

Throughput-Optimal Multi-hop Broadcast Algorithms

Abhishek Sinha
Laboratory for Information and
Decision Systems
MIT
sinhaa@mit.edu

Georgios Paschos
Mathematical and Algorithmic
Sciences Lab Huawei
Technologies Co. Ltd
georgios.paschos@huawei.com

Eytan Modiano
Laboratory for Information and
Decision Systems
MIT
modiano@mit.edu

ABSTRACT

In this paper we design throughput-optimal dynamic broadcast algorithms for multi-hop networks with arbitrary topologies. Most of the previous broadcast algorithms route packets along spanning trees, rooted at the source node. For large dynamic networks, computing and maintaining a set of spanning trees is not efficient, as the network-topology may change frequently. In this paper we design a class of dynamic algorithms which makes packet-by-packet scheduling and routing decisions and thus obviates the need for maintaining any global topological structures, such as spanning trees. Our algorithms may be conveniently understood as a non-trivial generalization of the familiar *back-pressure* algorithm which makes unicast packet routing and scheduling decisions, based on queue-length information, without maintaining end-to-end paths. However, in the broadcast problem, it is hard to define queuing structures due to absence of a work-conservation principle which results from packet duplications. We design and prove the optimality of a virtual-queue based algorithm, where a virtual-queue is defined for subsets of vertices. We then propose a multi-class broadcast policy which combines the above scheduling algorithm with a class-based *in-order* packet delivery constraint, resulting in significant reduction in complexity. Finally, we evaluate performance of the proposed algorithms via extensive numerical simulations.

1. INTRODUCTION

Packet broadcasting is used for efficiently disseminating messages to all recipients in a network. Its efficiency is measured in terms of *broadcast throughput*, i.e., the common rate of packet-reception by all nodes. Technically, the broadcast problem refers to finding a policy for duplicating and forwarding copies of packets such that the maximum broadcast throughput (also known as *broadcast-capacity*) is achieved. Solving the broadcast problem is challenging, especially for mobile wireless networks with time-varying connectivity and interference constraints. In this paper we focus on design-

ing dynamic broadcast algorithms. Such algorithms operate without the knowledge of network-topology or future arrivals, and hence, are robust. In this context, we derive provably throughput-optimal dynamic broadcast algorithms for networks with arbitrary topology.

Most of the existing broadcast algorithms are static by nature and operate by forwarding copies of packets along spanning trees [7]. In a network with time-varying topology, these static algorithms need to re-compute the trees every time the underlying topology changes, which could be quite cumbersome and inefficient. Recent works [8] and [13] consider the problem of throughput-optimal broadcasting in Directed Acyclic Graphs (DAG). Here the authors propose dynamic policies by exploiting the properties of DAG. However, it is not clear how to extend their algorithms to networks with arbitrary (non-DAG) topology. The authors in [4] propose a randomized packet-forwarding policy for wireless networks, which is shown to be throughput-optimal under some assumptions. However, their algorithm potentially needs to use unbounded amount of memory and can not be easily generalized to wireless networks with activation constraints. A straight-forward extension of their algorithm, proposed in [10], uses activation oracle, which is not practically feasible.

In this paper we study the broadcasting problem in arbitrary networks, including wireless. We propose algorithms that do not require the construction of global topological structures, like spanning trees. Leveraging the work in [8], we propose a novel multi-class heuristic, which simplifies the operational complexity of the proposed algorithm. Our main technical contributions in this paper are as follows: (1) We first identify a state-space representation of the network-dynamics, in which the broadcast-problem reduces to a “virtual-queue” stability problem. By utilizing techniques from Lyapunov-drift methodology, we derive a throughput-optimal broadcast policy. (2) Next, we introduce a multi-class heuristic policy, by combining the above scheduling rule with *in-class in-order* packet delivery, where the number of classes is a tunable parameter, which may be used as a trade-off between efficiency and complexity. (3) Finally, we validate the theoretical ideas through extensive numerical simulations. (4) An equivalent *mini-slot* model is proposed, which simplifies the analysis and may be of independent theoretical interest.

The rest of the paper is organized as follows. In section 2 we describe the operational network model and characterize its broadcast-capacity. In section 3 we derive our throughput-optimal broadcast policy. In section 4 we propose a multi-

arXiv:1604.00446v1 [cs.DS] 2 Apr 2016

class heuristic policy which uses the scheduling scheme from section 3. In section 6 we validate our theoretical results via extensive numerical simulations. Finally in section 7 we conclude the paper with some directions for future work.

2. SYSTEM MODEL

For simplicity, we first consider the problem in a wireline setting. The wireless model will be considered in section 5.

2.1 Network Model

Consider a graph $\mathcal{G}(V, E)$, V being the set of vertices and E being the set of edges, with $|V| = n$ and $|E| = m$. Time is slotted and the edges are directed. Transmission capacity of each edge is one packet per slot. External packets arrive at the source node $\mathbf{r} \in V$. The arrivals are i.i.d. at every slot with expected arrival of λ packets per slot.

For sake of convenience, we alter the slotted-time assumption and adopt a slightly different but equivalent *mini-slot* model. A slot consists of m consecutive mini-slots. As will be evident from what follows, our dynamic broadcast algorithms are conceptually easier to derive, analyze and understand in the mini-slot model. However, the algorithms can be easily applied in the more traditional slotted model.

Mini-slot model: In this model, the basic unit of time is called a *mini-slot*. At each mini-slot t , an edge $e = (a, b) \in E$ is chosen for activation, independently and uniformly at random from the set of all m edges. All other $m-1$ edges remain idle for that mini-slot. A packet can be transmitted over an active edge only. A single packet transmission takes one mini-slot for completion. This random edge-activity process is represented by the i.i.d. sequence of random variables $\{S(t)\}_{t=1}^{\infty}$, such that, if an edge $e \in E$ is chosen for activation at the mini-slot t , we have $S(t) = e$. Thus,

$$\mathbb{P}(S(t) = e) = 1/m, \quad \forall e \in E, \quad \forall t$$

External packets arrive at the source \mathbf{r} with expected arrival of λ/m packets per mini-slot.

The main operational advantage of the mini-slot model is that only a single packet transmission takes place at a mini-slot, which makes it easier to express the system-dynamics. However, as we show in Lemma (1), these two models are equivalent from the point-of-view of broadcast-capacity.

2.2 Broadcast-Capacity of a Network

Informally, a network supports a broadcast-rate of λ if external packets arrive at the source at the rate of λ and there exists a scheduling policy under which all nodes receive distinct packets at the rate of λ . The broadcast-capacity λ^* is the maximally achievable broadcast-rate in the network.

Formally, we consider a class Π of scheduling policies which executes the following two actions at every mini-slot t

- The policy observes the currently active edge $e = (a, b)$.
- The policy transmits (at most) one packet from node a to node b over the active edge e .

The policy-class Π includes policies that have access to all past and future information, and may forward any packet present at node a at time t to node b .

Recall that, a slot consists of m consecutive mini-slots. Let $R_i^\pi(t)$ be the number of distinct packets received by node $i \in V$ up to slot t , under a policy $\pi \in \Pi$. The time average $\liminf_{T \rightarrow \infty} R_i^\pi(T)/T$ is the rate at which distinct packets are received at node i , under the action of the policy π .

DEFINITION 1 (BROADCAST POLICY). A policy π is called a “broadcast policy of rate λ ” if all nodes in the network receive distinct packets at the rate of λ packets per slot, i.e.,

$$\min_{i \in V} \liminf_{T \rightarrow \infty} \frac{1}{T} R_i^\pi(T) = \lambda, \quad \text{in probability,} \quad (1)$$

when external packets arrive at the source node \mathbf{r} at rate λ .

DEFINITION 2. The broadcast capacity λ^* of a network is the supremum of all arrival rates λ for which there exists a broadcast policy $\pi \in \Pi$ of rate λ .

In the slotted-time model, the broadcast capacity λ^* of a network \mathcal{G} follows from the Edmonds’ tree-packing theorem [6], and is given by the following:

$$\lambda^* = \min_{\mathbf{t} \in V \setminus \{\mathbf{r}\}} \text{Max-Flow}(\mathbf{r} \rightarrow \mathbf{t}) \quad \text{per slot,} \quad (2)$$

where $\text{Max-Flow}(\mathbf{r} \rightarrow \mathbf{t})$ denotes the maximum value of flow that can be feasibly sent from the node \mathbf{r} to the node \mathbf{t} in the graph $\mathcal{G}(V, E)$ [1]. Edmonds’ theorem also implies that there exist λ^* edge-disjoint arborescences¹ or directed spanning trees, rooted at \mathbf{r} in the graph. By examining the flow from the source to every node and using (2), it follows that by sending unit flow over each edge-disjoint tree, we may achieve the capacity λ^* .

As an illustration, consider the graph shown in Figure 1. It follows from Eqn. (2) that the broadcast capacity of the graph is $\lambda^* = 2$. Edges belonging to a set of two edge-disjoint spanning trees \mathcal{T}_1 and \mathcal{T}_2 are shown in blue and red in the figure.

