Capacity of Cellular Wireless Network

Rahul Vaze

School of Technology and Computer Science
Tata Institute of Fundamental Research, Mumbai
email: vaze@tcs.tifr.res.in

Srikanth K. Iyer

Department of Mathematics
Indian Institute of Science, Bangalore
email: skiyer@math.iisc.ernet.in

***Abstract*—In cellular networks, under ARQ and SINR model of transmission, the effective downlink rate of packet transmission is the reciprocal of the expected delay (number of retransmissions needed till success). We define the cellular network capacity as the ratio of the basestation (BS) density and the expected delay. Exact characterization of this natural and practical but non-trivial (because of SINR temporal correlations) capacity metric is derived. The capacity is shown to first increase polynomially with the BS density and then scale inverse exponentially with the increasing BS density. Two distinct upper bounds are derived that are relevant for the low and the high BS density regimes. A single power control strategy is shown to achieve the upper bounds in both the regimes upto constants. Our result is fundamentally different than the transport and transmission capacity for ad hoc networks that scale as the square root of the (high) BS density. Our results show that the strong temporal correlations of SINRs with PPP distributed BS locations model for cellular networks is limiting, and the realizable capacity is much smaller than previously thought.**

***Index Terms*—Capacity, Cellular Wireless Networks, Poisson point process, ARQ.**

I. Introduction

Finding the Shannon capacity of a wireless network is perhaps one of the most well-studied problem that has remained unsolved. For ad hoc wireless networks, two slightly relaxed capacity notions have been defined, transport [1] and transmission [2], for which theoretical results have been possible mainly because of two important simplifications; SINR model of communication, and assuming random locations for nodes [3]. Under the SINR model, communication between two nodes is deemed successful if the SINR between them is larger than a threshold that depends on the rate of transmission. Even under these simplifications, as far as we know, there has been no fundamental characterization of the maximum throughput possible (under any reasonable capacity definition) in *cellular* wireless networks that are structured, rather than being ad hoc.

In this paper, we consider the well-accepted (tractable) model of a cellular wireless network [4], that has basestation (BS) locations distributed as a homogenous Poisson point process (PPP), and the focus is on characterizing the 'capacity' in the downlink. The tractable model is a reasonable abstraction in the modern scenario, where multiple layers of BSs (macro, micro, femto) are overlaid over each other. Moreover, we consider the widely used BS-MU (mobile user) association rule, where each MU connects to its nearest BS, i.e., BS serves all the MUs located in its Voronoi cell. The mobile user (MU) locations are also assumed to be distributed as a homogenous PPP, independent of the BS locations process.

We consider the SINR model of transmission for each BS-MU communication with ARQ, where a packet is retransmitted from the BS until the SINR seen at the MU is above a fixed threshold. Also, each BS serves all its MUs in a round-robin manner by dividing its slots/bandwidth equally among them to closely model the 'fair' practical implementation.

BSs in real-life cellular wireless networks are limited in their ability to coordinate their transmissions in order to control inter-cell interference. We consider that each BS is only allowed to use local strategies, i.e., each BS's transmission decisions are only based on local channel conditions (path-loss or fading gain) at the MU or some feedback (ack/nacks) from the MU it is serving. Our results also apply for small-scale BS coordination, where a fixed number of neighbouring basestations can schedule their transmissions together, e.g. CoMP.

With ARQ, let D be the number of retransmissions needed for a packet from BS x to be successful at its MU y, i.e.,

$$D = \min\{t : \mathsf{SINR}_{xy}(t) \ge \beta\}.$$

D has the interpretation of delay, the time (or retransmissions) needed to receive the packet successfully. Let λ be the BS density (per m^2) of the cellular network. With ARQ, a natural throughput metric is the reciprocal of the expected delay. We use this motivation and consider a natural definition of **capacity** (in the downlink), where under the round-robin policy, the per-BS capacity with BS strategy $\mathcal{S}$ is defined as $C_b(\mathcal{S}) = \frac{1}{\mathbb{E}\{D\}}$, and the network wide capacity with $\mathcal{S}$ is $C(\mathcal{S}) = \frac{\lambda}{\mathbb{E}\{D\}}$ packets/sec/m^2. Hence, our capacity definition is

$$C = \sup_{\mathcal{S}} C(\mathcal{S}), \quad (1)$$

i.e., we are looking for the best possible local (adaptive) BS strategy $\mathcal{S}$ that achieves the maximum network-wide throughput. With round-robin scheduling from each BS, C is independent of the MU density.

A. Prior work on Network Capacity

Two related and well-known metrics of capacity; transport and transmission, have been defined for ad hoc networks and for both, exact results have been obtained. Under the path-loss model, where fading gain is neglected, transport capacity has been shown to be $\Theta(\sqrt{\lambda})$ [1], [5]. Some extensions of

978-3-901882-90-6/17 ©2017 IFIP

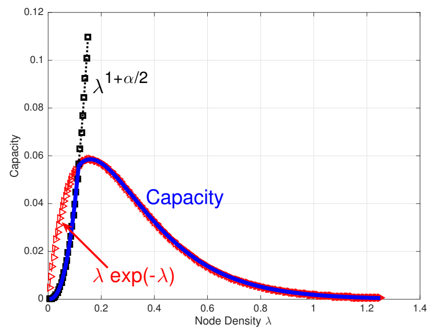

Fig. 1. Capacity of the cellular wireless network as a function of the BS density for path-loss exponent $\alpha = 3$.

transport capacity are also known [6]. Distance (transmitter-receiver) scaled transmission capacity also scales as $\Theta(\sqrt{\lambda})$ [2], [7] for both the path-loss and the path-loss-plus-fading model.

With retransmissions/ARQ, a generalization of the transmission capacity, called the *delay normalized* transmission capacity $\mathcal{S}$ [8], is defined as the number of successfully delivered packets in the network under the SINR model subject to a maximum limit on the number of retransmissions. For simplifying analysis, [8] made a limiting assumption that SINRs across time slots are independent. A more rigorous approach was taken in [9], but yielded limited analytical results. The basic problem in studying the delay normalized transmission capacity or our capacity definition (1) is the complicated correlation of SINRs across time slots under the PPP assumption on BS locations [10]. It is worth mentioning that a ALOHA strategy has capacity (1) $C = 0$, for any BS density λ following the result [11], that shows that the ALOHA strategy has infinite expected delay.

B. Our results

In this paper, we avoid making any simplifying/limiting assumptions for studying the joint distribution of SINRs across time slots for deriving bounds on the expected delay seen at any MU. The main result of our paper is that, when each BS is only allowed a local strategy or BSs can accomplish only small-scale coordination, upto constants, [1]

$$\min_{\mathcal{S}} \mathbb{E}\{D\} = \max\left\{\frac{1}{\lambda^{\frac{\alpha}{2}}}, \exp(\lambda)\right\},$$

and consequently, the capacity is

$$C = \min\left\{\lambda^{\frac{\alpha}{2}+1}, \lambda\exp(-\lambda)\right\}, \text{ packets/sec/m}^2,$$

where $\alpha > 2$ is the path-loss exponent. In Fig. 1, we illustrate the capacity (upto a constant) as the blue curve for $\alpha = 3$. The main result is derived via the following sub-results.

