Markov-Chain Modeling for Multicast Signaling Delay Analysis

Xi Zhang, *Senior Member, IEEE,* and Kang G. Shin, *Fellow, IEEE*

***Abstract*—Feedback signaling plays a key role in flow control because the traffic source relies on the signaling information to make *correct* and *timely* flow-control decisions. However, it is difficult to design an efficient signaling algorithm since a signaling message can tolerate neither error nor latency. Multicast flow-control signaling imposes two additional challenges: *scalability* and *feedback synchronization*. Previous research on multicast signaling has mainly focused on the development of algorithms without analyzing their delay performance. To remedy this deficiency, we have previously developed a binary-tree model and an *independent*-marking statistical model for multicast-signaling delay analysis. This paper considers a general scenario where the congestion markings at different links are *dependent*—a more accurate but complex case. Specifically, we develop a Markov-chain model defined by the link-marking state on each path in the multicast tree. The Markov chain can not only capture link-marking dependencies, but also yield a tractable analytical model. We also develop a Markov-chain dependency-degree model to evaluate all possible Markov-chain dependency degrees without any prior knowledge of them. Using the above two models, we derive the general probability distributions of each path becoming the multicast-tree bottleneck. Also derived are the first and second moments of multicast signaling delays. The proposed Markov chain is also shown to asymptotically reach an equilibrium, and its limiting distribution converges to the marginal link-marking probabilities when the Markov chain is irreducible. Applying the two models, we analyze and contrast the delay scalability of two representative multicast signaling protocols: *Soft-Synchronization Protocol* (SSP) and *Hop-By-Hop* (HBH) algorithms.**

***Index Terms*—Markov chain, multicast feedback synchronization/consolidation, multicast flow-control signaling, soft-synchronization protocol (SSP).**

I. Introduction

A. Background and Motivation

A FLOW-CONTROL algorithm consists of two basic components: rate control and flow-control signaling. These two components are conceptually separate from the flow-control theory's standpoint, but are often blended together in most flow-control algorithms. Rate control adjusts the source rate to the variation of bandwidth available in the network. Flow-control signaling conveys the congestion and rate-control information between the source and the network/receivers. This signaling is crucial to flow control because the source relies solely on the signaling information in making flow-control decisions. However, it is difficult to design an efficient flow-control signaling protocol because the signaling messages, unlike data or audio/video traffic, can tolerate neither error nor latency. A signaling message could be useless or even harmful if it is not accurate or its delay is unbounded. The delivery of signaling traffic must therefore be timely and reliable. For example, in ATM ABR service, flow-control signaling uses resource management (RM) cells to convey the rate-control and congestion information among the traffic source, switches, and the receivers.

Signaling for multicast flow control imposes two new problems: *scalability*[1] and *feedback synchronization* [1], [2]. These two problems are closely interwoven in the signaling protocol for multicast flow control. First, simultaneous feedback arrivals from all branches can cause *feedback implosion* [1], [2] at the source and branch nodes, especially when the multicast tree is large. Hence, each branch node need to consolidate the congestion feedbacks from its downstream paths and then forward only one consolidated feedback to its upstream node. Second, we need a feedback-synchronization signaling algorithm for each branch node to consolidate feedbacks, because they may arrive at significantly different times. To solve these problems, several multicast flow-control signaling algorithms have been proposed.

B. HBH and SSP Schemes for Multicast Signaling

The first-generation feedback consolidation algorithms [3]–[5] for multicast ABR flow control employ a simple *Hop-by-Hop* (HBH) mechanism to deal with the feedback-implosion problem. The HBH scheme works as follows. On receipt of one forward RM cell, each branch node sends only one consolidated feedback RM cell upward by a *single* hop, ensuring that at each node, the ratio of feedback RM cells to forward RM cells is no larger than 1. To reduce the RM-cell round-trip time (RTT) and improve multicast signaling accuracy, the authors of [6] proposed a different feedback-synchronization algorithm by accumulating feedbacks from *all* branches at each branch node. The authors of [7] proposed an algorithm to speed up the transient response by sending fast congestion feedback without waiting for all branches' feedbacks during the transient phase.

In [1] and [2], we proposed a feedback-synchronization signaling algorithm, called the *Soft-Synchronization Protocol* (SSP), which derives a single consolidated RM cell at each branch node from feedback RM cells of different downstream branches that are not necessarily responses to the same forward RM cell in each synchronization cycle. Fig. 1 shows the pseudocode of the SSP algorithm. When receiving a feed-

Manuscript received July 8, 2001; revised July 2, 2003; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor E. Biersack. This work was supported in part by the U.S. Office of Naval Research under Grant N00014-99-1-0465. A subset of the materials in this paper was presented at the IEEE INFOCOM, Anchorage, AK, 2001.

X. Zhang is with the Networking and Information Systems Laboratory, Department of Electrical Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: xizhang@ee.tamu.edu).

K. G. Shin is with the Real-Time Computing Laboratory, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: kgshin@eecs.umich.edu).

Digital Object Identifier 10.1109/TNET.2004.828943

[1]By saying that a multicast signaling protocol is scalable, we mean that the multicast signaling delay performance and cost/complexity do not get worse significantly as the multicast-tree size and structure scale up.

```
00. On receipt of a returned feedback RM cell from the i-th branch:
01.   if (conn_patt_vec(i) ≠ 1) { ! Only process connected branches
02.     resp_branch_vec(i) := 1; ! Mark connected and responsive branch
03.     MCI := MCI ∨ CI; ! CI is randomly marked and consolidated
04.     MER := min{MER, ER}; ! MER is the consolidated ER
05.     if (conn_patt_vec ⊕ resp_branch_vec = 1) { ! Soft Synchronization
06.       send RM cell (dir := back, ER := MER, CI := MCI);
07.       no_resp_timer := N_nrt; ! Reset non-responsive timer
08.       resp_branch_vec := 0); ! Reset the responsive branch vector;
09.       MCI := 0;  MER := ER;}}; ! Reset RM-cell control variable.
```

Fig. 1. Pseudocode for the switch feedback synchronization algorithm (SSP).

back RM-cell, the switch first marks its corresponding bit in the responsive branch vector $resp\_branch\_vec$ and then conducts RM-cell consolidations. If its modulo-2 addition with the connection pattern vector (which uses the negative logic): $conn\_patt\_vec \oplus resp\_branch\_vec$ equals $\underline{1}$, an all 1's vector—all feedback RM cells are synchronized, then a fully consolidated feedback RM cell is generated and sent upward; otherwise the switch needs to await other feedback RM cells for synchronization. Notice that since SSP allows feedback RM cells corresponding to different forward RM cells to be consolidated, the feedback RM cells are "softly synchronized" at branch nodes, making SSP not only scale well with the multicast-tree size/structure, but also readily detect/remove the nonresponsive branches [1].

C. Delay Modeling of Dependent-Marking Multicast Signaling

All of the above-referenced work only focused on the design and implementation of feedback synchronization signaling algorithms, without addressing their delay performances. To remedy this deficiency, we have previously developed a binary-tree deterministic model [1] and an *independent*-marking statistical model [1] to study the delay performance of various multicast feedback-synchronization signaling algorithms. The independent-marking statistical model [1] for multicast singling delay analysis builds on the recently proposed Random Early Marking (REM) [8], and the widely cited Random Early Detection (RED) [9] flow-control schemes.[2] The REM and RED—originally proposed for unicasts—can also be extended to multicast. Moreover, unicast and multicast transmissions usually co-exist in a network. In RED or REM, each router marks a packet's explicit congestion notification (ECN) bit with a probability that is exponential in REM, or proportional in RED, to the average queue length at the output link.

The independent-marking statistical model is suitable for signaling delay analysis for multicast flow control based on REM or RED, where link markings are assumed to be independent at different links/routers. However, there are also many cases where link markings may be dependent. In such a case, the independent-marking algorithm and analysis can only offer approximate results, and their performance and accuracy will be affected by the "dependency degree" between link markings. This paper addresses the general case of *dependent* link markings. Including dependence in the analysis is usually much more difficult than that with the independent-marking assumption.

In this paper, we develop a Markov-chain model over the link-marking/congestion states at different levels in a multicast tree, and a Markov-chain dependency-degree model which can capture all possible Markov-chain dependency degrees between different link-congestion markings. Using the Markov chain and Markov-chain dependency-degree models, we derive the probability distribution for a path to become the multicast-tree bottleneck path. We also derive the first and second moments of the multicast signaling delay for SSP and HBH, respectively.

[2]The analytical techniques developed in this paper are also applicable to cases where a link's random congestion state is caused by the flow-control schemes other than REM and RED.

D. Paper Organization

In Section II, we develop a Markov-chain model to derive the multicast-tree bottleneck probability. Section III proposes a Markov-chain dependency-degree model to compute transition probabilities. In Section IV, we derive the statistics of multicast signaling delays. Section V explores the asymptotical behavior of the link-marking Markov chain and its dependency-degree models. Section VI numerically evaluates and compares the delay performances of SSP and HBH. The paper concludes with Section VII.

II. Markov-Chain Model for Dependent Markings

In random-marking schemes like REM/RED, and any other flow-control schemes, the marking/congestion state of a link is a function of its queue length. However, the queue lengths of different links carrying the same flows are generally not independent of each other. For instance, if a large, or small, queue is built up at a congested upstream link in a multicast tree, the downstream links carrying the same flows are more likely to have large, or small, queues. For multicast flow control with *dependent* marking probabilities, we develop a Markov-chain model and a Markov-chain dependency-degree model for measuring and evaluating the degree of the Markov-chain dependency, in order to study the various statistical characteristics of multicast feedback-synchronization delay.

A. Binary Multicast-Tree-Based Dependent Statistical Model

A.1. Unbalanced Binary Multicast-Tree Signaling Model: To simplify the analysis of RM-cell RTTs, we measure/quantify each multicast-path's RTT based on the path length, i.e., number of hops along that path [1] by assuming that each switch-hop has a uniform delay, which only includes the propagation delays. We can only consider the propagation delay, excluding the queueing delay, because the flow-control signaling messages can tolerate neither latency nor error, and all network/flow-control signaling messages are typically delivered via an *out-of-band* channel,[3] bypassing buffering and queueing required for the normal data packets. The uniform hop-length assumption can be easily relaxed because the difference in the link-lengths in different switch-hops can be translated into different numbers of switch-hops, each with the same length.

In a real network, the structure of a multicast tree can be arbitrary, ranging from a balanced tree to an unbalanced tree. This paper aims at the modeling and analysis of signaling feedback *delay* and *delay variations* caused by dynamically changing location of the bottleneck path within a multicast tree, and the unbalanced multicast tree defined in this paper yields the largest

[3]For instance, in TCP the urgent-data packets are transmitted through "out-of-band" notification by bypassing buffering and queueing required for ordinary/normal data packets [10]. Similarly, the ATM network's RM-cells, carrying the time-sensitive flow-control feedback information, are also given the highest priority to bypass queueing delay in the normal/ordinary data buffers [11].

delay variation, representing the worst case (lower bound of signaling delay performance) in terms of delay variations. Since the above multicast-signaling delay analysis is only based on the path length (the number of hops on each path in a tree), we only need to consider the binary-tree and the delay modeling results based on the tree structure with the fan-out degree larger than 2 do not differ from those based on the binary-tree structure.

Our binary-tree model also considers the more realistic scenarios where multiple bottleneck links/paths can co-exist in a multicast tree. Moreover, to make the analysis complete, our binary-tree model allows the multicast-tree height m to be arbitrarily large and include ∞ as its limiting case. Allowing $m \to \infty$ enables us to study the scalability of signaling delay when the multicast tree becomes large, prove the normalization condition of bottleneck-path distribution, and study the various asymptotic properties of the proposed Markov-chain model. In addition, the proposed modeling techniques based on the unbalanced binary-tree model are generic, and thus, can be applied to any other multicast-tree structures.

A.2. Link Congestion Marking Dependency: While the congestion link-marking states can have either upstream or downstream dependency, the formulation of the proposed Markov-chain model (see the proof of Theorem 1 in Appendix A) only needs to consider one-way dependency to derive the probability distribution of a path becoming the dominant bottleneck path. We choose the upstream-dependency—i.e., congestion state in the downstream link depends only on its upstream-links' congestion states—to capture the congestion dependency caused by the multicast connection itself, which is clearly the major contributor to congestion dependency within the multicast tree. The link-congestion dependency by other cross-traffic can be taken into account by properly choosing/tuning the values of the dependency factors, α_i, $\forall i$.[4] We choose upstream, instead of downstream, dependency also based on the following observations, facts, and the simulation results described in Section II-A.3.