The following lemma establishes the equivalence of the *mini-slot* model and the *slotted-time* model in terms of broadcast-capacity.

LEMMA 1 (Invariance of Capacity). The broadcast capacity λ^* is the same for both the *mini-slot* and the *slotted-time* model and is given by Eqn. (2).

PROOF. See Appendix (9.2) \square

3. A THROUGHPUT-OPTIMAL BROADCAST POLICY π^*

In this section we design a throughput-optimal broadcast algorithm $\pi^* \in \Pi$, for networks with arbitrary topology. This algorithm is of *Max-weight* type and is reminiscent of the famous back-pressure policy for the corresponding unicast problem [9]. However, because of packet duplications, the usual per-node queues cannot be defined, unlike the unicast case. We get around this difficulty by defining certain virtual-queues, corresponding to subsets of nodes. We show that a scheduling policy in Π^* , that *stochastically stabilizes* these virtual queues for all arrival rates $\lambda < \lambda^*$, constitutes a throughput-optimal broadcast policy. Based on this result, we derive a Max-Weight policy π^* , by minimizing the drift of a quadratic Lyapunov function of the virtual queues.

¹An arborescence is a directed graph such that there is a unique directed path from the root \mathbf{r} to all other vertices in it. Thus, an arborescence is a directed form of a rooted tree. From now onwards, the terms arborescence and directed spanning tree (or simply, spanning tree) will be used interchangeably.

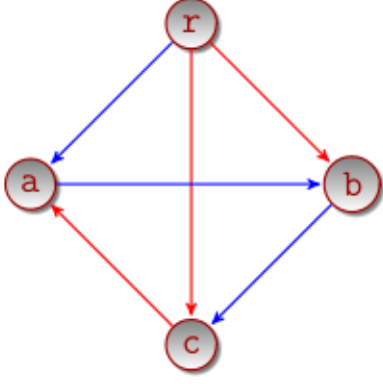


Figure 1: The four-node diamond network \mathcal{D}_4

3.1 Definitions and Notations

To describe our proposed algorithm, we first introduce the notion of *reachable sets* and *reachable sequence of sets* as follows.

DEFINITION 3 (Reachable Set). A subset of vertices $F \subset V$ is said to be reachable if the induced graph $F(\mathcal{G})$ contains a directed arborescence, rooted at source \mathbf{r} , which spans the node set F .

In other words, a subset of vertices $F \subset V$ is reachable if and only if there is a broadcast policy such that, a packet p may be duplicated exactly at the subset F in its course of broadcast. Note that the set of all reachable sets may be strict subset of the set of all subsets of vertices. This is true because all reachable sets, by definition, must contain the source node \mathbf{r} .

In fact, we may completely describe the trajectory of a packet during its course of broadcast, using the notion of *Reachable Sequences*, defined as follows:

DEFINITION 4 (Reachable Sequence). An ordered sequence of $n - 1$ (reachable set, edge) tuples $\{(F_j, e_j), j = 1, 2, \dots, n - 1\}$ is called a Reachable Sequence if the following properties hold:

- $F_1 = \{\mathbf{r}\}$ and for all $j = 1, 2, \dots, n - 1$:
- $F_j \subset F_{j+1}$
- $|F_{j+1}| = |F_j| + 1$.
- $e_j = (a, b) \in E : a \in F_j, b \in F_{j+1} \setminus F_j$

\mathcal{F} is defined to be the set of all reachable sequences.

A reachable sequence denotes a valid sequence of transmissions for broadcasting a particular packet to all nodes, where the j^{th} transmission of a packet takes place across the edge $e_j, j = 1, 2, \dots, n - 1$. By definition, every reachable set must belong to at least one reachable sequence. A trivial upper-bound on $|\mathcal{F}|$ is n^{2n} . An example illustrating the notions of reachable sets and reachable sequences for a simple graph is provided below.

²For a graph $\mathcal{G}(V, E)$ and a vertex set $F \subset V$, the induced graph $F(\mathcal{G})$ is defined as the sub-graph containing only the vertices F with the edges whose both ends lie in the set F .

Example: Consider the graph shown in Figure 1. A reachable sequence for this graph is given by \mathcal{S} below:

$$\mathcal{S} = \{(\{\mathbf{r}\}, \mathbf{ra}), (\{\mathbf{r}, \mathbf{a}\}, \mathbf{ab}), (\{\mathbf{r}, \mathbf{a}, \mathbf{b}\}, \mathbf{bc})\}$$

This reachable sequence is obtained by adding nodes along the tree with blue edges in Figure 1. Similarly, an example of a reachable set F in this graph is

$$F = \{\mathbf{r}, \mathbf{a}, \mathbf{b}\}$$

For a reachable set F , define its out-edges $\partial^+ F$ and in-edges $\partial^- F$ as follows:

$$\partial^+ F = \{(a, b) \in E : a \in F, b \notin F\} \quad (3)$$

$$\partial^- F = \{(a, b) \in E : a \in F, b \in F\} \quad (4)$$

For an edge $e = (a, b) \in \partial^+ F$, define

$$F + e = F \cup \{b\} \quad (5)$$

Similarly, for an edge $e = (a, b) \in \partial^- F$, define

$$F \setminus \{e\} = F \setminus \{b\} \quad (6)$$

For a sequence of random variables $\{X_n\}_1^\infty$ and another random variable X , defined on the same probability space, by the notation $X_n \xrightarrow{p} X$ we mean that the sequence of random variables $\{X_n\}_1^\infty$ converges in probability to the random variable X [2].

3.2 System Dynamics

Consider any broadcast policy π in action. For any reachable set $F \subsetneq V$, denote the number of packets, replicated *exactly* at the vertex-set F at mini-slot t , by $Q_F(t)$. A packet p , which is replicated exactly at the set F by time t , is called a *class- F packet*. Hence, at a given time t , the reachable sets $F \in \mathcal{F}$ induce a disjoint partition of all the packets in the network. The variable $Q_F(t)$ denotes the number of packets in the partition corresponding to the reachable set F .

Because of our mini-slot model, a class- F packet can make a transition only to a class $F + e$ (where $e \in \partial^+ F$) during a mini-slot. Let the rate allocated to the edge e , for transmitting a class- F packet at time t , be denoted by $\mu_{e,F}(t)$ ³. Here $\mu_{e,F}(t)$ is a binary-valued control variable, which assumes the value 1 if the edge e (if active) is allocated to transmit a class- F packet at mini-slot t . The allocated rates are constrained by the underlying random edge-activation process $\{S(t)\}_0^\infty$. In particular, $\mu_{e,F}(t)$ is zero unless $S(t) = e$. In the following we argue that, for any reachable set F , the variable $Q_F(t)$ satisfies following one-step queuing-dynamics (Lindley recursion) [3]:

$$Q_F(t+1) \leq \left(Q_F(t) - \sum_{e \in \partial^+ F} \mu_{e,F}(t) \right)^+ + \sum_{(e,G): e \in \partial^- F, G=F \setminus \{e\}} \mu_{e,G}(t), \quad \forall F \neq \{\mathbf{r}\} \quad (7)$$

$$Q_{\{\mathbf{r}\}}(t+1) \leq \left(Q_{\{\mathbf{r}\}}(t) - \sum_{e \in \partial^+(\{\mathbf{r}\})} \mu_{e,\{\mathbf{r}\}}(t) \right)^+ + A(t)$$

³Note that $\mu_{e,F}(t)$ and consequently, $Q_F(t)$ depends on the algorithm π in use and should be denoted by $\mu_{e,F}^\pi(t)$ and $Q_F^\pi(t)$. Here we drop the superscript π to simplify notation.

The dynamics in Eqn. (7) may be explained as follows: because of the mini-slot model, only one packet can be transmitted in the entire network at any mini-slot. Hence, for any reachable set F , the value of the corresponding state-variable $Q_F(t)$ may go up or down by at most one in a mini-slot. Now, $Q_F(t)$ decreases by one when any of the out-edges $e \in \partial^+ F$ is activated at mini-slot t and it carries a class- F packet, provided $Q_F(t) > 0$. This explains the first term in Eqn. (7). Similarly, the variable $Q_F(t)$ increases by one when a packet in some set $G = F \setminus \{e\}$ (or an external packet, in case $F = \{\mathbf{r}\}$), is transmitted to the set F over the (active) edge $e \in \partial^- F$. This explains the second term in Eqn. (7). In the following, we slightly abuse the notation by setting $\sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mu_{e,G}(t) \equiv A(t)$, when $F = \{\mathbf{r}\}$. Thus the system dynamics is completely specified by the first inequality in (7), which constitutes a discrete time **Lindley recursion** [3].

3.3 Relationship between Stability and Efficiency

The following lemma shows equivalence between system-stability and throughput-optimality for a Markovian policy.

