This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

# Transition from Heavy to Light Tails in Retransmission Durations

Jian Tan Ness B. Shroff
Department of Electrical and Computer Engineering
The Ohio State University, Columbus, 43210

***Abstract*— Retransmissions serve as the basic building block that communication protocols use to achieve reliable data transfer. Until recently, the number of retransmissions were thought to follow a light tailed (in particular, a geometric) distribution. However, recent work seems to suggest that when the distribution of the packets have infinite support, retransmission-based protocols may result in heavy tailed delays and even possibly zero throughput. While this result is true even when the distribution of packet sizes are light-tailed, it requires the assumption that the packet sizes have infinite support. However, in reality, packet sizes are often bounded by the Maximum Transmission Unit (MTU), and thus the aforementioned result merits a deeper investigation.**

**To that end, in this paper, we allow the distribution of the packet size $L$ to have finite support. This packet is sent over an on-off channel $\{(A_i, U_i)\}$ with alternating available $A_i$ and unavailable $U_i$ periods. If $L \geq A_i$, the transmission fails and we wait for the next period $A_{i+1}$ to retransmit the packet. The transmission duration is thus measured from the first attempt to a point when a channel available period larger than $L$.**

**Under mild conditions, we show that the transmission duration distribution exhibits a transition from a power law main body to an exponential tail with Weibull type distributions between the two. The time scale to observe the power law main body is roughly equal to the average transmission duration of the longest packet. Both the power law main body and the exponential tail could dominate the overall performance. For example, the power law main body, if significant, may cause the channel throughput to be very close to zero. On the other hand, the exponential tail, if more evident, may imply that the system operates in a benign environment. These theoretical findings provide an understanding on why some empirical measurements suggest heavy tails and light tails for others (e.g., wireless networks). We use these results to further highlight the engineering implications from distributions with power law main bodies and light tails by analyzing two cases: (1) The throughput of on-off channels with retransmissions, where we show that even when packet sizes have small means and bounded support the variability in their sizes can greatly impact system performance. (2) The distribution of the number of jobs in an $M/M/\infty$ queue with server failures. Here we show that retransmissions can cause long-range dependence and quantify the impact of the maximum job sizes on the long-range dependence.**

## I. Introduction

Retransmissions are fundamental in ensuring reliable data transfer over communication networks with channel errors. Traditionally, retransmissions were assumed to result in light-tailed (rapidly decaying tail distribution) transmission delays. The conventional belief was that the number of retransmissions follows a geometric distribution [1], which is true when the errors are independent of the size of the transmitting packet. However, recent work [2]–[6] shows that when the probability of packet errors is a function of the packet length, which is often true in communication networks, the number of retransmissions do not follow a geometric distribution. In fact, it has been shown, under the assumption that the packet size distribution has infinite support, that all retransmission-based protocols could cause heavy-tailed behavior (specifically, power law transmission durations) and possibly even zero throughput. These heavy tails can result entirely from retransmissions, even when the data units and channel characteristics are light-tailed. However, in practice, packet sizes are bounded by the maximum transmission unit (MTU). For example, WaveLAN's maximum transfer unit is 1500 bytes. This fact motivates us to more carefully investigate the impact that retransmissions have on network performance by allowing the packet sizes to have finite support.

We consider a system where the channel dynamics are modeled by an *on-off* process $\{(A_i, U_i)\}_{i\geq 1}$ where $A_i$ corresponds to the time when the channel is available and $U_i$ the time period when the channel is not available, as in [2]. Let $L$ be the random variable that denotes the length of a generic packet. At the beginning of each available period $A_i$, we attempt to transmit the packet. If $L < A_i$, we say that the transmission is successful; otherwise, we wait until the beginning of the next available period $A_{i+1}$ and retransmit the packet from the beginning. As mentioned earlier, we focus on the situation of practical interest, i.e., when the distribution of $L$ has finite support on the interval $[0, b]$. We study the asymptotic properties of the distributions of the total transmission time and number of retransmissions. Our main contributions in this paper can be summarized as follows:

(I) Under the condition that there is a polynomial relationship between the packet size distribution and channel available periods (see Theorem 1), we show that, even when the packet size has an upper limit, the transmission duration distribution is characterized by a power law main body. This power law behavior spans over a time scale that is approximately equal to the average transmission duration of the longest packet. Additionally, we show that this distribution eventually becomes light-tailed. We characterize the transition of the transmission delay distribution from a power law main body to an exponential tail with Weibull type distributions between the two under the aforementioned conditions. Thus, depending on the probability of interests and the sys-

978-1-4244-5837-0/10/$26.00 ©2010 IEEE

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

tem parameters, the transmission delays may experience heavy or light-tailed distributions. More importantly, both the power law main body and the exponential tail could dominate performance. When this power law main body is significant, it could possibly cause the channel throughput to be very close to zero (as shown in Theorem 4), implying that some careful re-examination and adjustment of system parameters are needed. On the other hand, if the exponential tail is more evident, this often suggests that the system is operating in a benign environment. Similar phenomenon of power law up to a certain threshold followed by an exponential decay has been observed for inter-contact time distributions between mobile devices [7].

(II) Using the afore-mentioned results, we study two cases of interest. First, we investigate the system throughput when the packet lengths have an upper limit $b$. Our results show that under certain conditions the channel throughput may be very close to zero for large $b$ even when the average packet size is very small. Next, we study an $M/M/\infty$ queue with server failures. When active servers fail according to i.i.d. Poisson point processes, we observe that the number of jobs in the system exhibits long-range dependence. This effect can be eliminated if job sizes are upper bounded. However, we find that there may still be a strong autocorrelation for the number of jobs in the system that spans over a large time interval for bounded job sizes, implying that the system may exhibit long-range dependence over operating regions of interest.

These theoretical findings provide a new understanding on the controversy in empirical measurements why heavy tails are observed for certain measurements and light tails for others (e.g., wireless networks). Since the first discoveries in the early 1990s [8] of the presence of heavy-tailed statistical characteristics of traffic streams in modern computer networks, there has been a large amount of research on the issue of power laws in information networks. For example, it was suggested in [9] that the transmission delay distribution in IEEE 802.11 wireless ad hoc networks can be expressed as a power law. There are also different views advocating other distributions as providing the correct description of the system. For example, the empirical measurement in [10] shows that link delays over the wireless mesh network are fitted by either gamma or logistic distributions. These seemingly contradicting results have been addressed in [11], which suggests that (1) some claims on the heavy/light tails may not be legitimate due to the lack of sufficient measurements for the hypothesis testing, and (2) engineers should focus on the behavior of a distribution's "waist" that refers to the portion for which there are enough data to summarize the distributional information. Our results provide the mathematical basis for understanding these competing claims and show that indeed depending on the operating points and parameters of interest, either heavy or light tail phenomenon may dominate performance.

Also, from an engineering perspective, our results further emphasize the insight developed in [2] that retransmissions may significantly amplify the packet size variability to much larger variability in transmission delays. More precisely, if there is a polynomial functional relationship between the statistical characteristics of channel dynamics and packet size variability, the transmission duration is very close to a power law distribution over the time scale of order $o\left(1/\mathbb{P}[A > b]\right)$. Thus, even for packet lengths with small MTUs, the small variability in the packet size can still be amplified by the retransmission based protocols, causing potentially poor performance. This observation is in agreement with empirical measurements in [12], which claims that the utilization of the 802.11 protocol is only $40\%$ basically due to retransmissions.

Our results also provide insights in designing control algorithms. For example, in physical layer, power control can change the rate at which the packet is transmitted. In this sense, power control can be thought as a way to change the relationship between the channel dynamics and the units in which packets should be transmitted in order to achieve the best network performance.

## II. Model description and preliminary results

In this section, we formally describe our model and provide necessary definitions and notation. Some preliminary results are also presented in this part.

