Opportunistic scheduling in large-scale wireless networks

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Opportunistic Scheduling in Large-Scale Wireless Networks

Mehdi Ansari Sadrabadi
Dept. of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: mahdi@mit.edu

Alireza Bayesteh
Dept. of Electrical Engineering
University of Waterloo
Waterloo, ON, N2L 3G1
Email: alireza@shannon2.uwaterloo.ca

Eytan Modiano
Dept. of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: modiano@mit.edu

Abstract—In this paper, we consider a distributed one-hop wireless network with \( n \) pairs of transmitters and receivers. It is assumed that each transmitter/receiver node is only connected to \( k \) receiver/transmitter nodes which are defined as neighboring nodes. The channel between the neighboring nodes is assumed to be Rayleigh fading. The objective is to find the maximum achievable sum-rate of the network in the asymptotic case of \( n, k \rightarrow \infty \). It is shown that the asymptotic throughput of the system scales as \( \frac{n \log k}{k} \). An opportunistic on-off scheduling is proposed and shown to be asymptotically throughput optimal.

I. INTRODUCTION

Throughput maximization in multi-user wireless networks has been addressed from different perspectives; resource allocation [1], scheduling [2], routing by using relay nodes [3], exploiting mobility of the nodes [4] and exploiting channel characteristics (e.g., power decay-versus-distance law [5]–[7], geometric pathloss and fading [8], [9]).

In recent years, power and spectrum allocation schemes have been extensively studied in cellular and multihop wireless networks [1], [10]–[12]. Much of these works rely on centralized and cooperative algorithms. Clearly, centralized resource allocation schemes provide a significant improvement in the network throughput over decentralized (distributed) approaches. However, they require extensive knowledge of the network configuration. In particular, when the number of nodes is large, deploying such centralized schemes may not be practically feasible. Decentralized resource allocation schemes have been extensively studied as alternatives to centralized schemes [13]–[16].

In decentralized schemes, the decisions concerning network parameters (e.g., rate and/or power) are made by the individual nodes based on their local information. Most of the works on the decentralized throughput maximization target the Signal-to-Interference-plus-Noise Ratio (SINR) parameter by using iterative algorithms [14], [15]. This leads to the use of game theoretic concepts [17] where the main challenge is the convergence issue. A more practical approach to avoid the extra amount of overhead in iterative algorithms is to rely on the channel gains as local decision parameters. References [18] and [19] consider a multihop ad hoc network model with random connections and devise routing schemes that maximize the network throughput. In [20], a wireless network with \( n \) pairs of transmitters/receivers is considered in which the transmission between each transmitter and its corresponding receiver takes place in one hop and the channel between each two nodes is modeled as Rayleigh fading. A distributed power allocation scheme called threshold-based on-off scheme (i.e., links with a direct channel gain above certain threshold transmit at full power and the rest remain silent) is introduced and shown to be order-optimal in the asymptotic case of \( n \rightarrow \infty \). Furthermore, the sum-rate throughput of the network is shown to scale as \( \Theta(\log n) \).

Distributed one-hop networks are extensively studied and have been considered in wireless standards. Local Area Networks (LAN) using unlicensed spectrum (e.g. Wi-Fi systems based on IEEE 802.11b standard [21]) are a typical example of such networks. In a LAN, there are several fixed nodes, called access points (APs). Mobile users can connect to the internet through APs. In the downlink phase, APs acts as transmitters and mobile terminals act as receivers. Each receiver observes the dominant part of the interference from the neighboring active transmitters in the network. In practice, one of the main parameters that influence the performance of the network is the typical range of the transmission, which is referred as coverage. Having more direct neighbors can increase the probability of a successful reception from at least one of the neighbors. However, it can also decrease the chance of successful decoding due to the overall larger interference.

