# Relay Load Balancing in Queued Cooperative Wireless Networks with Rateless Codes

Gaurav Bansal Vinod Sharma Neelesh B. Mehta Eitan Altman

***Abstract*—We consider a cooperative communication system where transmission is done via relay nodes using rateless codes. The relays have buffers that can queue messages that arrive from the source. While queuing the messages substantially increases the system throughput, it also increases the end-to-end message transmission times due to the additional queuing delays at the relay nodes. In this paper we suggest a novel load balancing scheme at the relay queues that exploits the unique properties of rateless codes. We show that it can substantially reduce the queuing delays while retaining most of the throughput. We provide – and verify with simulations – a theoretical analysis of this system, a simplifying and accurate approximation, and stability conditions for the system.**

***Index Terms*—Cooperative communications, queuing, relays, load balancing, relay selection, rateless codes, fading channels.**

## I. Introduction

In a cooperative relay communication system, a source collaborates with a relay to transmit information to a destination. Cooperation improves the system performance by exploiting the broadcast nature of the wireless channel from the source and its inherent spatial diversity. Given its promising gains, it has attracted considerable attention recently [1]–[5]. In case multiple candidate relays are present, relay selection methods can be employed to enable the source to select the most suitable relay on the basis of instantaneous channel conditions [6], [7].

Fountain codes and rateless codes, in general, have recently been shown to be well suited for cooperative relay networks [8]–[11]. In [8], a quasi-synchronous mode is considered, in which one or more relays that first receive the packet, transmit to the destination using rateless codes. This reduces the time required by the destination to receive the message. In [9], a rateless coding framework for space-time collaboration over a single relay channel was proposed.

Unlike conventional codes, which generate a finite number of parity bits, rateless codes generate an unbounded number of parity bits, which are transmitted until an acknowledgement is received from the recipient. The codes provide a natural way for accumulating information from one or more relays. With them, a receiver can recover the original information from *unordered* subsets of *one or more* rateless code-streams transmitted by multiple sources so long as the total mutual information accumulated marginally exceeds the entropy of the source information [12]–[14].

G. Bansal, V. Sharma, and N. B. Mehta are with the Dept. of Electrical Communication Eng. at the Indian Institute of Science (IISc), Bangalore, India. E. Altman is with INRIA, France.

The papers above approach the problem from a communication-theoretic perspective, in which the goal is to reduce the transmission time of the packet as much as possible so as to bolster throughput. It is implicitly assumed that the source transmits the next message only after the current message has been received by the destination. However, [15] showed that combining relay selection with buffering of messages at the relays dramatically increases the throughput of the system. This occurs because the source can start transmitting its next message earlier – and to the relay with the best channel condition – while the current one is en route to the destination. But, this gain comes at the expense of a larger average end-to-end (E-E) delay in the network due to queuing delays at the relays.

In this paper, we propose a novel method that reduces the queuing delays with only a marginal reduction in overall throughput. Thus, the significantly higher throughput of queuing and cooperation can be enjoyed with a much smaller increase in average delay. In our method, relays with longer queues reduce the rate at which they receive new packets by ignoring a fraction of the incoming rateless encoded bits. Specifically, a relay ignores incoming bits with a probability $p$ if its queue length exceeds a threshold of $q_{\text{th}}$ packets. Benefits are shown using both analysis and simulations.

Intuitively speaking, the method increases the odds that a relay with a smaller queue length and a good (but perhaps not the best) channel from the source is the first to receive the packet in its queue and send back an acknowledgement. It exploits the unique unordered binning property of rateless codes and also the spatial diversity inherent in a cooperative communication system. It succeeds because it balances the queuing load across relays and ensures that the source-relay (SR) link condition and the current queue length together determine the relay that receives the packet. In effect, the method performs relay selection implicitly on the basis of the above parameters (without requiring channel knowledge).

The paper is organized as follows. We describe the system model in Section II, and develop the analysis in Section III. Simulation results are presented in Section IV, and are followed by our conclusions in Section V.

## II. System Model

Shown in Fig. 1 is a two-hop network in which the source, $\mathcal{S}$, has a continuous stream of messages to transmit to the destination, $\mathcal{D}$, via $M$ decode-and-forward relays, $\mathcal{R}_1, \ldots, \mathcal{R}_M$. Each message has a bandwidth-normalized payload of $B$ nats/Hz. The signal transmitted by a source is received by multiple relays due to the broadcast nature of the wireless channel.

