k}{2\mu} + \varepsilon_{57}. \tag{57}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = 1 - \frac{z}{2\mu}\frac{\lambda f_1(\lambda)}{f_2(\lambda)} - \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} + \frac{y}{2\mu}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{58}$$

$$\mathbb{E}[\mu' - \mu \mid \mathbf{b}, \mathbf{d}] = -1 - \sum_{k\geq 2} \frac{kz_k}{2\mu}(k-1) + \varepsilon_{59}. \tag{59}$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - \frac{z}{2\mu}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{60}$$

Proof The verification of (12) – (59) is long but straightforward. We will verify (12) and (13) and add a few comments and hope that the reader is willing to accept or check the remainder by him/herself.

Suppose without loss of generality that $\mathbf{x}$ is such that $x_1 = v = 1 \in Y_1$. The remainder of $\mathbf{x}$ is a random permutation of $2m - 2\mu$ $\star$'s and $2\mu - 1$ values from $[n]$ where the number of times j occurs is $\mathbf{d}_{\mathbf{x}}(j)$ for $j \in [n]$. The term -1 accounts for the deletion of v from Γ. There is a probability $\frac{y_1}{2\mu-1} = \frac{y_1}{2\mu} + O\left(\frac{1}{\mu}\right)$ that $x_2 \in Y_1$ and this accounts for the second term in (12). Observe next that there is a probability $\frac{kz_k}{2\mu-1}$ that $x_2 \in Z_k, k \geq 2$. In which case another $k-1$ edges will be deleted. In expectation, the number of vertices in Y_1 lost by the deletion of one such edge is $\frac{y_1-1}{2\mu-3}$ and this accounts for the third term. On the other hand, each such edge has a $\frac{2y_2}{2\mu-3}$ probability of being incident with a vertex in Y_2. The deletion of such an edge will create a vertex in Y_1 and this explains the fourth term. We collect the errors from replacing μ by $\mu-1$ etc. into the last term. This gives a contribution of order $1/N$. The above analysis ignored the extra contributions due to multiple edges. We can bound this by

$$\eta_{12} = \sum_{k\geq 3} \frac{kz_k}{2\mu} \sum_{\ell \geq 3} \frac{\ell y_\ell}{2\mu-1} \binom{k-1}{\ell-1} \left(\frac{\ell}{2\mu-k}\right)^{\ell-2}. \tag{61}$$

To explain this, we assume $x_2 \in Z_k$, which is accounted for by the first sum over k. Now, to create a vertex in Y_1, the removal of x_2 must delete $\ell-1$ of the edges incident with some vertex y in Y_ℓ. The term $\frac{\ell y_\ell}{2\mu-1}$ is the probability that the first of the chosen $\ell-1$ edges is incident with $y \in Y_\ell$ and the factor $\left(\frac{\ell}{2\mu-k}\right)^{\ell-2}$ bounds the probability that the remaining $\ell-2$ edges are incident with y.

To go from conditioning on $\mathbf{b}$,$\mathbf{d}$ to conditioning on $|\mathbf{v}|$ we need to use the expected values of y_k, z_l etc., conditional on $\mathbf{v}$. For this we use (4) and (5).

We have, up to an error term $O\left(\frac{\log^2 N}{\lambda N}\right)$,

$$\mathbb{E}\left[\sum_{k\geq 3} ky_k \middle| |\mathbf{v}|\right] = \sum_{k\geq 3} k\, y\, \frac{\lambda^k}{k! f_3(\lambda)} = \frac{y\lambda}{f_3(\lambda)} \sum_{j\geq 2} \frac{\lambda^j}{j!} = y\, \frac{\lambda f_2(\lambda)}{f_3(\lambda)}, \tag{62}$$

$$\mathbb{E}\left[\sum_{k\geq 2} kz_k \middle| |\mathbf{v}|\right] = \sum_{k\geq 2} k\, z\, \frac{\lambda^k}{k! f_2(\lambda)} = \frac{z\lambda}{f_2(\lambda)} \sum_{j\geq 1} \frac{\lambda^j}{j!} = z\, \frac{\lambda f_1(\lambda)}{f_2(\lambda)}, \tag{63}$$

$$\mathbb{E}\left[\sum_{k\geq 3} k(k-1)y_k \middle| |\mathbf{v}|\right] = \sum_{k\geq 3} k(k-1)\, y\, \frac{\lambda^k}{k! f_3(\lambda)} = \frac{y\lambda^2}{f_3(\lambda)} \sum_{j\geq 1} \frac{\lambda^j}{j!} = y\, \frac{\lambda^2 f_1(\lambda)}{f_3(\lambda)}, \tag{64}$$

$$\mathbb{E}\left[\sum_{k\geq 2} k(k-1)z_k \middle| |\mathbf{v}|\right] = \sum_{k\geq 2} k(k-1)\, z\, \frac{\lambda^k}{k! f_2(\lambda)} = \frac{z\lambda^2}{f_2(\lambda)} \sum_{j\geq 0} \frac{\lambda^j}{j!} = z\, \frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)}. \tag{65}$$

In particular, using (65) in (12) we get (13). The other terms are obtained in a similar fashion. We remark that we need to use (5) when we deal with products $z_k y_\ell$, $k \geq 2$ and $\ell \geq 3$.

Since, $k, \ell \leq \log n$ in (61) we see, with the aid of (62) – (65) that $\mathbb{E}[\eta_{12} \mid \mathbf{v}] = O(1/N)$. This bound is true for all other $\varepsilon_?$. □

6.1 Negative drift for y_1, y_2, z_1

Algorithm 2GREEDY tries to keep y_1, y_2, z_1 small by its selection in Step 1. We now verify that there is a negative drift in

$$\zeta = \zeta(t) = y_1 + 2y_2 + z_1$$

in all cases of Step 1. This will enable us to show that w.h.p. ζ remains small throughout the execution of 2Greedy. Let

$$Q = Q(\mathbf{v}) = \frac{yz}{4\mu^2}\frac{\lambda^3}{f_3(\lambda)}\frac{\lambda^2 f_0(\lambda)}{f_2(\lambda)} + \frac{z^2}{4\mu^2}\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2}. \tag{66}$$

Then simple algebra gives us

$$\mathbb{E}[\zeta' - \zeta \mid |\mathbf{v}|] = -(1-Q) - \zeta\left(\frac{1}{2\mu} + \frac{z\lambda^2 f_0(\lambda)}{4\mu^2 f_2(\lambda)}\right) + O\left(\frac{\log^2 N}{\lambda N}\right) \qquad \textit{Case } 1(a) \tag{67}$$

$$\mathbb{E}[\zeta' - \zeta \mid |\mathbf{v}|] = -(1-Q) - \zeta\left(\frac{1}{2\mu} + \frac{z\lambda^2 f_0(\lambda)}{4\mu^2 f_2(\lambda)}\right) + O\left(\frac{\log^2 N}{\lambda N}\right) \qquad \textit{Case } 1(b) \tag{68}$$

$$\mathbb{E}[\zeta' - \zeta \mid |\mathbf{v}|] = -(1-Q) - \zeta\left(\frac{1}{2\mu} + \frac{z\lambda^2 f_0(\lambda)}{4\mu^2 f_2(\lambda)}\right) + O\left(\frac{\log^2 N}{\lambda N}\right) \qquad \textit{Case } 1(c) \tag{69}$$

We will show

Lemma 6.2. *[Pittel]*

$$\lambda > 0 \textit{ implies } Q < 1 \tag{70}$$

and

$$1 - Q = \begin{cases} 1 - O(\lambda^{-1}), & \lambda \to \infty, \\ \Omega(\lambda^2), & \lambda \to 0. \end{cases} \tag{71}$$

Proof Now, by (11), $Q < 1$ is equivalent to

$$yz\frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)} + z^2\frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} < \left(y\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + z\frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right)^2,$$

or, introducing $x = y/z$,

$$F(x,\lambda) := \frac{x\frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)} + \frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2}}{\left(x\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right)^2} < 1, \quad \forall\, \lambda > 0, x \geq 0. \tag{72}$$

In particular, $F(\infty, \lambda) = 0$. Now

$$F_x(x,\lambda) = \left(x\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right)^{-4} G(x,\lambda),$$

where

$$G(x,\lambda) = \frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)}\left(x\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right)^2 - 2\left(x\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right)\frac{\lambda f_2(\lambda)}{f_3(\lambda)}\left(x\frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)} + \frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2}\right). \tag{73}$$

Notice that

$$G(0,\lambda) = \lambda^6 f_0(\lambda) f_1(\lambda) f_2(\lambda)^{-3} f_3(\lambda)^{-1}\big(\lambda f_1(\lambda) - 2f_2(\lambda)\big) > 0,$$

as $\lambda f_1(\lambda) - 2f_2(\lambda) > 0$. Whence $F_x(0,\lambda) > 0$ and as a function of x, $F(x,\lambda)$ attains its maximum at the root of $G(x,\lambda) = 0$, which is

$$\bar{x} = \frac{f_3(\lambda)\big(\lambda f_1(\lambda) - 2f_2(\lambda)\big)}{\lambda f_2(\lambda)^2}. \tag{74}$$

Now, (73) implies that $\bar{x}$ satisfies

$$\bar{x}\frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)} + \frac{\lambda^4 f_0(\lambda)}{f_2(\lambda)^2} = \frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)}\left(\bar{x}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right) \times \frac{f_3(\lambda)}{2\lambda f_2(\lambda)} \tag{75}$$

and (74) implies that

$$\bar{x}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)} = \frac{2(\lambda f_1(\lambda) - f_2(\lambda))}{f_2(\lambda)}. \tag{76}$$

Substituting (75), (76) into (72), we see that

$$F(\bar{x},\lambda) = \frac{\frac{\lambda^5 f_0(\lambda)}{f_2(\lambda)f_3(\lambda)}\frac{f_3(\lambda)}{2\lambda f_2(\lambda)}}{\left(\bar{x}\frac{\lambda f_2(\lambda)}{f_3(\lambda)} + \frac{\lambda f_1(\lambda)}{f_2(\lambda)}\right)} = \frac{\lambda^4 f_0(\lambda)}{4f_2(\lambda)\big(\lambda f_1(\lambda) - f_2(\lambda)\big)}.$$

Thus,

$$1 - F(\bar{x},\lambda) = \frac{D(\lambda)}{4f_2(\lambda)\big(\lambda f_1(\lambda) - f_2(\lambda)\big)}, \tag{77}$$

where

$$\begin{aligned} D(\lambda) =& 4f_2(\lambda)\big(\lambda f_1(\lambda) - f_2(\lambda)\big) - \lambda^4 f_0(\lambda) \\ =& -4 - 4\lambda - (\lambda^4 + 4\lambda^2 - 8)e^{\lambda} + (4\lambda - 4)e^{2\lambda}. \end{aligned}$$

In particular,

$$1 - F(\bar{x},\lambda) = 1 - O(\lambda^{-1}), \quad \lambda \to \infty. \tag{78}$$

Expanding $e^{\lambda}$and $e^{2\lambda}$, we obtain after collecting like terms that

$$D(\lambda) = \sum_{j\geq 6} \frac{d_j}{j!}\,\lambda^j,$$

where

$$d_j = 2^{j+1}(j-2) - (j)_4 - 4(j)_2 + 8.$$

Here $d_j = 0$ for $0 \leq j \leq 5$ and $d_6 = 40, d_7 = 280, d_8 = 1176, d_9 = 3864, d_{10} = 10992$ and $d_j > 0$ for $j \geq 11$ is clear. Therefore $D(\lambda)$ is positive for all $\lambda > 0$. Since $D(\lambda) \sim d_6\lambda^6$ and $4f_2(\lambda)\big(\lambda f_1(\lambda) - f_2(\lambda)\big) \sim \lambda^4$ as $\lambda \to 0$, we see that

$$1 - F(\bar{x},\lambda) \sim d_6\lambda^2, \quad \lambda \to 0. \tag{79}$$

This completes the proof of Lemma 6.2. □

It follows from (67), (68), (69) and Lemma 6.2 that, regardless of case,

$$\zeta > 0 \text{ implies } \mathbb{E}[\zeta' - \zeta \mid |\mathbf{v}|] \leq -c_1(1\wedge\lambda)^2 + O\left(\frac{\log^2 N}{\lambda N}\right) \tag{80}$$

for some absolute constant $c_1 > 0$, where $1 \wedge \lambda = \min\{1, \lambda\}$.

To avoid dealing with the error term in (80) we introduce the stopping time,

$$T_{er} = \min\left\{t : \lambda^2 \leq \frac{\log^3 N}{\lambda N}\right\}.$$

(This is well defined, since eventually $N = 0$).

The following stopping time is also used:

$$T_0 = \min\{t : \lambda \leq 1 \text{ or } N \leq n/2\} < T_{er}.$$

So we can replace (80) by

$$\zeta > 0 \text{ implies } \mathbb{E}[\zeta' - \zeta \mid |\mathbf{v}|] \leq -c_1/2, \qquad 0 \leq t < T_0, \tag{81}$$

which holds for n sufficiently large.

There are several places where we need a bound on λ:

Lemma 6.3. Whp *$\lambda \leq 3ce$ for $t \leq T_0$.*

Proof We will show later that w.h.p. $y_1 + 2y_2 + z_1 = o(n)$ throughout. It follows from (11) and the inequalities in Section 8.0.2 ((204) in particular), that if $\lambda(t) \geq \Lambda$ then $Y \cup Z$ contains $y+z$ vertices and at least $\Lambda(y+z)/2$ edges and hence has total degree at least $\Lambda(y+z)$. We argue that w.h.p. G does not contain such a sub-graph, when $\Lambda = 3ce$. We will work in the random sequence model. We can assume that $|Y \cup Z| \geq n/2$. Now fix a set $S \subseteq [n]$ where $s = |S| \geq n/3$. Let D denote the total degree of vertices in S. Then

$$\mathbb{P}(D = d) \leq O(n^{1/2}) \sum_{\substack{d_1+\cdots+d_s=d\\ d_j \geq 3}} \prod_{i=1}^{s} \frac{\lambda^{d_i}}{f_3(\lambda) d_i!} \leq O(n^{1/2}) \frac{\lambda^d}{d! f_3(\lambda)^s} \sum_{\substack{d_1+\cdots+d_s=d\\ d_j \geq 0}} \frac{d!}{d_1! \cdots d_s!}$$
$$= O(n^{1/2}) \frac{\lambda^d s^d}{d! f_3(\lambda)^s}. \tag{82}$$

Here $\lambda = \lambda(0)$ and we are using Lemma 3.1. The factor $O(n^{1/2})$ accounts for the conditioning that the total degree is $2cn$. Now $\lambda(0) \leq 2c$ and $f_3(\lambda(0)) \geq 1$. It follows that

$$\mathbb{P}(\exists S : D = d, d \geq \Lambda s) \leq O(n^{1/2}) \sum_{s \geq n/3} \sum_{d \geq \Lambda s} \binom{n}{s} \frac{(2c)^d s^d}{d!} \leq O(n^{1/2}) \sum_{s \geq n/3} \sum_{d \geq \Lambda s} \left(\frac{ne}{s}\right)^s \frac{(2c)^d s^d}{d!}$$

The terms involving d in the second sum are $u_d = \frac{(2cs)^d}{d!}$ and for $d \geq \Lambda s$ large we have $u_{d+1}/u_d \leq 2/3$ and so we can put $d = \Lambda s$ in the second expression. After substituting $d! \geq (d/e)^d$ this gives

$$\mathbb{P}(\exists S : D = d, d \geq \Lambda s) \leq O(n^{1/2}) \sum_{s \geq n/2} \left(\frac{3e(2ce)^\Lambda}{\Lambda^\Lambda}\right)^s = o(1)$$

if $\Lambda \geq 3ce$. □

Our aim now is to give a high probability bound on the maximum value that ζ will take during the process. We first prove a simple lemma involving the functions $\phi_j(x) = \frac{x f_{j-1}(x)}{f_j(x)}$, $j = 2, 3$.

Lemma 6.4.

$$\phi_j(x) \text{ is convex and increasing and } j \le \phi_j(x) \text{ and } \frac{1}{j+1} \le \phi_j'(x) \le 1 \text{ for } j = 2,3. \tag{83}$$

Proof Now, if $H(x) = \frac{xF(x)}{G(x)}$ then

$$H'(x) = \frac{G(x)(xF'(x) + F(x)) - xF(x)G'(x)}{G(x)^2}$$

and

$$H''(x) = \frac{2xF(x)G'(x)^2 + G(x)^2(2F'(x) + xF''(x)) - G(x)(2xF'(x)G'(x) + F(x)(2G'(x) + xG''(x)))}{G(x)^3}.$$

Case $j = 2$**:**

$$\phi_2'(x) = \frac{e^{2x} - (x^2+2)e^x + 1}{(e^x - 1 - x)^2}. \tag{84}$$

But,

$$e^{2x} - (x^2+2)e^x + 1 = \sum_{j\ge 4} \frac{2^j - j(j-1) - 2}{j!} x^j$$

and so $\phi_2'(x) > 0$ for $x > 0$.

$$\phi_2''(x) = \frac{e^{2x}(x^2 - 4x + 2) + e^x(x^3 + x^2 + 4x - 4) + 2}{(e^x - 1 - x)^3}. \tag{85}$$

But

$$e^{2x}(x^2 - 4x + 2) + e^x(x^3 + x^2 + 4x - 4) + 2 = \sum_{j\ge 6} \frac{2^{j-2}(j(j-1) - 8j + 8) + j(j-1)(j-2) + j(j-1) + 4j - 4}{j!} x^j$$

and so $\phi_2''(x) > 0$ for $x > 0$.

