Cost of Path Loss and Local Cooperation in Capacity Scaling of Extended Wireless Networks

Jinfeng Du
Nokia Bell Labs
Holmdel, 07733, NJ
Email: jinfeng.du@bell-labs.com

Muriel Médard
MIT
Cambridge, 02139, MA
Email: medard@mit.edu

Shlomo Shamai (Shitz)
Technion
Technion City, Haifa 32000, Israel
Email: sshlomo@ee.technion.ac.il

***Abstract*—Given a large wireless network consisting of randomly deployed nodes, where each of the nodes wants to transmit to a random destination node within the network at some equal rate, how fast can the sum rate grow as the number of nodes scales up at fixed density? This question is important because it captures the bottleneck of message exchanging among randomly deployed Internet-of-Things (IoT) devices, and it models nicely the wireless backhaul communication among access points or within airborne communication systems. Previous work has shown that, given an extended network with fixed density, multi-hop routing based approach provides sum rate that scales at most as the square root of network size, where as hierarchical cooperation protocols have the potential to support linear scaling. With limited power, the SNR decreases at least inverse proportional to the network size, and therefore the benefit of hierarchical cooperation will be curbed by the combined effects of path loss and local communication cost. We show in this paper how the path loss and local cooperation cost reshape the capacity scaling law results.**

***Index Terms*—Scalability, distributed cooperation, wireless networks, virtual MIMO, massive IoT**

I. Introduction

Internet-of-Things (IoT) devices that are capable of transmitting and receiving information through wireless links have the potential to transform the way we live and work. In the past few years, the proliferation of IoT devices and applications in wearable electronics, healthcare, environmental monitoring, and industrial environments has called the attention to the underlining communication infrastructure to support the communication demand of the massive IoT devices [1]–[3]. Given an area with many randomly deployed IoT devices where information exchange are carried out using device-to-device communication (i.e., without dedicated communication infrastructure), how will the information exchange rate scale as the number of IoT devices increases? This question is especially relevant for the design and deployment of wide area sensor networks for environmental/industrial monitoring, sensing, and localization services [4], [5]. On the other hand, the wireless communication infrastructure itself is facing big challenges to provide high-speed wide-area coverage in a cost effective manner. Some high potential solutions, such as deploying wireless access points or airborne communication networks, would require massive amount of data/channel/control information exchange among the access/anchor points, which itself is a big challenge [6]–[8]. How does the aggregate throughput of such network increase as the coverage area increases?

The work of M. Médard has been supported by NSF grant CCF-1409228. The work of S. Shamai has been supported by the European Union's Horizon 2020 Research And Innovation Programme, grant agreement no. 694630.

To provide tractable performance analysis and closed-form scalability evaluation for the aforementioned two problems, we turn to a powerful model that was established in the last decade for evaluating wireless Ad-hoc and mesh networks [9]. In this model a large wireless network contains N randomly deployed nodes, each of which wants to transmit at some equal rate R to a random destination node within the network. We revisit this powerful model because it captures the message exchanging network of randomly deployed IoT devices, and it models nicely the wireless backhaul communication among access points. We are interested in the scalability of the sum rate as the number of nodes N scales up but the density of nodes is fixed. i.e., the *extended networks*. In the last decade the scalability of the sum rate has been extensively investigated. Xie and Kumar [10], [11] provided an upper bound on the capacity scaling and showed that the multiple-hop point-to-point transmission strategy proposed by Gupta and Kumar [9] is essentially order-optimal, approaching $O(\sqrt{N})$, when the path loss attenuation factor is high (path loss attenuation factor $\alpha>4$). A hierarchical cooperation scheme proposed by Özgür et al. [12] employs local communication among neighboring nodes to create virtual multiple-input multiple-output (MIMO) connections between source-destination pairs. It was shown in [12] that, with perfect channel state information at all nodes, the maximal sum rate can scale as $O(N^{2-\alpha/2})$ for $2\leq\alpha<3$. This hierarchical cooperation scheme was later refined in [13]–[15] to maximize the achievable rate. The tradeoff between capacity and delay scaling was investigated [16] and a comprehensive survey of research on the capacity and delay issues can be found in [17] .