The wireless channels between different nodes are assumed to be indepdendent, frequency-flat, block-fading channels. The channel is assumed to be constant over the duration of transmission of the message; it changes to an independent value thereafter [8]. For rateless codes, this assumption is valid in low to medium mobility scenarios of current high data rate wireless systems. This analytically tractable model is commonly used in the cooperative rateless code literature mentioned above and, in general, in the analysis of adaptive rate systems with fixed message payloads.[1] To simplify analysis, we also assume that the direct source-destination link is very weak.

The source as well as the relays use different rateless codes to transmit their messages. When a relay receives a sufficient number of bits, it can successfully decode the message, which is then queued in the relay's buffer for transmission in a first come first serve fashion to the destination. The time taken by relay $i$ to decode a message from $\mathcal{S}$ is[2] $B/\log_e(1+\gamma_i)$ since the transmission rate is $\log_e(1+\gamma_i)$ bits/sec/Hz. The source transmits until it receives an acknowledgement back from *any one* of the relays. The other relays thereafter play no further role in the transmission of that message, and the source starts transmitting the next message. This was studied in [15] and was called Relay Selection (RS).

In order for the receiver to separate the various transmissions that occur in a decentralized manner, we assume that the relays and the source transmit their signals using different, a priori assigned, spreading sequences. It has been shown in [8] that a typical system can operate well with low spreading gains. We assume infinitely long buffers at each relay, to keep the analysis tractable. No channel knowledge is required at the transmitters, which is one of the advantages of using rateless codes. The relays are assumed to be full-duplex capable. This is practically feasible using a duplexer when the carrier frequency used for the SR channels is sufficiently apart from that for the relay-destination (RD) channels. The analysis for the half-duplex case is beyond the scope of this paper.

As seen from [15], the queueing analysis of the Relay Selection (RS) system over fading channels is quite involved. The modification we suggest in this paper further complicates the analysis because the queue state itself modulates the arrival rate into the queue. *Thus, for analytical tractability, we will initially assume that the service (transmission) times at the relay nodes are exponentially distributed.* However, the SR links still undergo fading. This assumption captures the essential aspects of the load balancing scheme proposed and provably demonstrates the gains obtained via this method. We shall refer to this as the *SR fading & RD exponential rate* scenario. Furthermore, this approach can be generalized to approximate fading with an arbitrary distribution by means of exponential mixture models [16]. The efficacy of the proposed method when the RD channels also undergo fading is demonstrated using simulations. We shall refer to this as the *SR fading & RD fading* scenario.

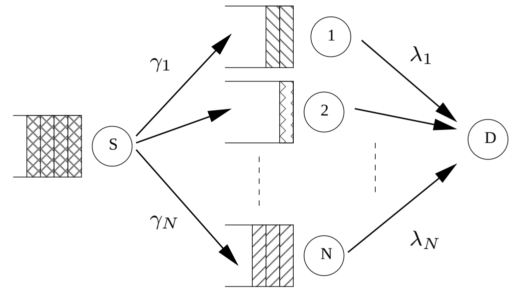


Fig. 1. Schematic of two-hop queued cooperative relay network

The key performance indices of this system are [15]: (i) *system throughput*, $\lambda$, which is the average rate at which the destination successfully receives messages, and (ii) *mean E-E delay* of a message from the source to the destination.

### A. Notation

Let $t_n$ be the epoch when the $n$th packet (message) arrives from the source to one of the relay queues. Then, $T_n = t_n - t_{n-1}$ is the time taken to transmit the $n$th message from the source to a relay. Let $q_n(i)$ be the queue length (including the message in service) at $\mathcal{R}_i$'s queue, $\mathcal{Q}_i$, at the time $t_n^+$. Let $\mathbf{q}_n = (q_n(1), ..., q_n(M))$. Let $r_n(i)$ and $r'(i)$ be the transmission rates of the $\mathcal{S}$–$\mathcal{R}_i$ and $\mathcal{R}_i$–$\mathcal{D}$ channels during duration $T_n$. $\mathbf{E}[A|B]$ and $P(A|B)$ denote the expectation and probability of event $A$ conditioned on event $B$.