Throughout this paper, a positive measurable function $f$ is called regularly varying (at infinity) with index $\rho$ if $\lim_{x\to\infty} f(\lambda x)/f(x) = \lambda^{\rho}$ for all $\lambda > 0$. It is called slowly varying if $\rho = 0$ [13]. For any two real functions $f(t)$ and $g(t)$ we use $f(t) \sim g(t)$ as $t \to \infty$ to denote $\lim_{t\to\infty} f(t)/g(t) = 1$. Similarly, we say that $f(t) \gtrsim g(t)$ as $t \to \infty$ if $\underline{\lim}_{t\to\infty} f(t)/g(t) \geq 1$; $f(t) \lesssim g(t)$ has a complementary definition. We use "$\overset{d}{=}$" and "$\overset{d}{\leq}(\overset{d}{\geq})$" to denote equal in distribution and less (greater) or equal in distribution, respectively. We use $\vee$ to denote $\max$, i.e., $x \vee y \equiv \max\{x, y\}$.

In this paper, we adopt the retransmission model that was proposed in [2]. The channel dynamics are modeled as an on-off process $\{(A_i, U_i)\}_{i\geq 1}$ that alternates between available $A_i$ and unavailable $U_i$ periods, respectively. Let $L$ denote the random length of a generic packet. At the beginning of each time period $A_i$ when the channel becomes available, we attempt to transmit the packet. If $L < A_i$, we say that the transmission is successful; otherwise, we wait until the beginning of the next available period $A_{i+1}$ and retransmit the packet from the beginning. This process continues until the packet is successfully transmitted over the channel. In this paper, we assume that $\{U_i\}_{i\geq 1}$ and $\{A_i\}_{i\geq 1}$ are two mutually independent sequences of i.i.d. random variables with $U_i \overset{d}{=} U$, $A_i \overset{d}{=} A$ and $U$ independent of $A$. A sketch of the model depicting the system is drawn in Figure 1; see also in [2].

As mentioned earlier, unlike [2], we allow the packet length $L$ to take values on finite interval $[0, b], b > 0$. Our goal will be to study the behavior of the number of retransmissions $N(b)$ and the total transmission delay $T(b)$ as $b$ scales with the number of retransmissions.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

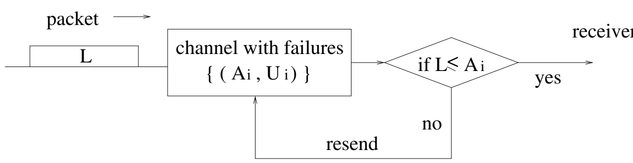


Fig. 1. Packets sent over channels with failures

*Definition 1:* The total number of (re)transmissions for a packet of length $L$ is defined as

$$N(b) \triangleq \inf\{n : A_n > L\},$$

and, the total transmission time for the packet is defined as

$$T(b) \triangleq \sum_{i=1}^{N(b)-1} (A_i + U_i) + L. \quad (1)$$

*Case I (heavy tails):* When the distribution function of $L$ has an infinite support ($b = \infty$), it has been shown in Lemma 1 of [2] and Proposition 1.2 of [6] that both the transmission time and the number of transmissions follow subexponential distributions, as given by the following result.

*Proposition 1 (from [2]):* If $\mathbb{P}[L > x] > 0$ for all $x \geq 0$, then both $N(\infty)$ and $T(\infty)$ are subexponential in the following sense that, for any $\epsilon > 0$, as $n \to \infty$ and $t \to \infty$,

$$e^{\epsilon n}\mathbb{P}[N(\infty) > n] \to \infty, \quad e^{\epsilon t}\mathbb{P}[T(\infty) > t] \to \infty.$$

This class of heavy-tailed distributions has a rich structure, including power laws, heavy-tailed Weibull distributions and nearly exponential distributions. For a detailed analysis of this class of distributions induced by retransmissions see [4]. Since power law distributions are closely related to long range dependency (see Section IV-B) and channel stability (see Section IV-A), we focus on power law delays. Below we quote Theorem 2 of [2] as Proposition 2 (see also Theorem 2.5 in [4] and Theorem 2.2 in [6]), which show that both $N(\infty)$ and $T(\infty)$ can follow power law distributions.

*Proposition 2:* If there exists $\alpha > 0$ such that

$$\lim_{x\to\infty} \frac{\log \mathbb{P}[L > x]}{\log \mathbb{P}[A > x]} = \alpha,$$

then,

$$\lim_{n\to\infty} \frac{\log \mathbb{P}[N(\infty) > n]}{\log n} = -\alpha.$$

Additionally, if $\mathbb{E}\left[U^{(\alpha\vee 1)+\theta}\right] < \infty$ and $\mathbb{E}\left[A^{1+\theta}\right] < \infty$ for some $\theta > 0$, then,

$$\lim_{t\to\infty} \frac{\log \mathbb{P}[T(\infty) > t]}{\log t} = -\alpha.$$

*Case II (light tails):* On the other hand, when $b$ is finite, unlike the preceding case that can cause subexponential delays, both the number of retransmissions $N(b)$ and the total transmission time $T(b)$ have exponential tails, as shown in the following proposition. When $U_i \equiv 0$ and $A$ has a density function, this case has been studied in Corollary 3.1 of [6].

*Proposition 3:* For $b < \infty$ with $0 < \mathbb{P}[A \leq b] < 1$, if $\mathbb{P}[A + U > x | A \leq b]e^{\gamma x}$ is nonlattice and directly Riemann integrable with $\gamma$ being the solution of $\int_0^\infty e^{\gamma s} d\mathbb{P}[A + U \leq s | A \leq b] = 1/\mathbb{P}[A \leq b]$, then

$$\mathbb{P}[N(b) > n] \leq (\mathbb{P}[A \leq b])^n, \quad \mathbb{P}[T(b) > t] \lesssim Ce^{-\gamma(t-b)},$$

where

$$C = \frac{\mathbb{P}[A > b]\mathbb{P}[A \leq b]^{-1}\gamma^{-1}}{\int_0^\infty s e^{\gamma s} d\mathbb{P}\left[A + U \leq s | A \leq b\right]}.$$

*Proof:* See Appendix. ■

Since heavy tails and light tails have very different characteristics, the preceding cases are two extremes that correspond to a finite and infinite $b$, respectively, which motivates us to study the transition from heavy tails to light tails when $b$ scales as well. To that end, we introduce a hidden random variable $L^*$ that has an infinite support, and the packet size $L$ satisfies the following condition

$$\mathbb{P}[L > x] = \begin{cases} \mathbb{P}[b \geq L^* > x]/\mathbb{P}[L^* \leq b] & x \leq b \\ 0 & x > b, \end{cases} \quad (2)$$

where $\mathbb{P}[L^* \leq b] > 0$. Clearly, when $b$ changes, the distribution of $L$ changes accordingly with respect to $b$. Thus, we also use the notation $L_b \equiv L$ for increased clarity when necessary.

We use the following simulation to further illustrate the effect of transition from the heavy-tailed main body to the light tail for the transmission delay distribution.

*Example 1:* This example shows that the delay distribution has a power law main body and an exponential tail when $L$ takes values on a finite interval $[0, b]$. Assume that $U_i = 0$ and both $L^*$ and $A$ follow exponential distributions with rate 0.8 and 1.0, respectively. We plot $\mathbb{P}[T(b) > t]$ on the log-log scale by changing $b$ from 4 to 10. It is clear from the figure that the support of the power law main body increases very quickly with respect to $b$: a small increment of $b$ from 4 to 10 results in a big expansion of the power law support from nearly 10 to 1000, which is a dramatic amplification.