- We show that for any local strategy $\mathcal{S}$, $\mathbb{E}\{D\} \geq \left(\frac{c_1}{\lambda^{\frac{\alpha}{2}}}\right)$, that is relevant for low-BS density regime, and $\mathbb{E}\{D\} \geq c_4\left(\exp(c_2\lambda)\right)$ that is tighter in the high-BS density regime.
- The upper bound on the capacity holds even if small scale BS coordination is allowed, where a fixed number (independent of the BS density) of BSs can schedule their transmissions jointly.
- We show that a simple non-adaptive strategy $\mathcal{S}$ (first proposed by us in [12]), where each BS transmits power to completely nullify the path-loss based signal attenuation at the MU it is serving (using only the knowledge of the distance to the MU and is independent of the fading gains), achieves

$$\mathbb{E}\{D\} \leq \sqrt{\left(1+\frac{\Gamma(\alpha+1)}{(\pi\lambda)^{\alpha}}\right)\exp\left(c_5\lambda\right)}, \text{ and}$$

$$C(\mathcal{S}) \geq \min\left\{c_8\lambda^{\frac{\alpha}{2}+1}\exp(-c_6\lambda), \frac{\lambda\exp(-c_6\lambda)}{c_7}\right\},$$

matching the upper bound on the capacity for any strategy $\mathcal{S}$. Remarkably, a single policy achieves the capacity upper bound for the both the low and the high-BS density regimes, for which the capacity scaling behaviour is quite different. The policy is similar to full-power control used by MUs in uplink for LTE systems.

Our result shows that in the low-BS density regime, the capacity increases polynomially with the BS density, where interference is weak and lower signal power can be compensated using the power control strategy [12]. In the high-BS density regime, where communication is interference-limited, our result shows that the 'realistic' capacity of a cellular wireless network is fundamentally different and significantly smaller (falls exponentially with λ) than earlier results on transport or the transmission capacity (both scale as $\Theta(\sqrt{\lambda})$) in the high-BS density regime. The capacity behaviour in both the low and the high-BS density cases is tightly governed by the temporal SINR correlations with PPP distributed BSs.

Our definition of capacity is simple yet practically realistic for cellular wireless networks, where ARQ is universally implemented. It is also mathematically rich because of temporal correlations in SINRs with PPP distributed BS process [10]. As far as we know, our result is a first simple and exact characterisation of a reasonable capacity definition for cellular wireless networks, without requiring any limiting assumptions.

We show that the capacity (both upper and lower bound) is indifferent to the value of path-loss exponent α for $\lambda > 1$, while for low-BS density regime $\lambda < 1$, we show that the capacity decreases with α as $\lambda^{1+\alpha/2}$. Thus, the effect of α is only visible at low-BS densities, where interference is weak, and larger nearest BS distance together with high value of α significantly limits the signal power of interest.

One important design implication of our result is in terms of cell densification [13]–[17], where number of basestations are increased to decrease the individual cell-sizes in the hope of improving connectivity and communication rate. Our results characterize the exact effect of cell densification on the capacity/long-term throughput.

[1]The lower and upper bounds only differ by constants, and all c_i's used to describe bounds are constants.

C. Comparison with Prior Work

To compare our capacity metric with transport capacity or transmission capacity (relevant for high BS density), one has to multiply the average distance between the MU and its nearest BS distance d_0 (that scales as $\Theta\left(\frac{1}{\sqrt{\lambda}}\right)$) with the capacity, which again yields $Cd_0 = \Theta\left(\exp(-\lambda)\right)$. The reasons for arriving at such fundamentally different scaling compared to the transport capacity and the transmission capacity, that both scale as $\Theta\left(\sqrt{\lambda}\right)$, are that transport capacity has stricter reliability requirement (SINR has to be $\geq \beta$ for each successful counted bit on realization basis) but allows for large scale BS/node scheduling, while transmission capacity has one-shot reliability constraint of $1-\epsilon$, and does not allow for retransmissions. In cellular networks, large scale BS scheduling is not possible, while retransmissions are an integral part of any deployment. The inability of large scale BS scheduling ensures that interference seen at any MU increases with BS density, while with retransmissions, the temporal correlation of SINRs degrades the delay performance.

II. System Model

We consider a cellular network model, that consists of basestations $\{T_n\}$, whose locations are distributed according to a homogeneous PPP $\Phi = \{T_n\}$ with density λ, popularly called the *tractable* model starting from [4]. The MUs are also located according to an independent PPP Φ_R with density μ with $\mu >> \lambda$, and each MU connects to its nearest basestation. Therefore, by the definition of Voronoi regions with respect to basestation locations, all MUs in a Voronoi cell/region connect to its representative basestation.

We consider a typical MU m, and to study the capacity C, focus on the expected delay encountered by its packets transmitted from its nearest basestation $\mathsf{n}(\mathsf{m})$. We consider that time is slotted, and each BS transmits to all the MUs that lie in its Voronoi region/cell in a round-robin manner, i.e., equally sharing time slots between them. We consider the distance based path-loss (signal attenuation) function to be $\ell(d) = \min\{1, d^{-\alpha}\}$, for $\alpha > 2$. A simple function $\ell(d) = d^{-\alpha}$ is used widely in literature, however, for $d < 1$, it produces signal amplification which is unrealizable.

We assume that each node in the network is equipped with a single antenna, and at time t, the fading gain between BS x and MU y is denoted by $h_{x,y}(t)$ that is assumed to be exponentially distributed with parameter 1. Moreover, $h_{x,y}(t)$ is assumed to be independent and identically distributed for all time t, and all location pairs x, y. We assume that each BS has an average power constraint of M.

Without loss of generality, we assume that the typical MU m is located at the origin, and the distance to the nearest basestation from origin is d_0, while any other BS z is located at distance z (abuse of notation) from the origin. Since we focus on only the MU m located at the origin, we abbreviate $h_{z,0}(t)$ to just $h_z(t)$. For BS z, let $P_z(t)$ be the transmit power at time t, and $\mathbf{1}_z(t)$ be the indicator variable denoting whether BS z is transmitting at time t or not, with $\mathbb{P}(\mathbf{1}_z(t) = 1) = p_z(t)$.

The received signal at m located at the origin that is connected to its nearest basestation $\mathsf{n}(\mathsf{m})$ is given by

$$\begin{aligned} y(t) =& \sqrt{P_{\mathsf{n}(\mathsf{m})}(t)\ell(d_0)h_{\mathsf{n}(\mathsf{m})}(t)}\mathbf{1}_{\mathsf{n}(\mathsf{m})}(t)s_t(\mathsf{n}(\mathsf{m})) \\ &+ \sum_{z\in\Phi\backslash\{\mathsf{n}(\mathsf{m})\}} \sqrt{\gamma P_z(t)\ell(z)h_z(t)}\mathbf{1}_z(t)s_t(z) + w, \quad (2) \end{aligned}$$

where $s_t(z)$ is the signal transmitted from BS z with power $P_z(t)$ at time t, w is the AWGN with variance N, and $0 < \gamma \leq 1$ is the processing gain of the system or the interference suppression parameter.

Thus, the SINR at m from its nearest BS $\mathsf{n}(\mathsf{m})$ in time slot t is given by

$$\mathsf{SINR}_{\mathsf{n}(\mathsf{m}),\mathsf{m}}(t) = \frac{P_{\mathsf{n}(\mathsf{m})}(t)h_{\mathsf{n}(\mathsf{m})}(t)\ell(d_0)\mathbf{1}_{\mathsf{n}(\mathsf{m})}(t)}{\gamma\sum_{z\in\Phi\backslash\{\mathsf{n}(\mathsf{m})\}} P_z(t)\mathbf{1}_z(t)h_z(t)\ell(z) + \mathsf{N}}. \quad (3)$$

The transmission from BS x to MU y is deemed *successful* at time t, if $\mathsf{SINR}_{xy}(t) > \beta$, where $\beta > 0$ is a fixed threshold depending on the rate of information transfer. Let

$$e_{xy}(t) = \begin{cases} 1 & \text{if } \mathsf{SINR}_{xy}(t) > \beta, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Since $h_{(.)}(t)$ is a random variable, multiple transmissions may be required for a packet to be successfully received at any node. Thus, a measure of delay, i.e., the number of retransmissions needed to successfully receive packets, is required, which is defined as follows.