LEMMA 2 (Stability implies Efficiency). *A Markovian policy π , under which the induced Markov Chain $\{\mathbf{Q}^\pi(t)\}_{t=0}^\infty$ is Positive Recurrent for all arrival rate $\lambda < \lambda^*$, is a throughput optimal broadcast policy.*

PROOF. Under the action of a Markovian Policy π , the total number of packets $D^\pi(T)$ delivered to all nodes by slot T is given by

$$D^\pi(T) = \sum_{t=1}^T A(t) - \sum_F Q_F^\pi(T)$$

Hence, the rate of packet broadcast is given by

$$\begin{aligned} \liminf_{T \rightarrow \infty} \frac{D^\pi(T)}{T} &= \liminf_{T \rightarrow \infty} \left(\frac{1}{T} \sum_{t=1}^T A(t) - \sum_F \frac{Q_F^\pi(T)}{T} \right) \\ &\stackrel{p}{\Rightarrow} \lambda - \sum_F \limsup_{T \rightarrow \infty} \frac{Q_F^\pi(T)}{T} \end{aligned} \quad (8)$$

$$\stackrel{p}{\Rightarrow} \lambda \quad (9)$$

Eqn. (8) follows from the Weak Law of Large Numbers for the arrival process. To justify Eqn. (9), note that for any $\delta > 0$ and any reachable set F , we have

$$\lim_{T \rightarrow \infty} \mathbb{P} \left(\frac{Q_F^\pi(T)}{T} > \delta \right) = \lim_{T \rightarrow \infty} \mathbb{P} \left(Q_F^\pi(T) > T\delta \right) = 0, \quad (10)$$

where the last equality follows from the definition of positive recurrence. Eqn. (10) implies that $\frac{Q_F^\pi(T)}{T} \xrightarrow{p} 0, \forall F$. This justifies Eqn. (9) and proves the lemma. \square

3.3.1 Stochastic Stability of the Process $\{\mathbf{Q}(t)\}_{t \geq 1}$

Equipped with Lemma (2), we now focus on finding a Markovian policy π^* , which stabilizes the chain $\mathbf{Q}^{\pi^*}(t)$ ⁴.

⁴The time-index t denotes time in mini-slots.

To accomplish this goal, we use the Lyapunov drift methodology [5], and derive a dynamic policy π^* which minimizes the one-minislot drift of a certain Lyapunov function. We then show that the proposed policy π^* has negative drift outside a bounded region in the state-space. Upon invoking the Foster-Lyapunov criterion [12], this proves positive recurrence of the chain $\{\mathbf{Q}(t)\}_0^\infty$.

To apply the scheme outlined above, we start out by defining the following Quadratic Lyapunov Function $L(\mathbf{Q}(t))$:

$$L(\mathbf{Q}(t)) = \sum_F Q_F^2(t), \quad (11)$$

where the sum extends over all reachable sets. Recall that, the r.v. $S(t)$ denotes the currently active edge at the mini-slot t . The one-minislot drift is defined as:

$$\Delta_t(\mathbf{Q}(t), S(t)) \equiv L(\mathbf{Q}(t+1)) - L(\mathbf{Q}(t)) \quad (12)$$

From the dynamics (7), we have

$$\begin{aligned} Q_F^2(t+1) &\leq Q_F^2(t) + \mu_{\max}^2 \\ &- 2Q_F(t) \left(\sum_{e \in \partial^+ F} \mu_{e,F}(t) - \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mu_{e,G}(t) \right), \end{aligned}$$

where $\mu_{\max} = 1$ is the maximum capacity of a link per mini-slot. Thus the one mini-slot drift may be upper-bounded as follows:

$$\begin{aligned} \Delta_t(\mathbf{Q}(t), S(t)) &\leq 2^n \mu_{\max}^2 \\ &- 2 \sum_{F \subseteq V} Q_F(t) \left(\sum_{e \in \partial^+ F} \mu_{e,F}(t) - \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mu_{e,G}(t) \right). \end{aligned}$$

Interchanging the order of summation, we have

$$\begin{aligned} \Delta_t(\mathbf{Q}(t), S(t)) &\leq 2^n \mu_{\max}^2 \\ &- \sum_{(e,F):e \in \partial^+ F} \mu_{e,F}(t) \left(Q_F(t) - Q_{F+e}(t) \right) \end{aligned}$$

Taking expectation of both sides of the above inequality with respect to the edge-activation process $S(t)$ and the arrival process $A(t)$, we obtain the following upper-bound on the conditional Lyapunov drift $\Delta_t(\mathbf{Q}(t))$, defined as follows:

$$\Delta_t(\mathbf{Q}(t)) \equiv \mathbb{E}_{S(t)} \Delta_t(\mathbf{Q}(t), S(t)) \quad (13)$$

$$\leq 2^n \mu_{\max}^2 -$$

$$\sum_{(e,F):e \in \partial^+ F} \left(Q_F(t) - Q_{F+e}(t) \right) \mathbb{E}(\mu_{e,F}(t) | \mathbf{Q}(t), S(t))$$

Due to the activity constraint, if $S(t) = e$, we must have $\mu_{l,G}(t) = 0, \forall l \neq e$, for all reachable sets G . In other words, a packet can only be transmitted along the *active* edge for the mini-slot t . Eqn. (13) immediately leads us to Algorithm 1, which is obtained by minimizing the right hand side of the above upper-bound point-wise. For a reachable set F and an out-edge $e \in \partial^+ F$, define the weight

$$w_{F,e}(t) = Q_F(t) - Q_{F+e}(t) \quad (14)$$

We now state the main theorem of this paper.

Algorithm 1 A Dynamic Broadcast Policy π^*

At each mini-slot t , the network-controller observes the state-vector $\mathbf{Q}(t)$ and the currently active edge $S(t) = e = (i, j)$ and executes the following steps

- 1: Compute all reachable sets F such that $e \in \partial^+ F$.
 - 2: Transmit a class- F packet over the edge e , such that the corresponding weight $w_{F,e}(t) = Q_F(t) - Q_{F+e}(t)$ is positive and achieves the maximum over all such reachable sets F under consideration in step 1.
 - 3: Idle, if no such F exists.
-

THEOREM 1 (THROUGHPUT-OPTIMALITY OF π^*).
The dynamic policy π^ is a throughput-optimal broadcast policy for any network with arbitrary topology.*

PROOF. See Appendix (9.1). \square

4. A MULTI-CLASS BROADCASTING HEURISTIC

We note that, the policy π^* makes dynamic routing and scheduling decision for each individual packets, based on the current network-state information $\mathbf{Q}(t)$. In particular, its operation does not depend on the global topology information of the network. This robustness property makes the policy π^* suitable for use in mobile adhoc wireless networks (MANET), where the underlying topology may change frequently. However, a potential difficulty in implementing the policy π^* is that, one needs to maintain a state-variable $Q_F(t)$, corresponding to each reachable set F , and keep track of the particular reachable set $F_p(t)$ to which packet p belongs. For large networks, without any additional structure in the scheduling policy, maintaining such detailed state-information is quite cumbersome. To alleviate this problem, we next propose a heuristic policy which combines the Max-weight scheduling algorithm designed for π^* , with the novel idea of *in-class in-order delivery scheme*. The introduction of class-based in-order delivery imposes additional structure in the packet scheduling, which in turn, substantially reduces the complexity of the state-space.

Motivation. To motivate the heuristic policy, we begin with a simple policy-space $\Pi^{\text{in-order}}$, first introduced in [8] for throughput-optimal broadcasting in wireless Directed Acyclic Graphs (DAG). In the space $\Pi^{\text{in-order}}$, the packets are delivered to nodes according to their order of arrival at the source. Unfortunately, as shown in [8], although $\Pi^{\text{in-order}}$ is sufficient for achieving throughput-optimality in a DAG, it is not necessarily throughput-optimal for arbitrary networks, containing directed cycles. To tackle this problem, we generalize the idea of in-order delivery by proposing a k -class policy-space $\Pi_k^{\text{in-order}}$, $k \geq 1$, which generalizes the space $\Pi^{\text{in-order}}$. In this space, the packets are separated in k classes. The in-order delivery constraint is imposed in each class but not across classes. Thus, in $\Pi_k^{\text{in-order}}$, the scheduling constraint of $\Pi^{\text{in-order}}$ is relaxed by requiring that packets belonging to *each individual class* be delivered to nodes according to their order of arrival at the source.

However, the space $\Pi_k^{\text{in-order}}$ does not impose any such restrictions on packet-delivery from different classes. Combining it with the max-weight scheduling scheme, designed earlier for the throughput-optimal policy π^* , we propose a multi-class heuristic policy $\pi_k^H \in \Pi_k^{\text{in-order}}$ which is *conjectured* to be throughput-optimal for large-enough number of classes k . Extensive numerical simulations have been carried out to support this conjecture.

The following section gives detailed description of this heuristic policy, outlined above.