Case $j = 3$**:**

$$\phi_3'(x) = \frac{2e^{2x} - e^x\left(x^3 - x^2 + 4x + 4\right) + x^2 + 4x + 2}{2(e^x - 1 - x - \frac{x^2}{2})^2}. \tag{86}$$

But,

$$2e^{2x} - e^x\left(x^3 - x^2 + 4x + 4\right) + x^2 + 4x + 2 = \sum_{j\ge 6} \frac{2^{j+1} - j(j-1)(j-2) + j(j-1) - 4j - 4}{j!} x^j$$

and so $\phi_3'(x) > 0$ for $x > 0$.

$$\phi_3''(x) = \frac{x(e^{2x}(2x^2 - 12x + 12) + e^x(x^4 + 8x^2 - 24) + 2x^2 + 12x + 12)}{4(e^x - 1 - x - \frac{x^2}{2})^3}. \tag{87}$$

But

$$e^{2x}(2x^2-12x+12)+e^x(x^4+8x^2-24)+2x^2+12x+12$$
$$\sum_{j\geq 9}\frac{2^{j-1}(j(j-1)-12j+24)+j(j-1)(j-2)(j-3)+8j(j-1)-24}{j!}x^j.$$

and so $\phi_3''(x)>0$ for $x>0$.

So ϕ_2,ϕ_3 are convex and so we only need to check that $\phi_2(0)=2,\phi_2'(0)=1/3,\phi_3(0)=3,\phi_3'(0)=1/4$ and $\phi_2'(\infty)=\phi_3'(\infty)=1$. □

Consider λ as a function of $\mathbf{v}$, defined by

$$y\phi_3(\lambda)+z\phi_2(\lambda)=\Pi \tag{88}$$

where $\Pi=2\mu-y_1-2y_2-z_1$.

We now prove a lemma bounding the change in λ as we change $\mathbf{v}$.

Lemma 6.5. *Let $\lambda_i=\lambda(\mathbf{v}_i)$ and $N_i=N(\mathbf{v}_i)$ for $i=1,2$.*

$$|\lambda_1-\lambda_2|\leq\frac{4\max\{\phi_3(\lambda_1),\phi_3(\lambda_2)\}}{\min\{N_1,N_2\}}||\mathbf{v}_1-\mathbf{v}_2||_1.$$

Proof We see in fact that $\lambda=\lambda(\mathbf{w})$ where $\mathbf{w}=(y,z,\Pi)$, We use $\mathbf{w}_1,\mathbf{w}_2$ to denote the corresponding vectors at $\mathbf{v}_1,\mathbf{v}_2$. Differentiating and using (83) we obtain

$$\frac{\partial\lambda}{\partial\Pi}=\frac{1}{y\phi_3'(\lambda)+z\phi_2'(\lambda)}\leq\frac{4}{y+z}$$
$$\frac{\partial\lambda}{\partial y}=-\frac{\phi_3(\lambda)}{y\phi_3'(\lambda)+z\phi_2'(\lambda)}\leq\frac{4\phi_3(\lambda)}{y+z}$$
$$\frac{\partial\lambda}{\partial z}=-\frac{\phi_2(\lambda)}{y\phi_3'(\lambda)+z\phi_2'(\lambda)}\leq\frac{4\phi_2(\lambda)}{y+z}$$

Now by the mean value theorem,

$$\lambda(\mathbf{v}_1)-\lambda(\mathbf{v}_2)=\left(\frac{\partial\lambda}{\partial\Pi},\frac{\partial\lambda}{\partial y},\frac{\partial\lambda}{\partial z}\right)\cdot(\mathbf{w}_1-\mathbf{w}_2)$$

where the gradient $\left(\frac{\partial\lambda}{\partial\Pi},\frac{\partial\lambda}{\partial y},\frac{\partial\lambda}{\partial z}\right)$ is evaluated at a point on the line segment between $\mathbf{w}_1$ and $\mathbf{w}_2$. To obtain the lemma we use $\phi_2\leq\phi_3$ and we use the convexity of ϕ_3 to bound ϕ_3 by its value at the endpoints. □

Lemma 6.6. *If $c\geq 10$ then q.s.*

$$\not\exists 1\leq t<T_0:\ \zeta(t)>\log^2 n.$$

Proof Define a sequence

$$X_i=\begin{cases}\min\{\zeta(i+1)-\zeta(i),\log n\} & 0\leq i<T_0\\ -c_1/2 & T_0\leq i\leq n\end{cases}$$

The variables $X_1, X_2, \ldots, X_n$ are not independent. On the other hand, conditional on an event that occurs q.s., we see that

$$X_s + \ldots + X_{t-1} = \zeta(t) - \zeta(s) \text{ for } 0 \le s < t \le T_0$$

and

$$\mathbb{E}[X_t \mid X_1, \ldots, X_{t-1}] \le -c_1/2 \text{ for } t \le n.$$

Next, for $0 \le s \le t \le T_0$ let

$$\bar{\lambda}(s,t) = \sum_{\tau=s}^{t-1} \lambda(\tau)^2.$$

Note that

$$\bar{\lambda}(s,t) \ge t - s. \tag{89}$$

We argue as in the proof of the Azuma-Hoeffding inequality that for any $1 \le s < t \le n$ and $u \ge 0$,

$$\mathbb{P}(X_s + \cdots + X_{t-1} \ge u - c_1\bar{\lambda}(s,t)/2) \le \exp\left\{-\frac{2u^2}{(t-s)\log^2 n}\right\}. \tag{90}$$

We deduce from this that for $1 \le s < t \le T_0$ we have

$$\mathbb{P}(\zeta(s) = 0 < \zeta(\tau), s < \tau \le t) \le \exp\left\{-\frac{2\max\left\{0, c_1\bar{\lambda}(s,t)/2 - \log n\right\}^2}{(t-s)\log^2 n}\right\}. \tag{91}$$

If $L_1 = \log^2 n$ then we see from (91) that q.s.

$$\not\exists 1 \le s < t - L_1 \le T_0 - L_1 : \zeta(s) = 0 < \zeta(\tau), s < \tau \le t. \tag{92}$$

Suppose now that there exists $\tau \le T_0$ such that $\zeta(\tau) \ge L_1$. Then q.s. there exists $t_1 \le \tau \le t_1 + L_1$ such that $\zeta(t_1) = 0$. But then given t_1,

$$\mathbb{P}(\exists t_1 \le \tau \le t_1 + L_1 : \zeta(\tau) \ge L_1) \le \exp\left\{-\frac{2(c_1L_1/2 - \log n)^2}{L_1 \log^2 n}\right\}.$$

Here we are using the generalisation of Hoeffding-Azuma that deals with $\max_{i \le L_1} X_1 + \cdots + X_i$. And then we get that q.s.

$$\not\exists t \le T_0 : \zeta(\tau) \ge L_1. \tag{93}$$

We do this in two stages because of the condition $\zeta > 0$ in (81). Remember here that $\zeta(0) = 0$ and (92) says that ζ cannot stay positive for very long. $\square$

7 Associated Equations.

The expected changes conditional on $\mathbf{v}$ lead us to consider the following collection of differential equations: Note that we do not use any scaling. We will put hats on variables i.e. $\hat{y}_1$ etc. will be the deterministic counterpart of y_1. Also, as expected, the hatted equivalent of (88) holds:

$$\frac{\hat{y}\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} + \frac{\hat{z}\lambda f_1(\hat{\lambda})}{f_2(\hat{\lambda})} = 2\hat{\mu} - \hat{y}_1 - 2\hat{y}_2 - \hat{z}_1. \tag{94}$$

Step 1(a). $\hat{y}_1 > 0$.

$$\frac{d\hat{y}_1}{dt} = -1 - \frac{\hat{y}_1}{2\hat{\mu}} - \frac{\hat{y}_1\hat{z}}{4\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + \frac{\hat{y}_2\hat{z}}{2\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{95}$$

$$\frac{d\hat{y}_2}{dt} = -\frac{\hat{y}_2}{\hat{\mu}} - \frac{\hat{y}_2\hat{z}}{2\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{96}$$

$$\frac{d\hat{z}_1}{dt} = -\frac{\hat{z}_1}{2\hat{\mu}} - \frac{\hat{z}_1\hat{z}}{4\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{97}$$

$$\frac{d\hat{y}}{dt} = -\frac{\hat{y}}{2\hat{\mu}}\frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{98}$$

$$\frac{dz}{dt} = \frac{\hat{y}}{2\hat{\mu}}\frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{z}}{2\hat{\mu}}\frac{\hat{\lambda} f_1(\hat{\lambda})}{f_2(\hat{\lambda})} - \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{99}$$

$$\frac{d\hat{\mu}}{dt} = -1 - \frac{\hat{z}}{2\hat{\mu}}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}. \tag{100}$$

Step 1(b). $\hat{y}_1 = 0, \hat{y}_2 > 0$.

$$\frac{d\hat{y}_1}{dt} = \frac{\hat{y}_2\hat{z}}{2\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{101}$$

$$\frac{d\hat{y}_2}{dt} = -1 - \frac{\hat{y}_2}{\hat{\mu}} - \frac{\hat{y}_2\hat{z}}{2\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{102}$$

$$\frac{d\hat{z}_1}{dt} = 1 - \frac{\hat{z}_1}{2\hat{\mu}} - \frac{\hat{z}_1\hat{z}}{4\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{103}$$

$$\frac{d\hat{y}}{dt} = -\frac{\hat{y}}{2\hat{\mu}}\frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{104}$$

$$\frac{d\hat{z}}{dt} = \frac{\hat{y}}{2\hat{\mu}}\frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{z}}{2\hat{\mu}}\frac{\hat{\lambda} f_1(\hat{\lambda})}{f_2(\hat{\lambda})} - \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{105}$$

$$\frac{d\hat{\mu}}{dt} = -1 - \frac{\hat{z}}{2\hat{\mu}}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}. \tag{106}$$

Step 1(c). $\hat{y}_1 = \hat{y}_2 = 0, \hat{z}_1 > 0$.

$$\frac{d\hat{y}_1}{dt} = 0, \tag{107}$$

$$\frac{d\hat{y}_2}{dt} = \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{108}$$

$$\frac{d\hat{z}_1}{dt} = -1 - \frac{\hat{z}_1}{2\hat{\mu}} - \frac{\hat{z}_1\hat{z}}{4\hat{\mu}^2}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{109}$$

$$\frac{d\hat{y}}{dt} = -\frac{\hat{y}}{2\hat{\mu}}\frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{110}$$

$$\frac{d\hat{z}}{dt} = \frac{\hat{y}}{2\hat{\mu}}\frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{z}}{2\hat{\mu}}\frac{\hat{\lambda} f_1(\hat{\lambda})}{f_2(\hat{\lambda})} - \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{111}$$

$$\frac{d\hat{\mu}}{dt} = -1 - \frac{\hat{z}}{2\hat{\mu}} \frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}. \tag{112}$$

Step 2. $\hat{y}_1 = \hat{y}_2 = \hat{z}_1 = 0$.

$$\frac{d\hat{y}_1}{dt} = 0, \tag{113}$$

$$\frac{d\hat{y}_2}{dt} = \frac{\hat{y}\hat{z}}{8\hat{\mu}^2} \frac{\hat{\lambda}^3}{f_3(\hat{\lambda})} \frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{114}$$

$$\frac{d\hat{z}_1}{dt} = \frac{\hat{z}^2}{4\hat{\mu}^2} \frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}, \tag{115}$$

$$\frac{d\hat{y}}{dt} = -1 - \frac{\hat{y}}{2\hat{\mu}} \frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} - \frac{\hat{y}\hat{z}}{8\hat{\mu}^2} \frac{\hat{\lambda}^3}{f_3(\hat{\lambda})} \frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}, \tag{116}$$

$$\frac{d\hat{z}}{dt} = 1 - \frac{\hat{z}}{2\hat{\mu}} \frac{\hat{\lambda} f_1(\hat{\lambda})}{f_2(\hat{\lambda})} - \frac{\hat{z}^2}{4\hat{\mu}^2} \frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2} + \frac{\hat{y}}{2\hat{\mu}} \frac{\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})}, \tag{117}$$

$$\frac{d\hat{\mu}}{dt} = -1 - \frac{\hat{z}}{2\hat{\mu}} \frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}. \tag{118}$$

We will show that w.h.p. the process defined by 2GREEDY can be closely modeled by a suitable weighted sum of the above four sets of equations. Let these weights be $\theta_a, \theta_b, \theta_c$ and $1 - \theta_a - \theta_b - \theta_c$ respectively. It has been determined that y_1, y_2, z_1 are all $O(\log^2 n)$ w.h.p.. We will only need to analyse our process up till the time $y = 0$ and we will show that at this time, $z = \Omega(n)$ w.h.p.. Thus y_1, y_2, z_1 are "negligible" throughout. In which case $\hat{y}_1, \hat{y}_2, \hat{z}_2$ should also be negligible. It makes sense therefore to choose $\theta_a = 0$. The remaining weights should be chosen so that the weighted derivatives of $\hat{y}_1, \hat{y}_2, \hat{z}_1$ are zero. This has all been somewhat heuristic and its validity will be verified in Section 7.2.

7.1 Sliding trajectory

Conjecturally we need to mix Steps 1(a), 1(b) 1(c) and 2 with nonnegative weights $\theta_a = 0$, θ_b, θ_c, $\theta_2 = 1 - \theta_b - \theta_c$ respectively, chosen such that the resulting system of differential equations admits a solution such that $\hat{y}_2(t) \equiv 0$ and $\hat{z}_1(t) \equiv 0$.

We will write the multipliers in terms of

$$\hat{A} = \frac{\hat{y}\hat{z}\hat{\lambda}^5 f_0(\hat{\lambda})}{8\hat{\mu}^2 f_2(\hat{\lambda}) f_3(\hat{\lambda})}, \quad \hat{B} = \frac{\hat{z}^2 \hat{\lambda}^4 f_0(\hat{\lambda})}{4\hat{\mu}^2 f_2(\hat{\lambda})^2}, \quad \hat{C} = \frac{\hat{y}\hat{\lambda} f_2(\hat{\lambda})}{2\hat{\mu} f_3(\hat{\lambda})}, \quad \hat{D} = \frac{\hat{z}\hat{\lambda}^2 f_0(\hat{\lambda})}{2\hat{\mu} f_2(\hat{\lambda})}. \tag{119}$$

Using, (95), (101), (107) and (113) we see $\hat{y}_1(t) \equiv 0$ implies that

$$0 \equiv \frac{d\hat{y}_1}{dt} = \theta_a.$$

Equivalently

$$\theta_a = 0. \tag{120}$$

Using (96), (108) and (114), we see that $\hat{y}_2(t) \equiv 0$ implies that

$$0 \equiv \frac{d\hat{y}_2}{dt}$$

$$= \theta_b \left[-1 + \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})}\right] + \theta_c \frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})} + (1-\theta_b-\theta_c)\frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^3}{f_3(\hat{\lambda})}\frac{\hat{\lambda}^2 f_0(\hat{\lambda})}{f_2(\hat{\lambda})},$$
$$= -\theta_b + \hat{A}. \tag{121}$$

Equivalently

$$\theta_b = \hat{A}. \tag{122}$$

Likewise, using (97), (109) and (115), $z_1(t) \equiv 0$ implies

$$0 \equiv \frac{d\hat{z}_1}{dt}$$
$$= \theta_b \left[1 + \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}\right] + \theta_c \left[-1 + \frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2}\right] + (1-\theta_b-\theta_c)\frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4 f_0(\hat{\lambda})}{f_2(\hat{\lambda})^2},$$
$$= \theta_b - \theta_c + \hat{B}. \tag{123}$$

Equivalently

$$\theta_c = \hat{A} + \hat{B}. \tag{124}$$

From (123) and (124) it follows that $1 - \theta_b - \theta_c \geq 0$ iff

$$2\hat{A} + \hat{B} \leq 1. \tag{125}$$

We conclude from (66) and Lemma 6.2 that $\theta_b, \theta_c, 1 - \theta_b - \theta_c \in [0, 1]$.

