Analyzing Random Network Coding With Differential Equations and Differential Inclusions

Dan Zhang and Narayan B. Mandayam, *Fellow, IEEE*

***Abstract*—We develop a framework based on differential equations (DE) and differential inclusions (DI) for analyzing Random Network Coding (RNC) in an arbitrary wireless network. The DEDI framework serves as a powerful numerical and analytical tool to study RNC. For demonstration, we first build a system of DE's with this framework, under the fluid approximation, to model the means of the rank evolution processes. By converting this system to DI's and explicitly solving them, we show that the average multicast throughput is equal to the min-cut bound. We then turn to the precise system of DE's regarding the means and variances of the rank evolution processes. By analyzing this system, we show that the rank evolution processes asymptotically concentrate to the solution of the DI's obtained previously. From this result, it immediately follows that the min-cut bound can be achieved as the number of source packets becomes large. We demonstrate the numerical accuracy and flexibility in performance analysis enabled by the DEDI framework via illustrative examples of networks with multiple multicast sessions, complex topology and correlated reception. We also briefly discuss its application in MAC and PHY adaptation and the extension to Random Coupon Selection.**

***Index Terms*—Capacity achievability, concentration, differential equation, differential inclusion, dynamical system, random coupon selection, random network coding.**

I. Introduction

SINCE the pioneering work by Ahlswede *et al.* [1] that established the benefits of coding in routers and provided theoretical bounds on the capacity of such networks, the breadth of areas that have been touched by network coding is vast and includes not only the traditional disciplines of information theory, coding theory and networking, but also topics such as routing algorithms[2], distributed storage[3], [4], network monitoring, content delivery[5], [6], and security[7]. Among other variants, random network coding (RNC) [8], [9] has received extensive interest in particular. By allowing routers to perform random linear operations, RNC is shown to be capacity achieving and fault tolerant. In spite of all the excellent progress previous studies have made in the area of RNC, what is still missing is a simple framework that can be used to describe the evolution of state in a wireless network where RNC is employed. In this paper we present a framework called DEDI based on differential equations (DE) and differential inclusions (DI), which are a generalization of DE's to allow for discontinuous right-hand sides. The DEDI serves as a powerful numerical and analytical tool to study RNC. We demonstrate this by presenting theoretical analysis of information flows with RNC as well as numerical examples. We will setup under the fluid approximation a system of DI's that approximately characterizes the means of the rank evolution processes and solve it explicitly. The solution shows that the average throughput is given by the min-cut bound. Next we prove that the actual rank evolution processes concentrate to the previously obtained solution in probability, hence proving the well known result that RNC achieves the min-cut bound, all in the context of general lossy wireless networks. The flexibility of DEDI in performance analysis will also be shown via illustrative examples of networks with multiple multicast sessions, user cooperation and arbitrary topologies.

Using the DEDI framework, we present results similar to Ho *et al.* [10] which for the first time characterized the achievable throughput of RNC by analyzing the codes algebraically; and also to Lun *et al.* [9] which later studied the same problem with a Jackson network approach, analyzing the achievable capacity by treating the propagation of innovative packets through the network as concatenated queuing systems. While the coding strategy considered in this paper is similar to [9], there are a number of notable differences with the prior work. Unlike [10], our work makes no assumption on the size of the underlying field when proving achievability. Rather, we show that achievability is a direct consequence of the convergence of the fluid model in this particular case, which does not hold in general. In [9], the fluid approximation is also used to characterize a fictitious queueing system whose throughput lower bounds the real process of innovative packet propagation. However, our contribution mainly focuses on a direct and compact description of innovative packet propagation through an arbitrary lossy wireless network using differential equations. This framework has several advantages:

1) It is easy to manipulate. Many problems related to RNC which have been previously studied with a mixture of information theory and queueing theory now become problems of analyzing and solving systems of differential equations. Previous studies of RNC usually begin their analysis with acyclic networks, and then extend it to a general

Manuscript received April 18, 2010; revised April 07, 2011; accepted July 04, 2011. Date of current version December 07, 2011. This work was supported in part by the NSF under Grant CCF-1016551. This work was presented in part at the 2010 IEEE International Symposium on Information Theory and in part at the 2010 IEEE Wireless Network Coding Workshop.

The authors are with the Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, NJ 08901 USA (e-mail: zhangdan@gmail.com, narayan@winlab.rutgers.edu).

Communicated by M. Médard, Associate Editor for the Special Issue on Facets of Coding Theory: From Algorithms to Networks.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2011.2170107

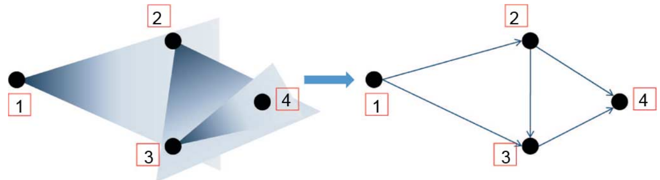

Fig. 1. Hypergraph model of a wireless network of four nodes and its arrow-dot representation.

topology, because cycles often impose an obstacle which must be circumvented this way. With our framework, this obstacle never arises because the topological information has been naturally embedded in the system of DE's.
2) It is a powerful computational tool that can help network designers evaluate their RNC schemes with dynamically adjusted parameters. Algebraic, information and queueing theoretic approaches often do not lend themselves directly amenable to use for this purpose. We will highlight this aspect in Section VII.
3) It can be generalized to other transmission schemes besides RNC. We will demonstrate in Section VII its possible extension to Random Coupon (RC) [11], an alternative to RNC.

In what follows, Section II introduces the hypergraph model for a wireless network first proposed in [9], concepts such as cut sets, the min cut and connectivity for a hypergraph, and the basic operations of RNC, the definition of rank evolution for RNC; Section III presents the setup of the DEDI framework for RNC using a fluid approximation as well as a systems of DE's and DI's that describe rank evolution of RNC. In Section IV we solve the system of DI's explicitly to obtain the average throughput of RNC, which is given by the min-cut bound. In Section V we further study the means and variances to give an achievability proof of RNC that makes no assumption on field size using the DEDI framework described here. Section VI presents extensive numerical examples to illustrate the application of DEDI to situations of DEDI to situations of multiple multicast sessions, complex topology and joint reception. Section VII gives a brief discussion on the possible application of the DEDI framework and its extension to Random Coupon. We conclude in Section VIII.

II. Review of the Hypergraph Model and RNC

A generic wireless network is modeled as a hypergraph $G = (\mathcal{N}, \mathcal{E})$ consisting of N nodes $\mathcal{N} = \{1, 2, \ldots, N\}$ and hyperarcs $\mathcal{E} = \{(i, \mathcal{K}) | i \in \mathcal{N}, \mathcal{K} \subset \mathcal{N}\}$. Each hyperarc captures the fact that, as any wireless transmission is inherently a broadcast, a packet sent from node i can be received by some or all the nodes in a set $\mathcal{K} \subset \mathcal{N}$. This idea is shown in Fig. 1 where the hypergraph of a four-node network is shown. The transmission from node 1 can be overheard by node 2 and 3, while the transmission from node 3 can only be overheard by node 4, all with a probability. This relationship between nodes can be conveniently represented with arrows. One should not, however, confuse the arrow representation with the digraph of a wired network. Assume some underlying MAC is operating in its steady state such that each node i is transmitting according to an independent Poisson process with the intensity of λ_i packets per second. We say that a packet is successfully received by a set $\mathcal{K}$ of nodes if the packet is successfully received by at least one node in $\mathcal{K}$, which happens with a probability $P_{i,\mathcal{K}}$. Note the definition of $P_{i,\mathcal{K}}$ is general and does not assume independent receptions among the nodes in $\mathcal{K}$. This generality allows channel correlation or user cooperation (e.g., joint detection) to be analyzed in a unified framework. We define the effective transmission rate $z_{i,\mathcal{K}}$ for $(i, \mathcal{K})$ (i.e., from i to $\mathcal{K}$) as

$$z_{i,\mathcal{K}} = \lambda_i P_{i,\mathcal{K}}. \tag{1}$$

which is the intensity of the Poisson process of packets from node i successfully arriving/being received by $\mathcal{K}$. $z_{i,\mathcal{K}}$ also can be regarded as the extended concept of link capacity from node i to the set $\mathcal{K}$. When $\mathcal{K} \subset \mathcal{T} \subset \mathcal{N}$, we must have

$$z_{i,\mathcal{K}} \leq z_{i,\mathcal{T}} \tag{2}$$

because $P_{i,\mathcal{K}} \leq P_{i,\mathcal{T}}$. Suppose $\mathcal{S}, \mathcal{K} \subset \mathcal{N}$ and $\mathcal{S} \cap \mathcal{K} = \emptyset$. Define a cut for the pair $(\mathcal{S}, \mathcal{K})$ as a set $\mathcal{T}$ satisfying $\mathcal{K} \subset \mathcal{T} \subset \mathcal{S}^c$. Let $C(\mathcal{S}, \mathcal{K})$ denote the collection of all cuts for $(\mathcal{S}, \mathcal{K})$. The size of $\mathcal{T}$ is defined as $c(\mathcal{T}) = \sum_{i \in \mathcal{T}^c} z_{i,\mathcal{T}}$. A min cut $\mathcal{T}_{\min}$ for $(\mathcal{S}, \mathcal{K})$, whose size is denoted as $c_{\min}(\mathcal{S}, \mathcal{K})$ is a cut satisfying

$$c(\mathcal{T}_{\min}) = \min_{\mathcal{T}' \in C(\mathcal{S}, \mathcal{K})} c(\mathcal{T}'). \tag{3}$$

We denote the collection of cuts for $(\mathcal{S}, \mathcal{K})$ that satisfy (3) as $C_{\min}(\mathcal{S}, \mathcal{K})$. Conventionally, we have

$$C_{\min}(\emptyset, \mathcal{K}) = \{\mathcal{T} | \mathcal{K} \subset \mathcal{T} \subset \mathcal{N}\} \text{ and } c_{\min}(\emptyset, \mathcal{K}) = 0. \tag{4}$$

We say G is connected if, for any $\emptyset \neq \mathcal{T} \subsetneq \mathcal{N}$, $c(\mathcal{T}) > 0$.

When RNC is employed in unicast/multicast sessions, a group of nodes work together by sending out coded packets that are generated from the received (coded) packets or the packets they deliver as the sources. The operation of RNC is different from that of the deterministic network coding [1] or randomized network coding [12] in that a coding coefficient vector is generated for each coded packet. Without loss of generality, we also assume that every node in the network executes RNC in a cooperative manner to carry an information flow comprised of one or more multicast sessions, otherwise we confine our discussion to the part of the network (and refer to it as "the network") in which every node participates in RNC cooperatively. What follows will only be concerned with a single information flow. Admittedly, we may have multiple independent information flows separately coded with RNC, which may traverse the same nodes. That means, if node i is such a node, it will linearly mix packets that belong to the same flow, but never mix packets from different flows with RNC. Consequently, λ_i will be divided among these flows and we can safely confine our attention to each flow individually, taking into consideration only the portion of λ_i that is allocated to the flow.

With RNC, each packet $\underline{w}$ is a row vector from $\mathbb{F}^L$ where $\mathbb{F}$ is a given finite field of size q and L is a positive constant that denotes the length of the packet. Every node maintains a reservoir consisting of all the packets the node holds as a source plus all the packets received thus far during a coded session. The reservoir is ever growing and purged only after the associated information flow is completed. Whenever a node gets to transmit, a coded packet is formed and sent out. Suppose at a time instant node i needs to form a coded packet $\underline{v}$ from its reservoir $\{\underline{w}_{i,1}, \underline{w}_{i,2}, \dots, \underline{w}_{i,n}\}$, $\underline{v}$ will have the form $\underline{v} = a_1\underline{w}_{i,1} + a_2\underline{w}_{i,2} + \cdots + a_n\underline{w}_{i,n}$, where $a_1, \dots, a_n \in \mathbb{F}$ are randomly generated. Since the coding operation is entirely linear, we have $\underline{v} = b_{i,1}\underline{w}_1 + b_{i,2}\underline{w}_2 + \cdots + b_{i,m}\underline{w}_m$ where $\underline{w}_1, \dots, \underline{w}_m$ are the ensemble of m source packets, possibly belong to multiple source nodes and multiple sessions. $[b_{i,1}, b_{i,2}, \dots, b_{i,m}] \in \mathbb{F}^m$ is called the global coefficient vector associated with $\underline{v}$. Each node sends the global coefficient vector along with its associated coded packet in order to enable the receiving nodes to calculate the global coefficient vectors for their own coded packets. Let S_i be the vector space spanned by the global coefficient vectors associated with the packets in node i's reservoir and define $N_i = \dim S_i$, which we call the *rank* of node i. S_i and N_i are time dependent as the coded transmissions evolve and once $N_i = m$, decoding can be carried out with a linear inverse operation. Further, for any set $\mathcal{K} \subset \mathcal{N}$, define

$$S_{\mathcal{K}} = \sum_{i \in \mathcal{K}} S_i = \mathrm{span}\left\{\bigcup_{i \in \mathcal{K}} S_i\right\}, \quad N_{\mathcal{K}} = \dim S_{\mathcal{K}} \tag{5}$$

and call $N_{\mathcal{K}}$ the rank of $\mathcal{K}$. The question we are interested in answering is how the rank N_i or $N_{\mathcal{K}}$ increases over time, i.e., how the ranks evolve.

III. DEDI Framework for RNC

In this section we will develop the DEDI framework for studying rank evolution of RNC and show its use via illustrative examples[13].

