LAYERED PERCOLATION ON THE COMPLETE GRAPH

LINCOLN CHAYES AND S. ALEX SMITH

ABSTRACT. We present a generalized (multitype) version of percolation on the complete graph, in which we divide the vertices of the graph into a fixed number of sets (called *layers*) and perform percolation where the probability of $\{u, v\}$ being in our edge set depends on the respective layers of u and v. Many results analogous to usual percolation on the complete graph are determined. In addition, we determine the exponential rate function for the probability that a giant component occupies a fixed fraction of the graph, while all other components are small. We also determine the exponential rate function for the probability that a particular exploration process on the random graph will discover a certain fraction of vertices in each layer, without encountering a giant component.

1. INTRODUCTION

The principle object of study in this paper is a generalized version of the Erdős-Rènyi random graph. The generalization is that instead of percolation occuring homogenously throughout the graph, we divide the graph into "layers", and perform percolation which depends on the layers of the involved vertices. For an examination of equivalent results on the non-layered complete graph, and for motivating methodology – such as it exists – we point the reader to [2].

For the purposes of this paper, we define a *layered* set to be a set in which each element v has an associated integer, which we designate by $\mathrm{layer}(v)$. We similarly define a layered graph to be a graph whose vertex set is a layered set. Since we think of $\mathrm{layer}(v)$ as having something to do with the position of v, we will freely use language such as "v is in the i^{th} layer" to mean that $\mathrm{layer}(v) = i$. If S is a layered set, we let $S_\ell = \{v \in S : \mathrm{layer}(v) = \ell\}$ be the set of elements of S in layer ℓ. If G is a layered graph with vertex set V, we let G_ℓ be the subgraph of G restricted to V_ℓ.

We will typically use L to indicate the number of layers in a graph and we will use the hat symbol ($\hat{}$) to indicate vectors of such length. If $\hat{\eta}$ is such a vector, we will let $|\hat{\eta}|$ be the $\mathscr{L}^1$ norm on $\hat{\eta}$. If S is a set or a graph, we let $|S|$ be the number of elements or vertices in S, respectively. We will let $\langle S \rangle$ be the vector of dimension L whose ℓ^{th} component is $|S_\ell|$. We will also use the componentwise partial ordering on vectors, so (for example) $\hat{\eta} > 0$ will mean that $\hat{\eta} \in (0, \infty)^L$. If v is a vertex in a graph G, we will let $C(v) = C_G(v)$ be the component of G containing v.

We now introduce the main model of the paper: Let $L \in \mathbb{N}$ and $\hat{n} \in \mathbb{N}^L$ be given, and consider the layered vertex set

$$\mathscr{V} = \{(\ell, k) : 1 \le \ell \le L, 1 \le k \le \hat{n}_L\}, \tag{1.1}$$

where $\mathrm{layer}\big((\ell, k)\big) = \ell$. Then given probabilities $(p_{ij})_{i,j \in \{1,\dots,L\}}$ with $p_{ij} = p_{ji}$ for all i and j, let $\mathscr{E}$ be the (random) edge set so that each edge $\{u, v\}$ appears (or not) in $\mathscr{E}$ independently, and $P(\{u, v\} \in \mathscr{E}) = p_{\mathrm{layer}(u)\,\mathrm{layer}(v)}$, so that the probability that an edge exists between given vertices in layers i and j is p_{ij}. In this paper we will restrict ourselves to the case in which the number of vertices in each layer scales proportionally to the number of vertices in every other layer; so let us for the remainder of the paper take $\hat{n} = (\lfloor \hat{\rho}_1 n \rfloor, \ldots, \lfloor \hat{\rho}_L n \rfloor)$ for some $\hat{\rho} \in (0, \infty)^L$ and $n > 0$. Additionally, let us take $\mathrm{A} = (\alpha_{ij})$ to be a symmetric, non-negative, irreducible $L \times L$ matrix, and let $p_{ij} = \frac{\alpha_{ij}}{n}$. We will generally consider L and A to be fixed throughout the paper and therefore we designate the resulting random graph by $\mathscr{G}(n, \hat{\rho})$ and the corresponding probability measure by $P_{n,\hat{\rho}}$ – although we will allow n and $\hat{\rho}$ to be implicit when it is clear from context. We let $E_{n,\hat{\rho}}$ indicate expectation with respect to $P_{n,\hat{\rho}}$.

2. Main Results

We will start with a mean field approach, which will give us quick insight into the size of giant components in $\mathscr{G}(n, \hat{\rho})$. Intuitively, the probability of a vertex $v \in \mathscr{V}$ not being in a "large" component of $\mathscr{G}(n, \hat{\rho})$ is approximately the probability that none of v's neighbors are in "large" components in $\mathscr{G}(n, \hat{\rho})$ restricted to $\mathscr{V} \setminus \{v\}$. If v is in layer k, then the number of neighbors of v in layer i is distributed like $\mathrm{Bin}\big(\lfloor \hat{\rho}_i n \rfloor - \delta_{ik}, \frac{\alpha_{ik}}{n}\big)$, which converges to $\mathrm{Poisson}(\alpha_{ik}\hat{\rho}_i)$ as $n \to \infty$. Thus letting the random vector $\hat{\theta}$ be the fraction of sites in each layer of $\mathscr{G}(n, \hat{\rho})$ which belong to "large" components, we expect

$$1 - \hat{\theta}_k \approx \prod_{i=1}^{L} \left[\sum_{j=0}^{\infty} \frac{(\alpha_{ik}\hat{\rho}_i)^j \mathrm{e}^{-\alpha_{ik}\hat{\rho}_i}}{j!} (1 - \hat{\theta}_i)^j \right]. \tag{2.1}$$

Pleasantly, this intuition is correct and, moreover, (2.1) simplifies to

$$\hat{\theta}_k = 1 - \exp\left(- \sum_{i=1}^{L} \alpha_{ik}\hat{\rho}_i\hat{\theta}_i \right), \tag{2.2}$$

which leads us to the main theorem of the paper:

Theorem 2.1. *Let $\hat{\theta}(r, n, \hat{\rho})$ be the (random) fraction of sites in each layer of $\mathscr{G}(n, \hat{\rho})$ which are in components of size greater than r. Then there exists a $\hat{\theta}^\star(\hat{\rho})$ so that*

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \hat{\theta}(\epsilon n, n, \hat{\rho}) = \hat{\theta}^\star(\hat{\rho}) \tag{2.3}$$

almost surely. Furthermore, $\hat{\theta}^\star(\hat{\rho})$ is the maximum solution to the system of equations given by (2.2).

There may be some question as to what Theorem 2.1 means by the maximum solution, since we are using only a partial ordering on $[0, \infty)^L$. It turns out that what we mean is "the nonzero solution if there is one, and zero otherwise". To wit:

Lemma 2.2. *The system of equations given by* (2.2) *has a maximal solution in $\mathbb{R}^L$ for all $\hat{\rho} > 0$, which is either zero or strictly positive. This maximal solution is an increasing function of $\hat{\rho}$. If a strictly positive solution to* (2.2) *exists, it is unique.*

We will say that $\hat{\rho}$ is a *supercritical* density if $\hat{\theta}^\star(\hat{\rho})$ is nonzero, a *subcritical* density if there is a neighborhood of $\hat{\rho}$ in $(0, \infty)^L$ on which $\hat{\theta}^\star$ is zero, and a *critical* density otherwise. In order to characterize these three regions of density further, we introduce a susceptibility matrix, whose ijth entry is the expected number of edges from a vertex in layer i to vertices in layer j (up to $O(1/n)$ corrections):

$$\mathrm{B}_{\hat\rho} = \begin{bmatrix} \alpha_{11}\hat\rho_1 & \alpha_{12}\hat\rho_1 & \cdots & \alpha_{1L}\hat\rho_1 \\ \alpha_{21}\hat\rho_2 & \alpha_{22}\hat\rho_2 & \cdots & \alpha_{2L}\hat\rho_2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{L1}\hat\rho_L & \alpha_{L2}\hat\rho_L & \cdots & \alpha_{LL}\hat\rho_L \end{bmatrix}. \tag{2.4}$$

We then have:

Lemma 2.3. *Let $\kappa_0(\hat\rho)$ be the maximal eigenvalue of $\mathrm{B}_{\hat\rho}$. Then $\hat\rho$ is subcritical if $\kappa_0(\hat\rho) < 1$, critical if $\kappa_0(\hat\rho) = 1$, and supercritical if $\kappa_0(\hat\rho) > 1$.*

Lemma 2.3 has a quite intuitive interpretation: Let $\hat\eta$ be the (strictly positive, as per Perron-Frobenius) eigenvector of $\mathrm{B}_{\hat\rho}$ corresponding to $\kappa_0(\hat\rho)$. Then if we taken a set $\mathscr{W} \subset \mathscr{V}$ with $\langle\mathscr{W}\rangle \approx \hat\eta$, the expected number of neighbors of $\mathscr{W}$ in each layer of $\mathscr{G}(n, \hat\rho)$ is approximately $\kappa_0(\hat\rho)\hat\eta$. Thus if $\kappa_0(\hat\rho) < 1$, we expect any exploration process of $\mathscr{G}(n, \hat\rho)$ to die off at an exponential rate. Conversely, if $\kappa_0(\hat\rho) > 1$ we expect that the exploration will survive with positive probability – unless the process grows to the point where these approximations are no longer valid. In Section 4, we will use this idea to prove:

Theorem 2.4. *Let L, A, and $\hat\rho > 0$ be given. Let $\hat c \geq 0$ be given with $|\hat c| = 1$, and for each n let v be a random vertex of $\mathscr{G}(n, \hat\rho)$, independent from $\mathscr{E}$, with $P_{n,\hat\rho}(\mathrm{layer}(v) = i) = \hat c_i$. Then if $\hat\rho$ is subcritical,*

$$\lim_{n\to\infty} E_{n,\hat\rho}[\langle C(v)\rangle] = (\mathrm{I} - \mathrm{B}_{\hat\rho})^{-1}\hat c. \tag{2.5}$$

If $\hat\rho$ is critical then

$$\lim_{n\to\infty} E_{n,\hat\rho}[|C(v)|] = \infty, \tag{2.6}$$

but

$$\lim_{n\to\infty} P_{n,\hat\rho}\big(|C(v)| > \epsilon n\big) = 0 \tag{2.7}$$

for all $\epsilon > 0$. If $\hat\rho$ is supercritical then there is some $\epsilon > 0$ such that

$$\liminf_{n\to\infty} P_{n,\hat\rho}\big(|C(v)| > \epsilon n\big) > 0. \tag{2.8}$$

Note that although we state the theorem in a general form, for this paper we will always take v to be a uniformly chosen vertex of $\mathscr{G}(n, \hat\rho)$, and thus will take $\hat c = \hat\rho/|\hat\rho|$.

To state the next theorem, we let S be the entropy function given by

$$S(\hat\eta, \hat\rho) = \sum_{k=1}^{L} \big(\hat\rho_k \log\hat\rho_k - \hat\eta_k \log\hat\eta_k - (\hat\rho_k - \hat\eta_k)\log(\hat\rho_k - \hat\eta_k)\big), \tag{2.9}$$

and let

$$\Xi(\hat\eta) = \sum_{k=1}^{L} \hat\eta_k \log\left(1 - \mathrm{e}^{-\sum_{i=1}^{L}\alpha_{ik}\hat\eta_i}\right). \tag{2.10}$$

With these, we can calculate the large deviation rate of large components existing in $\mathscr{G}(n, \hat\rho)$:

Theorem 2.5. *Let L, A, and $\hat\rho$ be given, and for each n let v be a random vertex of $\mathscr{G}(n,\hat\rho)$ chosen by any distribution desired which is independent from $\mathscr{E}$. Also let $\mathscr{S}\subset\prod_i[0,\hat\rho_i]$ be a set of densities which is open in $\prod_i[0,\hat\rho_i]$. Then*

$$\lim_{n\to\infty}\frac{1}{n}\log P_{n,\hat\rho}\left(\frac{1}{n}\langle C(v)\rangle\in\mathscr{S}\right)=\sup_{\hat\eta\in\mathscr{S}}\left[S(\hat\eta,\hat\rho)+\Xi(\hat\eta)-\hat\eta^{\mathrm{T}}\mathrm{A}(\hat\rho-\hat\eta)\right]. \tag{2.11}$$

We are also interested in the rate of $\mathscr{G}(n,\hat\rho)$ *not* having large components:

Theorem 2.6. *Let L, A, and $\hat\rho>0$ be given. Also define the continuous function*

$$\Psi(\hat\rho)=\begin{cases}\sum_{i=1}^{L}\left[\hat\rho_i\log\frac{\hat\rho_i}{\hat\rho_i^\star}-\frac12\left(\hat\rho_i-\hat\rho_i^\star\right)\left(1+\sum_{j=1}^{L}\alpha_{ij}\hat\rho_j\right)\right] & \text{if } \hat\rho \text{ is supercritical}\\ 0 & \text{otherwise}\end{cases}, \tag{2.12}$$

where for $\hat\rho$ supercritical, $\hat\rho^\star$ is the unique critical density such that $(\mathrm{I}-\mathrm{B}_{\hat\rho^\star})(\hat\rho-\hat\rho^\star)=0$. Then if $\mathbb{S}_r$ is the event that all components in $\mathscr{G}(\hat\rho,n)$ are of size smaller than r, we have

$$\lim_{r\to\infty}\lim_{n\to\infty}\frac{1}{n}\log P_{n,\hat\rho}(\mathbb{S}_r)=\lim_{\epsilon\to0}\lim_{n\to\infty}\frac{1}{n}\log P_{n,\hat\rho}(\mathbb{S}_{\epsilon n}) \tag{2.13}$$

$$=\Psi(\hat\rho). \tag{2.14}$$

Moreover, the convergence is uniform for $\hat\rho$ bounded above.

As a kind of combination of Theorem 2.5 and Theorem 2.6, we find the exponential rate of $\hat\theta$ taking a specified value:

Theorem 2.7. *Let L, A, and $\hat\rho>0$ be given, and let $\mathscr{S}\subset\prod_i[0,\hat\rho_i]$ be a set of densities open in $\prod_i[0,\hat\rho_i]$. Then*

$$\lim_{\epsilon\to0}\lim_{n\to\infty}\frac{1}{n}\log P\left(\hat\theta(\epsilon n,n,\hat\rho)\in\mathscr{S}\right)=\sup_{\hat\eta\in\mathscr{S}}\left[S(\hat\eta,\hat\rho)+\Xi(\hat\eta)+\Psi(\hat\rho-\hat\eta)-\hat\eta^{\mathrm{T}}\mathrm{A}(\hat\rho-\hat\eta)\right]. \tag{2.15}$$

Finally, we demonstrate that the giant component is unique:

Theorem 2.8. *Let L, A, $\hat\rho>0$, and $\delta>0$ be given. Let Q_ϵ be the event that $\mathscr{G}(n,\hat\rho)$ has two or more components of size at least ϵn. Then*

$$\limsup_{n\to\infty}\frac{1}{n}\log P_{n,\hat\rho}\left(Q_\epsilon\,\middle|\,\hat\theta(\epsilon n,n,\hat\rho)\in[\delta,1]^L\right)<0. \tag{2.16}$$

Thus the giant component is unique with exponential probability.

Proof of Lemma 2.2. We start by proving uniqueness of the nonzero solution, and are thus interested in studying fixed points of the transformation $\hat T:\mathbb{R}^L\mapsto\mathbb{R}^L$ given by

$$\hat T_k(\hat\theta)=1-\exp\left(-\sum_{i=1}^{L}\alpha_{ik}\hat\rho_i\hat\theta_i\right). \tag{2.17}$$

We claim that $\hat T$ has a maximal fixed point. To see this we let $\hat b=(1,1,\dots,1)$, noting that this is an upper bound on $\hat T$. Since $\hat T$ is order preserving and 0 is a fixed point of $\hat T$, we have that $\hat T^{(k)}(\hat b)$ is decreasing in k and bounded below by 0, and thus converges. Moreover, since $\hat T$ is continuous,

$\hat{c} = \lim_{k\to\infty} \hat{T}^{(k)}(\hat{b})$ must be a fixed point. Again using that $\hat{T}$ preserves order, $\hat{a} \leq \hat{T}^{(k)}(\hat{b})$ for all k and fixed points $\hat{a}$, and therefore $\hat{c}$ is a maximal fixed point of $\hat{T}$.

Since 0 is clearly a fixed point of $\hat{T}$, we must have $\hat{c} \geq 0$. Supposing that $\hat{c} \neq 0$, we note that $\hat{\rho} > 0$ and A irreducible means that all nonzero fixed points of $\hat{T}$ in $[0,\infty)^L$ must be strictly positive. Then consider $\hat{T}(\hat{c}t)$, and notice that with $\hat{c} > 0$, $\hat{T}(\hat{c}t)$ is a *strictly* concave function of t in each component. Since 0 and $\hat{c}$ are fixed points of $\hat{T}$, this means that $T(\hat{c}t) > \hat{c}t$ for all $t \in (0,1)$. Furthermore, if $\hat{a} \neq \hat{c}$ is given with $0 < \hat{a} \leq \hat{c}$, we can pick $t_0 \in (0,1)$ so that $\hat{c}t_0 \leq \hat{a}$ and $\hat{c}_j t_0 = \hat{a}_j$ for some j. Then $\hat{T}(\hat{c}t_0) \leq \hat{T}(\hat{a})$ but $\hat{T}_j(\hat{c}t_0) > \hat{a}_j$ (by the strict concavity), so $\hat{a}$ cannot be a fixed point. This proves the uniqueness of nonzero solution to (2.2) in $[0,\infty)^L$, if one exists.