Definition 1. *Let the minimum time (delay) taken by any packet to be successfully received at* m *located at the origin from its nearest BS* $\mathsf{n}(\mathsf{m})$ *be* $D = \min\left\{t > 0 : e_{\mathsf{n}(\mathsf{m}),\mathsf{m}}(t) = 1\right\}$. *Consequently, we define the network capacity to be*

$$C = \sup_{\mathcal{S}} \frac{\lambda}{\mathbb{E}\{D\}},$$

where $\mathcal{S}$ *is any BS strategy followed by all BSs.*

III. Upper Bound on Capacity

Definition 2. *Let* $\mathcal{I}$ *be a countable index set. Let* $\mathcal{S} = \{(p_i, P_i), i \in \mathcal{I} : p_i P_i \in [M/\tau, M]\}$, *where* $\tau \geq 1$ *is a constant, be a collection of probability of transmission* p_i *and power transmission* P_i *pairs. A strategy for BS is to choose any element of* $\mathcal{S}$ *at any given time slot, where the current choice can be adaptive, i.e., it can depend on entire history of earlier choices.* [2] *Note that with our definition, the average power constraint of* M *is satisfied automatically.*

We begin by restricting ourselves to local (easily implementable) strategies that are defined as follows.

Definition 3. *A BS strategy is called* **local**, *if it only depends on either the fading gain or the distance between itself and the*

[2] This strategy can closely emulate any transmission policy used by BS under an average power constraint M. If p_s, P_s is chosen for a slot, then BS transmits with probability p_s and power P_s. The lower bound on $p_i P_i \geq M/\tau$ reflects the considered setting of $\lambda << \mu$, where each BS has at least one MU in its Voronoi cell.

connected MU m*, or the history of success/failure event* $e_{\mathsf{n(m)},\mathsf{m}}$ *at* m *i.e., ack/nack signal sent back by MU* m*, or history of earlier transmitted power and probability of transmission. In particular, BS* $\mathsf{n(m)}$*, can choose* $p_{\mathsf{n(m)}}(t), P_{\mathsf{n(m)}}(t)$ *pair for time slot* t*, depending on* $h_{\mathsf{n(m)}}(t)$*,* d_0*, and history of* $e_{\mathsf{n(m)},\mathsf{m}}(s)$*, and* $p_{\mathsf{n(m)}}(s), P_{\mathsf{n(m)}}(s)$ *for* $1 \leq s < t$*.*

Theorem 1. *For any local strategy* $\mathcal{S}$ *employed by a BS, the expected delay at the typical MU* m *satisfies*

$$\mathbb{E}\{D\} \geq \left(\frac{c_1}{\lambda^{\frac{\alpha}{2}}}\right), \tag{5}$$

as well as

$$\mathbb{E}\{D\} \geq c_4 \left(\exp(c_2\lambda)\right). \tag{6}$$

Consequently, for any $\mathcal{S}$*, the network wide capacity is*

$$C = \frac{\lambda}{\mathbb{E}\{D\}} \leq \min\left\{\frac{\lambda^{\frac{\alpha}{2}+1}}{c_1}, c_3\lambda\exp(-c_2\lambda)\right\}.$$

Theorem 1 shows that for low-BS densities, the expected delay decreases polynomially with the BS density, where (5) dominates. Lower bound (5) is derived by removing the interference, which is anyway weak in low-BS density regime. Without intereference, the nearest BS-MU distance completely controls the expected delay, and increasing BS density decreases the nearest BS-MU distance and boosts the signal power. For moderate and high BS densities, (6) dominates, and our result shows that the expected delay grows at least exponentially with the BS density. Theorem 1 is valid (with different constants) even if some fixed number k of BSs (independent of λ) can coordinate their transmissions, where interference comes from BSs lying outside of radius d_k (the k^{th} nearest BS), rather than d_0 as is the case in Theorem 1. For lack of space, we omit the proof.

For large BS densities, Theorem 1 is essentially a negative result that shows that even when each BS has all the local information, that can be used adaptively, the expected delay increases exponentially with the density of BSs. Consequently, the network wide capacity decreases exponentially with the increase in the density of BSs in the high-BS density regime.

Theorem 1 also suggests that both transport and transmission capacity definitions $\left(\text{both scale as } \Theta\left(\sqrt{\lambda}\right)\right)$ overestimate the practically realizable capacity of a cellular wireless network in the high-BS density regime. The reason for this is that transport capacity allows for large scale BS coordination, while with transmission capacity, communication is one-shot with relaxed reliability constraint and no retransmissions.

IV. Achievability

In this section, we consider a simple BS strategy (proposed by us in [12]) to achieve the capacity upper bound (Theorem 1) upto the same order.

Strategy: Let for MU u, the distance to its nearest BS $\mathsf{n}(u)$ be d_u. Let each BS know d_u for all the users connected to it. For a slot t designated for a particular MU u, BS $\mathsf{n}(u)$ transmits with probability $p_{\mathsf{n}(u)}(t)$ with power $P_{\mathsf{n}(u)}(t)$ given by

$$P_{\mathsf{n}(u)}(t) = c\ell^{-1}(d_u), \qquad p_{\mathsf{n}(u)}(t) = M(P_{\mathsf{n}(u)}(t))^{-1}, \tag{7}$$

where $c = M(1-\epsilon)^{-1}$, $0 < \epsilon < 1, \beta\gamma(1-\epsilon) < 1$ is a constant, and $M = P_{\mathsf{n}(u)}(t)p_{\mathsf{n}(u)}(t)$ is the average power constraint. Condition $\beta\gamma(1-\epsilon) < 1$ is technical. Thus, in each time slot, with (7), each BS makes transmission attempts with transmission power proportional to the distance to the MU it is serving, to completely nullify the path-loss. The transmission probability is chosen so as to satisfy the average power constraint of M. It is worthwhile noting that the strategy does not use the knowledge of fading gain $h_{\mathsf{n(m)}}(t)$, and is not an adaptive strategy.

Theorem 2. *The power control strategy (7) achieves the following performance* $\mathbb{E}\{D\} \leq \sqrt{\left(1+\frac{c_4}{(\pi\lambda)^\alpha}\right)\exp\left(c_3\lambda\right)}$*, where* c_3 *and* c_4 *are constants. Thus,*

$$C \geq \frac{\lambda}{\mathbb{E}\{D\}} \geq \min\left\{c_6\lambda^{\frac{\alpha}{2}+1}\exp(-c_5\lambda), c_7\lambda\exp(-c_5\lambda)\right\}.$$

Theorem 2 shows that a simple non-adaptive strategy that does not need to learn the local fading gain is capable of achieving (order-wise) the upper bound on the capacity C. The only local information it needs is the distance d_0 that can be learned easily via ranging or RSSI measurements, making the power control strategy easily implementable in practice. The proof of Theorem 2 is similar to the one derived in [12], where the focus was only to show that the expected delay is finite, while here we need the exact scaling result.

V. Conclusion

In this paper, we proposed a natural and practical definition of capacity for cellular wireless networks. Our capacity metric is weaker than the Shannon capacity, however, it closely matches the throughput measure observed in a real-life implementation of cellular wireless networks with ARQ. Most importantly, we were able to derive the exact dependence of the BS density on the capacity of cellular networks, which has generally escaped analytical tractability. Conventional wisdom suggests that there is an advantage in increasing BS density; via increasing the SINR for the cell-edge users or improving the frequency reuse. We showed that that is true only for low-BS densities, where the capacity increases polynomially with the BS density, while as BS density is increased further the capacity starts to decrease exponentially.