A. Rank Evolution Modeled With DE

The DEDI framework begins with the following lemma that describes the mean of $N_{\mathcal{K}}(t)$:

Lemma 1:

$$\mathrm{d}E\left[N_{\mathcal{K}}(t)\right]/\mathrm{d}t = \sum_{i \in \mathcal{K}^c} z_{i,\mathcal{K}} E\left[1 - q^{N_{\mathcal{K}}(t) - N_{\{i\}\cup\mathcal{K}}(t)}\right]. \tag{6}$$

Proof: Let $\Delta_{\mathcal{K}}(t)$ denote the increment in the number of innovative packets in $(t, t+\Delta t)$, then

$$N_{\mathcal{K}}(t+\Delta t) = N_{\mathcal{K}}(t) + \Delta_{\mathcal{K}}(t). \tag{7}$$

Notice that every node i sends packets according to an independent Poisson process with intensity λ_i, we can calculate $E[\Delta(t)]$ as

$$E\left[\Delta_{\mathcal{K}}(t)\right] = \sum_{i \in \mathcal{K}^c} E\left[\Delta_{i,\mathcal{K}}(t)\right] \tag{8}$$

where $\Delta_{i,\mathcal{K}}(t)$ is the number of innovative packets (either 0 or 1) sent from node $i \in \mathcal{K}^c$ in $[t, t+\Delta t)$. Using the chain rule, we have

$$\begin{aligned} E\left[\Delta_{\mathcal{K}}(t)\right] &= \sum_{i \in \mathcal{K}^c} \lambda_i \Delta t E\left[\Delta_{i,\mathcal{K}}(t) | i \text{ sends a packet}\right] + o(\Delta t) \\ &= \sum_{i \in \mathcal{K}^c} \lambda_i \Delta t P_{i,\mathcal{K}} E\big[\Delta_{i,\mathcal{K}}(t) | i \text{ sends a packet} \\ &\qquad \text{which is received in } \mathcal{K}\big] + o(\Delta t). \end{aligned} \tag{9}$$

A packet sent from $i \in \mathcal{K}^c$ is innovative to $\mathcal{K}$ (i.e., $\Delta_{i,\mathcal{K}} = 1$) if and only if it comes from $S_i \setminus (S_i \cap S_{\mathcal{K}})$. Since

$$|S_i \cap S_{\mathcal{K}}| = q^{\dim S_i \cap S_{\mathcal{K}}} = q^{N_i + N_{\mathcal{K}} - N_{\mathcal{K}\cup\{i\}}} \tag{10}$$
$$\text{and } |S_i| = q^{\dim S_i} = q^{N_i} \tag{11}$$

it follows that the probability that the received packet is innovative is given by

$$\left(|S_i| - |S_i \cap S_{\mathcal{K}}|\right)/|S_i| = 1 - q^{N_{\mathcal{K}} - N_{\mathcal{K}\cup\{i\}}}. \tag{12}$$

Averaged over all possible values of $(N_{\{i\}\cup\mathcal{K}}, N_{\mathcal{K}})$, we have

$$E\left[\Delta_{\mathcal{K}}(t)\right] = \sum_{i \in \mathcal{K}^c} \lambda_i \Delta t P_{i,\mathcal{K}} E\left[1 - q^{N_{\mathcal{K}}(t) - N_{\{i\}\cup\mathcal{K}}(t)}\right] + o(\Delta t). \tag{13}$$

Therefore, we have a precise differential equation for $E[N_{\mathcal{K}}(t)]$ as follows:

$$\frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} = \sum_{i \in \mathcal{K}^c} z_{i,\mathcal{K}} E\left[1 - q^{N_{\mathcal{K}}(t) - N_{\{i\}\cup\mathcal{K}}(t)}\right]. \tag{14}$$

Let $V_i(t) = E[N_i(t)]$ and $V_{\mathcal{K}}(t) = E[N_{\mathcal{K}}(t)]$. We want to build a system of differential equations that (approximately) describe $V_i(t)$ and $V_{\mathcal{K}}(t)$. Though Lemma 1 does not precisely provide the equations we want (the right-hand sides are not functions of the unknowns), we can turn them into such via a fluid approximation argument: when m is large, the stochastic process $N_{\mathcal{K}}(t)$ behaves on a macro scale like a deterministic function which is $V_{\mathcal{K}}(t)$. This leads us to make the following approximation

$$E\left[1 - q^{N_{\mathcal{K}}(t) - N_{\{i\}\cup\mathcal{K}}(t)}\right] \approx 1 - q^{V_{\mathcal{K}}(t) - V_{\{i\}\cup\mathcal{K}}(t)} \tag{15}$$

and consequently we have

$$\dot{V}_{\mathcal{K}} \approx \sum_{i \notin \mathcal{K}} z_{i\mathcal{K}} \left(1 - q^{V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}}}\right). \tag{16}$$

The solution of (16) gives the expectation of the rank of a set $\mathcal{K}$ at any given time instant t. It actually stands for a system of $2^N - 1$ equations, each for an nonempty $\mathcal{K} \subset \mathcal{N}$. They collectively give a complete description of rank evolution in the system. Note $V_{\mathcal{K}}$ is solely determined by $\{V_{\mathcal{K}\cup\{i\}}\}_{i \notin \mathcal{K}}$. This dependency can be explored to arrange (16) into a partial order "$\lesssim$" such that $V_{\mathcal{K}} \lesssim V_{\mathcal{L}}$ if and only if $\mathcal{K} \subset \mathcal{L}$. This partial order can be pictorially represented as a layered structure, for which an example is shown in Fig. 2 for $N = 3$. To determine a quantity on any particular layer, one only needs to know the the quantities on the layer immediately above indicated by arrows.

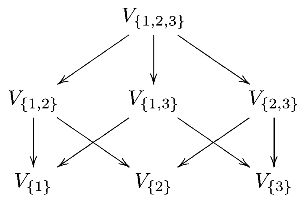

Fig. 2. Layered structure for the rank evolution of a 3-node network.

The layered structured will be exploited in Section IV to facilitate the proofs.

Theoretically, with appropriate boundary condition, (16) can be solved. The instantaneous throughput is then obtained as $\dot{V}_{\mathcal{K}}$ or $\dot{V}_i$. For example, assuming node 1 is the unique source with m packets to deliver, the boundary conditions (B.C.) for this systems of DE's are

$$V_{\mathcal{K}}(0) = \begin{cases} m, & 1 \in \mathcal{K} \\ 0, & \text{o.w.} \end{cases} \tag{17}$$

If only part of the nodes, say $\mathcal{L} \subset \mathcal{N}$, participate in carrying the flow, (16) still holds, except that we should replace $\mathcal{K}$ with $\mathcal{K} \cap \mathcal{L}$ and the top layer in the layered structure consists of $V_{\mathcal{L}}$ alone.

In practice, q is usually chosen to be an integral power of 2, not only because arithmetic in a field of characteristic q is then particularly amenable to machines, but also because they are the natural granularity used in storage and communication, e.g., bits, bytes, words, etc. With such choices of q, $q^{V_{\mathcal{K}} - V_{\{i\}\cup\mathcal{K}}}$ could be equal to 0 to within the precision of standard numerical software when $|V_{\mathcal{K}} - V_{\{i\}\cup\mathcal{K}}|$ is even moderately large. As $V_{\mathcal{K}}$ can never exceed $V_{\mathcal{K}\cup\{i\}}$, $V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}}$ is nonpositive and in this case we may approximate $1 - q^{V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}}}$ by

$$1 - q^{V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}}} \approx \begin{cases} 1, & V_{\mathcal{K}} < V_{\mathcal{K}\cup\{i\}} \\ 0, & V_{\mathcal{K}} = V_{\mathcal{K}\cup\{i\}}. \end{cases} \tag{18}$$

The approximation for different values of q is shown in Fig. 3. It is evident that, when $q = 256$ the approximation is very close for every nonpositive integer. Even when $q = 2$, the approximation is very good when $V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}} < -6$ or $V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}} = 0$. For other nonpositive integer values, the approximation has an error bounded by 1. When $q \to \infty$, (18) becomes more accurate. However, as will be shown in numerical examples, when the total rank is large the approximation rarely fails even for $q = 2$. Consequently we may rewrite (16) as

$$\dot{V}_{\mathcal{K}} = \sum_{i \notin \mathcal{K}} z_{i\mathcal{K}} (V_{\mathcal{K}\cup\{i\}} \ominus V_{\mathcal{K}}), \quad \forall \mathcal{K} \subset \mathcal{N} \tag{19}$$

with the same boundary conditions as in (17). The binary operation $\ominus$ is defined as

$$x \ominus y = \begin{cases} 1, & x > y \\ 0, & \text{o.w.} \end{cases} \tag{20}$$

Though the simplified DE's shown in (19) have discontinuous right-hand sides due to the $\ominus$ operation, they are no longer subject to the same precision problem. Numerical solution of (19) can be obtained by any DE solvers fairly efficiently. To demonstrate this, consider the following example illustrating a simplistic wireless network that employs RNC in a P2P-like transmission scheme, shown in Fig. 4. Assume $\lambda_i = 1$ packet/second, $i = 1, 2, 3, 4$ and $q = 2$. The labels attached to the arrows show reception probabilities, which are independent to each other. This means, for example, $P_{1,2} = 0.4$, but $P_{1,\{2,3\}} = 1 - (1 - 0.4)(1 - 0.2) = 0.52$. We assume that node 1 is the server which has 400 packets to be downloaded to node 2, 3 and 4 with RNC. Like a typical wired P2P network, node 2, 3 and 4 broadcast to each other to enhance efficiency. Fig. 5 shows the rank evolution at the four nodes, through both simulation and the solution to the corresponding simplified DE's. It is evident that the DE solution fits the simulated curves nicely.

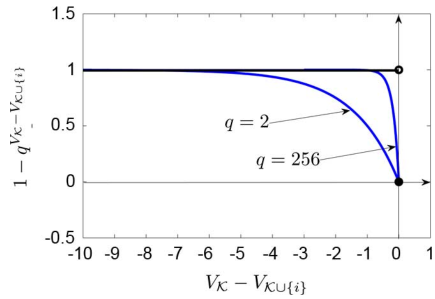

Fig. 3. Approximate $1 - q^{V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}}}$ with $V_{\mathcal{K}\cup\{i\}} \ominus V_{\mathcal{K}}$ when $q = 2$ and $q = 256$.

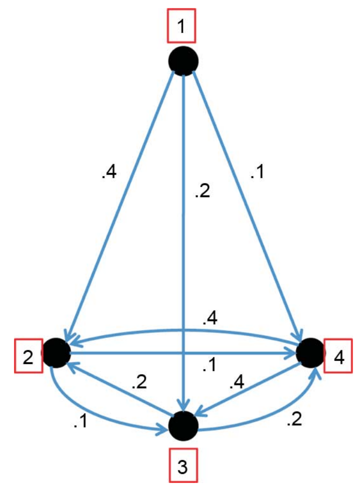

Fig. 4. Simplistic wireless P2P network.

B. Rank Evolution Modeled With DI

While (19) can be numerically evaluated with any DE solver, it is not amenable to analysis due to the discontinuous right-hand sides. Besides, the approximation shown in(18) becomes most inaccurate when $V_{\{i\}\cup\mathcal{K}} - V_{\mathcal{K}} \to 0^+$. In fact, no matter what q is, if $V_{\{i\}\cup\mathcal{K}} - V_{\mathcal{K}}$ is sufficiently small, $1 - q^{V_{\mathcal{K}} - V_{\mathcal{K}\cup\{i\}}}$ can take on any value in $[0, 1)$. This discrepancy prompts us to modify the right-hand side of (19) to incorporate semicontinuity [14], which allows a range of values for $V_{\{i\}\cup\mathcal{K}}(t) \ominus V_{\mathcal{K}}(t)$ to choose

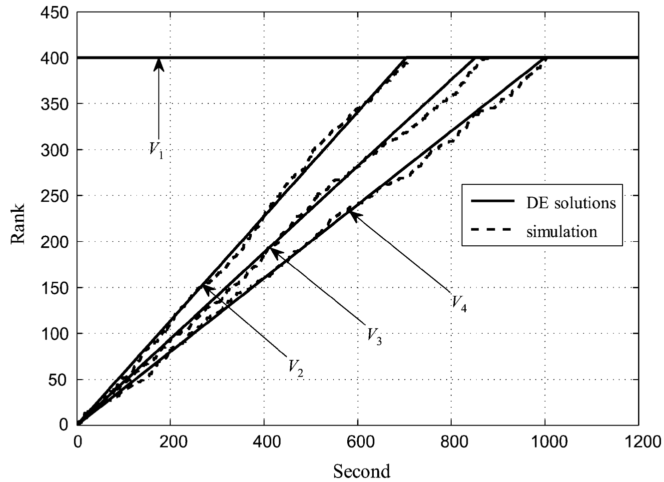

Fig. 5. Rank evolution of the simplistic wireless P2P network is obtained through simulation as well as solution to the corresponding DE's.

from when $V_{\mathcal{K}}(t) = V_{\{i\}\cup\mathcal{K}}(t)$. Specifically, we define an upper semicontinuous function $\mathrm{Sgn}^+ : \mathbb{R} \to 2^{\mathbb{R}}$

$$\mathrm{Sgn}^+(x) = \begin{cases} \{0\}, & x < 0 \\ [0,1], & x = 0 \\ \{1\}, & x > 0 \end{cases} \tag{21}$$

to replace the "$\ominus$" operation

$$\dot{V}_{\mathcal{K}} \in \sum_{i \notin \mathcal{K}} z_{i\mathcal{K}} \mathrm{Sgn}^+(V_{\mathcal{K}\cup\{i\}} - V_{\mathcal{K}}),\ \forall \mathcal{K} \subset \mathcal{N}. \tag{22}$$

Fig. 6 illustrates the conversion, where the Sgn^+ function shown in 6(b) has apparently re-acquired certain continuity compared to the jump discontinuity shown in 6(a). The same boundary condition in (17) still holds. To be compatible with (19), when $\mathcal{K} = \mathcal{N}$, we define the right-hand side of (22) to be $\{0\}$ instead of $\emptyset$. In mathematical literature, the system of inclusions in (22) plus the same boundary condition in(17) is called a system of differential inclusions (DI), first systematically studied by A. F. Filippov [15] followed by many analysts. DI is a generalization of the dynamical system described by DE's, allowing them, in particular, to have discontinuous right-hand sides, which is exactly the case in (19). Such dynamical systems with derivative discontinuities arise extensively in mechanics, electronics and biology. For example, an initial value problem on a time interval $[0,\infty)$ for DI takes the following form:

$$\dot{\underline{x}} \in F(t, \underline{x}), \quad \underline{x}(0) = \underline{x}_0. \tag{23}$$

where $\underline{x}(t) \in \mathbb{R}^d$ is the state vector, $F : [0,\infty) \times \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is a set-valued function and d is the dimension of the dynamical system. Its solution is defined to be an absolutely continuous function $\underline{y}(t)$ such that $\underline{y}(0) = \underline{x}_0$ and $\dot{\underline{y}}(t) \in F(t, \underline{y}(t))$ almost everywhere in $[0,\infty)$. In this article, however, owing to the particular form of the Sgn^+ function, we will be dealing with a special collection of DI's such that the solutions only need to be continuous functions satisfying the inclusion at all but finitely many points in $[0,\infty)$.It is clear that any solutions to (19) are necessarily solutions to (22). It is possible that the reformulation via DI's could enlarge the set of solutions. However, as we will see, for our specific problem, (22) turns out to have a unique solution in our discussion.

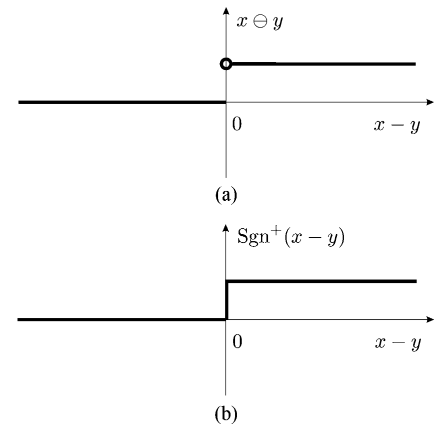

Fig. 6. (a) Plot of $x \ominus y$ as a function of $x - y$. (b) Plot of Sgn^+ as a set-valued function of $x - y$.

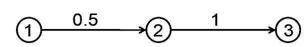

Fig. 7. Three node network with linear topology. Node 1 tries to deliver 100 packets to node 2 and node 3.