To show that the maximal solution is increasing as a function of $\hat{\rho}$, we let $\hat{R}$ and $\hat{S}$ be copies of $\hat{T}$ with $\hat{\rho} = \hat{\rho}^{(1)}$ and $\hat{\rho} = \hat{\rho}^{(2)}$ respectively, where $0 < \hat{\rho}^{(1)} \leq \hat{\rho}^{(2)}$. Then $\hat{R}(\hat{a}) \leq \hat{S}(\hat{a})$ for all $\hat{a}$. In particular, we must have $\lim_{k\to\infty} \hat{R}^k(\hat{b}) \leq \lim_{k\to\infty} \hat{S}^k(\hat{b})$. Since these limits are the respective maximal fixed points of $\hat{R}$ and $\hat{S}$, we have that the maximal fixed point of $\hat{T}$ is increasing in $\hat{\rho}$. □

We also note (without proof) that the ordering of the fixed points is strict.

Proof of Lemma 2.3. Let $\hat{T}$ be defined as in the proof of Lemma 2.2, and note

$$\hat{T}(\hat{\theta}) = \mathrm{B}_{\hat{\rho}}^{\mathrm{T}} \hat{\theta} + O\big(|\hat{\theta}|^2\big), \tag{2.18}$$

where $O\big(|\hat{\theta}|^2\big)$ is negative for $\hat{\theta} > 0$. Since A is irreducible, $\mathrm{B}_{\hat{\rho}}^{\mathrm{T}}$ must be as well. Then by the Perron-Frobenius theorem we know that all eigenvalues of $\mathrm{B}_{\hat{\rho}}^{\mathrm{T}}$ – which are also eigenvalues of $\mathrm{B}_{\hat{\rho}}$ – are bounded in magnitude by $\kappa_0(\hat{\rho})$, and $\mathrm{B}_{\hat{\rho}}^{\mathrm{T}}$ has a eigenvector $\hat{\eta}$ corresponding to $\kappa_0(\hat{\rho})$ with $\hat{\eta} > 0$. Because the $O\big(|\hat{\theta}|^2\big)$ term is negative, we have $\hat{T}(c\hat{\eta}) < c\kappa_0(\hat{\rho})\hat{\eta}$ for all $c > 0$. Then if $\kappa_0(\hat{\rho}) \leq 1$, for any $\hat{\theta} > 0$ we can find a $c > 0$ so that $c\hat{\eta} \geq \hat{\theta}$ and $c\hat{\eta}_i = \hat{\theta}_i$ for some i. By the monotonicity of $\hat{T}$,

$$\hat{T}_i(\hat{\theta}) \leq \hat{T}_i(c\hat{\eta}) < \hat{\eta}_i = \hat{\theta}_i, \tag{2.19}$$

which shows $\hat{\theta}$ is not a fixed point of $\hat{T}$. On the other hand, if $\kappa_0(\hat{\rho}) > 1$ then $\hat{T}(c\hat{\eta}) > c\hat{\eta}$ for small c. Thus since $\hat{T}$ is increasing, continuous, and bounded by $(1,1,\ldots,1)$, it has a nonzero fixed point in $(0,1)^L$.

Having established that $\hat{T}$ has a positive fixed point if and only if $\kappa_0(\hat{\rho}) > 1$, the results follow from $\kappa_0(\hat{\rho})$ being continuous and strictly increasing on $(0,\infty)^L$. □

3. All Sites Connected

The goal of this section is to prove Theorem 2.5 and the following:

Theorem 3.1. *Let L and* A *be given, and let K_m be the event that $\mathscr{G}(n,\hat{\rho})$ has m or fewer components. Then for each $m \geq 1$ and $\hat{\rho} > 0$,*

$$\lim_{n\to\infty} \frac{1}{n} \log P_{n,\hat{\rho}}(K_m) = \Xi(\hat{\rho}), \tag{3.1}$$

and the convergence is uniform for $\hat{\rho}$ bounded above and bounded away from zero.

Notice that if m is chosen to be 1, Theorem 3.1 gives the exponential rate of $\mathscr{G}(n,\hat{\rho})$ being connected.

In order to prove Theorem 3.1, we will convert the problem into an equivalent problem for directed graphs. To this end, for all vertex sets $\mathscr{U}$ and $\mathscr{W}$, and all directed edge sets $\vec{\mathscr{E}}$, let $F\big(\mathscr{U},\mathscr{W},\vec{\mathscr{E}}\big)$ be the event that for every vertex in $\mathscr{U}$ there is a path of (strictly) positive length in $\vec{\mathscr{E}}$ to a vertex in $\mathscr{W}$.

Lemma 3.2. *Let vertex sets $\mathscr{U}$ and $\mathscr{W}$, and $(p_v)_{v\in\mathscr{U}\cup\mathscr{W}}$ be given with $\sum_{v\in\mathscr{U}\cup\mathscr{W}} p_v = 1$. Let $(\omega_v)_{v\in\mathscr{U}}$ be i.i.d. random elements of $\mathscr{U}\cup\mathscr{W}$ with*

$$P(\omega_v = u) = p_u. \tag{3.2}$$

Then let $\vec{\mathscr{E}}$ be the random edge set $\vec{\mathscr{E}} = \big\{(v,\omega_v) : v\in\mathscr{U}\big\}$. Then

$$P\left[F\big(\mathscr{U},\mathscr{W},\vec{\mathscr{E}}\big)\right] = \sum_{v\in\mathscr{W}} p_v. \tag{3.3}$$

Proof. Given our construction of $\vec{\mathscr{E}}$ (in which each vertex has an outgoing edge), we have that

$$F\big(\mathscr{U},\mathscr{W},\vec{\mathscr{E}}\big) = F\big(\mathscr{U}\setminus\mathscr{W},\mathscr{W},\vec{\mathscr{E}}\big). \tag{3.4}$$

Indeed, in the left hand event each vertex in $\mathscr{W}$ is either directly connected to $\mathscr{W}$, or is indirectly connected to $\mathscr{W}$ through $\mathscr{U}\setminus\mathscr{W}$. Thus we may assume without loss of generality that $\mathscr{U}$ and $\mathscr{W}$ are disjoint. We will proceed by induction on $|\mathscr{U}|$; the size of $\mathscr{W}$ is unimportant. The case with $|\mathscr{U}| = 1$ is trivial, so we assume that the lemma holds for $|\mathscr{U}|\le N$, and take $|\mathscr{U}| = N+1$. Let $u\in\mathscr{U}$, and note that if $\omega_u = u$ then $F\big(\mathscr{U},\mathscr{W},\vec{\mathscr{E}}\big)$ does not occur. On the other hand – that is, if $\omega_u\neq u$ – we let

$$\omega'_v = \begin{cases} \omega_u & \text{if } \omega_v = u \\ \omega_v & \text{otherwise} \end{cases}, \tag{3.5}$$

for each $v\in\mathscr{U}\setminus\{u\}$, and let $\vec{\mathscr{E}}' = \{(v,\omega'_v) : v\in\mathscr{U}\setminus\{u\}\}$. Then every path in $\vec{\mathscr{E}}$ which does not start or end with u has a naturally corresponding path in $\vec{\mathscr{E}}'$: simply remove all instances of u from the path. By this correspondence, we see that $F\big(\mathscr{U}\setminus\{u\},\mathscr{W},\vec{\mathscr{E}}'\big)$ occurs if and only if $F\big(\mathscr{U},\mathscr{W},\vec{\mathscr{E}}\big)$ occurs. We note further that in the case $\omega_u\in\mathscr{W}$ we have $P(\omega'_v\in\mathscr{W}) = p_u + \sum_{y\in\mathscr{W}} p_y$ for $v\in\vec{\mathscr{U}}\setminus\{u\}$.

Thus we have

$$\begin{aligned} P\left[F\big(\mathscr{U},\mathscr{W},\vec{\mathscr{E}}\big)\right] &= \left(\sum_{y\in\mathscr{W}} p_y\right) P\left[F\big(\mathscr{U}\setminus\{u\},\mathscr{W},\vec{\mathscr{E}}'\big)\,\middle|\,\omega_u\in\mathscr{W}\right] \\ &\quad + \left(\sum_{y\in\mathscr{U}\setminus\{u\}} p_y\right) P\left[F\big(\mathscr{U}\setminus\{u\},\mathscr{W},\vec{\mathscr{E}}'\big)\,\middle|\,\omega_u\in\mathscr{U}\setminus\{u\}\right]. \end{aligned} \tag{3.6}$$

Since $(\omega'_v)_{v\in\mathscr{U}\setminus\{u\}}$ are i.i.d. once ω_u is conditioned upon, our induction assumption gives that

$$P\left[F\big(\mathscr{U}\setminus\{u\},\mathscr{W},\vec{\mathscr{E}}'\big)\,\middle|\,\omega_u = y\right] = \begin{cases} p_u + \sum_{v\in\mathscr{W}} p_v & \text{if } y\in\mathscr{W} \\ \sum_{v\in\mathscr{W}} p_v & \text{if } y\in\mathscr{U}\setminus\{u\} \end{cases}. \tag{3.7}$$

Substituting these into (3.6) gives the result for $|\mathscr{U}| = N+1$, and completes the inductive step of the proof. $\square$

Lemma 3.3. *Let* L, A, *and* $\hat{\rho} > 0$ *be given, and let* $\vec{\mathscr{E}}$ *be the random directed edge set on vertex set* $\mathscr{V}$ *so that each directed edge* (u,v) *appears (or not) in* $\vec{\mathscr{E}}$ *independently with* $P\big((u,v) \in \vec{\mathscr{E}}\big) = p_{\mathrm{layer}(u)\,\mathrm{layer}(v)}$. *Then, letting* $\mathbb{P}_{n,\hat{\rho}}$ *be the probability measure associated with* $\vec{\mathscr{E}}$*, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{n,\hat{\rho}} \left[F\big(\mathscr{V}, \{v\}, \vec{\mathscr{E}}\big) \right] = \Xi(\hat{\rho}) \tag{3.8}$$

for all $v \in \mathscr{V}$. *Moreover, if* $\hat{\rho}$ *is bounded above and bounded away from zero, the convergence is uniform.*

Proof. We will prove the lemma by showing

$$\lim_{n\to\infty} \mathbb{P}_{n,\hat{\rho}} \left[F\big(\mathscr{V}, \{v\}, \vec{\mathscr{E}}\big) \right]^{1/n} = \lim_{n\to\infty} \mathbb{P}_{n,\hat{\rho}} \left[F\big(\mathscr{V}, \mathscr{V}, \vec{\mathscr{E}}\big) \right]^{1/n}, \tag{3.9}$$

noting that the event on the right hand side is the event that each vertex in $\mathscr{V}$ has an outgoing edge in $\vec{\mathscr{E}}$.

First, note that since $F\big(\mathscr{V}, \mathscr{W}, \vec{\mathscr{E}}\big)$ is increasing in $\mathscr{W}$, in the sense that $F\big(\mathscr{V}, \mathscr{W}, \vec{\mathscr{E}}\big) \subset F\big(\mathscr{V}, \mathscr{U}, \vec{\mathscr{E}}\big)$ whenever $\mathscr{W} \subset \mathscr{U}$. In particular, we have

$$\mathbb{P}_{n,\hat{\rho}} \left[F\big(\mathscr{V}, \{v\}, \vec{\mathscr{E}}\big) \right] \le \mathbb{P}_{n,\hat{\rho}} \left[F\big(\mathscr{V}, \mathscr{V}, \vec{\mathscr{E}}\big) \right]. \tag{3.10}$$

To get a lower bound, we let $\vec{\mathscr{F}}$ be a subset of $\vec{\mathscr{E}}$ gotten by discarding all but one outgoing edge (randomly selected) from each vertex. Then let $\tilde{\mathbb{P}}(\cdot) = \mathbb{P}_{n,\hat{\rho}} \left[\cdot \middle| F\big(\mathscr{V}, \mathscr{V}, \vec{\mathscr{E}}\big) \right]$, and note that since $F\big(\mathscr{V}, \{v\}, \vec{\mathscr{E}}\big)$ is an increasing event with respect to $\vec{\mathscr{E}}$, we have

$$\tilde{\mathbb{P}} \left[F\big(\mathscr{V}, \{v\}, \vec{\mathscr{E}}\big) \right] \ge \tilde{\mathbb{P}} \left[F\big(\mathscr{V}, \{v\}, \vec{\mathscr{F}}\big) \right]. \tag{3.11}$$

We also note that under the measure $\tilde{\mathbb{P}}$, the set $\vec{\mathscr{F}}$ contains exactly one outgoing edge for each vertex, and that if we let

$$b_{ij} = \lim_{n\to\infty} \tilde{\mathbb{P}} \left[(x,y) \in \vec{\mathscr{F}} \text{ for some } y \in \mathscr{V}_j \right] \tag{3.12}$$

where $x \in \mathscr{V}_i$, we have

$$b_{ij} \ge \alpha_{ij} \hat{\rho}_j \mathrm{e}^{-\sum_{k=1}^{L} \alpha_{ik} \hat{\rho}_k}. \tag{3.13}$$

Specifically, if the probability of x having an edge in $\vec{\mathscr{E}}$ to a given layer is uniformly bounded away from zero, the probability of x having such an edge in $\vec{\mathscr{F}}$ is also uniformly bounded away from zero.

We will now use induction on L to show that

$$\tilde{\mathbb{P}} \left[F\big(\mathscr{V}, \{v\}, \vec{\mathscr{F}}\big) \right] = c/n + O(1/n^2), \tag{3.14}$$

for some constant $c > 0$ depending on L, A, and $\hat{\rho}$. If $L = 1$, this is a straightforward application of Lemma 3.2. We then assume that (3.14) holds for $L \le M$ and take $L = M+1$. Assume without loss of generality that v is not in layer $M+1$ and apply Lemma 3.2 to conclude

$$\tilde{\mathbb{P}}\left[F\left(\mathscr{V}_{M+1},\bigcup_{i=1}^{M}\mathscr{V}_i,\vec{\mathscr{F}}'\right)\right]=1-b_{(M+1)(M+1)}+o(1). \tag{3.15}$$

Here the fact that A is irreducible (and $\hat{\rho}>0$) implies that $1-b_{(M+1)(M+1)}>0$. Next, if $F\left(\mathscr{V}_{M+1},\bigcup_{i=1}^{M}\mathscr{V}_i,\vec{\mathscr{F}}\right)$ occurs, we let $\vec{\mathscr{F}}'$ be a copy of $\vec{\mathscr{F}}$ in which the edges terminating in $\mathscr{V}_{M+1}$ are remapped to edges terminating in $\bigcup_{i=1}^{M}\mathscr{V}_i$ á la the proof of Lemma 3.2. The rates of connections in $\vec{\mathscr{F}}'$ are given by (b'_{ij}), where $b'_{ij}=b_{ij}+b_{i(M+1)}b_{j(M+1)}$ for $i,j\le M$. Once again A being irreducible implies (b'_{ij}) must be as well. Thus we may use our inductive assumption to prove that (3.14) holds for all L.