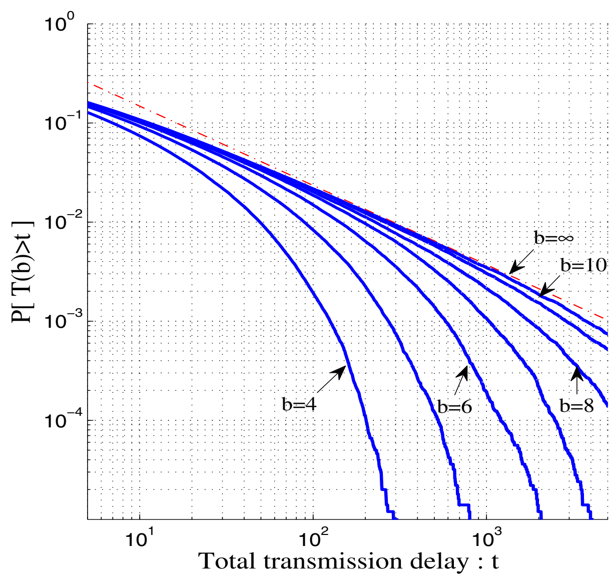


Fig. 2. The transmission delay distribution

We use the following notations to denote the complementary cumulative distribution functions for $A$ and $L^*$, respectively,

$$\bar{G}(x) \triangleq \mathbb{P}[A > x]$$

and

$$\bar{F}(x) \triangleq \mathbb{P}[L^* > x].$$

*Lemma 1:* If $0 < b_1 \le b_2$, then $N(b_1) \overset{d}{\le} N(b_2)$ and $T(b_1) \overset{d}{\le} T(b_2)$.

*Remark 1:* This result is supported by the preceding example. As easily seen from Figure 2, $\mathbb{P}[T(b) > t]$ is monotonically increasing as $b$ increases.

*Proof:* First, we want to show that $L_{b_1} \overset{d}{\le} L_{b_2}$. Recalling (2), it is sufficient to show that, for $x \ge 0$,

$$\frac{\mathbb{P}[x < L^* \le b_1]}{\mathbb{P}[L^* \le b_1]} \le \frac{\mathbb{P}[x < L^* \le b_2]}{\mathbb{P}[L^* \le b_2]}.$$

It is easy to verify the preceding inequality by checking

$$(\mathbb{P}[L^* \le b_1] - \mathbb{P}[L^* \le x])\, \mathbb{P}[L^* \le b_2] \le (\mathbb{P}[L^* \le b_2] - \mathbb{P}[L^* \le x])\, \mathbb{P}[L^* \le b_1].$$

Now, by Definition 1, both $N(b)$ and $T(b)$ is monotonically increasing in $L$. Therefore, we prove that $N(b_1) \overset{d}{\le} N(b_2)$ and $T(b_1) \overset{d}{\le} T(b_2)$. ■

## III. Main Results

This section presents our main results. Here, we assume that $\bar{F}(x)$ is a continuous function with support on $[0, \infty)$; note that $L$ has a finite support on $[0, b]$ by Equation (2).

We now present our main results that characterize the transition of the distributions of $N(b)$ and $T(b)$ from power law main bodies to exponential tails.

*Theorem 1:* If for some $\alpha > 0$,

$$\lim_{x \to \infty} \frac{\log \mathbb{P}[L^* > x]}{\log \mathbb{P}[A > x]} = \alpha, \tag{3}$$

then, for fixed $0 < \eta < 1, \eta_2 > \eta_1 > 1$ and $\epsilon > 0$, there exist $n_0$ and $b_0$ such that for any $b > b_0$, we have

i) for $n_0 < n < \left(\bar{G}(b)\right)^{-\eta}$,

$$1 - \epsilon < \frac{\log \mathbb{P}\left[N\left(b\right) > n\right]}{-\alpha \log n} < 1 + \epsilon, \tag{4}$$

ii) for $\left(\bar{G}(b)\right)^{-\eta_1} < n < \left(\bar{G}(b)\right)^{-\eta_2}$,

$$-n^{1 - \frac{1}{\eta_2 + \epsilon}} \le \log \mathbb{P}\left[N\left(b\right) > n\right] \le -n^{1 - \frac{1}{\eta_1 - \epsilon}}. \tag{5}$$

In addition, if $\bar{G}(x)$ is left-continuous at $b^* \triangleq \sup\{x : \mathbb{P}[L > x] > 0\} > 0$, we have,

$$\lim_{n \to \infty} \frac{\log \mathbb{P}\left[N\left(b\right) > n\right]}{n} = \log\left(1 - \bar{G}(b^*)\right). \tag{6}$$

*Remark 2:* This result shows that when $b$ is large, albeit finite, the distributions of $N(b)$ and $T(b)$ consist of different regions: a power law main body, an exponential tail, and Weibull type distributions in the middle. Equation (4) implies that $\mathbb{P}[N(b) > n]$ is approximately a power law distribution with index $\alpha$ when $n$ is smaller than $\left(\bar{G}(b)\right)^{-1}$. Equation (4) says that $\mathbb{P}[N(b) > n]$ is sandwiched between Weibull distributions when $n$ is larger than $\left(\bar{G}(b)\right)^{-1}$ but smaller than a polynomial function of $\left(\bar{G}(b)\right)^{-1}$. Equation (6) suggests an exponential tail distribution when $n$ is large. Note that $\left(\bar{G}(b)\right)^{-1} = (\mathbb{P}[A > b])^{-1}$ is the expected number of transmissions for packet of size $b$.

*Proof:* See Appendix. ■

*Theorem 2:* Under the condition (3) and $\mathbb{E}\left[e^{\theta(A+U)}\right] < \infty$ for some $\theta > 0$, we have, there exist $t_0$ and $b_0$ such that for any $b > b_0$,

i) for $t_0 < t < \left(\bar{G}(b)\right)^{-\eta}$,

$$1 - \epsilon < \frac{\log \mathbb{P}\left[T\left(b\right) > t\right]}{-\alpha \log t} < 1 + \epsilon. \tag{7}$$

ii) for $\left(\bar{G}(b)\right)^{-\eta_1} < t < \left(\bar{G}(b)\right)^{-\eta_2}$,

$$-t^{1 - \frac{1}{\eta_2 + \epsilon}} \le \log \mathbb{P}\left[T\left(b\right) > t\right] \le -t^{1 - \frac{1}{\eta_1 - \epsilon}}. \tag{8}$$

In addition, if $\bar{G}(x)$ is left-continuous at $b^* = \sup\{x : \mathbb{P}[L > x] > 0\} > 0$, we have,

$$\lim_{t \to \infty} \frac{\log \mathbb{P}\left[T\left(b\right) > t\right]}{t} = -\gamma, \tag{9}$$

where $\gamma$ is the solution of

$$\int_0^\infty e^{\gamma s} d\mathbb{P}\left[A + U \le s | A \le b^*\right] = 1/(1 - \bar{G}(b^*)). \tag{10}$$

*Remark 3:* This theorem basically says that the distribution tail of $T(b)$ is dominated by $N(b)$ if $L$, $A$ and $U$ have lighter tails than $N(b)$. In order for (7) to hold, we only need the existent of a finite moment larger than $\alpha$; see the condition in Proposition 2. By the same reasoning, we only need a Weibull distribution bound on $A$ and $L$ for (8) to hold. However, in order to provide a unified framework to study all the three cases, we use the condition $\mathbb{E}\left[e^{\theta(A+U)}\right] < \infty$ for some $\theta > 0$ instead.

*Proof:* See Appendix. ■

Under a more restrictive condition, we can obtain the following more precise result that characterizes the exact asymptote for the power law main body. Since the proof of this theorem basically combines the proof of Theorem 2.7 in [4] and the same techniques used in proving Theorems 1 and 2 above, we omit the details.