4.1 The In-order Policy-Space $\Pi^{\text{in-order}}$

Now we formally define the policy-space $\Pi^{\text{in-order}}$:

DEFINITION 5 (POLICY-SPACE $\Pi^{\text{in-order}}$ [8]). *A broadcast policy π belongs to the space $\Pi^{\text{in-order}}$ if all incoming packets at the source r are serially indexed $\{1, 2, 3, \dots\}$, according to their order of arrivals and a node $i \in V$ is allowed to receive a packet p at time t only if the node i has received the packets $\{1, 2, \dots, p-1\}$ by time t .*

As a result of the *in-order* delivery property of policies in the space $\Pi^{\text{in-order}}$, it follows that the configuration of the packets in the network at time t may be completely represented by the n -dimensional vector $\mathbf{R}(t)$, where $R_i(t)$ denotes the highest index of the packet received by node $i \in V$ by time t . We emphasize that this succinct representation of network-state is valid only under the action of the policies in the space $\Pi^{\text{in-order}}$, and is not necessarily true in the general policy-space Π .

Due to the highly-simplified state-space representation, it is natural to try to find efficient broadcast-policies in the space $\Pi^{\text{in-order}}$ for arbitrary network topologies. It is shown in [8] that if the underlying topology of the network is restricted to DAGs, the space $\Pi^{\text{in-order}}$ indeed contains a throughput-optimal broadcast policy. However, it is also shown that the space $\Pi^{\text{in-order}}$ is not rich-enough to achieve broadcast capacity in networks with arbitrary topology. We re-state the following proposition in this connection.

PROPOSITION 1. (THROUGHPUT-LIMITATION OF THE SPACE $\Pi^{\text{in-order}}$ [8]) *There exists a network \mathcal{G} such that, no broadcast-policy in the space $\Pi^{\text{in-order}}$ can achieve the broadcast-capacity of \mathcal{G} .*

The proof of the above proposition is given in [8], where it is shown that no broadcast policy in the space $\Pi^{\text{in-order}}$ can achieve the broadcast-capacity in the diamond-network \mathcal{D}_4 , shown in Figure 1.

4.2 The Multi-class Policy-Space $\Pi_k^{\text{in-order}}$

To overcome the throughput-limitation of the space $\Pi^{\text{in-order}}$, we propose the following generalized policy-space $\Pi_k^{\text{in-order}}$, $k \geq 1$, which retains the efficient representation property of the space $\Pi^{\text{in-order}}$.

DEFINITION 6 (POLICY-SPACE $\Pi_k^{\text{in-order}}$). *A broadcast policy π belongs to the space $\Pi_k^{\text{in-order}}$ if the following conditions hold:*

- *There are k distinct “classes”.*

- A packet, upon arrival at the source, is labelled with any one of the k classes, uniformly at random. This label of a packet remains fixed throughout its course of broadcast.
- Packets belonging to each individual class $j \in [1, \dots, k]$, are serially indexed $\{1, 2, 3, \dots\}$ according to their order of arrival.
- A node $i \in V$ in the network is allowed to receive a packet p from class j at time t , only if the node i has received the packets $\{1, 2, \dots, p-1\}$ from the class j by time t .

In other words, in the policy-space $\Pi_k^{\text{in-order}}$, packets belonging to each individual class $j \in [1, \dots, k]$ are delivered to nodes in-order. It is also clear from the definition that

$$\Pi_1^{\text{in-order}} = \Pi^{\text{in-order}}$$

Thus, the space $\{\Pi_k^{\text{in-order}}, k \geq 1\}$ generalizes the space $\Pi^{\text{in-order}}$.

State-Space representation under $\Pi_k^{\text{in-order}}$. Since each class in the policy-space $\Pi_k^{\text{in-order}}$ obeys the in-order delivery property, it follows that the network-state at time t is completely described by the k -tuple of vectors $\{\mathbf{R}^j(t), 1 \leq j \leq k\}$, where $R_i^j(t)$ denotes the highest index of the packet received by node $i \in V$ from class j by time t . Thus the state-space complexity grows *linearly* with the number of classes used.

Following our development so far, it is natural to seek a throughput-optimal broadcast policy in the space $\Pi_k^{\text{in-order}}$ with a *small* class-size k . In contrast to Proposition (1), the following proposition gives a positive result in this direction.

PROPOSITION 2. (THROUGHPUT-OPTIMALITY OF THE SPACE $\Pi_k^{\text{in-order}}, k \geq n/2$) For every network \mathcal{G} , there exists a throughput-optimal broadcast policy in the policy-space $\Pi_k^{\text{in-order}}$ where $k \geq n/2$.

The proof of this proposition uses a static policy, which routes the incoming packets along a set of λ^* edge-disjoint spanning trees. For a network with broadcast-capacity λ^* , the existence of these trees are guaranteed by Edmonds' tree packing theorem [6]. Then we show that for any network with unit-capacity edges, its broadcast-capacity λ^* is upper-bounded by $n/2$, which completes the proof. The details of this proof are outlined in Appendix (9.5).

The policy-class $\Pi_k^{\text{in-order}}$ fixes *intra-class* packet scheduling, by definition. Finally, we need an *inter-class* scheduling policy to resolve contentions among packets from different classes. In the following section, we propose such a scheme.

4.3 A Multi-class Heuristic Policy $\pi_k^H \in \Pi_k^{\text{in-order}}$

In this sub-section, we propose a dynamic policy $\pi_k^H \in \Pi_k^{\text{in-order}}$, which uses the same Max-Weight packet scheduling rule, as the throughput-optimal policy π^* , for *inter-class* packet scheduling. As we will see, the computation of weights and packet scheduling in this case may be efficiently carried out by exploiting the special structure of the space $\Pi_k^{\text{in-order}}$.

We observe that, when the number of classes $k = \infty$ and every incoming packet to the source \mathbf{r} joins a new class, the *in-order* restriction of the space $\Pi_k^{\text{in-order}}$ is essentially no longer in effect. In particular, the throughput-optimal policy π^* of Section 3 belongs to the space $\Pi_\infty^{\text{in-order}}$. However, we conjecture that the space $\Pi_k^{\text{in-order}}$ is throughput-optimal even when $k = \mathcal{O}(\text{poly}(n))$. Numerical simulation results, supporting this conjecture will be shown subsequently.

The packet-scheduling algorithm of the policy π_k^H may be formally described in the following two parts:

Intra-class packet scheduling. As in all policies in the class $\Pi_k^{\text{in-order}}$, when a packet p arrives at the source \mathbf{r} , it is placed into one of the k classes *uniformly at random*. Packets belonging to any class $c = 1, 2, \dots, k$ are delivered to all nodes *in-order* (i.e. the order they arrived at the source \mathbf{r}). Let the state-variable $R_i^c(t)$ denote the number of packets belonging to the class c received by node i up to the mini-slot t , $i = 1, 2, \dots, n$, $c = 1, 2, \dots, k$. As discussed earlier, given the intra-class in-order delivery restriction, the state of the network at the mini-slot t is completely specified by the vector $\{\mathbf{R}^c(t), c = 1, 2, \dots, k\}$.

Again, because of the in-order packet-delivery constraint, when an edge $e = (i, j)$ is active at the mini-slot t , not all packets that are present at node i and not-present at node j are eligible for transmission. Under the policy $\pi_k^H \in \Pi_k^{\text{in-order}}$, only the next *Head-of-the-Line* (HOL) packet from each class, i.e., packet with index $R_j^c(t) + 1$ from the class c , $c = 1, 2, \dots, k$ are eligible to be transmitted to the node j , provided that the corresponding packet is also present at node i by mini-slot t . Hence, at a given mini-slot t , there are at most k contending packets for an active edge. This should be compared with the policy π^* , in which there are potentially $\mathcal{O}(\exp(n))$ contending packets for an active edge at a mini-slot.

Inter-class packet scheduling. Given the above intra-class packet-scheduling rule, which follows straight from the definition of the space $\Pi_k^{\text{in-order}}$, we now propose an inter-class packet scheduling, for resolving the contention among multiple contending classes for an active edge e at a mini-slot t . For this purpose, we utilize the same Max-Weight scheduling rule, derived for the policy π^* (step 2 of Algorithm 1). However, instead of computing the weights $w_{F,e}(t)$ in (14) for all reachable sets F , in this case we only need to compute the weights of the sets F_c corresponding to the HOL packets (if any) belonging to the class c . This amounts to a linear number of computations in the class-size k . Finally, we schedule the HOL packet from the class c^* having the maximum (positive) weight. By exploiting the structure of the space $\Pi_k^{\text{in-order}}$, the computation of the weights w_c can be done in linear-time in the number of classes k . It appears from our extensive numerical simulations that $k = \mathcal{O}(m)$ classes suffice for achieving the broadcast capacity in any network.

Pseudo code. The full pseudo code of the policy π_k^H is provided in Algorithm 2. In lines 4...10, we have used the in-order delivery property of the policy π_k^H to compute the sets F_c , to which the next HOL packet from the class c belongs. This property is also used in computing the number

of packets in the set $G = F_c, F_{c+e}$ in line 14. Recall that, the variable $Q_G(t)$ counts the number of packets that the reachable set G contains exclusively at mini-slot t . These packets can be counted by counting such packets from each individual classes and then summing them up. Again utilizing the in-class in-order delivery property, a little thought reveals that the number of packets $N_G^c(t)$ from class c , that belongs exclusively to the set G at time t is given by

$$N_G^c(t) = \left(\min_{i \in G} R_i^c(t) - \max_{i \in V \setminus G} R_i^c(t) \right)^+$$

Hence,

$$Q_G(t) = \sum_{c=1}^k N_G^c(t),$$

which explains the statement in line 14. In line 17, the *weights* corresponding to the HOL packets of each class is computed according to the formula (14). Finally, in line 19, the HOL packet with the highest positive weight is transmitted across the active edge e . The per mini-slot complexity of the policy π_k^H is $\mathcal{O}(nk)$.