### B. Load Balancing Method

The main reason for queueing delays in RS is the mismatch in the transmission rates of the $\mathcal{S}$–$\mathcal{R}_i$ and $\mathcal{R}_i$–$\mathcal{D}$ channels. The transmission rate of a $\mathcal{S}$–$\mathcal{R}_i$ channel may be high but that of the corresponding $\mathcal{R}_i$–$\mathcal{D}$ channel may be low, or vice versa.

In the load balancing scheme, each relay autonomously reduces the rate of its $\mathcal{S}$–$\mathcal{R}_i$ channel from $r_n(i)$ to $pr_n(i)$ when $q_n(i) \geq q_{th}$. Here, $0 < p < 1$ and $q_{th}$ are system parameters that need to be optimized. This reduction can be naturally done in a rateless code framework by simply letting the relay ignore, with probability $p$, the received signal corresponding to a coded bit. This increases the odds that the packet goes to a different relay with a smaller queue (even if it has a weaker $\mathcal{S}$–$\mathcal{R}_i$ channel). Intuitively, a larger value of $q_{th}$ increases the throughput but also the delays. Whereas, a larger $p$ reduces delays but also the throughput.

Note that the source needs to know neither the channel gains to the relays nor the queue lengths at the relays. Only a relay needs to know its *own* queue length at any time. Thus, no *extra signalling* is required. In general, the bit dropping probability $p$ and the queue thresholds $q_{th}$ can be different for different relays. But, for simplicity of notation and to illustrate the idea, we will keep these the same for all relays.

[1]In practice this is an approximation since successive messages sent by the same node will experience correlated channels when the channel gain is very high and the transmission durations are small.

[2]Inefficiencies that occur in any practical implementation of the rateless code can be factored into $B$ [8], and are thus not shown here.

## III. Analysis

Under our assumptions $\{r_n(i)\}$ will be independent, identically distributed (i. i. d.) sequences. With the exponential $\mathcal{R}_i$–$\mathcal{D}$ service time assumption, $\{\mathbf{q}_n\}$ becomes an embedded $M$-dimensional Markov chain. Furthermore, if $p$ is chosen appropriately, $\{\mathbf{q}_n\}$ becomes an irreducible, aperiodic Markov chain [16]. When the system is stable, $\mathbf{q}_n$ will have a unique stationary distribution $\pi$ (conditions for stability are discussed in the Appendix.) To compute system throughput and mean delay, $\mathbf{E}[D]$, we shall first compute $\pi$.

The transition matrix $\mathbf{P}$ of the Markov chain $\{\mathbf{q}_k\}$ can be constructed using the following observations. A detailed derivation and explicit characterization of $\mathbf{P}$ is omitted to conserve space. The $(n+1)$th packet from $\mathcal{S}$ will go to $\mathcal{Q}_i$ if its $\mathcal{S}$–$\mathcal{R}_i$ channel rate $\tilde{r}_n(i)$ exceeds that of all other relays, *i.e.*, $\tilde{r}_n(i) > \tilde{r}_n(k)$, $k \neq i$, where

$$\tilde{r}_n(i) = \begin{cases} r_n(i), & \text{if } q_n(i) < q_{\text{th}} \\ (1-p)r_n(i), & \text{if } q_n(i) \geq q_{\text{th}} \end{cases}. \quad (1)$$

Let $\tilde{r}_n^{\max} \triangleq \max_i(\tilde{r}_n(i))$. Consequently, the time taken for the source to transmit the $n$th packet equals

$$T_n = B/\tilde{r}_n^{\max}. \quad (2)$$

Conditioned on $T_n = t$, the number of possible packet departures from $\mathcal{Q}_i$ is a Poisson random variable with mean $r'(i)t$. Therefore, the probability that $m$ packet departures occur in $\mathcal{Q}_i$ equals $\frac{(r'(i)t)^m}{m!}e^{-r'(i)t}$. (To compute $\pi$ numerically, a truncation of the state space is required.)

### A. System Throughput

The expression for the system throughput depends on whether the queues are stable or not. We compute it for the two scenarios below.