References

[1] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar 2000.

[2] S. Weber, X. Yang, J. Andrews, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with outage constraints," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4091–4102, Dec. 2005.

[3] R. Vaze, *Random Wireless Networks*. Cambridge University Press, 2015.

[4] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[5] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "Closing the gap in the capacity of random wireless networks," in *Proc. Int. Symp. Information Theory ISIT 2004*, 2004.

[6] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3549–3572, Oct. 2007.

[7] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 421–436, 2006.

[8] J. Andrews, S. Weber, M. Kountouris, and M. Haenggi, "Random access transport capacity," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2101 –2111, June 2010.

[9] R. Vaze, "Throughput-delay-reliability tradeoff with ARQ in wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2142 –2149, Jul. 2011.

[10] R. Ganti and M. Haenggi, "Spatial and temporal correlation of the interference in aloha ad hoc networks," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 631 –633, sept. 2009.

[11] F. Baccelli, B. Błaszczyszyn, and M.-O. Haji-Mirsadeghi, "Optimal paths on the space-time SINR random graph," *Advances in Applied Probability*, pp. 131–150, 2011.

[12] S. K. Iyer and R. Vaze, "Achieving non-zero information velocity in wireless networks," *Annals of Applied Probability*, to appear, 2017.

[13] X. Zhang and J. G. Andrews, "Downlink cellular network analysis with multi-slope path loss models," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1881–1894, 2015.

[14] C. S. Chen, V. M. Nguyen, and L. Thomas, "On small cell network deployment: A comparative study of random and grid topologies," in *Vehicular Technology Conference (VTC Fall),*. IEEE, 2012, pp. 1–5.

[15] T. Samarasinghe, H. Inaltekin, and J. S. Evans, "On optimal downlink coverage in Poisson cellular networks with power density constraints," *IEEE Trans. Commun.*, vol. 62, no. 4, pp. 1382–1392, 2014.

[16] F. Baccelli and A. Biswas, "On scaling limits of power law shot-noise fields," *Stochastic Models*, vol. 31, no. 2, pp. 187–207, 2015.

[17] V. M. Nguyen and M. Kountouris, "Performance limits of network densification," subm., 2016. [Online]. Available: https://arxiv.org/abs/1611.07790

[18] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468–1489, 1999.

Appendix A
Proof of Theorem 1

We first prove the lower bound (5) via the following result.

Proposition 4. *[18] To minimize $\mathbb{P}(h_i P_i \le c)$, where h_i is i.i.d. $\sim EXP(1)$, the optimal power allocation is given by*

$$P_i = \begin{cases} c/h_i & h_i \ge \delta, \\ 0 & h_i < \delta, \end{cases}$$

where δ is chosen to satisfy the average power constraint M.

[Proof of Theorem 1] Consider that other than BS $\mathsf{n(m)}$ there is no other active BS in the network. Thus, the interference seen at m is zero. Then the SINR seen at m in any time slot t

$$\mathsf{SINR}_{\mathsf{n(m)},\mathsf{m}}(t) = \frac{P_{\mathsf{n(m)}} h_{\mathsf{n(m)}} \mathbf{1}_{\mathsf{n(m)}}(t) \ell(d_0)}{\mathsf{N}}. \tag{8}$$

Given a realization of the BS PPP Φ, the distance d_0 is fixed. Given Φ, we will find a lower bound on $\mathbb{E}\{D|\Phi\}$. With fixed M, β, for two different cases depending on d_0, we will consider stronger fading gains to lower bound the outage probabilities, and consequently lower bound the expected delay, as follows.

Case 1: $\ell(d_0) \ge \frac{\beta}{M}$. For this case, we first show that the outage probability $p_o = \mathbb{P}(\mathsf{SINR} \le \beta|\Phi)$ (8) with $h \sim EXP(1)$ is larger than when $h \equiv 1$ always (no fading/line of sight). With $h \sim EXP(1)$, from Proposition 4, for any M, the outage probability $p_o = \mathbb{P}(\mathsf{SINR} \le \beta|\Phi)$ is at least $1 - \exp(-\delta)$, where δ is chosen to satisfy the average power constraint $\int_\delta^\infty \frac{\beta}{\ell(d_0)x} \exp(-x)dx = M$. With $h \equiv 1$, when $\frac{M\ell(d_0)}{\beta} \ge 1$, just by choosing $P_{\mathsf{n(m)}} = \beta\ell(d_0)^{-1}$ always, we get outage probability $p_o = \mathbb{P}(\mathsf{SINR} \le \beta|\Phi) = 0$, while satisfying the average power constraint of M.

Case 2: $\ell(d_0) < \frac{\beta}{M}$. To derive a lower bound in this case, we replace $h_{\mathsf{n(m)}} \sim EXP(1)$ with 'stronger' $h'_{\mathsf{n(m)}}$ that has PDF $f_{h'_{\mathsf{n(m)}}}(x) = x\exp(-x)$ and CDF $\mathbb{P}(h'_{\mathsf{n(m)}} < x) = 1 - (x + 1)\exp(-x)$. With $h_{\mathsf{n(m)}} \sim EXP(1)$, the CDF is $\mathbb{P}(h_{\mathsf{n(m)}} < x) = 1 - \exp(-x)$. Stronger h' implies lower outage probability, allowing us to lower bound the expected delay.

With the stronger fading gain $h'_{\mathsf{n(m)}}$, conditioned on d_0, from Proposition 4, with an average power constraint of M, the outage probability $p_o = \mathbb{P}\left(\frac{P_{\mathsf{n(m)}} h'_{\mathsf{n(m)}} \mathbf{1}_{\mathsf{n(m)}}(t)\ell(d_0)}{\mathsf{N}} \le \beta|\Phi\right)$ is minimized when $P_{\mathsf{n(m)}} = \frac{\beta}{\ell(d_0)h'_{\mathsf{n(m)}}}$ for $h'_{\mathsf{n(m)}} > \delta$ and $P_{\mathsf{n(m)}} = 0$ otherwise, where δ is such that the average power constraint $\int_\delta^\infty \frac{\beta}{\ell(d_0)x} x\exp(-x)dx = M$ is satisfied. Thus, we get that $\exp(-\delta) = \frac{M\ell(d_0)}{\beta}$, where recall that $\frac{M\ell(d_0)}{\beta} < 1$. Hence, the resulting outage probability, $p_o = \mathbb{P}(h'_{\mathsf{n(m)}} \le \delta) = 1 - \left(1 - \ln\frac{M\ell(d_0)}{\beta}\right)\frac{M\ell(d_0)}{\beta}$, and the success probability $p_s = 1 - p_o = \left(1 - \ln\frac{M\ell(d_0)}{\beta}\right)\frac{M\ell(d_0)}{\beta}$.

Thus, combining case 1 and 2, given Φ, the expected delay is $\mathbb{E}\{D|\Phi\} \ge \mathbf{1}_{\frac{M\ell(d_0)}{\beta} \ge 1} 1 + \mathbf{1}_{\frac{M\ell(d_0)}{\beta} < 1} \frac{1}{p_s}$, since given Φ, the success events $(h'_{\mathsf{n(m)}} > \delta)$ are independent across time slots. Thus,

$$\mathbb{E}\{D|\Phi\} \ge \mathbf{1}_{\frac{M\ell(d_0)}{\beta} < 1} \frac{\beta\ell(d_0)^{-1}}{\left(1 - \ln\frac{M\ell(d_0)}{\beta}\right)M}. \tag{9}$$

From Proposition 5, we know that $f_{d_0}(y) = 2\pi\lambda y \exp\left(-\lambda\pi y^2\right)$. Recall that $\ell(d_0) = \min\{1, d_0^{-\alpha}\}$, taking the expectation of (9) with respect to d_0, we get $\mathbb{E}\{D\}$

$$\begin{aligned} &\ge \frac{\beta}{M}\left(\int_{\ell(y) < \frac{\beta}{M}} \frac{\ell(y)^{-1}}{\left(1 - \ln\frac{M\ell(y)}{\beta}\right)} 2\pi\lambda y \exp\left(-\lambda\pi y^2\right) dy\right), \\ &\ge \frac{\beta}{M}\frac{c}{(\pi\lambda)^{\frac{\alpha}{2}}}, \end{aligned} \tag{10}$$

where c is a constant. This proves (5).