Now using (3.11) and taking limits we have

$$\lim_{n\to\infty}\mathbb{P}_{n,\hat{\rho}}\left[F\big(\mathscr{V},\{v\},\vec{\mathscr{E}}\big)\right]^{1/n}\ge\lim_{n\to\infty}\mathbb{P}_{n,\hat{\rho}}\left[F\big(\mathscr{V},\mathscr{V},\vec{\mathscr{E}}\big)\right]^{1/n}, \tag{3.16}$$

which proves (3.9); and since nonzero b_{ij} are uniformly bounded away from zero for $\hat{\rho}$ bounded above and bounded away from zero, we have uniform convergence. The lemma now follows from recalling

$$\mathbb{P}_{n,\hat{\rho}}\left[F\big(\mathscr{V},\mathscr{V},\vec{\mathscr{E}}\big)\right]=\prod_{k=1}^{L}\left(1-\prod_{i=1}^{L}\Big(1-\frac{\alpha_{ik}}{n}\Big)^{\lfloor\hat{\rho}_i n\rfloor}\right)^{\lfloor\hat{\rho}_k n\rfloor} \tag{3.17}$$

and taking limits. □

We will also use Lemma 3.2 from [2], which we reproduce here (with slightly modified notation):

Lemma 3.4. *For a collection of vertices $\mathscr{W}=\{1,\dots,n\}$ with an associated set of edge probabilities $(p_{kl})_{1\le k<l\le n}$, let $\mathscr{G}$ be the inhomogeneous undirected random graph over $\mathscr{W}$. Similarly, let $\vec{\mathscr{G}}$ denote the inhomogeneous* directed *graph with the restriction that the two possible (directed) edges between k and l occur independently, each with probability p_{kl}. Letting $\vec{\mathscr{E}}_{\vec{\mathscr{G}}}$ be the edge set of $\vec{\mathscr{G}}$, we have $P(\mathscr{G}$ is connected$)=P\big(F(\mathscr{W}\setminus\{1\},\{1\},\vec{\mathscr{E}}_{\vec{\mathscr{G}}})\big)$.*

Proof of Theorem 3.1. For a vertex set $\mathscr{W}$, let $\mathscr{E}^+(\mathscr{W})$ be the complete set of edges on $\mathscr{W}$. We then note for $n\ge m$

$$K_m=\bigcup_{\substack{\mathscr{W}\subset\mathscr{V}\\|\mathscr{W}|=m}}\left\{\mathscr{G}(n,\hat{\rho})\cup\mathscr{E}^+(\mathscr{W})\text{ is connected}\right\}, \tag{3.18}$$

and thus

$$P_{n,\hat{\rho}}(K_m)=n^{O(m)}\sup_{\substack{\mathscr{W}\subset\mathscr{V}\\|\mathscr{W}|=m}}P_{n,\hat{\rho}}\left(\mathscr{G}(n,\hat{\rho})\cup\mathscr{E}^+(\mathscr{W})\text{ is connected}\right). \tag{3.19}$$

Moreover, since the vertices are a priori interchangeable within a given layer, we need only consider a finite number of $\mathscr{W}$ in the supremum. Thus we will have proved the theorem once we have shown that $\frac{1}{n}\log P_{n,\hat{\rho}}\big(\mathscr{G}(n,\hat{\rho})\cup\mathscr{E}^+(\mathscr{W})\text{ is connected}\big)$ converges uniformly to $\Xi(\hat{\rho})$ uniformly, for $\hat{\rho}$ bounded above and bounded away from zero. Thus let us fix $\mathscr{S}\subset\mathscr{V}$ with

$|\mathscr{S}| = m$, and let $v \in \mathscr{S}$. Since adding an edge (a,b) to $\mathscr{G}(n,\hat{\rho})$ is equivalent to setting $p_{ab} = 1$, and since $F\big(\mathscr{V} \setminus \{v\}, \{v\}, \vec{\mathscr{E}} \cup \vec{\mathscr{E}}^{+}(\mathscr{S})\big) = F\big(\mathscr{V} \setminus \mathscr{S}, \mathscr{S}, \vec{\mathscr{E}}\big)$ – where $\vec{\mathscr{E}}^{+}$ is the full set of directed edges on $\mathscr{S}$ – Lemma 3.4 shows that

$$P_{n,\hat{\rho}}\left(\mathscr{G}(n,\hat{\rho}) \cup \mathscr{E}^{+}(\mathscr{S}) \text{ is connected}\right) = \mathbb{P}_{n,\hat{\rho}}\left[F\big(\mathscr{V} \setminus \mathscr{S}, \mathscr{S}, \vec{\mathscr{E}}\big)\right]. \tag{3.20}$$

Then we note

$$F\big(\mathscr{V}, \mathscr{S}, \vec{\mathscr{E}}\big) = F\big(\mathscr{V} \setminus \mathscr{S}, \mathscr{S}, \vec{\mathscr{E}}\big) \bigcap F\big(\mathscr{S}, \mathscr{V}, \vec{\mathscr{E}}\big). \tag{3.21}$$

and point out that events on the right hand side are independent. Since the rightmost event occurs with (uniform) probability of order unity and the event on the left hand side is bounded between $F\big(\mathscr{V}, \{v_1\}, \vec{\mathscr{E}}\big)$ and $F\big(\mathscr{V}, \mathscr{V}, \vec{\mathscr{E}}\big)$, we can use Lemma 3.3 to conclude

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{n,\hat{\rho}}\left[F\big(\mathscr{V} \setminus \mathscr{S}, \mathscr{S}, \vec{\mathscr{E}}\big)\right] = \Xi(\hat{\rho}). \tag{3.22}$$

Since this holds for all choices of $\mathscr{S}$, and the convergence is uniform, we use (3.20) and (3.19) to finish the proof. □

Corollary 3.5. *Let $m \geq 1$ and $\hat{\rho}^{(0)}$ be given. Then,*

$$\frac{1}{n} \log P_{n,\hat{\rho}}(K_m) \leq \Xi(\hat{\rho}) + o(1), \tag{3.23}$$

where the $o(1)$ term is uniformly bounded for $(1/n, \ldots, 1/n) \leq \hat{\rho} \leq \hat{\rho}^{(0)}$. Furthermore, if $m = 1$ then the $o(1)$ term is uniformly bounded for $0 \leq \hat{\rho} \leq \hat{\rho}^{(0)}$.

Proof. We first note that the lower bound on $\hat{\rho}$ is chosen to guarantee that each layer has at least one vertex. For $m > 1$, this guarantees that no vertices are strictly isolated. For $m = 1$ such a restriction is unnecessary – indeed, the existence of vertices which are almost surely isolated will cause the left hand side of (3.23) to be negative infinity. Thus for the remainder of the proof we assume that each vertex has at least one potential neighbor.

Following the proof of Theorem 3.1 up to around (3.19), we get

$$\frac{1}{n} \log P_{n,\hat{\rho}}(K_m) = o(1) + \sup_{\substack{\mathscr{W} \subset \mathscr{V} \\ |\mathscr{W}| = m}} \frac{1}{n} \log P_{n,\hat{\rho}}\left(\mathscr{G}(n,\hat{\rho}) \cup \mathscr{E}^{+}(\mathscr{W}) \text{ is connected}\right), \tag{3.24}$$

where $o(1)$ is independent of $\hat{\rho}$. Then taking $\mathscr{S}$ as in the proof of Theorem 3.1, we have by (3.21)

$$\mathbb{P}_{n,\hat{\rho}}\left[F\left(\mathscr{V} \setminus \mathscr{S}, \mathscr{S}, \vec{\mathscr{E}}\right)\right] \leq O(n)^m \prod_{k=1}^{L} \left(1 - \prod_{i=1}^{L} \left(1 - \frac{\alpha_{ik}}{n}\right)^{\lfloor \hat{\rho}_i n \rfloor}\right)^{\lfloor \hat{\rho}_k n \rfloor}, \tag{3.25}$$

where the $O(n)^m$ term comes from $\mathbb{P}_{n,\hat{\rho}}\left[F\big(\mathscr{S}, \mathscr{V}, \vec{\mathscr{E}}\big)\right]$ – here we have used the assumption that each vertex has at least one potential neighbor – and the remainder of the right hand side comes from bounding the event $F\big(\mathscr{V}, \mathscr{S}, \vec{\mathscr{E}}\big)$ by $F\big(\mathscr{V}, \mathscr{V}, \vec{\mathscr{E}}\big)$ and using (3.17). The result then follows by applying (3.20) and taking limits. □

Proof of Theorem 2.5. We note that for all $\hat{\eta}$ with $0 \leq \hat{\eta} \leq \hat{\rho}$ we have

$$P_{n,\hat{\rho}}\big(\langle C(v)\rangle = \lfloor \hat{\eta} n\rfloor\big) = \prod_{i=1}^{L} \binom{\lfloor \hat{\rho}_i n\rfloor}{\lfloor \hat{\eta}_i n\rfloor} P_{n,\hat{\eta}}(K_1) \prod_{i,j=1}^{L} \left(1 - \frac{\alpha_{ij}}{n}\right)^{\lfloor \hat{\eta}_i n\rfloor \left(\lfloor \rho_j n\rfloor - \lfloor \hat{\eta}_j n\rfloor\right)}. \tag{3.26}$$

Thus for a lower bound we note that for each $\hat{\eta} \in \mathscr{S}$ we have $\frac{1}{n}\lfloor \hat{\eta} n\rfloor \in \mathscr{S}$ for sufficiently large n – following from $\mathscr{S}$ open. Thus for each $\hat{\eta} \in \mathscr{S}$ we use (3.26) and Theorem 3.1 to get

$$\liminf_{n\to\infty} \frac{1}{n} \log P_{n,\hat{\rho}} \left(\frac{1}{n}\langle C(v)\rangle \in \mathscr{S}\right) \geq S(\hat{\eta},\hat{\rho}) + \Xi(\hat{\eta}) - \hat{\eta}^{\mathrm{T}}\mathrm{A}(\hat{\rho} - \hat{\eta}). \tag{3.27}$$

For an upper bound, note that $\langle C(v)\rangle$ can take only polynomially many values, and thus

$$P_{n,\hat{\rho}} \left(\frac{1}{n}\langle C(v)\rangle \in \mathscr{S}\right) = \mathrm{e}^{o(n)} \sup_{\hat{\eta}\in\mathscr{S}} P_{n,\hat{\rho}}\big(\langle C(v)\rangle = \lfloor \hat{\eta} n\rfloor\big). \tag{3.28}$$

Now using (3.26) and Corollary 3.5, we get an upper bound which matches (3.27). This proves the Theorem. □

4. Locally a Tree

The goal of this section is to prove Theorem 2.4, which we shall do by bounding (in a manner of speaking) $\mathscr{G}(n,\hat{\rho})$ between random layered trees, which is essentially a flushing out of the ideas following Lemma 2.3. To do this we introduce a random tree which locally resembles $\mathscr{G}(n,\hat{\rho})$, and will prove a lemma similar to Theorem 2.4 on the tree. Let $\hat{c} \geq 0$ be given with $|\hat{c}| = 1$, and let $\mathscr{T}(n,\hat{\eta})$ be a random layered tree constructed so that its root r is in layer j with probability $\hat{c}_j$; and so that each vertex $w \in \mathscr{T}_i(n,\hat{\eta})$ has a random, independent number of children, distributed as $\mathrm{Bin}\left(\lfloor \hat{\eta}_i n\rfloor, \frac{\alpha_{ij}}{n}\right)$ in each of layers $j = 1,\ldots,L$. Then we have:

Lemma 4.1. *Let $\hat{\eta} > 0$ be given. If $\hat{\eta}$ is subcritical, then $\mathscr{T}(n,\hat{\eta})$ is almost surely finite, and*

$$E[\langle \mathscr{T}(n,\hat{\eta})\rangle] = (\mathrm{I} - \mathrm{B}_{\hat{\eta}})^{-1}\hat{c} \tag{4.1}$$

for all n, and

$$\lim_{j\to\infty}\lim_{n\to\infty} E[\langle \mathscr{T}(n,\hat{\eta})\rangle 1_{|\mathscr{T}(n,\hat{\eta})|\leq j}] = (\mathrm{I} - \mathrm{B}_{\hat{\eta}})^{-1}\hat{c}. \tag{4.2}$$

If $\hat{\eta}$ is critical then $\mathscr{T}(n,\hat{\eta})$ is almost surely finite for all n and

$$\lim_{j\to\infty}\lim_{n\to\infty} P(|\mathscr{T}(n,\hat{\eta})| \leq j) = 1 \tag{4.3}$$

but

$$E\big[|\mathscr{T}_i(n,\hat{\eta})|\big] = \infty \tag{4.4}$$

for all i,n. Finally, if $\hat{\eta}$ is supercritical then $P\big(|\mathscr{T}_i(n,\hat{\eta})| = \infty\big)$ is uniformly bounded away from zero for all i,n.

The majority – if not the entirety – of Lemma 4.1 exists as general knowledge in the literature (see, for example [1], and references therein). For this reason we will attempt to be succinct in our proof.

Proof. To prove (4.1), we let $Y(n, \hat{\eta}, k)$ be the set of vertices in $\mathscr{T}(n, \hat{\eta})$ which are distance k from the root. Then

$$E\big[\langle Y(n, \hat{\eta}, k+1)\rangle \,\big|\, \langle Y(n, \hat{\eta}, k)\rangle\big] = \mathrm{B}_{\hat{\eta}}\langle Y(n, \hat{\eta}, k)\rangle. \tag{4.5}$$

Thus we find

$$E\big[\langle \mathscr{T}(n, \hat{\eta})\rangle\big] = \sum_{k=0}^{\infty} \mathrm{B}_{\hat{\eta}}^{k}\hat{c}. \tag{4.6}$$

Since all eigenvalues of $\mathrm{B}_{\hat{\eta}}$ are bounded by $\kappa_0(\hat{\eta})$ in magnitude, if $\kappa_0(\hat{\eta}) < 1$ then $\sum_{k=0}^{\infty} \mathrm{B}_{\hat{\eta}}^{k} = (I - B(\hat{\eta}))^{-1}$, which gives us (4.1). On the other hand, if $\kappa_0(\hat{\eta}) = 1$ then the sum is non-convergent. Moreover, since $\mathrm{B}_{\hat{\eta}}$ is irreducible, we have (4.4).

Recalling once more that $\mathrm{Bin}(\lfloor an \rfloor, \frac{b}{n})$ converges in distribution to a $\mathrm{Poisson}(ab)$ as n tends to infinity, $\langle Y(n, \hat{\eta}, k)\rangle$ converges in distribution to a random vector $\hat{Z}(\hat{\eta}, k)$ as $n \to \infty$, and $\hat{Z}(\hat{\eta}, k)$ satisfies $E[\hat{Z}(\hat{\eta}, k)] = \mathrm{B}_{\hat{\eta}}^{k}\hat{c}$. Then for all k, using the dominated convergence theorem followed by the monotone convergence theorem, we have

$$\lim_{j\to\infty}\lim_{n\to\infty} E\left[\langle Y(n, \hat{\eta}, k)\rangle 1_{|\mathscr{T}(n,\hat{\eta})|\le j}\right] = \mathrm{B}_{\hat{\eta}}^{k}\hat{c} \tag{4.7}$$

if $\hat{\eta}$ is subcritical. To prove (4.2), we note that

$$\lim_{j\to\infty}\lim_{n\to\infty} E\big[\langle \mathscr{T}(n, \hat{\eta})\rangle 1_{|\mathscr{T}(n,\hat{\eta})|\le j}\big] = \lim_{j\to\infty}\lim_{n\to\infty} \sum_{k=0}^{\infty} E\big[\langle Y(n, \hat{\eta}, k)\rangle 1_{|\mathscr{T}(n,\hat{\eta})|\le j}\big], \tag{4.8}$$

and use dominated and monotone convergence to move the limits inside the sum on the right hand side.

From (4.5) we also have that $E\big[\langle Y(n, \hat{\eta}, k)\rangle\big]$ is uniformly bounded as $k \to \infty$ if $\kappa_0(\hat{\eta}) \le 1$. In particular if $\hat{\eta}$ is not supercritical, then for almost every instance of $\langle Y(n, \hat{\eta}, \cdot)\rangle$ there is some r such that $|\langle Y(n, \hat{\eta}, k)\rangle| \le r$ for infinitely many k. Since zero is an absorbing state and $P\big(\langle Y(n, \hat{\eta}, k{+}1)\rangle = 0 \,\big|\, \langle Y(n, \hat{\eta}, k)\rangle = \hat{a}\big) > 0$ for all $\hat{a}$, we must have that $\lim_{k\to\infty}\langle Y(n, \hat{\eta}, k)\rangle = 0$. Thus $\mathscr{T}(n, \hat{\eta})$ is almost surely finite if $\hat{\eta}$ is not supercritical. Since the same argument holds for $\hat{Z}(\hat{\eta}, k)$, we also have (4.3).