It may be of some use to picture the equations defining $\theta_a, \theta_b, \theta_c, \theta_2$:

$$\begin{array}{ccccl} \theta_a & & & & = 0 \\ & \theta_b & & & = \hat{A} \\ & -\theta_b & \theta_c & & = \hat{B} \\ \theta_a & +\theta_b & +\theta_c & +\theta_2 & = 1. \end{array} \tag{126}$$

If in the notation of Lemma 6.6 we let $\Omega_1 = \{\mathbf{v} : \zeta \leq L_1\}$ then we may restrict our attention to $\mathbf{v}$ in (12) – (59) such that $\mathbf{v} \in \Omega_1$. In which case, the terms involving y_1, y_2, z can be absorbed into the error term for $t \leq T_0$. The relevant equations then become, with

$$A = \frac{yz\lambda^5 f_0(\lambda)}{8\mu^2 f_2(\lambda) f_3(\lambda)}, \quad B = \frac{z^2\lambda^4 f_0(\lambda)}{4\mu^2 f_2(\lambda)^2}, \quad C = \frac{y\lambda f_2(\lambda)}{2\mu f_3(\lambda)}, \quad D = \frac{z\lambda^2 f_0(\lambda)}{2\mu f_2(\lambda)}.$$

Step 1(a). $y_1 > 0$.

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = -1 + O\left(\frac{\log^2 N}{\lambda N}\right) \tag{127}$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = A + O\left(\frac{\log^2 N}{\lambda N}\right) \tag{128}$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{129}$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -C - A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{130}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = C - (1-C) - B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{131}$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - D + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{132}$$

Step 1(b). $y_1 = 0, y_2 > 0$.

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = O\left(\frac{\log^2 N}{\lambda N}\right). \tag{133}$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = -1 + A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{134}$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = 1 + B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{135}$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -C - A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{136}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = C - (1-C) - B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{137}$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - D + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{138}$$

Step 1(c). $y_1 = y_2 = 0, z_1 > 0$.

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = O\left(\frac{1}{N}\right). \tag{139}$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{140}$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = -1 + B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{141}$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -C - A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{142}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = C - (1-C) - B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{143}$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - D + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{144}$$

Step 2. $y_1 = y_2 = z_1 = 0$.

$$\mathbb{E}[y_1' - y_1 \mid |\mathbf{v}|] = O\left(\frac{1}{N}\right). \tag{145}$$

$$\mathbb{E}[y_2' - y_2 \mid |\mathbf{v}|] = A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{146}$$

$$\mathbb{E}[z_1' - z_1 \mid |\mathbf{v}|] = B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{147}$$

$$\mathbb{E}[y' - y \mid |\mathbf{v}|] = -1 - C - A + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{148}$$

$$\mathbb{E}[z' - z \mid |\mathbf{v}|] = 1 + C - (1 - C) - B + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{149}$$

$$\mathbb{E}[\mu' - \mu \mid |\mathbf{v}|] = -1 - D + O\left(\frac{\log^2 N}{\lambda N}\right). \tag{150}$$

7.2 Closeness of the process and the differential equations

We already know that y_1, y_2, z_1 are small w.h.p. up to time T_0. We now show that w.h.p. y, z, μ are closely approximated by $\hat{y}, \hat{z}, \hat{\mu}$, which are the solutions to the weighted sum of the sets of equations labelled Step 1(b), Step 1(c) and Step 2. These equations will be simplified by putting $y_1 = y_2 = z_1 = 0$. First some notation. We will use $\psi_{\eta,\xi}$ to denote the expression we have obtained for the derivative of ξ in Case 1 (η) or Case 2 in the case of $\eta = 2$. We are then led to consider the equations:

Sliding Trajectory:

$$\begin{aligned}
\frac{d\hat{y}}{dt} &= \theta_b \psi_{b,y}(\hat{y}, \hat{z}, \hat{\mu}) + \theta_c \psi_{c,y}(\hat{y}, \hat{z}, \hat{\mu}) + (1 - \theta_b - \theta_c)\psi_{2,y}(\hat{y}, \hat{z}, \hat{\mu}) \\
&= \theta_b(-(\hat{C} + \hat{A})) + \theta_c(-(\hat{C} + \hat{A})) + (1 - \theta_b - \theta_c)(-(1 + \hat{C} + \hat{A})) \\
&= -(\hat{C} + \hat{A})(\theta_b + \theta_c + 1 - \theta_b - \theta_c) - (1 - \theta_b - \theta_c) \\
&= \hat{A} + \hat{B} - \hat{C} - 1. \\
\frac{d\hat{z}}{dt} &= \theta_b \psi_{b,z}(\hat{y}, \hat{z}, \hat{\mu}) + \theta_c \psi_{c,z}(\hat{y}, \hat{z}, \hat{\mu}) + (1 - \theta_b - \theta_c)\psi_{2,z}(\hat{y}, \hat{z}, \hat{\mu}) \\
&= \theta_b((\hat{C} - (1 - \hat{C}) - \hat{B})) + \theta_c(\hat{C} - (1 - \hat{C}) - \hat{B}) + (1 - \theta_b - \theta_c)(1 + \hat{C} - 1 + \hat{C} - \hat{B}) \\
&= 2\hat{C} - \hat{B} - 1 + 1 - \theta_b - \theta_c \\
&= 2\hat{C} - 2\hat{A} - 2\hat{B}. \\
\frac{d\hat{\mu}}{dt} &= \theta_b \psi_{b,\mu}(\hat{y}, \hat{z}, \hat{\mu}) + \theta_c \psi_{c,\mu}(\hat{y}, \hat{z}, \hat{\mu}) + (1 - \theta_b - \theta_c)\psi_{2,\mu}(\hat{y}, \hat{z}, \hat{\mu}) \\
&= \theta_b(-(1 + \hat{D})) + \theta_c(-(1 + \hat{D})) + (1 - \theta_b - \theta_c)(-(1 + \hat{D})) \\
&= -(1 + \hat{D})(\theta_b + \theta_c + 1 - \theta_b - \theta_c) \\
&= -1 - \hat{D}.
\end{aligned}$$

The initial conditions are

$$\hat{y}(0) = n, \hat{z}(0) = 0, \hat{\mu}(0) = cn. \tag{151}$$

Summarising:

$$\frac{d\hat{y}}{dt} = \hat{A} + \hat{B} - \hat{C} - 1; \quad \frac{d\hat{z}}{dt} = 2\hat{C} - 2\hat{A} - 2\hat{B}; \quad \frac{d\hat{\mu}}{dt} = -1 - \hat{D}. \tag{152}$$

and

$$\frac{\hat{y}\hat{\lambda} f_2(\hat{\lambda})}{f_3(\hat{\lambda})} + \frac{\hat{z}\hat{\lambda} f_1(\hat{\lambda})}{f_2(\hat{\lambda})} = 2\hat{\mu}. \tag{153}$$

We remark for future reference that (152) implies that

$$\hat{\mu} \text{ is decreasing with } t \text{ as long as } \hat{\lambda} > 0 \tag{154}$$

and (153) implies that

$$\hat{y} + \hat{z} \leq \frac{2\hat{\mu}}{\hat{\lambda}}. \tag{155}$$

Let

$$\hat{T}_0 = \min\left\{t : \hat{\lambda} \leq 8 \text{ or } \hat{y} + \hat{z} \leq n/2\right\}.$$

(The 8 is the 8 from Lemma 8.1 below).

Let $\mathbf{u} = \mathbf{u}(t)$ denote $(y(t), z(t), \mu(t))$ and let $\hat{\mathbf{u}} = \hat{\mathbf{u}}(t)$ denote $(\hat{y}(t), \hat{z}(t), \hat{\mu}(t))$. We now show that $\mathbf{u}$ and $\hat{\mathbf{u}}$ remain close:

Lemma 7.1.

$$||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1 \leq n^{8/9}, \qquad \textit{for } 1 \leq t \leq \min\left\{T_0, \hat{T}_0\right\} \; \textit{w.h.p.}.$$

Proof Let $\delta_\eta(\mathbf{v}), \eta = a, b, c, 2$ be the 0/1 indicator for the process 2GREEDY applying Step 1(η) for $\eta = a, b, c$ or Step 2 if $\eta = 2$ when the current state is $\mathbf{v}$. For times $t_1 < t_2$ we use the notation

$$\Delta_\eta(\mathbf{v}(t_1, t_2)) = \sum_{t=t_1}^{t_2} \delta_\eta(\mathbf{v}(t))$$

Now let $\rho = n^\alpha$ where $\alpha = 1/4$. It follows from Lemma 6.5 and our bound on λ in Lemma 6.3 that q.s. for $t \leq T_0 - \rho$,

$$|\lambda(t) - \lambda(t+\rho)| \leq \frac{\phi_3(3ce)}{n/2} ||\mathbf{v}(t+\rho) - \mathbf{v}(t)||_1 = O\left(\frac{\rho \log n}{n}\right). \tag{156}$$

Because λ changes very little, simple estimates then give

Claim 2.2.

$$|A(t) - A(t+\rho)| = O\left(\frac{\rho \log n}{n}\right) \qquad |B(t) - B(t+\rho)| = O\left(\frac{\rho \log n}{n}\right) \tag{157}$$

$$|C(t) - C(t+\rho)| = O\left(\frac{\rho \log n}{n}\right) \qquad |D(t) - D(t+\rho)| = O\left(\frac{\rho \log n}{n}\right) \tag{158}$$

If $||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1 \leq n^{8/9}$ *then*

$$|A(t) - \hat{A}(t)| = O\left(\frac{||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1}{n}\right) \qquad |B(t) - \hat{B}(t)| = O\left(\frac{||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1}{n}\right) \tag{159}$$

$$|C(t) - \hat{C}(t)| = O\left(\frac{||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1}{n}\right) \qquad |D(t) - \hat{D}(t)| = O\left(\frac{||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1}{n}\right) \tag{160}$$

Proof The first expressions in (157) and (158) are easy to deal with as the functions f_j are smooth and λ is bounded throughout, see Lemma 6.3. Thus each f_j changes by $O(\rho \log n/n)$ and y, z, μ change by $O(\rho \log n)$ and $y + z = \Omega(n)$.

For (159) and (160) we use Lemma 6.5 to argue that

$$|\lambda(t) - \hat{\lambda}(t)| = O\left(\frac{||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1}{n}\right).$$

Our assumption $t \leq \hat{T}_0$ implies that $\hat{y} + \hat{z} = \Omega(n)$ and we get a bound on $\phi_3(\hat{\lambda})$ by using Lemma 8.1. We can then argue as for (157) and (158).

End of proof of Claim 2.2

Now fix t and define for $\xi = y_1, y_2, z_1$,

$$X_i(\xi) = \begin{cases} \xi(t+i+1) - \xi(t+i) & t+i < \min\left\{T_0, \hat{T}_0\right\} \\ \mathbb{E}[\xi(t+1) - \xi(t) \mid \mathbf{v}(t)] & t+i \geq \min\left\{T_0, \hat{T}_0\right\} \end{cases}$$

Then,

$$\mathbb{E}[X_i(\xi) \mid \mathbf{v}(t+i)] = \sum_{\eta \in \{a,b,c,2\}} \delta_\eta(t+i)\psi_{\eta,\xi}(\mathbf{u}(t+i)) + O\left(\frac{\log^2 n}{n}\right) \tag{161}$$

It follows from (156) – (158) that for all η, ξ and $i \leq \rho$,

$$\psi_{\eta,\xi}(\mathbf{u}(t+i)) = \psi_{\eta,\xi}(\mathbf{u}(t)) + O\left(\frac{\log^3 n}{n^{1-\alpha}}\right).$$

It then follows from Lemma 6.6 and (161) that conditional on an event that occurs q.s.

$$\log^2 n \geq \mathbb{E}[\xi(t+\rho) - \xi(t) \mid \mathbf{u}(t)] = \sum_{\eta \in \{a,b,c,2\}} \Delta_\eta(\mathbf{u}(t, t+\rho))\psi_{\eta,\xi}(\mathbf{u}(t)) + O\left(\frac{\log^3 n}{n^{1-\alpha}}\right)$$

This can be written as follows: We let $\Delta_a = \Delta_a(\mathbf{u}(t, t+\rho))/\rho$ etc. and $A = A(t), B = B(t)$.

$$\begin{array}{lllll} \Delta_a & & & & = O(\rho^{-1}\log^2 n) \\ & \Delta_b & & & = A + O(\rho^{-1}\log^2 n) \\ & -\Delta_b & +\Delta_c & & = B + O(\rho^{-1}\log^2 n) \\ \Delta_a & +\Delta_b & +\Delta_c & +\Delta_2 & = 1. \end{array} \tag{162}$$

In comparison with (126) we see, using (159), (160) that

$$|\theta_\xi(\hat{\mathbf{u}}(t)) - \Delta_\xi| = O\left(\rho^{-1}\log^2 n + \frac{||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1}{n}\right) \text{ for } \xi = a, b, c, 2. \tag{163}$$

We now consider the difference between $\hat{\mathbf{u}}$ and $\mathbf{u}$ at times $\rho, 2\rho, \ldots$. We write

$$\xi(i\rho) - \hat{\xi}(i\rho) = \xi((i-1)\rho) - \hat{\xi}((i-1)\rho) + \sum_{t=(i-1)\rho+1}^{i\rho} ([\xi(t) - \xi(t-1)] - [\hat{\xi}(t) - \hat{\xi}(t-1)]) \tag{164}$$

where $\xi = y, z, \mu$ and $\hat{\xi} = \hat{y}, \hat{z}, \hat{\mu}$ in turn. Then we write

$$\xi(t) - \xi(t-1) = \alpha_t + \beta_t \text{ and } \hat{\xi}(t) - \hat{\xi}(t-1) = \hat{\alpha}_t + \hat{\beta}_t \tag{165}$$

where

$$\alpha_t = \sum_{\eta \in \{a,b,c,2\}} \delta_{\eta,\xi}(\mathbf{u}(t-1))\psi_{\eta,\xi}(\mathbf{u}(t-1)) \text{ and } \beta_t = \xi(t) - \xi(t-1) - \alpha_t$$

and

$$\hat{\alpha}_t = \sum_{\eta \in \{a,b,c,2\}} \theta_{\eta,\hat{\xi}}(\hat{\mathbf{u}}(t-1))\psi_{\eta,\hat{\xi}}(\hat{\mathbf{u}}(t-1)) \text{ and } \hat{\beta}_t = \hat{\xi}(t) - \hat{\xi}(t-1) - \hat{\alpha}_t.$$

It follows from (130), (131) etc. and Claim 2.2 that

$$\mathbb{E}[\beta_t \mid \mathbf{u}(t-1)] = O\left(\frac{\log^2 n}{n}\right).$$

An easy bound, which is a consequence of the Azuma-Hoeffding inequality, is that

$$\mathbb{P}\left(\sum_{t=(i-1)\rho+1}^{i\rho} \beta_t \geq \rho^{1/2}\log^2 n\right) \leq e^{-\Omega(\log^2 n)}. \tag{166}$$

We see furthermore that

$$\begin{aligned}
&\sum_{t=(i-1)\rho+1}^{i\rho} \hat{\beta}_t = \\
&= \sum_{t=(i-1)\rho+1}^{i\rho} \left(\hat{\xi}'(\hat{\mathbf{u}}(t-1+\varsigma_t)) - \sum_{\eta\in\{a,b,c,2\}} \theta_{\eta,\hat{\xi}}(\hat{\mathbf{u}}(t-1))\hat{\xi}_\eta'(\hat{\mathbf{u}}(t-1))\right) \\
&= \sum_{t=(i-1)\rho+1}^{i\rho} \left(\hat{\xi}'(\hat{\mathbf{u}}(t-1)) + O\left(\frac{\log n}{n}\right) - \sum_{\eta\in\{a,b,c,2\}} \theta_{\eta,\hat{\xi}}(\hat{\mathbf{u}}(t-1))\hat{\xi}_\eta'(\hat{\mathbf{u}}(t-1))\right) \\
&= O\left(\frac{\rho\log^2 n}{n}\right) = o(\rho^{1/2}\log^2 n),
\end{aligned} \tag{167}$$

where $0 \leq \varsigma_t \leq 1$ and $\hat{\xi}_\eta'(t)$ is the derivative of $\hat{\xi}$ in Case η.