*1) Stable System:* If the system is stable with stationary distribution $\pi$, the system throughput is given by

$$\lambda = \frac{1}{\mathbf{E}_\pi[T_n]} \quad \text{messages/sec.} \quad (3)$$

Here, $\mathbf{E}_\pi[T_n]$ is given by

$$\mathbf{E}_\pi[T_n] = \sum_{(m_1,\ldots,m_M)} \pi(m_1,\ldots,m_M) \times \mathbf{E}[T_n|\mathbf{q}_{n-1} = (m_1,\ldots,m_M)]. \quad (4)$$

From (2), $\mathbf{E}[T_n|\mathbf{q}_{n-1} = (m_1,\ldots,m_M)] = \mathbf{E}\left[\frac{B}{\tilde{r}_n^{\max}}\right]$. Note that $\tilde{r}_n(i)$, and consequently $\tilde{r}_n^{\max}$, depends on $r_n(i)$ and $q_n(i)$ as per (1).

The throughput, $\lambda_i$, of relay queue $\mathcal{Q}_i$, measured in terms of the number of messages that depart from $\mathcal{Q}_i$ every second, can now be written down in terms of $p_i$, the stationary probability that a packet transmitted by the source joins $\mathcal{Q}_i$, as follows:[3]

$$p_i = \sum \pi(m_1,\ldots,m_M)p[\tilde{r}_n^{\max} = \tilde{r}_n(i)|\mathbf{q}_n = (m_1,\ldots,m_M)]. \quad (5)$$

Then, $\lambda_i = p_i\lambda$, $i = 1,\ldots,M$.

[3]If more than one $\mathcal{Q}_i$ satisfies $\tilde{r}_n = \tilde{r}_n(i)$, then the message will go to any of these queues with equal probability.

*2) Unstable System:* The throughput calculation when the system is unstable is a little more involved since some (but not all) of the queues might still be stable.

If $\mathcal{Q}_i$ is unstable, then, under steady state, its queue length will be infinite. Hence, $\tilde{r}_n(i) = pr_n(i)$, for all large $n$. Otherwise, for a stable queue, $\tilde{r}_n(i)$ is given by (1). Let $U$ be the set of relay indices that have unstable queues. (see Appendix to identify $U$) To calculate the system throughput, we proceed as follows: (i) Each of these queues contributes a throughput $\lambda_i = r'(i)$, $i \in U$. (ii) Compute the stationary distribution $\pi'$ of only the remaining set of queues $U^C$. In this computation, the rate of the unstable queues is set as $\tilde{r}_n(i) = pr_n(i)$, for all $i \in U$. (iii) For $i \in U^C$, the throughput $\lambda_i$ of $\mathcal{Q}_i$ can then be computed as done above.

The system throughput in the unstable case then equals

$$\lambda = \sum_{i=1}^{M} \lambda_i.$$

### B. Mean End-to-End Delay

If any of the queues in the system is unstable then mean E-E delay is unbounded. Therefore, the E-E delay needs to be evaluated only for the case when the system is stable and has a stationary distribution $\pi$.

Let $\mathbf{E}[D_i]$ be the stationary mean E-E delay for a message that passed through $\mathcal{Q}_i$. Then, from the law of total probability, $\mathbf{E}[D] = \sum_{i=1}^M p_i\mathbf{E}[D_i]$. Also, $\mathbf{E}[D_i] = \mathbf{E}\left[\frac{B}{\tilde{r}_i}|\tilde{r}_i > \tilde{r}_j,\ j \neq i\right] + \mathbf{E}[W_i]$, where $\mathbf{E}[W_i]$ is the mean sojourn time of messages in $\mathcal{Q}_i$.

Next we compute $\mathbf{E}[W_i]$. From the exponential RD service time assumption, we have

$$\mathbf{E}[W_i] = \sum \pi(m_1,m_2,\ldots,m_M)P\big[\mathcal{R}_i \text{ is selected and next state is } (m_1',m_2',\ldots,m_M')|(m_1,\ldots,m_M)\big]\frac{m_i'}{r'(i)p_i}.$$

The above probability term that $\mathcal{R}_i$ is selected and the next state is $(m_1',\ldots,m_M')$ conditioned on the present state being $(m_1,\ldots,m_M)$ equals

$$P[\tilde{r}_n(i) > \tilde{r}_n(j), j \neq i|(m_1,\ldots,m_M)] \times P[(m_1',\ldots,m_M')|\tilde{r}_n(i) > \tilde{r}_n(j), j \neq i, (m_1,\ldots,m_M)].$$

If $m_1 + 1 = m_1' \geq 2$ and $0 < m_i' \leq m_i$, $2 \leq i \leq M$, the second term in the above expression (conditioned on $T_n = t$) equals $e^{-r'(1)t}\prod_{j=2}^M \frac{e^{-r'(j)t}(r'(j)t)^{m_j - m_j'}}{(m_j - m_j')!}$.