To prove the final statement of the lemma, we fix an i and assume for the remainder of the proof that $\hat{\eta}$ is supercritical. We also assume without loss of generality that $r \in \mathscr{T}_i(n, \hat{\eta})$. We then define a distance function $d(u, v)$ on $\mathscr{T}_i(n, \hat{\eta})$ to be the number of vertices in $\mathscr{T}_i(n, \hat{\eta})$ on the (unique) path – which lies in the full tree – from u to v, minus 1. Let $X(n, \hat{\eta}, k)$ be the number of vertices $v \in \mathscr{T}_i(n, \hat{\eta})$ for which $d(v, r) = k$. By the construction of $\mathscr{T}(n, \hat{\eta})$, if $v \in \mathscr{T}_i(n, \hat{\eta})$ then the random variable

$$\big|\{u \in \mathscr{T}_i(n, \hat{\eta}) : u \text{ is a descendant of } v, d(u, v) = 1\}\big| \tag{4.9}$$

is an i.i.d. copy of $X(n, \hat{\eta}, 1)$, which means that $X(n, \hat{\eta}, \cdot)$ is a branching process. Additionally, since $\sum_{k=0}^{\infty} X(n, \hat{\eta}, k) = |\mathscr{T}_i(n, \hat{\eta})|$, we have that $\sum_{k=0}^{\infty} E[X(n, \hat{\eta}, k)] = \infty$. Then writing

$$X(n, \hat{\eta}, k+1) = \sum_{j=1}^{X(n,\hat{\eta},k)} \zeta_{n,\hat{\eta}}^{(j,k)}, \tag{4.10}$$

where $\zeta_{n,\hat{\eta}}^{(j,k)}$ are i.i.d. copies of $\zeta_{n,\hat{\eta}} \overset{d}{=} X(n,\hat{\eta},1)$, we see in particular that $E[\zeta_{n,\hat{\eta}}] > 1$. Moreover, $\zeta_{n,\hat{\eta}}$ converges in distribution to some $\zeta_{\infty,\hat{\eta}}$ with $E[\zeta_{\infty,\hat{\eta}}] > 1$ as $n \to \infty$. Then we can find a $\zeta'_{\hat{\eta}}$ with $E[\zeta'_{\hat{\eta}}] > 1$ so that $\zeta'_{\hat{\eta}}$ is stochastically dominated by $\zeta_{n,\hat{\eta}}$ for sufficiently large n. Using standard proofs about the non-convergence of branching processes shows that a branching process $X'(\hat{\eta},\cdot)$ constructed from $\zeta'_{\hat{\eta}}$ diverges with positive probability. Thus for large n

$$P\big(|\mathscr{T}_i(n,\hat{\eta})| = \infty\big) \geq P\Big(\lim_{k\to\infty} X(n,\hat{\eta},k) = \infty\Big) \geq P\Big(\lim_{k\to\infty} X'(\hat{\eta},k)\Big) > 0, \tag{4.11}$$

which shows that $P\big(|\mathscr{T}_i(n,\hat{\eta})| = \infty\big)$ is bounded away from 0 uniformly in n. □

Proof of Theorem 2.4. We now introduce an exploration process on $\mathscr{G}(n,\hat{\rho})$. Let v_0 be a vertex in $\mathscr{G}(n,\hat{\rho})$, and let $\mathscr{B}(0) = \{v_0\}$. Then for each t let $\mathscr{C}(t) = \bigcup_{i=0}^{t} \mathscr{B}(i)$ and construct $\mathscr{B}(t+1)$ from $\mathscr{B}(t)$ and $\mathscr{C}(t)$ as follows: if $\mathscr{B}(t) = \emptyset$, then $\mathscr{B}(t+1) = \emptyset$. Otherwise let v_t be an element of $\mathscr{B}(t)$ chosen by any mechanism so desired, and let

$$\mathscr{B}(t+1) = \mathscr{B}(t) \cup \{u \notin \mathscr{C}(t) : (u,v_t) \text{ is an edge in } \mathscr{G}(n,\hat{\rho})\} \setminus \{v_t\}. \tag{4.12}$$

Here $\mathscr{B}(t)$ should be thought of as the set of vertices on the boundary of the exploration process, and $\mathscr{C}(t)$ as the set of all vertices seen so far. Then at time t we explore vertex v_t, which moves it to the interior and moves its previously unseen neighbors to the boundary. Let $\sigma = \inf\{t : \mathscr{B}(t) = \emptyset\}$. Then $\sigma = |C(v)|$ and $C(v) = \mathscr{C}(\sigma)$.

Let $\mathscr{D}_{\hat{\eta}}$ be the "exploration process" on $\mathscr{T}(n,\hat{\eta})$ corresponding to $\mathscr{B}$.

We note that if j is the layer of v_t then $\langle \mathscr{B}(t+1)\rangle_i$ is distributed like

$$\langle \mathscr{B}(t+1)\rangle_i \overset{d}{=} \langle \mathscr{B}(t)\rangle_i + \mathrm{Bin}\left(\lfloor \hat{\rho}_i n\rfloor - |\mathscr{C}_i(t)|, \frac{\alpha_{ij}}{n}\right), \tag{4.13}$$

whereas

$$\langle \mathscr{D}_{\hat{\eta}}(t+1)\rangle_i \overset{d}{=} \langle \mathscr{D}_{\hat{\eta}}(t)\rangle_i + \mathrm{Bin}\left(\lfloor \hat{\eta}_i n\rfloor, \frac{\alpha_{ij}}{n}\right). \tag{4.14}$$

In particular, $\langle \mathscr{B}(t)\rangle$ is stochastically dominated by $\langle \mathscr{D}_{\hat{\rho}}(t)\rangle$. Moreover, letting $c = \min_i \hat{\rho}_i > 0$ and $\epsilon > 0$ be fixed, we have that $\langle \mathscr{B}(t)\rangle$ stochastically dominates $\langle \mathscr{D}_{(1-\epsilon)\hat{\rho}}(t)\rangle$ for $t \leq \inf\{t : |\mathscr{C}(t)| > c\epsilon n\}$. This means that we can couple $\mathscr{T}(n,(1-\epsilon)\hat{\rho})$, $\mathscr{G}(n,\hat{\rho})$, and $\mathscr{T}(n,\hat{\rho})$ so that

$$|\mathscr{T}(n,(1-\epsilon)\hat{\rho})| \wedge c\epsilon n \leq |C(v)| \tag{4.15}$$

and

$$\langle C(v)\rangle \leq \langle \mathscr{T}(n,\hat{\rho})\rangle. \tag{4.16}$$

In order to use Lemma 4.1, we use (4.15) to get

$$|\mathscr{T}(n,(1-\epsilon)\hat{\rho})|\, 1_{|\mathscr{T}(n,\hat{\eta})|\leq j} \leq |C(v)| \tag{4.17}$$

for $j < c\epsilon n$. Taking expectations and limits, and applying (4.2), we find

$$\big|\big[\mathrm{I} - B\big((1-\epsilon)\hat{\rho}\big)\big]^{-1}\hat{c}\big| \leq \lim_{n\to\infty} E_{n,\hat{\rho}}\big[|C(v)|\big]. \tag{4.18}$$

Taking ϵ to zero gives (2.6) if $\hat{\rho}$ is critical; or

$$|(\mathrm{I} - \mathrm{B}_{\hat{\rho}})^{-1}\hat{c}| \leq \lim_{n\to\infty} E_{n,\hat{\rho}}\big[|C(v)|\big]. \tag{4.19}$$

if $\hat{\rho}$ is subcritical. Applying (4.1) to (4.16) gives

$$E_{n,\hat{\rho}}[\langle C(v)\rangle] \leq (\mathrm{I} - \mathrm{B}_{\hat{\rho}})^{-1}\hat{c} \tag{4.20}$$

for all n, which proves (2.5). Furthermore, combining (4.16) and (4.3) gives (2.7). Finally, if $\hat{\rho}$ is supercritical, we can choose ϵ small enough so that $(1-\epsilon)\hat{\rho}$ is also supercritical. Then $P\big(|\mathscr{T}(n,(1-\epsilon)\hat{\rho})| = \infty\big)$ is uniformly bounded away from zero, and thus (4.15) gives us (2.8). □

Corollary 4.2. *Let $\hat{\rho} > 0$ subcritical be given. If v is a uniformly chosen vertex from $\mathscr{G}(n,\hat{\rho})$ and $\hat{b} = \lim_{n\to\infty} E_{n,\hat{\rho}}[\langle C(v)\rangle]$, we have*

$$\hat{b}_i = \frac{\hat{\rho}_i}{|\hat{\rho}|} + \hat{\rho}_i \sum_{j=1}^{L} \alpha_{ij}\hat{b}_j. \tag{4.21}$$

Proof. From (2.5) we have $\hat{b} = (\mathrm{I} - \mathrm{B}_{\hat{\rho}})^{-1}\frac{\hat{\rho}}{|\hat{\rho}|}$. Multiplying both sides of this by $(\mathrm{I} - \mathrm{B}_{\hat{\rho}})$ and expanding gives the result. □

Corollary 4.2 can be thought of as a statement about the stationary distribution of an exploration process on $\mathscr{G}(n,\hat{\rho})$. Suppose we have a process on $\mathscr{G}(n,\hat{\rho})$ which visits a vertex v_t at every step according to the rules:

(1) v_{t+1} is uniformly chosen amongst the unvisited neighbors of v_t, if there are any.
(2) v_{t+1} is chosen uniformly amongst all unvisited vertices of $\mathscr{G}(n,\hat{\rho})$ otherwise.

Then we could reasonably expect that $\mathrm{layer}(v_t)$ would have an approximate stationary distribution for $t \ll n$. If we let $\hat{d}$ be this stationary distribution, so that $P(\mathrm{layer}(v_t) = i) \approx \hat{d}_i$, we expect

$$\hat{d}_i = \frac{\hat{\rho}_i}{|\hat{\rho}| E_{n,\hat{\rho}}\big[|C(v)|\big]} + \sum_{j=1}^{L} \hat{\rho}_i \alpha_{ij} \hat{d}_j, \tag{4.22}$$

where v is distributed according to the stationary distribution. As we see, this is just a normalized version of (4.21).

5. Non-Percolating Supercriticality

5.1 Overview.

In order to prove Theorems 2.6 & 2.7, (as well as Theorem 2.1) we will have to consider certain paths in density parameter space; roughly speaking, these represent progress in the *reduction* of the density by the extraction of existing components. In this subsection, we will first define the relevant sorts of paths and introduce a cost function for motion along these paths. Our primary result of this section – Theorem 5.2 – relates the large deviation rate of density reduction to the optimal cost among all paths which connect the densities. We shall turn to some preliminary definitions that will culminate in a statement of Theorem 5.2 – after which we can conclude this overview.

Let $\hat{\Gamma} : (0,\infty) \mapsto (0,\infty)^L$ denote a path in density parameter space. Of exclusive interest will be *assents*, which are Lipschitz continuous paths emanating from the origin and which are nondecreacing in all components and increasing in at least one component. For convenience, we shall parameterize the assents in such a way that $|\hat{\Gamma}(t)| = t$ for all $t > 0$. Finally, we denote by $\mathscr{P}(\hat{\eta})$ the set of assents that pass through the point $\hat{\eta}$.

For $\hat{\gamma}, t \geq 0$ with $|\hat{\gamma}| = 1$, let us define

$$\chi^{-1}(\hat{\gamma}, t) = \inf\left\{x \in (0,1] : \sum_{i=1}^{L} \frac{\hat{\gamma}_i}{x + t\sum_{j=1}^{L} \alpha_{ij}\hat{\gamma}_j} \leq 1\right\} \tag{5.1}$$

where, for historical reasons we have adhered to the inverse rather than the object. Further, we define $\hat{\psi}$ by

$$\hat{\psi}_i(\hat{\gamma}, t) = \frac{\hat{\gamma}_i t}{\chi^{-1}(\hat{\gamma}, t) + t\sum_{j=1}^{L} \alpha_{ij}\hat{\gamma}_j} \tag{5.2}$$

If χ is finite, then $|\hat{\psi}| = t$ and these quantities represent a density/average cluster pairing with $\hat{\psi}$ the density and $\chi\hat{\gamma}$ corresponding to the average sizes, in each layer, of a cluster attached to a uniformly selected site. On the other hand, if t is too large, $\hat{\psi}$ is a critical density and $\hat{\gamma}$ represents the limiting (or critical) ratio of the average cluster sizes. This is summarized in the following:

Lemma 5.1. *If v is a vertex chosen uniformly at random from $\mathscr{V}$ then*

$$\chi(\hat{\gamma}, t) = \lim_{n\to\infty} E_{n,\hat{\psi}(\hat{\gamma},t)}\big[|C(v)|\big] \tag{5.3}$$

with both sides infinity if $\hat{\psi}$ is critical. Further, for all $\hat{\gamma}$ and t for which $\hat{\psi}(\hat{\gamma}, t)$ is subcritical.

$$\lim_{n\to\infty} \frac{E_{n,\hat{\psi}(\hat{\gamma},t)}\big[\langle C(v)\rangle\big]}{E_{n,\hat{\psi}(\hat{\gamma},t)}\big[|C(v)|\big]} = \hat{\gamma}. \tag{5.4}$$

Finally, for critical $\hat{\psi}$, $\hat{\gamma}$ is precisely the maximum eigenvector of $\mathrm{B}_{\hat{\psi}}$.

Proof. Note that for any choice of $\hat{\gamma}$ and t we have from the definition of $\hat{\psi}(\hat{\gamma}, t)$ that

$$\left[\left(\mathrm{I} - \mathrm{B}_{\hat{\psi}(\hat{\gamma},t)}\right)\hat{\gamma}\right]_i = \hat{\gamma}_i - \hat{\psi}_i(\hat{\gamma}, t)\sum_{j=1}^{L} \alpha_{ij}\hat{\gamma}_j = \hat{\gamma}_i\chi^{-1}(\hat{\gamma}, t). \tag{5.5}$$

Thus if $\hat{\psi}(\hat{\gamma}, t)$ is subcritical, (2.5) gives us both (5.3) and (5.4). If $\hat{\psi}(\hat{\gamma}, t)$ is critical then (5.3) is given by (2.6) or (2.8), respectively while (5.5) with $\chi^{-1} = 0$ implies the final statement. □

In light of the above, we may associate with each point on an ascending path the critical or subcritical density $\hat{\psi}\big(\hat{\Gamma}'(t)/|\hat{\Gamma}'(t)|, |\hat{\Gamma}(t)|\big)$; although due to our normalization this simplifies to $\hat{\psi}(\hat{\Gamma}'(t), t)$. We next set

$$\xi(\hat{\gamma}, \hat{\eta}) = \sum_{i=1}^{L} \hat{\gamma}_i \log \hat{\eta}_i - \hat{\gamma}^{\mathrm{T}}\mathrm{A}\hat{\eta}. \tag{5.6}$$

and on the basis of ξ and $\hat{\psi}$ we define a free-energy like object for assents:

$$H\big(\hat{\Gamma}, a, b\big) = \int_a^b \Big(\xi\big(\hat{\Gamma}'(t), \hat{\Gamma}(t)\big) - \xi\big(\hat{\Gamma}'(t), \hat{\psi}(\hat{\Gamma}'(t), t)\big)\Big)\,\mathrm{d}t. \tag{5.7}$$

Most often, we will be interested in maximizing H over paths with fixed endpoints $\hat{\Gamma}(a), \hat{\Gamma}(b)$, which will be clear from context and suppressed from our notation.

Finally, let $C^{(1)}, C^{(2)}, \ldots$ be the components of $\mathscr{G}(n, \hat\rho)$ in a random size-biased order. Then for $0 \le \hat\eta \le \hat\rho$, let

$$\Upsilon(\hat\rho, \hat\eta, r, n) = P\left(\exists k : \sum_{i=1}^{k} \langle C^{(i)} \rangle = \lfloor \hat\rho n \rfloor - \lfloor \hat\eta n \rfloor, |C^{(i)}| \le r \,\forall i \le k\right). \tag{5.8}$$

Put simply, $\Upsilon(\hat\rho, \hat\eta, r, n)$ is the probability that, starting with the system at parameter $\hat\rho$ if we pluck out components at random we arrive at the system with parameter $\hat\eta$ (and do so without ever having selected a component of size larger than r).

We are now in a position to state the following, which is the central object of this section:

Theorem 5.2. *Let $\hat\rho$ and $\hat\eta$ be given with $\hat\rho \ge \hat\eta \ge 0$. Then*

$$\lim_{r\to\infty} \lim_{n\to\infty} \frac{1}{n} \log \Upsilon(\hat\rho, \hat\eta, r, n) = \lim_{\epsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \log \Upsilon(\hat\rho, \hat\eta, \epsilon n, n) = \sup_{\hat\Gamma \in \mathscr{P}(\hat\eta) \cap \mathscr{P}(\hat\rho)} H(\hat\Gamma, |\hat\eta|, |\hat\rho|). \tag{5.9}$$

Moreover, the convergence is uniform for $\hat\rho$ bounded above.

Theorem 5.2 has an interpretation that is not without appeal. Assume that $|\hat\Gamma(a) - \hat\Gamma(b)| \ll 1$ so that we may envision the integrand for a minimizing path of $H(\hat\Gamma, a, b)$ as essentially constant. The cost in (5.7) is now seen as $[\hat\Gamma(b) - \hat\Gamma(a)]$ times the probability – measured on the exponential scale – of observing the system at parameter $\hat\psi$ if the actual system has parameter $\hat\Gamma$. Then, $\hat\psi$ is chosen so that in a typical selection of *its* clusters, the size ratio is such that, typically, when these clusters are selected, the density decreases in the (desired) direction of $\hat\Gamma(a) - \hat\Gamma(b)$.

We now proceed with the overview: In the next subsection, we will present a variety of results which concern the distribution of cluster sizes in these systems and in Subsection 5.3 these will be assembled into a proof of Theorem 5.2. In Subsection 5.4, we will study the H–functional in its own right. In particular, we will define paths called *natural assents* which characterize the maximizers of the cost–functional. Subsection 5.5, noted for its brevity, will contain the proofs of Theorems 2.6 & 2.7 as well as Theorem 2.1.