*Theorem 3:* If

$$\mathbb{P}[L^* > x]^{-1} \sim \Phi\left(\mathbb{P}[A > x]^{-1}\right) \tag{11}$$

where $\Phi(\cdot)$ is regularly varying with index $\alpha > 0$, then, for $\epsilon > 0$ and $0 < \eta < 1$, there exist $n_0, t_0$ and $b_0$ such that for any $b > b_0$, we have

$$\sup_{n_0 < n < \left(\bar{G}(b)\right)^{-\eta}} \left| \frac{\mathbb{P}\left[N\left(b\right) > n\right] \Phi(n)}{\Gamma(\alpha + 1)} - 1 \right| < \epsilon$$

where $\Gamma(\cdot)$ is the gamma function. Additionally, if $\mathbb{E}\left[U^{(\alpha \vee 1) + \theta}\right] < \infty$, $\mathbb{E}\left[A^{1+\theta}\right] < \infty$ and $\mathbb{E}\left[(L^*)^{\alpha+\theta}\right] < \infty$

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

for some $\theta > 0$, then,

$$\sup_{t_0 < t < (\bar{G}(b))^{-\eta}} \left| \frac{\mathbb{P}\left[T(b) > t\right] \Phi(t)}{\Gamma(\alpha+1)} - 1 \right| < \epsilon. \qquad (12)$$

*Remark 4:* The preceding results imply Equations (4) and (7), and characterize the exact asymptote for the power law main body.

### A. Simulation results

Now, we use simulation to support our results. First, we verify the support region of the power law main body that is characterized by Theorem 2.

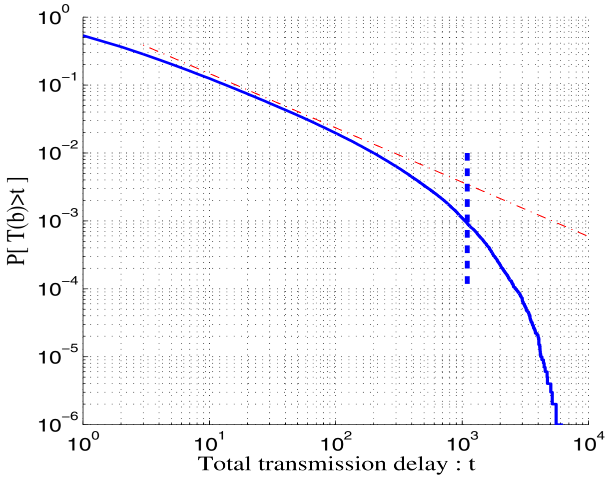


Fig. 3. Power law main body

Under the setting of Example 1, we plot the distribution of the total transmission delay $T(b)$ for $b = 7$ in Figure 3. The dotted line represents the asymptote $\Gamma(1.0/0.8)t^{-1.0/0.8}$ that is computed using (12). The dashed vertical line corresponds to the average transmission time for a packet of size $b = 7$ over the channel, i.e., $\mathbb{E}[A]/\bar{G}(b) = e^7 = 1.096 \times 10^3$. It is easy to see a power law main body that is close to the computed asymptote on the interval $[0, 1.096 \times 10^3]$.

Next, we investigate the exponential tail under the setting of Example 1. We plot the distribution of the number of retransmissions $N(b)$ for $b = 3$ in Figure 4. It can be observed that, after taking logarithm with base 10, the tail distribution is asymptotically a straight line. The slope of the line is $-0.022$, which matches our theoretical result $\log_{10}\left(1 - \bar{G}(b)\right) = \log_{10}\left(1 - e^{-3}\right)$ that is computed by equation (6).

## IV. Related Models

In this section, using the results obtained in Section III, we study two related models that further highlight the engineering implications. In Section IV-A, we show that, even when the packet has a small mean with a bound that does not depend on its upper limit $b$, due to the power law main body caused by the polynomial functional relationship between the packet size variability and channel dynamics, the channel throughput may still be very close to zero. Therefore, both choosing an appropriate upper limit $b$ and designing the packet variability

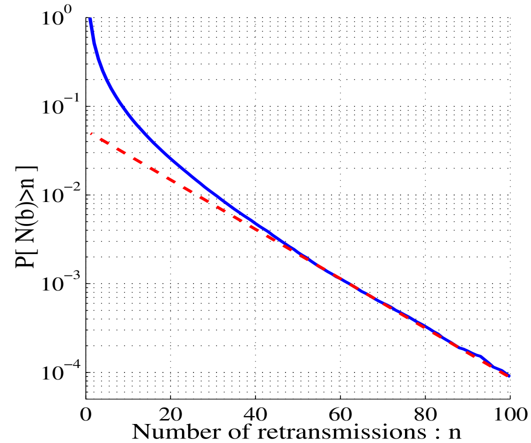


Fig. 4. Exponential tail

are important for the system performance. In Section IV-B, we investigate the autocorrelation function of the number of jobs in a $M/M/\infty$ queue with server failures. If a server failure occurs during the processing of a job, this job has to restart from the beginning on the same server. We show that, under the condition that the job sizes follow exponential distribution (with infinite support), this model can cause long range dependence. Introducing an upper limit to all job sizes can eliminate the long range dependence; however, the strong dependence can span over a large interval, which implies a performance very close to long range dependence.

### A. Throughput of the on-off channel

Consider the same on-off channel model as in Section II. Now, suppose that the source has an infinite number of backlogged packets to be sent. Let $\{L_i\}_{i\geq 1}$ be the i.i.d. sequence of packet sizes. Define $\{T_i\}$ to be the duration for transmitting packet $L_i$. Note that the channel is still available immediately after the successful transmission of packet $L_i$, thus we can start transmitting $L_{i+1}$ without waiting for the next available period. Due to this effect, the durations $\{T_i\}$ are not independent random variables. We are interested in studying the throughput $\Lambda_n(b)$ of this system for the first $n$ packets,

$$\Lambda_n(b) \triangleq \frac{\sum_{i=1}^{n} L_i}{\sum_{i=1}^{n} T_i}.$$

Since it is not clear whether $\Lambda_n(b)$ converges to a limit as $n$ goes to infinity, we use $\overline{\underline{\lim}}$ to denote both $\overline{\lim}$ and $\underline{\lim}$.

*Theorem 4:* Under the conditions described in Theorem 2, if $0 < \alpha < 1$, then, the throughput of the on-off channel $\Lambda_n(b)$ satisfies, as $b \to \infty$,

$$\overline{\underline{\lim}}_{n\to\infty} \frac{\log \Lambda_n(b)}{\log \bar{G}(b)} \sim (1-\alpha). \qquad (13)$$

*Proof:* See Appendix. ■

*Remark 5:* It is clear from the preceding result that, if both $L^*$ and $A$ follows exponential distributions with rate $\mu$ and $\nu$, respectively, then, as the upper limit $b$ goes to $\infty$, the throughput vanishes to zero roughly with speed $e^{-(\nu-\mu)b}$ for

$\nu > \mu$. Note that the mean packet size $\mathbb{E}[L]$ is bounded by a constant $\mathbb{E}[L^*] = 1/\mu$ that does not depend on $b$. This suggests that some special care needs to be taken when engineering retransmission based protocols. Not only the maximum packet size but also the variability of the packet size can impact the throughput.

### B. Long-range dependence in M/M/∞ queues with failures

Consider a $M/M/\infty$ queue with Poisson arrival rate $\lambda$. Let $L^*$ be an exponential random variable with rate $\mu$. Suppose that all the active servers fail independently according to Poisson point processes with the same rate $\nu$. Immediately after a server fails in the middle of processing a job that runs on this server, this job has to restart on the same server from the beginning. Denote by $T_i$ the processing time for job $i$. Let $M(t)$ be the number of jobs in the system at time $t$. Theorem 5 shows that $M(t)$ may be long range dependent if all the job sizes $\{L_i\}_{i>-\infty}$ follow an exponential distribution with rate $\mu$; $L_i \stackrel{d}{=} L^*$. In the rest of this section, we assume that the system has reached stationarity.