### C. Simplifying Approximation

The analysis above can be considerably simplified by considering an approximate system in which the queue evolution of different relays is decoupled. In this system, we focus on a specific relay (say relay 1), and assume that none of the other $M-1$ relays ever drop bits. Thus, decoupling between queues occurs because the arrival rate to a queue now no longer depends on the state of the other queues. This decoupling is advantageous because only a one-dimensional

Markov chain needs to be handled unlike the original system's $M$-dimensional chain. This change not only decouples the $\mathcal{Q}_1$ from other queues, but also reduces the arrival process to $\mathcal{Q}_1$. We use this to get a lower bound on mean sojourn time in $\mathcal{Q}_1$. Due to decoupling, this lower bound is much easier to compute. The above modified system also possibly increases the departure rate from the source. Thus, to get a lower bound on the mean transmission time from the source, we consider another system in which the relays never drop the bits. The mean transmission time from the source in this system is easy to compute and can be used along with the lower bound on mean sojourn time obtained above to obtain a good approximation to mean end to end time in the system as we will verify via simulations.

## IV. Simulation Results

We now plot the throughput and mean E-E delay as a function of the bit dropping probability, $p$, and the queue length threshold, $q_{\text{th}}$. The simulations assume a system bandwidth = 2MHz, message length = 4096 bits and 3 relays. We first study the behavior of the system for the SR fading and RD exponential rate scenario, and verify our analysis with Monte Carlo simulations. Here we take SR link SNR = 17 dB and mean RD service time = 1.6 ms which provide a system utilitzation of $\rho = 0.79$ at $p = 0$. Thereafter, we consider the more realistic SR fading and RD fading scenario where we assume a timeout of 10ms at the relays if no acknowledgement is received within this time. Here the SR link SNR is 13 dB and RD link SNR is 10 dB and hence $\rho = 0.80$ at $p = 0$. The fading distribution is rayleigh.

### *A. SR Fading & RD Exponential Rate Scenario*

Figure 2 plots the throughput as a function of $p$ for different values of $q_{\text{th}}$. The throughput decreases as either $p$ increases or $q_{\text{th}}$ decreases since the source transmission time increases. This occurs because for larger values of $q_{\text{th}}$, the odds that the number of packets in the queue exceeds $q_{\text{th}}$, and consequently the probability that a relay accepts packets at a reduced rate, decreases. The simulations results, which are shown using markers ($\diamond$, etc.), match the analysis curves (–) well.

The corresponding E-E delay is studied in Figure 3. As $p$ increases, the E-E delay starts decreasing because the queue lengths and the queuing delays decrease. As expected, the smaller the value of $q_{\text{th}}$, the larger the decrease (except for small p). However, for small $q_{\text{th}}$, the E-E delay increases again for larger $p$ values since the source transmission time increases and becomes the dominant term in the E-E delay. The simulation results, which are shown using markers, again match the analysis curves (lines) well. The figure also plots, using dashed lines (- -) the results from the simplifying approximation of Sec. III-C. It can be seen that it is reasonably accurate.

The two figures together show that load balancing reduces the E-E delay substantially with only a marginal decrease in throughput. For example, at $p = 0.28$ and $q_{\text{th}} = 2$, the E-E delay decreases by 52% (while the throughput decreases by 10%) compared to the case without load balancing ($p = 0$). The reduction in throughput is, in fact, a minor issue since introducing queuing at the relays itself leads to throughput gains of 100% or more [15].

### *B. SR Fading & RD Fading Scenario*

We now study, using simulations, the realistic scenario in which the RD links also undergo Rayleigh fading (in which case the relay transmission time is not exponentially distributed). As before, Figures 4 and 5 plot the throughput and E-E delay as a function of $p$ for different $q_{\text{th}}$.