O1. Congestion caused by a multicast connection always originates and then propagates from the multicast source to the receivers at downstream, rather than from the downstream receivers to the multicast source at upstream. This implies that time-sequence-wise, the multicast congestion always occurs at the upstream link first if any, and then the congestion state of the same multicast flow propagates to the downstream links.

O2. Upward propagation of the congestion state can occur in the network using the hop-by-hop flow-control scheme, such as the back-pressure [12, pp. 506–507] and others [13]. By the backpressure scheme, for instance, the congestion state propagates backward from the congested node to the multicast sources through the back-pressure signal. However, our multicast delay modeling mainly targets at the networks using the end-to-end flow control, such as the Internet, and thus does not consider downstream dependency.

O3. For the multicast networks using the end-to-end flow-control, if an upstream link is marked congested, i.e., its arrival rate is larger than its available bandwidth, then the immediate-next downstream link carrying the same flows will also likely (with a larger probability) be congested, as long as the immediate-next upstream link gets more available bandwidth/service rate for these connections such that the large arrival rate/traffic will go through, or overflood the immediate-next downstream node/link, thus propagating the congestion state to the downstream nodes/links.

[4]The definition of α_i will be specified by Claim 2 in Theorem 2.

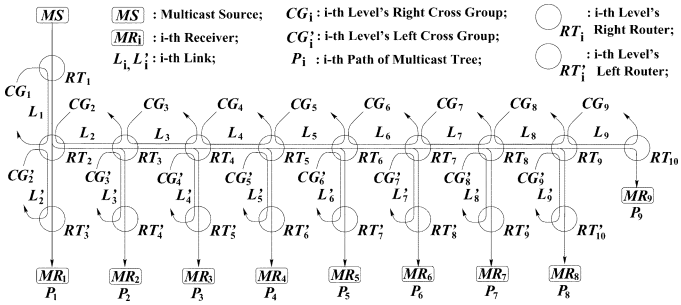


Fig. 2. Simulation model for link-marking dependency in a multicast tree with tree height $m = 10$; link-marking probability $\Pr\{X_i = 1\} = p_i = 1 - e^{-\xi \overline{q}_i}$ ($\overline{q}_i$ is average queue size and $\xi = 0.02$); link delay of L_i or $L'_i = 7$ ms, $\forall i$.

TABLE I
SIMULATED LINK-MARKING DEPENDENCY CAUSED BY THE MULTICAST CONNECTION

μ_7	μ_8	p_7	p_8	$p_7 \times p_8$	$p_{7,8}(1,1)$	α_7	$p_7 \geq p_8$
40	70	0.2092	0.1308	0.0274	0.0360	0.0835	Yes
40	60	0.2019	0.1949	0.0394	0.0575	0.1167	Yes
40	50	0.2147	0.2121	0.0455	0.0658	0.1216	Yes
40	40	0.2211	0.1209	0.0267	0.0397	0.1377	Yes
40	30	0.2323	0.3382	0.0786	0.1160	0.2435	No
40	20	0.1997	0.4531	0.0905	0.1207	0.2767	No
40	10	0.2342	0.5063	0.1186	0.1417	0.2000	No

A.3. Simulation Experiments on the Upstream Link-Marking Dependency: To further verify the *upstream-dependency* of link-marking states in the multicast tree, using ns-2 [14] we simulated the probabilities $p_7 = \Pr\{X_7 = 1\}$ and $p_8 = \Pr\{X_8 = 1\}$ of link L_7 and link L_8 being marked, respectively, and the joint probability $p_{7,8}(1,1) \triangleq \Pr\{X_7 = 1, X_8 = 1\}$ of links L_7 and L_8 both being marked, in the unbalanced binary multicast-tree network model as shown in Fig. 2. The simulation methods and parameters are detailed in [15]. The simulated results are listed in Table I where μ_7 and μ_8 are the bottleneck bandwidths (in Mbps) at links L_7 and L_8, respectively, and α_7 is the the Markov-chain dependency degree factor of event $\{X_8 = 1\}$ (link L_8 is marked) on event $\{X_7 = 1\}$ (link L_7 is marked). We observe that

$$\begin{aligned} p_{7,8}(1,1) &= \Pr\{X_8 = 1, X_7 = 1\} \\ &= \Pr\{X_8 = 1 \,|\, X_7 = 1\} \Pr\{X_7 = 1\} \\ &\neq \Pr\{X_8 = 1\} \Pr\{X_7 = 1\} = p_7 p_8 \end{aligned} \tag{1}$$

implying that events $\{X_7 = 1\}$ and $\{X_8 = 1\}$ are *not* independent. Furthermore, we also observe that

$$\begin{aligned} p_{7,8}(1,1) &= \Pr\{X_8 = 1, X_7 = 1\} \\ &= \Pr\{X_7 = 1\} \Pr\{X_8 = 1 \,|\, X_7 = 1\} \\ &> \Pr\{X_7 = 1\} \Pr\{X_8 = 1\} = p_7 p_8 \end{aligned} \tag{2}$$

always holds for all (μ_7, μ_8) combinations regardless whether $p_7 \leq p_8$ (link L_8 becomes the bottleneck link) or $p_7 > p_8$ (link L_7 becomes the bottleneck link), implying that

$$\Pr\{X_8 = 1 \,|\, X_7 = 1\} > \Pr\{X_8 = 1\} \tag{3}$$

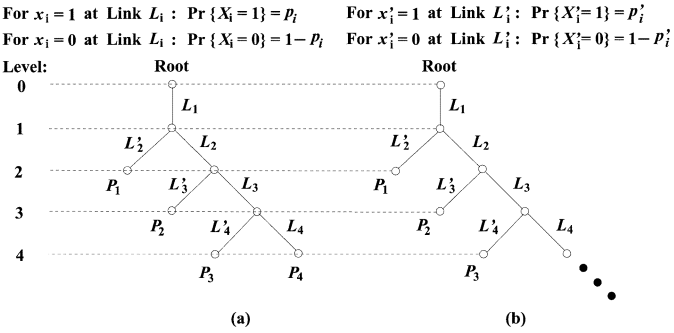


Fig. 3. Dependent random-marking unbalanced binary multicast-tree model. (a) Unbalanced tree: height $m = 4$. (b) Unbalanced tree: height $m = \infty$.

which represents the positive[5] Markov-chain dependency of the downstream link-marking state $\{X_8 = 1\}$ on its immediate upstream marking state $\{X_7 = 1\}$, because given the upstream link L_7 is marked, the downstream link L_8 is more likely to be marked than the case where $\{X_7 = 1\}$ is not given. So, the downstream link-marking state $\{X_8 = 1\}$ depends statistically on its immediate upstream marking state $\{X_7 = 1\}$, regardless whether the upstream link (L_7) or the downstream link (L_8) is the bottleneck link.

A.4. Markov-Chain Model for Link-Marking States: Based on the above observations/analysis, we can construct the following Markov-chain model over link-marking states in an unbalanced binary tree (see Fig. 3) to analyze the multicast signaling delay with dependent marking probabilities.

Definition 1: Markov-Chain Model for Link-Marking States: A *dependent* random-marking unbalanced binary-tree of height m consists of a set, $\mathcal{L}$, of links which satisfy the following conditions:

C1. All links in $\mathcal{L}$ are so labeled as in Fig. 3(a) and (b) for $m < \infty$ and for $m = \infty$, respectively, that

$$\mathcal{L} \triangleq \begin{cases} \{L_1, L'_2, L_2, L'_3, L_3, \ldots, L'_m, L_m\}, & \text{if } m < \infty \\ \{L_1, L'_2, L_2, L'_3, L_3, \ldots, L'_\infty, L_\infty\}, & \text{if } m = \infty. \end{cases} \tag{4}$$

The link set $\mathcal{L}$ contains m paths, $P_1, \ldots, P_m$, each represented by its component links as

$$\begin{cases} P_k \triangleq \{L_1, L_2, \ldots, L_k, L'_{k+1}\}, & \text{if } 1 \le k \le m-1 \\ P_m \triangleq \{L_1, L_2, \ldots, L_m\}, & \text{if } k = m. \end{cases} \tag{5}$$

Define P_m as the *main-stream path* taking only right branches at branch nodes, and define each P_k, for $1 \le k \le m-1$, as a *branch-stream path* consisting of k right branches and one left branch at the last branch node (see Fig. 3). Links L_i and $L'_i, \forall i \ge 2$, are at the same tree level.

C2. The marking state of link L_i, $(i = 1, 2, \ldots)$,[6] is represented by a random variable X_i which takes value in $\{0, 1\}$ such that (see the top part of Fig. 3)

$$\Pr\{X_i = x_i\} = \begin{cases} p_i, & \text{for } x_i = 1 \\ 1 - p_i, & \text{for } x_i = 0 \end{cases} \tag{6}$$

where $0 < p_i < 1$ is L_i's marking probability, which measures the traffic-load level on L_i. Thus, we use "marking probability" or "traffic load" interchangeably for p_i.

[5]Definition 3 details the positive/negative dependency of link-marking states.

[6]The marking state on L'_i is defined in the same way by using X'_i, x'_i, and p'_i.

C3. The congestion marking states at all links are *dependent*, and satisfy the following Markovian properties:

$$\begin{aligned} &\Pr\{X_i = x_i \,|\, X_{i-1} = x_{i-1}, X'_{i-1} = x'_{i-1}, X_{i-2} = x_{i-2}, \\ &\qquad X'_{i-2} = x'_{i-2}, \ldots, X_1 = x_1\} \\ &\quad = \Pr\{X_i = x_i \,|\, X_{i-1} = x_{i-1}\} \end{aligned} \tag{7}$$

$$\begin{aligned} &\Pr\{X'_i = x'_i \,|\, X_{i-1} = x_{i-1}, X'_{i-1} = x'_{i-1}, X_{i-2} = x_{i-2}, \\ &\qquad X'_{i-2} = x'_{i-2}, \ldots, X_1 = x_1\} \\ &\quad = \Pr\{X'_i = x'_i \,|\, X_{i-1} = x_{i-1}\}. \end{aligned} \tag{8}$$

C4. The congestion marking states within the same level are also *dependent* and satisfy the following properties:

$$\begin{aligned} &\Pr\{X_i = x_i \,|\, X'_i = x'_i, X_{i-1} = x_{i-1}, X'_{i-1} = x'_{i-1}, \\ &\qquad X_{i-2} = x_{i-2}, X'_{i-2} = x'_{i-2}, \ldots, X_1 = x_1\} \\ &\quad = \Pr\{X_i = x_i \,|\, X_{i-1} = x_{i-1}\} \end{aligned} \tag{9}$$

$$\begin{aligned} &\Pr\{X'_i = x'_i \,|\, X_i = x_i, X_{i-1} = x_{i-1}, X'_{i-1} = x'_{i-1}, \\ &\qquad X_{i-2} = x_{i-2}, X'_{i-2} = x'_{i-2}, \ldots, X_1 = x_1\} \\ &\quad = \Pr\{X'_i = x'_i \,|\, X_{i-1} = x_{i-1}\} \end{aligned} \tag{10}$$

where the random variables X_i and $X'_i \in \{0, 1\}, \forall i$. ■

*Remarks on Definition 1 (**C3** and **C4**):* While the congestion state of a link depends on *all its upstream*[7] and same-level links' congestion states, our Markov-chain formulation/modeling ensures that this link's congestion state is *conditionally independent* of all its upstream and same-level links' congestion states as long as its immediate-next upstream link congestion state is given [see (7)–(10)]. This is because the congestion information on the links above the immediate-next upstream link (see **C3**) is all concentrated into, and carried over by, the given congestion information on the immediate-next upstream link. **C3** and **C4** are reasonable since one link's congestion state depends *most* on its immediate-next upstream link's congestion state in the same multicast tree. The upstream's influence on a downstream link's congestion state propagates through its immediate-next upstream link which carries the same flows.