*Theorem 5:* If $1 < \mu/\nu < 2$, then $M(t)$ is long range dependent in the sense that

$$\int_0^\infty \mathrm{Cov}(M(t), M(t+s))ds = \infty.$$

*Remark 6:* If $\mu/\nu \leq 1$, then $\mathbb{V}ar[M(t)] = \infty$, and $\mathrm{Cov}(M(t), M(t+s))$ is not defined for this case. If $\mu/\nu = 2$, the result in this theorem also holds, but we omit the proof.

*Proof:* By the well-known result on the autocorrelation function of the number of jobs in a $M/G/\infty$ queue (e.g., see [14]), we obtain

$$\mathrm{Cov}(M(t), M(t+s)) = \lambda \int_s^\infty \mathbb{P}[T_1 > x]dx.$$

Proposition 2 implies that, for $0 < \epsilon < 2 - \mu/\nu$,

$$\mathbb{P}[T_1 > x] \gtrsim \frac{1}{x^{\mu/\nu+\epsilon}},$$

as $x \to \infty$. Therefore, we obtain, as $s \to \infty$,

$$\mathrm{Cov}(M(t), M(t+s)) \gtrsim \frac{\lambda}{s^{\mu/\nu-1+\epsilon}},$$

which, using the fact that $0 < \mu/\nu - 1 + \epsilon < 1$, completes our proof. ■

This long range dependence can be eliminated by restricting the support of the job size distribution. In the following, similar to Section III, assume that all the i.i.d. job sizes $\{L_i\}$ follow a truncated exponential distribution with $\mathbb{P}[L_1 \geq x] = \mathbb{P}[L^* \geq x | L^* \leq b]$ for $b > 0$. We show in Theorem 6 that the integration of the autocorrelation function may still be very large, indicating a performance close to long-range dependence. Since the proof is based on Theorem 2 and similar arguments as in deriving (34), we omit the proof.

*Theorem 6:* If $1 < \mu/\nu < 2$ and job sizes have a finite support on $[0, b]$, then, as $b \to \infty$,

$$\log \int_0^\infty \mathrm{Cov}(M(t), M(t+s))ds \sim (2\nu - \mu)b,$$

which goes to infinity exponentially fast as $b \to \infty$.

## V. Appendix

Since $\bar{F}(x)$ is a continuous function on $[0, \infty)$, we can define its inverse function $\bar{F}^{\leftarrow}(x) \triangleq \inf\{y : \bar{F}(y) > x\}$.

### A. Proof of Proposition 3

First we proof the result for $N(b)$. Since the sequence $\{A_i\}$ is i.i.d., upper bounding the packet size $L$ by $b$ yields $\mathbb{P}[N(b) > n] \leq \mathbb{P}[A \leq b]^n$.

Next, we study $T(b)$. For a random variable $\bar{N}$ that is independent of $\{A_i, U_i\}$ with $\mathbb{P}\left[\bar{N} > n\right] = \mathbb{P}[A \leq b]^n$, $n = 0, 1, 2, \cdots$ (thus $\bar{N} \geq 1$), we have

$$T(b) \leq \sum_{i=1}^{\bar{N}-1} (\bar{A}_i + U_i) + b, \tag{14}$$

where $\bar{A}_i, i \geq 1$ are i.i.d. and dependent of $\bar{N}, \{U_i\}$ with $\mathbb{P}\left[\bar{A}_i > t\right] = \mathbb{P}[A > t | A \leq b]$. Noting that the first term on the righthand side of (14) is a geometric sum of i.i.d. random variables, we derive the following defective renewal equation, for $t > 0$,

$$\mathbb{P}\left[\sum_{i=1}^{\bar{N}-1} (\bar{A}_i + U_i) > t\right] = \mathbb{P}\left[\bar{N} > 1, \bar{A}_1 + U_1 > t\right] + \mathbb{P}\left[\bar{N} > 1, \bar{A}_1 + U_1 \leq t, \sum_{i=1}^{\bar{N}-1} (\bar{A}_i + U_i) > t\right]$$

$$= \mathbb{P}[A \leq b]\mathbb{P}[\bar{A}_1 + U_1 > t] + \mathbb{P}[A \leq b] \times \int_0^t \mathbb{P}\left[\sum_{i=1}^{\bar{N}-1} \bar{A}_1 + U_1 > t - u\right] d\mathbb{P}\left[\bar{A}_1 + U_1 \leq u\right]. \tag{15}$$

Applying Theorem 7.1 of [15], we derive

$$\lim_{t\to\infty} e^{\gamma(t-b)}\mathbb{P}\left[\sum_{i=1}^{\bar{N}-1} (\bar{A}_i + U_i) + b > t\right] = \frac{\int_0^\infty e^{\gamma s}\mathbb{P}\left[A + U > s | A \leq b\right] ds}{\mathbb{P}[A \leq b] \int_0^\infty s e^{\gamma s} d\mathbb{P}\left[A + U \leq s | A \leq b\right]},$$

which, using $\int_0^\infty e^{\gamma s} d\mathbb{P}[A + U \leq s | A \leq b] = 1/\mathbb{P}[A \leq b]$ to compute the numerator, completes the proof. ■

### B. Proof of Theorem 1

Notice that the number of retransmissions is geometrically distributed given the unit size $L$,

$$\mathbb{P}[N(b) > n \mid L] = (1 - \bar{G}(L))^n,$$

and therefore,

$$\mathbb{P}[N(b) > n] = \mathbb{E}[(1 - \bar{G}(L))^n]. \tag{16}$$

I) First, we prove (4). Using Lemma 1 and Proposition 2, we obtain the upper bound

$$\lim_{n\to\infty}\frac{\log\mathbb{P}[N(b)>n]}{\log n}\le\lim_{n\to\infty}\frac{\log\mathbb{P}[N(\infty)>n]}{\log n}=-\alpha,$$

which implies that, for $\epsilon>0$, there exists $n_0$ such that

$$\inf_{n>n_0}\frac{\log\mathbb{P}[N(b)>n]}{-\alpha\log n}-1>-\epsilon. \tag{17}$$

Next, we derive a lower bound. The condition described in (3) implies that for any $0<\epsilon<1/\alpha$, there exists $x_\epsilon$, such that for all $x>x_\epsilon$, we have

$$\bar{F}(x)^{\frac{1}{\alpha}+\epsilon}\le\bar{G}(x)\le\bar{F}(x)^{\frac{1}{\alpha}-\epsilon}. \tag{18}$$

The condition $n<\left(\bar{G}(b)\right)^{-\eta}$ implies $n<\left(\bar{F}(b)\right)^{-\left(\frac{1}{\alpha}-\epsilon\right)\eta}$. Let $\zeta\triangleq(1/\alpha-\epsilon)\eta<1/\alpha$, and we obtain $b>\bar{F}^{\leftarrow}(n^{-1/\zeta})$. Thus, there exists $b_0>x_\epsilon$ such that for all $b>b_0$,

$$\begin{aligned}
&\mathbb{E}[(1-\bar{G}(L))^n]=\mathbb{E}\left[(1-\bar{G}(L^*))^n|L^*\le b\right]\\
&\ge\frac{1}{\mathbb{P}[L^*\le b]}\mathbb{E}\left[\left(1-\bar{G}(L^*)\right)^n\mathbf{1}(x_\epsilon<L^*<b)\right]\\
&\ge\mathbb{E}\Big[\left(1-\bar{F}(L^*)^{\frac{1}{\alpha}-\epsilon}\right)^n\mathbf{1}\left(\bar{F}(b)<\bar{F}(L^*)<\bar{F}(x_\epsilon)\right)\Big]\\
&\ge\mathbb{E}\left[\left(1-\bar{F}(L^*)^{\frac{1}{\alpha}-\epsilon}\right)^n\mathbf{1}\left(\bar{F}(b)<\bar{F}(L^*)\right)\right]\\
&\quad-\left(1-\bar{F}(x_\epsilon)^{\frac{1}{\alpha}-\epsilon}\right)^n.
\end{aligned}$$