Now write

$$\begin{aligned}
\sum_{t=(i-1)\rho+1}^{i\rho} \alpha_t &= \sum_{t=(i-1)\rho+1}^{i\rho} \sum_{\eta\in\{a,b,c,2\}} \delta_{\eta,\xi}(\mathbf{u}(t)) \left(\psi_{\eta,\xi}(\mathbf{u}((i-1)\rho) + O\left(\frac{\rho\log^2 n}{n}\right)\right) \\
&= \sum_{\eta\in\{a,b,c,2\}} \Delta_\xi(\mathbf{u}((i-1)\rho+1, i\rho))\psi_{\eta,\xi}(\mathbf{u}((i-1)\rho) + O\left(\frac{\rho^2\log^2 n}{n}\right)
\end{aligned} \tag{168}$$

and

$$\begin{aligned}
&\sum_{t=(i-1)\rho+1}^{i\rho} \hat{\alpha}_t = \\
&\quad \sum_{t=(i-1)\rho+1}^{i\rho} \sum_{\eta\in\{a,b,c,2\}} \left(\theta_{\eta,\xi}(\hat{\mathbf{u}}((i-1)\rho)) + O\left(\frac{\rho}{n}\right)\right)\left(\psi_{\eta,\xi}(\hat{\mathbf{u}}((i-1)\rho) + O\left(\frac{\rho}{n}\right)\right) \\
&\qquad = \sum_{\eta\in\{a,b,c,2\}} \rho\theta_{\eta,\xi}(\hat{\mathbf{u}}((i-1)\rho))\psi_{\eta,\xi}(\hat{\mathbf{u}}((i-1)\rho) + O\left(\frac{\rho^2}{n}\right).
\end{aligned} \tag{169}$$

It follows that

$$\sum_{t=(i-1)\rho+1}^{i\rho} (\hat{\alpha}_t - \alpha_t) = A_1 + A_2 + o\left(\rho^{1/2}\log^2 n\right) \tag{170}$$

where

$$A_1 = \sum_{\eta\in\{a,b,c,2\}} (\Delta_\xi(\mathbf{u}((i-1)\rho+1,i\rho)) - \rho\theta_{\eta,\xi}(\hat{\mathbf{u}}((i-1)\rho)))\psi_{\eta,\xi}(\mathbf{u}((i-1)\rho)$$

$$= O\left(\log^2 n + \frac{\rho||\mathbf{u}((i-1)\rho) - \hat{\mathbf{u}}((i-1)\rho)||_1}{n}\right). \tag{171}$$

$$A_2 = \rho\sum_{\eta\in\{a,b,c,2\}} \psi_{\eta,\xi}(\mathbf{u}((i-1)\rho)(\psi_{\eta,\xi}(\mathbf{u}((i-1)\rho) - \psi_{\eta,\xi}(\hat{\mathbf{u}}((i-1)\rho))$$

$$= O\left(\frac{\rho||\mathbf{u}((i-1)\rho) - \hat{\mathbf{u}}((i-1)\rho)||_1}{n}\right). \tag{172}$$

It follows from (164) to (172) that w.h.p., $i\rho \le T_0$ implies that with

$$a_i = ||\mathbf{u}(i\rho) - \hat{\mathbf{u}}(i\rho)||_1 \tag{173}$$

that for some $C_1 > 0$,

$$a_i \le a_{i-1}\left(1 + \frac{C_1\rho}{n}\right) + 2\rho^{1/2}\log^2 n.$$

Putting

$$\Pi_i = \prod_{j=0}^{i}\left(1 + \frac{C_1\rho}{n}\right) \le e^{C_1 i\rho/n}$$

we see by induction that

$$a_i \le 2\rho^{1/2}\log^2 n\sum_{j=0}^{i}\frac{\Pi_i}{\Pi_j} \le 2\rho^{1/2}\log^2 n(i+1)e^{C_1 i\rho/n}. \tag{174}$$

Since $i \le n/\rho$ we have

$$||\mathbf{u}(i\rho) - \hat{\mathbf{u}}(i\rho)||_1 = O(n\rho^{-1/2}\log^2 n).$$

Going from $\rho\lfloor T_0/\rho\rfloor$ to T_0 adds at most $\rho\log n$ to the gap and the lemma follows. □

8 Approximate equations

The equations (152) are rather complicated and we have not made much progress in solving them. Nevertheless, we can obtain information about them from a simpler set of equations that closely approximate them when c is sufficiently large. The important observation is that when $\hat{\lambda}$ is large,

$$\hat{A} \ll 1; \quad \hat{B} \ll 1; \quad \hat{C} \approx \frac{\hat{y}\hat{\lambda}}{2\hat{\mu}}; \quad \hat{D} \approx \frac{\hat{z}\hat{\lambda}^2}{2\hat{\mu}}; \quad \hat{\lambda} \approx \frac{2\hat{\mu}}{\hat{y}+\hat{z}}. \tag{175}$$

We will therefore approximate equations (152) by the following equations in variables $\tilde{y}, \tilde{z}, \tilde{\mu}, \tilde{\lambda}$:

$$\tilde{y}' = -\frac{\tilde{y}}{\tilde{y}+\tilde{z}} - 1 \tag{176}$$

$$\tilde{z}' = \frac{2\tilde{y}}{\tilde{y}+\tilde{z}} \tag{177}$$

$$\tilde{\mu}' = -1 - \frac{2\tilde{z}\tilde{\mu}}{(\tilde{y}+\tilde{z})^2} \tag{178}$$

$$\tilde{\lambda} = \frac{2\tilde{\mu}}{\tilde{y}+\tilde{z}}. \tag{179}$$

The initial conditions for $\tilde{y}, \tilde{z}, \tilde{\mu}, \tilde{\lambda}$ are that they start out equal to $\hat{y}, \hat{z}, \hat{\mu}, \hat{\lambda}$ at time $t = 0$ i.e.

$$\tilde{y}(0) = n; \quad \tilde{z}(0) = 0; \quad \tilde{\mu}(0) = cn; \quad \tilde{\lambda} = 2c. \tag{180}$$

8.0.1 Analysis of the approximate equations

The first two approximate equations imply $(\tilde{y} + \tilde{z}/2)' = -1$, so that

$$\tilde{y} + \frac{\tilde{z}}{2} = n - t.$$

Using the second approximate equation and $\tilde{y} = n - t - \tilde{z}/2$, we obtain

$$\tilde{z}' = \frac{2(n - t - \tilde{z}/2)}{n - t + \tilde{z}/2},$$

or, introducing $\tau = n - t$ and

$$X = \frac{\tilde{z}}{2(n-t)} = \frac{\tilde{z}}{2\tau},$$

we get

$$\frac{X+1}{X^2+1}\, dX = -\frac{1}{\tau}\, d\tau. \tag{181}$$

Integrating,

$$\frac{1}{2}\ln(X^2+1) + \arctan X = -\ln\tau + C.$$

Now, at $t = 0$ we have $\tau = n$ and $X = 0$. So $C = \ln n$, i.e.

$$\frac{1}{2}\ln(X^2+1) + \arctan X = -\ln(\tau/n).$$

Let $\tilde{T}$ satisfy $\tilde{y}(\tilde{T}) = 0$. At $t = \tilde{T}$, we have $X = 1$, so

$$\ln n - \ln(n - \tilde{T}) = \frac{1}{2}\ln 2 + \frac{\pi}{4}$$

which implies

$$\tilde{T} = \left(1 - \frac{1}{2^{1/2}} e^{-\pi/4}\right) n \approx 0.677603n. \tag{182}$$

Note that

$$\begin{aligned}
\tilde{\lambda}' &= \frac{2\tilde{\mu}'}{\tilde{y}+\tilde{z}} - \frac{2\tilde{\mu}(\tilde{y}'+\tilde{z}')}{(\tilde{y}+\tilde{z})^2} \\
&= -\frac{2}{\tilde{y}+\tilde{z}} - \frac{4\tilde{z}\tilde{\mu}}{(\tilde{y}+\tilde{z})^3} - \frac{2\tilde{\mu}}{(\tilde{y}+\tilde{z})^2}\left(\frac{\tilde{y}}{\tilde{y}+\tilde{z}} - 1\right) \\
&= -\frac{2}{\tilde{y}+\tilde{z}} - \frac{2\tilde{z}\tilde{\mu}}{(\tilde{y}+\tilde{z})^3}
\end{aligned}$$

$$= -\frac{2}{\tilde{y}+\tilde{z}} - \frac{\tilde{z}\tilde{\lambda}}{(\tilde{y}+\tilde{z})^2},$$

which implies that $\tilde{\lambda}$ is decreasing with t, at least as long as $\tilde{y}, \tilde{z}, \tilde{\lambda} > 0$. (183)

Here

$$\begin{aligned}\frac{\tilde{z}}{(\tilde{y}+\tilde{z})^2} &= \frac{\tilde{z}}{(n-t+\tilde{z}/2)^2} \\ &= \frac{\tilde{z}}{(n-t)^2(1+X)^2} \\ &= \frac{2X}{(n-t)(1+X)^2}.\end{aligned}$$

Likewise

$$-\frac{2}{\tilde{y}+\tilde{z}} = -\frac{2}{(n-t)(1+X)}.$$

So $\tilde{\lambda}$ satisfies

$$\tilde{\lambda}' = -\frac{2}{(n-t)(1+X)} - \frac{2X}{(n-t)(1+X)^2}\,\tilde{\lambda}, \quad \tilde{\lambda}(0) = 2c.$$

Using (181), we obtain

$$\frac{d\tilde{\lambda}}{dX} = -\frac{2}{1+X^2} - \frac{2X}{(1+X)(1+X^2)}\,\tilde{\lambda}, \quad \tilde{\lambda}(X)\Big|_{X=0} = 2c.$$

Integrating this first-order, linear ODE, we obtain

$$\tilde{\lambda}(X) = \frac{(1+X)e^{-\arctan X}}{\sqrt{1+X^2}}\left[2c - \int_0^X \frac{2e^{\arctan x}}{(1+x)\sqrt{1+x^2}}\,dx\right]. \tag{184}$$

In which case

$$\tilde{\lambda}(\tilde{T}) = \alpha_1 c - \alpha_2 \approx 1.53c - 1.418. \tag{185}$$

We have seen that $\tilde{\lambda}$ decreases with t, (unfortunately) we also need to verify that $\hat{\lambda}$ decreases with t, at least for $\hat{\lambda}$ sufficiently large.

Lemma 8.1. *[Pittel]*

$$\frac{d\hat{\lambda}}{dt} < 0 \text{ when } \hat{\lambda} \geq 8.$$

Proof Recall that

$$\hat{y}\frac{\hat{\lambda}f_2(\hat{\lambda})}{f_3(\hat{\lambda})} + \hat{z}\frac{\hat{\lambda}f_1(\hat{\lambda})}{f_2(\hat{\lambda})} = 2\hat{\mu}. \tag{186}$$

Therefore taking derivative with respect to t we have

$$\hat{y}'\frac{\hat{\lambda}f_2(\hat{\lambda})}{f_3(\hat{\lambda})} + \hat{z}'\frac{\hat{\lambda}f_1(\hat{\lambda})}{f_2(\hat{\lambda})} + \left(\hat{y}\frac{d}{d\hat{\lambda}}\frac{\hat{\lambda}f_2}{f_3} + \hat{z}\frac{d}{d\hat{\lambda}}\frac{\hat{\lambda}f_1}{f_2}\right)\hat{\lambda}' = 2\hat{\mu}'.$$

Now $\hat{\lambda}f_{i-1}(\hat{\lambda})/f_i(\hat{\lambda})$ are strictly increasing with $\hat{\lambda}$; so if $\hat{\lambda}'(t) \geq 0$ for some t, then it follows that

$$\hat{y}'\frac{\hat{\lambda}f_2(\hat{\lambda})}{f_3(\hat{\lambda})} + \hat{z}'\frac{\hat{\lambda}f_1(\hat{\lambda})}{f_2(\hat{\lambda})} \leq 2\hat{\mu}'. \tag{187}$$

Here $\hat{y}'$, $\hat{z}'$, $\hat{\mu}'$ are given by (152). Plugging their values into (187), and writing f_i instead of $f_i(\hat{\lambda})$, and re-arranging, we obtain

$$\frac{\hat{y}}{2\hat{\mu}}\left[\frac{\hat{\lambda}f_1}{f_2}\frac{\hat{\lambda}f_2}{f_3}-\left(\frac{\hat{\lambda}f_2}{f_3}\right)^2\right]+\frac{\hat{z}}{2\hat{\mu}}\left[2\frac{\hat{\lambda}^2f_0}{f_2}-\left(\frac{\hat{\lambda}f_1}{f_2}\right)^2\right]+\frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^5f_0}{f_2f_3}\left(\frac{\hat{\lambda}f_1}{f_2}-\frac{\hat{\lambda}f_2}{f_3}\right)$$
$$\leq\frac{\hat{\lambda}f_2}{f_3}-\frac{\hat{\lambda}f_1}{f_2}-2+\left(2\frac{\hat{\lambda}f_1}{f_2}-\frac{\hat{\lambda}f_2}{f_3}\right)\left(\frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^5f_0}{f_2f_3}+\frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4f_0}{f_2^2}\right).$$

Since $(\hat{y}/2\hat{\mu})\hat{\lambda}f_2/f_3+(\hat{z}/2\hat{\mu})\hat{\lambda}f_1/f_2\equiv 1$, we can multiply $(\hat{\lambda}f_2/f_3-\hat{\lambda}f_1/f_2-2)$ by this linear combination of $\hat{y}/2\hat{\mu}$ and $\hat{z}/2\hat{\mu}$ and re-arrange to obtain an equivalent inequality

$$\frac{\hat{y}}{2\hat{\mu}}\frac{2\hat{\lambda}f_2}{f_3}\left(\frac{\hat{\lambda}f_1}{f_2}+1-\frac{\hat{\lambda}f_2}{f_3}\right)+\frac{\hat{z}}{2\hat{\mu}}\frac{\hat{\lambda}f_1}{f_2}\left(2\frac{\hat{\lambda}f_0}{f_1}+2-\frac{\hat{\lambda}f_2}{f_3}\right)+\frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^5f_0}{f_2f_3}\left(\frac{\hat{\lambda}f_1}{f_2}-\frac{\hat{\lambda}f_2}{f_3}\right)$$
$$\leq\left(2\frac{\hat{\lambda}f_1}{f_2}-\frac{\hat{\lambda}f_2}{f_3}\right)\left(\frac{\hat{y}\hat{z}}{8\hat{\mu}^2}\frac{\hat{\lambda}^5f_0}{f_2f_3}+\frac{\hat{z}^2}{4\hat{\mu}^2}\frac{\hat{\lambda}^4f_0}{f_2^2}\right).$$

Multiplying the first two terms on the LHS by $(\hat{y}/2\hat{\mu})\hat{\lambda}f_2/f_3+(\hat{z}/2\hat{\mu})\hat{\lambda}f_1/f_2$, and introducing $x=\hat{y}/\hat{z}$, we obtain a quadratic inequality $\Phi(x,\hat{\lambda})=A_1x^2+A_2x+A_3\leq 0$, where

$$A_1=\frac{2\hat{\lambda}^2f_2^2}{f_3^2}\left(\frac{\hat{\lambda}f_1}{f_2}+1-\frac{\hat{\lambda}f_2}{f_3}\right).$$
$$A_2=\frac{2\hat{\lambda}^2f_1}{f_3}\left(\frac{\hat{\lambda}f_1}{f_2}+1-\frac{\hat{\lambda}f_2}{f_3}\right)+\frac{\hat{\lambda}f_2}{f_3}\left(\frac{2\hat{\lambda}f_0}{f_1}+2-\frac{\hat{\lambda}f_2}{f_3}\right)-\frac{\hat{\lambda}^6f_0f_1}{2f_2^2f_3}.$$
$$A_3=\frac{\hat{\lambda}f_1}{f_2}\left(2\frac{\hat{\lambda}f_0}{f_1}+2-\frac{\hat{\lambda}f_2}{f_3}\right)-\frac{\hat{\lambda}^4f_0}{f_2^2}\left(2\frac{\hat{\lambda}f_1}{f_2}-\frac{\hat{\lambda}f_2}{f_3}\right).$$

If we show that $\Phi(x,\hat{\lambda})>0$ for all $x\geq 0$, and $\hat{\lambda}\geq 8$, we will be able to conclude that $d\hat{\lambda}/dt<0$ for $\hat{\lambda}\geq 8$. This will follow from the fact that $A_1,A_2,A_3>0$ when $\hat{\lambda}\geq 8$.

For this observe that $f_{i-1}(x)/f_i(x)$ decreases monotonically to one, for all i and for all $i,k>0$, $\hat{\lambda}^k/f_i(\hat{\lambda})$ decreases for $\hat{\lambda}\geq k$. We have

$$\frac{f_0(8)}{f_1(8)}\approx 1.00034;\ \frac{f_1(8)}{f_2(8)}\approx 1.00269;\ \frac{f_1(8)}{f_2(8)}\approx 1.01088;\ \frac{8^4}{f_2(8)}\approx 1.37822;\ \frac{8^6}{2f_3(8)}\approx 44.583.$$

With these values, it is easy to see that $A_1,A_2,A_3>0$ for $\hat{\lambda}\geq 8$. □

It is of course likely that the condition $\hat{\lambda}\geq 8$ is unnecessary in the above lemma. We do not need to consider $\hat{\lambda}<8$ in the subsequent calculations and we do not deem it worthwhile at present to sharpen the lemma.

8.0.2 Simple Inequalities

We will use the following to quantify (175):

$$1\leq\frac{f_2(\hat{\lambda})}{f_3(\hat{\lambda})}=1+\varepsilon_1,\qquad 1\leq\frac{f_0(\hat{\lambda})}{f_2(\hat{\lambda})}=1+\varepsilon_2,\qquad 1\leq\frac{f_0(\hat{\lambda})}{f_3(\hat{\lambda})}=1+\varepsilon_3.$$

where

$$\varepsilon_1 = \frac{\hat{\lambda}^2}{2f_3(\hat{\lambda})}, \qquad \varepsilon_2 = \frac{1+\hat{\lambda}}{f_2(\hat{\lambda})}, \qquad \varepsilon_3 = \frac{\hat{\lambda}^2+2\hat{\lambda}+2}{2f_3(\hat{\lambda})}.$$

Note that $\varepsilon_2 < \varepsilon_1$ for $\hat{\lambda} \geq 3$ which follows from $f_2(\hat{\lambda}) > f_3(\hat{\lambda})$ and $\hat{\lambda}^2 > 2(1+\hat{\lambda})$ for $\hat{\lambda} \geq 3$.