B. Probability Distribution of the Dominant Bottleneck Path

B.1. Bottleneck Path and Dominant Bottleneck Path: For reliable data transmission, unicast ABR service adjusts its sending rate by the feedback from the *most* congested link having the minimum *available* bandwidth on the path [16]. A natural extension of this strategy to multicast ABR service is to adjust the source rate to the minimum available bandwidth share on the *most* congested path across the multicast tree. Clearly, by OR rules (see Fig. 1), the *shortest* bottleneck path in a multicast tree governs source's flow-control decisions and feedback RTT. To explicitly model these features, we define the following.

Definition 2: 1) The *bottleneck path* is the multicast path containing at least one congestion-marked link whose congestion feedback received at the source dictates the source flow control (if it is the only bottleneck path). 2) The *bottleneck path RM-cell RTT* is the RM-cell RTT experienced on this *bottleneck path*. 3) Among all bottleneck paths concurrently existing in a multicast tree, the bottleneck path with the *minimum* length is called the *dominant bottleneck path* or *multicast-tree bottleneck path*, and its RM-cell RTT is called the *multicast-tree bottleneck RM-cell RTT* or *multicast-tree RTT*. ■

[7]As mentioned in Section II-A.2, the proposed Markov-chain model only considers the *upstream* dependency (see **C3**).

B.2. Probability Distribution of the Dominant Bottleneck Path: Based on Definition 1, Proposition 1 lays a foundation for the dominant bottleneck path probability distributions derived in Theorem 1, which are described as follows.

Proposition 1: The sequence of random link-marking states $\{X_1,\ldots,X_m\}$ (for tree height $m<\infty$ and $m\to\infty$) in Definition 1 defines a two-state discrete-indexed Markov chain over the links on the main-stream path $P_m=\{L_1,\ldots,L_m\}$, and each sequence of link-marking states $\{X_1,\ldots,X_k,X'_{k+1}\}$ in Definition 1 on each branch-stream path $P_k=\{L_1,\ldots,L_k,L'_{k+1}\}$, for $k=1,\ldots,m-1$, also defines a two-state Markov chain.

Proof: The proof follows from **C3** of Definition 1 and notice that unlike the traditional Markov chain indexed by time, we define the Markov chain over each multicast path and index it by the link sequence numbers associated with that path. ■

Theorem 1: If a *dependent*-marking multicast tree of height m defined by Definition 1 is flow-controlled under SSP or HBH and its link-marking states are not always *perfectly dependent* (see Claim 2 in Theorem 2), then the following claims hold.

Claim 1. If $m\to\infty$, there is *one and only one* dominant bottleneck path, and the probability, $\psi_d(P_k,\infty)$, of P_k becoming the dominant bottleneck path, is given by

$$\psi_d(P_k,\infty)=\begin{cases}1-\Pr\{X_1=0\}\Pr\{X'_2=0\mid X_1=0\}, & \text{if } k=1\\ \Pr\{X_1=0\}\Pr\{X'_k=0\mid X_{k-1}=0\}\\ \quad\cdot\Big[\Pr\{X_k=1\mid X_{k-1}=0\}+\Pr\{X_k=0\mid X_{k-1}=0\}\\ \quad\cdot\Pr\{X'_{k+1}=1\mid X_k=0\}\Big]\prod_{i=1}^{k-2}\Big\{\Pr\{X_{i+1}=0\mid X_i=0\}\\ \quad\cdot\Pr\{X'_{i+1}=0\mid X_i=0\}\Big\}, & \text{if } k\ge 2\end{cases}\tag{11}$$

which satisfies

$$\lim_{m\to\infty}\sum_{k=1}^{m}\psi_d(P_k,\infty)=1.\tag{12}$$

Claim 2. If $m<\infty$, then there exists *at most one* dominant bottleneck path, and the probability, $\psi_d(P_k,m)$, that P_k becomes the dominant bottleneck path, is given by

$$\psi_d(P_k,m)=\begin{cases}1-\Pr\{X_1=0\}\Pr\{X'_2=0\mid X_1=0\}, & \text{if } k=1\\ \Pr\{X_1=0\}\Pr\{X'_k=0\mid X_{k-1}=0\}\\ \quad\cdot\Big[\Pr\{X_k=1\mid X_{k-1}=0\}+\Pr\{X_k=0\mid X_{k-1}=0\}\\ \quad\cdot\Pr\{X'_{k+1}=1\mid X_k=0\}\Big]\prod_{i=1}^{k-2}\Big\{\Pr\{X_{i+1}=0\mid X_i=0\}\\ \quad\cdot\Pr\{X'_{i+1}=0\mid X_i=0\}\Big\}, & \text{if } k\ge 2\\ \Pr\{X_1=0\}\Pr\{X_m=1\mid X_{m-1}=0\}\\ \quad\cdot\Pr\{X'_m=0\mid X_{m-1}=0\}\prod_{i=1}^{m-2}\Big\{\Pr\{X_{i+1}=0\mid X_i=0\}\\ \quad\cdot\Pr\{X'_{i+1}=0\mid X_i=0\}\Big\}, & \text{if } k=m.\end{cases}\tag{13}$$

Proof: The proof is provided in Appendix A. ■

Remarks on Theorem 1: Equation (11) implies that $\lim_{k\to\infty}\psi_d(P_k,\infty)=0$. This is expected, since a longer bottleneck path is always dominated by a co-existing shorter bottleneck path, if any. Thus, when $k\to\infty$ as $m\to\infty$, P_∞ is always dominated by a shorter bottleneck path for $0<p_i,p'_i<1, i=1,\ldots,\infty$, i.e., $\psi_d(P_\infty,\infty)=0$. Moreover, by (12) we have $\sum_{k=1}^{\infty}\psi_d(P_k,\infty)=1$, which also makes sense because as tree height $m\to\infty$ and $0<p_i,p'_i<1$, there always exists (with probability 1) one and only one dominant bottleneck path in a multicast tree. For $m<\infty$, by (13) and (12) we have $\sum_{k=1}^{m}\psi_d(P_k,m)\le 1$, implying the possibility of no dominant bottleneck path in the multicast tree with $m<\infty$. This is also expected because $0<p_i,p'_i<1$.

III. Model of Markov-Chain Dependency Degree

In order to use (11) and (13), we need to derive explicit expressions for $\Pr\{X_i=x_i\,|\,X_{i-1}=x_{i-1}\}$ and $\Pr\{X'_i=x'_i\,|\,X_{i-1}=x_{i-1}\}$, the fundamental conditional distributions used in (11) and (13). However, it is difficult to know/compute the accurate dependency between two random variables. To solve this problem, we propose to use a real-valued Markov-chain *dependency-degree factor* $\alpha\in[0,1]$ to quantify all possible degrees of dependency between the random variables in the Markov chain's one-step transition probabilities. Using this dependency-degree factor, one can evaluate the degree of Markov chain's dependency ranging from *independent* to *perfectly dependent*,[8] without knowing *a priori* the dependency degree of the two random variables.

Two dependent random events can affect each other either *positively* or *negatively*. If occurrence of one event is likely to trigger another, they are said to be *positively dependent*; if occurrence of one event makes another event unlikely to occur, they are said to be *negatively dependent*. As discussed above, an upstream link's congested (uncongested) state makes the downstream links carrying the same flows more likely (unlikely) congested. So, the positive dependence can accurately characterize the marking dependence. Thus, we define the following.

Definition 3: Two dependent link-marking states X_i and X_{i+1} are said to be positively (negatively) dependent if $\Pr\{X_{i+1}=x\,|\,X_i=x\}>\Pr\{X_{i+1}=x\,|\,X_i=\overline{x}\}$ ($\Pr\{X_{i+1}=x\,|\,X_i=x\}<\Pr\{X_{i+1}=x\,|\,X_i=\overline{x}\}$), where $x\in\{0,1\}$. ■

Using Definition 3, the theorem given below models the dependency-degree between the random variables of the Markov chain. Notice that this theorem only gives the results for the case of $\Pr\{X_{i+1}=x_{i+1}\,|\,X_i=x_i\}$ and $\Pr\{X_i=1\}=p_i$, and it also holds for the case of $\Pr\{X'_{i+1}=x'_{i+1}\,|\,X_i=x_i\}$ and $\Pr\{X'_i=1\}=p'_i$ with the similar results that we omitted.

Theorem 2: Consider the Markov chain $\{X_i\}$ defined on link-marking states on every path (for both main-stream and branch-stream) in the multicast tree specified by Definition 1. If $\{X_i\}$ is *positively dependent*, and the link marking-probability equals $\Pr\{X_i=1\}=p_i$, then the following claims hold.

[8]The typical examples of *perfectly dependent* events are (1) two events are *identical*—where the occurrence of either event implies the occurrence of the other event; and (2) one event is a *sub-event* of the other— if the first event occurs then the other event which subsumes the first event as a sub-event also occurs, and thus "perfectly depends" on the first event.

Claim 1. $\Pr\{X_{i+1} = x_{i+1} \mid X_i = x_i\}$, with $x_i, x_{i+1} \in \{0,1\}$, is upper- and lower-bounded by

$$1 - p_{i+1} \le \Pr\{X_{i+1} = 0 \mid X_i = 0\} \le \begin{cases} 1, & \text{if } p_i \ge p_{i+1} \\ (1-p_{i+1})/(1-p_i), & \text{if } p_i < p_{i+1} \end{cases} \quad (14)$$

$$p_{i+1} \ge \Pr\{X_{i+1}=1 \mid X_i=0\} \ge \begin{cases} 0, & \text{if } p_i \ge p_{i+1} \\ (p_{i+1}-p_i)/(1-p_i), & \text{if } p_i < p_{i+1} \end{cases} \quad (15)$$

$$1-p_{i+1} \ge \Pr\{X_{i+1}=0 \mid X_i=1\} \ge \begin{cases} (p_i-p_{i+1})/p_i, & \text{if } p_i \ge p_{i+1} \\ 0, & \text{if } p_i < p_{i+1} \end{cases} \quad (16)$$

$$p_{i+1} \le \Pr\{X_{i+1}=1 \mid X_i=1\} \le \begin{cases} p_{i+1}/p_i, & \text{if } p_i \ge p_{i+1} \\ 1, & \text{if } p_i < p_{i+1}. \end{cases} \quad (17)$$

Claim 2. $\exists\, \alpha_i\ (\alpha_i') \in [0,1]$ such that all possible dependency degrees between X_i and X_{i+1} (X_{i+1}') can be measured by the real-valued Markov-chain *dependency-degree factor*[9] α_i (α_i'), and

$$\begin{cases} \alpha_i = 0, & \textbf{iff } X_i \text{ and } X_{i+1} \text{ are independent} \\ \alpha_i = 1, & \textbf{iff } X_i \text{ and } X_{i+1} \text{ are } \textbf{perfectly} \text{ dependent} \end{cases} \quad (18)$$

$$\begin{cases} \alpha_i' = 0, & \textbf{iff } X_i \text{ and } X_{i+1}' \text{ are independent} \\ \alpha_i' = 1, & \textbf{iff } X_i \text{ and } X_{i+1}' \text{ are } \textbf{perfectly} \text{ dependent.} \end{cases} \quad (19)$$

Claim 3. $\Pr\{X_{i+1} = x_{i+1} \mid X_i = x_i\}$, with $x_i, x_{i+1} \in \{0,1\}$, are determined by

$$\Pr\{X_{i+1}=0 \mid X_i=0\} = \begin{cases} 1-(1-\alpha_i)\,p_{i+1}, & \text{if } p_i \ge p_{i+1} \\ (1-\alpha_i)(1-p_{i+1}) + \dfrac{\alpha_i(1-p_{i+1})}{1-p_i}, & \text{if } p_i < p_{i+1} \end{cases} \quad (20)$$

$$\Pr\{X_{i+1}=1 \mid X_i=0\} = \begin{cases} (1-\alpha_i)\,p_{i+1}, & \text{if } p_i \ge p_{i+1} \\ (1-\alpha_i)\,p_{i+1} + \dfrac{\alpha_i(p_{i+1}-p_i)}{1-p_i}, & \text{if } p_i < p_{i+1} \end{cases} \quad (21)$$

$$\Pr\{X_{i+1}=0 \mid X_i=1\} = \begin{cases} (1-\alpha_i)(1-p_{i+1}) + \dfrac{\alpha_i(p_i-p_{i+1})}{p_i}, & \text{if } p_i \ge p_{i+1} \\ (1-\alpha_i)(1-p_{i+1}), & \text{if } p_i < p_{i+1} \end{cases} \quad (22)$$

$$\Pr\{X_{i+1}=1 \mid X_i=1\} = \begin{cases} (1-\alpha_i)p_{i+1} + \dfrac{\alpha_i p_{i+1}}{p_i}, & \text{if } p_i \ge p_{i+1} \\ p_{i+1} + \alpha_i(1-p_{i+1}), & \text{if } p_i < p_{i+1} \end{cases} \quad (23)$$

where $i = 1, 2, \ldots; \alpha_i$ is defined in Claim 2 of Theorem 2.