5.2 Cluster Distributions.

Lemma 5.3. *Let A, L, and $\hat\rho^{(0)}$ be given, and for each $\hat\rho$, let $p(\hat k, \hat\rho, n) = P_{n,\hat\rho}(\langle C(v) \rangle = \hat k)$, where v is a vertex chosen uniformly at random from $\mathscr{G}(n, \hat\rho)$. Then there exists a $c = c(\hat\rho^{(0)})$ and a $p(\hat k, \hat\rho)$ such that for all $\hat\rho \le \hat\rho^{(0)}$,*

$$\exp\left(-c \sum_{i=1}^{L} \frac{\hat k_i^2}{(1 \wedge \hat\rho_i) n}\right) \le \frac{p(\hat k, \hat\rho, n)}{p(\hat k, \hat\rho)} \le \exp\left(c\big(|\hat k|^2/n\big) + \frac{c}{|\hat\rho| n}\right), \tag{5.10}$$

where the lower bound holds only for $\hat k \le \hat\rho n$. In this bound, $\hat k_i^2/\hat\rho_i$ is considered to be zero if both $\hat k_i$ and $\hat\rho_i$ are zero, and $p(\hat k, \hat\rho, n)/p(\hat k, \hat\rho)$ is considered to be one if both $p(\hat k, \hat\rho, n)$ and $p(\hat k, \hat\rho)$ are zero. Furthermore, for all $\hat\rho, \hat\eta > 0$,

$$\frac{p(\hat k, \hat\rho)}{p(\hat k, \hat\eta)} = \frac{|\hat\eta|}{|\hat\rho|} \prod_{i=1}^{L} \left(\frac{\hat\rho_i}{\hat\eta_i} \mathrm{e}^{-\sum_{j=1}^{L} \alpha_{ij}(\hat\rho_j - \hat\eta_j)}\right)^{\hat k_i}. \tag{5.11}$$

Proof. Recall that $\mathscr{V}$ is the vertex set of $\mathscr{G}(n,\hat{\rho})$ and let $\mathscr{W} \subset \mathscr{V}$ with $\langle \mathscr{W} \rangle = \hat{k}$ so that

$$g(\hat{k},\hat{\rho},n) = \prod_{i,j=1}^{L} \Big(1 - \frac{\alpha_{ij}}{n}\Big)^{\hat{k}_i(\lfloor \hat{\rho}_j n \rfloor - \hat{k}_j)} = \mathrm{e}^{-\hat{\rho}^{\mathrm{T}} \mathrm{A} \hat{k} + O(|\hat{k}|^2/n)} \tag{5.12}$$

is the probability that $\mathscr{W}$ is disconnected from $\mathscr{V} \setminus \mathscr{W}$ in $\mathscr{G}(n,\hat{\rho})$. Note that the error here is uniformly bounded for $\hat{\rho}$ bounded above. Let $(s_{ij})_{1\le i\le j\le L}$ be given and let T be a tree on $\mathscr{W}$ which has s_{ij} edges between $\mathscr{W}_i$ and $\mathscr{W}_j$ for each i and j. Then if v is a vertex uniformly chosen from $\mathscr{V}$, we have

$$P(T \text{ spans } C(v)) = \frac{|\hat{k}|}{\sum_{i=1}^{L} \lfloor \hat{\rho}_i n \rfloor} g(\hat{k},\hat{\rho},n) \prod_{1\le i\le j\le L} \Big(\frac{\alpha_{ij}}{n}\Big)^{s_{ij}}, \tag{5.13}$$

where by 'spans' we means that T and $C(v)$ have the same vertex set, and the edge set of T is contained in that of $C(v)$. Additionally, we have $\sum_{1\le i\le j\le L} s_{ij} = |\hat{k}| - 1$, so

$$P(C(v) = T \big| T \text{ spans } C(v)) = \left[\prod_{i=1}^{L} \Big(1 - \frac{\alpha_{ii}}{n}\Big)^{\binom{\hat{k}_i}{2} - s_{ii}}\right] \left[\prod_{1\le i<j\le L} \Big(1 - \frac{\alpha_{ij}}{n}\Big)^{\hat{k}_i\hat{k}_j - s_{ij}}\right] \tag{5.14}$$

$$= \exp\Big[O\Big(|\hat{k}|^2/n\Big)\Big] \tag{5.15}$$

whenever $P(T \text{ spans } C(v)) > 0$. We note that this error term is independent of $\hat{\rho}$.

Now let $N(\hat{k},(s_{ij}))$ be the number of trees on $\hat{k}$ layered vertices that have s_{ij} edges between layers i and j for each i and j. We define

$$f(\hat{k}) = \frac{|\hat{k}|}{\prod_{i=1}^{L} \hat{k}_i!} \sum_{(s_{ij}):\sum_{i\le j} s_{ij} = |\hat{k}|-1} N(\hat{k},(s_{ij})) \prod_{1\le i\le j\le L} \alpha_{ij}^{s_{ij}}. \tag{5.16}$$

Thus we have

$$p(\hat{k},\hat{\rho},n) = \exp\Big[O\Big(|\hat{k}|^2/n\Big) + O(1/(|\hat{\rho}n|))\Big] \left(\prod_{i=1}^{L} \hat{k}_i!\right) f(\hat{k}) \frac{g(\hat{k},\hat{\rho},n)}{|\hat{\rho}| n^{|\hat{k}|}} \prod_{i=1}^{L} \binom{\lfloor \hat{\rho}_i n \rfloor}{\hat{k}_i}, \tag{5.17}$$

where the $O\big(1/(|\hat{\rho}|n)\big)$ is non-negative, and comes from replacing $\sum_i \lfloor \hat{\rho}_i n \rfloor$ with $|\hat{\rho}|n$ in the denominator. For an upper bound, we expand the binomials and bound $\frac{\lfloor \hat{\rho}_i n \rfloor!}{\lfloor \hat{\rho}_i n - \hat{k}_i \rfloor!}$ above by $(\hat{\rho}_i n)^{\hat{k}_i}$. For a lower bound, we bound this same term below by $(\hat{\rho}_i n)^{\hat{k}_i} \exp\big(-O(\hat{k}_i^2/(\hat{\rho}_i n))\big)$. Now defining $p(\hat{k},\hat{\rho}) = \lim_{n\to\infty} p(\hat{k},\hat{\rho},n)$, we get both (5.10) and (5.11). Note that without knowing $N(\hat{k},(s_{ij}))$, we cannot be more specific about $p(\hat{k},\hat{\rho})$. □

Now define

$$\mathscr{M}(\hat{\eta}) = \Big\{\mu : \mathbb{N}^L \mapsto [0,\infty) \Big| \sum_{\hat{k}} \hat{k}\mu(\hat{k}) = \hat{\eta}\Big\} \tag{5.18}$$

$$\mathscr{M}(\hat{\eta},r) = \Big\{\mu \in \mathscr{M}(\hat{\eta}) \Big| \mu(\hat{k}) = 0 \text{ for } |\hat{k}| > r\Big\}. \tag{5.19}$$

This brings us to

Lemma 5.4. *Given $\delta > 0$, there exists a $c > 0$ so that for all $\hat\eta^{(1)}, \hat\eta^{(2)}$ with $\hat\eta^{(1)} \le \hat\eta^{(2)}$, $\delta \le |\hat\eta^{(1)}| \le |\hat\eta^{(2)}| \le 1/\delta$, and $|\hat\eta^{(2)} - \hat\eta^{(1)}|$ sufficiently small, all $\mu \in \mathscr{M}(\hat\eta^{(2)} - \eta^{(1)}, \epsilon n)$ for some ϵ, and any function $\hat\eta(\hat k)$ with $\hat\eta^{(1)} \le \hat\eta(\hat k) \le \hat\eta^{(2)}$,*

$$\sum_{\hat k} \mu(\hat k) \log \frac{p(\hat k, \hat\eta(\hat k), n)}{p(\hat k, \hat\eta^{(2)})} \le c|\hat\eta^{(2)} - \hat\eta^{(1)}| \left(\epsilon + |\hat\eta^{(2)} - \hat\eta^{(1)}| + \frac{1}{n} \right). \tag{5.20}$$

Also, given $\hat\rho^{(2)} \ge \hat\rho^{(1)} > 0$, there exists a c so that for all $\hat\eta^{(1)}, \hat\eta^{(2)}$ with $\hat\rho^{(1)} \le \hat\eta^{(1)} \le \hat\eta^{(2)} \le \hat\rho^{(2)}$ and $|\hat\eta^{(2)} - \hat\eta^{(1)}|$ sufficiently small, all $\mu \in \mathscr{M}(\hat\eta^{(2)} - \hat\eta^{(1)}, \epsilon n)$ for some ϵ, and any function $\hat\eta(\hat k)$ with $\hat\eta^{(1)} \le \hat\eta(\hat k) \le \hat\eta^{(2)}$,

$$\sum_{\hat k} \mu(\hat k) \log \frac{p(\hat k, \hat\eta(\hat k), n)}{p(\hat k, \hat\eta^{(2)})} \ge -c|\hat\eta^{(2)} - \hat\eta^{(1)}| \left(\epsilon + |\hat\eta^{(2)} - \hat\eta^{(1)}| \right). \tag{5.21}$$

Note that the difference in the requirements of the first and second half of the lemma is that in the first half we require that $\hat\eta$ be bonded away from zero in magnitude, whereas in the second half we require that $\hat\eta$ be bounded away from zero in all components.

Proof. Using (5.10) and the fact that $\mu(\hat k) = 0$ for $|\hat k| > \epsilon n$, we have that

$$- c'\epsilon \sum_{i=1}^{L} \frac{\hat\eta_i^{(2)} - \hat\eta_i^{(1)}}{1 \wedge \hat\eta_i^{(1)}} \le \sum_{\hat k} \mu(\hat k) \log \frac{p(\hat k, \hat\eta(\hat k), n)}{p(\hat k, \hat\eta(\hat k))} \le c' \left(\epsilon + \frac{1}{|\hat\eta^{(1)}| n} \right), \tag{5.22}$$

where c' is the c in (5.10). Then using (5.11) we have

$$\begin{aligned} \sum_i (\hat\eta_i^{(2)} - \hat\eta_i^{(1)}) \log \left(\frac{\hat\eta_i^{(1)}}{\hat\eta_i^{(2)}} \right) &\le \sum_{\hat k} \mu(\hat k) \log \frac{p(\hat k, \hat\eta(\hat k))}{p(\hat k, \hat\eta^{(2)})} \\ &\le \log \left(\frac{|\hat\eta^{(2)}|}{|\hat\eta^{(1)}|} \right) + (\hat\eta^{(2)} - \hat\eta^{(1)})^{\mathrm{T}} \mathrm{A} (\hat\eta^{(2)} - \hat\eta^{(1)}). \end{aligned} \tag{5.23}$$

Combining (5.22) and (5.23) gives the result. □

Lemma 5.5. *Let $\hat\rho, \hat\gamma$ be given with $\hat\rho, \hat\gamma \ge 0$ componentwise and $|\hat\gamma| = 1$. Then we have*

$$\sup_{\mu \in \mathscr{M}(\hat\gamma)} \sum_{\hat k} \mu(\hat k) \log \frac{p(\hat k, \hat\rho) \sum_{\hat\ell} \mu(\hat\ell)}{\mu(\hat k)} = \xi(\hat\gamma, \hat\rho) - \xi(\hat\gamma, \hat\psi(\hat\gamma, |\hat\rho|)). \tag{5.24}$$

Proof. To reduce clutter, let us write $\hat{\psi}$ in place of $\hat{\psi}(\hat{\gamma}, |\hat{\rho}|)$. Then using (5.11) to rewrite $p(\hat{k}, \hat{\rho})$ in terms of $p(\hat{k}, \hat{\psi})$, we have

$$\sum_{\hat{k}} \mu(\hat{k}) \log \frac{p(\hat{k}, \hat{\rho}) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})} = \Bigg[\sum_{\hat{k}} \mu(\hat{k}) \log \frac{\frac{|\hat{\psi}|}{|\hat{\rho}|} p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})} + \sum_{\hat{k}} \mu(\hat{k}) \sum_{i=1}^{L} \hat{k}_i \left(\log \frac{\hat{\rho}_i}{\hat{\psi}_i} - \sum_{j=1}^{L} \alpha_{ij}(\hat{\rho}_j - \hat{\psi}_j)\right)\Bigg]. \tag{5.25}$$

Due to the requirement that $\sum \hat{k}\mu(\hat{k}) = \hat{\gamma}$ for each $\mu \in \mathscr{M}(\hat{\gamma})$, the second sum on the right hand side evaluates to $\xi(\hat{\gamma}, \hat{\rho}) - \xi(\hat{\gamma}, \hat{\psi})$. Thus it remains to be proved that

$$\sup_{\mu \in \mathscr{M}(\hat{\gamma})} \sum_{\hat{k}} \mu(\hat{k}) \log \frac{\frac{|\hat{\psi}|}{|\hat{\rho}|} p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})} = 0, \tag{5.26}$$

which we shall do in two parts. We start by showing that the supremum is at least zero.

We recall that $\hat{\psi}$ is either subcritical or critical. In the former case we take $\mu(\hat{k}) = \frac{p(\hat{k}, \hat{\psi})}{\sum_{\hat{\ell}} |\hat{\ell}| p(\hat{k}, \psi)}$, in which case (5.4) shows that $\mu \in \mathscr{M}(\hat{\gamma})$. Then using $|\hat{\psi}| = |\hat{\rho}|$, we have

$$\sum_{\hat{k}} \mu(\hat{k}) \log \frac{\frac{|\hat{\psi}|}{|\hat{\rho}|} p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})} = 0. \tag{5.27}$$

For the critical case, let ϵ be given and define

$$\nu_\epsilon(\hat{k}) = \frac{p(\hat{k}, (1-\epsilon)\hat{\rho})}{\sum_{\hat{\ell}} |\hat{\ell}| p(\hat{\ell}, (1-\epsilon)\hat{\psi})}, \tag{5.28}$$

and note that (5.4) gives that $\lim_{\epsilon \to 0} \sum_{\hat{k}} \hat{k} \nu_\epsilon(\hat{k}) = \hat{\gamma}$. Thus for small ϵ we can pick a uniformly bounded $\delta_\epsilon(\hat{k})$ that is $o(1)$ as $\epsilon \to 0$ such that $\sum_{\hat{k}} \big(1 + \delta_\epsilon(\hat{k})\big) \hat{k} \nu_\epsilon(\hat{k}) = \hat{\gamma}$. Then defining $\mu_\epsilon = (1 + \delta_\epsilon)\nu_\epsilon$, we have that $\mu_\epsilon \in \mathscr{M}(\hat{\gamma})$. Furthermore,

$$\left| \sum_{\hat{k}} \mu_\epsilon(\hat{k}) \log \frac{p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu_\epsilon(\hat{\ell})}{\mu_\epsilon(\hat{k})} \right| \leq o(1) + \sum_{\hat{k}} \mu_\epsilon(\hat{k}) \left| \log \frac{p(\hat{k}, \hat{\psi})}{p\big(\hat{k}, (1-\epsilon)\hat{\psi}\big)} \right|. \tag{5.29}$$

Using (5.11) we can see that the summand on the right hand side is uniformly $O(\epsilon)|\hat{k}|\mu_\epsilon(\hat{k})$, so recalling that $\sum_{\hat{k}} |\hat{k}| \mu_\epsilon(\hat{k}) = 1$ and taking limits, we get

$$\lim_{\epsilon \to 0} \left| \sum_{\hat{k}} \mu_\epsilon(\hat{k}) \log \frac{p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu_\epsilon(\hat{\ell})}{\mu_\epsilon(\hat{k})} \right| = 0. \tag{5.30}$$

Thus we've shown that the supremum in (5.26) is at least zero for all $\hat{\gamma}$ and $\hat{\rho}$.

To show that the supremum is at most zero, we will drop the $\frac{|\hat{\psi}|}{|\hat{\rho}|} \le 1$ term and show

$$\sum_{\hat{k}} \mu(\hat{k}) \log \frac{p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})} \le 0. \tag{5.31}$$

For all μ (whether in $\mathscr{M}(\hat{\gamma})$ or not). To do this, we fix μ and consider

$$f(t) = \sum_{\hat{k}} \left[(1-t)\mu(\hat{k}) + t\nu(\hat{k})\right] \log \frac{\nu(\hat{k})}{(1-t)\mu(\hat{k}) + t\nu(\hat{k})}, \tag{5.32}$$

where $\nu(\hat{k}) = p(\hat{k}, \hat{\psi}) \sum_{\hat{\ell}} \mu(\hat{\ell})$. Then noting that $\sum_{\hat{k}} \mu(\hat{k}) = \sum_{\hat{k}} \nu(\hat{k})$, we find

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = \sum_{\hat{k}} \left[\mu(\hat{k}) - \nu(\hat{k})\right] \log \left[t + (1-t)\frac{\mu(\hat{k})}{\nu(\hat{k})}\right], \tag{5.33}$$

in which each summand is non-negative for all $t \in [0,1]$. Thus f is increasing on the interval $[0,1]$, and a quick check confirms that $f(0)$ is the left hand side of (5.31) while $f(1)$ is the right hand side. This proves (5.31), which finishes the proof of (5.26), which finishes the proof of the lemma. □

For the proof of Theorem 5.2, we will need to introduce a discrete version of $\mathscr{M}(\hat{\gamma}, r)$. Let $M(\hat{b}, r)$ be the set of sequences indexed by elements of $\mathbb{N}^L \setminus \{0\}$ given by

$$M(\hat{b}, r) = \Big\{(m_{\hat{k}})_{\hat{k} \in \mathbb{N}^L \setminus \{0\}} : \sum_{\hat{k}} \hat{k} m_{\hat{k}} = \hat{b};\, m_{\hat{k}} = 0 \text{ for } |\hat{k}| > r\Big\}. \tag{5.34}$$

We are then in need of a bound on the size of M:

Lemma 5.6. $|M(\hat{b}, r)| = \exp\Big(O\Big(n^{\frac{L+1}{L+2}} (\log n)^2\Big)\Big)$, *where* $n = |\hat{b}|$, *uniformly in* r. *In particular,* $\frac{1}{n} \log |M(\hat{b}, r)| = o(1)$.