The hierarchical cooperation protocol rests on a layered structure, where the transmission on each layer is divided into three stages: two local cooperation stages among the transmit cluster and the receive cluster, respectively, and one virtual MIMO transmission between the two clusters. The local cooperation tasks of an upper layer in the hierarchy are treated as the communication problem to be solved by the lower layer, which itself also consists of three stages partitioned in the same fashion. Such recursion is applied repeatedly until the communication task for the lower layer becomes trivial or can be solved by local transmission. The cost of local cooperation, which is inevitable to establish the virtual MIMO transmission, is shown in [18] to grow exponentially with the number of layers in the cooperation hierarchy, and an upper bound for the sum rate scaling shows that, linear scaling of sum rate is out of reach even when the number of nodes within the network is extremely large. For example, scaling of $O(N^{0.7})$ requires a network of $N{=}10^9$ nodes and scaling of $O(N^{0.8})$ requires $N{=}10^{24}$.

In *extended networks* the distance between source and destination nodes grows as $O(\sqrt{N})$, the received signal to noise ratio (SNR) becomes inevitably low and the channel knowledge can no longer be assumed available for free, which hinders the feasibility of quantization based joint detection in the distributed receivers [19]. Therefore the number of nodes N cannot be too large, and the benefit of hierarchical cooperation will be curbed by the combined effects of path loss and local cooperation cost. We show in this paper how the combined effect of path loss and local cooperation cost reshapes the capacity scaling law results.

II. System Model

We inherit the grid topology from [12] such that there are N nodes evenly distributed in a square[1]. Every node wants to transmit the same amount data to a destination node randomly chosen from the rest of the $N-1$ nodes, and the aggregate throughput of all the N source-destination pairs is referred as the sum rate. As illustrated in Fig. 1, for a virtual MIMO transmission between two clusters, each of M nodes, in Stage I the source node s_i distributes Q_s bits to all its $M-1$ neighboring nodes via local communication. In Stage II all the M nodes in the transmit cluster transmit to all the M nodes in the destination cluster using the same channel (i.e., in the same time and frequency resource block). In Stage III all nodes in the destination cluster first independently quantize their observation into Q_d bits and then forward them to the destination node d_j via local communication. The virtual MIMO transmission is assumed to be capacity achieving in the sense that we can recover the transmitted message successfully if decoding is done based on the original observations (without quantization) of all the receiving nodes in the cluster.

The SNR for the transmission between node i in the transmit cluster and node j in the receive cluster can be written as

$$\gamma_{i,j} \triangleq P_i(d_{ij}) d_{ij}^{-\alpha} E[|h_{i,j}|^2] \approx \gamma(d_{st}),\ \forall i,j, \tag{1}$$

where d_{st} is the average distance between the two clusters, d_{ij} is the actual distance between node i and node j, $\alpha{\geq}2$ is the path loss attenuation factor, and $P_i(d_{ij})$ accounts for all other factors such as the transmit power of node i, the noise power at node j, antenna gains, and other loss. We assume that all channels $h_{i,j}$ are i.i.d. with $E[|h_{i,j}|^2]{=}1$. The approximation in (1) is due to minor variation of SNR, after proper power

[1]The scaling law gap between random networks and regular-grid networks is closed by [20] using percolation theory.

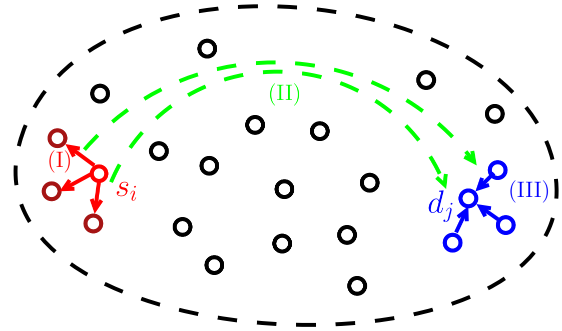

Fig. 1. Illustration of the hierarchical virtual MIMO transmission proposed by [12] between a source node s_i and a destination node d_j, where in Stage I the local cooperation among the transmit cluster creates a virtual multiple-antenna transmitter, Stage II is dedicated for the virtual MIMO transmission between the transmit cluster and the receive cluster, and in Stage III all the nodes in the receive cluster conveys their observation to the destination node for joint processing. Such virtual MIMO technique is then applied recursively to solve the local communication problems by formulating new virtual MIMO connections at a smaller scale.

control, across different transmit-receive pairs from the two clusters that are d_{st} apart, as we are aiming at the capacity scaling behavior rather than the exact capacity.