The generalization from (19) to (22) not only paves the way for easy analysis of RNC, but also furnishes a better interpretation to the solution of (19), which is illustrated by the example shown in Fig. 7. Suppose we wish to use RNC in a network consisting of three nodes 1, 2 and 3 to deliver $m = 100$ packets from node 1 to node 2 and 3. The network has a linear topology shown in Fig. 7. Based on the underlying MAC, node 1 transmits at 0.5 packets/second to node 2 which transmits at 1 packets/second to node 3, i.e., $z_{12} = z_{1,\{23\}} = 0.5$ packet/ second, $z_{23} = 1$ packet/ second. We wish to know at what rates $V_2(t)$ and $V_3(t)$ increase by solving the corresponding system of DE's as given in (19)

$$\dot{V}_2 = z_{12}(m \ominus V_2) \tag{24}$$
$$\dot{V}_3 = z_{23}(V_{\{23\}} \ominus V_3) \tag{25}$$
$$\dot{V}_{\{23\}} = z_{1,\{23\}}(m \ominus V_{\{23\}}) \tag{26}$$
$$\text{B.C.} \quad V_2 = V_3 = V_{\{23\}} = 0 \tag{27}$$

for which the solutions obtained by a numerical DE solver are shown in Fig. 8. We are not particularly interested in $V_{\{23\}}$ per se, but by comparing (24) and (26) we observe that they have the same solutions, i.e., $V_2(t) = V_{\{23\}}(t), \forall t$. In fact, Fig. 8 shows $V_2(t) = V_{\{23\}}(t) = V_3(t), \forall t \geq 0$ and $\dot{V}_2(t) = \dot{V}_{\{23\}}(t) = \dot{V}_3(t) = 0.5, \forall t \in [0, 200)$. However, if we plug the solution back into (25), we get $\dot{V}_3(t) = 0, \forall t \in [0, 200)$. This discrepancy arises due to the discontinuous right-hand sides of the system of DE's in (24)–(26). This can be explained if we recast (24)–(26) into differential inclusions as follows:

$$\begin{aligned} \dot{V}_2 &\in z_{12}\mathrm{Sgn}^+(m - V_2) = 0.5\mathrm{Sgn}^+(100 - V_2) \\ \dot{V}_3 &\in z_{23}\mathrm{Sgn}^+(V_{\{23\}} - V_3) = \mathrm{Sgn}^+(V_{\{23\}} - V_3) \\ \dot{V}_{\{23\}} &\in z_{1,\{23\}}\mathrm{Sgn}^+(m - V_{\{23\}}) = 0.5\mathrm{Sgn}^+(100 - V_{\{23\}}) \\ \text{B.C.} &\quad V_2 = V_3 = V_{\{23\}} = 0. \end{aligned}$$

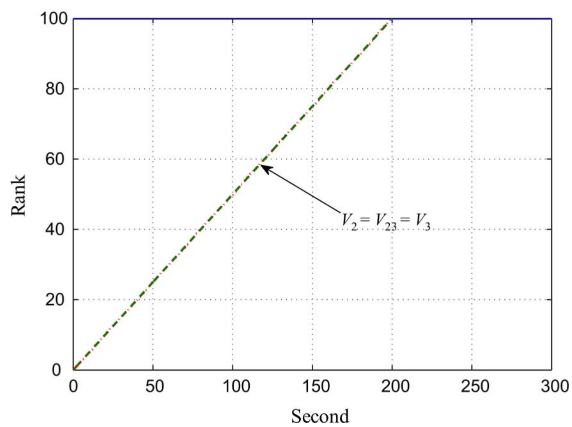

Fig. 8. Bottle neck phenomenon: rank at node 3 follows rank at node 2. Numerical solver exhibits fluctuation of V_3 around V_2.

By doing so, it is trivial to see that $V_2(t) = V_{\{23\}}(t) = V_3(t) = 0.5t$ is a solution for the system of differential inclusions for $t \in [0, 200)$.

IV. Analyzing Information Flows With DEDI—The Average Case

In this section, using the fluid approximation and the DEDI framework described in Section III, we setup the system of DI's that describes the average behavior of RNC applied to multiple concurrent multicast sessions. We derive the explicit DI solution as done in our earlier work [16]. We show from the solution that the average throughput of RNC is determined by the min-cut bound. The concentration behavior is presented in the next section.

For the average case, we begin by explicitly solving the deterministic DE(22). We will directly deal with multiple multicast sessions and the general topology. Then we will specialize the results to show that the average throughput of a single multicast session is determined by the min-cut bound. In general, suppose we have a wireless network $G = (\mathcal{N}, \mathcal{E})$ and J independent multicast sessions and session j originates from a set of source nodes

$$\mathcal{S}_j = \left\{s_{j,1}, s_{j,2}, \dots, s_{j,n_j}\right\},\ j = 1, 2, \dots, J \tag{28}$$

where each node in $\mathcal{S}_j$ contains the same set of m_j packets to be delivered to the rest of the network or part of it. Note it is possible that a node serves more than one multicast session and it contains as many sets of packets. To identify the source for any nonempty $\mathcal{K} \subset \mathcal{N}$, define

$$\mathrm{Src}(\mathcal{K}) = \left\{j | \mathcal{S}_j \cap \mathcal{K} = \emptyset, j = 1, 2, \dots, J\right\}. \tag{29}$$

For the coding scheme, we let each node generate a coded packet by randomly linearly mixing all the packets it holds, regardless which multicast sessions these packets belong to. Suppose all the multicast sessions start synchronously from time 0 as an integral information flow. This scenario is captured by the following system of DI's:

$$\dot{V}_{\mathcal{K}} \in \sum_{i \notin \mathcal{K}} z_{i,\mathcal{K}} \mathrm{Sgn}^{+}(V_{\{i\}\cup\mathcal{K}} - V_{\mathcal{K}})$$
$$\text{B.C.} \quad V_{\mathcal{K}}(0) = \sum_{\substack{1 \le j \le J \\ j \notin \mathrm{Src}(\mathcal{K})}} m_j. \tag{30}$$

In what follows, Theorem 1 explicitly solves (30). Its proof also shows the uniqueness of the solution.

Theorem 1: The solution to (30) is given recursively as

$$V_{\mathcal{K}}(t) = \min_{\mathcal{K}' \supset \mathcal{K}} \{V'_{\mathcal{K}}(0) + c(\mathcal{K}')t\} \tag{31}$$
$$= \min\{V_{\mathcal{K}}(0) + c(\mathcal{K})t, \min_{\mathcal{K}' \supset \mathcal{K}} \{V_{\mathcal{K}'}(t)\}\} \tag{32}$$
$$= \min\{V_{\mathcal{K}}(0) + c(\mathcal{K})t, \min_{\ell \notin \mathcal{K}} \{V_{\{\ell\}\cup\mathcal{K}}(t)\}\} \tag{33}$$

and

$$V_{\mathcal{N}}(t) = \sum_{j=1}^{J} m_j. \tag{34}$$

Besides, for each $\mathcal{K}$, there is a sequence

$$0 = t_0 < t_1 < \dots < t_{n_{\mathcal{K}}-1} < t_{n_{\mathcal{K}}} = \infty \tag{35}$$

such that over $[t_p, t_{p+1})$, $p = 0, 1, \dots, n_{\mathcal{K}} - 1$, $V_{\mathcal{K}}$ is affine

$$V_{\mathcal{K}}(t) = V_{\mathcal{K}'}(0) + c(\mathcal{K}')t,\ t \in [t_p, t_{p+1}) \tag{36}$$

where $\mathcal{K}'$ satisfies

$$\mathcal{K}' \in C_{\min}(\cup_{j \in \mathrm{Src}(\mathcal{K}')} \mathcal{S}_j, \mathcal{K}). \tag{37}$$

We need a few preliminaries for the proof of Theorem 1. We begin with Lemma 2 which gives a solution to (30) on an interval.

Lemma 2: Suppose $V_{\mathcal{K}}(t_1)$ is known, (30) has a solution on $[t_1, t_2)$ given as

$$V_{\mathcal{K}}(t) = V_{\mathcal{K}}(t_1) + z(t - t_1) \tag{38}$$

if there is a set of nodes $\mathcal{P}$ such that $\mathcal{P} \cap \mathcal{K} = \emptyset$ and $z \ge 0$ satisfying

1) $V_{\{i\}\cup\mathcal{K}}(t) = V_{\mathcal{K}}(t_1) + z(t - t_1), \forall t \in [t_1, t_2), \forall i \in \mathcal{P}$;
2) $V_{\{j\}\cup\mathcal{K}} \ge V_{\mathcal{K}}(t_1) + z(t - t_1), \forall t \in [t_1, t_2), \forall j \notin \mathcal{K} \cup \mathcal{P}$;
3) $\sum_{j \notin \mathcal{K}\cup\mathcal{P}} z_{j,\mathcal{K}} \le z \le \sum_{i \notin \mathcal{K}} z_{i,\mathcal{K}}$.

Proof: Suppose this is not true, there is $t'' \in [t_1, t_2)$ such that

$$V_{\mathcal{K}}(t'') \neq V_{\mathcal{K}}(t_1) + z(t - t_1).$$

Let

$$t' = \sup\{t < t''; V_{\mathcal{K}}(t) = V_{\mathcal{K}}(t_1) + z(t - t_1)\} \tag{39}$$

then t' exists, $t_1 \leq t' < t'' < t_2$ and $V_{\mathcal{K}}(t) \neq V_{\mathcal{K}}(t_1) + z(t - t_1)$ $\forall t \in (t', t'']$ by definition. Because $V_{\mathcal{K}}(t)$ is continuous, it is either

$$V_{\mathcal{K}}(t) > V_{\mathcal{K}}(t_1) + z(t - t_1),\ \forall t \in (t', t''] \tag{40}$$

or

$$V_{\mathcal{K}}(t) < V_{\mathcal{K}}(t_1) + z(t - t_1),\ \forall t \in (t', t'']. \tag{41}$$

If (40) holds, by assumption 1, $V_{\{i\}\cup\mathcal{K}}(t) < V_K(t)$, $\forall t \in (t', t'']$, $\forall i \in \mathcal{P}$. So, with assumption 2

$$\dot{V}_{\mathcal{K}}(t) = \sum_{i \notin \mathcal{K}} z_{i,\mathcal{K}} \mathrm{Sgn}^+(V_{\{i\}\cup\mathcal{K}} - V_{\mathcal{K}}) \leq \sum_{j \notin \mathcal{K}\cup\mathcal{P}} z_{j,\mathcal{K}} \tag{42}$$

thus

$$\begin{aligned} V_{\mathcal{K}}(t'') &= V_{\mathcal{K}}(t') + \int_{t'}^{t''} \dot{V}_{\mathcal{K}}(t)dt \\ &= V_{\mathcal{K}}(t_1) + z(t' - t_1) + \int_{t'}^{t''} \dot{V}_{\mathcal{K}}(t)dt \\ &\leq V_{\mathcal{K}}(t_1) + (t' - t_1) + \sum_{j \notin \mathcal{K}\cup\mathcal{P}} z_{j,\mathcal{K}}(t'' - t') \\ &\leq V_{\mathcal{K}}(t_1) + z(t' - t_1) + z(t'' - t') \quad \text{(assumption 3)} \\ &= V_{\mathcal{K}}(t_1) + z(t'' - t_1) \end{aligned} \tag{43}$$

which is a contradiction to (40). If (41) holds, by assumption 1 and 2

$$\dot{V}_{\mathcal{K}} = \sum_{i \notin \mathcal{K}} z_{i,\mathcal{K}} \mathrm{Sgn}^+(V_{\{i\}\cup\mathcal{K}} - V_{\mathcal{K}}) = \sum_{i \notin \mathcal{K}} z_{i,\mathcal{K}}. \tag{44}$$

Then by assumption 3

$$\begin{aligned} V_{\mathcal{K}}(t'') &= V_{\mathcal{K}}(t_1) + z(t' - t_1) + \int_{t'}^{t''} \dot{V}_{\mathcal{K}}(t)dt \\ &= V_{\mathcal{K}}(t_1) + z(t' - t_1) + \sum_{i \notin \mathcal{K}} z_{i,\mathcal{K}}(t'' - t') \\ &\geq V_{\mathcal{K}}(t_1) + z(t' - t_1) + z(t'' - t') \quad \text{(assumption3)} \\ &= V_{\mathcal{K}}(t_1) + z\,(t'' - t_1) \end{aligned} \tag{45}$$

which is a contradiction to (41). ■

Now we can give

Proof to Theorem 1: Since (33) and(32) are simply the recursive forms of(31), they are equivalent; hence, it suffices to prove (31). We prove this via induction on $|\mathcal{K}^c|$. When $|\mathcal{K}^c| = 0$, $\mathcal{K} = \mathcal{N}$, $V_{\mathcal{N}}(0) = \sum_{j=1}^{J} m_j$, $\forall t \geq 0$. Equation (31)–(32) are trivially true. Assume it is true when $|\mathcal{K}^c| \leq k - 1$, we prove it is also true for $|\mathcal{K}^c| = k$. Let

$$U_{\mathcal{K}}(t) = \min_{\mathcal{K}' \supset \mathcal{K}} \{V_{\mathcal{K}'}(0) + c(\mathcal{K}')t\} \tag{46}$$

then $U_{\mathcal{K}}(t)$ is piecewise linear (since it is the minimum of a finitely many affine functions) and there is a sequence

$$0 = t_0 < t_1 < \cdots < t_{n_{\mathcal{K}}-1} < t_{n_{\mathcal{K}}} = \infty \tag{47}$$

such that for each $p = 0, 1, \ldots, n_{\mathcal{K}} - 1$

$$U_{\mathcal{K}}(t) = V_{\mathcal{K}'}(0) + c(\mathcal{K}')t, \quad t \in [t_p, t_{p+1}) \tag{48}$$

for some $\mathcal{K}'$. We claim $\mathcal{K}' \in C_{\min}(\cup_{j\in\mathrm{Src}(\mathcal{K}')}\mathcal{S}_j, \mathcal{K})$ (it is apparent that $\mathcal{K}' \in C(\cup_{j\in\mathrm{Src}(\mathcal{K}')}\mathcal{S}_j, \mathcal{K})$). Otherwise, let $\mathcal{K}'' \in C_{\min}(\cup_{j\in\mathrm{Src}(\mathcal{K}')}\mathcal{S}_j, \mathcal{K})$ but $\mathcal{K}'' \neq \mathcal{K}'$. By definition of min cut for the hypergraph model, we have

$$c(\mathcal{K}'') < c(\mathcal{K}'). \tag{49}$$

Since $(\cup_{j\in\mathrm{Src}(\mathcal{K}')}\mathcal{S}_j) \cap \mathcal{K}'' = \emptyset$, $\mathrm{Src}(\mathcal{K}') \subset \mathrm{Src}(\mathcal{K}'')$, so

$$V_{\mathcal{K}''}(0) \leq V_{\mathcal{K}'}(0). \tag{50}$$

Therefore, $\forall t \in (t_p, t_{p+1})$

$$V_{\mathcal{K}''}(0) + c(\mathcal{K}'')t < V_{\mathcal{K}'}(0) + c(\mathcal{K}')t \tag{51}$$

which is a contradiction to (46).