We use the above to verify the following sequence of inequalities for $\hat{y}, \hat{z}, \hat{\mu}, \hat{\lambda}$:

$$\frac{2\hat{\mu}}{\hat{y}+\hat{z}} \cdot \frac{1}{1+\varepsilon_1} \leq \hat{\lambda} \leq \frac{2\hat{\mu}}{\hat{y}+\hat{z}}. \tag{188}$$

$$0 \leq \hat{A} \leq \varepsilon_A$$

$$0 \leq \hat{B} \leq \varepsilon_B$$

$$\frac{\hat{y}\hat{\lambda}}{2\hat{\mu}} \leq \hat{C} = \frac{y\hat{\lambda}}{2\hat{\mu}}(1+\varepsilon_1).$$

$$\frac{z\hat{\lambda}^2}{2\hat{\mu}} \leq \hat{D} = \frac{\hat{z}\hat{\lambda}^2}{2\hat{\mu}}(1+\varepsilon_2).$$

where

$$\varepsilon_A = \frac{(1+\varepsilon_2)(1+\varepsilon_3)\hat{\lambda}^3}{8f_0(\hat{\lambda})}, \qquad \varepsilon_B = \frac{\hat{\lambda}^2(1+\varepsilon_2)^2}{f_0(\hat{\lambda})}.$$

(We use (155) to get $\hat{y}\hat{z} \leq \hat{\mu}^2/\hat{\lambda}^2$ for use in defining ε_A).

For (188) we use

$$\frac{\hat{\lambda}(\hat{y}+\hat{z})}{2\hat{\mu}} \geq \frac{1}{\max\left\{\frac{f_2(\hat{\lambda})}{f_3(\hat{\lambda})}, \frac{f_1(\hat{\lambda})}{f_2(\hat{\lambda})}\right\}} = \frac{f_3(\hat{\lambda})}{f_2(\hat{\lambda})} = \frac{1}{1+\varepsilon_1}.$$

It follows from (188) that the initial value $\hat{\lambda}_0$ of $\hat{\lambda}$ satisfies

$$\frac{2c}{1+\varepsilon_1} \leq \hat{\lambda}_0 \leq 2c.$$

Now $\hat{\lambda}_0$ is the solution to $\hat{\lambda} f_2(\hat{\lambda})/f_3(\hat{\lambda}) = 2c$. It follows that $\hat{\lambda}_0 > 17$ for $c \geq 10$. Furthermore, $\varepsilon_1 \leq .0001$ for $\hat{\lambda} \geq 17$ and so

$$2c(1-.001) \leq \hat{\lambda}_0 \leq 2c. \tag{189}$$

for $c \geq 10$.

8.0.3 Main Goal

Lemma 8.2 (below) in conjunction with Lemma 7.1, will enable us to argue that w.h.p. in the process 2Greedy, at some time $T \leq \min\left\{T_0, \hat{T}_0\right\}$ we will have

$$y(T) = 0, z(T) = \Omega(n) \text{ and } \lambda(t) = \Omega(1) \text{ for } t \leq T. \tag{190}$$

Define

$$T_+ = \min\left\{t > 0 : \hat{y}(t) \leq 0 \text{ or } \hat{z}(t) \leq 0 \text{ or } \tilde{y}(t) \leq 0 \text{ or } \tilde{z}(t) \leq 0\right\}.$$

We can bound this from below by a small constant as follows: Initially $\hat{A}, \hat{B}$ are small $\hat{C}$ is close to one for $c \geq 10$ and so (152) implies that $\hat{z}$ is strictly increasing at the beginning. Also, $\hat{y}, \tilde{y}$ start out large ($= n$) and so remain positive initially.

Next define

$$T_1 = \min\left\{T_+, \max\left\{t : \hat{\lambda}(\tau)) \geq \lambda^* \text{ and } \min\{\tilde{y}(\tau)+\tilde{z}(\tau), \hat{y}(\tau)+\hat{z}(\tau)\} \geq \beta n \text{ for } \tau \leq t\right\}\right\} \quad (191)$$

where

$$\beta = -.01 + \tilde{z}(\tilde{T})/n = -.01 + 2(n-\tilde{T})/n \approx .63$$

and

$$8.8513 < \lambda^* < 8.8514 \quad (192)$$

is defined to be the solution to

$$2\hat{\lambda}\varepsilon_1 = \frac{1}{10}.$$

(This definition of λ^* means that $\hat{\mu}$ can easily be seen to be a decreasing function for $\lambda \geq \lambda^*$, see (205).)

Comparing (185) and (189) we see that if $c \geq 10$ then $\hat{\lambda}_0 > \lambda^*$.

Note that

$$T_1 \leq T_+ \leq \tilde{T}. \quad (193)$$

which follows from $\tilde{y}(\tilde{T}) = 0$.

Lemma 8.2. *For $c \geq 10$,*

$$\hat{y}(T_1) = 0 < \hat{z}(T_1) = \Omega(n) \text{ and } \hat{\lambda}(T_1) = \Omega(1). \quad (194)$$

Proof It follows from (152) and Section 8.0.2 that

$$\hat{y}' \leq \varepsilon_A + \varepsilon_B - \frac{\hat{y}\hat{\lambda}}{2\hat{\mu}} - 1 \leq -\frac{\hat{y}}{\hat{y}+\hat{z}} - 1 + \left(\frac{\varepsilon_1}{1+\varepsilon_1} + \varepsilon_A + \varepsilon_B\right)$$

$$\hat{y}' \geq -\frac{\hat{y}\hat{\lambda}}{2\hat{\mu}}(1+\varepsilon_1) - 1 \geq -\frac{\hat{y}}{\hat{y}+\hat{z}} - 1 - \varepsilon_1. \quad (195)$$

$$\hat{z}' \leq \frac{2\hat{y}\hat{\lambda}}{2\hat{\mu}}(1+\varepsilon_1) \leq \frac{2\hat{y}}{\hat{y}+\hat{z}} + 2\varepsilon_1.$$

$$\hat{z}' \geq \frac{2\hat{y}}{\hat{y}+\hat{z}} - (2\varepsilon_A + 2\varepsilon_B) \quad (196)$$

$$\hat{\lambda} \leq \frac{2\hat{\mu}}{\hat{y}+\hat{z}} \quad (197)$$

$$\hat{\lambda} \geq \frac{2\hat{\mu}}{\hat{y}+\hat{z}} \cdot \frac{1}{1+\varepsilon_1} \geq \frac{2\hat{\mu}}{\hat{y}+\hat{z}} - \hat{\lambda}\varepsilon_1 \quad (198)$$

$$\hat{\mu}' = -1 - \frac{\hat{z}\hat{\lambda}^2}{2\hat{\mu}} \leq -1 - \frac{2\hat{z}\hat{\mu}}{(\hat{y}+\hat{z})^2} \cdot \frac{1}{(1+\varepsilon_1)^2} \leq -1 - \frac{2\hat{z}\hat{\mu}}{(\hat{y}+\hat{z})^2} + 2\hat{\lambda}\varepsilon_1.$$

$$\hat{\mu}' \geq -1 - \frac{\hat{z}\hat{\lambda}^2}{2\hat{\mu}}(1+\varepsilon_2) \geq -1 - \frac{2\hat{z}\hat{\mu}}{(\hat{y}+\hat{z})^2} - \hat{\lambda}\varepsilon_2. \quad (199)$$

When $t=0$ we have $\hat{y}=n, \hat{z}=0, \hat{\mu}=cn$ and $\hat{\lambda}$ satisfying (189), we see that $T_1>0$ for $c \geq 10$. We can write $\hat{y}(0)=n, \hat{z}(0)=0, \hat{\mu}(0)=cn$ and

$$\hat{y}' = -\frac{\hat{y}}{\hat{y}+\hat{z}} - 1 + \theta_1 \qquad \text{where } \theta_1 \in \left[-\varepsilon_1, \frac{\varepsilon_1}{1+\varepsilon_1} + \varepsilon_A + \varepsilon_B\right]. \tag{200}$$

$$\hat{z}' = \frac{2\hat{y}}{\hat{y}+\hat{z}} + \theta_2 \qquad \text{where } \theta_2 \in [-2\varepsilon_A - 2\varepsilon_B, 2\varepsilon_1]. \tag{201}$$

$$\hat{\mu}' = -1 - \frac{2\hat{z}\hat{\mu}}{(\hat{y}+\hat{z})^2} + \theta_3 \qquad \text{where } \theta_3 \in \left[-\hat{\lambda}\varepsilon_2, 2\hat{\lambda}\varepsilon_1\right] \tag{202}$$

$$\hat{\lambda} = \frac{2\hat{\mu}}{\hat{y}+\hat{z}} + \theta_4 \qquad \text{where } \theta_4 \in [-\hat{\lambda}\varepsilon_1, 0]. \tag{203}$$

Next let

$$\delta^* = \delta^*(\hat{\lambda}) = \max\left\{\varepsilon_1, \varepsilon_A + \varepsilon_B + \frac{\varepsilon_1}{1+\varepsilon_1}, \varepsilon_A + \varepsilon_B\right\}$$

so that we have

$$|\theta_1| \leq \delta^* \text{ and } |\theta_2| \leq 2\delta^*.$$

It can easily be checked that the functions $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_A, \varepsilon_B$ are all monotone decreasing for $\hat{\lambda} \geq \lambda^*$. Also,

$$\hat{\lambda} \geq 10 \text{ implies } \delta^* \leq .011. \tag{204}$$

It follows from (178) and (202) that

$$\hat{\mu}, \tilde{\mu} \text{ both decrease for } t \leq T_1, \text{ since } \theta_3 < 1 \text{ for } \hat{\lambda} \geq \lambda^*. \tag{205}$$

The ensuing calculations involve many constants and expressions that are tedious to justify. It is unrealistic to expect the reader to check these calculations. Instead, we have provided mathematica output in an appendix that will be seen to justify our claims.

The reader will notice the similarity between these equations and the approximation (176) – (179). We will now refer to the equations (200) – (203) as the *true* equations and (176) – (179) as the *approximate* equations.

8.0.4 $\tilde{y}, \tilde{z}$ and $\hat{y}, \hat{z}$ are close

We claim next that

$$\max\{|\hat{y}(t) - \tilde{y}(t)|, |\hat{z}(t) - \tilde{z}(t)|\} \leq \delta^* F_1(t/n)n \text{ for } 0 \leq t \leq T_1. \tag{206}$$

where $\delta^* = \delta^*(\hat{\lambda}(t))$ and

$$F_a(x) = \beta(e^{2ax/\beta} - 1) \text{ for } x \leq \frac{\tilde{T}}{n}. \tag{207}$$

for $a > 0$.

Note that

$$F_a'(t) = 2a(F_a(t)/\beta + 1).$$

In the proof of (206), think of n as fixed and h as a parameter that tends to zero. Think of ε as small, but fixed until the end of the proof. In the display beginning with equation (209), only h is the quantity going to zero. Let

$$\hat{u}_i = \hat{y}(ih), \hat{v}_i = \hat{z}(ih), \tilde{u}_i = \tilde{y}(ih), \tilde{v}_i = \tilde{z}(ih) \text{ for } 0 \le i \le n/h.$$

Assume inductively that for $i < i_0 = T_1/h$

$$|\hat{u}_i - \tilde{u}_i|, |\hat{v}_i - \tilde{v}_i| \le \delta^* F_{1+\varepsilon}(ih/n)n. \tag{208}$$

This is true for $i = 0$.

Suppose that

$$F_{1+\varepsilon}((i+1)h/n) = F_{1+\varepsilon}(ih/n) + \frac{h}{n} F'_{1+\varepsilon}((i+\theta)h/n)$$

for some $0 \le \theta \le 1$.

Then by the inductive assumption and the Taylor expansion and uniform boundedness of second derivatives,

$$\begin{aligned}
\hat{v}_{i+1} - \tilde{v}_{i+1} &= \hat{v}_i - \tilde{v}_i + h\left(\frac{2\hat{u}_i}{\hat{u}_i + \hat{v}_i} - \frac{2\tilde{u}_i}{\tilde{u}_i + \tilde{v}_i} + \theta_2(ih)\right) + O(h^2) \qquad (209)\\
&= \hat{v}_i - \tilde{v}_i + h\left(\frac{2\hat{u}_i(\tilde{v}_i - \hat{v}_i) - 2\hat{v}_i(\tilde{u}_i - \hat{u}_i)}{(\hat{u}_i + \hat{v}_i)(\tilde{u}_i + \tilde{v}_i)} + \theta_2(ih)\right) + O(h^2)\\
&\le \hat{v}_i - \tilde{v}_i + h\left(\frac{2(\hat{u}_i + \hat{v}_i)\max\{|\hat{v}_i - \tilde{v}_i|, |\hat{u}_i - \tilde{u}_i|\}}{(\hat{u}_i + \hat{v}_i)(\tilde{u}_i + \tilde{v}_i)} + \theta_2(ih)\right) + O(h^2)\\
&\le \delta^* F_{1+\varepsilon}(ih/n)n + 2\delta^* h\left(F_{1+\varepsilon}(ih/n)/\beta + 1\right) + O(h^2)
\end{aligned}$$

Here we have used Lemma 8.1 to replace $\theta_2(\hat{\lambda}(ih))$ by $2\delta^*$. We see from (191) and (192) and induction on i that $\hat{\lambda}(ih) \ge \hat{\lambda}(T_1) > 8$.

$$\begin{aligned}
&= \delta^*(F_{1+\varepsilon}(ih/n)n + hF'_{1+\varepsilon}(ih/n) - 2h\varepsilon(F_{1+\varepsilon}(ih/n)/\beta + 1)) + O(h^2)\\
&\le \delta^*(F_{1+\varepsilon}((i+1)h/n)n + h(F'_{1+\varepsilon}(ih/n) - F'_{1+\varepsilon}((i+\theta)h/n)) - \varepsilon) + O(h^2)\\
&\le \delta^*(F_{1+\varepsilon}((i+1)h/n)n,
\end{aligned}$$

completing the induction, for small enough h.

The remaining three cases are proved similarly. This completes the inductive proof of (208). Letting $\varepsilon \to 0$ we see for example that $\hat{y}(t) - \tilde{y}(t) \le \delta^* F_1(t)$ for $t \le T_1$. This completes the proof of (206).

Let

$$\alpha_0 = F_1(\tilde{T}/n).$$

Observe next that

$$(\tilde{y} + \tilde{z})' = \frac{\tilde{y}}{\tilde{y} + \tilde{z}} - 1 \le 0. \tag{210}$$

So for $t \le T_1$ we have

$$\tilde{y} + \tilde{z} \ge \tilde{y}(\tilde{T}) + \tilde{z}(\tilde{T}) = \tilde{z}(\tilde{T}) = 2(n - \tilde{T}) = (\beta + .01)n. \tag{211}$$

8.0.5 Lower bounding $\hat{\lambda}$

We now show that $\tilde{\lambda} - \hat{\lambda}$ is small. We now use (202) and (206) to write for $t \le T_1$,

$$
\begin{aligned}
&|\tilde{\mu}' - \hat{\mu}'| \\
&\le |\theta_3| + \left| \frac{2\hat{\mu}\hat{z}((\hat{y}+\hat{z})^2 + 4\delta^* F_1(t/n)(\hat{y}+\hat{z})n + 4\delta^{*2} F_1(t/n)^2 n^2) - 2\tilde{\mu}(\hat{y}+\hat{z})^2(\hat{z} - \delta^* F_1(t/n)n)}{(\hat{y}+\hat{z})^2(\tilde{y}+\tilde{z})^2} \right| \\
&\le 2\hat{\lambda}\delta^* + \left| \frac{2\hat{z}(\hat{y}+\hat{z})^2(\hat{\mu} - \tilde{\mu}) + 2\delta^* F_1(t/n)n(\hat{y}+\hat{z})(4\hat{\mu}\hat{z} + \tilde{\mu}(\hat{y}+\hat{z})) + 8\hat{\mu}\hat{z}\delta^{*2} F_1(t/n)^2 n^2}{(\hat{y}+\hat{z})^2(\tilde{y}+\tilde{z})^2} \right|.
\end{aligned}
$$

Now, using (205),

$$\frac{4\hat{\mu}\hat{z} + \tilde{\mu}(\hat{y}+\hat{z})}{(\hat{y}+\hat{z})(\tilde{y}+\tilde{z})^2} \le \frac{4\hat{\mu} + \tilde{\mu}}{(\tilde{y}+\tilde{z})^2} \le \frac{5c}{\beta^2 n}$$

and

$$\frac{8\hat{\mu}\hat{z}}{(\hat{y}+\hat{z})^2(\tilde{y}+\tilde{z})^2} \le \frac{8c}{\beta^3 n^2}.$$

So,

$$|\tilde{\mu}' - \hat{\mu}'| \le \frac{2|\hat{\mu} - \tilde{\mu}|}{\beta n} + \alpha_3\delta^*$$

where

$$\alpha_3 = \alpha_3(c, \hat{\lambda}) = \frac{10\alpha_0 c}{\beta^2} + \frac{8c\alpha_0^2\delta^*}{\beta^3} + \frac{4c}{\beta},$$

where the third term follows from $\hat{\lambda} \le 2c/\beta$, see (197).