Proposition 5. *The cumulative distribution function and probability distribution function of nearest BS distance d_0 is*

$$\mathbb{P}(d_0 > y) = \exp(-\lambda\pi y^2), \implies f_{d_0}(y) = 2\lambda\pi y \exp(-\lambda\pi y^2).$$

To obtain the second lower bound on the expected delay (6), we set $\ell(d_0) = 1$ for the typical user m, to maximize the signal power in terms of path-loss, since $\ell(.) \le 1$. This will allow us to remove the dependence of $(p_{\mathsf{n(m)}}(t), P_{\mathsf{n(m)}}(t))$ on d_0. Note that we are not putting any restriction on d_0, since that would impact the interference term in the SINR expression.

Let $\mathcal{U}_z$ be the set of MUs connected to the BS z. As before, we consider the practical setting where the MU density is much larger than the BS density, and $|\mathcal{U}_z| \ge 1$ for all BSs z. [3] For BS $z \in \Phi\backslash\{\mathsf{n(m)}\}$ other than $\mathsf{n(m)}$, we let all the MUs

[3]Otherwise, since MU and BS processes are independent, we get a thinned BS process, for which the derived results apply directly.

$u \in \mathcal{U}_z$ connected to BS z be at most a unit distance away from z, i.e., $d_{zu} \leq 1, \forall\, z, u \in \mathcal{U}_z$, which implies $\ell(d_{zu}) = 1$. This minimizes the path-loss between z and $u \in \mathcal{U}_z$, and gives the largest gain $\ell(d_{zu}) = 1$ for each BS-MU pair (z, u) for $u \in \mathcal{U}_z$. Moreover, we also let that that each $u \in \mathcal{U}_z$ receives no interference from any basestation. Let this be called **restriction** 1 that applies to all non-typical BSs $z \neq \mathsf{n(m)}$ and its connected users in $\mathcal{U}_z$. Thus, the SINR seen at any MU $u \in \mathcal{U}_z, z \neq \mathsf{n(m)}$ under restriction 1 is $\mathsf{SINR}_{z,u}(t) = \frac{P_z(t)h_{zu}(t)\mathbf{1}_z(t)}{\mathsf{N}}$.

As a function of time, let $\tilde{p}_z(t)$ and $\tilde{P}_z(t)$ (where $\tilde{p}_z(t)\tilde{P}_z(t) \in [M/\tau, M]\}$ from strategy Definition 2), be the power profile used by BS $z \in \Phi\backslash\{\mathsf{n(m)}\}$ under restriction 1 to get the same expected delay/capacity without restriction 1 while using the optimal (unknown) power profile $p_z(t)$ and $P_z(t)$. Clearly, since there is no path-loss and no interference seen at any $u \in \mathcal{U}_z$ with restriction 1, power profile $\tilde{p}_z(t)$ and $\tilde{P}_z(t)$ is stochastically dominated by $p_z(t)$ and $P_z(t)$. Equivalently, the interference seen at the typical MU m from BS $z \neq \mathsf{n(m)}$ is stochastically dominated with restriction 1.[4]

Recall that the SINR (3) seen at the typical user m critically depends on the power profile ($p_z(t)$ and $P_z(t)$) of BS $z \in \Phi\backslash\{\mathsf{n(m)}\}$. Thus, using the stochastically dominated power profile $\tilde{p}_z(t)$ and $\tilde{P}_z(t)$ with restriction 1 for each BS $z \neq \mathsf{n(m)}$ cannot increase the delay at m. Thus, we work under restriction 1 to lower bound the expected delay seen at m.

It is important to note that under restriction 1, power profile $\tilde{\mathbf{1}}_z(t)$ and $\tilde{P}_z(t)$ only depend on $h_{zu}(t)$. Since $h_{zu}(t)$'s are independent for each BS-MU pair (z, u) $u \in \mathcal{U}_z$ and across time t, restriction 1 helps simplify the ensuing analysis significantly.

We also also neglect additive noise in the SINR expression (3) for deriving the lower bound on the expected delay.

Under this setup ($\ell(d_0) = 1$, no noise, and restriction 1 for BSs $z \neq \mathsf{n(m)}$), from (3) for the $\mathsf{n(m)} - \mathsf{m}$ link

$$\tilde{\mathsf{SINR}}_{\mathsf{n(m)},\mathsf{m}}(t) = \frac{P_{\mathsf{n(m)}}(t)h_{\mathsf{n(m)}}(t)\mathbf{1}_{\mathsf{n(m)}}(t)}{\gamma\sum_{z\in\Phi\backslash\{\mathsf{n(m)}\}}\tilde{P}_z(t)\tilde{\mathbf{1}}_z(t)h_z(t)\ell(z)}. \quad (11)$$

Note that we have kept the path-loss $\ell(z)$ from BS z to the typical MU m as it is with restriction 1. Thus, we have the tail probability $\mathbb{P}(D > n|\Phi, \Phi_R)$

$$\begin{aligned} &= \mathbb{P}\left(\tilde{\mathsf{SINR}}_{\mathsf{m}}(1) < \beta, \ldots, \tilde{\mathsf{SINR}}_{\mathsf{m}}(n) < \beta \middle| \Phi, \Phi_R\right), \\ &= \prod_{t=1}^{n} \mathbb{P}\left(\tilde{\mathsf{SINR}}_{\mathsf{m}}(t) < \beta \middle| \Phi, \Phi_R\right), \end{aligned} \quad (12)$$

where the second equality follows since given Φ, Φ_R, SINRs are independent

Under restriction 1, both $\tilde{P}_z(t)$ and $\tilde{\mathbf{1}}_z(t)$ are independent across time, and are unknown at m. Moreover, their distributions are also unknown. Thus, to derive a lower bound on $\mathbb{P}\left(\tilde{\mathsf{SINR}}_{\mathsf{m}}(t) < \beta \middle| \Phi, \Phi_R\right)$, we let BS $\mathsf{n(m)}$ choose the optimal transmission policy to minimize $\mathbb{P}\left(P_{\mathsf{n(m)}}(t)h_{\mathsf{n(m)}}(t)\mathbf{1}_{\mathsf{n(m)}}(t) < \gamma\beta\right)$ from Proposition 4 that only depends on $h_{\mathsf{n(m)}}(t)$ and γ, β. Thus, the outage probability $\mathbb{P}\left(\tilde{\mathsf{SINR}}_{\mathsf{m}}(t) < \beta|\Phi, \Phi_R\right)$

$$\begin{aligned} &= \mathbb{P}\left(h_{\mathsf{n(m)}}(t) < \delta\right)\mathbb{P}\left(\frac{0}{\tilde{I}(t)} < \gamma\beta \middle| \Phi, \Phi_R\right) \\ &\quad + \mathbb{P}\left(h_{\mathsf{n(m)}}(t) \geq \delta\right)\mathbb{P}\left(\frac{\gamma\beta}{\tilde{I}(t)} < \gamma\beta \middle| \Phi, \Phi_R\right), \\ &\geq \mathbb{P}\left(h_{\mathsf{n(m)}}(t) \geq \delta\right)\mathbb{P}\left(\tilde{I}(t) > 1 \middle| \Phi, \Phi_R\right), \\ &= \exp(-\delta)\mathbb{P}\left(\tilde{I}(t) > 1 \middle| \Phi, \Phi_R\right), \end{aligned} \quad (13)$$