Proof. We will prove this by instead bounding the size of the larger set

$$Q(n) = \big\{(m_{\hat{k}})_{\hat{k} \in \mathbb{N}^L \setminus \{0\}} : \sum_{\hat{k}} |\hat{k}| m_{\hat{k}} \le n\big\}, \tag{5.35}$$

which does not depend on r at all. Let $J_j = \{\hat{k} \in \mathbb{N}^L : 2^j \le |\hat{k}| < 2^{j+1}\}$, and let $w(a,b) = |\{\hat{k} \in \mathbb{N}^b : |\hat{k}| = a\}| = \binom{a+b-1}{b-1}$. We then note that for each $(m_{\hat{k}}) \in Q(n)$, we must have $\sum_{\hat{k} \in J_j} m_{\hat{k}} \le n 2^{-j}$ for all j. Thus we have

$$|Q(n)| \le \prod_{j=0}^{\infty} \sum_{a=0}^{\lfloor n 2^{-j} \rfloor} w\big(a, |J_j|\big) \tag{5.36}$$

$$\le \prod_{j=0}^{\lfloor \log_2 n \rfloor} n w\big(\lfloor n 2^{-j} \rfloor, |J_j|\big). \tag{5.37}$$

We next note that $w(a,b) \leq (1+a/b)^b(1+b/a)^a$, and from this we get the useful bound $w(a,b) \leq (2\mathrm{e}m_1/m_2)^{m_2}$, where $m_1 = \max(a,b)$ and $m_2 = \min(a,b)$. This allows us to estimate $|J_j| \leq c_L 2^{j(L+1)}$, and so

$$w\big(\lfloor n2^{-j}\rfloor, |J_j|\big) \leq n^{O\left(n^{\frac{L+1}{L+2}}\right)} \tag{5.38}$$

for $j = 0, \ldots, \lfloor \log_2 n\rfloor$. Combined with (5.37), this gives us the result. □

5.3 The Rate for Υ.

Proof of Theorem 5.2. We will prove this by getting an upper and lower bound, although we will only show the upper bound explicitly. The proof of the upper bound requires a bit of boot strapping, and the first half of the process is to show that for any $\delta > 0$ there is a $c = c(\hat{\eta})$ so that

$$\frac{1}{n}\log \Upsilon\big(\hat{\eta}^{(2)}, \hat{\eta}^{(1)}, \epsilon n, n\big) \leq |\hat{\gamma}|\left[o(1) + c|\hat{\gamma}| + \Big(\xi(\hat{\gamma}, \hat{\eta}^{(2)}) - \xi\big(\hat{\gamma}, \hat{\psi}(\hat{\gamma}/|\hat{\gamma}|, |\hat{\eta}^{(2)}|)\big)\Big)\right], \tag{5.39}$$

for all $\hat{\rho} \geq \hat{\eta}^{(2)} \geq \hat{\eta}^{(1)}$ with $|\hat{\eta}^{(1)}| \geq \delta$ and sufficiently small $\hat{\gamma} = \hat{\eta}^{(2)} - \hat{\eta}^{(1)}$, where the $o(1)$ term is uniformly bounded for $\hat{\eta}^{(1)}, \hat{\eta}^{(2)}$ in this domain, and tends to zero as n tends to infinity and ϵ tends to zero.

We begin by setting $\hat{b} = \lfloor \hat{\eta}^{(2)} n\rfloor - \lfloor \hat{\eta}^{(1)} n\rfloor$, and noting that the event upon which Υ is based occurs if and only if there is some $(m_{\hat{k}}) \in M(\hat{b}, \epsilon n)$ so that exactly $m_{\hat{k}}$ of $C^{(1)}, C^{(2)}, \ldots, C^{(\sum m_{\hat{\ell}})}$ have $\langle C^{(\cdot)}\rangle = \hat{k}$. Thus we have

$$\Upsilon(\hat{\rho}, \hat{\eta}, \epsilon n, n) = \sum_{(m_{\hat{k}}) \in M(\hat{b}, \epsilon n)} \frac{(\sum m_{\hat{k}})!}{\prod m_{\hat{k}}!}\left[\prod p\big(\hat{k}, \hat{\eta}^{(2)} - O\big(|\hat{\gamma}|\big), n\big)^{m_{\hat{k}}}\right], \tag{5.40}$$

where the $O\big(|\hat{\gamma}|\big)$ term is strictly bounded in magnitude by $|\hat{\gamma}|$. Now let $(m_{\hat{k}}) \in M(\hat{b}, \epsilon n)$ be chosen to maximize the summand. Since Lemma 5.6 gives us $|M(\hat{b}, \epsilon n)| \leq \mathrm{e}^{|\hat{\gamma}| o(n)}$, we have

$$\log \Upsilon(\hat{\rho}, \hat{\eta}, \epsilon n, n) \leq |\hat{\gamma}| o(n) + \log \frac{(\sum m_{\hat{k}})!}{\prod m_{\hat{k}}!}\left[\prod p\big(\hat{k}, \hat{\eta}^{(2)} - O\big(|\hat{\gamma}|\big), n\big)^{m_{\hat{k}}}\right], \tag{5.41}$$

where the $o(n)$ term is uniformly bounded in $\hat{\eta}^{(1)}$, $\hat{\eta}^{(2)}$, and ϵ as $n \to \infty$. Now dividing by n and using Stirling's approximation, we get that there exists a $\mu \in \mathscr{M}\big(\frac{1}{n}\hat{b}, \epsilon n\big)$ with

$$\frac{1}{n}\log \Upsilon(\hat{\rho}, \hat{\eta}, \epsilon n, n) \leq |\hat{\gamma}| o(1) + \sum_{\hat{k}} \mu(\hat{k}) \log \frac{p\big(\hat{k}, \hat{\eta}^{(2)} - O\big(|\hat{\gamma}|\big), n\big) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})}. \tag{5.42}$$

Note that the dangerous looking $\prod m_{\hat{k}}!$ in (5.41) does not cause us trouble – because it is in the denominator – and the $o(1)$ term is still uniformly bounded. For the equivalent step in the lower bound, one would use the fact that $(m_{\hat{k}}) \in M(\hat{b}, r)$ to bound the error. Now applying (5.20) from Lemma 5.4 with $\delta = \min(|\hat{\eta}|, 1/|\hat{\rho}|)$, we get

$$\frac{1}{n}\log \Upsilon(\hat{\rho}, \hat{\eta}, \epsilon n, n) \leq |\hat{\gamma}| o(1) + c|\hat{\gamma}|(\epsilon + |\hat{\gamma}| + 1/n) + \sum_{\hat{k}} \mu(\hat{k}) \log \frac{p(\hat{k}, \hat{\rho}) \sum_{\hat{\ell}} \mu(\hat{\ell})}{\mu(\hat{k})} \tag{5.43}$$

for some c depending only on δ. By scaling μ by a factor of $n/|\hat{b}|$ we can apply Lemma 5.5 and use the continuity of ξ – which covers the $O(1/n)$ gap between $\hat{\gamma}$ and $\hat{b}/|\hat{b}|$ – to get (5.39).

We next note that if all components (discovered in the Υ process) in $\mathscr{G}(n,\hat{\rho})$ are of size smaller that ϵn, then for every $x \in (|\hat{\eta}|, |\hat{\rho}|)$ there must be at least one k so that

$$\big(|\hat{\rho}| - x\big)n \le \sum_{i=1}^{k} |C^{(i)}| \le \big(|\hat{\rho}| - x\big)n + \epsilon n. \tag{5.44}$$

Moreover, $P(|C^{(i)}| = 1)$ is uniformly (in $\hat{\rho}$, for $\hat{\rho}$ bounded above) bounded away from zero for all i and so we have for any given collection of integers $m_1, \ldots, m_j \in \big[|\hat{\eta}n|, |\hat{\rho}n|\big]^1$ there exist corresponding $k_1, \ldots, k_j$ so that $\sum_{i=1}^{k_\ell} |C^{(i)}| = m_\ell$ for each ℓ with probability $\mathrm{e}^{jO(\epsilon n)}$. Thus for all j the independence of edges in $\mathscr{G}(n,\hat{\rho})$ lets us write

$$\Upsilon(\hat{\rho}, \hat{\eta}, \epsilon n, n) \le \mathrm{e}^{j[O(\log n) + O(\epsilon n)]} \sup_{\substack{\hat{\rho} = \hat{\rho}^{(j)} \ge \ldots \ge \hat{\rho}^{(0)} = \hat{\eta} \\ \left|\hat{\rho}^{(i)}\right| = |\hat{\eta}| + \frac{i}{j}|\hat{\rho} - \hat{\eta}|}} \prod_{i=0}^{j-1} \Upsilon\Big(\hat{\rho}^{(i+1)}, \hat{\rho}^{(i)}, \epsilon n, n\Big), \tag{5.45}$$

where the $\mathrm{e}^{O(\log n)}$ term – which comes from the number of choices for each $\hat{\rho}^{(i)}$ – is uniformly bounded for $\hat{\rho}$ bounded above. We note at this point that each Υ term has a trivial upper bound of 1, and so – in preparation for the application of (5.39) – we will discard terms on the right hand side of (5.45) for which $|\hat{\rho}^{(i)}| < \delta$. Furthermore, to simplify notation, let us consider only $\hat{\rho}$ for which $|\hat{\rho}| \ge \delta$ at the moment. Then, because Υ discretizes the $\hat{\rho}^{(i)}$ terms, the supremum can be considered to be taken over a finite set, and thus there is a specific choice of $\big(\hat{\rho}^{(i)}\big)$ which achieves the supremum. Thus with

$$f(\hat{\rho}, \hat{\eta}, r, n) = \frac{1}{n} \log \Upsilon(\hat{\rho}, \hat{\eta}, r, n), \tag{5.46}$$

we have for all ϵ, n, j the existence of a partition $\hat{\eta} = \hat{\rho}^{(0)} \le \ldots \le \hat{\rho}^{(j)} = \hat{\rho}$ with $\left|\hat{\rho}^{(i)}\right| = |\hat{\eta}| \vee \delta + \frac{i}{j}\big(|\hat{\rho}| - |\hat{\eta}| \vee \delta\big)$ such that

$$f(\hat{\rho}, \hat{\eta}, \epsilon n, n) \le jo(1) + \sum_{i=0}^{j-1} f\Big(\hat{\rho}^{(i+1)}, \hat{\rho}^{(i)}, \epsilon n, n\Big). \tag{5.47}$$

Interpolating by straight lines, is is clear that $(\hat{\rho}^{(i)} \mid i = 0, \ldots j)$ defines an assent restricted to the interval $[|\hat{\eta}| \vee \delta, |\hat{\rho}|]$ which we denote by $\hat{\Gamma}$ (remembering that $\hat{\Gamma}$ depends on ϵ, n, and j). Let $t^{(j,i)} = |\hat{\eta}| \vee \delta + \frac{i}{j}\big(|\hat{\rho}| - |\hat{\eta}| \vee \delta\big)$ We now apply (5.39) to the right hand side of (5.47) to get

$$\begin{aligned} f(\hat{\rho}, \hat{\eta}, \epsilon, n) \le jo(1) + O(1/j) + \big(|\hat{\rho}| - |\hat{\eta}| \vee \delta\big) \frac{1}{j} \sum_{i=0}^{j-1} \Bigg[\xi \left(\frac{\hat{\Gamma}\big(t^{(j,i+1)}\big) - \hat{\Gamma}\big(t^{(j,i)}\big)}{j\big(|\hat{\rho}| - |\hat{\eta}| \vee \delta\big)}, \hat{\Gamma}\big(t^{(j,i)}\big) \right) \\ - \xi \left(\frac{\hat{\Gamma}\big(t^{(j,i+1)}\big) - \hat{\Gamma}\big(t^{(j,i)}\big)}{j\big(|\hat{\rho}| - |\hat{\eta}| \vee \delta\big)}, \hat{\psi}\bigg(\frac{\hat{\Gamma}\big(t^{(j,i+1)}\big) - \hat{\Gamma}\big(t^{(j,i)}\big)}{j\big(|\hat{\rho}| - |\hat{\eta}| \vee \delta\big)}, t^{(j,i)} \bigg) \right) \Bigg], \end{aligned} \tag{5.48}$$

[1]The range we are actually interested in is $\big[|\lfloor \hat{\eta}n \rfloor|, |\lfloor \hat{\rho}n \rfloor|\big]$, but the floors seem to distract from the presented idea.

where both error terms are uniformly bounded for $\hat{\rho}$ bounded above. It is noted that the substantive term in (5.48) is perfectly well defined, e.g. $\hat{\Gamma}_k\big(t^{(j,i+1)}\big) - \hat{\Gamma}_k\big(t^{(j,i)}\big)$ vanishes if and only if the the k^{th} component of the corresponding ψ vanishes – thence no difficulty with interpretation of logs.

We next claim that, for all intents and purposes, the "substantive" term is just $H(\hat{\Gamma}, |\hat{\eta}| \vee \delta, |\hat{\rho}|)$. This is not quite exact since if $\hat{\psi}(\hat{\Gamma}', t)$ is subcritical, at time $t^{(j,i)}$ it will change a bit over the course of $[t^{(j,i)}, t^{(j,i+1)}]$. However, in these circumstance, the change in $\hat{\psi}$ is (at most) proportional to the change in t itself and/or the change in χ^{-1} – which in turn is bounded by the change in t. Moreover, the object of proportion for the k^{th} component is the k^{th} component itself which again alleviates any concerns about the singularities associated with logarithms. On this basis it is seen that, in the course of each increment, the error incurred by replacing the appropriate term in (5.48) by the integration of $\xi(\hat{\Gamma}', \hat{\Gamma}) - \xi(\hat{\Gamma}', \hat{\psi})$ along the corresponding portion of the path is, in fact, bounded by a uniform constant times $t^{(j,i+1)} - t^{(j,i+1)}$. We thus have

$$f(\hat{\rho}, \hat{\eta}, \epsilon n, n) \le j o(1) + O(1/j) + H(\hat{\Gamma}, |\hat{\eta}| \vee \delta, |\hat{\rho}|) \tag{5.49}$$

where the additional small error terms have been incorporated into the $O(1/j)$ term. Obviously we may replace $H(\hat{\Gamma}, |\hat{\eta}| \vee \delta, |\hat{\rho}|)$ by the supremum over available $\hat{\Gamma}$ and, allowing j to be a considered as a function of ϵ and n, we can replace both error terms with a single $o(1)$ term which tends to zero uniformly – for $\hat{\rho}$ bounded above – as ϵ tends to zero and n tends to infinity. Furthermore, it is straightforward to see that the maximal difference between $H(\hat{\Gamma}, |\hat{\eta}| \vee \delta, |\hat{\rho}|)$ and $H(\hat{\Gamma}, |\hat{\eta}|, |\hat{\rho}|)$ is uniformly bounded – that is, for all $\hat{\eta}$, $\hat{\rho}$, and assents $\hat{\Gamma}$ – by an $O(\delta)$ correction, for small δ. Since we may retroactively declare δ to have been picked as small as desired – after which we may take ϵ to be small and n large, we find

$$\frac{1}{n} \log \Upsilon\big(\hat{\rho}, \hat{\eta}, \epsilon n, n\big) \le o(1) + \sup_{\hat{\Gamma} \in \mathscr{P}(\hat{\eta}) \cap \mathscr{P}(\hat{\rho})} H(\hat{\Gamma}, |\hat{\eta}|, |\hat{\rho}|), \tag{5.50}$$

where the $o(1)$ term converges uniformly to zero as $\epsilon \to 0$ and $n \to \infty$.