Algorithm 2 The Multi-class Scheduling Policy π_k^H

At each mini-slot t , the network-controller observes the state-variables $\{R_j^c(t), j \in V, c = 1, 2, \dots, k\}$, the currently active edge $S(t) = e = (i, j)$ and executes the following steps

```

1: for all classes  $c = 1 : k$  do
2:   // Determine the index of the next in-order
   // (HOL) packet  $p_c$  from the class  $c$  for node  $j$ 
3:    $p_c \leftarrow R_j^c(t) + 1$ .
4:   // Find the subset  $F_c \subset V$  where the packet  $p_c$ 
   // is currently present:
5:    $F_c \leftarrow \phi$ 
6:   for all node  $i = 1 : n$  do
7:     if  $R_i^c(t) \geq p_c$  then
8:        $F_c \leftarrow F_c \cup \{i\}$ 
9:     end if
10:  end for
11:   $F_{c+e} = F_c \cup \{j\}$ 
12:  // Determine  $Q_{F_c}(t)$  and  $Q_{F_{c+e}}(t)$ 
13:  for all  $G = F_c, F_{c+e}$  do
14:     $Q_G(t) \leftarrow \sum_{c=1}^k \left( \min_{i \in G} R_i^c(t) - \max_{i \in V \setminus G} R_i^c(t) \right)^+$ 
15:  end for
16:  // Compute the weight  $w_c$  for packet  $p_c$ 
17:   $w_c \leftarrow (Q_{F_c}(t) - Q_{F_{c+e}}(t))$ 
18: end for
19: Schedule the packet  $p^* \in \arg \max_c w_c$ , when  $\max w_c > 0$ , else idle.

```

5. WIRELESS INTERFERENCE

A wireless network is modeled by a graph $\mathcal{G}(V, E)$, along with a set of subset (represented by the corresponding binary characteristic vector) of edges \mathcal{M} , called the set of *feasible activations* [9]. The structure of the set \mathcal{M} depends on the underlying interference constraint, e.g., under the *primary*

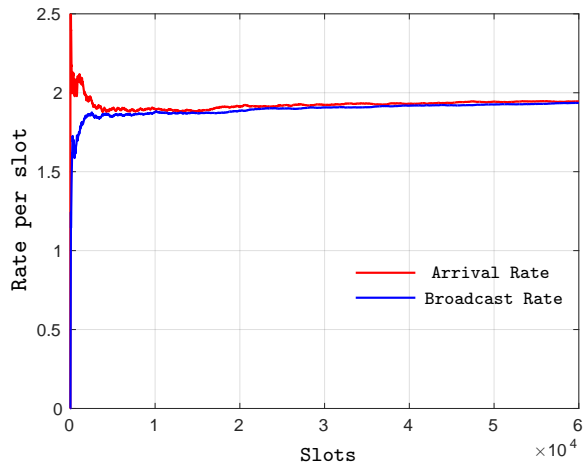


Figure 2: Packet Arrival and Broadcast Rate in the Diamond Network in Figure 1, under the action of the throughput-optimal policy π^*

interference constraint, the set \mathcal{M} consists of all *matchings* of the graph \mathcal{G} [11]. Any subset of edges $\mathbf{s} \in \mathcal{M}$ can be activated simultaneously at a given slot. For broadcasting in wireless networks, we first activate a feasible set of edges from \mathcal{M} and then forward packets on the activated edges. Since the proposed broadcast algorithms in sections 3 and 4 are *Max-Weight* by nature, they extend straight-forwardly to wireless networks with activation constraints [5]. In particular, from Eqn. (14), at each slot t , we first compute the weight of each edge, defined as $w_e(t) = \max_{F: e \in \partial^+ F} w_{e, F}(t)$. Next, we activate the subset of edges $\mathbf{s}^*(t)$ from the activation set \mathcal{M} , having the maximum weight, i.e.,

$$\mathbf{s}^*(t) = \arg \max_{\mathbf{s} \in \mathcal{M}} \sum_{e \in E} w_e(t) s_e$$

Packet forwarding over the activated edges remains the same as before. The above activation procedure carries over to the multi-class heuristic π_k^H in wireless networks.

6. NUMERICAL SIMULATIONS

6.1 Simulating the Throughput-optimal broadcast policy π^*

We simulate the policy π^* on the Diamond network \mathcal{D}_4 , shown in Figure 1. The broadcast-capacity of the network is 2 packets per slot. External packets arrive at the source node \mathbf{r} according to a Poisson process of a slightly lower rate of $\lambda = 1.95$ packets per slot. A packet is said to be broadcast when it reaches all the nodes in the network. The rate of packet arrival and packet broadcast by policy π^* , is shown in Figure 2. This plot exemplifies the throughput-optimality of the policy π^* for the diamond network.

6.2 Simulating the Multi-class Heuristic Policy π_k^H

The multi-class heuristic policy π_k^H has been numerically simulated with ~ 500 random networks. We have obtained similar qualitative results in all such instances. One representative sample is discussed here.

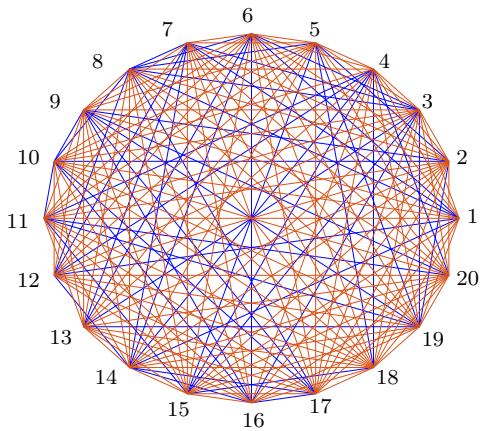


Figure 3: A network \mathcal{G} with $N = 20$ nodes. The colors of the edges indicate their directions (e.g., blue edge $\implies i \rightarrow j : i > j$ and vice versa). The broadcast capacity λ^* of the network is computed to be 6, with node 1 being the source node.

Consider running the broadcast-policy π_k^H on the network shown in Figure 3, containing $n = 20$ nodes and $m = 176$ edges. The directions of the edges in this network is chosen arbitrarily. With node 1 as the source node, we first compute the broadcast-capacity λ^* of this network using Eqn. (2) and obtain $\lambda^* = 6$. External packets arrive at the source node according to a Poisson process, with a slightly smaller rate of $\lambda = 5.95$ packets per slot. The rate of broadcast under the multi-class policy π_k^H for different values of k is shown in Figure 4. As evident from the plot, the achievable broadcast rate, obtained by the policy π_k^H , is non-decreasing in the number of classes k . Also, the policy π_k^H empirically achieves the broadcast-capacity of the network for a relatively small value of $k = 20$.

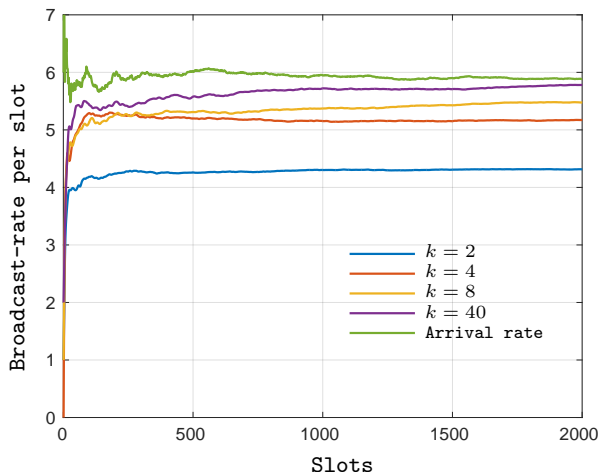


Figure 4: Achievable broadcast-rate with the multi-class heuristic broadcast-policies π_k^H , for $k = 2, 4, 8, 40$. The underlying network-topology is given in Figure 3.

7. CONCLUSION AND FUTURE WORK

In this paper we studied the problem of efficient, dynamic packet broadcasting in data networks with arbitrary underlying topology. We derived a throughput-optimal Max-weight broadcast policy that achieves the capacity, albeit at the expense of exponentially many counter-variables. To get around this problem, we next proposed a multi-class heuristic policy which combines the idea of in-order packet delivery with a Max-weight scheduling, resulting in drastic reduction in the implementation-complexity. The proposed heuristic with polynomially many classes is conjectured to be throughput-optimal. An immediate next step along this line of work would be to prove this conjecture. A problem of practical interest is to find the minimum number of classes $k^*(\epsilon)$ required to achieve a fraction $(1 - \epsilon)$ of the capacity.