where the last inequality follows since $h_{\mathsf{n(m)}}(t) \sim EXP(1)$. Now we derive bounds on $\mathbb{P}\left(\tilde{I}(t) > 1 \middle| \Phi, \Phi_R\right)$ to derive a lower bound on the expected delay. Recall that $\mathbb{P}\left(\tilde{I}(t) > 1 \middle| \Phi, \Phi_R\right)$

$$\begin{aligned} &= \mathbb{P}\left(\sum_{z\in\Phi\backslash\{\mathsf{n(m)}\}} \tilde{P}_z(t)\tilde{\mathbf{1}}_z(t)h_z(t)\ell(z) > 1 \middle| \Phi, \Phi_R\right), \\ &\geq 1 - \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}} \mathbb{P}\left(\tilde{P}_z(t)\tilde{\mathbf{1}}_z(t)h_z(t)\ell(z) < 1 \middle| \Phi, \Phi_R\right), \\ &= 1 - \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}} \mathbb{E}\left\{\left(1 - \exp\left(\frac{-1}{\tilde{P}_z(t)\tilde{\mathbf{1}}_z(t)\ell(z)}\right)\right)|\Phi, \Phi_R\right\}, \end{aligned}$$

where the second equality follows since $h_z(t)$'s are independent for z, t, and $\tilde{P}_z(t), \tilde{\mathbf{1}}_z(t)$ are independent for z under restriction 1, and the last inequality follows by taking the expectation with respect to $h_z(t) \sim EXP(1)$.

Taking the expectation with respect to $\tilde{\mathbf{1}}_z(t)$, we get $\mathbb{P}\left(\tilde{I}(t) > 1 \middle| \Phi, \Phi_R\right)$

$$\begin{aligned} &= 1 - \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}} \mathbb{E}\left\{\left(1 - \tilde{p}_z(t)\exp\left(\frac{-1}{\tilde{P}_z(t)\ell(z)}\right)\right)|\Phi, \Phi_R\right\}, \\ &\geq 1 - \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}} \mathbb{E}\left\{\left(1 - \tilde{p}_z(t)\exp\left(\frac{-1}{\frac{M/\tau}{\tilde{p}_z(t)}\ell(z)}\right)\right)|\Phi, \Phi_R\right\}, \\ &\geq 1 - \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}} \left(1 - \mathbb{E}\left\{\tilde{p}_z(t)|\Phi, \Phi_R\right\}\exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right), \end{aligned}$$

where the second inequality follows since $\tilde{p}_z(t)\tilde{P}_z(t) \geq M/\tau$ for each BS z, and the final inequality follows since $\tilde{p}_z(t) \leq 1$, and the fact that $\tilde{p}_z(t)$ are independent for different BSs z.

Remark 1. *Recall that the BS density is much smaller than MU density, hence $|\mathcal{U}_z| >> 1$, i.e., each BS z is transmitting to multiple MUs in its Voronoi cell. Thus, for $\tilde{p}_z(t)$ (that is independent across z and t) that only depends on fading gains from z to $\mathcal{U}_z$, we let $\mathbb{E}\{\tilde{p}_z(t)|\Phi, \Phi_R\} \geq \eta > 0$ (where η is a constant).*

[4]For two random variables X and Y, X is defined to be stochastically dominated by Y if $\mathbb{P}(X \leq x) \geq \mathbb{P}(Y \leq x)$ for all x. Essentially, what we are doing is that let $S^\star = (p(t), P(t))$ be the (unknown) optimal strategy (power profile) that achieves minimum delay $\mathbb{E}\{D^\star\}$ for each BS z. Then, under restriction 1, the power profile needed $\tilde{S}_z = (\tilde{p}_z(t), \tilde{P}_z(t))$ for all BSs $z \neq \mathsf{n(m)}$ to achieve $\mathbb{E}\{D^\star\}$ at their respective MUs, is stochastically dominated by $S^\star$. Thus, with each BS using $\tilde{S}_z$, the expected delay at the typical MU m is at most $\mathbb{E}\{D^\star\}$.

Thus, it follows that $\mathbb{P}\left(\tilde{I}(t) > 1 \middle| \Phi, \Phi_R\right)$

$$\geq 1 - \prod_{z \in \Phi \backslash \{\mathsf{n(m)}\}} \left(1 - \eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right). \quad (14)$$

Substituting (14) into (13), we get $\mathbb{P}\left(\tilde{\mathsf{SINR}}_{\mathsf{m}}(t) < \beta | \Phi, \Phi_R\right) \geq \exp(-\delta)\left(1 - \prod_{z \in \Phi \backslash \{\mathsf{n(m)}\}} \left(1 - \eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right)\right)$, which from (12), gives $\mathbb{P}(D > n | \Phi, \Phi_R)$

$$\geq \prod_{t=1}^{n} \exp(-\delta)\left(1 - \prod_{z \in \Phi \backslash \{\mathsf{n(m)}\}} \left(1 - \eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right)\right) \quad (15)$$

Let $g = \prod_{z \in \Phi \backslash \{\mathsf{n(m)}\}} \left(1 - \eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right)$. Then

$$\begin{aligned} \mathbb{E}\{D|\Phi, \Phi_R\} &= \sum_{n=0}^{\infty} \mathbb{P}(D > n | \Phi, \Phi_R), \\ &\geq \sum_{n=0}^{\infty} (\exp(-\delta)(1-g))^n, \text{ from (15)} \\ &= \frac{1}{1 - \exp(-\delta)(1-g)} \approx \frac{1}{\exp(-\delta) g}, \end{aligned}$$

where the approximation is tight for large average power constraints M for which δ is small as evident from Proposition 4.[5] Thus, using the definition of g,

$$\mathbb{E}\{D|\Phi, \Phi_R\} \geq \exp(\delta) \prod_{z \in \Phi \backslash \{\mathsf{n(m)}\}} \frac{1}{\left(1 - \eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right)}.$$

To find $\mathbb{E}\{D\}$ from $\mathbb{E}\{D|\Phi, \Phi_R\}$, we first use the probability generating functional (Proposition 6) for the PPP $\Phi \backslash \{\mathsf{n(m)}\}$ to get $\mathbb{E}\{D|d_0\}$

$$\geq \exp(\delta) \exp\left(2\pi\lambda \int_{\mathbb{R}, z > d_0} \frac{\eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)}{\left(1 - \eta \exp\left(\frac{-1}{(M/\tau)\ell(z)}\right)\right)} z \, dz\right).$$

Let $c_1 = \frac{\eta \exp\left(\frac{-1}{(M/\tau)}\right)}{\left(1 - \eta \exp\left(\frac{-1}{(M/\tau)}\right)\right)}$. Splitting the integral into two parts for $d_0 \leq 1$ and $d_0 > 1$, and keeping only the part with $d_0 \leq 1$ which implies $(z < d_0 < 1)$ for which $\ell(z) = 1$, we have

$$\begin{aligned} \mathbb{E}\{D|d_0\} &= \exp(\delta) \exp\left(2\pi\lambda c_1 \int_{d_0}^{1} z \, dz\right), \\ &\geq \exp(\delta) \exp\left(2\pi\lambda c_1 \left(\frac{1 - d_0^2}{2}\right)\right). \end{aligned}$$

Taking the expectation with respect to the nearest BS distance d_0, whose PDF is given by Proposition 5, we get $\mathbb{E}\{D\}$

$$\geq \exp(\delta) \frac{\exp(\pi\lambda c_1)}{1 + c_1} \left(1 - \exp(-\pi\lambda(c_1 + 1))\right). \quad (16)$$

Thus, (10) together with (16) gives the required lower bound on the expected delay stated in Theorem 1.