The opposite bound, namely

$$\frac{1}{n} \log \Upsilon(\hat{\rho}, \hat{\eta}, r, n) \ge o(1) + \sup_{\hat{\Gamma} \in \mathscr{P}(\hat{\eta}) \cap \mathscr{P}(\hat{\rho})} H\big(\hat{\Gamma}, |\hat{\eta}|, |\hat{\rho}|\big), \tag{5.51}$$

where the $o(1)$ term tends to zero uniformly as $r, n \to \infty$, is derived by the similar methods. First, an analog opposite bound of the form in (5.39) is derived for small $|\hat{\gamma}|$ – although due to the weaker lower bound of Lemma 5.4 the bound only holds uniformly for $\hat{\eta}^{(1)}$ larger that δ in all components. Then, at the point of (5.45) we can simply pick any particular assent from $\hat{\eta} \vee (\delta, \ldots, \delta)$ to $\hat{\rho}$ and place the points $\hat{\rho}^{(i)}$ along this path. The rest of the argument is identical with the final step being an optimization over assents, and using brute force to traverse the distance from $\hat{\eta} \vee (\delta, \ldots, \delta)$ to $\hat{\eta}$. □

5.4 Natural Assents.

Consider an assent, $\hat{\Gamma}(t)$ which starts at the origin and goes at least some distance into the supercritical region. Then (since at least one component increases) there is a unique $t^{\star}$ such that for $t < t^\star$, $\hat{\Gamma}(t)$ is subcritical while for $t > t^\star$, it is supercritical. To define a *natural assent*, we shall treat separately the subcritical and super critical behaviors. Starting with the former, it is stipulated that

$$\hat{\Gamma}'(t) = \frac{\left[\mathrm{I} - \mathrm{B}_{\hat{\Gamma}(t)}\right]^{-1} \hat{\Gamma}(t)}{\left|\left[\mathrm{I} - \mathrm{B}_{\hat{\Gamma}(t)}\right]^{-1} \hat{\Gamma}(t)\right|} \tag{5.52}$$

at least for $\hat{\Gamma} \neq 0$. (As is not hard to see, if we actually wish to start the assent at the origin, an initial direction must also be specified.) Notice that (5.52) implies (c.f. (5.1) – (5.4)) that $\hat{\psi}(\hat{\Gamma}', \hat{\Gamma}) = \hat{\Gamma}$ (while, of course, the associated average cluster size ratios, is proportional to $\hat{\Gamma}'$). In the limit as $t \uparrow t^\star$ we find that $\hat{\Gamma}'$ tends to the limiting size ratio associated with the critical density. Denoting these objects by $\hat{\gamma}_c$ and $\hat{\psi}_c$ respectively (so that $\hat{\gamma}_c$ is the maximum eigenvector of $\mathrm{B}_{\hat{\psi}_c}$) we define, for $t > t^\star$ $\hat{\Gamma}' \equiv \hat{\gamma}_c$ – with $\hat{\Gamma}'' \equiv 0$ – i.e.

$$\hat{\Gamma}(t) = \hat{\psi}_c + (t - t^\star)\hat{\gamma}_c. \tag{5.53}$$

Lemma 5.7. *Let $\hat{\rho}^{(2)} > \hat{\rho}^{(1)} \geq 0$ be given, and let $\hat{\Gamma} \in \mathscr{P}\big(\hat{\rho}^{(1)}\big) \cap \mathscr{P}\big(\hat{\rho}^{(2)}\big)$ be an assent such that*

$$H\big(\hat{\Gamma}, |\hat{\rho}^{(1)}|, |\hat{\rho}^{(2)}|\big) = \sup_{\hat{\Phi} \in \mathscr{P}(\hat{\rho}^{(1)}) \cap \mathscr{P}(\hat{\rho}^{(2)})} H\big(\hat{\Phi}, |\hat{\rho}^{(1)}|, |\hat{\rho}^{(2)}|\big). \tag{5.54}$$

Then there is some $\hat{c}$ with $\sum_{i=1}^{L} \hat{c}_i = 0$ such that $\hat{\Gamma} + \hat{c}$ is a natural assent restricted to $\big[|\hat{\rho}^{(1)}|, |\hat{\rho}^{(2)}|\big]$.

Proof. Let us abbreviate $a = |\hat{\rho}^{(1)}|$ and $b = |\hat{\rho}^{(2)}|$. It is first noted that for any assent $\hat{\Gamma}$ connecting $\hat{\rho}^{(1)}$ to $\hat{\rho}^{(2)}$ the portion of H consisting of $\int_a^b \xi(\hat{\Gamma}', \Gamma)$ is a constant c which depends only on $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$ and hence need not be further discussed. Our first goal is to establish lower bounds on the speed. Of course it may be the case that $\hat{\rho}_i^{(2)} = \hat{\rho}_i^{(1)}$ for one or more values of i in which case, in all available choices of assents, $\hat{\Gamma}_i = \text{const}$. We claim that in all other circumstances, the speeds are, componentwise, uniformly bounded below (recalling once more that our parameterization gives assents whose total speed is always one). We adopt the notation $\hat{\Gamma}' = \hat{\gamma}$ and, subtracting and adding $\sum_i \hat{\gamma}_i \log \hat{\gamma}_i$, it is seen that the remains of the integrand (after the $-\hat{\gamma}_i \log \hat{\gamma}_i$ term) are non–singular, with non–singular derivatives as any particular $\hat{\gamma}_i \to 0$. However the "principal" term is concave with a singular derivative ($\propto -\log \hat{\gamma}_i$) which, as we shall see, does not permit any component of $\hat{\gamma}_i$ to get too small. Indeed, let γ_i be non–trivial and suppose there is a set of size Δt, on which $\hat{\gamma}_i$ does not exceed some ε with $\varepsilon \ll 1$. Let us find another portion of the path where γ_i exceeds, half the total required rate of assent:

$$\hat{\gamma}_i > \frac{1}{2} \frac{\hat{\rho}_i^{(2)} - \hat{\rho}_i^{(1)}}{b - a}.$$

Here we may have to assume that Δt is not too large which, obviously, we may do without loss of generality. Denoting the two sets by $\underline{b}$ and $\underline{s}$ (big and small) we may consider the canonical map from $\underline{b}$ to $\underline{s}$ and by this means, replacing $\hat{\gamma}_i|_{\underline{b}}$ with $(1 - \varepsilon)\hat{\gamma}_i|_{\underline{b}}$ we can increase $\hat{\gamma}_i|_{\underline{s}}$ by an amount of order ε. The gain from this transfer, by consideration of the principal part along the set $\underline{s}$, is of the order $|\varepsilon \log \varepsilon|$. Meanwhile the "losses" from the other parts of the functional on both $\underline{s}$ and $\underline{b}$ are bounded above by a constant times ε for all ε sufficiently small. Of course in addition, we must now recalibrate to unit speed but this causes changes in the above effects which are also only of the order of ε. Thus, along a minimizer for any i in which $\hat{\rho}_i^{(2)} - \hat{\rho}_i^{(1)} > 0$, we may conclude that there is an $\varepsilon_0 > 0$ – which will depend on $\hat{\rho}^{(2)}$ and $\hat{\rho}^{(1)}$ – such that $\hat{\gamma}_i > \varepsilon_0$ for a.e. t.

With the above in hand, we may now add a perturbation to the minimizing assent secure in the knowledge that, for sufficiently small perturbation, the resultant function is indeed an assent. The natural procedure is to derive Euler–Lagrange equations but, unfortunately, there is no *a priori* guarantee that all required partial derivatives exist. For this reason we deviate from the usual method at the point where an integration by parts would normally be performed.

Let $\hat{g} : (0, \infty)^L \mapsto \mathbb{R}^L$ be a Lipschitz continuous function with $\hat{g}(a) = \hat{g}(b) = 0$ and $\sum_{i=1}^L g(t) = 0$ for $t \in (a, b)$. Then for all sufficiently small δ, we may write

$$H(\hat{\Gamma} + \delta\hat{g}, a, b) = c - \int_a^b \xi\left((\hat{\Gamma} + \delta\hat{g})'(t), \hat{\psi}\big((\hat{\Gamma} + \delta\hat{g})'(t), t\big)\right) \mathrm{d}t. \tag{5.55}$$

Now suppose that $\hat{\gamma}, t > 0$ are given with $|\hat{\gamma}| = 1$. We recall that $\chi^{-1}(\hat{\gamma}, t)$ and $\hat{\psi}(\hat{\gamma}, t)$ are defined in such a way that either $|\hat{\psi}(\hat{\gamma}, t)| = t$ or $\hat{\chi}^{-1}(\hat{\gamma}, t) = 0$. Using this, we have

$$\xi(\hat{\gamma}, \hat{\psi}(\hat{\gamma}, t)) = -1 + \chi^{-1}(\hat{\gamma}, t) + \sum_{i=1}^L \hat{\gamma}_i \log \hat{\psi}_i(\hat{\gamma}, t), \tag{5.56}$$

but we here we encounter our first problem with differentiability. Let $\hat{\eta}$ be given with $\sum_{i=1}^L \hat{\eta}_i = 0$. Then while $\frac{\partial}{\partial\delta}\hat{\psi}(\hat{\gamma} + \delta\hat{\eta}, t)|_{\delta=0}$ and $\frac{\partial}{\partial\delta}\chi^{-1}(\hat{\gamma} + \delta\hat{\eta}, t)|_{\delta=0}$ are well defined in the region where $|\hat{\psi}| \neq t$ and the region where $\chi^{-1} > 0$, the limiting values may not agree on the boundary. Nevertheless, a careful calculation shows that the discontinuities in the derivative arising from the last two terms in (5.56) cancel each other exactly, yielding

$$\frac{\partial}{\partial\delta}\xi\left(\hat{\gamma} + \delta\hat{\eta}, \hat{\psi}\left(\hat{\gamma} + \delta\hat{\eta}, t\right)\right)\bigg|_{\delta=0} = \xi\left(\hat{\eta}, \hat{\psi}\left(\hat{\gamma}, t\right)\right). \tag{5.57}$$

For brevity, let us for the remained of the proof write $\hat{\psi}(t)$ for $\hat{\psi}(\hat{\Gamma}'(t), t)$. Then by combining (5.54) and (5.57), we see that any $\hat{\Gamma}$ which maximizes $H(\hat{\Gamma}, a, b)$ must have

$$\int_a^b \xi\left(\hat{g}'(t), \hat{\psi}(t)\right) \mathrm{d}t = 0. \tag{5.58}$$

Since this must hold for all Lipschitz $\hat{g}$ with $\hat{g}(a) = \hat{g}(b) = 0$ and $\sum_{i=1}^L \hat{g}_i(t) = 0$, we can take $\hat{g}_i(t) = -\hat{g}_j(t) = g(t)$ and conclude

$$\int_a^b g'(t)\left[\left(\log\hat{\psi}_i(t) - \sum_{k=1}^L \alpha_{ik}\hat{\psi}_k(t)\right) - \left(\log\hat{\psi}_j(t) - \sum_{k=1}^L \alpha_{jk}\hat{\psi}_k(t)\right)\right] \mathrm{d}t = 0 \tag{5.59}$$

for every Lipschitz function g with $g(a) = g(b) = 0$. Thus the difference in the integrand must be constant almost everywhere, meaning that there must exist some $\hat{d}$ and $f(t)$ such that

$$\log\hat{\psi}_i(t) - \sum_{k=1}^L \alpha_{ik}\hat{\psi}_k(t) = \hat{d}_i + f(t) \tag{5.60}$$

for almost every t, for each i.

Since it will clean up the proof without making a substantial difference, let us suppose that (5.60) holds for every t, instead of merely almost every t. We then claim (the proof of which shall be postponed) that if $\hat{T}(\hat{x})$ is defined for $\hat{x}$ critical or subcritical with

$$\hat{T}_i(\hat{x}) = \log \hat{x}_i - \sum_{k=1}^{L} \alpha_{ik}\hat{x}_k, \tag{5.61}$$

then $\hat{T}$ is invertible and $\hat{T}^{-1}$ is increasing. Thus if t_0 is chosen so that $\hat{\psi}(t_0)$ is critical, we must have that $f(t) \leq f(t_0)$ – and thus $\hat{\psi}(t) \leq \hat{\psi}(t_0)$ – for all t. Furthermore, if $\hat{\psi}(t)$ is subcritical, we have $t = |\hat{\psi}(t)| < |\hat{\psi}(t_0)| \leq t_0$. From this we conclude that if $t^\star = \inf\{t : \hat{\psi}(t) \text{ is critical}\}$, then $\hat{\psi}(t)$ is critical and constant for $t > t^\star$. Furthermore, since $|\hat{\psi}(t)| = t$ for $t < t^\star$, we must have that f is increasing, and thus $\hat{\psi}(t)$ is increasing in all components. Thus $\hat{\psi}$ is Lipschitz continuous.

Since $\hat{\psi}$ is Lipschitz continuous, it is almost everywhere differentiable. Then differentiating both sides of (5.60) and multiplying by $\hat{\psi}_i(t)$, we get

$$\left[\mathrm{I} - \mathrm{B}_{\hat{\psi}(t)}\right] \hat{\psi}'(t) = f'(t)\hat{\psi}(t). \tag{5.62}$$

Multiplying both sides by $\left[\mathrm{I} - \mathrm{B}_{\hat{\psi}(t)}\right]^{-1}$ and recalling that $\frac{\mathrm{d}}{\mathrm{d}t}|\hat{\psi}(t)| = 1$ for $t < t^\star$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\psi}(t) = \frac{\left[\mathrm{I} - \mathrm{B}_{\hat{\psi}(t)}\right]^{-1} \hat{\psi}(t)}{\left|\left[\mathrm{I} - \mathrm{B}_{\hat{\psi}(t)}\right]^{-1} \hat{\psi}(t)\right|} \tag{5.63}$$

for $t < t^\star$. From (5.4), we see that the right hand side of (5.63) is equal to $\hat{\Gamma}'(t)$ for some natural assent Γ. Integrating gives that $\hat{\Gamma}(t) + \hat{c} = \hat{\psi}(t)$ for $t < t^\star$, for some $\hat{c}$. From (5.60) we can see that $\hat{\Gamma}'(t)$ is continuous, and since $\hat{\psi}(t)$ is constant for $t > t^\star$, we have that $\hat{\Gamma}'(t)$ is the maximum eigenvector of $\mathrm{B}_{\hat{\psi}(t^\star)}$ for $t > t^\star$. Combined with the fact that (5.63) shows that $\hat{\psi}(t)$ restricted to $t < t^\star$ is a natural assent, this gives us the desired result. □

Proof that $\hat{T}$ given by (5.61) *is invertible and $\hat{T}^{-1}$ is increasing.* We first show that $\hat{T}$ is invertible. Let $\hat{a}$ and $\hat{b}$ be given with $\hat{a}$ subcritical or critical and $T(\hat{a}) = \hat{b}$. Then for any $\hat{x}$, we can write $\hat{x}_i = (1 - \hat{s}_i)\hat{a}_i$. Doing this, we find that $\hat{T}(\hat{x}) = \hat{b}$ if and only if

$$\hat{s}_i = 1 - \mathrm{e}^{-\sum_{j=1}^{L} \alpha_{ij}\hat{a}_j\hat{s}_j}. \tag{5.64}$$

Since $\hat{a}$ is not supercritical, Lemma 2.2 tells us that all solutions to (5.64) are bounded by zero; and thus $\hat{T}(\hat{x}) = \hat{b}$ implies $\hat{x} \geq \hat{a}$. Similarly, if $\hat{x}$ is a subcritical or critical solution to $\hat{T}(\hat{x}) = \hat{b}$, we must also have $\hat{a} \geq \hat{x}$. Thus $\hat{T}$ is injective from the set of subcritical and critical densities, and is thus invertible over this domain.

To see that $\hat{T}^{-1}$ is increasing, let us define $\hat{F}$ by

$$\hat{F}_i(\hat{x}) = \mathrm{e}^{\hat{y}_i + \sum_{j=1}^{L} \alpha_{ij}\hat{x}_j}, \tag{5.65}$$

and note that $\hat{T}^{-1}(\hat{y}) = \hat{F}^{(\infty)}(0)$. (Indeed, $\hat{T}^{-1}(\hat{y})$ must be the minimal fixed point of $\hat{F}$. Since $\hat{F}$ is order preserving and bounded below by zero, this minimal fixed point must be $\hat{F}^{(\infty)}(0)$.) Since $\hat{F}(\cdot)$ is increasing as a function of $\hat{y}$, we see that $\hat{T}^{-1}$ is increasing. □

5.5 Main Proofs.

Proof of Theorem 2.6. Given any $\hat{\rho}$, let $\hat{\Gamma}^{(1)}, \hat{\Gamma}^{(2)}, \ldots$ be a sequence of assents with

$$\lim_{i\to\infty} H(\hat{\Gamma}^{(i)}, 0, |\hat{\rho}|) = \sup_{\hat{\Gamma}\in\mathscr{P}(\hat{\rho})} H(\hat{\Gamma}, 0, |\hat{\rho}|). \tag{5.66}$$

Then since $\hat{\Gamma}^{(i)}$ are bounded functions with bounded derivatives on a compact space, by standard equicontinuity arguments the sequence $\hat{\Gamma}^{(i)}$ must have a convergent subsequence, and we denote the limit of this subsequence to be $\hat{\Gamma}$. Then by continuity and properties inherited from the sequence, $\hat{\Gamma}$ is an assent passing through $\hat{\rho}$ which maximizes $H(\hat{\Gamma}, 0, |\hat{\rho}|)$. By Lemma 5.7 and the fact that $\hat{\Gamma}$ emanates from the origin, $\hat{\Gamma}$ must be a natural assent. Then integrating shows that the right hand side of (5.9) with $\hat{\eta} = 0$ is $\Psi(\hat{\rho})$ (although at this point we have not shown that the choice of $\hat{\rho}^\star$ is unique). In addition, Theorem 5.2 gives uniform convergence for $\hat{\rho}$ bounded above.