8. REFERENCES

- [1] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms*. MIT press, 2009.
- [2] R. Durrett. *Probability: theory and examples*. Cambridge university press, 2010.
- [3] D. V. Lindley. The theory of queues with a single server. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 48, pages 277–289. Cambridge Univ Press, 1952.
- [4] L. Massoulié, A. Twigg, C. Gkantsidis, and P. Rodriguez. Randomized decentralized broadcasting algorithms. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 1073–1081. IEEE, 2007.
- [5] M. J. Neely. Stochastic network optimization with application to communication and queueing systems. *Synthesis Lectures on Communication Networks*, 3(1):1–211, 2010.
- [6] R. Rustin. *Combinatorial Algorithms*. Algorithmics Press, 1973.
- [7] S. Sarkar and L. Tassiulas. A framework for routing and congestion control for multicast information flows. *Information Theory, IEEE Transactions on*, 48(10):2690–2708, 2002.
- [8] A. Sinha, G. Paschos, C. ping Li, and E. Modiano. Throughput-optimal broadcast on directed acyclic graphs. In *Computer Communications (INFOCOM), 2015 IEEE Conference on*, pages 1248–1256, April 2015.
- [9] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *Automatic Control, IEEE Transactions on*, 37(12):1936–1948, 1992.
- [10] D. Towsley and A. Twigg. Rate-optimal decentralized broadcasting: the wireless case, 2008.
- [11] D. B. West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.
- [12] E. Wong and B. Hajek. *Stochastic processes in engineering systems*. Springer Science & Business Media, 2012.
- [13] S. Zhang, M. Chen, Z. Li, and L. Huang. Optimal distributed broadcasting with per-neighbor queues in acyclic overlay networks with arbitrary underlay capacity constraints. In *Information Theory*

9. APPENDIX

9.1 Proof of Throughput Optimality of π^*

In this subsection, we show that the induced Markov-Chain $\mathbf{Q}^{\pi^*}(t)$, generated by the policy π^* is positive recurrent, for all arrival rates $\lambda < \lambda^*$ packets per slot. This is proved by showing that the expected one-minislot drift of the Lyapunov function $L(\mathbf{Q}(t))$ is negative outside a bounded region in the non-negative orthant \mathbb{Z}_+^M , where M is the dimension of the state-space $\mathbf{Q}(t)$. To establish the required drift-condition, we first construct an auxiliary stationary randomized policy π^{RAND} , which is easier to analyze. Then we bound the one-minislot expected drift of the policy π^* by comparing it with the policy π^{RAND} .

We emphasize that the construction of the randomized policy π^{RAND} is highly non-trivial, because under the action of the policy π^* , a packet may travel along an arbitrary tree and as a result, *any* reachable set $F \in \mathcal{F}$ may potentially contain non-zero number of packets.

For ease of exposition, the proof of throughput-optimality of the policy π^* is divided into several parts.

9.1.1 Part I: Consequence of Edmonds' Tree-packing Theorem

From Edmond's tree-packing theorem [6], it follows that the graph \mathcal{G} contains λ^* edge-disjoint directed spanning trees, ⁵ $\{\mathcal{T}^i\}_{i=1}^{\lambda^*}$. From Proposition (1) and Lemma (2), it follows that, to prove the throughput-optimality of the policy π^* , it is sufficient to show stochastic-stability of the process $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ for an arrival rate of λ/m per minislot, where $\lambda < \lambda^*$.

Fix an arbitrarily small $\epsilon > 0$ such that,

$$\lambda \leq \lambda^* - \epsilon$$

Now we construct a stationary randomized policy π^{RAND} , which utilizes the edge-disjoint trees $\{\mathcal{T}^i\}_{i=1}^{\lambda^*}$ in a critical fashion.

9.1.2 Part II: Construction of a Stationary Randomized Policy π^{RAND}

The stationary randomized policy π^{RAND} allocates rates $\mu_{e,F}(t)$ randomly to different ordered pairs (e, F) , for transmitting packets belonging to reachable sets F , across an edge $e \in \partial^+ F$.⁶ Recall that $\mu_{e,F}(t)$'s are binary variables. Hence, conditioned on the edge-activity process $S(t) = e$, the allocated rates are fully specified by the set of probabilities that a packet from the reachable set F is transmitted across the active edge $e \in \partial^+ F$. Equivalently, we may specify the allocated rates in terms of their expectation w.r.t. the edge-activation process (obtained by multiplying the corresponding probabilities by $1/m$).

Informally, the policy π^{RAND} allocates most of the rates along the reachable sequences corresponding to the edge-disjoint spanning trees $\{\mathcal{T}^i\}_{i=1}^{\lambda^*}$, obtained in Part I. However,

⁵Note that, since the edges are assumed to be of unit capacity, λ^* is an integer. This result follows by combining Eqn. (2) with the Max-Flow-Min-Cut theorem [1].

⁶If $e \notin \partial^+ F$, naturally $\mu_{e,F}(t) = 0, \forall t$.

since the dynamic policy π^* is not restricted to route packets along the spanning trees $\{\mathcal{T}^i\}_{i=1}^{\lambda^*}$ only, for technical reasons which will be evident later, π^{RAND} is designed to allocate small amount of rates along other reachable sequences. This is an essential and non-trivial part of the proof methodology. An illustrative example of the rate allocation strategy by the policy π^{RAND} will be described subsequently for the diamond graph \mathcal{D}_4 of Figure 1.

Formally, the rate-allocation by the randomized policy π^{RAND} is given as follows:

- We index the set of all reachable sequences in a specific order.

- The first λ^* reachable sequences $\{\zeta^i\}_{i=1}^{\lambda^*}$ are defined as follows: for each edge-disjoint tree $\mathcal{T}^i, i = 1, 2, \dots, \lambda^*$ obtained from Part-I, recursively construct a reachable sequence $\zeta^i = \{(F_j^i, e_j^i)\}_{j=1}^{n-1}$, such that the induced sub-graphs $\mathcal{T}^i(F_j^i)$ are connected for all $j = 1, 2, \dots, n - 1$.

In other words, for all $1 \leq i \leq \lambda^*$ define $F_1^i = \{\mathbf{r}\}$ and for all $1 \leq j \leq n - 2$, the set F_{j+1}^i is recursively constructed from the set F_j^i by adding a node to the set F_j^i while traversing along an edge of the tree \mathcal{T}^i . Let the corresponding edge in \mathcal{T}^i connecting the $j + 1$ th vertex $F_{j+1}^i \setminus F_j^i$, to the set F_j^i , be e_j^i . Since the trees $\{\mathcal{T}^i\}_{i=1}^{\lambda^*}$ are edge disjoint, the edges e_j^i 's are **distinct** for all $i = 1, 2, \dots, \lambda^*$ and $j = 1, 2, \dots, n - 1$. The above construction defines the first λ^* reachable sequences $\zeta^i = \{F_j^i, e_j^i\}_{j=1}^{n-1}, 1 \leq i \leq \lambda^*$.

- In addition to the above, let $\{\zeta^i = (F_j^i, e_j^i)\}_{j=1}^{n-1}, \lambda^* + 1 \leq i \leq B$ be the set of all *other* reachable sequence in the graph \mathcal{G} , different from the previously constructed λ^* reachable sequences. Recall that, B is the cardinality of the set of all reachable sequences in the graph \mathcal{G} . Thus the set of *all* reachable sequences in the graph \mathcal{G} is given by $\bigcup_{i=1}^B \zeta^i$.

- To define the expected allocated rates $\mathbb{E}\mu_{e,F}(t)$, it is useful to first define some auxiliary variables, called *rate-components* $\mathbb{E}\mu_{e,F}^i(t), i = 1, 2, \dots, B$, corresponding to each reachable sequence. The rate $\mathbb{E}\mu_{e,F}(t)$ is simply the sum of the rate-components, as given in Eqn. (17).

At each slot t and $1 \leq i \leq \lambda^*$, the randomized policy allocates i^{th} *rate-component* corresponding to the reachable sequence $\zeta^i = \{e_j^i, F_j^i\}_{j=1}^{n-1}$ according to the following scheme:

$$\begin{aligned} \mathbb{E}(\mu_{e_j^i, F_j^i}^i(t)) &= 1/m - \epsilon(n-j)/n, \\ &\quad \forall 1 \leq j \leq n-1 \\ &= 0, \quad \text{o.w.} \end{aligned} \quad (15)$$

- In addition to the rate-allocation (15), the randomized policy π^{RAND} also allocates small amount of rates corresponding to other reachable sequences $\{\zeta^i\}_{i=\lambda^*+1}^B$ according to the following scheme: For $\lambda^* + 1 \leq i \leq B$, the randomized policy allocates i^{th} rate-component to

the ordered pairs (e, F) as follows:

$$\begin{aligned}\mathbb{E}(\mu_{e_j^i, F_j^i}(t)) &= \frac{\epsilon}{2nB} - \frac{\epsilon}{2nB} \frac{n-j}{n}, \\ &\quad \forall 1 \leq j \leq n-1, \\ &= 0, \quad \text{o.w.}\end{aligned}\quad (16)$$

The overall rate allocated to the pair (e, F) is simply the sum of the component-rates, as given below:

$$\mathbb{E}\mu_{e,F}(t) = \sum_{i=1}^B \mathbb{E}\mu_{e,F}^i(t) \quad (17)$$

In the following, we show that the above rate-allocation is feasible with respect to the edge capacity constraint.