For the $M{\times}M$ virtual MIMO transmission with average SNR γ, its cut-set bound $\mathcal{C}(M,M,\gamma)$, established in [18], is maximized by choosing

$$Q_s \gtrsim R^*(M\gamma),\ Q_d \gtrsim R^*(M\gamma), \tag{2}$$

where

$$R^*(M\gamma) \triangleq \frac{1}{M}\mathcal{C}(M,M,\gamma) {\simeq} \log(1+M\gamma). \tag{3}$$

Assuming at the initial stage there exists a scheme that can support N_0 source-destination pairs at an aggregate throughput $C_0{=}\frac{1}{S_0}N_0^{b_0}$ within a cluster of N_0 nodes, where $S_0{>}0$ and $b_0{\geq}0$ are two parameters depending on the scheme. The aggregate throughput (i.e., the sum rate) at layer ℓ (with cluster size M_ℓ and network size N_ℓ) can be written as [18]

$$C_\ell = \frac{M_\ell R^*(M_\ell \gamma_\ell)}{1{+}S_{\ell-1}(Q_s{+}Q_d)} {=} \frac{M_\ell}{S_\ell}, \tag{4}$$

where γ_ℓ is the inter-cluster SNR at layer ℓ, and

$$N_\ell = N_0^{(\ell+1)-\ell b_0},\ M_\ell = N_{\ell-1} = N_0^{\ell-(\ell-1)b_0}. \tag{5}$$

III. Capacity Scaling for Extended Networks

As network size N_ℓ scales up, the inter-cluster SNR γ_ℓ scales down proportional to $N_\ell^{\alpha/2}$, i.e.,

$$\gamma_\ell = \gamma_0 N_\ell^{-\alpha/2}, \tag{6}$$

where $\alpha{\geq}2$ is the path loss attention factor for most wireless channels, and γ_0 is some constant chosen to meet the power constraint. From (3) we have

$$\begin{aligned} R_\ell^* &\triangleq R^*(M_\ell\gamma_\ell) \simeq \log(1+M_\ell\gamma_\ell) \\ &\overset{(a)}{=} \log(1+M_\ell\gamma_0 N_\ell^{-\alpha/2}) \\ &\overset{(b)}{=} \log(1+\gamma_0 N_0^{b_0-\alpha/2} N_0^{\ell(1-b_0)(1-\alpha/2)}) \\ &\overset{(c)}{=} R_A^* N_0^{\ell(1-b_0)(1-\alpha/2)}, \end{aligned} \tag{7}$$

where equation (a) is obtained from (6), (b) is by substitution of (5), and (c) by defining

$$R_A^* = \begin{cases} \gamma_0 N_0^{b_0-\alpha/2}\log_2(e), & \text{low SNR;} \\ \log(1+\gamma_0 N_0^{b_0-\alpha/2}), & \text{peaky+duty cycle.} \end{cases} \quad (8)$$

Now the cut-set upper bound can be written as

$$\mathcal{C}(M_\ell, M_\ell, \gamma_\ell) = M_\ell R_\ell^* = R_A^* N_0^{\ell(1-b_0)(1-\alpha/2)} N_0^{\ell-(\ell-1)b_0} = R_A^* N_0^{b_0} N_0^{\ell(1-b_0)(2-\alpha/2)}. \quad (9)$$

Although the capacity upper bound in (9) increases exponentially with the number of hierarchy layers ℓ when the path loss exponent $\alpha<4$, the per-node rate R_ℓ^* shown in (7) decreases exponentially with ℓ for all $\alpha>2$, which is the case for almost all the outdoor wireless channels. For high loss channels with $\alpha>4$, the upper bound decreases exponentially with ℓ.

Remark 1: For extended networks consisting of many IoT devices which are power limited, even if we assume perfect CSI for free, large path loss exponent[2] ($\alpha>4$) will make the virtual MIMO cooperation counter-productive.