In this paper, we consider a distributed one-hop wireless network with \( n \) transmitters and \( n \) receivers. We simplify the interference model by introducing the notion of neighbor, i.e., a pair of transmitter and receiver are referred as neighbors, if there exists a communication channel between them. In practice, the model is justified by observing the simple fact that signals from far transmitter nodes is negligible due to the attenuation. In this set-up, coverage is defined from the perspective of each wireless node as the number of neighbors that it can communicate with directly, which is assumed to be \( k \) for all nodes. Assuming Rayleigh fading for the connected links, i) the scaling of the maximum sum-rate throughput of the network is derived and is shown to scale as \( \frac{n \log k}{k} \) in

This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC), National Science Foundation under ITR Grant CCR-0325401 and ARO Muri grant number W911NF-08-1-0238.
the asymptotic region of \( n, k \to \infty \), and ii) an opportunistic on-off strategy is introduced which is shown to achieve the maximum sum-rate throughput. The difference between this work and the works of [20], [22] is that either they consider only the case of coverage equal to \( n \), i.e., all-connected nodes, while we consider a more general network set-up, or the transmitter/receiver pairs are assumed to be dedicated, i.e., each transmitter aims to communicate with only one specific receiver, while in this work, each transmitter can send data to any receiver which is connected to it. This makes the problem more challenging as the scheduling is involved. However, it is demonstrated that in the case of \( k = n \), scheduling provides no gain in the asymptotic throughput.

The rest of the paper is organized as follows: In section II, the network model and assumptions are described. Section III is devoted to the asymptotic analysis of sum-rate throughput, and finally, section IV concludes the paper.

II. NETWORK MODEL

We consider a wireless communication network with \( n \) pairs of transmitters and receivers. We assume that each transmitter is a neighbor to \( k \) receivers and each receiver is a neighbor to \( k \) transmitters. It is assumed that both \( n \) and \( k \) tend to infinity. Each transmitter can send data to any of its corresponding neighboring receivers and the transmission takes place in one hop. Let us define \( \Phi_i \) as the set of the receivers which are neighbors to \( i \)th transmitter and \( \Psi_i \) as the set transmitters neighbors to \( j \)th receiver.

The channel between \( i \)th transmitter and \( j \)th receiver is characterized by the channel gain \( h_{ij} \). It is assumed that the channel gains are independent and identically distributed (i.i.d.) random variables with cumulative distribution function (CDF) \( F(.) \). We consider an additive white Gaussian noise (AWGN) with unit variance at the receivers.

We assume that receivers are equipped with single user detectors, i.e. each receiver decodes only the signal from the intended transmitter and consider the interference from other transmitters as noise. It is also assumed that the transmitters utilize on-off power scheme, i.e., they either transmit with full power, or remain silent. The power constraint of all transmitters is assumed to be equal to \( \rho \). Assuming Gaussian signal transmission from all the transmitters, the interference distribution is also Gaussian. Therefore, the maximum rate of the transmitter \( i \) to the receiver \( j \in \Phi_i \) is equal to

\[
\tau_{ij} = \log \left( 1 + \frac{h_{ij}}{\rho + I_{ij}} \right)
\]  

where \( I_{ij} = \sum_{l \in \mathcal{I}_j, l \neq j} h_{il} \), in which \( \mathcal{I}_j \) denotes the set of active transmitters in \( \Psi_j \), and \( \rho = \frac{1}{\theta} \) In this paper, the performance measure is the average system throughput which is defined as the average sum-rate of all links.

III. THROUGHPUT ANALYSIS

A. Lower-bound on the average throughput

In this part, we derive a lower bound on the average throughput of the system. Let \( \Phi_i^{(\tau)} \in \Phi_i \) denote the set of the receivers which their corresponding channel gains to the transmitter \( i \) are above \( \tau \). Similarly, let \( \Psi_i^{(\tau)} \in \Psi_i \) denote the set of transmitters which their corresponding channel gains to the receiver \( i \) are above \( \tau \). We give the following opportunistic scheme, which is called opportunistic on-off scheme:

At the receivers, the channel gains are estimated. A single-bit data is fed back to the transmitters acknowledging permission of transmission. If the channel gain of only one of the links is above \( \tau \), the receiver acknowledges the corresponding transmitter for the transmission permission. If the transmitter receives acknowledgment from only one receiver, it transmits to the corresponding receiver with full power. Otherwise, it remains silent.