Since $\bar{F}(x)$ is continuous, $\bar{F}(L^*)$ is a uniform random variable between 0 and 1, the preceding inequality implies

$$\begin{aligned}
\mathbb{E}[(1-\bar{G}(L))^n]\ge&\int_{\bar{F}(b)}^1\left(1-u^{\frac{1}{\alpha}-\epsilon}\right)^n du\\
&-\left(1-\bar{F}(x_\epsilon)^{\frac{1}{\alpha}-\epsilon}\right)^n,
\end{aligned}$$

which, letting $u^{1/\alpha-\epsilon}=v$ and noting $\bar{F}(b)<n^{-1/\zeta}$, yields

$$\begin{aligned}
\mathbb{E}[(1-\bar{G}(L))^n]\ge&\int_0^1(1-v)^n\frac{\alpha}{1-\alpha\epsilon}v^{\frac{\alpha}{1-\alpha\epsilon}-1}dv\\
&-\int_0^{n^{-1/\zeta}}\left(1-u^{\frac{1}{\alpha}-\epsilon}\right)^n du-\left(1-\bar{F}(x_\epsilon)^{\frac{1}{\alpha}-\epsilon}\right)^n\\
\triangleq\ &P_1-P_2-P_3.
\end{aligned} \tag{19}$$

Recalling the property of Beta function, for fixed $x$,

$$B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}dt\sim\Gamma(x)y^x \tag{20}$$

as $y\to\infty$, we obtain,

$$P_1\sim\frac{\alpha\Gamma(\alpha_\epsilon)}{1-\alpha\epsilon}n^{-\alpha_\epsilon},\quad\alpha_\epsilon=\frac{\alpha}{1-\alpha\epsilon}. \tag{21}$$

Choosing $\epsilon<1/\alpha-\zeta$, i.e., $1/\zeta>\alpha_\epsilon$, we know

$$P_2\le\frac{1}{n^{1/\zeta}},$$

which, in conjunction with (19) and the fact that $P_3$ is exponentially bounded, implies that, for $\epsilon>0$, there exists $n_0>0$ such that for $n_0<n<\left(\bar{G}(b)\right)^{-\eta}$,

$$\mathbb{E}[(1-\bar{G}(L))^n]\ge(1-\epsilon)\int_0^1(1-v)^n\frac{\alpha}{1-\alpha\epsilon}v^{\frac{\alpha}{1-\alpha\epsilon}-1}dv.$$

Using (21) and the preceding lower bound, we obtain, for any fixed $\epsilon_1>0$, there exists $n_0$ such that

$$\sup_{n_0<n<\left(\bar{G}(b)\right)^{-\eta}}\frac{\log\mathbb{P}[N(b)>n]}{-\alpha\log n}-1<\epsilon_1. \tag{22}$$

Combining (17) and (22), we finish the proof of (4).

II) Next, we prove (5). Since $\bar{F}(x)$ is eventually non-increasing, by equation (18), the condition $\left(\bar{G}(b)\right)^{-\eta_1}<n<\left(\bar{G}(b)\right)^{-\eta_2}$ implies that

$$\bar{F}^{\leftarrow}\left(n^{-1/\zeta_2}\right)<b<\bar{F}^{\leftarrow}\left(n^{-1/\zeta_1}\right) \tag{23}$$

where $\zeta_2\triangleq(1/\alpha+\epsilon)\eta_2$ and $\zeta_1\triangleq(1/\alpha-\epsilon)\eta_1$. Choosing $0<\epsilon<(\eta_1-1)/(\alpha\eta_1)$, we obtain $\zeta_2>\zeta_1>1/\alpha$.

For the upper bound, by Lemma 1 and (23), we obtain,

$$\begin{aligned}
\mathbb{P}[N(b)>n]=&\ \mathbb{E}\left[(1-\bar{G}(L^*))^n|L^*\le b\right]\\
\le&\ \frac{1}{1-n^{-1/\zeta_2}}\mathbb{E}\Big[\left(1-\bar{G}(L^*)\right)^n\\
&\qquad\mathbf{1}\left(x_\epsilon<L^*\le\bar{F}^{\leftarrow}\left(n^{-1/\zeta_1}\right)\right)\Big]\\
&+\frac{1}{1-n^{-1/\zeta_2}}(1-\bar{G}(x_\epsilon))^n.
\end{aligned}$$

Therefore, there exists $n_0$, such that for all $n>n_0$, $\mathbb{P}[N(b)>n]$ is upper bounded by

$$\begin{aligned}
&(1+\epsilon)\mathbb{E}\left[\left(1-V^{\frac{1}{\alpha}+\epsilon}\right)^n\mathbf{1}\left(\frac{1}{n^{1/\zeta_1}}<V<\bar{F}(x_\epsilon)\right)\right]\\
&\quad+(1+\epsilon)(1-\bar{G}(x_\epsilon))^n\\
&\le(1+\epsilon)\left(1-\frac{1}{n^{1/(\zeta_1\alpha)+\epsilon/\zeta_1}}\right)^n+(1+\epsilon)(1-\bar{G}(x_\epsilon))^n\\
&\le(1+2\epsilon)e^{-n^{1-1/(\zeta_1\alpha)-\epsilon/\zeta_1}},
\end{aligned}$$

which implies, for $n_0$ large enough,

$$\log\mathbb{P}[N(b)>n]\le-n^{1-\frac{1}{\alpha(\zeta_1-\epsilon_1)}} \tag{24}$$

where $0<\epsilon_1<\zeta_1-1/\alpha$.

To prove the lower bound, for $0<\epsilon<1/\alpha$, we obtain

$$\begin{aligned}
&\mathbb{P}[N(b)>n]=\mathbb{E}\left[(1-\bar{G}(L^*))^n|L^*\le b\right]\\
&\ge\mathbb{E}\Big[\left(1-\bar{G}(L^*)\right)^n\\
&\qquad\mathbf{1}\left(\bar{F}^{\leftarrow}\left(n^{-1/\zeta_2}\right)<L^*<\bar{F}^{\leftarrow}\left(n^{-1/\zeta_1}\right)\right)\Big]\\
&\ge\mathbb{E}\left[\left(1-V^{\frac{1}{\alpha}-\epsilon}\right)^n\mathbf{1}\left(\frac{1}{n^{1/\zeta_1}}<V<\frac{1}{n^{1/\zeta_2}}\right)\right]\\
&\ge\left(\frac{1}{n^{1/\zeta_2}}-\frac{1}{n^{1/\zeta_1}}\right)\left(1-\frac{1}{n^{1/(\alpha\zeta_2)-\epsilon/\zeta_2}}\right)^n\\
&\sim\frac{1}{n^{1/\zeta_2}}e^{-n^{1-1/(\alpha\zeta_2)-\epsilon/\zeta_2}},
\end{aligned}$$

which implies, for $n > n_0$ with $n_0$ large enough,

$$\log \mathbb{P}[N(b) > n] \geq -n^{1-\frac{1}{\alpha(\zeta_2+\epsilon_2)}} \tag{25}$$

where $\epsilon_2 > 0$. Combining (24), (25) and choosing the right $\epsilon_1, \epsilon_2$, we finish the proof of (5).

III) Finally, we prove (6). Using $\mathbb{P}[L \leq b^*] = 1$ and (16) yields

$$\mathbb{P}[N(b) > n] \leq (1 - \bar{G}(b^*))^n,$$

which proves the upper bound.