To maximize the capacity upper bounds for each of the virtual MIMO sessions at all levels, we choose

$$Q_{s,\ell} = Q_{d,\ell} = (1+\eta)R_\ell^*,$$

where $\eta > 0$ is the overhead as compared the minimum required rate for local communications. From (4), the aggregate throughput at layer ℓ can be written as

$$C_\ell = \frac{M_\ell R_\ell^*}{1+2(1+\eta)R_\ell^* S_{\ell-1}} = \frac{N_0^{\ell(1-b_0)+b_0}}{S_\ell}, \quad (10)$$

where

$$S_\ell = 2(1+\eta)S_{\ell-1} + \frac{1}{R_A^*} N_0^{\ell(1-b_0)(\frac{\alpha}{2}-1)} \quad (11)$$

$$= S_0(2+2\eta)^\ell + \frac{1}{R_A^*}\sum_{k=1}^{\ell}(2+2\eta)^{\ell-k} N_0^{k(1-b_0)(\frac{\alpha}{2}-1)} \quad (12)$$

$$= (2+2\eta)^\ell S_0 + \frac{1}{R_A^*} N_0^{\ell(1-b_0)(\frac{\alpha}{2}-1)} \sum_{k=0}^{\ell-1} \frac{(2+2\eta)^k}{N_0^{k(1-b_0)(\frac{\alpha}{2}-1)}}. \quad (13)$$

Theorem 1: For extended networks with large but finite network size N, the sum rate scales at most as $O(N^{\beta^*})$, with the optimal scaling factor

$$\beta^* = \begin{cases} 1+\frac{\log_N(2+2\eta)}{1-b_0} - 2\sqrt{\log_N(2+2\eta)}, & \alpha_0<\alpha<\alpha_1; \\ 2-\frac{\alpha}{2} - \frac{2-\frac{\alpha}{2}-b_0}{(1-b_0)(\frac{\alpha}{2}-1)}\log_N(2+2\eta), & \text{otherwise;} \end{cases} \quad (14)$$

where

$$\alpha_0 \triangleq 2+2(1-b_0)\log_N(2), \ \ \alpha_1 \triangleq 2+2\sqrt{\log_N(2+2\eta)}.$$

Proof: See Appendix A. ∎

Although the first scaling factor in (14) converges to 1 as the network size $N\to\infty$, this is not contradicting with the

[2]It is very likely to have $\alpha>4$ in scenarios where neither the transmitter nor the receiver is elevated above the propagation clutter.

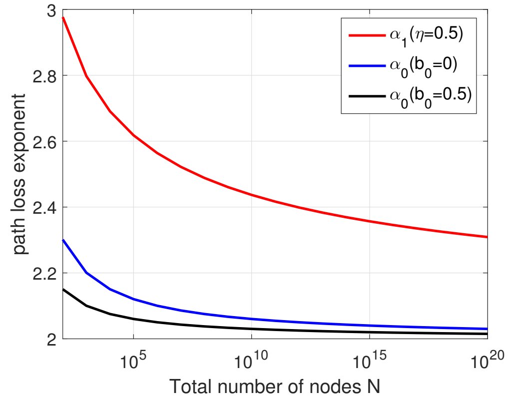

Fig. 2. The upper and lower bounds of path loss $\alpha_0<\alpha<\alpha_1$ as a function of network size to support the first scaling factor in (14).

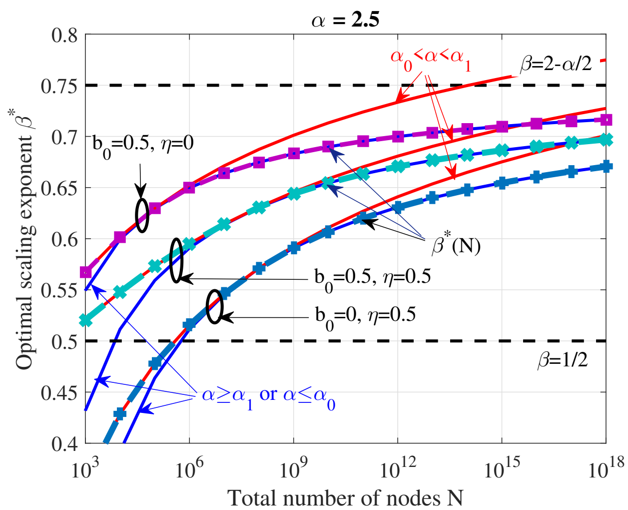

Fig. 3. The optimal scaling factor β^* as a function of the size of the extended networks, with path loss exponent α=2.5. The multi-hop scaling factor β=1/2 and the hierarchical virtual MIMO scaling factor β=2−α are also plotted as reference. The impact of the initial rate scaling factor b_0 and the local communication overhead η are also shown.