LEMMA 3 (Feasibility of Rate Allocation). *The rate allocation (17) by the randomized policy π^{RAND} is feasible.*

The reader is referred to Appendix (9.3) for the proof the lemma. An illustrative example for the above randomized rate-allocation scheme is given in Appendix (9.4).

9.1.3 Part III: Comparison of drifts under action of policies π^* and π^{RAND}

Recall that, from Eqn. (13) we have the following upper-bound on the one-minislot drift of the Lyapunov function $L(\mathbf{Q}(t))$, achieved by the policy π^* :

$$\begin{aligned}(\Delta^{\pi^*}(\mathbf{Q}(t)|S(t)) &\leq 2^n \mu_{\max}^2 - \\ &\sum_{(e,F):e \in \partial^+ F} \left(Q_F(t) - Q_{F+e}(t) \right) \mathbb{E}(\mu_{e,F}^{\pi^*}(t)|\mathbf{Q}(t), S(t))\end{aligned}$$

Since the policy π^* , by definition, transmits packets to maximize the weight $w_{F,e}(t) = Q_F(t) - Q_{F+e}(t)$ point wise, the following inequality holds

$$\begin{aligned}&\sum_{(e,F):e \in \partial^+ F} \left(Q_F(t) - Q_{F+e}(t) \right) \mathbb{E}(\mu_{e,F}^{\pi^*}(t)|\mathbf{Q}(t), S(t)) \geq \\ &\sum_{(e,F):e \in \partial^+ F} \left(Q_F(t) - Q_{F+e}(t) \right) \mathbb{E}(\mu_{e,F}^{\pi^{\text{RAND}}}(t)|\mathbf{Q}(t), S(t)),\end{aligned}$$

where the randomized rate-allocation $\mu^{\pi^{\text{RAND}}}$ is given by Eqn. (17). Noting that π^{RAND} operates independently of the ‘‘queue-states’’ $\mathbf{Q}(t)$ and dropping the super-script π^{RAND} from the control variables $\mu(t)$ on the right hand side, we can bound the drift of the policy π^* as follows:

$$\begin{aligned}(\Delta^{\pi^*}(\mathbf{Q}(t)|S(t)) &\leq 2^n \mu_{\max}^2 - \sum_{(e,F):e \in \partial^+ F} \left(Q_F(t) - Q_{F+e}(t) \right) \mathbb{E}(\mu_{e,F}(t)|S(t)) \\ &= 2^n \mu_{\max}^2 - \sum_F Q_F(t) \left(\sum_{e \in \partial^+ F} \mathbb{E}(\mu_{e,F}(t)|S(t)) \right. \\ &\quad \left. - \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mathbb{E}(\mu_{e,G}(t)|S(t)) \right)\end{aligned}$$

$$\stackrel{(a)}{=} 2^n \mu_{\max}^2 - \sum_F Q_F(t) \left(\sum_{e \in \partial^+ F} \left(\sum_{i=1}^B \mathbb{E}(\mu_{e,F}^i(t)|S(t)) \right) - \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \left(\sum_{i=1}^B \mathbb{E}(\mu_{e,G}^i(t)|S(t)) \right) \right),$$

where in (a) we have used Eqn. (17).

Taking expectation of both sides of the above inequality w.r.t the random edge-activation process $S(t)$ and interchanging the order of summation, we have

$$\begin{aligned}\Delta^{\pi^*}(\mathbf{Q}(t)) &\leq 2^n \mu_{\max}^2 - \sum_F Q_F(t) \sum_{i=1}^B \left(\sum_{e \in \partial^+ F} \mathbb{E}(\mu_{e,F}^i(t)) \right. \\ &\quad \left. - \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mathbb{E}(\mu_{e,G}^i(t)) \right), \quad (18)\end{aligned}$$

where the rate-components μ^i of the randomized policy π^{RAND} are defined in Eqns (15) and (16).

Fix a reachable set F , appearing in the outer-most summation of the above upper-bound (18). Now focus on the i^{th} reachable sequence $\zeta^i \equiv \{F_j^i, e_j^i\}_{j=1}^{n-1}$. We have two cases:

Case I: $F \notin \zeta^i$

Here, according to the allocations in (15) and (16), we have

$$\sum_{e \in \partial^+ F} \mathbb{E}(\mu_{e,F}^i(t)) \stackrel{(a)}{=} 0, \quad \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mathbb{E}(\mu_{e,G}^i(t)) \stackrel{(b)}{=} 0$$

Where the equality (a) follows from the assumption that $F \notin \zeta^i$ and equality (b) follows from the fact that positive rates are allocated only along the tree corresponding to the reachable sequence ζ^i . Hence, if no rate is allocated to drain packets outside the set F , π^{RAND} does not allocate any rate to route packets to the set F .

Case II: $F \in \zeta^i$

In this case, from Eqns. (15) and (16), it follows that

$$\begin{aligned}\left(\sum_{e \in \partial^+ F} \mathbb{E}(\mu_{e,F}^i(t)) - \sum_{(e,G):e \in \partial^- F, G=F \setminus \{e\}} \mathbb{E}(\mu_{e,G}^i(t)) \right) \\ = \begin{cases} \frac{\epsilon}{n}, & 1 \leq i \leq \lambda^* \\ \frac{\epsilon}{2n^2B}, & \lambda^* + 1 \leq i \leq B \end{cases} \quad (19)\end{aligned}$$

By definition, each reachable set is visited by at least one reachable sequence. In other words, there exists at least one $i, 1 \leq i \leq B$, such that $F \in \zeta^i$. Combining the above two cases, from the upper-bound (18) we conclude that

$$\Delta^{\pi^*}(\mathbf{Q}(t)) \leq 2^n \mu_{\max}^2 - \frac{\epsilon}{2n^2B} \sum_F Q_F(t), \quad (20)$$

where, the sum extends over *all* reachable sets. The drift is negative, i.e., $\Delta^{\pi^*}(\mathbf{Q}(t)) < -\epsilon$, when $\mathbf{Q}_F \in \mathcal{B}^c$, where

$$\mathcal{B} = \left\{ (Q_F \geq 0) : \sum_F Q_F \geq \frac{2n^2B}{\epsilon} (\epsilon + 2^n \mu_{\max}^2) \right\}$$

Invoking the Foster-Lyapunov criterion [12], we conclude that the Markov-Chain $\{\mathbf{Q}^{\pi^*}(t)\}_0^\infty$ is positive recurrent. Finally, throughput-optimality of the policy π^* follows from lemma 2. ■

9.2 Proof of Lemma (1)

PROOF. We prove this lemma in two parts. First, we upper-bound the achievable broadcast rate of the network under any policy in the mini-slot model by the broadcast capacity $\lambda^*(\mathcal{G})$ of the network in the usual slotted model, which is given by Eqn. (2). Next, in our main result in section (9.1), we constructively show that this rate is achievable, thus proving the lemma.

Let $\mathcal{C} \subsetneq V$ be a non-empty subset of the nodes in the graph \mathcal{G} such that $\mathbf{r} \in \mathcal{C}$. Since \mathcal{C} is a strict subset of V , there exists a node $i \in V$ such that $i \in \mathcal{C}^c$. Let the set $E(\mathcal{C})$ denote the set of all directed edges $e = (a, b)$ such that $a \in \mathcal{C}$ and $b \notin \mathcal{C}$. Denote $|E(\mathcal{C})|$ by $\text{Cut}(\mathcal{C})$. Using the MAX-FLOW-MIN-CUT theorem [1], the broadcast-capacity in the slotted model, given by Eqn. (2), may be alternatively represented as

$$\lambda^* = \min_{\mathcal{C} \subsetneq V, \mathbf{r} \in \mathcal{C}} \text{Cut}(\mathcal{C}) \quad (21)$$

Now let us proceed with the mini-slot model. Since all packets arrived at source \mathbf{r} that are received by the node i must cross some edge in the cut $E(\mathcal{C})$, it follows that, under any policy $\pi \in \Pi$, the total number of packets $R_i(t)$ that are received by node i up to mini-slot t is upper-bounded by

$$R_i(t) \leq \sum_{\tau=1}^t \sum_{e \in E(\mathcal{C})} \mathbb{1}(S(\tau) = e) = \sum_{e \in E(\mathcal{C})} \sum_{\tau=1}^t \mathbb{1}(S(\tau) = e) \quad (22)$$