Integrating, we get that if $\gamma = \hat{\mu} - \tilde{\mu}$ then

$$\gamma' - \frac{2|\gamma|}{\beta n} \le \alpha_3\delta^*.$$

Let f be the solution to

$$f(0) = 0 \text{ and } f'(x) = \frac{2f(x)}{\beta n} + \alpha_3\delta^*.$$

Then we have

$$\gamma(t) \le f(t) = \alpha_3\delta^* e^{2t/\beta n} \int_{\tau=0}^{t} e^{-2\tau/\beta n} d\tau = \alpha_3\delta^* \frac{\beta n}{2}(e^{2t/\beta n} - 1) \le \alpha_4\delta^* n.$$

for $t \le T_1$, where

$$\alpha_4 = \alpha_4(c, \hat{\lambda}) = \alpha_0\alpha_3/2.$$

The same inequality will hold for $-\gamma$ and so we have $|\gamma(t)| \le \alpha_4\delta^* n$ for $t \le T_1$.

It then follows that as long as $t \le T_1$,

$$
\begin{aligned}
\tilde{\lambda} - \hat{\lambda} &= -\theta_4 + \frac{2\tilde{\mu}(\hat{y}+\hat{z}) - 2\hat{\mu}(\tilde{y}+\tilde{z})}{(\hat{y}+\hat{z})(\tilde{y}+\tilde{z})} \\
&\le \hat{\lambda}\varepsilon_1 + \frac{2(\hat{\mu} + \alpha_4\delta^* n)(\hat{y}+\hat{z}) - 2\hat{\mu}(\hat{y}+\hat{z} - 2\alpha_0\delta^* n)}{(\hat{y}+\hat{z})(\tilde{y}+\tilde{z})} \\
&\le \hat{\lambda}\varepsilon_1 + \frac{2\alpha_4\delta^*}{\beta} + \frac{4c\alpha_0\delta^*}{\beta^2}.
\end{aligned}
$$

It follows from (183) that for $t \leq T_1$ we have

$$\hat{\lambda}(t) \geq \tilde{\lambda}(T_1) - \alpha_5 \delta^* \tag{212}$$

where

$$\alpha_5 = \alpha_5(c, \hat{\lambda}) = \frac{2c}{\beta} + \frac{2\alpha_4}{\beta} + \frac{4c\alpha_0}{\beta^2}.$$

We now argue that $\hat{y}(T_1) = 0$ and $\hat{\lambda}(T_1) \geq \lambda^*$. This proves Lemma 8.2, since $\hat{y}(T_1) + \hat{z}(T_1) \geq \beta n$. Suppose then to the contrary that $\hat{y}(T_1) > 0$. Recall that $T_1 \leq \tilde{T}$ (see (193)) and suppose first that $T_1 < \tilde{T}$. Now let

$$T_2 = \min\left\{T_1 + \varepsilon n, (T_1 + \tilde{T})/2\right\}$$

where $0 < \varepsilon < 10^{-10}$ is such that

$$\max\left\{\tau \in [T_1, T_2] : \varepsilon \max\left\{|\hat{\lambda}'(\tau)|, |\hat{y}'(\tau)|, |\hat{z}'(\tau)|\right\} \leq 10^{-10}\right\}. \tag{213}$$

The existence of such an ε follows by elementary propositions in real analysis.

We will argue that $\tau \in [T_1, T_2]$ implies

$$\hat{\lambda}(\tau) \geq \lambda^* \text{ and } \min\{\hat{y}(\tau) + \hat{z}(\tau), \tilde{y}(\tau) + \tilde{z}(\tau)\} \geq \beta n \text{ and } \min\{\hat{y}(\tau), \hat{z}(\tau), \tilde{y}(\tau), \tilde{z}(\tau)\} > 0, \tag{214}$$

which contradicts the definition of T_1.

Fix $\tau \in [T_1, T_2]$. Now $\tau < \tilde{T}$ implies that $\tilde{y}(\tau) > 0$. Together with (177) we see that $\tilde{z}$ increases for $t \leq \tilde{T}$ and hence $\tilde{z}(\tau) > 0$. We have $\tilde{y}'(t) \geq -2$ (see (176)) and $\tilde{z}'(t) \geq 0$ for $t \leq T_2$ (see (177)) and so for some $\tau_1, \tau_2 \in [T_1, T_2]$

$$\tilde{y}(\tau) = \tilde{y}(T_1) + (\tau - T_1)\tilde{y}'(\tau_1) \geq \tilde{y}(T_1) - 2\varepsilon n.$$
$$\tilde{z}(\tau) = \tilde{z}(T_1) + (\tau - T_1)\tilde{z}'(\tau_2) \geq \tilde{z}(T_1).$$

It follows that

$$\tilde{y}(\tau) + \tilde{z}(\tau) \geq \tilde{y}(T_1) + \tilde{z}(T_1) - 2\varepsilon n \geq \tilde{y}(\tilde{T}) + \tilde{z}(\tilde{T}) - 2\varepsilon n > \beta n. \tag{215}$$

We have, for some $\tau_3, \tau_4 \in [T_1, T_2]$,

$$\begin{aligned}\hat{y}(\tau) + \hat{z}(\tau) &= \hat{y}(T_1) + \hat{z}(T_1) + (\tau - T_1)(\hat{y}(\tau_3) + \hat{z}(\tau_4))' \\ &\geq \tilde{y}(T_1) + \tilde{z}(T_1) - 2F_1(\tilde{T}/n)\delta^* n - 2 \times 10^{-10} n \\ &\geq \left(\beta + .01 - 2\left(\alpha_0 \delta^* + 10^{-10}\right)\right) n \\ &\geq \beta n. \end{aligned} \tag{216}$$

We now argue that $\hat{z}(\tau) > 0$. Equation (177) shows that $\tilde{z}$ is strictly increasing initially. Also, if $\hat{\lambda} \geq \lambda^*$ then $\theta_3 \leq 1/10$. From (201) we see that $\hat{z}$ is strictly increasing at least until a time τ_0 when $\hat{y}(\tau_0) \leq \beta \delta^*$. On the other hand, we see from (216) that if $\hat{y}(\tau) \leq \beta \delta^*$ then $\hat{z}(\tau) > 0$. So,

$$\min\{\hat{y}(\tau), \hat{z}(\tau), \tilde{y}(\tau), \tilde{z}(\tau)\} > 0. \tag{217}$$

Now we write

$$\begin{aligned}\hat{\lambda}(\tau) &= \hat{\lambda}(T_1) + (\tau - T_1)\hat{\lambda}'(\tau_3)\\ &\geq \tilde{\lambda}(T_1) - (\tilde{\lambda}(T_1) - \hat{\lambda}(T_1)) - 10^{-10}, \qquad \text{using (213).}\\ &\geq \tilde{\lambda}(\tilde{T}) - \alpha_5\delta^* - 10^{-10},\\ &\geq \alpha_1 c - \alpha_2 - \alpha_5\delta^* - 10^{-10}, \qquad \text{where } \alpha_5 = \alpha_5(c, \hat{\lambda}(\tilde{T})) \text{ and } \delta^* = \delta^*(\hat{\lambda}(\tilde{T})).\end{aligned} \tag{218}$$

Now,

$$\alpha_1 \geq 1.531 \text{ and } \alpha_2 \leq 1.419 \text{ and so } \alpha_1 c - \alpha_2 \geq 13.5 \text{ for } c \geq 10. \tag{219}$$

Furthermore,

$$\alpha_5(10, 13.5) \leq 9800 \text{ and } \delta^*(13.5) \leq .0005 \tag{220}$$

and so the RHS of (218) is at least $8.9 > \hat{\lambda}^*$, see (192), when $c = 10$. We claim next that $\alpha_5(c, \hat{\lambda}(\tilde{T})) \times \delta^*(\hat{\lambda}(\tilde{T}))$ decreases monotonically with c. This follows from the fact that $c\delta^*(\hat{\lambda}(\tilde{T}))$ is monotonically decreasing. For this note that $\hat{\lambda}(\tilde{T}) \geq c$ for $c \geq 10$ and it is easy to check that $c\delta^*(c)$ decreases for $c \geq 10$. This verifies (214).

We must now deal with the case where $T_1 = \tilde{T}$. Here we can just use (206) to argue that $\hat{z}(T_1) > \tilde{z}(T_1) - \alpha_0\delta^* n > 0$ and $\hat{y}(T_1) + \hat{z}(T_1) > \tilde{y}(T_1) + \tilde{z}(T_1) - \alpha_0\delta^* n > (\beta + .01 - \alpha_0\delta^*)n > \beta n$ and $\hat{\lambda}(\tilde{T}) \geq \tilde{\lambda}(\tilde{T}) - \alpha_5\delta^* > \lambda^*$.

The above calculations imply that $T_1 < \hat{T}_0$, justifying the use of Lemma 7.1.

This completes the proof of Lemma 8.2. □

It follows from Lemma 7.1 that w.h.p. $y(T_1) \leq n^{8/9}$, $z(T_1) \geq \beta n - n^{8/9}$ and $\lambda(T_1) \geq \lambda^*$. We claim that q.s., y becomes zero within the next $\nu = n^{9/10}$ steps of 2GREEDY. Suppose not. It follows from Lemma 6.5 that λ changes by $o(1)$ and by (6) that z changes by $o(n)$ during these ν steps. Thus $T_1 + \nu \leq T_0$. It follows from (92) that q.s. at least $\nu \log^{-2} n$ of these steps will be of type Step 2. But each such step reduces y by at least one, contradiction.

This verifies (190).

9 The number of components in the output of the algorithm

We will tighten our bound on ζ from Lemma 6.6.

Lemma 9.1. *If $c \geq 10$ then for every positive constant K there exists a constant $c_2 = c_2(K)$ such that*

$$\mathbb{P}\left(\exists 1 \leq t \leq T_1 : \ \zeta(t) > c_2 \log n\right) \leq n^{-K}.$$

Proof We now need to use a sharper inequality than (90) to replace L_1 by what is claimed in the statement of the lemma. This sharper inequality uses higher moments of the X_t's and we can estimate them now that we have the estimate of the maximum of $\zeta(t)$ given in (90). So, we now have to estimate terms of the form

$$\Psi_j(\xi \mid \boldsymbol{\eta}) = \mathbb{E}[\ |(\xi' - \xi) - \mathbb{E}[(\xi' - \xi) \mid \boldsymbol{\eta}]|^j \ \mid \boldsymbol{\eta}).$$

for $\xi = y_1, y_2, z_2$, $2 \le j \le \log n$ and $\boldsymbol{\eta} = \mathbf{v}$ or $\mathbf{b}, \mathbf{d}$.

We use the inequality

$$(a+b+c+d)^j \le 4^j(|a|^j + |b|^j + |c|^j + |d|^j)$$

for $j \ge 1$.

We will also need to estimate, for $2 \le j \le \log n$,

$$\sum_{k\ge 2} \frac{k(k-1)^j \lambda^k}{k!} = \lambda^2 \sum_{k\ge 0} \frac{(k+1)^{j-1}\lambda^k}{k!} < 2^j \lambda^2 \sum_{k\ge 0} \frac{k^j \lambda^k}{k!} = 2^j \lambda^2 \sum_{k\ge 0}\sum_{\ell=0}^{j} \left\{ {j \atop \ell} \right\} \frac{(k)_\ell \lambda^k}{k!}$$
$$= 2^j \lambda^2 \sum_{\ell=0}^{j} \left\{ {j \atop \ell} \right\} \lambda^\ell \sum_{k\ge \ell} \frac{\lambda^{k-\ell}}{(k-\ell)!} \le 2^j \lambda^{j+2} e^\lambda \sum_{\ell=0}^{j} \left\{ {j \atop \ell} \right\} \le 2^j j! \lambda^{j+2} e^\lambda.$$

Here $\left\{ {j \atop \ell} \right\}$ is a Stirling number of the second kind and it is easy to verify by induction on j that the Bell number $\sum_{\ell=0}^{j} \left\{ {j \atop \ell} \right\} \le j!$.

Step 1. $y_1 + y_2 + z_1 > 0$.

Step 1(a). $y_1 > 0$.

$$\Psi_j(y_1 \mid \mathbf{b},\mathbf{d}) \le 4^j \left(\frac{y_1}{2\mu} + \sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{y_1}{2\mu} + \sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{2y_2}{2\mu} + \varepsilon_{221} \right). \tag{221}$$

$$\Psi_j(y_i \mid \mathbf{v}) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \tag{222}$$

$$\Psi(y_2 \mid \mathbf{b},\mathbf{d}) \le 4^j \left(\frac{2y_2}{2\mu} + \sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{2y_2}{2\mu} + \sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{3y_3}{2\mu} + \varepsilon_{223} \right). \tag{223}$$

$$\Psi_j(y_2 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \tag{224}$$

$$\Psi(z_1 \mid \mathbf{b},\mathbf{d}) \le 4^j \left(\frac{z_1}{2\mu} + \sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{z_1}{2\mu} + \sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{2z_2}{2\mu} + \varepsilon_{225} \right). \tag{225}$$

$$\Psi(z_1 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \tag{226}$$

(227)

Step 1(b). $y_1 = 0, y_2 > 0$.

$$\Psi_j(y_1 \mid \mathbf{b},\mathbf{d}) \le 4^j \left(2\sum_{k\ge 2} \frac{kz_k}{2\mu}(k-1)^j \frac{2y_2}{2\mu} + \varepsilon_{228} \right). \tag{228}$$

$$\Psi_j(y_1 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \tag{229}$$

$$\Psi_j(y_2 \mid \mathbf{b}, \mathbf{d}) \leq 4^j \left(\frac{2y_2}{2\mu} + 2 \sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{2y_2}{2\mu} + 2 \sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{3y_3}{2\mu} + \varepsilon_{230} \right). \quad (230)$$

$$\Psi_j(y_2 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \quad (231)$$

$$\Psi_j(z_1 \mid \mathbf{b}, \mathbf{d}) \leq 4^j \left(\frac{z_1}{\mu} + 2 \sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{z_1}{2\mu} + 2 \sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{2z_2}{2\mu} + \varepsilon_{232} \right). \quad (232)$$

$$\Psi_j(z_1 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \quad (233)$$

$$(234)$$

Step 1(c). $y_1 = y_2 = 0, z_1 > 0$.

$$\Psi_j(y_1 \mid \mathbf{b}, \mathbf{d}) = \varepsilon_{235}. \quad (235)$$

$$\Psi_j(y_1 \mid |\mathbf{v}|) = \varepsilon_{236}. \quad (236)$$

$$\Psi_j(y_2 \mid \mathbf{b}, \mathbf{d}) \leq 2^j \left(\sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{3y_3}{2\mu} + \varepsilon_{237} \right). \quad (237)$$

$$\Psi_j(y_2 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \quad (238)$$

$$\Psi_j(z_1 \mid \mathbf{b}, \mathbf{d}) \leq 4^j \left(\frac{z_1}{2\mu} + \sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{z_1}{2\mu} + \sum_{k \geq 2} \frac{kz_k}{2\mu} (k-1)^j \frac{2z_2}{2\mu} + \varepsilon_{239} \right). \quad (239)$$

$$\Psi_j(z_1 \mid |\mathbf{v}|) = O\left(2^{3j} \lambda^j e^\lambda j! \left(\frac{\zeta}{N} + \frac{\log^2 N}{\lambda N} \right) \right). \quad (240)$$

Now let $\mathcal{E}_t = \left\{ \zeta(\tau) \leq \log^2 n : 1 \leq \tau \leq t \right\}$. Then let

$$Y_i = \begin{cases} (\zeta(i+1) - \zeta(i)) 1(\mathcal{E}_i) & 0 \leq i \leq T_1 \\ -c_1/2 & T_1 < i \leq n \end{cases}$$

where c_1 is from (81).

Then, q.s.

$$Y_{s+1} + \ldots + Y_t = \zeta(t) - \zeta(s) \text{ for } 0 \leq s < t \leq T_1.$$

For some absolute constant c_2, and with $\theta = \frac{c_1}{100ce^{3ce+1}(3ce)^3}$ and $i \leq L_1$,

$$\mathbb{E}[e^{\theta Y_{s+i}} \mid Y_{s+1}, \ldots, Y_{s+i-1}] = \sum_{k=0}^{\infty} \theta^k \mathbb{E}\left[\frac{Y_{s+i}^k}{k!} \middle| Y_{s+1}, \ldots, Y_{s+i-1}\right]$$
$$\leq 1 - \theta c_1/2 + c_2 \sum_{k=2}^{\infty} \theta^k 2^{3k} \lambda(i)^k e^{\lambda(i)} \leq e^{-\theta c_1/3},$$

where we have used (81) and we have used Lemma 6.3 to bound $\lambda(i)$.