To prove the uniqueness of $\hat{\rho}^\star$ and the continuity of $\Psi(\hat{\rho})$, we consider the transformation $\hat{T}$ given by

$$\hat{T}_i(\hat{\eta}) = \frac{\hat{\rho}_i}{1 + \sum_{j=1}^{L} \alpha_{ij}(\hat{\rho}_j - \hat{\eta}_j)}. \tag{5.67}$$

We claim that $\hat{T}^{(k)}(\hat{0})$ converges to the minimal positive solution to

$$(\mathrm{I} - \mathrm{B}_{\hat{\eta}})(\hat{\rho} - \hat{\eta}) = 0. \tag{5.68}$$

Indeed, $\hat{T}$ is increasing (and $\hat{T}(\hat{0}) \geq 0$), and solutions to (5.68) are exactly the fixed points of $\hat{T}$. Since the eigenvalues of $(\mathrm{I} - \mathrm{B}_{\hat{\eta}})$ are strictly positive for $\hat{\eta}$ subcritical, the minimal solution to (5.68) cannot be subcritical if $\hat{\rho}$ is supercritical. Since we have already seen the existence of a critical solution to (5.68) if $\hat{\rho}$ is supercritical, this must be the minimal solution, and thus must be unique. Additionally, we see that for small ϵ the subcritical density $(1-\epsilon)\hat{\rho}^\star$ is below the fixed point of $\hat{T}$ for any sufficiently small perturbation of $\hat{\rho}$ (that is, we fix $\hat{\rho}^\star$ and slightly perturb the $\hat{\rho}$ used in $\hat{T}$). By this we see that $\hat{\rho}^\star$ is continuous as a function of $\hat{\rho}$, and therefore $\Psi(\hat{\rho})$ is a continuous function. □

Proof of Theorem 2.7. Let $\mathbb{B}_r$ be the event that all components in $\mathscr{G}(n, \hat{\rho})$ are of size r or bigger, and recall that $\mathbb{S}_r$ is the event that all components in $\mathscr{G}(n, \hat{\rho})$ are of size smaller than r. Then for all $\hat{\eta}$ with $0 \le \hat{\eta} \le \hat{\rho}$,

$$P\left(\hat{\theta}(\epsilon n, n, \hat{\rho}) = \frac{1}{n}\lfloor \hat{\eta} n \rfloor\right) = \mathrm{e}^{O(1)} \left[\prod_{i=1}^{L} \binom{\lfloor \hat{\rho}_i n \rfloor}{\lfloor \hat{\eta}_i n \rfloor}\right] P_{n,\hat{\eta}}(\mathbb{B}_{\epsilon n}) P_{n,\hat{\rho}-\hat{\eta}}(\mathbb{S}_{\epsilon n}) \prod_{i,j=1}^{L} \left(1 - \frac{\alpha_{ij}}{n}\right)^{\hat{\rho}_i n(\hat{\rho}_j n - \hat{\eta}_j n)}, \quad (5.69)$$

where the error term comes from rounding (and is uniformly bounded). Following the proof of Theorem 2.5, we note that for each $\hat{\eta} \in \mathscr{S}$ we have $\frac{1}{n}\lfloor \hat{\eta} n \rfloor \in \mathscr{S}$ for sufficiently large n. We will also use that for $m > |\hat{\rho}|/\epsilon$, we have $K_1 \subset \mathbb{B}_{\epsilon n} \subset K_m$. Now for a lower bound we fix a $\hat{\eta} \in \mathscr{S}$ with $\hat{\eta} > 0$, combining Theorem 3.1, Theorem 2.6, and direct calculation gives

$$\liminf_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log P\left(\hat{\theta}(\epsilon, n, \hat{\rho}) = \frac{1}{n}\lfloor \hat{\eta} n \rfloor\right) \ge S(\hat{\eta}, \hat{\rho}) + \Xi(\hat{\eta}) + \Psi(\hat{\rho} - \hat{\eta}) - \hat{\eta}^{\mathrm{T}} \mathrm{A}(\hat{\rho} - \hat{\eta}). \quad (5.70)$$

For an upper bound, since $\hat{\theta}(\epsilon n, n, \hat{\rho})$ can take only polynomially many values, we have

$$P\big(\hat{\theta}(\epsilon n, n, \hat{\rho}) \in \mathscr{S}\big) = \mathrm{e}^{o(n)} \sup_{\hat{\eta} \in \mathscr{S}} P\left(\hat{\theta}(\epsilon n, n, \hat{\rho}) = \frac{1}{n}\lfloor \hat{\eta} n \rfloor\right). \quad (5.71)$$

Note also that for any nonzero $\hat{\eta}$, we have

$$P\big(\hat{\theta}(\epsilon, n, \hat{\rho}) = \hat{\eta} \vee (1/n, \ldots, 1/n)\big) = \mathrm{e}^{O(1)} P\big(\hat{\theta}(\epsilon, n, \hat{\rho}) = \hat{\eta}\big), \quad (5.72)$$

which stems from the entropy considerations balancing out the cost of adding an edge (although we would be willing to accept any sub-exponential correction). Thus using Corollary 3.5 we may replace the $P_{n,\hat{\eta}}(\mathbb{B}_{\epsilon n})$ term in (5.69) with a uniform upper bound of $\exp\big(n\Xi(\hat{\eta}) + o(n)\big)$. Furthermore, using Theorem 2.6, we may replace the $P_{n,\hat{\rho}-\hat{\eta}}(\mathbb{S}_{\epsilon n})$ term with a uniform bound of $\exp\big(o(n) + n\Psi(\hat{\rho} - \hat{\eta})\big)$, in which the $o(n)$ term requires us to first take $n \to \infty$ followed by $\epsilon \to 0$.

Thus from (5.69) and (5.71) we have

$$\frac{1}{n} \log P\big(\hat{\theta}(\epsilon, n, \hat{\rho}) \in \mathscr{S}\big) \le o(1) + \sup_{\hat{\eta} \in \mathscr{S}} \big[S(\hat{\eta}, \hat{\rho}) + \Xi(\hat{\eta}) + \Psi(\hat{\rho} - \hat{\eta}) - \hat{\eta}^{\mathrm{T}} \mathrm{A}(\hat{\rho} - \hat{\eta})\big]. \quad (5.73)$$

Taking $n \to \infty$, followed by $\epsilon \to 0$, we get the desired upper bound, and prove the result. □

Proof of Theorem 2.1. By Theorem 2.7, we will have proved Theorem 2.1 if we can show that

$$S(\hat{\eta}, \hat{\rho}) + \Xi(\hat{\eta}) + \Psi(\hat{\rho} - \hat{\eta}) - \hat{\eta}^{\mathrm{T}} \mathrm{A}(\hat{\rho} - \hat{\eta}) \quad (5.74)$$

is maximized only when $\hat{\eta}_i / \hat{\rho}_i = \hat{\theta}^{\star}_i(\hat{\rho})$ for all i. The easiest way to see this is to use an idea mentioned in [2] following Theorem 2.1: Dropping the Ψ term and taking exponentials, we find

$$\mathrm{e}^{n[S(\hat{\eta},\hat{\rho}) + \Xi(\hat{\eta}) - \hat{\eta}^{\mathrm{T}} \mathrm{A}(\hat{\rho} - \hat{\eta})]} = \mathrm{e}^{o(n)} \prod_{i=1}^{L} \left[\binom{\lfloor \hat{\rho}_i n \rfloor}{\lfloor \hat{\eta}_i n \rfloor} \left(1 - \mathrm{e}^{-\sum_{j=1}^{L} \alpha_{ij} \hat{\eta}_j}\right)^{\hat{\eta}_i n} \left(\mathrm{e}^{-\sum_{j=1}^{L} \alpha_{ij} \hat{\eta}_j}\right)^{(\hat{\rho}_i - \hat{\eta}_i) n}\right]. \quad (5.75)$$

Well known results regarding binomials give us that the right hand side is exponentially small unless $\frac{\hat{\eta}_i}{\hat{\rho}_i} \approx 1 - \mathrm{e}^{-\sum_{j=1}^{L} \alpha_{ij}\hat{\eta}_j}$, which is to say that $\hat{\eta}_i/\hat{\rho}_i$ must satisfy (2.2) for all i. We can also get this result by maximizing $S(\hat{\eta}, \hat{\rho}) + \Xi(\hat{\eta}, \hat{\rho}) - \hat{\eta}^{\mathrm{T}}\mathrm{A}(\hat{\rho} - \hat{\eta})$ directly, and we note that the maximum is zero (as, indeed, it must be). If $\hat{\rho}$ is critical or subcritical, this finishes the proof, since $\hat{\Psi}(\hat{\rho}) = 0$ and Lemma 2.2 shows that $\hat{\theta}^{\star}(\hat{\rho}) = 0$. For $\hat{\rho}$ supercritical, the result follows by noting that $\Psi(\hat{\rho}) < 0$, while $\Psi(\hat{\rho} - \hat{\eta}) = 0$ for $\hat{\eta}_i = \hat{\rho}_i\hat{\theta}^{\star}_i(\hat{\rho})$ – that is to say, only the maximal solution to (2.2) maximizes (5.74). □

6. The Final Stage

We finish this note with a proof of Theorem 2.8 which requires one additional preliminary lemma:

Lemma 6.1. *Let $\mathcal{K}_\epsilon$ be the event that all components of $\mathscr{G}(n, \hat{\rho})$ are of size at least ϵn, and recall that K_1 is the event that $\mathscr{G}(n, \hat{\rho})$ is connected. Then for all $\hat{\rho}^{(1)}$ and $\hat{\rho}^{(2)}$ with $\hat{\rho}^{(2)} \geq \hat{\rho}^{(1)} > 0$ and sufficiently small $\epsilon > 0$,*

$$\limsup_{n\to\infty} \sup_{\hat{\rho}:\hat{\rho}^{(1)}\leq\hat{\rho}\leq\hat{\rho}^{(2)}} \frac{1}{n} \log P_{n,\hat{\rho}}\big(K_1^{\mathrm{c}} \,\big|\, \mathcal{K}_\epsilon\big) < 0. \tag{6.1}$$

Proof. Given vertex sets $\mathscr{X}, \mathscr{Y} \subset \mathscr{V}$, let $K_m\big|_{\mathscr{X}}$ represent the event that $\mathscr{G}(n, \hat{\rho})$ has m or fewer components when restricted to $\mathscr{X}$, and let $\mathscr{X} \not\leftrightarrow \mathscr{Y}$ represent the event that there are no edge between vertices of $\mathscr{X}$ and $\mathscr{Y}$ in $\mathscr{G}(n, \hat{\rho})$. Then with $m = |\hat{\rho}|/\epsilon$, we have

$$\mathcal{K}_\epsilon \cap K_1^{\mathrm{c}} \subset \bigcup_{\mathscr{A}\subset\mathscr{V}:|\mathscr{A}|\geq\epsilon n,|\mathscr{V}\setminus\mathscr{A}|\geq\epsilon n} \Big[(\mathscr{A} \not\leftrightarrow \mathscr{V} \setminus \mathscr{A}) \;\&\; K_m\big|_{\mathscr{A}} \;\&\; K_m\big|_{\mathscr{V}\setminus\mathscr{A}}\Big], \tag{6.2}$$

and so

$$P_{n,\hat{\rho}}\left(\mathcal{K}_\epsilon \cap K_1^{\mathrm{c}}\right) \leq \mathrm{e}^{o(n)} \sup_{\hat{\eta}:|\hat{\eta}|\geq\epsilon n,|\hat{\rho}-\hat{\eta}|\geq\epsilon n} \binom{\lfloor\hat{\rho}_i n\rfloor}{\lfloor\hat{\eta}_i n\rfloor} P_{n,\hat{\rho}}\big(K_m\big|_{\mathscr{A}}\big) P_{n,\hat{\rho}}\big(K_m\big|_{\mathscr{V}\setminus\mathscr{A}}\big) \mathrm{e}^{-\hat{\eta}^{\mathrm{T}}\mathrm{A}(\hat{\rho}-\hat{\eta})}, \tag{6.3}$$

where $\mathscr{A} \subset \mathscr{V}$ is a set with $\langle\mathscr{A}\rangle = \lfloor\hat{\eta}n\rfloor$. Since the "cost" of moving a vertex from $\mathscr{A}$ to $\mathscr{V} \setminus \mathscr{A}$ (or the reverse direction) is at most polynomial, we can add the requirement that both $\hat{\eta}n$ and $\lfloor\hat{\rho}n\rfloor - \lfloor\hat{\eta}n\rfloor$ are bounded below by $(1, \ldots, 1)$ without altering the error term. By applying Corollary 3.5 (and immediately dropping the added requirement on $\hat{\eta}$), we have

$$P_{n,\hat{\rho}}(\mathcal{K}_\epsilon \cap K_1^{\mathrm{c}}) \leq \mathrm{e}^{o(n)} \sup_{\hat{\eta}:|\hat{\eta}|\geq\epsilon,|\hat{\rho}-\hat{\eta}|\geq\epsilon} \prod_{i=1}^{L} \binom{\lfloor\hat{\rho}_i n\rfloor}{\lfloor\hat{\eta}_i n\rfloor} \mathrm{e}^{\Xi(\hat{\eta})+\Xi(\hat{\rho}-\hat{\eta})-(\hat{\rho}-\hat{\eta})^{\mathrm{T}}A\hat{\eta}}. \tag{6.4}$$

From this and Theorem 3.1, if we let

$$h(\hat{\gamma}) = \sum_{i=1}^{L} \hat{\gamma}_i \log \frac{1 - \mathrm{e}^{-\sum_{j=1}^{L}\alpha_{ij}\hat{\gamma}_j}}{\hat{\gamma}_i} + \frac{1}{2}\hat{\gamma}^{\mathrm{T}}A\hat{\gamma}, \tag{6.5}$$

we get (using uniform convergence)

$$\limsup_{n\to\infty} \sup_{\hat\rho:\hat\rho^{(1)}\le\hat\rho\le\hat\rho^{(2)}} \frac{1}{n}\log P_{n,\hat\rho}(K_1^{\mathrm{c}}|K_{\epsilon,2}) \le \sup_{\substack{\hat\rho:\hat\rho^{(1)}\le\hat\rho\le\hat\rho^{(2)}\\ \hat\eta:|\hat\eta|\ge\epsilon n,|\hat\rho-\hat\eta|\ge\epsilon n}} h(\hat\eta)+h(\hat\rho-\hat\eta)-h(\hat\rho). \tag{6.6}$$

By writing $\hat x_i = t\sum_{i=1}^L \alpha_{ij}\hat\eta_j$ and $\hat y_i = t\sum_{i=1}^L \alpha_{ij}\hat\rho_j$, we find

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{t}\big(h(t\hat\eta)+h(t(\hat\rho-\hat\eta))-h(t\hat\rho)\big)\right]\\ &\qquad= \frac{1}{2t}\sum_{i=1}^L\left[\hat\eta_i\hat x_i\frac{1+\mathrm{e}^{\hat x_i}}{1-\mathrm{e}^{-\hat x_i}}+(\hat\rho_i-\hat\eta_i)(\hat x_i-y_i)\frac{1+\mathrm{e}^{\hat x_i-\hat y_i}}{1-\mathrm{e}^{-\hat x_i-\hat y_i}}-\hat\rho_i\hat y_i\frac{1+\mathrm{e}^{\hat y_i}}{1-\mathrm{e}^{-\hat y_i}}\right].\end{aligned} \tag{6.7}$$

A brief calculation shows that $x\frac{1+\mathrm{e}^x}{1-\mathrm{e}^{-x}}$ is strictly increasing on $(0,\infty)$, and so (6.7) shows that $h(t\hat\eta)+h(t(\hat\rho-\hat\eta))-h(t\hat\rho)$ is decreasing in t for any applicable choice of $\hat\eta$. Then by taking $t\to 0$, we find

$$h(\hat\eta)+h(\hat\rho-\hat\eta)-h(\hat\rho) < g(\hat\eta)+g(\hat\rho-\hat\eta)-g(\hat\rho). \tag{6.8}$$

where $g(\hat\gamma)=\sum_{i=1}^L \hat\gamma_i \log\frac{\sum_{i=1}^L \alpha_{ij}\hat\gamma_j}{\hat\gamma_i}$. As a final step, we differentiate with respect to α_{kk}, which yields

$$\frac{\mathrm{d}}{\mathrm{d}\alpha_{kk}}\left(g(\hat\eta)+g(\hat\rho-\hat\eta)-g(\hat\rho)\right) = \frac{(\hat\eta_k\hat y_k-\hat\rho_k\hat x_k)^2}{\hat x_k(\hat y_k-\hat x_k)\hat y_k}. \tag{6.9}$$

Thus $g(\hat\eta)+g(\hat\rho-\hat\eta)-g(\hat\rho)$ is increasing as $\alpha_{kk}\to\infty$, but a brief examination shows

$$\lim_{\alpha_{11},\alpha_{22},\dots,\alpha_{LL}\to\infty} g(\hat\eta)+g(\hat\rho-\hat\eta)-g(\hat\rho)=0. \tag{6.10}$$

Combined with (6.8), this means that $h(\hat\eta)+h(\hat\rho-\hat\eta)-h(\hat\rho)<0$ for all applicable $\hat\rho$ and $\hat\eta$. Since h is continuous and we are working on a compact set, this proves the lemma. □

Proof of Theorem 2.8. From the independence of edges in $\mathscr{G}(n,\hat\rho)$, we have

$$P_{n,\hat\rho}\left(Q_\epsilon\,\middle|\,\hat\theta(\epsilon n,n,\hat\rho)=\frac{1}{n}\lfloor\hat\eta n\rfloor\right) = P_{n,\hat\eta}\left(K_1^{\mathrm{c}}\,\middle|\,\mathcal{K}_\epsilon\right). \tag{6.11}$$

The result then follows from Lemma 6.1. □

ACKNOWLEDGMENTS

This research was partially supported by the NSF grant DMS-0306167.

REFERENCES

[1] K. B. Athreya and A. N. Vidyashankar, *Branching Processes* http://citeseer.ist.psu.edu/211428.html (1999)

[2] M. Biskup, L. Chayes and S.A. Smith, *Large-deviations/thermodynamic approach to percolation on the complete graph*, Random Structures & Algorithms 31 (2007), no. 3, 354-370.