Proof: The proof is detailed in Appendix B. ■

Remarks on Theorem 2: Claim 1 finds the upper and lower bounds of all four possible two-state Markov chain one-step transition probabilities as functions of the marginal link-marking probabilities p_i and p_{i+1} specified by networks. Claim 2 ensures the existence of a real-valued dependence-degree factor $\alpha_i \in [0,1]$. It also proves the completeness of the Markov-chain dependence-degree factor by mapping all possible dependency degrees onto the real-valued interval [0, 1]. Claim 3 derives all four possible two-state Markov-chain one-step conditional transition probabilities as the functions of their marginal distributions.

Applying Theorem 2 and (20)–(23) to Theorem 1, Corollary 1 derives the multicast-tree bottleneck path probabilities.

Corollary 1: Let a dependent-marking multicast tree of height m given by Definition 1 be flow-controlled under SSP or HBH. If the one-step transition probability of Markov chain $\{X_i\}$ defined over each path (main- or branch-stream path) is specified by the dependency-factor vector $\vec{\alpha} \triangleq (\alpha_1, \alpha_1', \alpha_2, \alpha_2', \alpha_3, \alpha_3', \ldots)$ which is derived in Theorem 2, and further, denote the link-marking probability vector by $\vec{p} \triangleq (p_1, p_1', p_2, p_2', p_3, p_3', \ldots)$, respectively, then the following claims hold.

Claim 1. If $m \to \infty$, then there exists *one and only one* dominant bottleneck path, and the probability, $\psi_d(P_k, \vec{\alpha}, \vec{p}, \infty)$, that P_k becomes the dominant bottleneck path, is given by

$$\psi_d(P_k, \vec{\alpha}, \vec{p}, \infty) = \begin{cases} 1-(1-p_1)\left[1-(1-\alpha_1')p_2'\right], & \text{if } k=1 \\ (1-p_1)\left[1-(1-\alpha_{k-1}')p_k'\right]\Big[(1-\alpha_{k-1})p_k \\ \quad +\left[1-(1-\alpha_{k-1})p_k\right](1-\alpha_k')p_{k+1}'\Big] \\ \quad \cdot\displaystyle\prod_{i=1}^{k-2}\Big\{\left[1-(1-\alpha_i)p_{i+1}\right]\left[1-(1-\alpha_i')p_{i+1}'\right]\Big\}, & \text{if } k \ge 2 \end{cases} \quad (24)$$

which satisfies

$$\lim_{m\to\infty} \sum_{k=1}^{m} \psi_d(P_k, \vec{\alpha}, \vec{p}, \infty) = 1. \quad (25)$$

Claim 2. If $m < \infty$, then there exists *at most one* dominant bottleneck path, and the probability, $\psi_d(P_k, \vec{\alpha}, \vec{p}, m)$, that P_k becomes the dominant bottleneck path, is given by

$$\psi_d(P_k, \vec{\alpha}, \vec{p}, m) = \begin{cases} 1-(1-p_1)\left[1-(1-\alpha_1')p_2'\right], & \text{if } k=1 \\ (1-p_1)\left[1-(1-\alpha_{k-1}')p_k'\right]\Big[(1-\alpha_{k-1})p_k \\ \quad +\left[1-(1-\alpha_{k-1})p_k\right](1-\alpha_k')p_{k+1}'\Big] \\ \quad \cdot\displaystyle\prod_{i=1}^{k-2}\Big\{\left[1-(1-\alpha_i)p_{i+1}\right]\left[1-(1-\alpha_i')p_{i+1}'\right]\Big\}, & \text{if } k \ge 2 \\ (1-p_1)(1-\alpha_{m-1})p_m\left[1-(1-\alpha_{m-1}')p_m'\right] \\ \quad \cdot\displaystyle\prod_{i=1}^{m-2}\Big\{\left[1-(1-\alpha_i)p_{i+1}\right]\left[1-(1-\alpha_i')p_{i+1}'\right]\Big\}, & \text{if } k = m. \end{cases} \quad (26)$$

[9]The actual value of dependency-degree factor (α) can be determined through either simulations or experiments.

Proof: The proof follows by plugging (20)–(23) of Theorem 2 into (11)–(13) of Theorem 1. ■

Remarks on Corollary 1: We can use (24) and (26), and tune up the dependence-degree factor $\vec{\alpha}$ to see how the system performs with different $\vec{\alpha}$'s. More importantly, the completeness of this approach guarantees that the actual unknown Markov-chain dependency degree imposed by the practical problems can always be covered by tuning $\alpha_i \in [0,1]$, $\forall i$ for heterogeneity.

IV. Statistics of Multicast Signaling Delays

Using Corollary 1 and the multicast signaling delay expression $\tau_u(j,\Delta)$ derived in [1] on each path, Theorem 3 derives the probabilities, their properties, and the means and variances of multicast signaling delays under SSP and HBH, respectively.

Theorem 3: Let a dependent-marking multicast tree of height m as defined in Definition 1 be flow controlled under SSP and HBH, respectively, with the RM-cell interval Δ. If $m<\infty, 0<p_i=p_i'=p<1$ and $0\le\alpha_i=\alpha_i'=\alpha\le 1, \forall i$ (the homogeneous case), then the following claims hold.

Claim 1. The probability for P_k to become the dominant bottleneck path, $\psi_d(P_k,\alpha,p,m)$, is given by

$$\psi_d(P_k,\alpha,p,m)= \begin{cases} 1-(1-p)\left[1-(1-\alpha)p\right], & \text{if } k=1 \\ (1-\alpha)(1-p)p\left[2-(1-\alpha)p\right] \\ \qquad \cdot\left[1-(1-\alpha)p\right]^{2k-3}, & \text{if } k\ge 2 \\ (1-\alpha)(1-p)p\left[1-(1-\alpha)p\right]^{2m-3}, & \text{if } k=m \end{cases} \tag{27}$$

Claim 2. For each path P_k and a given α, $\psi_d(P_k,\alpha,p,m)$ attains the *unique* maximum at

$$p^* \triangleq \arg\max_{0<p<1}\psi_d(P_k,\alpha,p,m)= \begin{cases} 1, & \text{if } k=1 \\ [(1-\alpha)(2m-1)]^{-1}\left\{m-(m-1)\alpha \right. \\ \quad \left. -\sqrt{[m-(m-1)\alpha]^2-(1-\alpha)(2m-1)}\right\}, & \text{if } k=m \\ \gamma_1\sqrt[3]{-\omega/2+\sqrt{(\omega/2)^2+(\upsilon/3)^3}} \\ \quad +\gamma_2\sqrt[3]{-\omega/2-\sqrt{(\omega/2)^2+(\upsilon/3)^3}} \\ \quad -\left[(2k-1)\alpha-6k\right]/\left[6k(1-\alpha)\right], & \text{if } 2\le k\le m-1 \end{cases} \tag{28}$$

where $\gamma_1=(-1+i\sqrt{3})/2$ and $\gamma_2=(-1-i\sqrt{3})/2$ are complex numbers with $i^2=-1$, and

$$\begin{cases} \upsilon=\left[1+2k-(2k-1)\alpha\right]/[k(1-\alpha)^2] \\ \qquad -\left[(2k-1)\alpha-6k\right]^2/[12k^2(1-\alpha)^2] \\ \omega=\left[(2k-1)\alpha-6k\right]^3/[72k^3(1-\alpha)^3] \\ \qquad +\left[(2k-1)\alpha-2k-1\right]\left[(2k-1)\alpha-6k\right] \\ \qquad \cdot[6k^2(1-\alpha)^3]^{-1}-[k(1-\alpha)^2]^{-1} \\ \qquad -\left[(2k-1)\alpha-6k\right]^3/[6k(1-\alpha)]^3. \end{cases} \tag{29}$$

Claim 3. For each path P_k and a given p, $\psi_d(P_k,\alpha,p,m)$ attains the *unique* maximum at

$$\alpha^* \triangleq \arg\max_{0<\alpha<1}\psi_d(P_k,\alpha,p,m)= \begin{cases} \dfrac{p-1}{p}+\dfrac{1}{p}\sqrt{1-\dfrac{2}{2k-1}}, & \text{if } 2\le k\le m-1 \\ & \text{and } k\ge\lceil 1/2+1/[p(2-p)]\rceil \\ 1-\dfrac{1}{2(m-1)p}, & \text{if } k=m \text{ and } k\ge\lceil 1+1/(2p)\rceil. \end{cases} \tag{30}$$

Claim 4. If $\alpha=\alpha_0>0$ for a given α_0, it shifts the probability of multicast-tree bottleneck path from shorter to longer paths. If the multicast-tree height m satisfies

$$m\ge\left\lfloor\frac{\log\sqrt{1/(1-\alpha_0)}}{\log\{[1-(1-\alpha_0)p]/(1-p)\}}+2.5\right\rfloor \tag{31}$$

then there exists the *unique dependency-balanced path* $P_{\widetilde{k}}$ such that $2\le\widetilde{k}\le m-1$ and

$$\begin{cases} \psi_d(P_k,\alpha,p,m)|_{\alpha=0}\ge\psi_d(P_k,\alpha,p,m)|_{\alpha=\alpha_0}, & \text{if } k\le\widetilde{k} \\ \psi_d(P_k,\alpha,p,m)|_{\alpha=0}<\psi_d(P_k,\alpha,p,m)|_{\alpha=\alpha_0}, & \text{if } k>\widetilde{k} \end{cases} \tag{32}$$

where $\psi_d(P_k,\alpha,p,m)$ is given by (27) and the *dependency-balanced path* number $\widetilde{k}$ is given by

$$\widetilde{k}=\left\lfloor\frac{\log\sqrt{(2-p)/\{(1-\alpha_0)[2-(1-\alpha_0)p]\}}}{\log\{[1-(1-\alpha_0)p]/(1-p)\}}+1.5\right\rfloor. \tag{33}$$

Claim 5. The means of multicast-tree bottleneck RM-cell RTT, denoted by $\overline{\tau}_{\mathrm{SSP}}(\alpha,p,m)$ and $\overline{\tau}_{\mathrm{HBH}}(\alpha,p,m)$ for the SSP and HBH schemes, respectively, are given by

$$\begin{aligned} \overline{\tau}_{\mathrm{SSP}}(\alpha,p,m)=&\left[p+(1-\alpha)(p-p^2)\right]\left\{2m-\left\lfloor\frac{2(m-2)}{\Delta}\right\rfloor\Delta\right\} \\ &+2m(1-p)\left[1-(1-\alpha)p\right]\left\{1+[1-(1-\alpha)p]^{2(m-2)}\right. \\ &\left.\cdot\left[(1-\alpha)p-1\right]\right\}-(1-\alpha)(1-p)p\left[2-(1-\alpha)p\right]\Delta \\ &\cdot\sum_{k=2}^{m-1}\left\{\left\lfloor\frac{2(m-k-1)}{\Delta}\right\rfloor\left[1-(1-\alpha)p\right]^{2k-3}\right\} \end{aligned} \tag{34}$$

$$\begin{aligned} \overline{\tau}_{\mathrm{HBH}}(\alpha,p,m)=&\frac{(1-p)\Theta(\Delta)}{(1-\alpha)p\left[2-(1-\alpha)p\right]}\left\{2\left[1-(1-\alpha)p\right]\right. \\ &-\left[1-(1-\alpha)p\right]^3-m\left[1-(1-\alpha)p\right]^{2m-3}+(m-1) \\ &\left.\cdot\left[1-(1-\alpha)p\right]^{2m-1}\right\}+(1-p)\left[1-(1-\alpha)p\right]^{2m-3} \\ &\cdot\left\{(1-\alpha)p\left[2+(m-1)\Theta(\Delta)\right]-2\right\}+(2+\Theta(\Delta)) \\ &\cdot\left[p+(1-\alpha)(p-p^2)\right]+2(1-p)\left[1-(1-\alpha)p\right] \end{aligned} \tag{35}$$

where $\Theta(\Delta)$ is defined by (2) in [1].