We want to show that $V_{\mathcal{K}}(t) = U_{\mathcal{K}}(t)$, $\forall t \in [t_p, t_{p+1})$, using Lemma 2, which amounts to checking three conditions. Let $\mathcal{K}'$ be as in (48). Let $\mathcal{P} = \mathcal{K}' \setminus \mathcal{K}$, $z = \sum_{i\notin\mathcal{K}'} z_{i,\mathcal{K}'} = c(\mathcal{K}')$. First note

$$V_{\{i\}\cup\mathcal{K}}(t) = U_{\mathcal{K}}(t), \quad \forall i \in \mathcal{P}, \forall t \in [t_p, t_{p+1}). \tag{52}$$

This is because, on one hand, $V_{\{i\}\cup\mathcal{K}}(t) \geq U_{\mathcal{K}}(t)$ by (46), while on the other hand, by induction assumption ($|\mathcal{N} \setminus (\{i\} \cup \mathcal{K})| = k - 1$)

$$\begin{aligned} V_{\{i\}\cup\mathcal{K}}(t) &= \min_{(\{i\}\cup\mathcal{K})\subset\mathcal{K}''} \{V_{\mathcal{K}''}(0) + c(\mathcal{K}'')t\} \\ &\leq V_{\mathcal{K}'}(0) + c(\mathcal{K}')t \quad (\text{since } \{i\} \cup \mathcal{K} \subset \mathcal{K}') \\ &= U_{\mathcal{K}}(t). \end{aligned} \tag{53}$$

Meanwhile, by (46) we have

$$V_{\{j\}\cup\mathcal{K}}(t) \geq U_{\mathcal{K}}(t), \quad \forall t \in [t_p, t_{p+1}), \forall j \notin \mathcal{K} \cup \mathcal{P}. \tag{54}$$

Because $\mathcal{K}' = C_{\min}(\cup_{j\in\mathrm{Src}(\mathcal{K}')}\mathcal{S}_j, \mathcal{K})$, $z \leq \sum_{i\notin\mathcal{K}} z_{i,\mathcal{K}}$. Because $\mathcal{K} \subset \mathcal{K}'$,

$$\sum_{i\notin\mathcal{K}'} z_{i,\mathcal{K}} \leq \sum_{i\notin\mathcal{K}'} z_{i,\mathcal{K}'} = z. \tag{55}$$

Thus, assumption 3 of Lemma 2 is checked for all t. We then check assumption 1 and 2 piecewise. Because for any $\mathcal{K}'' \supset \mathcal{K}$

$$\begin{aligned} V_{\mathcal{K}''}(0) &= \sum_{j=1}^{J} I(\mathcal{S}_j \cap \mathcal{K}'' \neq \emptyset) m_j \\ &\geq \sum_{j=1}^{J} I(\mathcal{S}_j \cap \mathcal{K} \neq \emptyset) m_j = V_{\mathcal{K}}(0) \end{aligned} \tag{56}$$

by (46) we have

$$V_{\mathcal{K}'}(0) = V_{\mathcal{K}}(0). \tag{57}$$

From (52) and (57)

$$\begin{aligned} V_{\{i\}\cup\mathcal{K}}(t) &= V_{\mathcal{K}'}(0) + c(\mathcal{K}')t \\ &= V_{\mathcal{K}}(0) + c(\mathcal{K}')t, \forall t \in [t_0, t_1), \forall i \in \mathcal{P} \end{aligned} \tag{58}$$

Hence, assumption 1 is checked for $[t_0, t_1)$. From (54) and (57)

$$\begin{aligned} V_{\{j\}\cup\mathcal{K}}(t) &\geq V_{\mathcal{K}'}(0) + c(\mathcal{K}')t \\ &= V_{\mathcal{K}}(0) + c(\mathcal{K}')t,\ \forall t \in [t_0, t_1), \forall j \notin \mathcal{K} \cup \mathcal{P}. \end{aligned} \tag{59}$$

Hence, assumption 2 is checked for $[t_0, t_1)$. Therefore, $V_{\mathcal{K}}(t) = U_{\mathcal{K}}(t)$, $\forall t \in [t_0, t_1)$. But this in turn implies that $V_{\mathcal{K}}(t_1) = U_{\mathcal{K}}(t_1)$ by continuity (cf. the definition of solution to DI in Section III-B), which implies that assumption 1 and 2 are also checked for $[t_1, t_2)$ (same argument as for $[t_0, t_1)$). Therefore, $V_{\mathcal{K}}(t) = U_{\mathcal{K}}(t), \forall t \in [t_1, t_2)$. Repeat this argument $n_{\mathcal{K}}$ times, we conclude that $V_{\mathcal{K}}(t) = U_{\mathcal{K}}(t), \forall t \geq 0$. This shows the validity of (31) for $|\mathcal{K}^c| = k$.

Essentially, Theorem 1 (c.f (31)) states that $V_{\mathcal{K}}(t)$ is the min-envelop of $2^{|\mathcal{K}^c|}$ affine functions corresponding to so many subsets of nodes that contain $\mathcal{K}$. The partial order "$\precsim$" illustrated by the layered structure also implies the usual linear order "$\leq$", i.e.,

$$\mathcal{K} \precsim \mathcal{K}' \Rightarrow V_{\mathcal{K}}(t) \leq V'_{\mathcal{K}}(t), \quad \forall t \geq 0. \tag{60}$$

Therefore, it is always true

$$V_{\mathcal{K}} \leq V_{\mathcal{N}} = \sum_{j=1}^{J} m_j. \tag{61}$$

A stronger statement than (61) can be made when G is connected, i.e.,

Corollary 1: If $G = (\mathcal{N}, \mathcal{E})$ is connected, then $\forall \mathcal{K} \neq \emptyset$

$$V_{\mathcal{K}}(t) = \sum_{j=1}^{J} m_j, \quad t \in [t_{n_{\mathcal{K}}-1}, \infty). \tag{62}$$

Proof: Because G is connected, $\forall \mathcal{K} \subset \mathcal{K}' \subsetneq \mathcal{N}$, $c(\mathcal{K}') > 0$. Therefore, when t is sufficiently large

$$V'_{\mathcal{K}}(0) + c(\mathcal{K}')t > \sum_{j=1}^{J} m_j = V_{\mathcal{N}}(t).$$

Hence, we have the conclusion from (31) of Theorem 1. ■

Corollary 2 implies that with the RNC scheme for multiple flows as described here, a node may have to wait until its rank reaches $\sum_{j=1}^{J} m_j$ to start decoding. This time is denoted as $T_{\mathcal{K}}^{\text{total}}$. Though there could be fairly large decoding delay for nodes only interested in one or few sessions, the intersession coding is optimal in the sense of min cut bound. Applying (36) in Theorem 1 to $[t_{n_{\mathcal{K}}-1}, t_{n_{\mathcal{K}}})$, it is clear that $T_{\mathcal{K}}^{\text{total}}$ is determined by one of the min cut bounds that $\mathcal{K}$ has to take into consideration. The min cut that determines the finish time can be regarded as the worst bottleneck for $\mathcal{K}$. More precisely, we have

Corollary 2: If $G = (\mathcal{N}, \mathcal{E})$ is connected, then

$$T_{\mathcal{K}}^{\text{total}} = \max_{S \subset \text{Src}(\mathcal{K})} \left\{ \sum_{j \in S} m_j / c_{\min}(\cup_{i \in S} \mathcal{S}_i, \mathcal{K})) \right\}. \tag{63}$$

Proof: Clearly $T_{\mathcal{K}}^{\text{total}} = t_{n_{\mathcal{K}}-1}$. By Theorem 1, there is $\mathcal{K}' \supset \mathcal{K}$ such that $\forall t \in [t_{n_{\mathcal{K}}-2}, t_{n_{\mathcal{K}}-1})$

$$\begin{aligned} V_{\mathcal{K}}(t) &= V_{\mathcal{K}'}(0) + c(\mathcal{K}')t \\ &= \sum_{j \notin \text{Src}(\mathcal{K}')} m_j + c_{\min}(\cup_{i \in \text{Src}(\mathcal{K}')} \mathcal{S}_i, \mathcal{K})t \end{aligned} \tag{64}$$

and by setting $V_{\mathcal{K}}(t_{n_{\mathcal{K}}-1}) = \sum_{j=1}^{J} m_j$, we get

$$\begin{aligned} T_{\mathcal{K}}^{\text{total}} &= t_{n_{\mathcal{K}}-1} \\ &= \left(\sum_{j=1}^{J} m_j - \sum_{\substack{1 \leq j' \leq J \\ j' \notin \text{Src}(\mathcal{K}')}} m_{j'} \right) / c_{\min}\left(\cup_{i \in \text{Src}(\mathcal{K}')} \mathcal{S}_i, \mathcal{K}\right) \\ &= \sum_{j \in \text{Src}(\mathcal{K}')} m_j / c_{\min}(\cup_{i \in \text{Src}(\mathcal{K}')} \mathcal{S}_i, \mathcal{K}) \\ &\leq \max_{S \subset \text{Src}(\mathcal{K})} \left\{ \sum_{j \in S} m_j / c_{\min}(\cup_{i \in S} \mathcal{S}_i, \mathcal{K})) \right\} \end{aligned} \tag{65}$$

where the last inequality holds because $\mathcal{K}' \supset \mathcal{K}$; hence, $\text{Src}(\mathcal{K}') \subset \text{Src}(\mathcal{K})$. However, if there is $S' \subset \text{Src}(\mathcal{K})$, such that

$$\sum_{j \in S'} m_j / c_{\min}(\cup_{i \in S'} \mathcal{S}_i, \mathcal{K}) > T_{\mathcal{K}}^{\text{total}} \tag{66}$$

let $\mathcal{K}'' = C_{\min}(\cup_{i \in S'} \mathcal{S}_i, \mathcal{K})$, then we have

$$V_{\mathcal{K}''}(0) \leq \sum_{\substack{1 \leq j'' \leq J \\ j'' \notin S'}} m_{j''} \tag{67}$$

because $S' \subset \text{Src}(\mathcal{K}'')$, and

$$\begin{aligned} &V_{\mathcal{K}''}(0) + c(\mathcal{K}'')T_{\mathcal{K}}^{\text{total}} \\ &= V_{\mathcal{K}''}(0) + c_{\min}\left(\cup_{i \in S'} \mathcal{S}_i, \mathcal{K}\right) T_{\mathcal{K}}\text{total} \\ &< \sum_{\substack{1 \leq j'' \leq J \\ j'' \notin S'}} m_{j''} + \sum_{j \in S'} m_j \\ &= \sum_{j=1}^{J} m_j = V_{\mathcal{K}}(T_{\mathcal{K}}^{\text{total}}) \end{aligned} \tag{68}$$

which contradicts (60). So

$$T_{\mathcal{K}}^{\text{total}} \geq \max_{S \subset \text{Src}(\mathcal{K})} \left\{ \sum_{j \in S} m_j / c_{\min}(\cup_{i \in S} \mathcal{S}_i, \mathcal{K})) \right\}. \tag{69}$$

Combine (65) with (69), we get (63). ■

From Theorem 1 and Corollary 2 we can readily show that the average throughput under the fluid approximation is given by the min-cut bound, i.e., we have.

Theorem 2: If $G = (\mathcal{N}, \mathcal{E})$ is connected and node 1 is the single source of a multicast session, the solution to (22) with B.C. Equation (17) that describes this scenario is given as:

1) $\forall \mathcal{K} \subset \mathcal{N}$ and $1 \in \mathcal{K}$

$$V_{\mathcal{K}}(t) = m, \ \forall t \in [0, \infty). \tag{70}$$

2) $\forall \mathcal{K} \subset \mathcal{N}$ and $1 \notin \mathcal{K}$

$$V_{\mathcal{K}}(t) = \begin{cases} c_{\min}(1,\mathcal{K})t, & \forall t \in [0, m/c_{\min}(1,\mathcal{K})) \\ m, & \forall t \in [m/c_{\min}(1,\mathcal{K}), \infty). \end{cases} \tag{71}$$

Proof: By Theorem 1, the solution to (22) for each $\mathcal{K}$ is piecewise linear. If $1 \in \mathcal{K}$, for any $\mathcal{K}'$ that satisfy (36) and (37), we must have $V_{\mathcal{K}'}(0) = m$ and $c(\mathcal{K}') = 0$. This implies $V_{\mathcal{K}}(t)$ is a constant. By Corollary 2, we have (70). If $1 \notin \mathcal{K}$, there are two possibilities:

1) $\mathcal{K}' \in C(1, \mathcal{K})$ and $c(\mathcal{K}') = c_{\min}(1, \mathcal{K})$;
2) $\mathcal{K}' \in C(\emptyset, \mathcal{K})$ and $c(\mathcal{K}') = 0$.

By Corollary 2, the second case applies to $t \in [t_1, \infty)$; hence, the first case applies to $t \in [t_0, t_1)$ as we know $V_{\mathcal{K}}(0) = 0$. We, therefore, have (71).

Theorem 2 states that the rank of $\mathcal{K}$ increases until it reaches m at the rate allowed by the min cut that separates $\mathcal{K}$ from the source.

Corollary 3: For $1 \notin \mathcal{K}$, $\dot{V}_{\mathcal{K}} = c_{\min}(1, \mathcal{K})$, when $V_{\mathcal{K}} < m$.

Specializing Corollary 3 to an arbitrary destination node i, we obtain:

Corollary 4: For $i \neq 1$, $\dot{V}_i = c_{\min}(1, i)$, when $V_i < m$.

Corollary 4 shows that if a unicast at average rate R exists for each destination i separately, i.e., $c_{\min}(1, i) \geq R$, then the proposed coding scheme is capable to implement a multicast at average rate R.

V. Concentration Behavior of Dedi Solution—the Asymptotic Case

Section IV presented an average analysis of RNC throughput based on the fluid approximation. In this section we show that this throughput can be achieved asymptotically with increasing number of source packets m. This asymptotic result was previously proven in [9] using a queueing approach and graph decomposition. In our paper, we begin with (6) and solely work with differential equations to show the same result. This perspective on RNC is new.

To motivate the achievability problem, we first prove a weak[1] version of min-cut max-flow theorem for RNC.

Theorem 3: Assume node 1 is the only source in the network and the transmission begins at $t = 0$. Let $N_{\mathcal{K}}(t)$ be the incremental process of innovative packets at $\mathcal{K} \subset \mathcal{N} \setminus \{1\}$, then

$$E[N_{\mathcal{K}}(t)] \leq c_{\min}(1, \mathcal{K})t. \tag{72}$$

[1]A stronger version would say $\mathrm{d}E[N_{\mathcal{K}}(t)]/\mathrm{d}t \leq c_{\min}(1, \mathcal{K})$.

Proof: Suppose $\mathcal{T} \in C_{\min}(1, \mathcal{K})$. We have, from Lemma 1

$$\frac{\mathrm{d}E[N_{\mathcal{T}}(t)]}{\mathrm{d}t} = \sum_{i \in \mathcal{T}^c} z_{i,\mathcal{T}} E\left[1 - q^{N_{\mathcal{T}}(t) - N_{\{i\}\cup\mathcal{T}}(t)}\right] \tag{73}$$

$$\leq \sum_{i \in \mathcal{T}^c} z_{i,\mathcal{T}} = c_{\min}(1, \mathcal{K}). \tag{74}$$

Since $N_{\mathcal{K}}(0) = 0$, we have

$$E[N_{\mathcal{K}}(t)] \leq E[N_{\mathcal{T}}(t)] \leq c_{\min}(1, \mathcal{K})t. \tag{75}$$

Though Theorem 3 indicates that the time average throughput of RNC is governed by the min-cut bound, we will show that the min-cut bound can be asymptotically achieved. In this section, we assume the hypergraph is connected and we give the asymptotic achievability proof of RNC within the DEDI framework. For any $\mathcal{K} \subset \mathcal{N}$, we let $D_{\mathcal{K}}(m_1, m_2)$ denote the time taken for $N_{\mathcal{K}}$ to increase from m_1 to m_2. We will prove $\forall 1 > \epsilon > 0$

$$\lim_{m \to \infty} P\left(m/D_{\mathcal{K}}(0, m) > \epsilon c_{\min}(1, \mathcal{K})\right) = 1. \tag{76}$$

In order to prove this, we will follow the strategy outlined as below: We begin with Lemma 3 that is fundamental for the argument. Lemma 4 builds a system of DE's for $\mathrm{Var}[N_{\mathcal{K}}(t)]$, from which we will prove Theorem 4 with the help of Lemma 3 that says the standard deviation of $N_{\mathcal{K}}(t)$ is upper bounded by a sublinear function. Then we scale $N_{\mathcal{K}}(t)$ to produce a new process $M_{\mathcal{K}}(t)$ and show in Theorem 5 that $M_{\mathcal{K}}(t)$ converges in probability to $c_{\min}(1, \mathcal{K})t$. While the following Corollary 5 implies that the throughput is given by $c_{\min}(1, \mathcal{K})$ for the entire course of transmission except the last few packets, Proposition 2 and Corollary 6 show that the last few packets take bounded time to get transmitted. Theorem 6 combines Corollary 5 and Corollary 6 to complete the achievability proof.