It follows that for $t - s \leq L_1$ and real $u > 0$

$$\mathbb{P}(Y_{s+1} + \cdots + Y_t \geq u) \leq e^{-\theta(u + c_1(t-s)/3)} \tag{241}$$

Suppose now that there exists $\tau \leq T_0$ such that $\zeta(\tau) \geq L_2 = \frac{6K \log n}{c_1}$. Now q.s. there exists $t_1 \leq \tau \leq t_1 + L_1$ such that $\zeta(t_1) = 0$. But then putting $u = -\log n$ in (241) we see that given t_1,

$$\mathbb{P}(\exists t_1 \leq \tau \leq t_1 + L_1 : \zeta(\tau) \geq L_2) \leq \mathbb{P}\left(\neg \bigcap_t \mathcal{E}_t\right) + e^{-\theta(c_1 L_2/3 - \log n)} \leq n^{-K}.$$

□

The number of paths in the output is bounded by the sum of the increases in $y_0 + y_1 + z_1$. If we look at equations (12) etc., then we see that the expected number added to $y_0 + y_1 + z_1$ at step t is $O(\zeta(t)/\mu(t))$. So if $Z_P(t)$ is the number of increases at time t and $Z_P = \sum_{t=0}^{T_3} Z_P(t)$, where T_3 is the time at the beginning of Step 3, then

$$\mathbb{E}[Z_P] = O\left(\mathbb{E}\left(\log n \sum_{t=0}^{T_3} \frac{1}{\mu(t)}\right)\right) = O\left(\log n \; \mathbb{E}\left[\log\left(\frac{\mu(0)}{\mu(T_3)}\right)\right]\right). \tag{242}$$

Now in our case $\mu(T_3) = \Omega(n)$ with probability $1 - o(n^{-2})$ in which case $\mathbb{E}[Z_P] = O(\log n)$. We will apply the Chebyshev inequality to show concentration around the mean. We will condition on $||\mathbf{u}(t) - \hat{\mathbf{u}}(t)||_1 \leq n^{8/9}$ for $t \leq T_1$ (see Lemma 7.1). With this conditioning, the expected value of $Z_P(t)$ is determined up to a factor $1 - O(n^{-1/9} \log^2 n)$ by the value of $\hat{\mathbf{u}}(t)$. In which case, $\mathbb{E}[Z_P(t) \mid Z_P(s)] = (1 + o(1))\mathbb{E}[Z_P(t)]$ and we can apply the Chebychev inequality to show that w.h.p. $Z_P = O(\log n)$. We combine this with Lemma 5.4 to obtain Theorem 1.

10 Hamilton cycles

We will now show how we can use Theorem 1(a) to prove the existence and construction of Hamilton cycles. We will first need to remove a few random edges X from $G = G_{n,cn}^{\delta \geq 3}$ in such a way that the pair $(G - X, X)$ is distributed very close to $(H = G_{n,cn-|X|}^{\delta \geq 3}, Y)$ where Y is a random set of edges disjoint from $E(H)$. In which case we can apply Theorem 1 to H and then we can use the edges of Y to close cycles in the extension-rotation procedure.

10.1 Removing a random set of edges

Let
$$s = n^{1/2}\log^{-2} n$$
and let
$$\Omega = \left\{(H,Y) : H \in \mathcal{G}^{\delta\geq 3}_{n,cn-s}, Y \subseteq \binom{[n]}{2}, |Y| = s \text{ and } E(H)\cap Y = \emptyset\right\}$$
where $\mathcal{G}^{\delta\geq 3}_{n,m} = \{G^{\delta\geq 3}_{n,m}\}$.

We consider two ways of randomly choosing an element of Ω.

(a) First choose G uniformly from $\mathcal{G}^{\delta\geq 3}_{n,cn}$ and then choose an s-set X uniformly from $E(G)\setminus E_3(G)$, where $E_3(G)$ is the set of edges of G that are incident with a vertex of degree 3. This produces a pair $(G-X,X)$. We let $\mathbb{P}_a$ denote the induced probability measure on Ω.

(b) Choose H uniformly from $\mathcal{G}^{\delta\geq 3}_{n,cn-s}$ and then choose an s-set Y uniformly from $\binom{[n]}{2}\setminus E(H)$. This produces a pair (H,Y). We let $\mathbb{P}_b$ denote the induced probability measure on Ω.

The following lemma implies that as far as properties that happen w.h.p. in G, we can use Method (b), just as well as Method (a) to generate our pair (H,Y).

Lemma 10.1. *There exists $\Omega_1 \subseteq \Omega$ such that*

(i) $\mathbb{P}_a(\Omega_1) = 1 - o(1)$.

(ii) $\omega = (H,Y) \in \Omega_1$ *implies that* $\mathbb{P}_a(\omega) = (1+o(1))\mathbb{P}_b(\omega)$.

Proof We first compute the expectation of the number $\mu_3 = \mu_3(G)$ of edges incident to a vertex of degree 3 in G chosen uniformly from $\mathcal{G}^{\delta\geq 3}_{n,cn}$. We will use the random sequence model of Section 3. We will show that μ_3 is highly concentrated in this model and then we can transfer this result to our graph model. Observe first that if ν_3 is the number of vertices of degree 3 in $G_{\mathbf{x}}$ then Lemma 3.3 implies that
$$\left|\nu_3 - \frac{\lambda^3}{3!f_3(\lambda)}n\right| = O(n^{1/2}\log n), \qquad \text{q.s..}$$
Here λ is the solution to $\lambda f_2(\lambda)/f_3(\lambda) = 2cn$.

To see how many edges are incident to these ν_3 vertices we consider the following experiment: Condition on $\nu_3 = \rho n$ where ρ will be taken to be close to $\rho_3 = \frac{\lambda^3}{3!f_3(\lambda)}$. We take a random permutation π of $[2cn]$ and compute the number Z of $i \leq cn$ such that $\{\pi(2i-1),\pi(2i)\}\cap[3\nu_3] \neq \emptyset$. This will give us the number of edges in $G_{\mathbf{x}}$ that are incident with a vertex of degree 3. Now
$$\mathbb{E}[Z] = cn\left(1 - \frac{2cn-3\rho n}{2cn}\frac{2cn-3\rho n-2}{2cn-1}\right) = cn\left(1-\left(1-\frac{3\rho}{2c}\right)^2 + O(1/n)\right).$$

Now interchanging two positions in π can change Z by at most one and so applying the Azuma-Hoeffding inequality for permutations (see for example Lemma 11 of Frieze and Pittel [16] or Section 3.2 of McDiarmid [19]) we see that
$$\mathbb{P}(|Z-\mathbb{E}[Z]| \geq u) \leq e^{-u^2/(cn)} \text{ for any } u \geq 0.$$

Putting this all together we see that with $u = Kn^{1/2}\log n$, for sufficiently large K,

$$\mathbb{P}(|\mu_3(G) - \rho_3(2-\rho_3)cn| \geq u) \leq e^{-u^2/2cn}.$$

Now let

$$\widehat{\mathcal{G}}^{\delta\geq 3}_{n,cn} = \left\{ G \in \mathcal{G}^{\delta\geq 3}_{n,cn} : \left| \mu_3(G) - cn\left(1 - \left(1 - \frac{3\rho_3}{2c}\right)^2\right)\right| \leq Kn^{1/2}\log n \right\}$$

and

$$\Omega_a = \left\{ (H,Y) \in \Omega : H + Y \in \widehat{\mathcal{G}}^{\delta\geq 3}_{n,cn} \right\}.$$

This satisfies requirement (a) of the lemma.

Suppose next that $\omega \in \Omega_a$. Then

$$\mathbb{P}_a(\omega) = \frac{1}{|\mathcal{G}^{\delta\geq 3}_{n,cn}|} \cdot \frac{1}{\binom{cn(1-3\rho_3/2c)^2 \pm Kn^{1/2}\log n}{s}} = \frac{1 + O(\log^{-1} n)}{|\mathcal{G}^{\delta\geq 3}_{n,cn}| \cdot \binom{cn(1-3\rho_3/2c)^2}{s}} \tag{243}$$

$$\mathbb{P}_b(\omega) = \frac{1}{|\mathcal{G}^{\delta\geq 3}_{n,cn-s}|} \cdot \frac{1}{\binom{\binom{n}{2} - cn}{s}} \tag{244}$$

One can see from this that one has to estimate the ratio $|\mathcal{G}^{\delta\geq 3}_{n,cn}|/|\mathcal{G}^{\delta\geq 3}_{n,cn-s}|$. For this we make estimates of

$$M = |\left\{ (G_1, G_2) \in \mathcal{G}^{\delta\geq 3}_{n,cn} \times \mathcal{G}^{\delta\geq 3}_{n,cn-s} : E(G_1) \supseteq E(G_2) \right\}|.$$

We have the following inequalities:

$$|\widehat{\mathcal{G}}^{\delta\geq 3}_{n,cn}| \binom{cn(1-3\rho_3/2c)^2 - Kn^{1/2}\log n}{s} \leq M \leq |\widehat{\mathcal{G}}^{\delta\geq 3}_{n,cn}| \binom{cn(1-3\rho_3/2c)^2 + Kn^{1/2}\log n}{s} + |\mathcal{G}^{\delta\geq 3}_{n,cn}| \sum_{|u|\geq Kn^{1/2}\log n} \binom{cn(1-3\rho_3/2c)^2 + u}{s} e^{-u^2/2cn} \tag{245}$$

$$M = |\mathcal{G}^{\delta\geq 3}_{n,cn-s}| \binom{\binom{n}{2} - cn}{s}. \tag{246}$$

We get (245) by summing $\mu_3(G_1)$ over $G_1 \in \mathcal{G}^{\delta\geq 3}_{n,cn}$ and bounding $\mu_3(G_1)$ according to whether or not G is in $\widehat{\mathcal{G}}^{\delta\geq 3}_{n,cn}$. Equation (246) is obtained by summing over $G_2 \in \mathcal{G}^{\delta\geq 3}_{n,cn-s}$, the number of ways of adding s edges to G_2.

Now

$$\sum_{|u|\geq Kn^{1/2}\log n} \binom{cn(1-3\rho_3/2c)^2 + u}{s} e^{-u^2/2cn} \leq 2 \sum_{u\geq Kn^{1/2}\log n} \binom{cn(1-3\rho_3/2c)^2}{s} e^{O(us/n)} e^{-u^2/2cn}$$
$$\leq 2\binom{cn(1-3\rho_3/2c)^2}{s} \sum_{u\geq Kn^{1/2}\log n} e^{-u^2/3cn} = O\left(\binom{cn(1-3\rho_3/2c)^2}{s} e^{-\Omega(\log^2 n)}\right).$$

It follows from this and (245) that

$$M = |\mathcal{G}^{\delta\geq 3}_{n,cn}| \binom{cn(1-3\rho_3/2c)^2}{s} \left(1 + O(\log^{-1} n)\right).$$

By comparing with (246) we see that

$$\frac{|\mathcal{G}^{\delta\geq 3}_{n,cn}|}{|\mathcal{G}^{\delta\geq 3}_{n,cn-s}|} = (1+o(1))\frac{\binom{\binom{n}{2}-cn}{s}}{\binom{cn(1-3\rho_3/2c)^2}{s}}.$$

The lemma follows by using this in conjunction with (243) and (244). □

10.2 Connectivity of $G^{\delta\geq 3}_{n,cn}$

Lemma 10.2. $G^{\delta\geq 3}_{n,cn}$ *is connected, w.h.p..*

Proof It follows from Lemma 10.1 that we can replace $G^{\delta\geq 3}_{n,cn}$ by $G^{\delta\geq 3}_{n,cn-s}$ plus s random edges. We use the random sequence model to deal with $G^{\delta\geq 3}_{n,cn-s}$. Let Fix $4 \leq k \leq n/\log^{20} n$. For $K \subseteq [n]$, $e(K)$ denotes the number of edges of $G_{\mathbf{x}}$ contained in K. Let $\ell_0 = \log n/(\log\log n)^{1/2}$. Then with λ the solution to $\lambda f_2(\lambda)/f_3(\lambda) = 2c$,

$$\mathbb{P}(\exists K \subseteq [n] : e(K) \geq 5k/4) \leq o(1) + \delta_k \binom{n}{k} \sum_{d=3k/2}^{\ell_0 k} \frac{\lambda^d k^d}{d! f_3(\lambda)^k} \binom{cn}{5k/4} \left(\frac{d}{cn}\right)^{5k/2} \tag{247}$$

$$\leq \delta_k \sum_{d=3k/2}^{\ell_0 k} \left(\frac{\lambda e k}{d}\right)^d \left(\frac{e^{9/4}\ell_0^{5/2}k^{1/4}}{(5c/4)^{5/4} f_3(\lambda) n^{1/4}}\right)^k \leq \delta_k \ell_0 k \left(\frac{e^{9/4}\ell_0^{5/2}k^{1/4}e^{\lambda}}{(5c/4)^{5/4} f_3(\lambda) n^{1/4}}\right)^k. \tag{248}$$

Explanation of (247)**:** Here $\delta_k = 1 + o(1)$ for $k \leq \log^2 n$ and $O(n^{1/2})$ for larger k. The term $\frac{\lambda^d k^d}{d! f_3(\lambda)^k}$ bounds the probability that the total degree of K is d, see (82). Given the degree sequence we take a random permutation π of the multi-set $\{d_{\mathbf{x}}(j) \times j : j \in [n]\}$ and bound the probability that there is a set of $5k/4$ indices i such that $\pi(2i-1), \pi(2i) \in K$. This expression assumes that vertex degrees are independent random variables. We can always inflate the estimate by $O(n^{1/2})$ to account for the degree sum being fixed. This is what δ_k does for $k > \log^2 n$. For smaller k we use (3). The bound of $d \leq \ell_0 k$ arises from Lemma 3.2(b).

Let σ_k denote the RHS of (248). Then, we have $\sum_{k=4}^{n/\log^{20} n} \sigma_k = o(1)$ and we can assume that there is no set of size $4 \leq k \leq n/\log^{20} n$ containing at least $5k/4$ edges.

But if G has minimum degree at least 3 and a set K contains at most $5|K|/4$ edges then there must be edges with one end in K. So, we see that w.h.p. the minimum component size in G will be at least $n/\log^{20} n$. We now use the result of Section 10.1. If we take $H = G^{\delta\geq 3}_{n,cn-s}$, $s = n^{1/2}\log^{-2} n$ then we know by the above that w.h.p. it only has components of size at least $n/\log^{20} n$. Now add s random edges Y. Then

$$\mathbb{P}(H + Y \text{ is not connected}) = o(1) + \log^{40} n \left(1 - \frac{1}{\log^{40} n}\right)^s = o(1).$$

Now apply Lemma 10.1. □

10.3 Extension-Rotation Argument

We will as in Section 10.2 replace $G^{\delta\geq 3}_{n,cn}$ by $G^{\delta\geq 3}_{n,cn-s}$ plus s random edges Y. Having run 2GREEDY we will w.h.p. have a 2-matching M_0 say such that M_0 has $O(\log n)$ components.

The main idea now of course is that of a *rotation*. Given a path $P = (u_1, u_2, \ldots, u_k)$ and an edge (u_k, u_i) where $i \leq k-2$ we say that the path $P' = (u_1, \ldots, u_i, u_k, u_{k-1}, \ldots, u_{i+1})$ is obtained from P by a rotation. u_1 is the *fixed* endpoint of this rotation. We now describe an algorithm, EXTEND-ROTATE that w.h.p. converts M_0 into a Hamilton cycle in $O(n^{1.5+o(1)})$ time.

Given a path P with endpoints a, b we define a *restricted rotation search* $RRS(\nu)$ as follows: Suppose that we have a path P with endpoints a, b. We start by doing a sequence of rotations with a as the fixed endpoint. Furthermore

R1 We only do a rotation if the endpoint of the path created is not an endpoint of the paths that have been created so far.

R2 We stop this process when we have either (i) created ν endpoints or (ii) we have found a path Q with an endpoint that has a neighbor outside Q. We say that we have found an *extension*.

Let $END(a)$ be the set of endpoints, other than a, produced by this procedure. The main result of [15] is that w.h.p., regardless of our choice of path P, either (i) we find an extension or (ii) we are able to generate $n^{1-o(1)}$ endpoints. We will run this procedure with $\nu = n^{3/4} \log^3 n$.

Assuming that we did not find an extension and having constructed $END(a)$, we take each $x \in END(a)$ in turn and starting with the path P_x that we have found from a to x, we carry out R1,R2 above with x as the fixed endpoint and either find an extension or create a set of ν paths with x as one endpoint and the other endpoints comprising a set $END(x)$ of size ν.

It follows from [2] that the above construction $RSS(\nu)$ can be carried out in $O(\nu^2 \log n)$ time.

Algorithm EXTEND-ROTATE

Step 1 Choose a path component P of the current 2-matching M, with endpoints a, b or if M is not a Hamilton cycle, choose a cycle C of M and delete an edge to create P. We choose P as large as possible here and note that P will have length $\Omega(n/\log n)$.

Step 2 Carry out $RSS(\nu)$ until either an extension is found or we have constructed $\nu + 1$ endpoint sets.

Case a: We find an extension. Suppose that we construct a path Q with endpoints x, y such that y has a neighbor $z \notin Q$.