Next, we prove the lower bound. For $\epsilon > 0$, note that

$$\begin{aligned}\mathbb{P}[N(b) > n] &\geq \mathbb{E}\left[(1 - \bar{G}(L)^n \mathbf{1}(b^* - \epsilon < L < b^*)\right] \\ &\geq \mathbb{P}[b^* - \epsilon < L < b^*](1 - \bar{G}(b^* - \epsilon))^n.\end{aligned}$$

The definition of $b^*$ implies $\mathbb{P}[b^* - \epsilon < L < b^*] > 0$, and therefore,

$$\lim_{n\to\infty} \frac{\mathbb{P}[N(b) > n]}{n} \geq -\log\left(1 - \bar{G}(b^* - \epsilon)\right),$$

which, by passing $\epsilon \to 0$ and using the continuity, finishes the proof of the lower bound. ■

### C. Proof of Theorem 2

First, we prove equations (7) and (8). Since the proof is based on the same approach as in proving Theorem 3 in [2], here we only discuss the proof of the upper bound.

For any $\delta > 0$, we have

$$\begin{aligned}\mathbb{P}[T(b) > t] &= \mathbb{P}\left[\sum_{i=1}^{N(b)-1}(U_i + A_i) + L > t\right] \\ &\leq \mathbb{P}\left[\sum_{i=1}^{N(b)}(U_i + A_i) > t, N(b) \leq \frac{t(1-\delta)}{\mathbb{E}[U+A]}\right] \\ &\quad + \mathbb{P}\left[N(b) > \frac{t(1-\delta)}{\mathbb{E}[U+A]}\right] + \mathbb{P}[L > t] \\ &\triangleq I_1 + I_2 + I_3.\end{aligned} \tag{26}$$

For $I_1$, let $X_i \triangleq (U_i + A_i) - \mathbb{E}[(U_i + A_i)]$ and $\zeta \triangleq (1 - \delta)/\mathbb{E}[U + A]$. Noting $\mathbb{E}e^{\theta X_1} < \infty$ and $\mathbb{E}X_1 = 0$, we obtain

$$I_1 \leq \mathbb{P}\left[\sum_{i \leq t(1-\delta)/\mathbb{E}[U+A]}(U_i + A_i) > t\right] = \mathbb{P}\left[\sum_{i\leq \zeta t} X_i > \delta t\right],$$

which, by Chernoff bound, is bounded by $he^{-\zeta t}$ for $h, \zeta > 0$.

Since $I_3$ also has an exponential tail, the upper bound is dominated by $I_2$, which implies the upper bounds for (7) and (8). The lower bound can be obtained by combining similar arguments as in the upper bound and the proof in [2].

Next, we prove (9). We begin with the upper bound. Recalling (1) and using Lemma 1, we upper bound the packet size $L$ by $b^*$ to obtain, almost surely,

$$T(b) \leq \sum_{i=1}^{\bar{N}-1}(\bar{A}_i + U_i) + b^*, \tag{27}$$

where $\mathbb{P}\left[\bar{N} > n\right] = \left(1 - \bar{G}(b^*)\right)^n$, $\mathbb{P}\left[\bar{A}_i > t\right] = \mathbb{P}[A > t|A \leq b^*]$, $\{\bar{A}_i\}$ is a sequence of i.i.d. random variables independent from $\{U_i\}$ and $\bar{N}$ is independent of $\{\bar{A}_i, U_i\}$. Using the same approach as in computing (15), we obtain the upper bound.

Now, we prove the lower bound. Observe that, for $0 < \epsilon < b^*$,

$$\begin{aligned}\mathbb{P}[T(b) > t] &\geq \mathbb{P}[T(b) > t, b^* - \epsilon < L < b^*] \\ &\geq \mathbb{P}[b^* - \epsilon < L < b^*]\,\mathbb{P}\left[\sum_{i=1}^{\underline{N}-1}(\underline{A}_i + U_i) + b^* - \epsilon > t\right],\end{aligned}$$

where $\mathbb{P}[\underline{N} > n] = \left(1 - \bar{G}(b^* - \epsilon)\right)^n$, $\mathbb{P}[\underline{A}_i > t] = \mathbb{P}[A > t|A \leq b^* - \epsilon]$, $\{\underline{A}_i\}$ is a sequence of i.i.d. random variables independent from $\{U_i\}$ and $\underline{N}$ is independent of $\{\underline{A}_i, U_i\}$. By the definition of $b^*$, we know $\mathbb{P}[b^* - \epsilon < L < b^*] > 0$. Using the same approach as in computing (15), we obtain

$$\lim_{t\to\infty} \frac{\mathbb{P}[T(b) > t]}{t} \geq -\gamma_\epsilon, \tag{28}$$

where $\gamma_\epsilon$ is the solution of

$$\int_0^\infty e^{\gamma s} d\mathbb{P}[A + U \leq s | A \leq b^* - \epsilon] = 1/\left(1 - \bar{G}(b^* - \epsilon)\right).$$

Passing $\epsilon \to 0$ and using the continuity, we finish the proof of the lower bound. ■

### D. Proof of Theorem 4

Since $\mathbb{E}[e^{\theta A}] < \infty$ for some $\theta > 0$, we can always find a random variable $X$ and $t_0, \delta > 0$ with $A \overset{d}{\leq} X$ and

$$\mathbb{P}[X > t] = \begin{cases} 1 & \text{if } t \leq t_0, \\ e^{-\delta(t-t_0)} & \text{if } t > t_0. \end{cases} \tag{29}$$

It can be checked that, $\mathbb{P}[X > t+s] \leq \mathbb{P}[X > t]\mathbb{P}[X > s]$ for any $t, s \geq 0$. Therefore, for $t \geq 0$ and any positive random variable $Y$ that is independent of $X$ with $\mathbb{P}[X > Y] > 0$, we obtain

$$\mathbb{P}[X > Y + t | X > Y] \leq \mathbb{P}[X > t]. \tag{30}$$

Denote by $\Phi(i), i \geq 1$ the index of the channel available period within which $L_i$ succeeds in transmission. Immediately after the successful transmission of the packet $L_i$, we denote by $\{\tau_i\}$ the remaining time and by $\{\sigma_i\}$ the elapsed time, respectively, within the available period $A_{\Phi(i)}$.

First, we prove the upper bound using the coupling argument. Using (29), we can construct in the same probability space an i.i.d. sequence $\{X_i^{(1)}\}_{i\geq 1}$ with $X_i^{(1)} \geq A_i$ and $X_i^{(1)} \overset{d}{=} X$ for all $i \geq 1$. Define $\bar{\tau}_i = X_{\Phi(i)}^{(1)} - \sigma_i$; obviously, $\bar{\tau}_i \geq \tau_i$. For a successful transmission with $L_i = \sigma_i$, we obtain, by (30),

$$\begin{aligned}\mathbb{P}[\tau_i > t | L_i = \sigma_i] &\leq \mathbb{P}[\bar{\tau}_i > t | L_i = \sigma_i] \\ &= \mathbb{P}[X_1^{(1)} > L_i + t | X_1^{(1)} > L_i] \leq \mathbb{P}[X > t].\end{aligned} \tag{31}$$