This opportunistic scheme constructs a one-to-one map from the set of transmitters to the set of receivers. In fact, the transmitter \( i \) and the receiver \( j \) communicate iff \( \Phi_i^{(\tau)} = \{ j \} \) and \( \Psi_i^{(\tau)} = \{ i \} \). Let us call such an event \( \mathcal{L} \). Assume that \( \pi_i \)'s are the indices of the active transmitters and \( \theta_i \)'s are the corresponding receivers. The average throughput of the opportunistic scheduling can be written as follows:

\[
T = \mathbb{E} \left\{ \sum_{i=1}^{n^*} \log \left( 1 + \frac{h_{\pi_i\theta_i}}{\nu + I_{\pi_i\theta_i}} \right) \right\},
\]  

where \( n^* \) is the number of active transmitters and the expectation is taken over \( n^*, h_{\pi_i\theta_i}, \) and \( I_{\pi_i\theta_i} \). The following theorem gives a lower-bound on the average throughput based on the proposed opportunistic on-off scheme:

Theorem 1 The asymptotic average throughput of the proposed opportunistic scheme in a Rayleigh fading environment can be lower-bounded as follows:

\[
T \geq \frac{n \log k}{k},
\]  

as \( n, k \to \infty \).

Proof: In order to derive the average throughput, we first derive the probability of activation for any of the transmitters. The probability of the activation event can be lower bounded by the event that the channel gain between the transmitter to exactly one receiver in its neighborhood is greater than \( \tau \) and the channel gain of this receiver to the rest of the transmitters in its neighborhood is less than \( \tau \). Hence, due to the independence of the channels, the probability that a transmitter becomes active can be bounded as follows

\[
\Pr(\mathcal{L}) \geq k(1 - F(\tau)) F(\tau)^{2k-1},
\]  

where \( F(.) \) denotes the CDF of the channel gain which is equal to \( F(\tau) = 1 - e^{-\tau} \) in the case of Rayleigh fading. Noting that for the active links the corresponding channel gains are above \( \tau \), the average throughput in (2) can be lower-bounded as follows:

\[
T \geq \mathbb{E} \left\{ \sum_{i=1}^{n^*} \log \left( 1 + \frac{\tau}{\nu + I_{\pi_i\theta_i}} \right) \right\} \\
\geq \mathbb{E} \left\{ n^* \log \left( 1 + \frac{\tau}{\nu + \frac{1}{n^*} \sum_{i=1}^{n^*} I_{\pi_i\theta_i}} \right) \right\},
\]
where the second inequality results from the convexity of \( \log(1 + \frac{x}{x + b}) \) with respect to \( x \) and applying Jensen’s inequality. By selecting \( \tau = \log(2k) \), we have \( F(\tau) = 1 - \frac{1}{2^\tau} \). Using (4), we have \( \Pr(\mathcal{L}) \geq \frac{1}{2^\tau(n - \frac{1}{2^\tau})^{2k-1}} \geq \frac{1}{k^2} \) for all \( k \). This implies that \( \mathbb{E}\{n^*\} \geq \frac{n}{2^\tau} \) and as a result, \( n^* \to \infty \) with probability one. Using Tchebychev’s inequality, we have
\[
\Pr\left( \frac{1}{n^*} \sum_{i=1}^{n^*} I_{\pi, \theta_{i}} \neq \mathbb{E}\{I_{\pi, \theta_{i}}\} \bigg| \geq \beta \right) \leq \frac{\sigma^2}{\beta^2},
\]
where \( \sigma^2 \) denotes the variance of the term \( \frac{1}{n^*} \sum_{i=1}^{n^*} I_{\pi, \theta_{i}} \).