(i) If z lies in a cycle C then let R be a path obtained from C by deleting one of the edges of C incident with z. Let now $P = x, Q, y, z, R$ and go to Step 1.

(ii) If $z = u_j$ lies on a path $R = (u_1, u_2, \ldots, u_k)$ where the numbering is chosen so that $j \geq k/2$ then we let $P = x, Q, y, z, u_{j-1}, \ldots, u_1$ and go to Step 1.

Case b: If there is no extension then we search for an edge $e = (p, q) \in Y$ such that $p \in END(a)$ and $q \in END(p)$. if there is no such edge then the algorithm fails. If there is such an edge, consider the cycle $P + e$. Now either C is a Hamilton cycle and we are done, or else there is a vertex $u \in C$ and a vertex $v \notin C$ such that (u, v) is an edge of H, assuming that H is connected, see Lemma 10.2. We now delete one of the edges, (u, w) say, of C incident with u to create a path Q from w to u and treat e as an extension of this path. We can now proceed as in Case a.

10.3.1 Analysis of EXTEND-ROTATE

We first bound the number of executions of $RSS(\nu)$. Suppose that M_0 has $\kappa \leq K_1 \log n$ components for some $K_1 > 0$. Each time we execute Step 2, we either reduce the number of components by one or we halve the size of one of the components not on the current path. So if the component sizes of M_0 are $n_1, n_2, \ldots, n_\kappa$ then the number of executions of Step 2 can be bounded by

$$\kappa + \sum_{i=1}^{\kappa} \log_2 n_i = O(\log^2 n).$$

An execution of Step 2 takes $O(\nu^2 \log n)$ time and so we are within the time bound claimed by Theorem 2.

We next argue that EXTEND-ROTATE succeeds w.h.p.. Suppose that the edges of Y are $e_1, e_2, \ldots, e_s$. We can allow the algorithm to access these edges in order, never going back to a previously examined edge. The probability that an e_i can be used in Case b is always at least $\frac{\binom{\nu}{2}-s}{\binom{n}{2}} \geq \frac{\log^6 n}{2n^{1/2}}$ (we have subtracted s because some of the useful edges might have been seen before the current edge in the order). So the probability of failure is bounded by the probability that the binomial $Bin\left(s, \frac{\log^6 n}{2n^{1/2}}\right)$ is less than $K_2 \log^2 n$ for some $K_2 > 0$. And this tends to zero. This completes the proof of Theorem 2.

11 Concluding remarks

The main open question concerns what happens when $c < 10$. Is it true that (194) holds all the way down to $c > 3/2$? We have done some numerical experiments and here are some results from these experiments:

c	y_{final}	z_{final}	μ_{final}	λ_{final}
3.0	0.000009	0.401133	0.626647	2.351035
2.5	0.000010	0.271633	0.362381	1.563165
2.0	0.000010	0.114914	0.131391	0.761156
1.9	0.000010	0.085062	0.094501	0.604995
1.8	0.000010	0.057356	0.061974	0.451806
1.7	0.000010	0.031910	0.033543	0.298630
1.6	0.000010	0.033080	0.034737	0.292363
1.55	0.000008	0.213928	0.264772	0.372103

These are the results of running Euler's method with step length 10^{-5} on the sliding trajectory (152). They indicate that (194) holds down to somewhere close to 2.0. The increase in z_{final} towards the end may just be due to numerical errors? Could there be some sort of phase transition in the performance of 2GREEDY at around this point. There is one for the Karp-Sipser matching algorithm and so we are led to conjecture there is one here too.

Can we prove anything for $c < 10$? At the moment we can not even show that at the completion of 2GREEDY the 2-matching M has $o(n)$ components. This will be the subject of further research. We recently showed that 2GREEDY only produces $O(n^{1/5+o(1)})$ components w.h.p. when run on a random cubic graph, see Bal, Bennet, Bohman and Frieze [4].

Finally, we mention once again, the possible use of the ideas of [11] to reduce the running time of our Hamilton cycle algorithm to $O(n^{1+o(1)})$ time.

Our list of problems/conjectures arising from this research can thus be summarised:

(a) Find a threshold c_1 such that 2GREEDY produces a 2-matching in $G_{n,cn}^{\delta\geq 3}$ with $O(\log n)$ components w.h.p. iff $c > c_1$. Is $c_1 = 3/2$?

(b) If $c_1 > 3/2$ then show that when $c \in (3/2, c_1)$, the number of components in the 2-matching produced is $O(n^\alpha)$ for some constant $\alpha < 1$. The value of α should be $1/5 + o(1)$?

(c) Analyse the performance of 2GREEDY on the random graph $G_{n,cn}$ i.e. do not condition on degree at least three. Is there a threshold c_2 such that if $c \leq c_2$ then w.h.p. only Steps 1a,1b,1c are needed, making the matching produced optimal.

(d) Can 2GREEDY be used to find a Hamilton cycle w.h.p. in $O(n^{1+o(1)})$ time when applied to $G_{n,cn}^{\delta\geq 3}$ and c sufficiently large?

(e) How much of this can be extended to find edge disjoint Hamilton cycles in $G_{n,cn}^{\delta\geq k}$ for $k \geq 4$.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, The first occurrence of Hamilton cycles in random graphs, *Annals of Discrete Mathematics* 27 (1985) 173–178.

[2] D. Angluin and L.G. Valiant, Fast probabilistic algorithms for Hamilton circuits and matchings, *Journal of Computer and System Sciences* 18 (1979) 155–193.

[3] J. Aronson, A.M. Frieze and B.G. Pittel, Maximum matchings in sparse random graphs: Karp-Sipser re-visited, *Random Structures and Algorithms* 12 (1998) 111–178.

[4] D. Bal, P. Bennett, T. Bohman and A.M. Frieze, A greedy algorithm for finding a large 2-matching on a random cubic graph, submitted for publication.

[5] T. Bohman and A.M. Frieze, Hamilton cycles in 3-out, *Random Structures and Algorithms* 35 (2009) 393–417.

[6] B. Bollobás, The evolution of sparse graphs, *In: Graph Theory and Combinatorics, Proc. Cambridge Combin. Conf. in honour of Paul Erdős (B. Bollobás, ed.), Academic Press* (1984) 35–57.

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

[8] B. Bollobás, C. Cooper, T.I. Fenner and A.M.Frieze, On Hamilton cycles in sparse random graphs with minimum degree at least k, *Journal of Graph Theory* 34 (2000) 42–59.

[9] B. Bollobás, T.I.Fenner and A.M.Frieze, An algorithm for finding Hamilton paths and cycles in random graphs, *Combinatorica* 7 (1987), 327–341.

[10] B. Bollobás and A.M.Frieze, On matchings and Hamiltonian cycles in random graphs, *Annals of Discrete Mathematics* 28 (1985) 23–46.

[11] P. Chebolu, A.M. Frieze and P. Melsted, Finding a Maximum Matching in a Sparse Random Graph in $O(n)$ Expected Time JACM 57, (2010)

[12] V. Chvátal, Almost all graphs with $1.44n$ edges are 3-colourable, *Random Structures and Algorithms* 2 (1991) 11–28.

[13] A.M. Frieze, Finding Hamilton cycles in sparse random graphs, *Journal of Combinatorial Theory B* 44 (1988) 230–250.

[14] A. M. Frieze and T. Łuczak, Hamiltonian cycles in a class of random graphs: one step further, in *Proceedings of Random Graphs '87*, Edited by M. Karonski, J. Jaworski and A. Rucinski, John Wiley and Sons, 53–59.

[15] A.M. Frieze and B. Pittel, On a sparse random graph with minimum degree three: Likely Posa's sets are large.

[16] A.M. Frieze and B. Pittel, Perfect matchings in random graphs with prescribed minimal degree, *Trends in Mathematics*, Birkhauser Verlag, Basel (2004) 95–132.

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

[18] J. Komlós and E. Szemerédi, *Limit distributions for the existence of Hamilton circuits in a random graph*, Discrete Mathematics 43 (1983), 55–63.

[19] C. McDiarmid, Concentration, *Probabilistic Methods for Algorithmic Discrete Mathematics* (M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, B. Reed eds.), Springer, Berlin (1998) 1–46.

[20] B. McKay, Asymptotics for 0-1 matrices with prescribed line sums, in *Enumeration and Design*, (Academic Press, 1984) 225-238.

[21] B. Pittel, On tree census and the giant component in sparse random graphs, *Random Structures and Algorithms* 1 (1990), 311–342.

[22] R.W. Robinson and N.C. Wormald, Almost all cubic graphs are Hamiltonian. *Random Structures and Algorithms* **3** (1992) 117–125.

[23] R.W. Robinson and N.C. Wormald, Almost all regular graphs are Hamiltonian. *Random Structures and Algorithms* **5** (1994) 363–374.

A Proof of (3)

To find a sharp estimate for the probabilities in (3) we have to refine a bit the proof of the local limit theorem, since in our case the variance of the Z_j are not always bounded away from zero.

However it is enough to consider the case where $N\sigma^2 \to \infty$. There is little loss of generality in assuming that $D = 0$ here. As usual, we start with the inversion formula

$$\begin{aligned}\mathbb{P}\left(\sum_{j=1}^{N} Z_j = \tau\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\tau x}\mathbb{E}\left(e^{ix\sum_{j=1}^{N} Z_j}\right) dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\tau x}\prod_{\ell=2}^{3}\left[\mathbb{E}(e^{ix\mathcal{P}_\ell})\right]^{N_\ell} dx, \qquad (249)\end{aligned}$$

where $\tau = 2M - k$. Consider first $|x| \geq (N\lambda)^{-5/12}$. Using an inequality (see Pittel [21])

$$|f_\ell(\eta)| \leq e^{(\mathrm{Re}\eta - |\eta|)/(\ell+1)} f_\ell(|\eta|),$$

we estimate

$$\begin{aligned}&\frac{1}{2\pi}\int_{|x|\geq(N\lambda)^{-5/12}} \left| e^{-i\tau x}\prod_{\ell=2}^{3}\left(\frac{f_\ell(e^{ix}\lambda)}{f_\ell(\lambda)}\right)^{N_\ell}\right| dx \\ &\leq \frac{1}{2\pi}\int_{|x|\geq(N\lambda)^{-5/12}} e^{N\lambda(\cos x - 1)/4}\, dx \\ &\leq e^{N\lambda[(\cos((N\lambda)^{-5/12}) - 1)/4]} \\ &\leq e^{-(N\lambda)^{1/6}/9}. \qquad (250)\end{aligned}$$

For $|x| \leq (N\lambda)^{-5/12}$, putting $\eta = \lambda e^{ix}$ and using

$$\sum_{\ell=2}^{3}\frac{N_\ell \lambda f_\ell'(\lambda)}{f_\ell(\lambda)} = 2M \text{ and } d/dx = i\eta d/d\eta$$

we expand $\sum_{\ell=2}^{3} N_\ell \log\left(\frac{f_\ell(\eta)}{f_\ell(\lambda)}\right)$ as a Taylor series around $x = 0$ to obtain

$$\begin{aligned}-i\tau x + \sum_{\ell=2}^{3} N_\ell \log\left(\frac{f_\ell(e^{ix}\lambda)}{f_\ell(\lambda)}\right) &= ikx - \frac{x^2}{2}\,\mathcal{D}\left(\sum_{\ell=2}^{3} N_\ell \frac{\eta f_\ell'(\eta)}{f_\ell(\eta)}\right)\Bigg|_{\eta=\lambda} \\ &\quad - \frac{ix^3}{3!}\,\mathcal{D}^2\left(\sum_{\ell=2}^{3} N_\ell \frac{\eta f_\ell'(\eta)}{f_\ell(\eta)}\right)\Bigg|_{\eta=\lambda} \\ &\quad + O\left(x^4\,\mathcal{D}^3\left(\sum_{\ell=2}^{3} N_\ell \frac{\eta f_\ell'(\eta)}{f_\ell(\eta)}\right)\Bigg|_{\eta=\tilde{\eta}}\right). \qquad (251)\end{aligned}$$

Here $\tilde{\eta} = \lambda e^{i\tilde{x}}$, with $\tilde{x}$ being between 0 and x, and $\mathcal{D} = \eta(d/d\eta)$. Now, the coefficients of $x^2/2$, $x^3/3!$ and x^4 are $N\sigma^2$, $O(N\sigma^2)$, $O(N\sigma^2)$ respectively, and σ^2 is of order λ. (Use (2) and consider the effect of $\mathcal{D}$ on a power of η.) So the second and the third terms in (251) are $o(1)$ uniformly for $|x| \leq (N\lambda)^{-5/12}$. Therefore

$$\frac{1}{2\pi}\int_{|x|\leq(N\lambda)^{-5/12}} = \int_1 + \int_2 + \int_3, \qquad (252)$$

where

$$
\begin{aligned}
\int_1 &= \frac{1}{2\pi}\int_{|x|\le(N\lambda)^{-5/12}} e^{ikx-N\sigma^2x^2/2}\,dx \\
&= \frac{1}{\sqrt{2\pi N\sigma^2}} + O\left(\frac{k^2+1}{(\lambda N)^{3/2}}\right), && (253) \\
\int_2 &= O\left(\mathcal{D}^2\left(\sum_{\ell=2}^{3} N_\ell \frac{\lambda f'_\ell(\lambda)}{f_\ell(\lambda)}\right)\int_{|x|\le(N\lambda)^{-5/12}} x^3 e^{-N\sigma^2x^2/2}\,dx\right) \\
&= O\left(N\lambda\int_{|x|\le(N\lambda)^{-5/12}} |x|^3 e^{-N\sigma^2x^2/2}\,dx\right) \\
&= O(e^{-\alpha(N\lambda)^{1/6}}), && (254)
\end{aligned}
$$

($\alpha > 0$ is an absolute constant), and

$$
\begin{aligned}
\int_3 &= O\left(N\lambda\int_{|x|\le(N\lambda)^{-5/12}} x^4 e^{-N\sigma^2x^2/2}\,dx\right) \\
&= o\left(\int_2\right). && (255)
\end{aligned}
$$

Using (249)-(255), we arrive at

$$
\mathbb{P}\left(\sum_\ell Z_\ell = \tau\right) = \frac{1}{\sqrt{2\pi N\sigma^2}} \times \left(1 + O\left(\frac{k^2+1}{N\lambda}\right)\right).
$$

B Mathematica Output

In the computations below, $\varepsilon_p(\hat{\lambda})$ is represented by $ep[x]$ and α_p is represented by Ap and β is represented by B. The computation C_1 is the evaluation of the RHS of (218).

f0[x_]:=Exp[x]

f1[x_]:=f0[x] − 1

f2[x_]:=f1[x] − x

f3[x_]:=f2[x] − $\frac{x^2}{2}$

e1[x_]:=$\frac{\text{f2}[x]}{\text{f3}[x]}$ − 1

e2[x_]:=$\frac{\text{f0}[x]}{\text{f2}[x]}$ − 1

e3[x_]:=$\frac{\text{f0}[x]}{\text{f3}[x]}$ − 1

eA[x_]:=$\frac{(1+\text{e2}[x])(1+\text{e3}[x])x^3}{8\text{f0}[x]}$

eB[x_]:=$\frac{x^2(1+\text{e2}[x])^2}{\text{f0}[x]}$

f[x_]:=Max $\left[\text{e1}[x], \text{eA}[x] + \text{eB}[x] + \frac{\text{e1}[x]}{1+\text{e1}[x]}, \text{eB}[x] + \text{eA}[x]\right]$

$T = 1 - \frac{1}{2^{1/2}\text{Exp[Pi/4]}}$

$1 - \frac{e^{-\pi/4}}{\sqrt{2}}$

$B = -.01 + 2(1 - T)$

0.634794

$\text{A0} = B(\text{Exp}[2T/B] - 1)$

4.73302

$\text{A1} = N\left[\frac{2(1+T)\text{Exp}[-\text{ArcTan}[T]]}{(1+T^2)^{1/2}}\right]$

1.5312

$\text{A2} = N\left[\frac{(1+T)\text{Exp}[-\text{ArcTan}[T]]}{(1+T^2)^{1/2}}\text{Integrate}\left[\frac{2\text{Exp}[\text{ArcTan}[x]]}{(1+x)(1+x^2)^{1/2}}, \{x, 0, 1\}\right]\right]$

1.41846

$M[\text{c\_}]{:=}\text{Abs}[\text{A1}c - \text{A2}]$

$\text{A3}[\text{c\_}, \text{x\_}]{:=}\frac{10\text{A0}c}{B^2} + \frac{8\text{A0}^2 c f[x]}{B^3} + \frac{4c}{B}$

$\text{A4}[\text{c\_}, \text{x\_}]{:=}\text{A0A3}[c, x]/2$

$\text{A5}[\text{c\_}, \text{x\_}]{:=}\frac{2c}{B} + \frac{2\text{A4}[c,x]}{B} + \frac{4c\text{A0}}{B^2}$

$\text{A5}[10, 13.5]$

9770.23

$f[13.5]$

0.000796502

$\text{C1}[\text{c\_}, \text{x\_}]{:=}M[c] - \text{A5}[c, x]f[x]$

$\text{C1}[10, M[10]]$

8.25371