Coding Achieves the Optimal Delay-Throughput Trade-off in Mobile Ad-Hoc Networks: Two-Dimensional I.I.D. Mobility Model with Fast Mobiles

Lei Ying, Sichao Yang and R. Srikant
Coordinated Science Lab
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
{lying,syang8,rsrikant}@uiuc.edu

***Abstract*— In this paper, we investigate the delay-throughput trade-off in mobile ad-hoc networks under two-dimensional i.i.d. mobility model with fast mobiles, and show that the optimal trade-off can be achieved using rate-less codes. Given a delay constraint D, we first prove that the maximum throughput per source-destination (S-D) pair is $O\left(\sqrt{D/n}\right)$, and then propose a joint coding-scheduling algorithm to achieve the maximum throughput. The result can be extended to two-dimensional i.i.d. mobility model with slow mobiles, one-dimensional mobility models, and hybrid random walk mobility models.**

I. Notations

The following notations are used throughout this paper, given non-negative functions $f(n)$ and $g(n)$:

(1) $f(n) = O(g(n))$ means there exist positive constants c and m such $f(n) \leq cg(n)$ for all $n \geq m$.
(2) $f(n) = \Omega(g(n))$ means there exist positive constants c and m such that $f(n) \geq cg(n)$ for all $n \geq m$. Namely, $g(n) = O(f(n))$.
(3) $f(n) = \Theta(g(n))$ means that both $f(n) = \Omega(g(n))$ and $f(n) = O(g(n))$ hold.
(4) $f(n) = o(g(n))$ means that $\lim_{n\to\infty} f(n)/g(n) = 0$.
(5) $f(n) = \omega(g(n))$ means that $\lim_{n\to\infty} g(n)/f(n) = 0$. Namely, $g(n) = o(f(n))$.

II. Introduction

The throughput of a random wireless network with n static nodes and n random S-D pairs was studied by Gupta and Kumar [9]. They showed that the maximum throughput per S-D pair is $O(1/\sqrt{n})$, and proposed a scheduling scheme achieving a throughput of $\Theta(1/\sqrt{n\log n})$ per S-D pair. The throughput decreases with n because each successful transmission from source to destination needs to take $\sqrt{n/\log n}$ hops. Later Grossglauser and Tse [8] considered mobile ad-hoc networks, and showed that $\Theta(1)$ throughput per S-D pair is achievable. The idea is to deliver a packet to its destination only when it is within distance $\Theta(1/\sqrt{n})$ from the destination. However, packets have to tolerate large delays to achieve this throughput.

We first review the results for i.i.d. mobility models. Neely and Modiano [13] studied the i.i.d. mobility model where the positions of nodes are totally reshuffled from one time slot to another, and showed that the mean delay of Grossglauser and Tse's algorithm is $\Theta(n)$. In the same paper, they also proposed an algorithm which generates multiple copies of each source packet to reduce the mean delay. Since more transmissions are required when we generate multiple copies, the throughput per S-D decreases with the number of copies per source packet. The delay-throughput trade-off is shown to be $\lambda = \Omega(D/n)$ in [13], where λ is the throughput per S-D pair, and D is the number of time slots taken to deliver packets from source to destination.

In [13], fast mobility is assumed. A different time-scale of mobility, slow mobility, was considered by Toumpis and Goldsmith in [18], and Lin and Shroff in [10]. For slow mobiles, node mobility is assumed to be much slower than data transmissions. So the packet size can be scaled down as n increases, and multi-hop transmissions are feasible in single time slot. The delay-throughput trade-off was shown to be $\lambda = \Omega\left(\sqrt{D/n}\log n\right)$ in [18]. The trade-off was improved in [10], where the maximum throughput per S-D pair for mean delay D was shown to be $\lambda = O\left(\sqrt[3]{D/n}\log n\right)$, and a scheme was proposed to achieve a trade-off of $\lambda = \Theta\left(\sqrt[3]{D/\left(n\log^{9/2} n\right)}\right)$.

Besides the i.i.d. mobility model, other mobility models have also been studied in the literature. The random walk model was introduced by El Gamal *et al* in [4], and later studied in [5], [6] and [16]. In [5] and [6], the throughput per S-D pair is shown to be $\Theta(1/\sqrt{n\log n})$ for $D = O(\sqrt{n/\log n})$, and $\Theta(D/n)$ for $D = \Omega(\sqrt{n/\log n})$, where [5] focused on the slow mobility and [6] focused on the fast mobility. Other mobility models, like Brownian motion, one dimensional mobility, and hybrid random walk models have been studied in [11], [3], [7] and [16].

The research was supported by a Vodafone Fellowship.

Although the delay-throughput trade-off has been widely studied for various mobility models, the optimal delay-throughput trade-off has not yet been established except for two cases of mobility models [5], [6], [11]. In this paper, we investigate ad-hoc networks with the two-dimensional i.i.d. mobility with fast mobiles. We first show that the maximum throughput per S-D pair is $O\left(\sqrt{D/n}\right)$ under a delay constraint D, and then propose a joint coding-scheduling algorithm to achieve the maximum throughput for D is both $\omega(\sqrt[3]{n})$ and $o(n)$. The result can be extended to slow mobility case, and other mobility models including one-dimensional mobility models and hybrid random walk models. The extensions can be found in [19] and [20].

We also would like to mention that there is a very recent result by Ozgur, Leveque, and Tse [14] where they showed a throughput of $\Theta(1)$ per S-D pair is achievable using node cooperation and MIMO communication; see also the earlier paper by Aeron and Saligrama in [1]. These schemes require sophisticated signal processing techniques, not considered in this paper.

The remainder of the paper is organized as follows: In Section III, we introduce the communication and mobility model. Main results along with some intuition into them are presented in Section IV. Then we analyze the two-dimensional i.i.d. mobility models with fast mobiles in Section V. Finally, the conclusions is given in Section VI. In the appendix, we collect some results that are frequently used in the paper.

III. Model

In this section, we first present the mobility and wireless interference models used in this paper. Then the definitions of delay and throughput are provided.

Mobile Ad-Hoc Network Model: Consider an ad-hoc network where wireless mobile nodes are positioned in a unit square. Assuming the time is slotted, we study the two-dimensional i.i.d. mobility model in this paper, which was introduced in [13] and defined as follows:

(i) There are n wireless mobile nodes positioned on a unit square. At each time slot, the nodes are uniformly, randomly positioned in the unit square.
(ii) The node positions are independent of each other, and independent from time slot to time slot. So the nodes are totally reshuffled at each time slot.
(iii) There are n S-D pairs in the network. Each node is both a source and a destination. Without loss of generality, we assume that the destination of node i is node $i+1$, and the destination of node n is node 1.

Communication Model: We assume the protocol model introduced in [9] in this paper. Let $\text{dist}(i,j)$ denote the Euclidean distance between node i and node j, and r_i to denote the transmission radius of node i. A transmission from node i can be successfully received at node j if and only if following two conditions hold:

(i) $\text{dist}(i,j) \leq r_i$;
(ii) $\text{dist}(k,j) \geq (1+\Delta)r_i$ for each node $k \neq i$ which transmits at the same time, where Δ is a protocol-specified guard-zone to prevent interference.