Since the terms \( I_{\pi, \theta_{i}} \) are independent of each other for different \( i \) and the variance of each term \( I_{\pi, \theta_{i}} \) can be shown to be \( k\Pr(\mathcal{L}) \), it follows that \( \sigma^2 = \frac{k\Pr(\mathcal{L})}{n} \). Noting that \( \mathbb{E}\{I_{\pi, \theta_{i}}\} = k\Pr(\mathcal{L}) \), and selecting \( \beta = \frac{k\Pr(\mathcal{L})}{n} \), we have
\[
\Pr\left( \frac{1}{n^*} \sum_{i=1}^{n^*} I_{\pi, \theta_{i}} > k\Pr(\mathcal{L})(1 + \epsilon) \right) \leq \frac{2e}{kn^*\epsilon^2},
\]
which approaches zero for some \( \epsilon > 0 \). This implies that
\[
\frac{1}{n^*} \sum_{i=1}^{n^*} I_{\pi, \theta_{i}} \leq k\Pr(\mathcal{L})(1 + \epsilon),
\]
with probability one. Substituting in (5), we have
\[
T \geq \mathbb{E}\{n^*\} \log\left( 1 + \frac{\log(2k)}{k\Pr(\mathcal{L})(1 + \epsilon)} \right)
= \frac{n k \Pr(\mathcal{L}) \log(1 + \frac{\log(2k)}{k\Pr(\mathcal{L})(1 + \epsilon)})}{k(1 + \epsilon)}
\approx \frac{n \log k}{k(1 + \epsilon)},
\]
where the last line follows from the fact that \( \log(1 + x) \approx x \), for \( x = o(1) \). Selecting small enough \( \epsilon \), the theorem is proved.

\[\blacksquare\]

B. Upper Bound on the average throughput

In the following, we derive an upper-bound by removing the constraint that each receiver should be served by at most one neighboring transmitter. In other words, we assume that each receiver can be served by more than one neighboring transmitter without imposing any interference on each other, which gives an upper bound on the performance of the system. By removing this constraint, we can assume that the transmitters operate independently which makes the analysis tractable.

It is known that in a broadcast block fading channel, the maximum sum-rate throughput is achieved by transmitting to the user with the highest channel gain at a time. Here, from the view point of each active transmitter, we have a single-antenna broadcast channel in which the statistics of the noise (capturing also the imposed interference from the other transmitters) is the same for all neighboring receivers. Therefore, to maximize the sum-rate for this channel, the transmitter should send data to the receiver with the highest direct channel gain. Note that however, due to the imposed interference from the active transmitters to their neighboring receivers who are served by other transmitters, activation of all transmitters may not be optimum. Since each transmitter is only aware of its local channel information, i.e., the channel gains to its corresponding neighboring receivers, the decision of being active or not is solely performed based on these information. In general, the transmission scheme can be expressed based on a function \( f(\cdot) \) such that for the channel realizations for which \( f(h_{ij}) \geq 0 \), the \( i^{th} \) transmitter is active and otherwise it remains silent, where \( h_{ij} = \{h_{ij}\}_{j \in \Phi} \). Note that because of the network symmetry, \( f(\cdot) \) is the same for all transmitters.

Based on the set-up introduced here, the following theorem gives an upper-bound on the system throughput.

\[\textbf{Theorem 2} \quad \text{The asymptotic average throughput of the system in a Rayleigh fading environment is upper bounded as follows:}\]
\[
T \leq \frac{(1 + \epsilon)n \log k}{k},
\]
for some \( \epsilon > 0 \).

\[\text{Proof:} \quad \text{Let us denote the probability that any transmitter} \ i \ 	ext{becomes active by} \ p \ \text{and the set of active transmitters by} \ S. \ \text{Since the activation of transmitters is performed independently by our relaxing assumption, it follows that}\ |S| \ \text{is a Binomial random variable with parameters}\ (n, p). \ \text{In the sequel, we consider two cases for}\ p: \]