Claim 6. The variances of multicast-tree bottleneck RM-cell RTT, denoted by $\sigma^2_{\mathrm{SSP}}(\alpha,p,m)$ and $\sigma^2_{\mathrm{HBH}}(\alpha,p,m)$ for the SSP and HBH schemes, respectively, are given by

$$\begin{aligned}\sigma^2_{\mathrm{SSP}}(\alpha,p,m)=&4m^2+4m^2(1-p)\left[1-(1-\alpha)p\right]^{2m-3}\\&\cdot\left[(1-\alpha)p-1\right]-(1-\alpha)(1-p)p\left[2-(1-\alpha)p\right]\Bigg\{4m\Delta\\&\cdot\sum_{k=2}^{m-1}\left\{\left\lfloor\frac{2(m-k-1)}{\Delta}\right\rfloor\left[1-(1-\alpha)p\right]^{2k-3}\right\}-\Delta^2\\&\cdot\sum_{k=2}^{m-1}\left\{\left\lfloor\frac{2(m-k-1)}{\Delta}\right\rfloor^2\left[1-(1-\alpha)p\right]^{2k-3}\right\}\Bigg\}\\&+p\Big[1+(1-\alpha)(1-p)\Big]\Bigg\{\Delta^2\left\lfloor\frac{2(m-2)}{\Delta}\right\rfloor^2-4m\Delta\\&\cdot\left\lfloor\frac{2(m-2)}{\Delta}\right\rfloor\Bigg\}-\overline{\tau}^2_{\mathrm{SSP}}(\alpha,p,m)\end{aligned}\tag{36}$$

$$\begin{aligned}\sigma^2_{\mathrm{HBH}}(\alpha,p,m)=&\left[1+(1-\alpha)(1-p)\right]p[2+\Theta(\Delta)]^2\\&+(1-\alpha)(1-p)p\left[1-(1-\alpha)p\right]^{2m-3}\left[2+(m-1)\right.\\&\left.\cdot\,\Theta(\Delta)\right]^2+4(1-p)\left[1-(1-\alpha)p\right]\\&\cdot\left\{1-\left[1-(1-\alpha)p\right]^{2(m-2)}\right\}\\&+\frac{4(1-p)\left[1-(1-\alpha)p\right]\Theta(\Delta)}{(1-\alpha)p\left[2-(1-\alpha)p\right]}\Bigg\{2-\left[1-(1-\alpha)p\right]^2\\&-m\left[1-(1-\alpha)p\right]^{2(m-2)}+\left[1-(1-\alpha)p\right]^{2(m-1)}\\&\cdot(m-1)\Bigg\}+\frac{(1-p)\Theta^2(\Delta)}{(1-\alpha)^2\left[2-(1-\alpha)p\right]^2\left[p^2-(1-\alpha)p^3\right]}\\&\cdot\Bigg\{1+\left[1-(1-\alpha)p\right]^2-\left[2-(1-\alpha)p\right]^3\left[(1-\alpha)p\right]^3\\&-m^2\left[1-(1-\alpha)p\right]^{2(m-1)}+(2m^2-2m-1)\\&\cdot\left[1-(1-\alpha)p\right]^{2m}-(m-1)^2\\&\cdot\left[1-(1-\alpha)p\right]^{2(m+1)}\Bigg\}-\overline{\tau}^2_{\mathrm{HBH}}(\alpha,p,m)\end{aligned}\tag{37}$$

where $\Theta(\Delta)$ is given by (2) in [1], and $\overline{\tau}_{\mathrm{SSP}}(\alpha,p,m)$ and $\overline{\tau}_{\mathrm{HBH}}(\alpha,p,m)$ are given by (34) and (35), respectively.

Proof: The proof is available online in [15]. ■

Remarks on Theorem 3: Claim 1 derives multicast bottleneck-path probabilities in terms of k,p,α, and m. Claim 2 examines the dynamics of $\psi_d(P_k,\alpha,p,m)$ as p varies, where $\psi_d(P_k,\alpha,p,m)$ attains the unique maximum at p^* given by (28)—the link-marking probability making P_k the most likely multicast bottleneck path. Claim 3 shows that $\psi_d(P_k,\alpha,p,m)$ can be either monotonic or nonmonotonic in terms of α: if k and p satisfy the conditions specified in (30), $\psi_d(P_k,\alpha,p,m)$ achieves the maximum at α^* given by (30). Claim 4 reveals that the Markov-chain dependency ($\alpha>0$) reduces the probabilities for shorter paths, while increasing the probabilities for longer ones, to become the bottleneck. This probability shift is balanced at the unique path, $P_{\widetilde{k}}$, where $\psi_d(P_{\widetilde{k}},\alpha,p,m)|_{\alpha=0}=\psi_d(P_{\widetilde{k}},\alpha,p,m)|_{\alpha>0}$, if m is large enough. This claim also finds the condition for the existence and uniqueness of $P_{\widetilde{k}}$ and the equation for $\widetilde{k}$ in terms of α_0 and p. Claim 5 and Claim 6 derive the multicast-signaling delay means and variances, respectively, of SSP and HBH as functions of Δ,p,α, and m.

V. Asymptotical Analysis of the Markov Chain

Theorem 4 investigates the long-term behavior of the link-marking Markov chains based on the proposed Markov-chain dependency-degree model when m is large. For simplicity, the following asymptotical analysis and its numerical analysis will only focus on the homogeneous Markov chain. However, the analytical techniques/results developed here can also be applied to the heterogeneous Markov chain, where $(\alpha_1,\alpha_1',\alpha_2,\alpha_2',\ldots)$ differ from each other, and so do $(p_1,p_1',p_2,p_2',\ldots)$.

Theorem 4: Consider the Markov chain $\{X_i\}$ defined by the link-marking states on both main- and branch-stream paths in the multicast tree specified by Definition 1. If: i) the dependency degree of $\{X_i\}$ is specified by the dependency-degree factor vector $\vec{\alpha}=(\alpha_1,\alpha_1',\alpha_2,\alpha_2',\alpha_3,\alpha_3',\ldots)$ derived in Theorem 2; ii) the link-marking probability vector is specified by $\vec{p}=(p_1,p_1',p_2,p_2',p_3,p_3',\ldots)$ defined in Definition 1; and iii) $\vec{p}$ and $\vec{\alpha}$ satisfy $0<p_i=p_i'=p<1$ and $0\le\alpha_i=\alpha_i'=\alpha\le 1,\forall i$, respectively, such that $\{X_i\}$ becomes a homogeneous Markov chain, then the following claims hold.

Claim 1. The n-step transition probability matrix, $P^{(n)}$, of the homogeneous Markov chain $\{X_i\}$ is determined by

$$\begin{aligned}P^{(n)}\triangleq\left\{p_{jk}^{(n)}\right\}&\triangleq\left\{\Pr\left\{X_{r+n}=k\mid X_r=j\right\}\right\}\\&=\begin{bmatrix}1-(1-\alpha^n)p & (1-\alpha^n)p\\(1-\alpha^n)(1-p) & \alpha^n(1-p)+p\end{bmatrix}\end{aligned}\tag{38}$$

where $j,k\in\{0,1\},n\in\{0,1,\ldots\},\forall r\ge 1,\Pr\{X_{r+n}=k\mid X_r=j\}|_{(r=i,n=1)}$ are given by (20)–(23), and the Markov-chain model for $P^{(n)}|_{n=1}$ is shown in Fig. 4.

Claim 2. If $\alpha\in[0,1],\{X_i\}$'s states are *ergodic*, and also

$$\limsup_{n\to\infty}p_{jj}^{(n)}=\lim_{n\to\infty}p_{jj}^{(n)}>0;\quad\lim_{n\to\infty}\sum_{r=1}^{n}p_{jj}^{(r)}=\infty\tag{39}$$

where $j\in\{0,1\}$ and the recurring probability converges as

$$\lim_{n\to\infty}p_{jj}^{(n)}=\begin{cases}\Pr\{X_k=j\}=1-p, & \text{if } j=0,\ \alpha\in[0,1)\\ \Pr\{X_k=j\}=p, & \text{if } j=1,\ \alpha\in[0,1)\\ 1, & \text{if } j\in\{0,1\},\ \alpha=1.\end{cases}\tag{40}$$

Claim 3. If $\alpha\in[0,1),\{X_i\}$ is *ergodic* and its limiting probabilities converge to the marginal link-marking probabilities as follows:

$$\lim_{n\to\infty}\vec{p}(n)=\vec{\pi}\triangleq\begin{bmatrix}\pi_0 & \pi_1\end{bmatrix}=\begin{bmatrix}(1-p) & p\end{bmatrix}.\tag{41}$$

Claim 4. If $\alpha=1,\{X_i\}$ converges to an equilibrium state where the equilibrium probabilities are not unique, but equal to the initial state probabilities $\begin{bmatrix}p_0(1) & p_1(1)\end{bmatrix}$.

Proof: The proof is provided in Appendix C. ■

Remarks on Theorem 4: Claim 1 specifies the long-term behavior of Markov chain $\{X_i\}$ by giving its n-step transition probabilities. Claim 2 classifies link-marking states as α varies, and shows that $\{X_i\}$'s state recurring probabilities converge asymptotically to the marginal link-marking probabilities [see

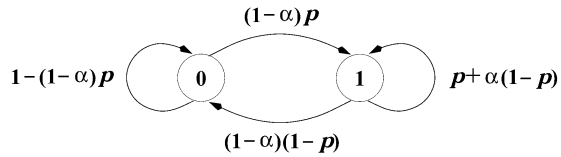


Fig. 4. Markov-chain model for dependent marking multicast flow control.

(40)], if $\{X_i\}$ is not perfectly dependent ($\alpha \neq 1$). Claim 3 ensures that the Markov-chain dependency-degree model converges asymptotically, and the long-term behavior of $\{X_i\}$ is stable. Also, the ergodicity of $\{X_i\}$ enables us to evaluate its various statistics (ensemble average) through the sample averages. Moreover, this claim shows that the limiting probabilities converge to the marginal link-marking probabilities $\Pr\{X_i = x_i\}$ ($x_i \in \{0,1\}$). This is expected because π_0 and π_1 represent the long-term proportion of $\{X_i\}$ remaining at state 0 and 1, respectively, being consistent with the definitions of $\Pr\{X_i = 0\}$ and $\Pr\{X_i = 1\}$, thus validating the Markov-chain dependency-degree model. Claim 4 says that if $\alpha = 1$ (perfectly dependent), the equilibrium distribution still exists, but is not unique, depending on initial state probabilities. This is expected since when $\alpha = 1$, $\{X_i\}$ has two isolated classes (see Fig. 4). So, $\{X_i\}$ is not irreducible and thus is no longer ergodic.

VI. Numerical Evaluations and Comparisons

A. Multicast-Tree Bottleneck Path Distribution $\psi_d(P_k, \alpha, p, m)$

Using the derived analytical results, Fig. 5(a) plots $\psi_d(P_k, \alpha, p, m)$ versus path length k while varying the Markov-chain dependency-degree factor α. $\psi_d(P_k, \alpha, p, m)$ is found to be a strictly monotonic decreasing function of k for both the independent ($\alpha = 0$) and dependent ($\alpha > 0$) cases. This is expected because the longer the bottleneck path, the more likely it will be dominated by shorter paths, as described in Definition 2.

Compared to the independent-marking case, the marking dependency is found to reduce the probability for shorter paths (with $k \leq 4$), while increasing the probability for longer paths (with $k > 5$), to become the bottleneck. This verifies Claim 4 [see (32)] of Theorem 3, and the dependency-balanced path number $\tilde{k}$ is found to be around 4–5. Fig. 5(a) also shows that the larger α, the more this probability shifts from shorter to longer paths. This is because the stronger the link-marking dependency, the larger the probability that all links stay in the *same* congestion state. This trend is also shown in Fig. 7(a), which plots $\psi_d(P_k, \alpha, p, m)|_{\alpha=0} - \psi_d(P_k, \alpha, p, m)|_{\alpha=\alpha_0>0}$ versus k for different values of $\alpha = \alpha_0$.