We further assume that at each time slot, at most W bits can be transmitted in a successful transmission.

Fast mobility: The mobility of nodes is at the same time-scale as the data transmission, so W is a constant independent of n and only one-hop transmissions are feasible in single time slot.

Delay and Throughput: We consider hard delay constraints in this paper. Given a delay constraint D, a packet is said to be successfully delivered if the destination obtains the packet within D time slots after it is sent out from the source.

Let $\Lambda_i[T]$ denote the number of bits successfully delivered to the destination of node i in time interval $[0,T]$. A throughput of λ per S-D pair is said to be feasible under the delay constraint D and loss probability constraint $\varepsilon > 0$ if there exists n_0 such that for any $n \geq n_0$, there exists a coding/routing/scheduling algorithm with the property that each bit transmitted by a source is received at its destination with probability at least $1-\varepsilon$, and

$$\lim_{T\to\infty} \Pr\left(\frac{\Lambda_i[T]}{T} \geq \lambda,\ \forall\ i\right) = 1. \tag{1}$$

IV. Main Results and Some Intuition

Recall that our objective is to maximize throughput in a wireless network subject to a delay constraint and a wireless interference constraint. More precisely, the constraints can be viewed as follows:

(1) Wireless interference: Throughput is limited due to the fact that transmissions interfere with each other.
(2) Mobility: A packet may not be delivered to its destination before the delay deadline since neither the packet's source nor any relay node may get close enough to the destination.

In this section, we present some heuristic arguments to obtain an upper bound on the maximum throughput subject to these two constraints and derive the key results of the paper. While the heuristic is far from precise derivations of the optimal delay-throughput trade-offs, it may be useful to the reader in understanding the main results. In addition, the heuristic argument provides the right order for the "hitting distance" (to be defined later) which plays a critical role in the optimal scheme used to achieve the delay-throughput trade-off.

Consider the two-dimensional i.i.d. mobility model with fast mobiles. We say that a packet *hits* its destination at time slot t if the distance between the packet and its destination is less than or equal to L. Under the two-dimensional i.i.d. mobility model, a packet hits its destination with probability πL^2 at each time slot. So given a delay constraint D, the probability that a packet hits its destination in one of D time slots is

$$1-\left(1-\pi L^2\right)^D.$$

Furthermore under the fast mobility, only one-hop transmissions are feasible at each time slot. So the transmission radius needs to be at least L to deliver packets to the destinations when their distance is L. Assume all nodes use a common transmission radius L and that all nodes wish to transmit at each time slot, then each node has $1/(c_1nL^2)$ fraction of time to transmit, and the throughput per S-D pair is no more than $1/(c_1nL^2)$ where c_1 is a positive constant independent of n. Thus the network can be regarded as a system where there are two virtual channels between each S-D pair as in Figure 1. The packets are first sent over the erasure channel with erasure probability

$$P_e = (1-\pi L^2)^D,$$

and then over the reliable channel with rate

$$R = \frac{1}{c_1L^2n}$$

bits per time slot. Each source can transmit at most W bits per time slot on average. So in this virtual system, the maximum throughput of a S-D pair is

$$\begin{aligned}\lambda &= \max_L \min\left\{W\left(1-\left(1-\pi L^2\right)^D\right), \frac{1}{c_1L^2n}\right\} \\ &= \sqrt{\frac{\pi W D}{c_1 n}},\end{aligned}$$

and the corresponding optimal hitting distance $L^* = b_1/\sqrt[4]{nD}$ where $b_1 = \sqrt[4]{c_1\pi W}$.

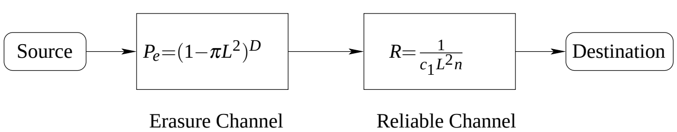

Fig. 1. Virtual-channel Representation for the Tow-Dimensional I.I.D. Mobility Model with Fast Mobiles

To achieve this throughput, we first need to use the optimal L. Furthermore, *a coding scheme achieving the capacity of the erasure channel* is needed. Since the erasure probability is determined by L and D, which are different under different delay constraints, rate-less codes become a reasonable choice. The key idea in this paper is to encode source packets using Raptor codes, which are near optimal rate-less codes with low complexity. We also note that the idea of using coding to improve reliability of packet delivery has also been considered by Shah and Shakkottai in [15] for ad hoc sensor networks in a different context. Our main result is as follows.

Main Result: Under the two-dimensional i.i.d. mobility model with fast mobiles, the throughput per S-D pair is $\lambda = O\left(\sqrt{D/n}\right)$ given a delay constraint D. For D is both $\omega(\sqrt[3]{n})$ and $o(n)$, this throughput can be achieved using a joint coding-scheduling algorithm.

Note that the heuristic argument leading up to the above result has many flaws. For example, it suggests that one can wait for the source to hit the destination to deliver the packet. In reality, such a scheme will not work since we deliver only one packet to the destination during each encounter between the S-D pair. Thus other packets at the source which are not delivered may violate their delay constraints. This problem in the heuristic argument is due to the fact that it assumes that we have an independent erasure channel for each packet despite the fact that the transmitting node is the same source. Despite the flaws, the heuristic argument surprisingly captures the delay-throughput trade-off and the optimal hitting distance correctly up to the right order. We use the virtual channel model to provide some insight into the results and the hitting distance we should use in the achievable algorithms, and rigorous proofs of the main results are provided in subsequent sections. We will show that the bound is achievable by exploiting the broadcast nature of the wireless channel to transmit each packet to several relay nodes and allowing relay nodes to independently attempt to deliver the packet the destination.

V. Two-Dimensional I.I.D. Mobility Model, Fast Mobiles

In this section, we investigate the two-dimensional i.i.d. mobility model with fast mobiles. Assuming that all mobiles have wireless communication and coding capability, we investigate the maximum throughput the network can achieve by using relaying and coding to recover packet loss as discussed in the heuristic arguments. Given a delay constraint D, we will first prove that the maximum throughput per S-D pair which can be supported by the network is $O\left(\sqrt{D/n}\right)$. Then a joint coding-scheduling scheme will be proposed to achieve the maximum throughput when D is both $\omega(\sqrt[3]{n})$ and $o(n)$.

A. Upper Bound

In this subsection, we show the maximum throughput the network can support without network coding, i.e., under the following assumption.

Assumption 1: Packets destined for different nodes cannot be encoded together.

Assumption 1 is the only significant restriction imposed on coding/routing/scheduling schemes. We also make the following two assumptions.

Assumption 2: Nodes have an infinite buffer, and store all packets they receive.

Assumption 3: A new coded packet is generated right before the packet is sent out. The node generating the coded packet does not store the packet in its buffer.

Assumption 3 is not restrictive since the information contained in the new packet is already available at the node.

Next we introduce following notations which will be used in our proof.

- b : Index of a bit stored in the network. Bit b could be either a bit of a source packet or a bit of a coded packet.
- d_b : The destination of bit b.
- c_b : The node storing bit b.
- t_b : The time slot at which bit b is generated.
- S_b : If bit b is delivered to its destination, then S_b is the transmission radius used to deliver b.
- $\mathscr{R}[T]$: The set of all bits stored at relay nodes at time slot T.
- $\Lambda[T]$: $\Lambda[T] = \sum_{i=1}^n \Lambda_i[T]$.