Thus the broadcast-rate $\lambda_{\text{mini-slot}}^\pi$ achievable in the mini-slot model is upper-bounded by

$$\begin{aligned} \lambda_{\text{mini-slot}}^\pi &\stackrel{(a)}{\leq} \liminf_{t \rightarrow \infty} \frac{R_i(t)}{t} \stackrel{(b)}{\leq} \liminf_{t \rightarrow \infty} \frac{1}{t} \sum_{e \in E(\mathcal{C})} \sum_{\tau=1}^t \mathbb{1}(S(\tau) = e) \\ &= \sum_{e \in E(\mathcal{C})} \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=1}^t \mathbb{1}(S(\tau) = e) \stackrel{(c)}{=} \frac{1}{m} \text{Cut}(\mathcal{C}), \text{ w.p.1} \quad (23) \end{aligned}$$

Where the inequality (a) follows from the definition of broadcast-rate (1), inequality (b) follows from Eqn. (22) and finally, the equality (c) follows from the Strong Law of Large Numbers [2]. Since the inequality (23) holds for any cut $\mathcal{C} \subsetneq V$ containing the source \mathbf{r} and any policy π , from Eqn. (21) we have

$$\lambda_{\text{mini-slot}}^* \leq \lambda_{\text{mini-slot}}^\pi \leq \frac{1}{m} \text{Cut}(\mathcal{C}) \leq \frac{1}{m} \lambda^* \text{ per mini-slot} \quad (24)$$

Since according to the hypothesis of the lemma, a slot is identified with m mini-slots, the above result shows that

$$\lambda_{\text{mini-slot}}^* \leq \lambda^* \text{ per slot} \quad (25)$$

This proves that the capacity in the mini-slot model (per slot) is at most the capacity of the slotted-time model (given by Eqn. (2)). In section (3), we show that there exists a broadcast policy $\pi^* \in \Pi$ which achieves a broadcast-rate of λ^* packets per-slot in the mini-slot model. This concludes the proof of the lemma. \square

9.3 Proof of Lemma (3)

PROOF. The rate allocation (17) will be feasible if the sum of the allocated probabilities that an active edge e carries a class- F packet, for all reachable sets F , is at most unity. Since an edge can carry at most one packet per mini-slot, this feasibility condition is equivalent to the requirement that the

total expected rate, i.e., $\mathbb{E}\mu_e(t) = \sum_F \mathbb{E}\mu_{e,F}(t)$, allocated to an edge $e \in E$ by the randomized policy π^{RAND} does not exceed $\frac{1}{m}$ (the expected capacity of the edge per mini-slot). Since an edge e may appear at most once in any reachable sequence, the total rate allocated to an edge e by the randomized-policy π^{RAND} is upper-bounded by $\frac{1}{m} - \frac{\epsilon}{n} + (B - \lambda^*) \frac{\epsilon}{2nB} \leq \frac{1}{m} - \frac{\epsilon}{2n} < 1/m$. Hence the rate allocation by the randomized policy π^{RAND} is feasible. \square

9.4 An Example of Rate Allocation by the Stationary policy π^{RAND}

As an explicit example of the above stationary policy, consider the case of the Diamond network \mathcal{D}_4 , shown in Figure 1. The edges of the trees $\{\mathcal{T}^i, i = 1, 2\}$ are shown in blue and red colors in the figure. Then the randomized policy allocates the following rate-components to the edges, where the expectation is taken w.r.t. random edge-activations per mini-slot.

First we construct a reachable sequence ζ^1 consistent with the tree \mathcal{T}^1 as follows:

$$\zeta^1 = \{(\{\mathbf{r}\}, \mathbf{ra}), (\{\mathbf{r}, \mathbf{a}\}, \mathbf{ab}), (\{\mathbf{r}, \mathbf{a}, \mathbf{b}\}, \mathbf{bc})\}$$

Next we allocate the following rate-components as prescribed by π^{RAND} :

$$\begin{aligned} \mathbb{E}\mu_{\mathbf{ra}, \{\mathbf{r}\}}^1(t) &= 1/6 - 3\epsilon/4 \\ \mathbb{E}\mu_{\mathbf{ab}, \{\mathbf{r}, \mathbf{a}\}}^1(t) &= 1/6 - 2\epsilon/4 \\ \mathbb{E}\mu_{\mathbf{bc}, \{\mathbf{r}, \mathbf{a}, \mathbf{b}\}}^1(t) &= 1/6 - \epsilon/4 \\ \mathbb{E}\mu_{e, F}^1(t) &= 0, \quad \text{o.w.} \end{aligned}$$

Similarly for the tree \mathcal{T}^2 , we first construct a reachable sequence ζ^2 as follows:

$$\zeta^2 = \{(\{\mathbf{r}\}, \mathbf{rb}), (\{\mathbf{r}, \mathbf{b}\}, \mathbf{rc}), (\{\mathbf{r}, \mathbf{b}, \mathbf{c}\}, \mathbf{ca})\}$$

Then we allocate the following component-rates to the (edge, set) pairs as follows:

$$\begin{aligned} \mathbb{E}\mu_{\mathbf{rb}, \{\mathbf{r}\}}^2(t) &= 1/6 - 3\epsilon/4 \\ \mathbb{E}\mu_{\mathbf{rc}, \{\mathbf{r}, \mathbf{b}\}}^2(t) &= 1/6 - 2\epsilon/4 \\ \mathbb{E}\mu_{\mathbf{ca}, \{\mathbf{r}, \mathbf{b}, \mathbf{c}\}}^2(t) &= 1/6 - \epsilon/4 \\ \mathbb{E}\mu_{e, F}^2(t) &= 0, \quad \text{o.w.} \end{aligned}$$

In this example $\lambda^* = 2$, thus these two reachable sequence accounts for a major portion of the rates allocated to the edges. The randomized policy π^{RAND} , however, allocates small rates to other reachable sequences too. As an example, consider the following reachable sequence ζ^3 , given by

$$\zeta^3 = \{(\{\mathbf{r}\}, \mathbf{ra}), (\{\mathbf{r}, \mathbf{a}\}, \mathbf{rb}), (\{\mathbf{r}, \mathbf{a}, \mathbf{b}\}, \mathbf{rc})\}$$

Then, as prescribed above, the randomized policy allocates the following rate-components

$$\begin{aligned} \mathbb{E}\mu_{\mathbf{ra}, \{\mathbf{r}\}}^3(t) &= \frac{\epsilon}{8B} - \frac{3\epsilon}{32B} \\ \mathbb{E}\mu_{\mathbf{rb}, \{\mathbf{r}, \mathbf{a}\}}^3(t) &= \frac{\epsilon}{8B} - \frac{2\epsilon}{32B} \\ \mathbb{E}\mu_{\mathbf{rc}, \{\mathbf{r}, \mathbf{a}, \mathbf{b}\}}^3(t) &= \frac{\epsilon}{8B} - \frac{\epsilon}{32B} \\ \mathbb{E}\mu_{e, F}^3(t) &= 0, \quad \text{o.w.} \end{aligned}$$

Here B is the number of all distinct reachable sequences, which is upper-bounded by 4^8 . The rate-components corresponding to other reachable sequences may be computed as above. Finally, the actual expected rate-allocation to the pair (e, F) is given by

$$\mathbb{E}\mu_{e,F}(t) = \sum_{i=1}^B \mathbb{E}\mu_{e,F}^i(t)$$

9.5 Proof of Proposition (2)

The proof of this proposition is conceptually simplest in the slotted-time model. The argument also applies directly to the mini-slot model.

Consider a network \mathcal{G} with broadcast-capacity λ^* . Assume a slotted-time model. By Edmonds' tree-packing Theorem [6], we know that there exists λ^* number of edge-disjoint directed spanning trees (arborescences) $\{\mathcal{T}_i\}_1^{\lambda^*}$ in \mathcal{G} , rooted at the source node \mathbf{r} . Now consider a policy $\pi \in \Pi_k^{\text{in-order}}$ with $k \geq \lambda^*$ which operates as follows:

- An incoming packet is placed in any of the classes $[1, 2, 3, \dots, \lambda^*]$, uniformly at random.
- Packets in a class i are routed to all nodes in the network *in-order* along the directed tree \mathcal{T}_i , where the packets are replicated in all non-leaf nodes of the tree \mathcal{T}_i , $1 \leq i \leq \lambda^*$.

Since the trees are edge-disjoint, the classes do not interact; i.e., routing in each class can be carried out independently. Also by the property of \mathcal{T}_i , there is a *unique* directed path from the source node \mathbf{r} to any other node in the network along the edges of the tree \mathcal{T}_i , $1 \leq i \leq \lambda^*$. Thus packets in every class can be delivered to all nodes in the network *in-order* in a pipe-lined fashion with the long-term delivery-rate of 1 packet per class. Since there are λ^* packet-carrying classes, it follows that the policy $\pi \in \Pi_k^{\text{in-order}}$ is throughput-optimal for $k \geq \lambda^*$.

Next we show that, $\lambda^* \leq n/2$ for a simple network. Since there exist λ^* number of edge-disjoint directed spanning trees in the network, and since each spanning-tree contains $n - 1$ edges, we have

$$\lambda^*(n - 1) \leq m \tag{26}$$

Where m is the number of edges in the network. But we have $m \leq n(n - 1)/2$ for a simple graph. Thus, from the above equation, we conclude that

$$\lambda^* \leq n/2. \tag{27}$$

This completes the proof of the Proposition.