Fig. 5(b) shows that $\psi_d(P_k, \alpha, p, m)$ is inversely proportional to path length k, also verifying the above observations. Fig. 5(b) also shows that $\psi_d^*(P_k, \alpha, p^*, m)$ has the unique maximum, given k, verifying Claim 2 of Theorem 3. Fig. 5(c) says that given α, the larger k, the smaller $\psi_d(P_k, \alpha, p, m)$. Fig. 5(c) also shows that $\psi_d(P_k, \alpha, p, m)$ is not always a monotonic function of α, but it can have the maximum $\psi_d^*(P_k, \alpha^*, p, m)$ if k and p satisfy the conditions in (30). As k gets larger, α^* increases. These validate Claim 3 of Theorem 3. Fig. 6(a) shows complete dynamic behaviors of $\psi_d(P_k, \alpha, p, m)$ as a function of (α, p). Fig. 6(a) shows that $\psi_d(P_k, \alpha, p, m)$ always has the maximum on p-axis as α varies from 0 to 1. In contrast, for given k, $\psi_d(P_k, \alpha, p, m)$ can have the maximum on α-axis only for some p's satisfying the conditions in (30) in Theorem 3.

B. Delay Statistics of HBH and SSP for Dependent Markings

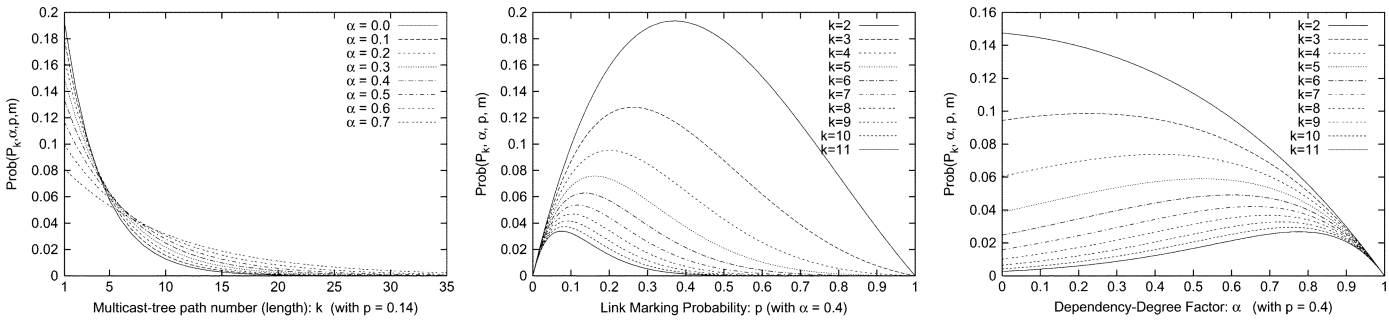
Fig. 6(b) plots $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ and $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ by (34) and (35) versus m with α varying. $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ is found much larger, and increasing much faster, than $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ [see Fig. 6(b)]. Moreover, $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ is more sensitive to α than $\overline{\tau}_{\rm SSP}(\alpha, p, m)$. Fig. 6(b) also shows that, compared to $\overline{\tau}_{\rm HBH}(\alpha, p, m)$, $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ is virtually independent of m and α. Fig. 6(b) also indicates that for longer paths ($m > 20$), the larger α, the larger the means while for shorter paths ($m < 12$), the larger α, the smaller the means, verifying that the bottleneck path probabilities shift from shorter to longer paths as α increases [see Fig. 7(a)].

Fig. 6(c) plots the standard deviations of $\sigma_{\rm SSP}(\alpha, p, m)$ and $\sigma_{\rm HBH}(\alpha, p, m)$ by (36) and (37), versus m while varying α. Fig. 6(c) says that $\sigma_{\rm HBH}(\alpha, p, m)$ is much larger, and increases much faster, than $\sigma_{\rm SSP}(\alpha, p, m)$ as m increases. Also, $\sigma_{\rm HBH}(\alpha, p, m)$ is much more sensitive to α than $\sigma_{\rm SSP}(\alpha, p, m)$. Thus, SSP's multicast RTT scales much better than HBH's in terms of multicast-tree height and structure. Fig. 6(c) also shows that SSP's RTT variation $\sigma_{\rm SSP}(\alpha, p, m)$ is virtually independent of m and α, compared to HBH's $\sigma_{\rm HBH}(\alpha, p, m)$. Fig. 6(c) says that for longer paths ($m \geq 10$), the larger α, the larger the variances while for shorter paths ($m < 8$), the larger α, the smaller the variances, also verifying that the bottleneck probabilities shift from shorter to longer paths as α increases [also see Fig. 7(a)].

C. Marking Dependency Impact on Multicast Signaling Delays

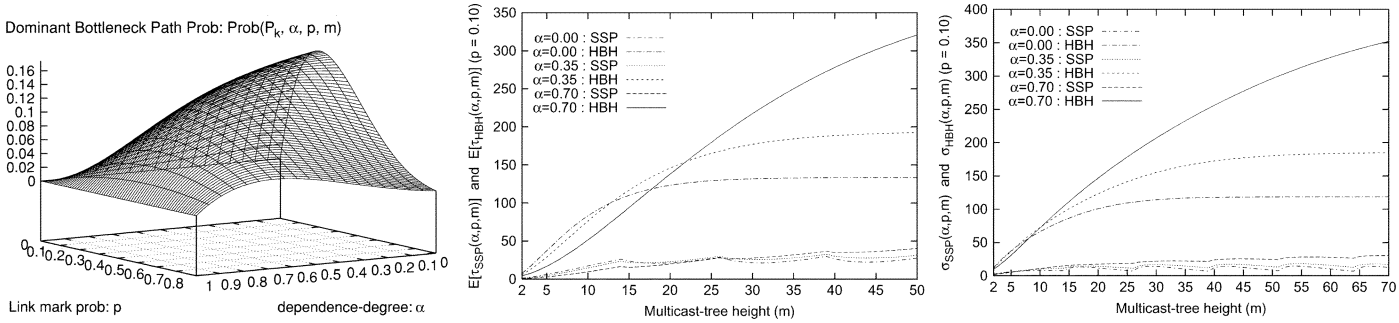
Fig. 7(b) and (c) plots the means of multicast signaling RTTs $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ and $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ versus traffic load p, while varying α. We observe: 1) there is a unique maximum for each $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ and $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ in terms of p, being consistent with the unique maximum of $\psi_d(P_k, \alpha, p, m)$ in Claim 2 of Theorem 3; 2) the maximizers for $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ and $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ shift from a small to large p as α increases; 3) as α increases, $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ and $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ become less sensitive to p; and 4) $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ is about 2 times larger than $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ for any p derived in our parameter settings.

To assess the approximation error of the multicast delay analysis caused by assuming independent marking, while the actual link markings are dependent, Fig. 8(a) and (b) plots the approximation errors of means: $\varepsilon_m^{\rm HBH}(\alpha, p, m) \triangleq \overline{\tau}_{\rm HBH}(\alpha, p, m)|_{\alpha=0} - \overline{\tau}_{\rm HBH}(\alpha, p, m)|_{\alpha>0}$ and $\varepsilon_m^{\rm SSP}(\alpha, p, m) \triangleq \overline{\tau}_{\rm SSP}(\alpha, p, m)|_{\alpha=0} - \overline{\tau}_{\rm SSP}(\alpha, p, m)|_{\alpha>0}$, respectively, between the multicast-signaling delay analyses under the *dependent* ($\alpha > 0$) and *independent* ($\alpha = 0$) markings. We observe: 1) the maxima of both $\varepsilon_m^{\rm HBH}(\alpha, p, m)$ and $\varepsilon_m^{\rm SSP}(\alpha, p, m)$ are monotonically increasing functions of α, showing that the approximation error increases as α increases; 2) $\varepsilon_m^{\rm HBH}(\alpha, p, m)$ and $\varepsilon_m^{\rm SSP}(\alpha, p, m)$ are not monotonic functions of p, but change from a positive to a negative value as p increases $0 \rightarrow 1$, indicating that the analysis assuming independent *overestimates* the mean delay for small p, while *underestimating* the mean delay for large p; 3) the approximation error for HBH is more than two times higher than that for SSP. Thus, the approximation error in multicast signaling-delay mean caused by the inaccurate independent assumption is not negligible, justifying the necessity of a Markov-chain-based marking-dependency



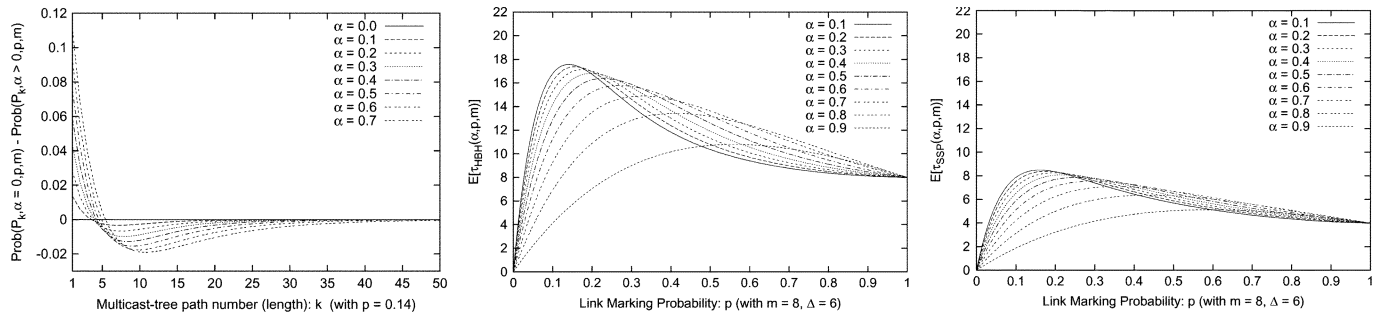
(a) (b) (c)

Fig. 5. Impact of path length k, link-marking probability p, and dependency-degree α on multicast-tree bottleneck path probability distribution $\psi_d(P_k, \alpha, p, m)$ shows the dynamics of multicast-tree bottleneck path probabilities. (a) $\psi_d(P_k, \alpha, p, m)$ versus k. (b) $\psi_d(P_k, \alpha, p, m)$ versus p. (c) $\psi_d(P_k, \alpha, p, m)$ versus α.



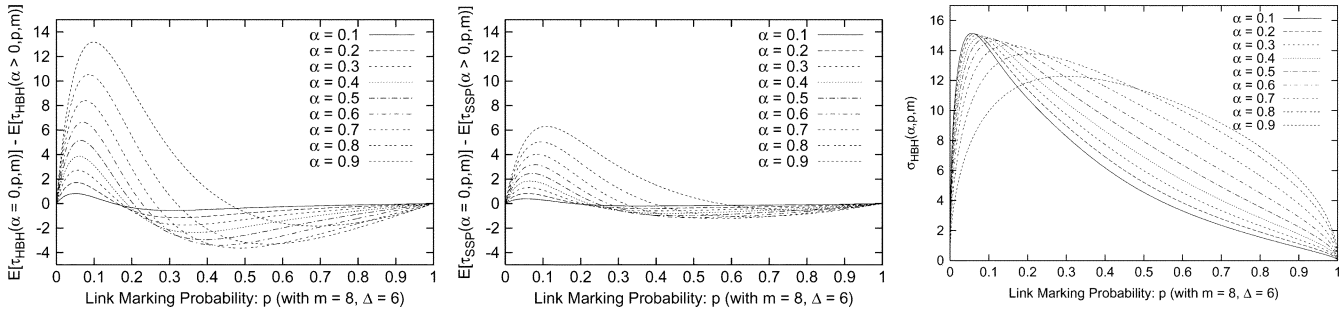
(a) (b) (c)

Fig. 6. Impact of dependency-degree factor α, link-marking probability p, and multicast-tree height m on multicast-tree bottleneck path probability $\psi_d(P_k, \alpha, p, m)$ and bottleneck RM-cell RTT means and standard deviations. (a) $\psi_d(P_k, \alpha, p, m)$ versus (α, p). (b) $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ and $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ versus m. (c) $\sigma_{\rm SSP}(\alpha, p, m)$ and $\sigma_{\rm HBH}(\alpha, p, m)$ versus m.