Assume that the delay constraint is D, and a source packet is processed by the source node at time slot t_p. Then the source packet is said to be active from time slot t_p to t_p+D-1. A bit b is said to be active if at least one source packet encoded into the packet containing bit b has not expired. It is easy to see that any bit expires at most D time slots after the bit is generated. Also a bit is said to be good if it is active when delivered to its destination. Now let $\tilde{\Lambda}[T]$ denote the number of good bits delivered to destinations in $[0,T]$. Without loss of generality, we assume good bits are indexed from 1 to $\tilde{\Lambda}[T]$. Note that expired bits might help decode good source bits but would not contribute to the total throughput, so we have

$$\tilde{\Lambda}[T] \geq \Lambda[T],$$

where $\Lambda[T]$ is the number of source bits successfully recovered at destinations.

Next we show three fundamental constraints. Inequalities (2) and (3) hold since the total number of bits transmitted or received in T time slots cannot exceed nWT. Inequality (4) holds since under the protocol model, discs of radius $\Delta r_i/2$ around the receivers should be mutually disjoint from each other.

Lemma 1: For any mobility model, the following inequalities hold,

$$\tilde{\Lambda}[T] \leq nWT \tag{2}$$

$$|\mathscr{R}[T]| \leq nWT \tag{3}$$

$$\sum_{b=1}^{\tilde{\Lambda}[T]} \frac{\Delta^2}{16}(S_b)^2 \leq \frac{WT}{\pi}, \tag{4}$$

where $|\mathscr{R}[T]|$ is the cardinality of the set $\mathscr{R}[T]$.

Proof: Since each node can transmit at most W bits per time slot, the total number of bits transmitted in T time slots is less than nWT which implies inequalities (2) and (3). Inequality (4) was proved in [2]. ■

We first consider the scenario where packet relaying is not allowed, i.e., packets need to be directly transmitted from sources to destinations. In the following lemma, we show that the throughput in this case is at most $\Theta(1/\sqrt{n})$ even without the delay constraint.

Lemma 2: Consider the two-dimensional i.i.d. mobility model with fast mobiles. Suppose that packets have to be directly transmitted to destinations from sources, then given $f(n) = \omega(1/\sqrt{n})$ and $\varepsilon > 0$, there exists n_0 such that for any $n \geq n_0$ and any $T > 0$,

$$\Pr(\Lambda[T] \geq nTf(n)) < \varepsilon. \tag{5}$$

Proof: First we have $\tilde{\Lambda}[T] \geq nTf(n)$ if $\Lambda[T] \geq nTf(n)$. Then from the Cauchy-Schwarz inequality and Lemma 1, we have

$$\begin{aligned}\left(\sum_{b=1}^{nTf(n)} S_b\right)^2 &\leq \left(\sum_{b=1}^{nTf(n)} 1\right)\left(\sum_{i=1}^{n}\sum_{b=1}^{nTf(n)} (S_b)^2\right) \\ &\leq \left(\sum_{b=1}^{nTf(n)} 1\right)\left(\sum_{b=1}^{\tilde{\Lambda}[T]} (S_b)^2\right) \\ &\leq nTf(n)\frac{16WT}{\pi\Delta^2},\end{aligned}$$

which implies

$$\sum_{b=1}^{nTf(n)} S_b \leq \frac{4T\sqrt{Wnf(n)}}{\Delta\sqrt{\pi}}.$$

Letting $\Gamma_{f(n)}$ denote the event that $\Lambda[T] \geq nTf(n)$, we have

$$E\left[\sum_{b=1}^{nTf(n)} S_b \middle| \Gamma_{f(n)}\right] \leq \frac{4T\sqrt{Wnf(n)}}{\Delta\sqrt{\pi}}. \tag{6}$$

Next define

$$L = \sqrt{\frac{\varepsilon f(n)}{4\pi W}}.$$

Under the two-dimensional i.i.d. mobility model, we have that for any i, j and t,

$$\Pr(\mathrm{dist}(i,j)(t) \leq L) = \pi L^2,$$

which implies that

$$E\left[\sum_{t=1}^{T}\left(1_{\mathrm{dist}(n,1)(t)\leq L} + \sum_{i=1}^{n-1} 1_{\mathrm{dist}(i,i+1)(t)\leq L}\right)\right] = \frac{\varepsilon nTf(n)}{4W}.$$

According to the Markov inequality, we thus have

$$\Pr\left(\sum_{t=1}^{T}\left(1_{\mathrm{dist}(n,1)(t)\leq L} + \sum_{i=1}^{n-1} 1_{\mathrm{dist}(i,i+1)(t)\leq L}\right) \geq \frac{nTf(n)}{2W}\right) \leq \frac{\varepsilon}{2}.$$

Let $\tilde{\Gamma}$ denote the event that

$$\sum_{t=1}^{T}\left(1_{\mathrm{dist}(n,1)(t)\leq L} + \sum_{i=1}^{n-1} 1_{\mathrm{dist}(i,i+1)(t)< L}\right) \leq \frac{nTf(n)}{2W}.$$

Note that at one time slot, at most W bits can be transmitted from a source to its destination. So if $\Gamma_{f(n)}$ and $\tilde{\Gamma}$ both occur, at least $nTf(n)/2$ good bits are transmitted over a distance at least L, and

$$\begin{aligned}E\left[\sum_{b=1}^{nTf(n)} S_b \middle| \Gamma_{f(n)}\right] &\geq \Pr(\tilde{\Gamma}|\Gamma_{f(n)}) E\left[\sum_{b=1}^{nTf(n)} S_b \middle| \Gamma_{f(n)}, \tilde{\Gamma}\right] \\ &\geq \Pr(\tilde{\Gamma}|\Gamma_{f(n)}) \frac{nTf(n)}{4}\sqrt{\frac{\varepsilon f(n)}{\pi W}}.\end{aligned} \tag{7}$$

Since $f(n) = \omega(1/\sqrt{n})$, there exists n_0 such that for any $n \geq n_0$,

$$f(n) > \frac{32W}{\Delta\varepsilon\sqrt{\varepsilon}}\frac{1}{\sqrt{n}}. \tag{8}$$

Now assume that for $n \geq n_0$, we have

$$\Pr(\Lambda[T] \geq nTf(n)) = \Pr(\Gamma_{f(n)}) \geq \varepsilon.$$

Then

$$\begin{aligned}\Pr(\tilde{\Gamma}|\Gamma_{f(n)}) &= \frac{\Pr\left(\tilde{\Gamma}\right)-\Pr\left(\tilde{\Gamma},\Gamma^c_{f(n)}\right)}{\Pr\left(\Gamma_{f(n)}\right)} \\ &\geq \Pr\left(\tilde{\Gamma}\right)-\Pr\left(\Gamma^c_{f(n)}\right) \\ &\geq 1-\frac{\varepsilon}{2}-(1-\varepsilon) \\ &= \frac{\varepsilon}{2}. \end{aligned} \tag{9}$$

From inequalities (6), (7) and (9), we can conclude that for $n \geq n_0$,

$$\frac{\varepsilon nTf(n)}{8}\sqrt{\frac{\varepsilon f(n)}{\pi W}} \leq \frac{4T\sqrt{Wnf(n)}}{\Delta\sqrt{\pi}},$$

which implies

$$f(n) \leq \frac{32W}{\Delta\varepsilon\sqrt{\varepsilon}}\frac{1}{\sqrt{n}},$$

and contradicts inequality (8). So lemma holds. ■

Next we investigate the maximum throughput the network can support using coding/routing/scheduling schemes. We have obtained an upper bound on the number of bits directly transmitted from sources to destinations in Lemma 2. To bound the maximum throughput with relaying, we will calculate the number of bits transmitted from relays to destinations in the following analysis. The reason we treat active bits at relays and active bits at sources differently is that we can bound the number of active bits at relays using inequality (3), but the number of active bits at sources could be infinitely many in the case where the data reservoirs at the sources are infinite.