The behavior in this case turns out to be the same as the Rayleigh-Exponential case analyzed and simulated earlier. Now, a 54% decrease in the E-E delay can be achieved with a 10% reduction in throughput at $p = 0.28$ and $q_{\text{th}} = 2$.

## V. Conclusions

We proposed a relay load balancing method in which a relay autonomously reduces the rate of arrival of packets into its queue by randomly ignoring a fraction of the coded bits that arrive at its receiver. We showed that over a wide range of operating points such a system fully reaps the benefits of spatial diversity provided by multiple relays since its throughput decreases marginally while the end-to-end delays decrease substantially. In addition to deriving an exact analysis that used embedded Markov chains, we also developed a good approximation that required considerably simpler analysis. While the ideas in this paper consider a basic two-hop network, they are useful for multihop networks as well. Generalizing these ideas to the multihop scenario is an interesting avenue for further research. Another interesting analysis aspect is to derive tight upper and lower bounds for the general set up in which the source-relay and relay-destination channels undergo fading according to an arbitrary distribution.

## Appendix

*Symmetric Case Stability*: A system is *symmetric* if $r(i)$ and $r'(i)$ have the same statistics as $r(j)$ and $r'(j)$, respectively, for all $i$ and $j$. Then a necessary and sufficient condition for stability is

$$(1-p)\mathbf{E}\left[\max_i r(i)\right] < M\mathbf{E}\left[r'(i)\right]. \quad (6)$$

The left hand side then is the maximum rate at which the source can pump messages. The right hand side is the maximum throughput obtainable by all the RD channels, since near the stability boundary, all queues will exceed $q_{\text{th}}$ with a large probability. The above result also implies that any system can be stabilized by choosing a large enough $p$.

*General Asymmetric Case Stability*: Obtaining necessary and sufficient conditions for stability in the general asymmetric case is difficult. However, sufficient conditions for individual queues can be derived.

For a relay $i$, the original system's time evolution can be written as:

$$q_{n+1}(i) = (q_n(i) - X_n(i))^+ + Y_n(i), \quad (7)$$

where $X_n(i)$ is the potential number of packets that can be transmitted from $\mathcal{Q}_i$ in time $T_n$ and $Y_n(i)$ is the number of arrivals at time $t_n$ at $\mathcal{Q}_i$. (From the definition of the embeded Markov chain, it follows that $Y_n(i) \leq 1$.)

Now conjure a new system whose evolution equation is

$$q'_{n+1}(i) = (q'_n(i) - X'_n(i))^+ + Y'_n(i), \quad (8)$$

where the departures from the source correspond to the case where each relay *does not* drop any bit and $T'_n$ denotes the time taken to transmit the $n$th message from the source to any relay. And, $Y'_n(i)$ is different from $Y_n(i)$ in that it is obtained from another scenario where $\mathcal{Q}_i$ never drops packets but other queues always drop packets with probability $p$.

One can then show that $T' \leq_{\text{st}} T$ since the transmission rate is higher for the conjured system.[4] Similarly, $Y'(i) \geq_{\text{st}} Y(i)$. It then follows that in this new system, $\mathcal{Q}_i$ is a GI/GI/1 queue for which $q'_n(i) \geq_{\text{st}} q_n(i)$ for each n. This will also hold for the limiting distribution if it exists. Thus, if the $i^{\text{th}}$ queue of the new system is stable, so will the one in the original system.

The stability region for each queue in the new system can be easily computed since $T'_n$, $X'_n(i)$ and $Y'_n(i)$ are all i. i. d.. The detailed derivation is not given for want of space.

## References


[1] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, pp. 68–73, 2004.

[2] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1099–1109, 2004.

[3] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, vol. 52, pp. 1470–1476, Sept. 2004.

[4] I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1080–1088, Aug. 2004.

[5] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 659–672, Mar. 2006.

[6] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3066–3078, Aug. 2008.

[7] D. S. Michalopoulos and G. K. Karagiannidis, "PHY-layer fairness in amplify and forward cooperative diversity systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1073–1083, Mar. 2008.

[8] A. F. Molisch, N. B. Mehta, J. Yedidia, and J. Zhang, "Performance of fountain codes in collaborative relay networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 4108–4119, Nov. 2007.