[5]Note that even for $M = 1$, $\delta \approx .267$. This approximation can be avoided, and exact expression can be derived but results in more tedious computations as will be evident next.

Proposition 6. *For a homogenous PPP Φ with density λ, the probability generating functional is given by* $\mathbb{E}\left\{\prod_{x_n \in \Phi} f(x_n)\right\} = \exp^{-\int (1 - f(x))\lambda dx}$.

APPENDIX B
PROOF OF THEOREM 2

Without loss of generality, we will derive the upper bound on the expected delay for the typical user m that is served by its nearest BS $\mathsf{n(m)}$ at a distance of d_0 from it. We consider k successive slots (not necessarily consecutive) that are dedicated for transmission to m by BS $\mathsf{n(m)}$, and are interested in probability $\mathbb{P}(D > k)$ to upper bound the expected delay, where delay D is as defined in Definition 1.

Typically, the number of MUs connected to different BSs are different. Thus, during the k considered slots for m, any BS other than $\mathsf{n(m)}$ transmits potentially to different MUs. Let $\mathcal{G}_k$ be the sigma field generated by the BS point process Φ and MU point process Φ_R and the choice (index) of MUs being served by BSs of Φ at the above described k slots $t = 1, 2, \ldots, k$.

With the power control strategy (7), the SINR seen at m in time slot t is

$$\mathsf{SINR}_{n(\mathsf{m}),\mathsf{m}}(t) = \frac{c h_{\mathsf{n(m)}}(t) \mathbf{1}_{\mathsf{n(m)}}(t)}{\gamma I(t) + \mathsf{N}}, \quad (17)$$

where $I(t) = \sum_{z \in \Phi \backslash \{\mathsf{n(m)}\}} \mathbf{1}_z(t) P_z(t) h_z(t) \ell(z)$.

With $e_{\mathsf{n(m)},\mathsf{m}}(t) = 1$ if $\mathsf{SINR}_{(\mathsf{m},\mathsf{n(m)})}(t) > \beta$, and 0 otherwise, we have $\mathbb{P}\left[D > k \middle| \mathcal{G}_k\right] = \mathbb{E}\left\{\mathbb{P}\left[e_{\mathsf{n(m)},\mathsf{m}}(t) = 0, \ \forall \, t = 1, \ldots, k \middle| \mathcal{G}_k\right]\right\}$.

Given $\mathcal{G}_k$, with the described strategy (7), the transmission events $\mathbf{1}_z(t)$, and the transmit powers $P_z(t)$ are independent across time slots t for all BSs z. Moreover, the fading gains $h_{(.)}(t)$ are all independent. Hence, we get $\mathbb{P}\left[D > k \middle| \mathcal{G}_k\right] = \mathbb{E}\left\{\prod_{t=1}^{k} \mathbb{P}\left[e_{\mathsf{n(m)},\mathsf{m}}(t) = 0 \middle| \mathcal{G}_k\right]\right\}$.

Let $A(t)$ be the event that $\mathbf{1}_{\mathsf{n(m)}}(t) = 0$, i.e., the BS $\mathsf{n(m)}$ does not transmit in the slot designated for user m, while $B(t)$ be the event that $\{\mathbf{1}_{\mathsf{n(m)}}(t) = 1\} \cap e_{\mathsf{n(m)},\mathsf{m}}(t) = 0$, i.e., BS $\mathsf{n(m)}$ transmits in slot t but the transmission fails. Then

$$\mathbb{P}\left[D > k \middle| \mathcal{G}_k\right] = \mathbb{E}\left\{\prod_{t=1}^{k} \mathbb{P}\left[A(t) \cup B(t) \middle| \mathcal{G}_k\right]\right\}. \quad (18)$$

With strategy (7), given Φ, transmission event $\mathbf{1}_{\mathsf{n(m)}}(t)$ and the success event $e_{\mathsf{n(m)},\mathsf{m}}(t)$ are independent, hence $\mathbb{P}(A(t)|\mathcal{G}_k) = 1 - p_{\mathsf{n(m)}}(t)$, while $\mathbb{P}(B(t)|\mathcal{G}_k)$

$$= p_{\mathsf{n(m)}}(t)\left(1 - \mathbb{E}\left\{\exp\left(-\frac{\beta}{c}(\mathsf{N} + \gamma I(t))\right) \middle| \mathcal{G}_k\right\}\right), \quad (19)$$

that follows by taking expectation with respect to $h_{\mathsf{n(m)}}(t) \sim EXP(1)$. Using the union bound, from (18), we get $\mathbb{P}\left[A(t) \cup B(t) \middle| \mathcal{G}_k\right] \leq 1 - p_{\mathsf{n(m)}}(t) + p_{\mathsf{n(m)}}(t)$

$$\begin{aligned} &\left(1 - \mathbb{E}\left\{\exp\left(-\frac{\beta}{c}(\mathsf{N} + \gamma I(t))\right) \middle| \mathcal{G}_k\right\}\right) \\ &\leq 1 - p_{\mathsf{n(m)}}(t) \exp\left(-\frac{\beta \mathsf{N}}{c}\right) \mathbb{E}\left\{\exp\left(-\frac{\beta\gamma}{c} I(t)\right) | \mathcal{G}_k\right\}. \end{aligned} \quad (20)$$

Let $a = \frac{\beta\gamma}{c}$, and we focus on finding a lower bound on $\mathbb{E}\{\exp(-aI(t))|\mathcal{G}_k\}$, that is independent of the choice of the MU being served by BS z. To this end, we first expand $\mathbb{E}\left\{\exp(-aI(t))\middle|\mathcal{G}_k\right\} = \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}}\mathbb{E}\left\{\exp\left(-a\mathbf{1}_z(t)P_z^{(u(z))}(t)h_z(t)\ell(z)\right)\middle|\mathcal{G}_k\right\}$, where BS $z \in \Phi\backslash\{\mathsf{n(m)}\}$ transmits to the MU $u(z)$ in time slot t. This fixes the transmission probability $p_z^{(u(z))}(t)$ and power $P_z^{(u(z))}(t)$ (where we have included the index u to make the dependence on the MU explicit). Then, taking the expectation with respect to $\mathbf{1}_z(t)$, we have

$$\begin{aligned}&\mathbb{E}\left\{e^{-a\mathbf{1}_z(t)P_z^{(u(z))}(t)h_z(t)\ell(z)}\middle|\mathcal{G}_k\right\}\\ =&(1-p_z^{(u(z))}(t))\\ &+p_z^{(u(z))}(t)\mathbb{E}\left\{\exp\left(-aP_z^{(u(z))}(t)h_z(t)\ell(z)\right)\middle|\mathcal{G}_k\right\}\\ =&(1-p_z^{(u(z))}(t))+p_z^{(u(z))}(t)\frac{1}{1+a\gamma\ell(z)P_z^{(u(z))}(t)},\end{aligned} \quad (21)$$

where the second equality follows by taking expectation with respect to the independent fading gains $h_z(t) \sim EXP(1)$.