asymptotic scaling results $O(N^{2-\alpha/2})$ established by [12] since the condition $\alpha_0<\alpha<\alpha_1$ of the first optimal scaling factor in (14) holds up to network size

$$N < (2+2\eta)^{1/(\alpha/2-1)^2}, \quad (15)$$

beyond which the second exponent factor in (14) applies, and it converges to $2-\frac{\alpha}{2}$ as $N\to\infty$. Since both α_0 and α_1 monotonically decrease to 2 as N increases, as shown in Fig. 2, for small network size as defined in (15), the local cooperation cost dominates the scaling factor. For larger networks, the path loss is dominant.

To provide a quantitative sense of how local communication cost may degrade the overall performance, in Fig. 3 we plot optimal capacity scaling exponent β^* specified in (14) as a function of the total number of nodes N. For illustration purpose, we choose path loss exponent α=2.5, and the multi-hop scaling factor β=1/2 and the hierarchical virtual MIMO scaling factor β=2−α are also plotted as reference.

We can see that better message exchanging methods at the initial stage (i.e., larger b_0) helps to improve the scaling exponent β^*, With local communication overhead η=0.5, i.e., $\frac{Q_{s,\ell}+Q_{d,\ell}}{R_\ell^*} = 2(1+\eta) = 3$, the scaling factor will be reduced considerably as compared to the "ideal" case with no overhead (η=0), which is unlikely since channel/controlling information are essential to ensure the delivery of desired data. In all three cases, we observe the clear transition of β^* from regime $\alpha_0<\alpha<\alpha_1$ (in red) to $\{\alpha\le\alpha_0\} \cup \{\alpha\ge\alpha_1\}$ (in blue). As network size increases, the combined impact of path loss and local cooperation imposes a smooth transition of the scaling factor, and their dependence for any finite network size is characterized in (14). The convergence of β^* to its upper bound $2-\alpha/2$ is very slow.

IV. Conclusion

In this paper we investigate the cost of path loss and local cooperation in the capacity scaling law of wireless networks under the *extended network* model where the density of nodes is fixed and the number of nodes scales up. For small networks, the penalty on capacity scaling factor β is first dominated by the cost of local cooperation. As network size increases, it becomes increasingly challenging to deliver power/information to longer distances and the penalty on capacity scaling is dominated by path loss. The combined impact of path loss and local cooperation imposes a smooth transition of the scaling factor between these two regimes, and their dependence for any finite network size is well characterized. Although the capacity scaling factor will converge to its upper bound as $N\to\infty$, the convergence speed is very slow and the gap between the actual scaling and its upper bound is large even for networks of extremely large size.

The desire to have high throughput among large number of wireless nodes without dedicated communication infrastructure is appealing, but the reward each node can get from such collaboration within ad-hoc networks might be too small to make sense, as the per-node throughput scales as $N^{\beta-1}$. Infrastructure assisted collaboration among wireless nodes might be the way to achieve better trade-off.

Appendix A
Proof of Theorem 1

S_ℓ contains the ℓ-th partial sum of a series whose value depends on the ratio between $N_0^{(1-b_0)(\alpha/2-1)}$ and $2(1+\eta)$, and we have

$$\frac{N_0^{(1-b_0)(\frac{\alpha}{2}-1)}}{2+2\eta}\begin{cases} =1, & \text{only if } \alpha>2; \\ >1, & \text{only if } \alpha>2; \\ <1, & \alpha\le 2 \text{ or small } N_0; \end{cases} \tag{16}$$

where the last condition is due to $b_0<1$. We will analyze the scaling exponents for each of these three scenarios.