(a) (b) (c)

Fig. 7. Impact of dependency-degree factor α and link-marking probability p on bottleneck path probability $\psi_d(P_k, \alpha, p, m)$ shift and bottleneck RM-cell RTT means. (a) $\psi_d(P_k, \alpha, p, m)|_{\alpha=0} - \psi_d(P_k, \alpha, p, m)|_{\alpha>0}$ versus k. (b) $\overline{\tau}_{\rm HBH}(\alpha, p, m)$ versus p. (c) $\overline{\tau}_{\rm SSP}(\alpha, p, m)$ versus p.



(a) (b) (c)

Fig. 8. Impact of dependency-degree factor α and link-marking probability p on approximation error under independent markings assumption and bottleneck RM-cell RTT standard deviations. (a) $\varepsilon_m^{\rm HBH}(\alpha, p, m)$ versus p. (b) $\varepsilon_m^{\rm SSP}(\alpha, p, m)$ versus p. (c) $\sigma_{\rm HBH}(\alpha, p, m)$ versus p.

analysis. Similar observations are drawn for the multicast signaling *delay variations* [see Fig. 8(c)] and their approximation errors caused by "independent" assumptions, as shown in [15].

VII. Conclusion

We modeled and analyzed the delay performance of a class of multicast feedback-synchronization signaling protocols. Specifically, we developed a Markov-chain model to characterize the multicast signaling delay when the congestion markings at different links are dependent. We also developed a Markov-chain dependency-degree model to quantify the dependency degree between different link markings. Using these two models we derived the first and second moments of the multicast-signaling delay for HBH and SSP schemes with dependent link markings. Our numerical evaluations showed that SSP outperforms HBH in terms of multicast signaling delay, and the marking dependency tends to shift the bottleneck from shorter to longer paths, being consistent with the definition of the positive link-marking dependency imposed by the nature of multicast signaling.

Appendix A
Proof of Theorem 1

Proof: For presenting convenience, we start with Claim 2.

Claim 2: Since $0 < p_i < 1$, it is possible that all $(2m-1)$ links in $\mathcal{L} = \{L_1, L_2', L_2, L_3', L_3, \ldots, L_m', L_m\}$ [see Fig. 3(a)] defined by Definition 1 are not marked—no dominant bottleneck path. If at least one of $(2m-1)$ links is marked, by Definition 2, the shortest path containing the marked link(s) is the dominant bottleneck. By Definitions 1 and 2, the dominant bottleneck path is unique. Thus, there is at most one dominant bottleneck path.

By Definitions 1 and 2, the probability that P_1 becomes the dominant bottleneck path equals the probability that $X_1 = 1$ or $X_2' = 1$, which yields the first part of (13) as follows:

$$\begin{aligned}\psi_d(P_1, m) &= \Pr\{X_1 = 1 \cup X_2' = 1\} = 1 - \Pr\{X_1 = 0 \cap X_2' = 0\} \\ &= 1 - \Pr\{X_1 = 0\}\Pr\{X_2' = 0 \mid X_1 = 0\}.\end{aligned}$$

Consider path $P_k, 2 \le k \le m-1$. Since the last two links are L_k and L_{k+1}' [see Fig. 3(a)], the probability that P_k becomes the dominant bottleneck path is equal to the probability that $X_i = 0, \forall i \in \{1, 2, \ldots, k-1\}$ and $X_i' = 0, \forall i \in \{2, 3, \ldots, k\}$, and $X_k = 1$ or $X_{k+1}' = 1$, which leads to

$$\begin{aligned}&\psi_d(P_k, m) \\ &= \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\} \cap \{X_k = 1 \cup X_{k+1}' = 1\}\right\} \\ &= \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_k = 1\right\} \\ &\quad + \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_{k+1}' = 1\right\} \\ &\quad - \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_k = 1, X_{k+1}' = 1\right\} \\ &= \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_k = 1\right\} \\ &\quad - \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_k = 1, X_{k+1}' = 1\right\} \\ &\quad + \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, \right. \\ &\qquad \left. \{X_k = 0 \cup X_k = 1\}, X_{k+1}' = 1\right\} \\ &= \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_k = 1\right\} \\ &\quad + \Pr\left\{\bigcap_{i=1}^{k-1}\{X_i = 0, X_{i+1}' = 0\}, X_k = 0, X_{k+1}' = 1\right\} \\ &= \Pr\{X_1 = 0\}\Pr\{X_2 = 0 \mid X_1 = 0\} \\ &\quad \cdot \Pr\{X_3 = 0 \mid X_2 = 0\} \\ &\quad \cdot \cdots \Pr\{X_{k-1} = 0 \mid X_{k-2} = 0\} \\ &\quad \cdot \Pr\{X_2' = 0 \mid X_1 = 0\} \\ &\quad \cdot \Pr\{X_3' = 0 \mid X_2 = 0\} \cdots \Pr\{X_k' = 0 \mid X_{k-1} = 0\} \\ &\quad \cdot \Pr\{X_k = 1 \mid X_{k-1} = 0\} + \Pr\{X_1 = 0\} \\ &\quad \cdot \Pr\{X_2 = 0 \mid X_1 = 0\}\Pr\{X_3 = 0 \mid X_2 = 0\} \cdots \\ &\quad \cdot \Pr\{X_k = 0 \mid X_{k-1} = 0\}\Pr\{X_2' = 0 \mid X_1 = 0\} \\ &\quad \cdot \Pr\{X_3' = 0 \mid X_2 = 0\} \cdots \Pr\{X_k' = 0 \mid X_{k-1} = 0\} \\ &\quad \cdot \Pr\{X_{k+1}' = 1 \mid X_k = 0\} \qquad (42) \\ &= \Pr\{X_1 = 0\}\Pr\{X_k' = 0 \mid X_{k-1} = 0\} \\ &\quad \cdot \Big[\Pr\{X_k = 1 \mid X_{k-1} = 0\} + \Pr\{X_k = 0 \mid X_{k-1} = 0\} \\ &\quad \cdot \Pr\{X_{k+1}' = 1 \mid X_k = 0\}\Big] \prod_{i=1}^{k-2}\Big\{\Pr\{X_{i+1} = 0 \mid X_i = 0\} \\ &\quad \cdot \Pr\{X_{i+1}' = 0 \mid X_i = 0\}\Big\} \qquad (43)\end{aligned}$$

where (42) is due to **C3** and **C4** of Definition 1. Thus, (43) yields the second part of (13).

The probability that P_m becomes the dominant bottleneck path is equal to the probability that $X_i = 0, \forall i \in \{1, 2, \ldots, m-1\}$ and $X_i' = 0, \forall i \in \{2, 3, \ldots, m\}$, and $X_m = 1$, implying

$$\begin{aligned}&\psi_d(P_k, m) \\ &= \Pr\left\{\bigcap_{i=1}^{m-1}\{X_i = 0, X_{i+1}' = 0\}, X_m = 1\right\} = \Pr\{X_1 = 0\} \\ &\quad \cdot \Pr\{X_m = 1 \mid X_{m-1} = 0\}\Pr\{X_m' = 0 \mid X_{m-1} = 0\} \\ &\quad \cdot \prod_{i=1}^{m-2}\Big\{\Pr\{X_{i+1} = 0 \mid X_i = 0\}\Pr\{X_{i+1}' = 0 \mid X_i = 0\}\Big\} \qquad (44)\end{aligned}$$

where (44) follows from the proof of the first term of (42) and is also due to **C3** and **C4** of Definition 1. Hence, the third part of (13) follows from (44).

Claim 1: Equation (11) follows from the proof of the first and second parts of (13). Now, we prove (12):

$$\begin{aligned}&\lim_{m\to\infty}\sum_{k=1}^{m}\psi_d(P_k, \infty) \\ &= \lim_{m\to\infty}\Big\{1 - \Pr\{X_1 = 0\}\Pr\{X_2' = 0 \mid X_1 = 0\} \\ &\quad + \Pr\{X_1 = 0\}\Pr\{X_2' = 0 \mid X_1 = 0\}\Big[\Pr\{X_2 = 1 \mid X_1 = 0\}\end{aligned}$$

$$
\begin{aligned}
&+\Pr\{X_2=0\mid X_1=0\}\Pr\{X_3'=1\mid X_2=0\}\Big]\\
&+\Pr\{X_1=0\}\Big[\Pr\{X_2=0|X_1=0\}\Pr\{X_2'=0|X_1=0\}\Big]\\
&\cdot\Pr\{X_3'=0\mid X_2=0\}\Big[\Pr\{X_3=1\mid X_2=0\}\\
&+\Pr\{X_3=0\mid X_2=0\}\Pr\{X_4'=1\mid X_3=0\}\Big]+\cdots\\
&+\Pr\{X_1=0\}\Big[\Pr\{X_2=0\mid X_1=0\}\\
&\cdot\Pr\{X_2'=0\mid X_1=0\}\Pr\{X_3=0\mid X_2=0\}\\
&\cdot\Pr\{X_3'=0\mid X_2=0\}\Pr\{X_4=0\mid X_3=0\}\\
&\cdot\Pr\{X_4'=0\mid X_3=0\}\cdots\Pr\{X_{m-1}=0\mid X_{m-2}=0\}\\
&\cdot\Pr\{X_{m-1}'=0\mid X_{m-2}=0\}\Big]\Pr\{X_m'=0\mid X_{m-1}=0\}\\
&\cdot\Big[\Pr\{X_m=1\mid X_{m-1}=0\}+\Pr\{X_m=0\mid X_{m-1}=0\}\\
&\cdot\Pr\{X_{m+1}'=1\mid X_m=0\}\Big]\Big\}\\
&=\lim_{m\to\infty}\Big\{1-\Pr\{X_1=0\}\Pr\{X_2'=0|X_1=0\}+\Pr\{X_1=0\}\\
&\cdot\Pr\{X_2'=0\mid X_1=0\}\Big[\Big(1-\Pr\{X_2=0\mid X_1=0\}\Big)\\
&+\Pr\{X_2=0\mid X_1=0\}\Big(1-\Pr\{X_3'=0\mid X_2=0\}\Big)\Big]\\
&+\Pr\{X_1=0\}\Big[\Pr\{X_2=0\mid X_1=0\}\Pr\{X_2'=0|X_1=0\}\Big]\\
&\cdot\Pr\{X_3'=0\mid X_2=0\}\Big[\Big(1-\Pr\{X_3=0\mid X_2=0\}\Big)\\
&+\Pr\{X_3=0|X_2=0\}\Big(1-\Pr\{X_4'=0|X_3=0\}\Big)\Big]+\cdots\\
&+\Pr\{X_1=0\}\Big[\Pr\{X_2=0|X_1=0\}\Pr\{X_2'=0|X_1=0\}\\
&\cdot\Pr\{X_3=0|X_2=0\}\Pr\{X_3'=0|X_2=0\}\Pr\{X_4=0|X_3=0\}\\
&\cdot\Pr\{X_4'=0\mid X_3=0\}\cdots\Pr\{X_{m-1}=0\mid X_{m-2}=0\}\\
&\cdot\Pr\{X_{m-1}'=0\mid X_{m-2}=0\}\Big]\Pr\{X_m'=0\mid X_{m-1}=0\}\\
&\cdot\Big[\Big(1-\Pr\{X_m=0|X_{m-1}=0\}\Big)+\Pr\{X_m=0|X_{m-1}=0\}\\
&\cdot\Big(1-\Pr\{X_{m+1}'=0\mid X_m=0\}\Big)\Big]\Big\}\\
&=\lim_{m\to\infty}\Big\{1-\Pr\{X_1=0\}\Pr\{X_2'=0|X_1=0\}+\Pr\{X_1=0\}\\
&\cdot\Pr\{X_2'=0\mid X_1=0\}\Big[1-\Pr\{X_2=0\mid X_1=0\}\\
&\cdot\Pr\{X_3'=0|X_2=0\}\Big]+\Pr\{X_1=0\}\Big[\Pr\{X_2=0|X_1=0\}\\
&\cdot\Pr\{X_2'=0\mid X_1=0\}\Big]\Pr\{X_3'=0\mid X_2=0\}\\
&\cdot\Big[1-\Pr\{X_3=0|X_2=0\}\Pr\{X_4'=0|X_3=0\}\Big]+\cdots\\
&+\Pr\{X_1=0\}\Big[\Pr\{X_2=0|X_1=0\}\Pr\{X_2'=0|X_1=0\}\\
&\cdot\Pr\{X_3=0|X_2=0\}\Pr\{X_3'=0|X_2=0\}\Pr\{X_4=0|X_3=0\}\\
&\cdot\Pr\{X_4'=0\mid X_3=0\}\cdots\Pr\{X_{m-1}=0\mid X_{m-2}=0\}\\
&\cdot\Pr\{X_{m-1}'=0\mid X_{m-2}=0\}\Big]\Pr\{X_m'=0|X_{m-1}=0\}\\
&\cdot\Pr\{X_m=0|X_{m-1}=0\}\Pr\{X_{m+1}'=0\mid X_m=0\}\Big\}\\
&=\lim_{m\to\infty}\Big\{1-\Pr\{X_1=0\}\Pr\{X_{m+1}'=0\mid X_m=0\}\\
&\cdot\prod_{i=1}^{m-1}\Big[\Pr\{X_{i+1}=0\mid X_i=0\}\Pr\{X_{i+1}'=0\mid X_i=0\}\Big]\Big\}\\
&=1-\Pr\{X_1=0\}\lim_{m\to\infty}\Pr\{X_{m+1}'=0\mid X_m=0\}\\
&\cdot\lim_{m\to\infty}\prod_{i=1}^{m-1}\Big[\Pr\{X_{i+1}=0|X_i=0\}\Pr\{X_{i+1}'=0|X_i=0\}\Big] \qquad (45)\\
&=1 \qquad (46)
\end{aligned}
$$