Theorem 3: Consider the two-dimensional i.i.d. mobility model with fast mobiles. Given a delay constraint D, $f(D,n)=\omega(\sqrt{D/n})$ and ε, there exits n_0 such that for any $n \geq n_0$ and $T>0$,

$$\Pr\left(\Lambda[T] \geq nTf(D,n)\right) < \varepsilon. \tag{10}$$

Proof: Let $\tilde{\Lambda}^s[T]$ denote the number of good bits from sources to destinations, and $\tilde{\Lambda}^r[T]$ denote the number of good bits from relays to destinations. Since

$$\frac{f(D,n)}{9} = \omega(\sqrt{D/n}) = \omega(\sqrt{1/n}),$$

from Lemma 2, we have that there exists $\tilde{n}_0$ such that for any $n \geq \tilde{n}_0$,

$$\Pr\left(\tilde{\Lambda}^s[T] > \frac{nTf(D,n)}{9}\right) < \frac{\varepsilon}{8}.$$

Also note that $f(D,n)=\omega(\sqrt{D/n})$ implies that there exists $\tilde{n}_1$ such that for any $n \geq \tilde{n}_1$,

$$f(D,n) > \frac{144\sqrt{2}W}{\Delta\varepsilon\sqrt{\varepsilon}}\sqrt{\frac{D}{n}}. \tag{11}$$

Now let $n_0=\max\{\tilde{n}_0,\tilde{n}_1\}$, and assume that for $n \geq n_0$,

$$\Pr\left(\Lambda[T] > nTf(D,n)\right) \geq \varepsilon$$

holds. Then we first have that

$$\begin{aligned} &\Pr\left(\tilde{\Lambda}^r[T] > \frac{8}{9}nTf(D,n)\right) \\ \geq\ &\Pr\left(\tilde{\Lambda}[T] > nTf(D,n)\right) - \Pr\left(\tilde{\Lambda}^s[T] > \frac{nTf(D,n)}{9}\right) \\ \geq\ &\frac{7\varepsilon}{8}. \end{aligned}$$

Without loss of generality, assume that bit 1 to bit $\Lambda^r[T]$ are the good bits from relays. Let $\Gamma_{f(D,n)}$ denote the event that

$$\tilde{\Lambda}^r[T] \geq \frac{8}{9}nTf(D,n).$$

Similar to inequality (6), it can be first shown that

$$E\left[\sum_{b=1}^{8nTf(n)/9} S_b \middle| \Gamma_{f(D,n)}\right] \leq \frac{4T\sqrt{Wnf(n)}}{\Delta\sqrt{\pi}}. \tag{12}$$

Next let $\tilde{L}_b$ denote the minimum distance between node d_b and node c_b from time slot t_b to time slot t_b+D-1, i.e.,

$$\tilde{L}_b = \min_{t_b \leq t \leq t_b+D-1} \text{dist}(d_b,c_b)(t).$$

Then define

$$L=\sqrt{\frac{\varepsilon f(D,n)}{2W\pi D}},$$

we have that for any bit $b \in \mathscr{R}[T]$,

$$\Pr\left(\tilde{L}_b \leq L\right) = 1-(1-\pi L^2)^D \leq \pi L^2 D,$$

which implies

$$E\left[\sum_{b\in\mathscr{R}[T]} 1_{\tilde{L}_b \leq L}\right] \leq nWT\pi L^2 D = \frac{\varepsilon nTf(D,n)}{2}$$

where the inequality follows from inequality (3) and Lemma 6 provided in Appendix C. According to the Markov inequality, we further have

$$\Pr\left(\sum_{b\in\mathscr{R}[T]} 1_{\tilde{L}_b \leq L} \geq \frac{2}{3}nTf(D,n)\right) \leq \frac{3\varepsilon}{4}.$$

Let $\tilde{\Gamma}$ denote the event that

$$\sum_{b\in\mathscr{R}[T]} 1_{\tilde{L}_b \leq L} < \frac{2}{3}nTf(D,n).$$

Note that when $\tilde{\Gamma}$ and $\Gamma_{f(D,n)}$ both occur, at least $2nTf(D,n)/9$ good bits are transmitted to destinations over a distance at least L. Thus, we have

$$\begin{aligned} &E\left[\sum_{b=1}^{8nTf(D,n)/9} S_b \middle| \Gamma_{f(D,n)}\right] \\ \geq\ &\Pr(\tilde{\Gamma}|\Gamma_{f(D,n)})E\left[\sum_{b=1}^{8nTf(D,n)/9} S_b \middle| \Gamma_{f(D,n)},\tilde{\Gamma}\right] \\ \geq\ &\Pr(\tilde{\Gamma}|\Gamma_{f(D,n)})\frac{2nTf(D,n)}{9}\sqrt{\frac{\varepsilon\lambda}{2W\pi D}}. \end{aligned} \tag{13}$$

Note that

$$\Pr(\tilde{\Gamma}|\Gamma_{f(D,n)}) \geq \Pr(\tilde{\Gamma}) - \Pr(\Gamma^c_{f(D,n)}) \geq \frac{\varepsilon}{8}. \quad (14)$$

From inequalities (12), (13) and (14), we can conclude that for any $n \geq n_0$,

$$\frac{\varepsilon}{8}\frac{2nT\,f(D,n)}{9}\sqrt{\frac{\varepsilon f(D,n)}{2W\pi D}} \leq \frac{4T\sqrt{Wnf(n)}}{\Delta\sqrt{\pi}},$$

which implies that

$$f(D,n) \leq \frac{144\sqrt{2}W}{\Delta\varepsilon\sqrt{\varepsilon}},$$

and contradicts inequality (11). So theorem holds. ∎

B. Joint Coding-Scheduling Algorithm

In Section IV, we motivated the need to first encode source packets. In this subsection, we use Raptor codes and propose a joint coding-scheduling scheme to achieve the maximum throughput obtained in Theorem 3. We first define four different types of packets.

- Source packets: Packets which have to be transmitted from source to destination.
- Coded packets: Packets generated by Raptor codes. We use (i,k) to denote the k^{th} coded packet of node i.
- Duplicate packets: Each coded packet could be broadcast to other nodes to generate multiple copies, called duplicate packets. We use (i,k,j) to denote a copy of (i,k) carried by node j, and (i,k,J) to denote the set of all copies of coded packet (i,k).
- Deliverable packets: Duplicate packets that happen to be within distance L from their destinations.