\[\text{•} \quad p = \omega\left(\frac{1}{T}\right): \quad \text{This case is referred to the strong interference scenario, as} \ \mathbb{E}\{I_{\pi, \theta_{i}}\} = kp = \omega(1). \ \text{In this case, one can easily show that} \ |S| \approx np \ \text{and the number of interfering transmitters for each active link} \ \pi_{ij}, \ \text{denoted by} \ |I_{\pi_{ij}}| \approx kp, \ \text{with probability one. More precisely, using the Gaussian approximation for the Binomial distribution one can show that} \ \Pr\{np(1 - \epsilon) \leq |S| \leq np(1 + \epsilon)\} \leq 1 - e^{-np\epsilon/2} \ \text{and} \ \Pr\{kp(1 - \epsilon) \leq |I_{\pi_{ij}}| \leq kp(1 + \epsilon)\} \leq 1 - e^{-kp\epsilon/2} \ \text{for some} \ \epsilon > 0, \ \text{such that} \ kp\epsilon = \omega(1). \ \text{As mentioned earlier, the upper bound is achieved if the link activation strategy leads to a one-to-one transmission map from the transmitters to the receivers. In other words, transmitter} \ i \ \text{sends to the corresponding receiver} \ \theta_{ij}, \ \text{where} \ h_{ij} = \max_{j \in \Phi} h_{ij} \ \text{and} \ \theta_{ij} \ \text{is a one-to-one map. Considering this assumption, and defining} \]
\[
\Upsilon_{\theta_{ij}} \equiv \frac{\max_{j \in \Phi} h_{ij}}{1/p + \sum_{l \in \mathcal{U}_{ij}, l \neq i} h_{il}},
\]
where \( \mathcal{U}_{ij} \) denotes the set of active transmitters in the neighborhood of the \( j^{th} \) receiver, we can bound the asymptotic average throughput of the system as follows:
\[
T \leq \mathbb{E}\left\{ \sum_{i=1}^{|S|} \log(1 + \Upsilon_{\theta_{ij}}) \right\}
\approx \frac{np\mathbb{E}\{\log(1 + \Upsilon_{\theta_{ij}})\}}{(a)}
\leq \frac{np\log(1 + \mathbb{E}\{\Upsilon_{\theta_{ij}}\})}{(b)},
\]
where \((a)\) follows from the fact that \(|S| \approx np, \) with
probability one\(^1\) and (b) follows from the concavity of \(\log(.)\) function and Jensen’s inequality. The following lemma, gives an upper-bound on \(\mathbb{E}\{\Upsilon_{i\theta_i}\}\).

**Lemma 1** There exists some \(\epsilon > 0\) for which
\[
\Pr \left( \Upsilon_{i\theta_i} > \frac{(1 + \epsilon) \log k}{kp} \right) \rightarrow 0, \quad (12)
\]
with probability one. This also implies that \(\mathbb{E}\{\Upsilon_{i\theta_i}\} \leq \frac{(1 + \epsilon) \log k}{kp}\).

**Proof:** See Appendix A.

Using the result of Lemma 1, the upper-bound on the throughput given in (11) can be written as
\[
T \leq n\log \left( 1 + \frac{(1 + \epsilon) \log k}{kp} \right) \\
\leq \frac{(1 + \epsilon) n \log k}{k}, \quad (13)
\]
where the second line comes from the fact that \(\log(1 + x) \leq x\).

* • \(p = O\left(\frac{1}{k}\right)\): In this case, \(p\) can be upper-bounded as \(c/k\) for some constant \(c\). An upper-bound on the average throughput can be given as
\[
T \overset{(a)}{=} \mathbb{E} \left\{ \sum_{i=1}^{S} \log \left( 1 + \frac{\max_{j \in \Phi} h_{ij}}{v} \right) \right\} \\
\overset{(b)}{=} \mathbb{E} \{ |S| \} \log \left( 1 + \frac{a \log k}{v} \right) \\
\leq \frac{cn \log \log k}{k}, \quad (14)
\]
for some constant \(c\), where (a) results from removing the interference term in the denominator of \(\Upsilon_{i\theta_i}\), and (b) follows from the fact that \(\max_{j \in \Phi} h_{ij} < a \log k\) with probability one for some \(a > 1\). It can be observed that this upper-bound is less than the one given in the case \(p = \omega(1/k)\). This completes the proof of Theorem 2.

**IV. DISCUSSION**

Combining the results of Theorems 1 and 2, it follows that the average sum-rate capacity of the network scales as \(\frac{n \log k}{k}\), which is achieved by opportunistic on-off scheme. Defining the **connectivity factor** of the network as \(\kappa \triangleq \frac{k}{n}\), it follows that \(T \sim \frac{\log k}{\kappa}\). This implies that the network throughput is inversely proportional to the **connectivity factor**. The factor \(\log k\) can be interpreted as the **scheduling diversity gain**, since it captures the effect of selecting the best transmission link for each transmitter.