Lemma 3: Let X be an r.v. and let $f : \mathbb{R} \to \mathbb{R}$ be a decreasing function, then

$$E[Xf(X)] \leq E[X]E[f(X)]. \tag{77}$$

In other words

$$\mathrm{Cov}[X, f(X)] \leq 0. \tag{78}$$

Proof: Let X' be an independent copy of X. Since f is decreasing, we have

$$(X - X')(f(X) - f(X')) \leq 0 \quad \text{a.s.} \tag{79}$$

so

$$E\left[(X - X')(f(X) - f(X')\right] \leq 0. \tag{80}$$

Expanding (80) we have (77). ■

It should be pointed out that a more general statement of Lemma 3 can be found in [17].

Lemma 4: For any set $\mathcal{K}$ of nodes, we have

$$\frac{\mathrm{d}\mathrm{Var}[N_{\mathcal{K}}(t)]}{\mathrm{d}t} = \frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} + 2\sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}}\mathrm{Cov}[N_{\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}]. \quad (81)$$

Proof: The proof is manifest in the following computation.

$$N_{\mathcal{K}}(t+\Delta t) = N_{\mathcal{K}}(t) + \Delta_{\mathcal{K}}(t) \quad (82)$$

so

$$N_{\mathcal{K}}^2(t) = N_{\mathcal{K}}^2(t) + \Delta_{\mathcal{K}}^2(t) + 2N_{\mathcal{K}}(t)\Delta_{\mathcal{K}}(t). \quad (83)$$

We have

$$E\left[\Delta_{\mathcal{K}}^2(t)\right] = \sum_{i\in\mathcal{K}^2} \lambda_i \Delta t P_{i,\mathcal{K}} E\Big[1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\Big] + o(\Delta t) \quad (84)$$

and

$$E\left[N_{\mathcal{K}}(t)\Delta_{\mathcal{K}}(t)\right] = \sum_{i\in\mathcal{K}^c} \lambda_i \Delta t P_{i,\mathcal{K}} E\Big[N_{\mathcal{K}}(t)\Big(1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\Big)\Big] + o(\Delta t). \quad (85)$$

Therefore

$$\begin{aligned}\frac{\mathrm{d}E\left[N_{\mathcal{K}}^2(t)\right]}{\mathrm{d}t} &= \sum_{i\in\mathcal{K}^2} \lambda_i P_{i,\mathcal{K}} E\left[1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right] \\ &\quad + \sum_{i\in\mathcal{K}^c} \lambda_i P_{i,\mathcal{K}} E\left[N_{\mathcal{K}}(t)(1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)})\right] \\ &= \frac{\mathrm{d}E\left[N_{\mathcal{K}}(t)\right]}{\mathrm{d}t} + \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} E\left[N_{\mathcal{K}}(t)\left(1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right)\right] \quad (86)\end{aligned}$$

where the last equality follows from Lemma 1. We also have

$$\begin{aligned}\frac{\mathrm{d}E^2[N_{\mathcal{K}}(t)]}{\mathrm{d}t} &= 2E[N_{\mathcal{K}}(t)]\frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} \\ &= 2\sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} E[N_{\mathcal{K}}(t)]E[1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}] \quad (87)\end{aligned}$$

so

$$\begin{aligned}\frac{\mathrm{d}\mathrm{Var}[N_{\mathcal{K}}(t)]}{\mathrm{d}t} &= \frac{\mathrm{d}E[N_{\mathcal{K}}^2(t)]}{\mathrm{d}t} - \frac{\mathrm{d}E^2[N_{\mathcal{K}}(t)]}{\mathrm{d}t} \\ &= \frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} + 2\sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}}\mathrm{Cov}\Big[N_{\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\Big]. \quad (88)\end{aligned}$$

■

Theorem 4 bounds $\sqrt{\mathrm{Var}[N_{\mathcal{K}}(t)]}$ with a sublinear function $d_{\mathcal{K}}t^{1-\delta_{\mathcal{K}}/2}$ in time. Since we know $E[N_{\mathcal{K}}(t)] \approx c_{\min}(1,\mathcal{K})t$ from Section IV, it later enables us to use Chebyshev inequality to show concentration of $N_{\mathcal{K}}(t)$.

Theorem 4: There exists constants $d_{\mathcal{K}} > 0$ and $1 \geq \delta_{\mathcal{K}} > 0$ for each $\mathcal{K}$ independent of q, m, such that

$$\limsup_{t\to\infty} \frac{\mathrm{Var}[N_{\mathcal{K}}(t)]}{t^{2-\delta_{\mathcal{K}}}} \leq d_{\mathcal{K}}. \quad (89)$$

Proof: We first make a trivial observation that, if we let

$$a_{\mathcal{K}} = \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} \quad (90)$$

then

$$\frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} \leq a_{\mathcal{K}} \quad (91)$$

by Lemma 1. The rest of the proof is by induction using the partial order. When $\mathcal{K} = \{1\}^c$, we have

$$\frac{\mathrm{d}\mathrm{Var}\left[N_{\mathcal{K}}(t)\right]}{\mathrm{d}t} = \frac{\mathrm{d}E\left[N_{\mathcal{K}}(t)\right]}{\mathrm{d}t} + 2\sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}}\mathrm{Cov}\left[N_{\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-m}\right]. \quad (92)$$

Since $1 - q^{N_{\mathcal{K}}(t)-m}$ is decreasing and concave in $N_{\mathcal{K}}(t)$, by Lemma 3, we know

$$\frac{\mathrm{d}\mathrm{Var}[N_{\mathcal{K}}(t)]}{\mathrm{d}t} \leq \frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} \quad (93)$$

so

$$\mathrm{Var}[N_{\mathcal{K}}(t)] \leq E[N_{\mathcal{K}}(t)] \leq a_{\mathcal{K}}t \quad (94)$$

by (91), as $\mathrm{Var}[N_{\mathcal{K}}(0)] = E[N_{\mathcal{K}}(0)] = 0$. Therefore, we can assign $d_{\mathcal{K}} = a_{\mathcal{K}}$ and $\delta_{\mathcal{K}} = 1$. Now suppose the statement is true $\forall\mathcal{K}' \supset \mathcal{K}$, $\mathcal{K}' \neq \mathcal{K}$, we prove it is also true for $\mathcal{K}$. First note

$$\begin{aligned}&\mathrm{Cov}\left[N_{\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right] \\ &= \mathrm{Cov}\left[N_{\mathcal{K}}(t) - N_{\{i\}\cup\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right] \\ &\quad + \mathrm{Cov}\left[N_{\{i\}\cup\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right]. \quad (95)\end{aligned}$$

The first term on the right-hand side of (95) is nonpositive by Lemma 3. By Cauchy–Schwarz inequality, the second term can be upper bounded as

$$\begin{aligned}&\mathrm{Cov}\left[N_{\{i\}\cup\mathcal{K}}(t), 1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right] \\ &\leq \sqrt{\mathrm{Var}\left[N_{\{i\}\cup\mathcal{K}}(t)\right]}\sqrt{\mathrm{Var}\left[1 - q^{N_{\mathcal{K}}(t)-N_{\{i\}\cup\mathcal{K}}(t)}\right]} \\ &\leq \sqrt{\mathrm{Var}\left[N_{\{i\}\cup\mathcal{K}}(t)\right]} \\ &\leq \sqrt{d_{\{i\}\cup\mathcal{K}}}\, t^{1-\delta_{\{i\}\cup\mathcal{K}}/2} \quad (96)\end{aligned}$$

so

$$\frac{\mathrm{dVar}[N_{\mathcal{K}}(t)]}{\mathrm{d}t} \leq \frac{\mathrm{d}E[N_{\mathcal{K}}(t)]}{\mathrm{d}t} + 2\sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}}\sqrt{d_{\{i\}\cup\mathcal{K}}}t^{1-\delta_{\mathcal{K}}/2}. \quad (97)$$

By induction, we have

$$\mathrm{Var}[N_{\mathcal{K}}(t)] \leq a_{\mathcal{K}}t + 2\sum_{i\in\mathcal{K}^c} \frac{z_{i,\mathcal{K}}\sqrt{d_{\{i\}\cup\mathcal{K}}}}{2-\delta_{\{i\}\cup\mathcal{K}}/2} t^{2-\delta_{\{i\}\cup\mathcal{K}}/2} \quad (98)$$

as $\mathrm{Var}[N_{\mathcal{K}}(0)] = E[N_{\mathcal{K}}(t)] = 0$. Therefore, we can pick

$$d_{\mathcal{K}} = a_{\mathcal{K}} + 2\sum_{i\in\mathcal{K}^c} \frac{z_{i,\mathcal{K}}\sqrt{d_{\{i\}\cup\mathcal{K}}}}{2-\delta_{\{i\}\cup\mathcal{K}}/2} \quad (99)$$

and

$$\delta_{\mathcal{K}} = \min\{1, \min_{i\in\mathcal{K}^c}\{\delta_{\{i\}\cup\mathcal{K}}/2\}\}. \quad (100)$$

The next result is more conveniently discussed in terms of the scaled process defined as

$$M_{\mathcal{K}}^{(m)}(t) \triangleq \frac{1}{m}N_{\mathcal{K}}(mt). \quad (101)$$

The scaled process $M_{\mathcal{K}}^{(m)}(t)$ contains essentially the same information as $N_{\mathcal{K}}(t)$ except its graph is scaled down with a fixed aspect ratio. We list some obvious properties of $M_{\mathcal{K}}^{(m)}(t)$.

Proposition 1: $M_{\mathcal{K}}^{(m)}(t)$ has the following properties:

1) $E[M_{\mathcal{K}}^{(m)}(t)]$ is increasing from 0 to 1 for $1\notin\mathcal{K}$.
2) $M_{\mathcal{K}}^{(m)}(t) = 1$ for $1\in\mathcal{K}$. For $1\notin\mathcal{K}$, $M_{\mathcal{K}}^{(m)}(t)$ satisfies

$$\frac{\mathrm{d}E\left[M_{\mathcal{K}}^{(m)}(t)\right]}{\mathrm{d}t} = \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}}E\left[1-(q^m)^{M_{\mathcal{K}}^{(m)}(t)-M_{\{i\}\cup\mathcal{K}}^{(m)}(t)}\right]. \quad (102)$$

3)

$$E\left[M_{\mathcal{K}}^{(m)}(t)\right] = \frac{1}{m}E[N_{\mathcal{K}}(mt)] \leq c_{\min}(1,\mathcal{K})t, \quad \forall t\in[0, m/c_{\min}(1,\mathcal{K})), \forall\mathcal{K}\not\ni 1. \quad (103)$$

4) $\{E[M_{\mathcal{K}}^{(m)}(t)]\}_m$ are uniformly Lipschitz continuous.
5) $\exists d_{\mathcal{K}} > 0$ and $\delta_{\mathcal{K}} > 0$ independent of m, such that

$$\limsup_{m\to\infty} \frac{\mathrm{Var}\left[M_{\mathcal{K}}^{(m)}(t)\right]}{m^{-\delta_{\mathcal{K}}}} \leq d_{\mathcal{K}}t^{2-\delta_{\mathcal{K}}} \quad (104)$$

for fixed t.

Proof: Property 1) is obvious. Property 2) follows from Lemma 1. Property 3) follows from Theorem 3. Property 4) follows from

$$\frac{\mathrm{d}E\left[N_{\mathcal{K}}(t)\right]}{\mathrm{d}t} \leq \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} \quad (105)$$

and

$$\frac{\mathrm{d}E\left[M_{\mathcal{K}}^{(m)}(t)\right]}{\mathrm{d}t} = E\frac{[N_{\mathcal{K}}(\tau)]}{\mathrm{d}\tau}\Big|_{\tau=mt}. \quad (106)$$

Property 5) follows from Lemma 4

$$\limsup_{m\to\infty} \frac{\mathrm{Var}\left[M_{\mathcal{K}}^{(m)}(t)\right]}{m^{-\delta_{\mathcal{K}}}} = \limsup_{m\to\infty} \frac{\mathrm{Var}\left[\frac{1}{m}N_{\mathcal{K}}(mt)\right]}{m^{-\delta_{\mathcal{K}}}} \leq d_{\mathcal{K}}t^{2-\delta_{\mathcal{K}}}. \quad (107)$$

Note that property 3) of Proposition 1 provides an upper bound on $E[M_{\mathcal{K}}^{(m)}(t)]$, we will show that in fact $M_{\mathcal{K}}^{(m)}(t)$ asymptotically concentrates to this upper bound, i.e.,

Theorem 5: Given $\mathcal{K}$, $M_{\mathcal{K}}^{(m)}(t) \xrightarrow{p} c_{\min}(1,\mathcal{K})t$, $\forall t\in(0, 1/c_{\min}(1,\mathcal{K})]$ as $m\to\infty$.

Proof: Due to property 4) of Proposition 1, $\{E[M_{\mathcal{K}}^{(m)}(t)]\}_m$ are Lipschitz continuous hence equicontinuous. They are also uniformly bounded by constant 1. By the Arzelà-Ascoli Theorem, $\forall\mathcal{K}\subset\mathcal{N}$, a subsequence $\{E[M_{\mathcal{K}}^{(m_\ell)}(t)]\}_{\ell=1}^{\infty}$ converges uniformly to some continuous function $M_{\mathcal{K}}(t)$ over $[0, 1/c_{\min}(1,\mathcal{K})]$. In fact, we can pick a single sequence $\{E[M_{\mathcal{K}}^{(m_\ell)}(t)]\}$ that converges uniformly to $M_{\mathcal{K}}(t)$ for all $\mathcal{K}$ because there are only a finite number of $\mathcal{K}\subset\mathcal{N}$. So we make this assumption and in what follows we prove $M_{\mathcal{K}}(t) = c_{\min}(1,\mathcal{K})t$ by contradiction and via induction on the partial order of $\mathcal{K}$.