[9] J. Castura and Y. Mao, "Rateless coding for wireless relay channels," in *Proc. IEEE Intl. Symp. Inf. Th.*, 2005.

[10] A. Eckford, J. Chu, and R. Adve, "Low-complexity cooperative coding for sensor networks using rateless and LDGM codes," in *Proc. ICC*, pp. 1537–1542, Jun. 2006.

[11] R. Nikjah and N. C. Beaulieu, "Achievable rates and fairness in rateless coded decode-and-forward half-duplex and full-duplex opportunistic relaying," in *Proc. ICC*, pp. 3701–3707, 2008.

[12] A. Shokrollahi, "Raptor codes," in *Proc. IEEE Intl. Symp. Inf. Th.*, p. 36, 2004.

[13] M. Mitzenmacher, "Digital fountains: A survey and look forward," in *Proc. IEEE Inform. Th. Workshop*, pp. 271–276, Oct. 2004.

[14] R. Palanki and J. Yedidia, "Rateless codes on noisy channels," in *Proc. IEEE Intl. Symp. Inf. Th.*, p. 37, 2004.

[15] N. B. Mehta, V. Sharma, and G. Bansal, "Queued cooperative wireless networks with rateless codes," *Proc. Globecom*, 2008.

[16] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Prentice Hall, 1989.


[4] A sequence of i. i. d. random variables $\{X_n\}$ is said to be stochastically smaller than another i. i. d. sequence $\{Y_n\}$ if $P(X_n \leq t) \geq P(Y_n \leq t)$. We shall denote this by $X \leq_{\text{st}} Y$. One can similarly define $X \geq_{\text{st}} Y$.

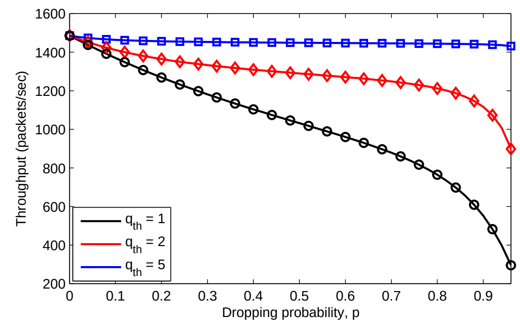


Fig. 2. SR Fading & RD Exponential Rate Scenario: Throughput

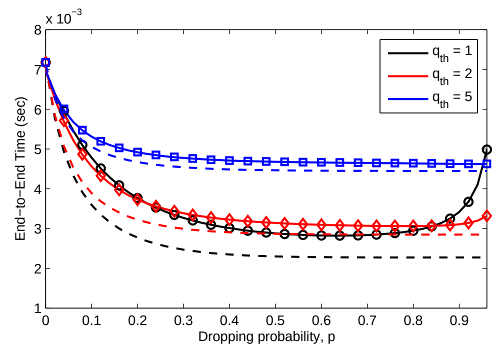


Fig. 3. SR Fading & RD Exponential Rate Scenario: End-To-End Delay

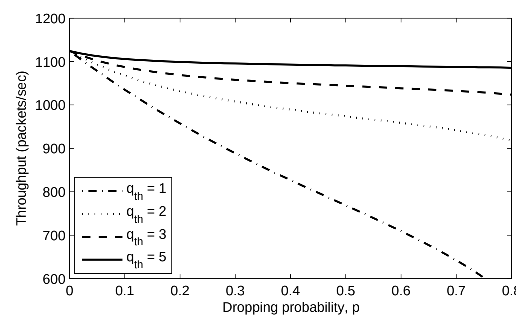


Fig. 4. SR Fading & RD Fading Scenario: Throughput

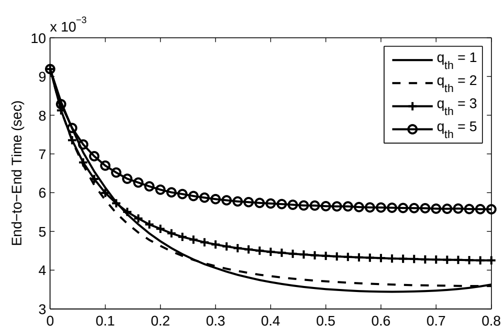


Fig. 5. SR Fading & RD Fading Scenario: End-To-End Delay