Motivated by the heuristic argument in Section IV, we divide the unit square into square cells with each side of length equal to $1/\sqrt[4]{nD}$, which is of the same order as the optimal hitting distance. In our scheme, we will allow final delivery of a packet to its destination only when a relay carrying the packet is in the same cell as the destination. Thus, a packet delivered only when the relay and destination are within a distance of $\sqrt{2}/\sqrt[4]{nD}$, which is also the same as the hitting distance calculated in Section IV except for a constant factor which does not play a role in the order calculations. The mean number of nodes in each cell will be denoted by M and is equal to $\sqrt{n/D}$. The transmission radius of each node is chosen to be $\sqrt{2}/\sqrt[4]{nD}$ so that any two nodes within a cell can communicate with each other. This means that, given the interference constraint, two nodes in a cell can communicate if all nodes in cells within a fixed distance from the given cell stay silent. Each time slot is further divided into C mini-slots and each cell is guarantee to be active in at least one mini-slot within each time slot. Assume $C = 9$. The reason we use nine mini-slots is that if a node in a cell is active, then no other nodes in any of its neighboring eight cells can be active, but nodes outside this neighborhood can be active. Further, we denote the packet size to be $W/(2C)$ so that two packets can be transmitted in each mini-slot.

Our algorithm consists of three steps. The first step is Raptor encoding. The second step is the broadcasting step. In this step, sources broadcast coded packets to the other nodes in the same cell. After the broadcasting step, there are multiple duplicate packets for each coded packet. The third step is the receiving step, and a duplicate packet will be delivered to its destination if the packet and its destination happen to be in the same cell.

Let $\tilde{A}$ denote the area of a cell (note that $\tilde{A}$ is equal to M/n in this section) and $\tilde{M}[t]$ to denote the number of nodes in the cell at time slot t. A cell is said to be a *good cell* at time t if

$$\frac{9}{10}\tilde{A}n + 1 \leq \tilde{M}[t] \leq \frac{11}{10}\tilde{A}n.$$

Joint Coding-Scheduling Scheme: We group every $6D$ time slots into a super time slot. At each super time slot, the nodes transmit packets as follows.

(1) **Raptor Encoding:** Each source takes $6D/(25M)$ source packets, and uses Raptor codes to generate D/M coded packets.

(2) **Broadcasting:** This step consists of D time slots. At each time slot, the nodes do the following:
   (i) In each good cell, one node is randomly selected. If the selected node has not already transmitted all of its D/M coded packets, then it broadcasts a coded packet that was not previous transmitted to $9M/10$ other nodes in the cell during the mini-slot allocated to that cell. Recall that our choice of packet size allows one node in every good cell to transmit during every time slot.
   (ii) All nodes check the duplicate packets they have. If more than one duplicate packets has the same destination, randomly keep one and drop the others.

(3) **Receiving:** This step consists of $5D$ time slots. At each time slot, if a cell contains no more than two deliverable packets, the deliverable packets are delivered to their destinations using one-hop transmissions during the mini-slot allocated to that cell. At the end of this step, all undelivered packets are dropped. The destinations decode the received coded packets using Raptor decoding.

Note that in describing the algorithm, we did not account for the delays in Raptor encoding and decoding. However, Raptor codes have linear encoding and decoding complexity. Hence, even if these delays are taken into account, our order results will not change.

Theorem 4: Consider the joint coding-scheduling algorithm. Suppose D is both $\omega(\sqrt[3]{n})$ and $o(n)$, and the delay constraint is $6D$. Then given any $\varepsilon > 0$, there exists n_0 such that for any $n \geq n_0$, every source packet sent out can be recovered at the destination with probability at least $1-\varepsilon$, and furthermore

$$\lim_{T\to\infty} \Pr\left(\frac{\Lambda_i[T]}{T} \geq \frac{9W}{500C}\sqrt{\frac{D}{n}}\ \forall i\right) = 1. \quad (15)$$

Proof: Let t_s denote the t_s^{th} super time slot. We say that a coded packet is successfully duplicated if there are at least $4M/5$ copies of it after the broadcasting step. Using $A_i[t_s]$ to denote the number of coded packets which are successfully duplicated in super time slot t_s, we will first show that there exists n_1 such that for any $n \geq n_1$,

$$\Pr\left(A_i[t_s] \geq \frac{16}{25}\frac{D}{M}\right) \geq 1 - 3e^{-\frac{D}{500M}}. \quad (16)$$

Next use $B_i[t_s]$ to denote the number of distinct coded packets delivered to destination $i+1$ in super time slot t_s, we will show there exists n_2 such that for all $n \geq n_2$,

$$\Pr\left(B_i[t_s] \geq \frac{7}{25}\frac{D}{M}\middle| A_i[t_s] \geq \frac{16}{25}\frac{D}{M}\right) \geq 1 - 2e^{-\frac{D}{180M}}. \quad (17)$$

Further define $\mathscr{E}_i[t_s]$ to the event such that all $6D/(25M)$ source packets are fully recovered. From Lemma 5 on the error probability of Raptor codes provided in Appendix A, we have

$$\Pr\left(\mathscr{E}_i[t_s] | B_i[t_s] \geq \frac{7}{25}\frac{D}{M}\right) \geq 1 - \left(\frac{M}{D}\right)^a \quad (18)$$

for some $a > 0$. Note that $D = \omega(\sqrt[3]{n})$ implies

$$\lim_{n\to\infty} \frac{M}{D} = \lim_{n\to\infty} \frac{\sqrt{n}}{D\sqrt{D}} = 0.$$

Combining inequalities (16)-(18), we can conclude that for any $\varepsilon \leq 1/19$, there exists $n_0 \geq \max\{n_1, n_2\}$ such that for $n \geq n_0$,

$$\Pr(\mathscr{E}_i[t_s]) \geq 1 - \varepsilon, \quad (19)$$

which implies that every source packet sent out can be recovered with probability at least $1-\varepsilon$ when $n \geq n_0$. Since $1-\varepsilon \geq 18/19$, from the Chernoff bound (see Lemma 7 provided in the Appendix B for convenience), we can conclude that for $n \geq n_0$,

$$\Pr\left(\sum_{t_s=1}^{T_s} 1_{\mathscr{E}_i[t_s]} \geq \frac{9}{10}T_s\right) \geq 1 - e^{-\frac{T_s}{800}},$$

where we choose $\delta = 1/20$. Note that $\sum_{t_s=1}^{T_s} 1_{\mathscr{E}_i[t_s]} \geq \frac{9}{10}T_s$ implies at least

$$\frac{9}{10}T_s \times \frac{6D}{25M} \times \frac{W}{2C} = \frac{27W}{250C}\frac{DT_s}{M} = \frac{27W}{250C}DT_s\sqrt{\frac{D}{n}}$$

bits are successfully transmitted from node i to node $i+1$ in T_s super time slots. Since each super time slot consists of $6D$ time slots, we can conclude that for $n \geq n_0$,

$$\Pr\left(\Lambda_i[6DT_s] \geq \frac{27W}{250C}DT_s\sqrt{\frac{D}{n}}\ \forall i\right) \geq 1 - ne^{-\frac{T_s}{800}},$$

which implies

$$\lim_{T\to\infty} \Pr\left(\frac{\Lambda_i[T]}{T} \geq \frac{9W}{500C}\sqrt{\frac{D}{n}}\ \forall i\right) = 1$$

for $n \geq n_0$.