A. Case I: When $N_0^{(1-b_0)(\alpha/2-1)} = 2(1+\eta)$

Since $N_0^{(1-b_0)(\alpha/2-1)} = 2(1+\eta)$, we must have $\alpha>2$ and N_0 fixed. We can rewrite S_ℓ in (13) as

$$S_\ell = N_0^{\ell(1-b_0)(\alpha/2-1)}\left(S_0+\frac{\ell}{R_A^*}\right), \tag{17}$$

and the capacity scaling exponent can be obtained as

$$\beta_\ell = \frac{\log_{N_0}(C_\ell)}{\log_{N_0}(N_\ell)} = \frac{\ell(1-b_0)+b_0-\log_{N_0}(S_\ell)}{\ell(1-b_0)+1} \tag{18}$$

$$= 2-\frac{\alpha}{2}-\frac{2-\frac{\alpha}{2}-b_0+\log_{N_0}(S_0+\frac{\ell}{R_A^*})}{\ell(1-b_0)+1}. \tag{19}$$

Since N_0 is fixed (determined by b_0, α, and η), for large but finite network size N, we get

$$\ell = \frac{\log_{N_0}(N)-1}{1-b_0} = (\frac{\alpha}{2}-1)\log_{2+2\eta}(N)-\frac{1}{1-b_0},$$

where the first equality is due to $N_0^{\ell(1-b_0)+1}=N$ and the second step is from $N_0^{(1-b_0)(\frac{\alpha}{2}-1)}=2+2\eta$. Substituting ℓ into (19), we obtain

$$\beta = 2-\frac{\alpha}{2}-\frac{2-\frac{\alpha}{2}-b_0+\log_{N_0}(S_0+\frac{\log_{N_0}(N)-1}{R_A^*(1-b_0)})}{\log_{N_0}(N)} \tag{20}$$

$$\simeq 2-\frac{\alpha}{2}-\frac{2-\frac{\alpha}{2}-b_0}{(1-b_0)(\frac{\alpha}{2}-1)}\log_N(2+2\eta), \tag{21}$$

where the last step is due to $S_0+\frac{\ell}{R_A^*}\ll N$.

B. Case II: When $N_0^{(1-b_0)(\alpha/2-1)} > 2(1+\eta)$

From $N_0^{(1-b_0)(\alpha/2-1)} > 2(1+\eta)$ we must have $\alpha>2$. Then S_ℓ can be rewritten as

$$S_\ell = N_0^{\ell(1-b_0)(\frac{\alpha}{2}-1)}\Omega_\ell, \tag{22}$$

where

$$\Omega_\ell \triangleq \frac{S_0(2+2\eta)^\ell}{N_0^{\ell(1-b_0)(\frac{\alpha}{2}-1)}}+\frac{1}{R_A^*}\frac{1-\frac{(2+2\eta)^{\ell+1}}{N_0^{(\ell+1)(1-b_0)(\alpha/2-1)}}}{1-\frac{(2+2\eta)}{N_0^{(1-b_0)(\frac{\alpha}{2}-1)}}}. \tag{23}$$

Note that Ω_ℓ can be upper bounded by

$$\Omega_\ell < \frac{S_0(2+2\eta)^\ell}{N_0^{\ell(1-b_0)(\frac{\alpha}{2}-1)}}+\frac{1}{R_A^*}\frac{1}{1-\frac{2+2\eta}{N_0^{(1-b_0)(\frac{\alpha}{2}-1)}}}, \tag{24}$$

and lower bounded by (applying the AM–GM inequality to the partial sum of the series)

$$\Omega_\ell > \frac{S_0(2+2\eta)^\ell}{N_0^{\ell(1-b_0)(\frac{\alpha}{2}-1)}}+\frac{\ell}{R_A^*}\frac{(2+2\eta)^{\frac{\ell-1}{2}}}{N_0^{(\frac{\ell-1}{2})(1-b_0)(\frac{\alpha}{2}-1)}}. \tag{25}$$

We can obtain from (19) that

$$\beta_\ell = 2-\frac{\alpha}{2}-\frac{2-\frac{\alpha}{2}-b_0}{\ell(1-b_0)+1}-\log_{N_\ell}(\Omega_\ell). \tag{26}$$

Since $N=N_0^{\ell(1-b_0)+1}$ and $N_0>(2+2\eta)^{\frac{1}{(1-b_0)(\alpha/2-1)}}$, we have

$$\ell < (\frac{\alpha}{2}-1)\log_{2+2\eta}(N)-\frac{1}{1-b_0}. \tag{27}$$

By substitution of (27) into (26) and and dropping the small term $\log_N(\Omega_\ell)$, we can obtain the optimal scaling factor β^* for each finite network size N as

$$\beta^*(N) \simeq 2-\frac{\alpha}{2}-\frac{2-\frac{\alpha}{2}-b_0}{(1-b_0)(\frac{\alpha}{2}-1)}\log_N(2+2\eta). \tag{28}$$