The more interesting case is the case of \(k = n\). The existing results in the literature [20], [23] indicate that the average network throughput scale as \(\log n\) for the case of dedicated links. Our results also show the same scaling in the case of opportunistic links, i.e., non-dedicated links. This implies that in the case of \(k = n\), scheduling provides no gain in the asymptotic network throughput. However, it should be noted that in the case of dedicated network, only a few portion of the transmitters must be active in order to achieve the maximum throughput, while in the proposed opportunistic scheme, it is possible to achieve the maximum throughput with the activation probability of \(\frac{1}{2\epsilon}\).

**APPENDIX A: PROOF OF LEMMA 1**

**Proof:** The CDF of the maximum channel gain among \(k\) channels is \(F_{\max}(x) = F(x)^k\). In the case of Rayleigh fading channel, we have \(F_{\max}(x) = 1 - (1 - e^{-x})^k\). Noting that the number of active transmitters in the neighborhood of \(\theta_j\) can be well approximated by \(kp^2\), the interference term \(I_{i\theta_i} = \sum_{j \in \mathcal{I}_i \cap \mathcal{I}_i} h_{ij}\) in the denominator of \(\Upsilon_{i\theta_i}\) in (10) has \(\chi^2(k + 2)\) distribution, where \(k^* \triangleq kp\). Hence, \(\eta \triangleq \Pr(\Upsilon_{i\theta_i} > x)\) can be written as
\[
\eta = \int_0^\infty \Pr(\Upsilon_{i\theta_i} > x | I_{i\theta_i} = y) \Pr(I_{i\theta_i} = y) dy \\
= \int_0^\infty (1 - (1 - e^{-x(v+y)})^k) \cdot \frac{y^{k^* - 2} e^{-y}}{(k^* - 2)!} dy \\
\overset{(a)}{=} \int_0^\infty \min\{1, ke^{-x(v+y)}\} \cdot \frac{y^{k^* - 2} e^{-y}}{(k^* - 2)!} dy, \quad (15)
\]
where (a) results from the fact that \(1 - (1 - e^{-z})^k \leq \min\{1, ke^{-z}\}\) for \(z > 0\). The integral in the last line can be written as the summation of two integrals as follows:
\[
\text{RH}(15) = \int_0^{\frac{-\log k}{v}} \frac{y^{k^* - 2} e^{-y}}{(k^* - 2)!} dy + \int_{\frac{-\log k}{v}}^\infty ke^{-x(v+y)} \cdot \frac{y^{k^* - 2} e^{-y}}{(k^* - 2)!} dy \\
\leq \int_0^{\frac{-\log k}{v}} \frac{y^{k^* - 2} e^{-y}}{(k^* - 2)!} dy + \frac{ke^{-xv}}{(x+1)^{k^* - 1}} \int_0^\infty \frac{y^{k^* - 2} e^{-y}}{(k^* - 2)!} dy, \quad (16)
\]
where \(\gamma(a, z) \triangleq \int_0^z e^{-t} t^{a-1} dt\) is the incomplete Gamma function which can be expanded as follows
\[
\gamma(a, z) = a^{-1} z^{a-z} \left( 1 + \frac{z}{a+1} + \frac{z^2}{(a+1)(a+2)} + \frac{z^3}{(a+1)(a+2)(a+3)} + \cdots \right). \quad (17)
\]
By choosing \(x = \frac{a \log k}{kp}\), where \(a > 1\) is a real number and substituting it in (16), the first term in the right hand side of

---

\(^1\)To be precise, however, we should also show that the contribution of the realizations in which \(|S| \not\in [np(1-\epsilon), np(1+\epsilon)]\) in the average throughput is negligible. This fact can be shown easily, however, due to the space limitations we do not bring the proof here.