To complete the proof, we now show inequalities (16) and (17).

Proof of inequality (16): Let $\mathscr{B}_i[t]$ denote the event that node i broadcast a coded packet at time slot t. So $\mathscr{B}_i[t]$ occurs when following two conditions hold:

(i) The cell node i in is a good cell;
(ii) Node i is selected to broadcast.

Since the nodes are uniformly randomly positioned, from the Chernoff bound we have

$$\Pr(\mathscr{B}_i[t]) \geq \frac{10}{11M}\left(1 - 2e^{-\frac{M}{300}}\right),$$

which implies that there exists $\tilde{n}_1$ such that for any $n \geq \tilde{n}_1$,

$$\Pr(\mathscr{B}_i[t]) \geq \frac{8}{9}\frac{1}{M}.$$

So from the Chernoff bound again, we have

$$\Pr\left(\sum_{t=1}^{D} 1_{\mathscr{B}_i[t]} \geq \frac{4}{5}\frac{D}{M}\right) \geq 1 - e^{-\frac{D}{300M}} \quad (20)$$

for $n \geq \tilde{n}_1$. With a high probability, more than $4D/(5M)$ coded packets are broadcast, and each broadcast generates $9M/10$ copies.

Duplicate packets might be dropped at step (ii) of the broadcasting step. We next calculate the number of duplicate packets of node i left after the broadcasting step. Assume node i broadcasts $\tilde{D}_i$ coded packets, so $\tilde{D}_i \leq D/M$. Note that when node i broadcasts a coded packet, all other nodes have equal opportunity to get a copy. Then the number of duplicate packets left after the broadcasting step is the same as the number of nonempty bins of following balls-and-bins problem, where the bins represent the mobile nodes other than node i, and the balls represent the duplicate packets broadcast from node i.

Balls-and-bins Problem: Assume we have $(n-1)$ bins. At each time slot, we select $9M/10$ bins and drop one ball in each of them. Repeat this $\tilde{D}_i$ times.

Using N_1 to denote this number, from Lemma 8 in Appendix B, we have

$$\Pr(N_1 \geq (1-\delta)(n-1)\tilde{p}_1) \geq 1 - 2e^{-\delta^2(n-1)\tilde{p}_1/3},$$

where

$$\tilde{p}_1 = \left(1 - e^{-\frac{9\tilde{D}_iM}{10n-10}}\right).$$

We have there exists $\tilde{n}_2$ such that for all $n \geq \tilde{n}_2$,

$$\begin{aligned}(n-1)\tilde{p}_1 &= (n-1)\left(1 - e^{-\frac{9\tilde{D}_iM}{10n-10}}\right) \\ &\geq \frac{9\tilde{D}_iM}{10} - \frac{81\tilde{D}_i^2M^2}{100n-100} \\ &\geq \frac{44}{49}\tilde{D}_iM,\end{aligned}$$

where the last inequality holds since $\tilde{D}_iM \leq D = o(n)$. Thus choose $\delta = 1/50$ and we can conclude for $n \geq \tilde{n}_2$,

$$\Pr\left(N_1 \geq \frac{22}{25}\tilde{D}_iM \middle| \sum_{t=1}^{D} 1_{\mathscr{B}_i[t]} = \tilde{D}_i\right) \geq 1 - 2e^{-\frac{\tilde{D}_iM}{10000}}. \quad (21)$$

Recall a coded packet is said to be successfully duplicated if it has at least $4M/5$ copies at the end of the broadcasting step. Inequality (21) implies for $n \geq \tilde{n}_2$,

$$\Pr\left(A_i \geq \frac{4}{5}\tilde{D}_i \middle| \sum_{t=1}^{D} 1_{\mathscr{B}_i[t]} = \tilde{D}_i\right) \geq 1 - 2e^{-\frac{\tilde{D}_i M}{10000}},$$

since otherwise, less than $22\tilde{D}_i M/25$ duplicate packets are left in the network. Thus we can conclude that for $n \geq \tilde{n}_2$,

$$\Pr\left(A_i \geq \frac{16}{25}\frac{D}{M} \middle| \sum_{t=1}^{D} 1_{\mathscr{B}_i[t]} \geq \frac{4}{5}\frac{D}{M}\right) \geq 1 - 2e^{-\frac{D}{20000}}. \qquad (22)$$

Letting $n_1 = \max\{\tilde{n}_1, \tilde{n}_2\}$, inequality (16) follows from inequalities (20) and (22) for $n \geq n_1$.

Proof of inequality (17): Assume coded packets $\{(i,1), \ldots, (i, 16D/(25M))\}$ are successfully duplicated. We use $\mathscr{D}_{(i,k)}[t]$ to denote the event that coded packet (i,k) is delivered at time slot t. Then $\mathscr{D}_{(i,k)}[t]$ occurs if two conditions hold:

- (i) One and only one duplicate packet of (i,k) becomes a deliverable packet. Let $\mathscr{D}^1_{(i,k)}[t]$ denote this event. Assume the duplicate packet is (i,k,j).
- (ii) There are no other deliverable packets in the cell containing node j except packet (i,k,j) and one possible duplicate packet to node j carried by node $i+1$. Let $\mathscr{D}^2_{(i,k)}[t]$ denote this event.

Note that duplicate packets of node i are carried by different nodes, and their mobilities are independent. Each node carries at most D duplicate packets, and they have different destinations. Now assume there are $\tilde{M}$ copies of (i,k), then

$$\Pr\left(\mathscr{D}^1_{(i,k)}[t]\right) = \frac{\tilde{M}M}{n}\left(1 - \frac{M}{n}\right)^{\tilde{M}-1}.$$

For a successfully duplicated packet, $\tilde{M} \geq 4M/5$, so there exists $\tilde{n}_3$ such that for any $n \geq \tilde{n}_3$,

$$\Pr\left(\mathscr{D}^1_{(i,k)}[t]\right) \geq \frac{7M^2}{10n}.$$

Suppose we have $\bar{M}$ nodes in the cell containing node j, from the Chernoff bound, we have

$$\Pr\left(\bar{M} \leq \frac{11}{10}M\right) \geq 1 - e^{-\frac{M}{300}}.$$

Note that condition (ii) is equivalent to the following event: Given node j and node $i+1$ in the cell, no more deliverable packets appear when we put another $\bar{M}-2$ nodes into the cell. Now given K nodes in the cell, the probability that no more deliverable appears when we put another node is at least

$$\left(1 - \frac{2KD}{n-K}\right).$$

This holds due to the following two facts:

- (a) The new node should not be the destination of any duplicate packets already in the cell (there are at most KD duplicate packets already in the cell).
- (b) The duplicate packets carried by the new node are not destined for any of the existing K nodes. Note that each source has no more than D duplicate packets, so there are at most KD nodes which carry the duplicate packet towards the K existing nodes.