Let $u^*(z)$ be the MU for which the right hand expression in (21) is minimized, i.e., BS z causes maximum interference at m when it is serving MU $u^*(z)$. Let p_z^*, P_z^* denote the corresponding transmission probability and power, respectively for BS z. Denote by $\mathbf{1}_z^*$ an independent Bernoulli random variable with $\mathbb{P}[\mathbf{1}_z^* = 1] = p_z^*$. Define $I^*(t) = \sum_{z\in\Phi\backslash\{\mathsf{n(m)}\}}\mathbf{1}_z^*P_z^*h_z(t)\ell(z)$. Essentially $I^*(t)$ dominates the actual interference $I(t)$ seen at m. Substituting $I^*(t)$ for I in (20) along with the observation that given Φ, Φ_R, $I^*(t) \stackrel{d}{=} I^*(1)$, we get $\mathbb{P}\left[A(t)\cup B(t)\middle|\mathcal{G}_k\right] \leq 1 - p_{\mathsf{n(m)}}(t)\exp\left(-\frac{\beta\mathsf{N}}{c}\right)\mathbb{E}\left\{\exp(-aI^*(1))\middle|\Phi,\Phi_R\right\}$. Let $\theta = \exp\left(-\frac{\beta\mathsf{N}}{c}\right)p_{\mathsf{n(m)}}(t)\,\mathbb{E}\left\{\exp(-aI^*(1))\middle|\Phi,\Phi_R\right\}$. Thus, from (18), we get

$$\mathbb{P}\left[D > k\middle|\Phi,\ \Phi_R\right] \leq (1-\theta)^k. \quad (22)$$

Then the expected delay can be written as

$$\mathbb{E}\{D\} = \sum_{k\geq 0}\mathbb{P}[D>k] = \mathbb{E}\left\{\sum_{k\geq 0}\mathbb{P}\left[D>k\middle|\Phi,\Phi_R\right]\right\} \leq \mathbb{E}\{\theta^{-1}\},$$

where the last inequality follows from (22). Using the Cauchy-Schwartz inequality, on the two random variables in θ^{-1},

$$\begin{aligned}\mathbb{E}\{D\} \quad \leq \quad & \exp(\beta\mathsf{N}/c)\left(\mathbb{E}\left\{\frac{1}{\left(\mathbb{E}\left\{\exp(-aI^*(1))\middle|\Phi,\Phi_R\right\}\right)^2}\right\}\right.\\ & \left.\mathbb{E}\left\{p_{\mathsf{n(m)}}(t)^{-2}\right\}\right)^{\frac{1}{2}}.\end{aligned} \quad (23)$$

From the definition of the transmission probability $p_{\mathsf{n(m)}}(t) = (M/c)\,\ell(d_0) = (M/c)\min\{1, d_0^{-\alpha}\}$, we get $\mathbb{E}[p_{\mathsf{n(m)}}(t)^{-2}]$

$$\begin{aligned}&= \left(\frac{c}{M}\right)^2\left(\int_0^1 1f_{d_0}(x)dx + \int_1^\infty d_0^{2\alpha}f_{d_0}(x)dx\right),\\ &\leq \left(\frac{c}{M}\right)^2\left(1 + \int_0^\infty d_0^{2\alpha}f_{d_0}(x)dx\right),\\ &= \left(\frac{c}{M}\right)^2\left(1 + \frac{\Gamma(\alpha+1)}{(\pi\lambda)^\alpha}\right),\end{aligned} \quad (24)$$

using the PDF of d_0 from Proposition 5.

Recall, $\mathbb{E}\left\{\exp(-aI^*(1))\middle|\Phi,\Phi_R\right\}$

$$= \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}}\mathbb{E}\left\{\exp(-a\mathbf{1}_z^*P_z^*h_z(1)\ell(z))\middle|\Phi,\Phi_R\right\}. \quad (25)$$

Taking expectations, first with respect to $\mathbf{1}_z^*$ and then with respect to fading gain $h_z(1)$, we get $\mathbb{E}\{\exp(-a\mathbf{1}_z^*P_z^*h_z(1)\ell(z))|\Phi,\Phi_R\}$

$$\begin{aligned}&= (1-p_z^*) + p_z^*\mathbb{E}\left\{e^{-aP_z^*h_z(1)\ell(z)}\middle|\Phi,\Phi_R\right\},\\ &\stackrel{(a)}{=} 1 - \frac{\beta\gamma p_z^*P_z^*\ell(z)}{c+\beta\gamma P_z^*\ell(z)},\\ &\stackrel{(b)}{\geq} 1 - \frac{\beta\gamma M\ell(z)}{c} = 1 - \beta\gamma(1-\epsilon)\ell(z),\end{aligned}$$

where (a) follows by resubstituting $a = \frac{\beta\gamma}{c}$, and (b) by invoking the average power constraint $p_zP_z \leq M$ for $\forall\ z$ and in particular $p_z^*P_z^* \leq M$ and $c = M(1-\epsilon)^{-1}$. Let $c_2 = \beta\gamma(1-\epsilon)$, where note that because of assumption on ϵ (ϵ satisfies $(1-\epsilon)\beta\gamma < 1$) in the power control strategy (7), $c_2 < 1$. Substituting the above bound in (25), we get $\mathbb{E}\left\{e^{-aI^*(1)}\middle|\Phi,\Phi_R\right\} \geq \prod_{z\in\Phi\backslash\{\mathsf{n(m)}\}}(1-c_2\ell(z))$.

Hence $\mathbb{E}\left\{\frac{1}{(\mathbb{E}\{\exp(-aI^*(1))|\Phi,\Phi_R\})^2}\right\}$

$$\leq \mathbb{E}\left\{\prod_{z\in\Phi\backslash\{\mathsf{n(m)}\})}\exp(-2\log(1-c_2\ell(z)))\right\}. \quad (26)$$

Once again using the probability generating functional (Proposition 6) for the PPP $\Phi\backslash\{\mathsf{n(m)}\}$, we get $\mathbb{E}\left\{\frac{1}{(\mathbb{E}\{\exp(-aI^*(1))|\Phi\})^2}\right\}$

$$\begin{aligned}\leq\ & \mathbb{E}_{d_0}\left\{\exp\left(\lambda\int_{\mathbb{R}^2\backslash\mathbf{B}(0,d_0)}(\exp(-2\log(1-c_2\ell(z)))\right.\right.\\ & \left.\left.-1)\,\mathrm{d}z\right)\right\},\\ \leq\ & \mathbb{E}_{d_0}\left\{\exp\left(\frac{2\lambda c_2}{(1-c_2)^2}\int_{\mathbb{R},z>d_0}z\ell(z)\mathrm{d}z\right)\right\},\end{aligned} \quad (27)$$

where $\mathbf{B}(0,d_0)$ is the disc with radius d_0 centered at the origin, and the last inequality follows by noting that $\ell(z) \leq 1$. With $\ell(d) = \min\{1, d^{-\alpha}\}$, we get $\mathbb{E}\left\{\frac{1}{(\mathbb{E}\{\exp(-aI^*(1))|\Phi\})^2}\right\}$

$$\begin{aligned}\leq\ & \int_0^1\exp\left(\frac{2\lambda c_2}{(1-c_2)^2}\left(\frac{1-x^2}{2}\right)\right)f_{d_0}(x)dx\\ & + \int_1^\infty\exp\left(\frac{2\lambda c_2}{(1-c_2)^2}\left(\frac{x^{2-\alpha}}{\alpha-2}\right)\right)f_{d_0}(x)dx,\end{aligned}$$

Since $\alpha > 2$, using Proposition 5, we get the following bound on the expectation $\mathbb{E}\left\{\frac{1}{(\mathbb{E}\{\exp(-aI^*(1))|\Phi\})^2}\right\} \leq \exp(c_3\lambda)$, where c_3 is a constant. Combining this with (24), from (23) we get

$$\mathbb{E}\{D\} \leq \sqrt{c_4\left(1+\frac{\Gamma(\alpha+1)}{(\pi\lambda)^\alpha}\right)\exp(c_3\lambda)},$$

where c_4 is a constant. This completes the proof of Theorem 2.