\(^2\)For simplicity, we assume that \(kp\) is an integer number.
(16) can be simplified as follows:

\[
\frac{\gamma(k^* - 1, \frac{\log k}{x})}{(k^* - 2)!} = \left(\frac{(k^*/\alpha)^{k^*-1}e^{-k^*/\alpha}}{(k^*-1)!}\right) \left(1 + \frac{1}{\alpha} + \frac{k^*}{\alpha^2(k^*+1)} + \cdots\right) \leq \frac{\alpha^2 k^* e^{-k^*/\alpha}}{\alpha^2 e^{-k^*(\log(\alpha)+1)}(\alpha-1)} \approx \frac{1}{\sqrt{2\pi k^*}(\alpha-1)}, \quad (18)
\]

where \((a)\) results from applying Stirling’s approximation, i.e. \(k! \approx \sqrt{2\pi k}k^ke^{-k}\). Also, defining \(t \triangleq \left(\frac{k^*}{\alpha} - x\right)(1+x)\), the second term in the right hand side of (16), denoted by \(S_2\), can be written as

\[
S_2 = \frac{k e^{-x^m}}{(x+1)^k-1} \int_0^\infty \frac{y^{k^*-2}e^{-y}}{(k^*-2)!} dy = \frac{k e^{-x^m}}{(x+1)^k} \sum_{m=1}^{k^*-2} \frac{t^m}{m!} \\
\leq \frac{k e^{-x^m}}{(x+1)^k} \sum_{m=1}^{k^*-2} \frac{t^m}{2}\approx \frac{k e^{-x^m}}{(x+1)^k} \left(1 + x\right)^{k^*-2} e^{-\frac{t}{\alpha}(1+x)} \\
\leq \frac{k e^{-x^m}}{(x+1)^k} \left(1 + x\right)^{k^*-2} e^{-\frac{t}{\alpha}(1+x)} \\
\leq \frac{e^\nu k e^{-x^m}}{(x+1)^k} \left(1 + x\right)^{k^*-2} e^{-\frac{t}{\alpha}(1+x)} \\
\leq \frac{c \sqrt{k^* e^{-k^*(\log(\alpha)+1)}(\alpha-1)}}{\sqrt{k^* \epsilon}}, \quad (19)
\]

for some constant \(c\). In the above equation, \((a)\) follows from the fact that as \(t > k^3\), we have \(\frac{t^m}{m!} \leq \frac{k^2}{(k^*-2)!}\), and \((b)\) follows from applying Stirling’s approximation in a similar manner as in \((18)\).

Setting \(\alpha = 1 + \epsilon\) and using the approximations \(\log(1+\epsilon) \approx \epsilon - \frac{\epsilon^2}{2}\) and \((1 + \epsilon)^{-1} \approx 1 - \epsilon + \frac{\epsilon^2}{2}\) for small enough \(\epsilon\), the right hand sides of \((18)\) and \((19)\) can be written as

\[
\text{RH}(18) \approx c_1 \frac{e^{-k^*\epsilon^2/2}}{\sqrt{k^* \epsilon}}, \quad (20)
\]

and

\[
\text{RH}(19) \approx c_2 \sqrt{k^* e^{-k^*\epsilon^2/2}}, \quad (21)
\]

respectively. Selecting \(\epsilon = 2\sqrt{\frac{\log(k^*)}{k}}\), it can be observed that both term approach to zero polynomially, as \(k^*\) increases, which gives us the desired result. Furthermore, we have

\[
\mathbb{E}\{Y_{i\theta_i}\} \leq x + \int_x^\infty tf_{X_{i\theta_i}}(t) dt = x + x \Pr(Y_{i\theta_i} > x) + \int_x^\infty \Pr(Y_{i\theta_i} > t) dt. \quad (22)
\]

\footnote{Note that one can select \(\alpha\) such that \(t > k^*\).}

Using \((18)\) and \((19)\), it follows that \(x \Pr(Y_{i\theta_i} > x) \rightarrow 0\) and \(\int_x^\infty \Pr(Y_{i\theta_i} > t) dt \rightarrow 0\), for \(x = \frac{(1+\epsilon)\log k}{k^*}\) and some \(\epsilon > 0\). This completes the proof. \(\blacksquare\)

REFERENCES