Note that $\lim_{n\to\infty} M = \infty$, so there exists $\tilde{n}_4$ such that for any $n \geq \tilde{n}_4$,

$$\begin{aligned} Pr\left(\mathscr{D}^2_{(i,k)}[t] \middle| \mathscr{D}^1_{(i,k)}[t]\right) &\geq \left(1 - e^{-\frac{M}{300}}\right)\Pi_{K=2}^{\frac{11M}{10}-1}\left(1 - \frac{2KD}{n-K}\right) \\ &\geq \left(1 - e^{-\frac{M}{300}}\right)\left(1 - \frac{22MD}{10n - 11M}\right)^{\frac{11M}{20}} \\ &\geq \frac{3}{11}. \end{aligned}$$

So we can conclude that for any $n \geq \max\{\tilde{n}_3, \tilde{n}_4\}$,

$$\Pr\left(\mathscr{D}_{(i,k)}[t]\right) \geq \frac{21M^2}{110n} = \frac{21}{110D},$$

which implies at each time slot, a successfully duplicated packet (i,k) will be delivered with probability at least $21/(110D)$. Note at each time slot, only one coded packet can be delivered to the destination of node i. So the number of distinct coded packets delivered to the destination of node i is the same as the number of nonempty bins of following balls-and-bins problem, where the bins represent the distinct coded packets, the balls represent successful deliveries, and a ball is dropped in a specific bin means the corresponding coded packet is delivered to the destination.

Balls-and-bins Problem: Suppose we have $16D/(25M)$ bins and one trash can. At each time slot, we drop a ball. Each bin receives the ball with probability $21/(110D)$, and the trash can receives the ball with probability $1-p$, where

$$p = \frac{21}{110D} \times \frac{16D}{25M} = \frac{168}{1375}\frac{1}{M}.$$

Repeat this $5D$ times.

Let N_2 denote nonempty bins of the above balls-and-bins problem and choose $\delta = 1/6$. From Lemma 8 in Appendix B, we have

$$\Pr\left(N_2 \geq \frac{7}{25}\frac{D}{M}\right) \geq 1 - 2e^{-\frac{D}{180M}},$$

and inequality (17) holds for $n \geq n_2$, where $n_2 = \max\{\tilde{n}_3, \tilde{n}_4\}$. ■

VI. Conclusion

In this paper, we investigated the optimal delay-throughput trade-off in ad-hoc networks with the two-dimensional i.i.d. mobility model and fast mobiles. The optimal trade-off was shown to be $\lambda = \Theta\left(\sqrt{D/n}\right)$ when D is both $\omega(\sqrt[3]{n})$ and $o(n)$, and a joint coding-scheduling scheme was proposed to achieve the optimal trade-off.

We now briefly comment on the conditions that we have imposed on the delay requirement to obtain the optimal delay-throughput trade-off. When D is $O(\sqrt[3]{n})$, the number of packets that can be transmitted in D time slots is a constant and hence one cannot use coding to ensure that the probability of packet loss is arbitrarily small. In this case, one can obtain a bound that is a logarithmic factor smaller than the upper bound using packet replication techniques as has been done in [10] for the slow-mobile case. However, the best achievable lower bound is unknown. Again the $o(n)$ requirement is not significant since a throughput of $\Theta(1)$ can be achieved if the delay requirement is larger [8].

Acknowledgment: The authors gratefully acknowledge the useful discussions with Prof. Geir E. Dullerud and Prof. Bruce Hajek, University of Illinois at Urbana-Champaign. We also thank Prof. Michael J. Neely, University of Southern California, for his comments on the earlier version of this paper.

References

[1] S. Aeron and V. Saligrama. Wireless ad hoc networks: strategies and scaling laws for the fixed SNR regime. Preprint, 2006.

[2] A. Agarwal and P. R. Kumar. Improved Capacity Bounds for Wireless Networks. In *Wireless Communications and Mobile Computing,* 4:251-261, 2004.

[3] S. N. Diggavi, M. Grossglauser and D. Tse. Even one-dimensional mobility increases ad hoc wireless capacity. In *Proc. of ISIT,* Lausanne, Switzerland, July 2002.

[4] A. El Gammal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proc. of IEEE Infocom,* 2004.

[5] A. El Gammal, J. Mammen, B. Prabhakar, and D. Shah. Optimal throughput-delay scaling in wireless networks - Part I: the fluid model. In *IEEE Transactions on Information Theory,* vol. 52, no. 6, pp. 2568-2592, 2006.

[6] A. El Gammal, J. Mammen, B. Prabhakar, and D. Shah. Optimal throughput-delay scaling in wireless networks - Part II: constant-size packets. In *IEEE Transactions on Information Theory,* vol. 52, no. 11, pp. 5111 - 5116, 2006.

[7] J. Mammen and D. Shah. Throughput and delay in random wireless networks with restricted mobility. Preprint, available at http://www.stanford.edu/˜jmammen/.

[8] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. In *Proc. of IEEE Infocom,* 2001.

[9] P. Gupta and P. Kumar. The capacity of wireless networks. In *IEEE Transactions of Information Theory,* vol. 46, no.2, pp. 388-404, 2000.

[10] X. Lin and N. B. Shroff. The fundamental capacity-delay tradeoff in large mobile ad hoc networks. In *Third Annual Mediterranean Ad Hoc Networking Workshop,* 2004.

[11] X. Lin, G. Sharma, R. R. Mazumdar and N. B. Shroff. Degenerate delay-capacity trade-offs in ad hoc networks with Brownian mobility. In *Joint Special Issue of IEEE Transactions on Information Theory and IEEE/ACM Transactions on Networking on Networking and Information Theory,* vol. 52, no. 6, pp 2777-2784, June 2006.

[12] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge, 2005.

[13] M.J. Neely and E. Modiano. Capacity and delay tradeoffs for ad-hoc mobile networks. In *IEEE Transactions on Information Theory,* vol. 51, No. 6, June 2005.

[14] A. Ozgur, O. Leveque, and D. Tse. How does the information capacity of ad hoc networks scale? In *Proc. of Allerton Conference,* 2006.

[15] D. Shah and S. Shakkottai. Oblivious routing with mobile fusion centers over a sensor network. Preprint.

[16] G. Sharma, R. Mazumdar, and N. Shroff. Delay and capacity trade-offs in mobile ad hoc networks: a global perspective. In *Proc. of IEEE Infocom,* 2006.

[17] A. Shokrollahi. Raptor Coes. In *IEEE Transactions on Information Theory,* vol. 52, No. 6, 2006.

[18] S. Toumpis and A. Goldsmith. Large wireless networks under fading, mobility, and delay constraints. In *Proc. of IEEE Infocom,* 2004.

[19] L. Ying, S. Yang and R. Srikant. Coding achieves the optimal delay-throughput trade-offs in mobile ad-hoc networks — Part I: two-dimensional i.i.d. mobility models. Preprint, 2006.

[20] L. Ying, S. Yang and R. Srikant. Coding achieves the optimal delay-throughput trade-offs in mobile ad-hoc networks — Part II: hybrid random walk and one-dimensional mobility models. Preprint, 2006.

Appendix A: Raptor Codes

A. Raptor Codes

Raptor codes are low-complexity, near-optimal rate-less codes for erasure channels. It was proposed by Shokrollahi in [17], and the following result was presented in [17].