When $\mathcal{K} = \{1\}^c$. Suppose at $t_2\in(0, 1/c_{\min}(1,\mathcal{K})]$, $c_{\min}(1,\mathcal{K})t_2 - M_{\mathcal{K}}(t_2) = h > 0$. Because $M_{\mathcal{K}}(0) = 0$, by continuity, $\exists 0 < t_1 < t_2$, such that $c_{\min}(1,\mathcal{K})t_1 - M_{\mathcal{K}}(t_1) = h/2$ and

$$c_{\min}(1,\mathcal{K})t - M_{\mathcal{K}}(t) \geq h/2, \quad \forall t\in[t_1, t_2]. \quad (108)$$

Let

$$\lambda_{\mathcal{K}}^{\ell} = m_\ell^{\delta_{\mathcal{K}}/4}, \text{ then } \lim_{\ell\to\infty}\frac{1}{(\lambda_{\mathcal{K}}^{\ell})^2} = 0,. \quad (109)$$

From property 5) of Proposition 1, we know

$$\lim_{\ell\to\infty}\lambda_{\mathcal{K}}^{\ell}\sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right]} = 0, \text{ uniformly } \forall t\in[t_1, t_2]. \quad (110)$$

Because $E[M_{\mathcal{K}}^{(m_\ell)}(t)] \xrightarrow{u} M_{\mathcal{K}}(t)$ on $[0, 1/c_{\min}(1,\mathcal{K})]$, $\exists L_1$ such that $\forall\ell > L_1$

$$|E[M_{\mathcal{K}}^{(m_\ell)}(t)] - M_{\mathcal{K}}(t)| < h/8, \quad \forall t\in[t_1, t_2] \quad (111)$$

so

$$E[M_{\mathcal{K}}^{(m_\ell)}(t)] < M_{\mathcal{K}}(t) + h/8, \forall t\in[t_1, t_2], \forall\ell > L_1 \quad (112)$$

and in particular $\forall\ell > L_1$

$$E[M_{\mathcal{K}}^{(m_\ell)}(t_1)] > c_{\min}(1,\mathcal{K})t_1 - 5h/8 \quad (113)$$
$$E[M_{\mathcal{K}}^{(m_\ell)}(t_2)] < c_{\min}(1,\mathcal{K})t_2 - 7h/8. \quad (114)$$

Let $p(\gamma)$ be the distribution function of $M_{\mathcal{K}}^{(m_\ell)}(t)$ and let A be the interval

$$A \triangleq \left[0, E\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right] + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right]}\right] \quad (115)$$

then we may calculate

$$\begin{aligned}
&\frac{\mathrm{d}E[M_{\mathcal{K}}^{(m_\ell)}(t)]}{\mathrm{d}t} \\
&= z_{1,\mathcal{K}} E\left[1 - (q^{m_\ell})^{M_{\mathcal{K}}^{(m_\ell)}(t)-1}\right] \\
&\geq \int_A p(\gamma) z_{1,\mathcal{K}} (1 - (q^{m_\ell})^{\gamma-1}) \mathrm{d}\gamma \\
&\geq \int_A p(\gamma) z_{1,\mathcal{K}} \left(1 - (q^{m_\ell})^{E\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right] + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right] - 1}}\right) \mathrm{d}\gamma \\
&> \int_A p(\gamma) z_{1,\mathcal{K}} \left(1 - (q^{m_\ell})^{M_{\mathcal{K}}(t) + h/8 + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right] - 1}}\right) \mathrm{d}\gamma \\
&= z_{1,\mathcal{K}} \left(1 - (q^{m_\ell})^{M_{\mathcal{K}}(t) + h/8 + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right] - 1}}\right) \\
&\cdot P\left(M_{\mathcal{K}}^{(m_\ell)}(t) < E[M_{\mathcal{K}}^{(m_\ell)}(t)] + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right]}\right).
\end{aligned} \quad (116)$$

But

$$P\left(M_{\mathcal{K}}^{(m_\ell)}(t) < E\left[M_{\mathcal{K}}^{(am_\ell)}(t)\right] + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right]}\right) \geq 1 - \frac{1}{(\lambda_{\mathcal{K}}^\ell)^2} \to 1, \quad (\ell \to \infty) \quad (117)$$

by Chebyshev inequality, which is true $\forall t \in [t_1, t_2]$ uniformly. Using (108), (110) and noticing that $t \leq t_2 \leq 1/c_{\min}(1, \mathcal{K})$, we have

$$\limsup_{\ell \to \infty} M_{\mathcal{K}}(t) + h/8 + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}[M_{\mathcal{K}}^{(m_\ell)}(t)]} - 1 \leq c_{\min}(1, \mathcal{K})t - 3h/8 - 1 \leq -3h/8 \quad (118)$$

so

$$\lim_{\ell \to \infty} 1 - (q^{m_\ell})^{M_{\mathcal{K}}(t) + h/8 + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}[M_{\mathcal{K}}^{(m_\ell)}(t)] - 1}} = 1 \quad (119)$$

which is also true $\forall t \in [t_1, t_2]$ uniformly. Therefore, given

$$\epsilon \in \left(0, \frac{h/4}{c_{\min}(1, \mathcal{K})(t_2 - t_1)}\right) \quad (120)$$

$\exists L_2 > L_1 > 0$, such that $\forall \ell > L_2$

$$\frac{\mathrm{d}E[M_{\mathcal{K}}^{(m_\ell)}(t)]}{\mathrm{d}t} > (1 - \epsilon) z_{1,\mathcal{K}} = (1 - \epsilon) c_{\min}(1, \mathcal{K}), \ \forall t \in [t_1, t_2]. \quad (121)$$

Therefore

$$\begin{aligned}
&E[M_{\mathcal{K}}^{(m_\ell)}(t_2)] \\
&\geq E[M_{\mathcal{K}}^{(m_\ell)}(t_1)] + \int_{t_1}^{t_2} (1 - \epsilon) c_{\min}(1, \mathcal{K}) \mathrm{d} \\
&> c_{\min}(1, \mathcal{K}) t_1 - 5h/8 + (1 - \epsilon) c_{\min}(1, \mathcal{K})(t_2 - t_1) \\
&= c_{\min}(1, \mathcal{K}) t_2 - 5h/8 - \epsilon c_{\min}(1, \mathcal{K})(t_2 - t_1) \\
&> c_{\min}(1, \mathcal{K}) t_2 - 7h/8
\end{aligned} \quad (122)$$

which is a contradiction to (114). This shows $M_{\mathcal{K}}(t) = c_{\min}(1, \mathcal{K})\ t$ for $t \in [0, 1/c_{\min}(1, \mathcal{K})]$ when $\mathcal{K} = \{1\}^c$. Suppose this statement is true $\forall \mathcal{K}' \supset \mathcal{K}, \mathcal{K}' \neq \mathcal{K}$. We claim that we still have $M_{\mathcal{K}}(t) = c_{\min}(1, \mathcal{K})t$. Otherwise, suppose at $t_2 \in (0, 1/c_{\min}(1, \mathcal{K})]$ we have

$$h = c_{\min}(1, \mathcal{K}) t_2 - M_{\mathcal{K}}(t_2) > 0 \quad (123)$$

then similarly, we can find $t_1 \in (0, t_2)$ such that

$$c_{\min}(1, \mathcal{K}) t_1 - M_{\mathcal{K}}(t_1) = h/2 \quad (124)$$

and

$$c_{\min}(1, \mathcal{K}) t - M_{\mathcal{K}}(t) \geq h/2, \quad \forall t \in [t_1, t_2]. \quad (125)$$

Reuse the definition of $\lambda_{\mathcal{K}}^\ell$ such that (109) and(110) still apply. Define $\lambda_{\{i\}\cup\mathcal{K}}^\ell$ in the similar way with respect to $\delta_{\{i\}\cup\mathcal{K}}$ and $M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)$. Because $E[M_{\mathcal{K}}^{(m_\ell)}(t)] \xrightarrow{u} M_{\mathcal{K}}(t)$ and $E[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)] \xrightarrow{u} M_{\{i\}\cup\mathcal{K}}(t)$ on $[0, 1/c_{\min}(1, \mathcal{K})]$, $\exists L_1$ such that $\forall \ell > L_1$ and $\forall t \in [t_1, t_2]$

$$\begin{aligned}
|E[M_{\mathcal{K}}^{(m_\ell)}(t)] - M_{\mathcal{K}}(t)| &< h/8 \\
E[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)] - M_{\{i\}\cup\mathcal{K}}(t)| &< h/8.
\end{aligned} \quad (126)$$

Let $p(\gamma_{\mathcal{K}}, \gamma_{\{i\}\cup\mathcal{K}})$ be the joint distribution function of $M_{\mathcal{K}}^{(m_\ell)}(t)$ and $M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)$. Let

$$\begin{aligned}
A \triangleq &\left[0, E[M_{\mathcal{K}}^{(m_\ell)}(t)] + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right]}\right] \\
&\times \left[E[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)] - \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]}, 1\right].
\end{aligned} \quad (127)$$

Then we may calculate as follows:

$$\begin{aligned}
&E\left[1 - (q^{m_\ell})^{M_{\mathcal{K}}^{(m_\ell)}(t) - M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)}\right] \\
&\geq \int_A p\left(\gamma_{\mathcal{K}}, \gamma_{\{i\}\cup\mathcal{K}}\right) \left(1 - (q^{m_\ell})^{\gamma_{\mathcal{K}} - \gamma_{\{i\}\cup\mathcal{K}}}\right).
\end{aligned} \quad (128)$$

But when $(\gamma_{\mathcal{K}}, \gamma_{\{i\}\cup\mathcal{K}}) \in A$, we have

$$\begin{aligned}
\gamma_{\mathcal{K}} - \gamma_{\{i\}\cup\mathcal{K}} \leq &E\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right] + \lambda_{\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\mathcal{K}}^{(m_\ell)}(t)\right]} \\
&- E\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right] + \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]}.
\end{aligned} \quad (129)$$

Using (125), (126) and the induction assumption $M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t) = c_{\min}(1, \{i\}\cup\mathcal{K})t$, we have

$$\begin{aligned}\gamma_\mathcal{K} - \gamma_{\{i\}\cup\mathcal{K}} &< M_\mathcal{K}(t) + h/8 + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}\left[M_\mathcal{K}^{(m_\ell)}(t)\right]} \\ &\quad - M_{\{i\}\cup\mathcal{K}}(t) + h/8 + \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]} \\ &\le c_{\min}(1,\mathcal{K})t - h/2 + h/8 + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}\left[M_\mathcal{K}^{(m_\ell)}(t)\right]} \\ &\quad - c_{\min}(1,\{i\}\cup\mathcal{K})t + h/8 + \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]} \\ &\le -h/4 + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}\left[M_\mathcal{K}^{(m_\ell)}(t)\right]} + \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]} \\ &\to -h/4, \quad (\ell\to\infty), \quad \text{uniformly } \forall t\in[t_1,t_2]\,. \end{aligned} \tag{130}$$

By Inclusion-Exclusion Principle, we have for any two events A and B

$$P(A\wedge B) \ge P(A) + P(B) - 1. \tag{131}$$

So

$$\begin{aligned}&P\Big(M_\mathcal{K}^{(m_\ell)}(t) < E\left[M_\mathcal{K}^{(m_\ell)}(t)\right] + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}\left[M_\mathcal{K}^{(m_\ell)}(t)\right]} \wedge \\ &M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t) > E\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right] - \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]}\Big) \\ &\ge P\left(M_\mathcal{K}^{(m_\ell)}(t) < E[M_\mathcal{K}^{(m_\ell)}(t)] + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}[M_\mathcal{K}^{(m_\ell)}(t)]}\right) \\ &+ P\big(M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t) > E[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)] \\ &- \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)]}\big) - 1 \\ &> \left(1 - \frac{1}{(\lambda_\mathcal{K}^\ell)^2}\right) + \left(1 - \frac{1}{(\lambda_{\{i\}\cup\mathcal{K}}^\ell)^2}\right) - 1 \\ &\quad \to 1, \quad (\ell\to\infty), \quad \text{uniformly } \forall t\in[t_1,t_2] \end{aligned} \tag{132}$$

by Chebyshev Inequality and (109).

Consequently

$$\begin{aligned}&E\left[1 - (q^{m_\ell})^{M_\mathcal{K}^{(m_\ell)}(t) - M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)}\right] \\ &> P\Big(M_\mathcal{K}^{(m_\ell)}(t) < E\left[M_\mathcal{K}^{(m_\ell)}(t)\right] + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}\left[M_\mathcal{K}^{(m_\ell)}(t)\right]} \wedge \\ &M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t) > E\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right] - \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]}\Big) \\ &\cdot (1 - (q^{m_\ell})^{-h/4 + \lambda_\mathcal{K}^\ell \sqrt{\mathrm{Var}\left[M_\mathcal{K}^{(m_\ell)}(t)\right]} + \lambda_{\{i\}\cup\mathcal{K}}^\ell \sqrt{\mathrm{Var}\left[M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)\right]}} \\ &\to 1, \quad (\ell\to\infty), \quad \text{uniformly } \forall t\in[t_1,t_2]. \end{aligned}$$

Therefore, given ϵ as defined in (120), $\exists L_2 > L_1$, such that $\forall \ell > L_2$

$$\begin{aligned}\frac{\mathrm{d}E[M_\mathcal{K}^{(m_\ell)}(t)]}{\mathrm{d}t} &= \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} E\left[1 - (q^{m_\ell})^{M_\mathcal{K}^{(m_\ell)}(t) - M_{\{i\}\cup\mathcal{K}}^{(m_\ell)}(t)}\right] \\ &\ge (1-\epsilon)\sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} \\ &\ge (1-\epsilon) c_{\min}(1,\mathcal{K}), \end{aligned} \tag{133}$$

from which we reach the same contradiction as in (122). This shows that $E[M_\mathcal{K}^{(m_\ell)}(t)] \xrightarrow{u} c_{\min}(1,\mathcal{K})t$ over $[0, 1/c_{\min}(1,\mathcal{K})]$. Now we claim that, in fact, $E[M_\mathcal{K}^{(m)}(t)] \to c_{\min}(1,\mathcal{K})t$ over $[0, 1/c_{\min}(1,\mathcal{K})]$. For otherwise $\exists t\in[0, 1/c_{\min}(1,\mathcal{K})]$, $\exists \epsilon > 0$ and $\exists \{m_\ell\}_\ell$ which is a subsequence of $m = 1, 2, \ldots$, such that $c_{\min}(1,\mathcal{K})t - E[M_\mathcal{K}^{(m_\ell)}(t)] > \epsilon$. Apply the previous set of induction arguments to $\{E[M_\mathcal{K}^{(m_\ell)}(t)]\}_\ell$, we know that a subsequence of $\{E[M_\mathcal{K}^{(m_\ell)}(t)]\}_\ell$ converges to $c_{\min}(1,\mathcal{K})t$ uniformly on $[0, 1/c_{\min}(1,\mathcal{K})]$, which is a contradiction.

Because $\forall t\in(0, 1/c_{\min}(1,\mathcal{K})]$, $E[M_\mathcal{K}^{(m)}(t)] \to c_{\min}(1,\mathcal{K})t$ and, from Property 5) of Proposition 1, $\mathrm{Var}[M_\mathcal{K}^{(m)}(t)] \to 0$, the conclusion follows from Chebyshev Inequality.