C. Case III: When $N_0^{(1-b_0)(\alpha/2-1)} < 2(1+\eta)$

When $N_0^{(1-b_0)(\alpha/2-1)} < 2(1+\eta)$, we have

$$S_\ell = (2{+}2\eta)^\ell \Theta_\ell, \tag{29}$$

where

$$\Theta_\ell \triangleq S_0 + \frac{1 - \frac{N_0^{\ell(1-b_0)(\alpha/2-1)}}{(2+2\eta)^\ell}}{R_A^*(\frac{(2+2\eta)}{N_0^{(1-b_0)(\alpha/2-1)}} - 1)}. \tag{30}$$

Similarly, we can upper and lower bound Θ_ℓ as follows

$$\Theta_\ell < S_0 + \frac{1}{R_A^*(\frac{(2+2\eta)}{N_0^{(1-b_0)(\alpha/2-1)}} - 1)}, \tag{31}$$

$$\Theta_\ell > S_0 + \frac{\ell}{R_A^*}\frac{N_0^{(1-b_0)(\alpha/2-1)(\ell+1)/2}}{(2+2\eta)^{(\ell+1)/2}}. \tag{32}$$

We obtain from (19) that

$$\begin{aligned}\beta_\ell &= 1 - \frac{\log_{N_0}(2{+}2\eta)}{1-b_0} - \frac{1{-}b_0{-}\log_{N_0^{1-b_0}}(2{+}2\eta)}{\ell(1-b_0)+1} - \log_{N_\ell}(\Theta_\ell) \\ &= 1 - \ell \log_{N_\ell}(2{+}2\eta) - \frac{1{-}b_0}{\ell(1{-}b_0){+}1} - \log_{N_\ell}(\Theta_\ell).\end{aligned} \tag{33}$$

For large but finite N, we can obtain the optimized scaling factor β^* by dropping the small term $\log_N(\Theta_\ell)$ from (33) and optimizing over ℓ, which reaches its maximum at

$$\ell^* = \sqrt{\log_{2+2\eta}(N)} - \frac{1}{1-b_0}. \tag{34}$$

When $\alpha > 2$, we have $2 \leq N_0 < (2+2\eta)^{\frac{1}{(1-b_0)(\alpha/2-1)}}$ and

$$(\frac{\alpha}{2} - 1)\log_{2+2\eta}(N) - \frac{1}{1-b_0} < \ell \leq \frac{\log_2(N) - 1}{1-b_0}. \tag{35}$$

As long as

$$2 < \alpha < 2 + 2\sqrt{\log_N(2+2\eta)}, \tag{36}$$

the optimal cooperation layer ℓ^* falls within the constraint (35) and therefore the optimal scaling factor is

$$\beta^*(N) \simeq 1 + \frac{\log_N(2+2\eta)}{1-b_0} - 2\sqrt{\log_N(2+2\eta)}. \tag{37}$$

When the path loss exponent $\alpha \geq 2 + 2\sqrt{\log_N(2+2\eta)}$, the optimal ℓ^* in (34) falls on the left outside the constraint (35). Therefore, the optimal scaling factor becomes

$$\beta^*(N) \simeq 2 - \frac{\alpha}{2} - \frac{2 - \frac{\alpha}{2} - b_0}{(1-b_0)(\frac{\alpha}{2}-1)}\log_N(2+2\eta), \tag{38}$$

with the optimal cooperation layer

$$\ell^* = (\frac{\alpha}{2} - 1)\log_{2+2\eta}(N) - \frac{1}{1-b_0}. \tag{39}$$

For $\alpha{\leq}2$, we have $N_0^{(1-b_0)(\alpha/2-1)}{<}2(1{+}\eta)$ for all feasible N_0. Therefore, the constraint on ℓ reduces to $\ell \leq \frac{\log_2(N)-1}{1-b_0}$, and the optimal cooperation layer ℓ^* is given by (34) and the optimal scaling factor β^* is given by (37).

Although the exact expressions of scaling factor β_ℓ stated in (20), (26), and (33) are chosen based on the conditions in (16), the optimized scaling factor $\beta^*(N)$ (after dropping small terms) given in (21), (28), (37) and (38) are mainly determined by the path loss exponent α and the local communication overhead η. Therefore, for extended networks with large but finite network size N, the optimal scaling factor can be written as (14), and the consolidation of results comes from the fact that, for $\alpha_0{<}\alpha{<}\alpha_1$, one can adjust the value of N_0 to switch between conditions specified in (16).