Lemma 5: The receiver can correctly decode the M source packets with probability at least $1-1/(M^{a(\varepsilon)})$ for some $a(\varepsilon)>0$ after it obtains $(1+\varepsilon)M$ coded packets generated by Raptor codes. The number of operations used for encoding and decoding is $O(M)$.

□

Appendix B: Probability Results

In this appendix, we present some standard results in probability for the reader's convenience. In addition, we also present some minor variations of standard results which do not seem to be available in any book to best of our knowledge.

Lemma 6: Suppose $\{X_i\}$ are random variables with the same expectation, and N is another random variable upper bounded by $N_{\max}$. Then

$$E\left[\sum_{i=1}^{N} X_i\right] \leq N_{\max} E[X_1].$$

Proof:

$$\begin{aligned} E\left[\sum_{i=1}^{N} X_i\right] &= \sum_{k=1}^{N_{\max}} E\left[\sum_{i=1}^{k} X_i \middle| N=k\right] \Pr(N=k) \\ &= \sum_{i=1}^{N_{\max}} \sum_{k=i}^{N_{\max}} E\left[X_i | N=k\right] \Pr\left(N=k\right) \\ &\leq \sum_{i=1}^{N_{\max}} E\left[X_i\right] \\ &= N_{\max} E[X_1]. \end{aligned}$$

■

The following lemma is a standard result in probability, which we provide here for convenience.

Lemma 7: Let $X_1,\ldots,X_n$ be independent $0-1$ random variables such that $\sum_i X_i = \mu$. Then, the following Chernoff bounds hold

$$\Pr\left(\sum_{i=1}^{n} X_i < (1-\delta)\mu\right) \leq e^{-\delta^2\mu/2}; \quad (23)$$

$$\Pr\left(\sum_{i=1}^{n} X_i > (1+\delta)\mu\right) \leq e^{-\delta^2\mu/3}. \quad (24)$$

Proof: A detailed proof can be found in [12]. ■

The next lemmas are variations of standard balls-and-bins problems. However, we have not seen the results for the particular variation that we need in this paper. So we present the lemmas along with brief proofs below.

Lemma 8: Assume we have m bins. At each time, choose h bins and drop one ball in each of them. Repeat this n times. Using N_1 to denote the number of bins containing at least one ball, the following inequality holds for sufficiently large n.

$$\Pr\left(N_1 \leq (1-\delta)m\tilde{p}_1\right) \quad \leq \quad 2e^{-\delta^2 m\tilde{p}_1/3}. \tag{25}$$

where $\tilde{p}_1 = 1 - e^{-\frac{nh}{m}}$.

Proof: At each time, bin i receives a ball with probability h/m. We use κ_i to denote the number of balls in bin i. Now consider a related balls-and-bins problem where the ball dropping procedure is replaced by a certain number of trials as dictated by a Poisson random variable. Specifically, define $\tilde{n}$ to be a Poisson random variable with mean n, and repeat the ball dropping procedure $\tilde{n}$ times. Let $\tilde{\kappa}_i$ denote the number of balls in bin i in this case. It is easy to see that $\{\tilde{\kappa}_i\}$ are i.i.d. Poisson random variables with mean nh/m. So we can conclude

$$\begin{aligned}
\Pr\left(N_1 \leq (1-\delta)m\tilde{p}_1\right) &= \Pr\left(\sum_{i=1}^{m} 1_{\kappa_i \geq 1} \leq (1-\delta)m\tilde{p}_1\right) \\
&\leq \Pr\left(\sum_{i=1}^{m} 1_{\tilde{\kappa}_i \geq 1} \leq (1-\delta)m\tilde{p}_1 \middle| \tilde{n} \geq n\right) \\
&\leq \frac{\Pr\left(\sum_{i=1}^{m} 1_{\tilde{\kappa}_i \geq 1} \leq (1-\delta)m\tilde{p}_1\right)}{\Pr(\tilde{n} \geq n)} \\
&= 2\Pr\left(\sum_{i=1}^{m} 1_{\tilde{\kappa}_i \geq 1} \leq (1-\delta)m\tilde{p}_1\right).
\end{aligned}$$

Since

$$\Pr\left(1_{\tilde{\kappa}_i \geq 1} = 1\right) = \Pr\left(\tilde{\kappa}_i \geq 1\right) = 1 - e^{-\frac{nh}{m}} = \tilde{p}_1,$$

from Lemma 7, we have

$$\Pr\left(N_1 \leq (1-\delta)m\tilde{p}_1\right) \quad \leq \quad 2e^{-\delta^2 m\tilde{p}_1/3}.$$

■

The above idea of using a Poission number of trials to bound the probability of the occurrence of an event in a fixed number of trials is called the Poisson heuristic in [12].

Lemma 9: Suppose n balls are independently dropped into m bins and one trash can. After a ball is dropped, the probability in the trash can is $1-p$, and the probability in a specific bin is p/m. Using N_2 to denote the number of bins containing at least 1 ball, the following inequality holds for sufficiently large n.

$$\Pr\left(N_2 \leq (1-\delta)m\tilde{p}_2\right) \quad \leq \quad 2e^{-\delta^2 m\tilde{p}_2/3}; \tag{26}$$

where $\tilde{p}_2 = 1 - e^{-\frac{np}{m}}$.

Proof: Let κ_i denote the number of balls in bin i. Next define $\tilde{n}$ to be a poisson random variable with mean n. We consider the case such that $\tilde{n}$ balls are independently dropped in m bins. Using $\tilde{\kappa}_i$ to be number of balls in bin i in this case, it is easy to see that $\{\tilde{\kappa}_i\}$ are i.i.d. poisson random variables with mean $\frac{np}{m}$.

Now given n_b, we first have

$$\begin{aligned}
\Pr\left(N_2 \leq (1-\delta)m\tilde{p}_2\right) &= \Pr\left(\sum_{i=1}^{m} 1_{\kappa_i \geq 1} \leq (1-\delta)m\tilde{p}_2\right) \\
&= \Pr\left(\sum_{i=1}^{m} 1_{\tilde{\kappa}_i \geq 1} \leq (1-\delta)m\tilde{p}_2 \middle| \tilde{n} \geq n\right) \\
&\leq \frac{\Pr\left(\sum_{i=1}^{m} 1_{\tilde{\kappa}_i \geq 1} \leq (1-\delta)m\tilde{p}_2\right)}{\Pr(\tilde{n} \geq n)}.
\end{aligned}$$

Since

$$\Pr\left(1_{\tilde{\kappa}_i \geq 1} = 1\right) = \Pr\left(\tilde{\kappa}_i \geq 1\right) = 1 - e^{-\frac{np}{m}} = \tilde{p}_2,$$

from Lemma 7, we have

$$\Pr\left(\sum_{i=1}^{m} 1_{\tilde{\kappa}_i \geq 1} \leq (1-\delta)m\tilde{p}_2\right) \leq e^{-\delta^2 m\tilde{p}_2/3}.$$

which implies for sufficiently large n,

$$\begin{aligned}
\Pr\left(N_2 \leq (1-\delta)m\tilde{p}_2\right) &\leq \sqrt{3\pi n}e^{-\delta^2 m\tilde{p}_2/3} \\
&\leq 2e^{-\delta^2 m\tilde{p}_2/3}.
\end{aligned}$$

■