Since $M_\mathcal{K}(t) \to c_{\min}(1,\mathcal{K})t$, we know that except for the last few packets, the transmission happens at the rate arbitrarily close to $c_{\min}(1,\mathcal{K})$, i.e., we have

Corollary 5: Given $\alpha\in(0,1)$, we have

$$\lim_{m\to\infty} P\left(D(0,\alpha^2 m) \le \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right) = 1. \tag{134}$$

Proof:

$$\begin{aligned}&\lim_{m\to\infty} p\left(d_\mathcal{K}(0,\alpha^2 M) \le \frac{\alpha M}{C_{\min}(1,\mathcal{K})}\right) \\ &= \lim_{m\to\infty} p\left(n_\mathcal{K}\left(\frac{\alpha M}{C_{\min}(1,\mathcal{K})}\right) \ge \alpha^2 M\right) \\ &= \lim_{m\to\infty} p\left(\frac{1}{M} n_\mathcal{K}\left(\frac{\alpha m}{C_{\min}(1,\mathcal{K})}\right) \ge \alpha^2\right) \\ &= \lim_{m\to\infty} p(m_\mathcal{K}^{(m)}\left(\frac{\alpha}{C_{\min}(1,\mathcal{K})}\right) \\ &\qquad \ge \alpha C_{\min}(1,\mathcal{K}) \cdot \frac{\alpha}{C_{\min}(1,\mathcal{K})}\Big) \to 1 \end{aligned} \tag{135}$$

according to Theorem 5 with

$$T = \frac{\alpha}{C_{\min}(1,\mathcal{K})}. \tag{136}$$

∎

However, the last few packets do not really affect the ensemble transmission rate because we can put a bound on the time it takes to transmit them. This bound is proportional to the amount of "slow" packets at the end of transmission and can be made negligible compared to the bulk of the transmission. Specifically, we have

Proposition 2: Given q and $1 > \alpha > 0$, $\exists b_i$ for each node i, such that

$$\lim_{m\to\infty} P\left(D_i(\alpha m, m) \le b_i(1-\alpha)m\right) = 1. \tag{137}$$

Proof: First consider the special case when $z_{1,i} > 0$ and we ignore packets received at i from nodes other than 1. When $N_i(t) < m$, any incoming packet is innovative with probability of $1 - q^{N_i(t)-m} \ge 1 - q^{-1}$. So let

$$b_i = \frac{3}{z_{1,i}(1-q^{-1})} \tag{138}$$

and apply Chebyshev inequality, we have

$$P\left(D_i(\alpha m, m) \le b_i(1-\alpha)m\right) \to 1. \tag{139}$$

If there is a path, say (WLOG), $1, 2, \dots, i-1$, such that $z_{j,j+1} > 0$ ($j = 1, 2, \dots, i-1$) (This must be true since we assume G is connected), we focus on the transmission along this path and ignore other transmissions. We use $D'_j(m_1, m_2)$ to denote the time $N_j(t)$ takes to increase from m_1 to m_2, assuming during that time $N_{j-1}(t) = m$. Let

$$b_i = \sum_{j=1}^{i-1} \frac{3}{z_{j,j+1}(1-q^{-1})} \tag{140}$$

then

$$\begin{aligned} &P(D_i(\alpha m, m) \le b_i(1-\alpha)m) \\ &\ge P\left(\sum_{j=1}^{i-1} D'_{j+1}(\alpha m, m) \le b_i(1-\alpha)m\right) \\ &\ge P\left(\bigwedge_{j=1}^{i-1} D'_{j+1}(\alpha m, m) \le \frac{3(1-\alpha)m}{z_{j,j+1}(1-q^{-1})}\right) \\ &\to 1 \end{aligned} \tag{141}$$

where the first inequality follows from the fact that for $j = 2, 3, \dots, i$, $N_{j-1}(t)$ reaches m before $N_j(t)$ reaches m; and the limit follows from the inclusion-exclusion principle and (139).

The following corollary trivially extends the conclusion to an arbitrary set $\mathcal{K}$.

Corollary 6: For any nonempty set $\mathcal{K} \subset \mathcal{N}$, $\exists b_{\mathcal{K}} > 0$ such that

$$\lim_{m\to\infty} P(D_{\mathcal{K}}(\alpha m, m) \le b_{\mathcal{K}}(1-\alpha)m) = 1. \tag{142}$$

Proof: Assume node $i \in \mathcal{K}$ and apply Proposition 2.

Now we are in a position to finish the achievability proof by combining Corollary 5 and Corollary 6.

Theorem 6: $\forall \mathcal{K} \subset \mathcal{N}$ and $\forall 0 < \epsilon < 1$

$$\lim_{m\to\infty} P\left(\frac{m}{D_{\mathcal{K}}(0,m)} > \epsilon c_{\min}(1,\mathcal{K})\right) = 1. \tag{143}$$

Proof: We note $\forall \alpha \in (0,1)$

$$\begin{aligned} &P\left(\frac{m}{D_{\mathcal{K}}(0,m)} > \epsilon c_{min}(1,\mathcal{K})\right) \\ &= P\left(D_{\mathcal{K}}(0,m) < \frac{m}{\epsilon c_{\min}(1,\mathcal{K})}\right) \\ &\ge P\left(D(0,\alpha^2 m) \le \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right. \\ &\qquad \left.\wedge D(\alpha^2 m, m) \le \frac{m}{\epsilon c_{\min}(1,\mathcal{K})} - \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right) \\ &\ge P\left(D(0,\alpha^2 m) \le \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right) \\ &\quad + P\left(D(\alpha^2 m, m) \le \frac{m}{\epsilon c_{\min}(1,\mathcal{K})} - \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right) - 1. \end{aligned} \tag{144}$$

Pick $\alpha \in (0,1)$ such that

$$\alpha > 1 - \frac{1/\epsilon - 1}{2 b_{\mathcal{K}} c_{\min}(1,\mathcal{K})} \tag{145}$$

then we have

$$\begin{aligned} \frac{1}{\epsilon c_{\min}(1,\mathcal{K})} &> \frac{1}{c_{\min}(1,\mathcal{K})} + b_{\mathcal{K}}(2-2\alpha) \\ &> \frac{\alpha}{c_{\min}(1,\mathcal{K})} + b_{\mathcal{K}}(2-2\alpha) \\ &\ge \frac{\alpha}{c_{\min}(1,\mathcal{K})} + b_{\mathcal{K}}(1-\alpha^2). \end{aligned} \tag{146}$$

By Corollary 6, we know

$$\lim_{m\to\infty} P\left(D(\alpha^2 m, m) \le \frac{m}{\epsilon c_{\min}(1,\mathcal{K})} - \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right) = 1. \tag{147}$$

By Corollary 5

$$\lim_{m\to\infty} P\left(D(0,\alpha^2 m) \le \frac{\alpha m}{c_{\min}(1,\mathcal{K})}\right) = 1. \tag{148}$$

Pluging (147) and(148) back in(144) and taking the limit $m \to \infty$, we get (143). ■

Theorem 6 is a little more general than the statements made in [9] and [18] since it reveals that, not only the rank at a single node, but also the rank at any subset $\mathcal{K} \subset \mathcal{N}$ increases at the rate determined by the min-cut bound $c_{\min}(1,\mathcal{K})$. It should also be pointed out that in the proof to Theorem 6, typical difficulties with cycles in the network topology do not arise due to the layered structure of the DI's that has encoded all topological information.

VI. Numerical Results of the DEDI Framework

In this section we present extensive numerical examples of the DEDI framework for RNC. We use $q = 2$ for all the RNC examples to show that, even for the small field size, the DEDI framework provides desirable accuracy. We give the simulation results for different network topologies that are described using the hypergraph model introduced in Section II (see Fig. 1). We remind the readers that the dots represent the nodes whose transmissions are according to independent Poisson processes; the arrows represent the reachability. We set the intensity of the Poisson processes uniformly to 1 packet/second. If multiple arrows emanate from the same node, it means when this node transmits, all the nodes on the other end of the arrows have a chance to receive this packet. In our simulation, unless indicated otherwise, the receptions are independent. Their independent reception probabilities are shown as the numbers attached to the arrows. We use a discrete event simulation model in which the transmission time instants at each node are determined by the associated independent Poisson process; the receptions are simulated with the assigned reception probabilities and the ranks are updated in real time. Because the convergence of the fluid approximation to $E[N_{\mathcal{K}}(t)]$ is extremely fast, we choose to show sample paths rather than the ensemble average to demonstrate the accuracy and versatility of the DEDI framework. We will demonstrate the ability of the DEDI framework to handle multiple sessions, complex network topology and correlated receptions by comparing the rank evolution processes at different nodes obtained from simulation with the DEDI solution. We assume all transmissions begin from $t = 0$.

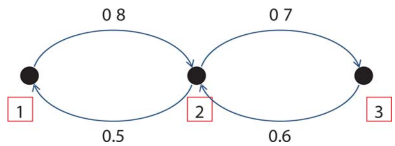

Fig. 9. Three node wireless network.

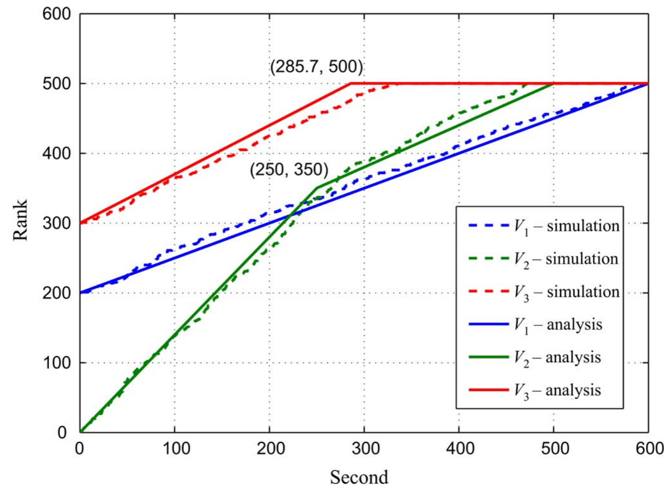

Fig. 10. Two multicast sessions with two sources using RNC.

A. Two Multicast Sessions

Consider a three node wireless network shown in Fig. 9. Assume the information flow is comprised of two multicast sessions originating from node 1 and node 3, respectively. Node 1 has 200 packets to deliver to node 2 and 3, while node 3 has 300 packets to deliver to node 1 and 2. We may write out the DI's that describe this scenario

$$\begin{aligned}\dot{V}_1 &= z_{2,1}\mathrm{Sgn}^+(V_{\{12\}} - V_1)\\ \dot{V}_2 &= z_{1,2}\mathrm{Sgn}^+(V_{\{12\}} - V_2) + z_{3,2}\mathrm{Sgn}^+(V_{\{23\}} - V_2)\\ \dot{V}_3 &= z_{2,3}\mathrm{Sgn}^+(V_{\{23\}} - V_3)\\ \dot{V}_{\{12\}} &= z_{2,3}\mathrm{Sgn}^+(V_{\{123\}} - V_{\{12\}})\\ \dot{V}_{\{23\}} &= z_{1,2}\mathrm{Sgn}^+(V_{\{123\}} - V_{\{23\}})\\ \dot{V}_{\{123\}} &= 0 \end{aligned} \tag{149}$$

with the B.C.

$$V_{\mathcal{K}}(0) = \begin{cases} 300, & 1 \in \mathcal{K},\ 3 \notin \mathcal{K}\\ 200, & 3 \in \mathcal{K},\ 1 \notin \mathcal{K}\\ 500, & \{1,3\} \subset \mathcal{K}\\ 0, & \text{o.w.} \end{cases} \tag{150}$$

Fig. 10 shows the analytical solution to (149) as well as the simulation results. The analysis matches the simulations closely. Clearly the rank increase at node 1 should be subject to its min cut bound $c_{\min}(3,1) = 0.5$ packet/second and node 3 subject to $c_{\min}(1,3) = 0.7$ packet/second. Consequently, $T_1^{\text{total}} = 300/0.5 = 600$ (seconds) and $T_3^{\text{total}} = 200/.7 = 285.7$ (seconds). For node 2, the flow from node 1 cannot exceed $c_{\min}(1,2)$; the flow from node 3 cannot exceed $c_{\min}(3,2)$; and the flow from the ensemble of node 1, 3 cannot exceed $c_{\min}(\{1,3\},2)$. Therefore

$$\begin{aligned} T_2^{\text{total}} = \max\{&m_1/c_{\min}(1,2), m_2/c_{\min}(3,2)\\ &(m_1+m_2)/c_{\min}(\{1,3\},2)\} = 500 \text{ (seconds)}.\end{aligned}$$

These calculations are readily verified in Fig. 10.

B. A Complex Topology

This example is intended to illustrate that the DEDI framework is capable of handling complex networks. The wireless network in Fig. 11 has 10 nodes and a fairly intricate connectivity. While the unidirectional arrows have the same meaning as in an arrow-dot representation of the hypergraph, the bidirectional arrows simply represent two unidirectional arrows whose reception probabilities are equal and as labeled. In this example we again assume independent reception at each node and the transmission rate $\lambda_i = 1$ packet/second, $i = 1,2,\ldots,10$. Node 1 is the only source node that has 100 packets to deliver.

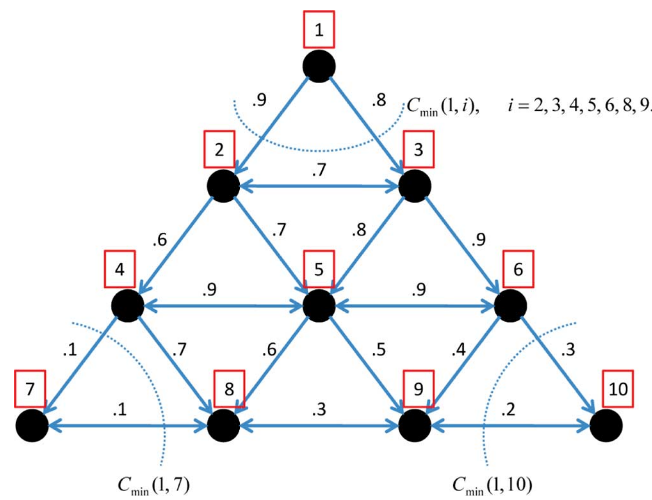

Fig. 11. A 10-node wireless network with node 1 being the unique source.

Fig. 12 shows the rank evolution at node 2, 4, 7 and 10. For node 7, the min cut is shown to be $0.1+0.1 = 0.2$ (packet/second). For node 10, it is shown to be $0.2+0.3 = 0.5$ (packet/second). For the other nodes, the min cut is $1-(1-0.9)\times(1-0.8) = 0.98$ (packet/second). These facts are reflected by the slope of the rank evolution curves on Fig. 12 where the simulated curves match the analytical solutions well.

C. Correlated Reception

The DEDI framework allows the analysis of rank evolution when receptions are not independent. The lack of independence could be due to correlated channels or joint reception by design and they are not uncommon in wireless communications. The ability to analyze the case of correlated reception is an advantage of the DEDI framework. Consider the four-node network shown in Fig. 13 where as usual the point-to-point reception probabilities are shown as labeled. Node 1 is the only source node that has 400 packets to deliver. However, we assume the receptions at node 2 and node 3 are not independent, i.e.,

$$P_{1,2} = 0.49, \quad P_{1,3} = 0.49, \quad P_{1,\{2,3\}} = 0.5. \tag{151}$$

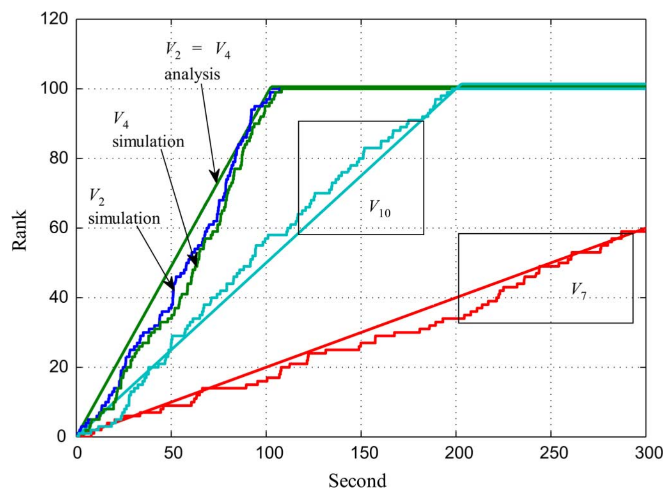

Fig. 12. Rank evolution for the network shown in Fig. 11.