References

[1] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, Applications and Research Challenges," *ELSEVIER Ad Hoc Networks*, vol. 10, pp .1497–1516, Sep. 2012.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," *ELSEVIER Future Generation Computer Systems*, vol. 29, pp. 1645–1660, Sep. 2013.

[3] M. T. Lazarescu, "Design of a WSN Platform for Long-Term Environmental Monitoring for IoT Applications," *IEEE Journal on Emerging & Selected Topics in Circuits and Systems*, vol. 3, pp. 45–54, Mar. 2013.

[4] A. Al-Fuqaha et al., "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys & Tutorials*, vol. 17, pp. 2347–2376, Nov. 2015.

[5] M. R. Palattella et al., "Internet of Things in the 5G Era: Enablers, Architecture, and Business Models," *IEEE Journal on Selected Areas in Communications*, vol. 34, pp. 510–527, Mar. 2016.

[6] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless Backhauling of 5G Small Cells: Challenges and Solution Approaches," *IEEE Wireless Communications*, vol. 22, pp. 22–31, Oct. 2015.

[7] N. Wang, E. Hossain and V. K. Bhargava, "Backhauling 5G small cells: A radio resource management perspective," *IEEE Wireless Communications*, vol. 22, pp. 41-49, Oct. 2015.

[8] Y. Zeng, R. Zhang and T. J. Lim, "Wireless Communications with Unmanned Aerial Vehicles: Opportunities and Challenges," *IEEE Communications Magazine*, vol. 54, pp. 36–42, May 2016.

[9] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388–404, Mar. 2000.

[10] L.-L. Xie and P. R. Kumar, "A Network Information Theory for Wireless Communication: Scaling Laws and Optimal Operation," *IEEE Transactions on Information Theory*, vol. 50, pp. 748–767, May 2004.

[11] L.-L. Xie and P. R. Kumar, "On the Path-Loss Attenuation Regime for Positive Cost and Linear Scaling of Transport Capacity in Wireless Networks," *IEEE Transactions on Information Theory*, vol. 52, pp. 2313–2328, Jun. 2006.

[12] A. Özgür, O. Lévêque, and D. N. C. Tse, "Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad Hoc Networks," *IEEE Transactions on Information Theory*, vol. 53, pp. 3549–3572, Oct. 2007.

[13] J. Ghaderi, L.-L. Xie, and X. Shen, "Hierarchical Cooperation in Ad Hoc Networks: Optimal Clustering and Achievable Throughput," *IEEE Transactions on Information Theory*, vol. 55, pp. 3425–3436, Aug. 2009.

[14] S.-N. Hong and G. Caire,"On the Performance of Optimized Dense Device-to-Device Wireless Networks," *arXiv:1402.1815v2*, Jun. 2015.

[15] A. Özgür, O. Lévêque, and D. N. C. Tse, "Spatial Degrees of Freedom of Large Distributed MIMO Systems and Wireless Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 202–214, Feb. 2013.

[16] A. El Gamal, J. Mammen, B. Prabhakar and D. Shah, "Optimal Throughput-Delay Scaling in Wireless Networks - Part I: the Fluid Model," *IEEE Transactions on Information Theory*, vol. 52, pp. 2568–2592, Jun. 2006.

[17] N. Lu and X. S. Shen, "Scaling Laws for Throughput Capacity and Delay in Wireless Networks A Survey," *IEEE Communications Surveys & Tutorials*, vol. 16, pp. 642–657, Aug. 2014.

[18] J. Du, M. Médard, and S. Shamai, "Cost of Local Cooperation in Hierarchical Virtual MIMO Transmission Schemes," in *proc. IEEE ITW*, Sep. 2016.

[19] N. Fawaz and M. Médard, "A Converse for the Wideband Relay Channel with Physically Degraded Broadcast," in *proc. IEEE ITW*, Oct. 2011.

[20] M. Franceschetti, O. Dousse, D. N. C. Tse and P. Thiran, "Closing the Gap in the Capacity of Wireless Networks Via Percolation Theory," *IEEE Trans. Information Theory*, vol. 53, pp. 1009–1018, Mar. 2007.