Now, we will construct a new system where we always have $L_i = \sigma_i$, and thus $\mathbb{P}[\tau_i > t | L_i = \sigma_i] = \mathbb{P}[\tau_i > t]$. We continue to use $\Phi(i), i \geq 1$ as the index of the channel available period within which $L_i$ succeeds in transmission in the newly constructed system. At the beginning of $A_1$, we replace $A_1$ by $X_1^{(1)}$ and start transmitting $L_1$. Then, immediately after each successful transmission, say, for packet $L_i$ in the available period $A_{\Phi(i)}$, in view of (31), we can construct a new available period by replacing $\tau_i$ with $X_i^{(2)}$, where $X_i^{(2)} \stackrel{d}{=} X$ and $X_i^{(2)} \geq \tau_i$. Note that in this construction we only change $A_{\Phi(i)}$ and other available periods are the same as the original ones. Then, let the system continue its operation by following this construction. Noting that $X$ has a constant hazard rate $\delta$ if $X \geq t_0$, the random variable $X_i^{(2)}$ is independent of all the random variables that appear before $X_i^{(2)}$ is generated in this new system. Denote by $\underline{T}_i, i \geq 1$ the transmission duration for packet $L_i$ in this new system excluding the time that this packet spends in the constructed interval $X_i^{(2)}$ and the unavailable period $U_{\Phi(i)}$ that immediately follows $X_i^{(2)}$; clearly $\sum_{i=1}^n \underline{T}_i \leq \sum_{i=1}^n T_i$ for all $n$. In addition, this construction implies that $\{\underline{T}_i\}_{i\geq 1}$ is an i.i.d. sequence. If $X_1^{(1)} > L_1$, the first transmission in the available period $X_1^{(1)}$ (replacing $A_1$) succeeds, and thus $\underline{T}_1 = 0$. If $X_1^{(1)} \leq L_1$, since the first transmission fails and we need to wait until the beginning of the second available period $A_2$ to retransmit $L_1$, $\underline{T}_1 \stackrel{d}{=} T_1$ where $T_1$ is the transmission time of the first packet in the original system. Hence, we obtain $\underline{T}_i \stackrel{d}{=} T_1 \mathbf{1}(L_1 > X)$ where $X$ is independent of $L_1$ and $\{A_i, U_i\}_{i\geq 1}$.

Therefore, as $n \to \infty$, we obtain, using the law of large numbers,

$$\Lambda_n(b) \leq \frac{\sum_{i=1}^n L_i}{\sum_{i=1}^n \underline{T}_i} = \frac{\sum_{i=1}^n L_i}{n} \frac{n}{\sum_{i=1}^n \underline{T}_i} \to \frac{\mathbb{E}[L]}{\mathbb{E}[\underline{T}_1]}. \quad (32)$$

Now, we need to compute, for $t > 0$,

$$\mathbb{P}[\underline{T}_1 > t] = \mathbb{P}[T_1 > t, L_1 > X].$$

Using the same $x_\epsilon$ chose in (18), we obtain, by the independence of $L_1$ and $X$,

$$\begin{aligned}\mathbb{P}[T_1 > t, L_1 > X] &\geq \mathbb{P}[T_1 > t, L_1 > X, L_1 > x_\epsilon] \\ &\geq \mathbb{P}[T_1 > t, L_1 > x_\epsilon, X < x_\epsilon] \\ &= \mathbb{P}[T_1 > t, L_1 > x_\epsilon]\mathbb{P}[X < x_\epsilon], \quad (33)\end{aligned}$$

where, choosing $x_\epsilon$ large enough, we can always make $\mathbb{P}[X < x_\epsilon] > 0$. Using the proof of Theorem 2, we know that, for any $0 < \epsilon < 1$, there exist $t_0$ and $b_0$ such that for $b > b_0$ and $t_0 < t < \bar{G}(b)^{-(1-\epsilon)}$,

$$\mathbb{P}[T_1 > t, L_1 > x_\epsilon] > \frac{1}{t^{(1-\epsilon)\alpha}},$$

which, by (33), implies,

$$\begin{aligned}\log \mathbb{E}[\underline{T}_1] &\geq \log \int_{t_0}^{\bar{G}(b)^{-(1-\epsilon)}} \frac{\mathbb{P}[X < x_\epsilon]}{t^{(1-\epsilon)\alpha}} dt \\ &\sim -(1-\epsilon)(1-(1-\epsilon)\alpha) \log \bar{G}(b), \quad (34)\end{aligned}$$

as $b \to \infty$. Using (34), (32), and passing $\epsilon \to 0$, we prove, as $b \to \infty$,

$$\overline{\lim_{n\to\infty}} \frac{\log \Lambda_n(b)}{\log \bar{G}(b)} \lesssim 1 - \alpha.$$

Next, we prove the lower bound. In each available period $A_{\Phi(i)}, i \geq 1$ that contains a successful transmission, we postpone the transmission of the new packet $L_{i+1}$ until the beginning of the next available period $A_{\Phi(i)+1}$. It is easy to see that, this construction increases the total transmission time, and also the durations for transmitting $L_i, i \geq 1$ are i.i.d. random variables. Thus, the law of large numbers can be applied. Based on similar arguments in deriving the upper bound, we can prove, as $b \to \infty$,

$$\underline{\lim}_{n\to\infty} \frac{\log \Lambda_n(b)}{\log \bar{G}(b)} \gtrsim 1 - \alpha.$$

Combining the lower and upper bounds, we complete the proof. ■

## References

[1] D. P. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.

[2] P. R. Jelenković and J. Tan, "Can retransmissions of superexponential documents cause subexponential delays?" in *Proceedings of IEEE INFOCOM'07*, Anchorage, Alaska, USA, 2007.

[3] ——, "Are end-to-end acknowlegements causing power law delays in large multi-hop networks?" in *14th Informs Applied Probability Conference*, Eindhoven, July 9-11 2007.

[4] ——, "Characterizing heavy-tailed distributions induced by retransmissions," Department of Electrical Engineering, Columbia University, EE2007-09-07, Tech. Rep., September 2007, eprint arXiv: 0709.1138v2.

[5] ——, "Is ALOHA causing power law delays?" in *Proceedings of the 20th International Teletraffic Congress,* Ottawa, Canada, June 2007*; Lecture Notes in Computer Science,* No 4516, pp. 1149-1160, Springer-Verlag, 2007.

[6] S. Asmussen, P. Fiorini, L. Lipsky, T. Rolski, and R. Sheahan, "Asymptotic behavior of total times for jobs that must start over if a failure occurs," *Mathematics of Operations Research*, vol. 33, no. 4, pp. 932–944, November 2008.

[7] T. Karagiannis, J.-Y. Le Boudec, and M. Vojnović, "Power law and exponential decay of inter contact times between mobile devices," in *MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2007, pp. 183–194.

[8] W. E. Leland, W. Willinger, M. S. Taqqu, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," *ACM SIGCOMM Computer Communication Review*, vol. 25, no. 1, pp. 202–213, 1995.

[9] P. Jacquet, A. M. Naimi, and G. Rodolakis, "Asymptotic delay analysis for cross-layer delay-based routing in ad hoc networks," *Advances in Multimedia*, p. 11, 2007, article ID 90879.

[10] N. A. Ali, E. Ekram, A. Eljasmy, and K. Shuaib, "Measured delay distribution in a wireless mesh network test-bed," *Computer Systems and Applications, ACS/IEEE International Conference on*, pp. 236–240, 2008.

[11] W.-B. Gong, Y. Liu, V. Misra, and D. Towsley, "Self-similarity and long range dependence on the Internet: a second look at the evidence, origins and implications," *Computer Networks*, vol. 48, pp. 377–399, June 2005.

[12] M. Rodrig, C. Reis, R. Mahajan, D. Wetherall, and J. Zahorjan, "Measurement-based characterization of 802.11 in a hotspot setting," in *Proceedings of ACM E-WIND '05*, Philadelphia, Pennsylvania, USA, 2005, pp. 5–10.

[13] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation*, G.-C. Rota, Ed. Cambridge University Press, 1987, vol. 27.

[14] M. Parulekar and A. Makowski, "$M/G/\infty$ input processes: a versatile class of models for network traffic," in *INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 2, April 1997, pp. 419–426.

[15] S. Asmussen, *Applied Probability and Queues*, 2nd ed. Springer-Verlag, 1987.