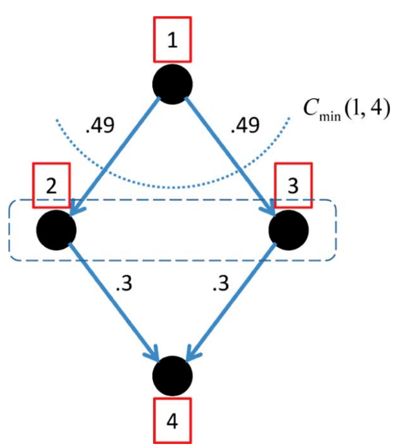

Fig. 13. Modeling correlated reception with DEDI in a four node wireless network.

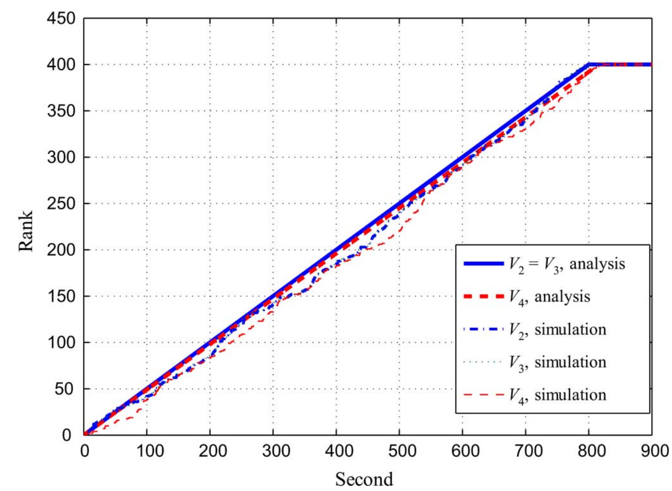

Fig. 14. Rank evolution if node 2 and 3 have correlated reception.

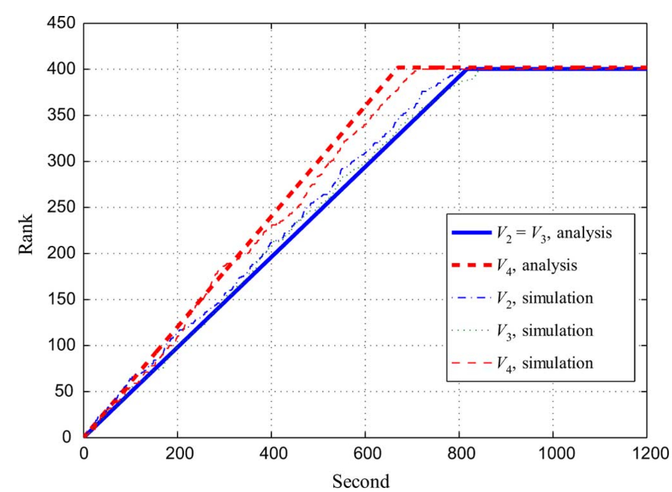

Fig. 15. Rank evolution if node 2 and 3 have independent reception.

This could happen when, for example, node 2 and 3 are in co-operation or the channels from node 1 are correlated. The rank evolution can still be accurately predicted by the DEDI framework as shown in Fig. 14. In this case, V_2 and V_3 increase at the same rate of $c_{\min}(1,2) = c_{\min}(1,3) = 0.49$ (packet/second) while V_4 increases at

$$c_{\min}(1,4) = z_{1,\{2,3\}} = \lambda_1 P_{1,\{2,3\}} = 0.5 \text{ (packet/second)}.$$

As a contrast, the results for independent receptions are shown in Fig. 15, where V_2 and V_3 increase at the same rate of $c_{\min}(1,2) = c_{\min}(1,3) = 0.49$ (packet/second) while V_4 increases at

$$\begin{aligned} c_{\min}(1,4) &= z_{1,\{2,3\}} = \lambda_1 P_{1,\{2,3\}} \\ &= \lambda_1(1-(1-P_{1,2})(1-P_{1,3})) = 0.74 \text{ (packet/second)}. \end{aligned}$$

VII. Discussion

A. The Dynamical System Point of View for Cross-Layer Design

The DEDI framework naturally presents a dynamical system point of view of RNC. The equations require two parameters to be specified, namely, the node transmission rate λ_i which is largely determined by MAC, and reception probability $P_{i,\mathcal{K}}$ which is largely determined by PHY. In a practical RNC application, both parameters are possibly subject to adaptive control depending on N_i or $N_{\mathcal{K}}$. This dependence could be characterized by functions or functionals synthesized via cross-layer optimization. In this case, assuming the fluid model is applicable so that $V_{\mathcal{K}}(t) \approx N_{\mathcal{K}}(t)$, we have the following dynamical system:

$$\begin{aligned} \dot{V}_{\mathcal{K}} &= \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}}(V_{\mathcal{K}\cup\{i\}} \ominus V_{\mathcal{K}}), \quad \forall\mathcal{K} \\ \lambda_i &= \lambda_i(\mathbf{V}), \quad \forall i \\ P_{i,\mathcal{K}} &= P_{i,\mathcal{K}}(\mathbf{V}), \quad \forall(i,\mathcal{K}) \end{aligned} \tag{152}$$

where $\mathbf{V}$ is the collection of all variables of V_i and $V_{\mathcal{K}}$. This dynamical system can be numerically evaluated with a DE solver, or optimized using variational methods. Therefore, the DEDI framework presents a possible tool for cross-layer design studies.

B. Extension to Random Coupon Selection

Random Coupon Selection (RC) [11] is another transmission scheme based on randomized operations. Instead of linear combination, the outgoing packet is randomly selected from the reservoir. With RC, an innovative packet is simply a distinct packet that has never been received. Let $N_i(t)$ denote the number of distinct packets node i has received at t. Let $N_{\mathcal{K}}(t) = \cup_{i\in\mathcal{K}} N_i(t)$, i.e., the number of distinct packets set $\mathcal{K}$ has received at t. Then N_i and $N_{\mathcal{K}}$ describe the propagation of innovative packets through the network with RC. Similar to Lemma 1, we have

Lemma 5: $\forall \mathcal{K} \subset \mathcal{N}$

$$\mathrm{d}E[N_{\mathcal{K}}(t)]/\mathrm{d}t = \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} E\left[\frac{N_{\mathcal{K}\cup\{i\}}(t) - N_{\mathcal{K}}(t)}{N_i(t)}\right]. \quad (153)$$

Proof: The proof is almost verbatim as the proof to Lemma 1, except that the probability a packet from i to $\mathcal{K}$ being innovative is given as

$$E\left[\frac{N_{\mathcal{K}\cup\{i\}}(t) - N_{\mathcal{K}}(t)}{N_i(t)}\right].$$

■

Assuming that the fluid approximation still applies, let $W_i(t) = E[N_i(t)]$ (resp. $W_{\mathcal{K}}(t) = E[N_{\mathcal{K}}(t)]$), we have the following system of differential equations that describes the approximate average behavior of RC

$$\dot{W}_{\mathcal{K}}(t) = \sum_{i\in\mathcal{K}^c} z_{i,\mathcal{K}} \frac{W_{\mathcal{K}\cup\{i\}}(t) - W_{\mathcal{K}}(t)}{W_i(t)}, \quad \forall \mathcal{K} \subset \mathcal{N}. \quad (154)$$

This system of DE's can be used in many ways to study transmission schemes based on RC, in a lossy wireless network with an arbitrary topology. We do not have a convergence proof for it as we do for RNC. However, the convergence can be quickly checked via simulation. We show in Fig. 16 the innovative packets propagation process, as well as the solution to (154), based on the network in Fig. 5 using RC, with varying m. It is evident from these simulations that the fluid approximation based DEDI framework can be applied to RC and (154) models its behavior well.

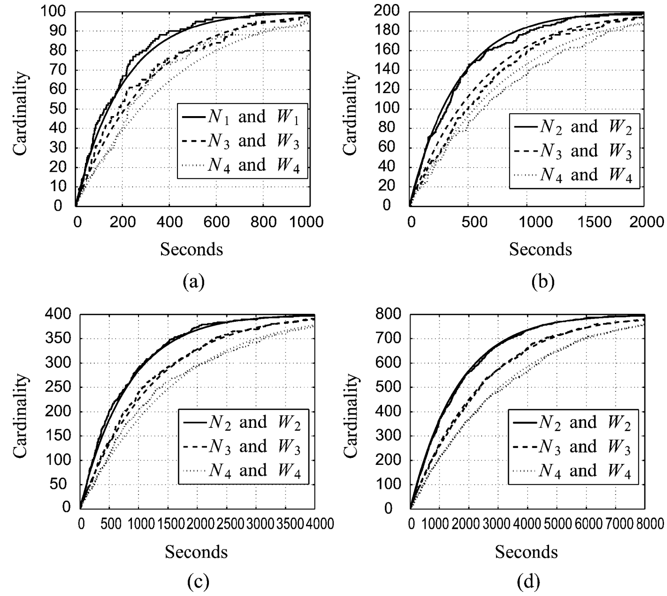

Fig. 16. Convergence behavior of (154) shown through simulation (a) $m = 100$ (b) $m = 200$ (c) $m = 400$ (d) $m = 800$.

VIII. Concluding Remarks

We presented the DEDI framework, based on differential equations and differential inclusions, for analyzing the rank evolution of RNC in an arbitrary wireless network. We showed that by adopting the fluid approximation, we can derive a system of deterministic DI's, the solution of which shows that the average throughput is given by the min-cut bound. We next showed that the rank evolution processes in fact concentrate to the DI solution in probability, by analyzing the exact system of DE's that characterizes the means and variances of the rank evolution processes, which in turn showed that the min-cut bound can be asymptotically achieved. We demonstrated the versatility of the DEDI framework by using it to analyze different networking scenarios including multiple multicast sessions, complex topology and correlated reception. We also discussed its advantages in practice as well as its possible extension to Random Coupon Selection.

Acknowledgment

The authors would like to thank Professor Daniel Ocone from the Mathematics Department at Rutgers University for his many valuable suggestions, especially regarding Lemma 3.

References

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
[2] D. D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," presented at the ACM MobiCom, San Diego, CA, Sep. 2003.
[3] A. Jiang, "Network coding for joint storage and transmission with minimum cost," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 1359–1363.
[4] A. G. Dimakis, P. B. Godfrey, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," in *Proc. Infocom*, Anchorage, AK, May 2007, pp. 2000–2008.
[5] C. Gkantsidis and P. Rodriguez, "Network coding for large scale content distribution," presented at the IEEE Conf. Computer Communications (INFOCOM), Miami, FL, Mar. 2005.
[6] C. Gkantsidis, J. Miller, and P. Rodriguez, "Comprehensive View of a Live Network Coding P2P System," in *Proc. 6th ACM SIGCOMM on Internet Measurement*, 2006, pp. 177–188.
[7] N. Cai and R. W. Yeung, "Secure network coding," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun./Jul. 2002, p. 323.
[8] T. Ho, R. Koetter, M. Médard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun./Jul. 2003, p. 442.

[9] D. S. Lun, M. Médard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Phys. Commun.*, vol. 1, no. 1, pp. 3–20, Mar. 2008.
[10] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
[11] C. Fragouli and E. Soljanin, "Network coding applications," *Found. Trends Netw.*, vol. 2, no. 2, pp. 135–269, 2007.
[12] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," presented at the 41st Annu. Allerton Conf. Communication, Control, and Computing, 2003.
[13] D. Zhang, N. Mandayam, and S. Parekh, "DEDI: A framework for analyzing rank evolution in random network coding," presented at the IEEE Int. Symp. Information Theory, Austin, TX, Jun. 2010.
[14] J. P. Aubin and A. Cellina*, Differential Inclusions: Set-Valued Maps and Viability Theory*. New York: Springer-Verlag, 1984.
[15] A. F. Filippov, "Differential Equations with Discontinuous Righthand Sides," in *Control Systems (Mathematics and its Applications)*, F. M. Arscott, Ed., 1st ed. New York: Springer, 1988, vol. 18.
[16] D. Zhang and N. Mandayam, "Analyzing multiple flows in a wireless network with differential equations and differential inclusions," presented at the IEEE Wireless Network Coding Workshop (WiNC), Boston, MA, Jun. 2010.
[17] H. Thorisson, "Coupling, Stationarity, and Regeneration," in *Probability and Its Applications*, 1st ed. New York: Springer, 2000, pp. 2–2.
[18] D. S. Lun, M. Médard, R. Koetter, and M. Effros, "Further results on coding for reliable communication over packet networks," in *Proc. Int. Symp. Information Theory (ISIT) IEEE*, Sep. 2005, pp. 1848–1852.

Dan Zhang received the B.S. (Hons.) degree in 2000 and M.S. degree in 2005 from Shanghai Jiaotong University, all in electrical engineering. In 2005, he joined the Ph.D. program of the ECE department at Rutgers University. Since 2007, he has been a graduate assistant of Prof. N. Mandayam at the Wireless Information Network Laboratory (WINLAB). His research interests include information theory, communication signal processing and wireless networking. He received his M.S. degree in Mathematics from Rutgers in 2011. Before he came to the USA, he worked for LSI Design & Integration Corp. as a digital engineer, and for Micronas as a system engineer.

Narayan B. Mandayam (F'09) is currently the Peter D. Cherasia Faculty Scholar at Rutgers University. He received the B.Tech (Hons.) degree in 1989 from the Indian Institute of Technology, Kharagpur, and the M.S. and Ph.D. degrees in 1991 and 1994 from Rice University, all in electrical engineering. From 1994 to 1996, he was a Research Associate at the Wireless Information Network Laboratory (WINLAB), Rutgers University before joining the faculty of the Electrical and Computer Engineering department at Rutgers where he became Associate Professor in 2001 and Professor in 2003. Currently, he also serves as Associate Director at WINLAB. He was a visiting faculty fellow in the Department of Electrical Engineering, Princeton University, in 2002 and a visiting faculty at the Indian Institute of Science in 2003. His research interests are in various aspects of wireless data transmission including system modeling and performance, signal processing and radio resource management with emphasis on techniques for cognitive radio networks.

Dr. Mandayam is a recipient of the Fred W. Ellersick Prize from the IEEE Communications Society in 2009 along with O. Ileri for their work on dynamic spectrum access models and spectrum policy. He is also a recipient of the Institute Silver Medal from the Indian Institute of Technology in 1989 and the National Science Foundation CAREER Award in 1998. He is a coauthor with C. Comaniciu and H. V. Poor of the book *Wireless Networks*: *Multiuser Detection in Cross-Layer Design* (New York: Springer). He has served as an Editor for the IEEE COMMUNICATION LETTERS and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has also served as a guest editor of the IEEE JSAC Special Issues on Adaptive, Spectrum Agile and Cognitive Radio Networks (2007) and Game Theory in Communication Systems (2008).