where (46) holds since the limiting terms of (45) satisfy

$$
\begin{aligned}
0&\le\lim_{m\to\infty}\Pr\{X_{m+1}'=0\mid X_m=0\}\\
&\quad\cdot\lim_{m\to\infty}\prod_{i=1}^{m-1}\Big[\Pr\{X_{i+1}=0|X_i=0\}\Pr\{X_{i+1}'=0|X_i=0\}\Big]\\
&\le\lim_{k\to\infty}\prod_{i=1}^{k-1}\Big[p_{\max}p'_{\max}\Big]=\lim_{k\to\infty}\Big[p_{\max}p'_{\max}\Big]^{k-1}=0 \qquad (47)
\end{aligned}
$$

where because the link-marking states are not always *perfectly dependent*, for $m=\infty$ there always exists a subsequence[10] for any given multicast tree defined by Definition 1 such that $\Pr\{X_{n_i+1}=0\mid X_{n_i}=0\}<1$ and $\Pr\{X'_{m_i+1}=0\mid X_{m_i}=0\}<1$ for $1\le n_1<n_2<\cdots$ and $1\le m_1<m_2<\cdots$, respectively, implying that

$$
\begin{cases}
p_{\max}\triangleq\max\limits_{i\in\{n_1,n_2,\cdots\}}\Big\{\Pr\{X_{i+1}=0|X_i=0\}\Big\}<1\\
p'_{\max}\triangleq\max\limits_{i\in\{m_1,m_2,\cdots\}}\Big\{\Pr\{X'_{i+1}=0|X_i=0\}\Big\}<1
\end{cases} \qquad (48)
$$

which proves (12). Thus, $\psi_d(P_k,\infty),\forall\, k\in\{1,2,\ldots,\infty\}$ defines a valid *pmf*. Moreover, (46) also implies that there is at least one dominant bottleneck path as $m\to\infty$. But, based on the tree structure defined by Definitions 1 and 2, there is at most one dominant bottleneck path. Thus, there exists one and only one dominant bottleneck path, completing the proof. ■

[10]For the trivial case of $\Pr\{X_{i+1}=0\mid X_i=0\}\equiv 1$ and $\Pr\{X'_{i+1}=0\mid X_i=0\}\equiv 1,\forall i$ (perfectly dependent), it is easy to prove that Theorem 1 still holds.

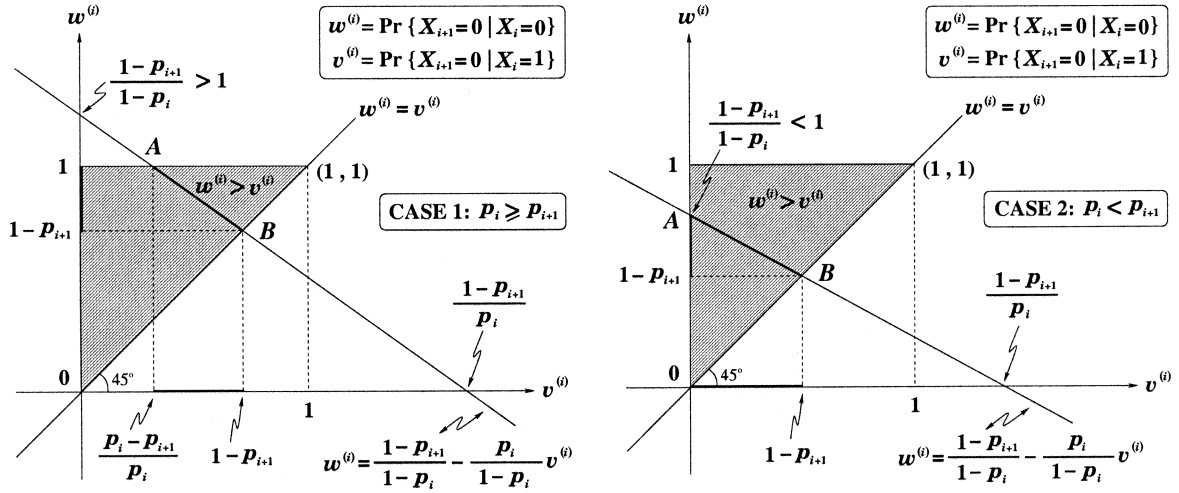


Fig. 9. Markov-chain dependency-degree modeling for CASE 1 and CASE 2. (a) CASE 1: $p_i \geq p_{i+1}$. (b) CASE 2: $p_i < p_{i+1}$.

APPENDIX B
PROOF OF THEOREM 2

Proof: Claim 1: Consider link-marking states X_i and $X_{i+1}, i = 1, 2, \ldots$. By the partition rule, we have

$$\Pr\{X_{i+1}=0\}=\Pr\{X_i=0\}\Pr\{X_{i+1}=0\,|\,X_i=0\} + \Pr\{X_i=1\}\Pr\{X_{i+1}=0\,|\,X_i=1\}. \quad (49)$$

By defining $w^{(i)} \triangleq \Pr\{X_{i+1} = 0\,|\,X_i = 0\}$ and $v^{(i)} \triangleq \Pr\{X_{i+1} = 0\,|\,X_i = 1\}$, (49) reduces to $w^{(i)} = (1-p_{i+1})/(1-p_i) - [p_i/(1-p_i)]v^{(i)}$, yielding a fundamental system functional $f(\cdot)$ between $w^{(i)}$ and $v^{(i)}$ as follows:

$$w^{(i)} \triangleq f(v^{(i)}) \triangleq \frac{1-p_{i+1}}{1-p_i} - \frac{p_i}{1-p_i}v^{(i)}. \quad (50)$$

Then, we can solve for the upper and lower bounds for $w^{(i)}$ and $v^{(i)}$ subject to the following three constraints:

- **C1.** $(w^{(i)}, v^{(i)}) \in \{(w^{(i)}, v^{(i)})\,|\,w^{(i)} = f(v^{(i)})\}$: where $f(\cdot)$ is defined in (50);
- **C2.** $w^{(i)} > v^{(i)}$: because the Markov chain $\{X_i\}$ is positively dependent (see Definition 3);
- **C3.** $0 \leq w^{(i)}, v^{(i)} \leq 1$: since $w^{(i)}$ and $v^{(i)}$ are both probabilities.

We need to consider the following two cases, depending on $p_i \geq p_{i+1}$ or $p_i < p_{i+1}$.

CASE 1: $p_i \geq p_{i+1}$: Fig. 9(a) plots the derived feasible solution regions, under **C1**, **C2**, and **C3**, for CASE 1 in the space spanned by $v^{(i)}$ and $w^{(i)}$ axes. **C1** requires that all solution points be on the line of $w^{(i)} = f(v^{(i)}) = (1-p_{i+1})/(1-p_i) - [p_i/(1-p_i)]v^{(i)}$; **C2** says all solution points must be within the region between the positive half axis of $w^{(i)}$ and the 45° line $w^{(i)} = v^{(i)}$ [shaded area in Fig. 9(a)]; **C3** confines all solution points within the unit square area $w^{(i)} \in [0,1]$ and $v^{(i)} \in [0,1]$. Applying **C1**, **C2**, and **C3**, the solution point set for $\{(w^{(i)}, v^{(i)})\}$ lies between points A and B on the line $w^{(i)} = f(v^{(i)}) = (1-p_{i+1})/(1-p_i) - [p_i/(1-p_i)]v^{(i)}$. After some algebra, the projection points of A and B onto $w^{(i)}$ and $v^{(i)}$ axes are $w_A^{(i)} = 1, w_B^{(i)} = 1-p_{i+1}$ and $v_A^{(i)} = (p_i - p_{i+1})/p_i, v_B^{(i)} = 1-p_{i+1}$, respectively. Then, the projections of A and B onto $w^{(i)}$ axis give $w^{(i)}$'s upper bound $w_{\max}^{(i)}$ and lower bound $w_{\min}^{(i)}$, respectively, while the projections of A and B onto the $v^{(i)}$ axis yield $v^{(i)}$'s lower bound $v_{\min}^{(i)}$ and upper bound $v_{\max}^{(i)}$, respectively, as follows:

$$\begin{cases} w_{\min}^{(i)} = w_B^{(i)} = 1-p_{i+1} \\ w_{\max}^{(i)} = w_A^{(i)} = 1 \end{cases} \Rightarrow \begin{cases} 1-p_{i+1} \leq \\ \Pr\{X_{i+1}=0\,|\,X_i=0\} \leq 1 \end{cases} \quad (51)$$

which proves the first part of (14). Similarly, the first part of (16) holds due to (52):

$$\begin{cases} v_{\min}^{(i)} = v_A^{(i)} = (p_i - p_{i+1})/p_i \\ v_{\max}^{(i)} = v_B^{(i)} = 1-p_{i+1} \end{cases} \Rightarrow \begin{cases} (p_i - p_{i+1})/p_i \leq \\ \Pr\{X_{i+1}=0\,|\,X_i=1\} \\ \leq 1-p_{i+1}. \end{cases} \quad (52)$$

CASE 2: $p_i < p_{i+1}$: Fig. 9(b) plots the derived feasible solution regions, by **C1**, **C2**, and **C3**, for CASE 2 in the same space. Due to $p_i < p_{i+1}$, the line $w^{(i)} = f(v^{(i)}) = (1-p_{i+1})/(1-p_i) - [p_i/(1-p_i)]v^{(i)}$ intersects with $w^{(i)}$ axis at point below 1 now. The new projections of line between A and B onto $w^{(i)}$ and $v^{(i)}$ axes are $w_A^{(i)} = (1-p_{i+1})/(1-p_i), w_B^{(i)} = 1-p_{i+1}$, and $v_A^{(i)} = 0, v_B^{(i)} = 1-p_{i+1}$, respectively, which yield CASE 2's upper and lower bounds for $w^{(i)}$ and $v^{(i)}$, proving the second part of (14) as follows:

$$\begin{cases} w_{\min}^{(i)} = w_B^{(i)} = 1-p_{i+1} \\ w_{\max}^{(i)} = w_A^{(i)} = \dfrac{1-p_{i+1}}{1-p_i} \end{cases} \Rightarrow \begin{cases} 1-p_{i+1} \leq \\ \Pr\{X_{i+1}=0\,|\,X_i=0\} \\ \leq \dfrac{1-p_{i+1}}{1-p_i} \end{cases} \quad (53)$$

Likewise, the second part of (16) holds due to (54):

$$\begin{cases} v_{\min}^{(i)} = v_A^{(i)} = 0 \\ v_{\max}^{(i)} = v_B^{(i)} = 1-p_{i+1} \end{cases} \Rightarrow \begin{cases} 0 \leq \Pr\{X_{i+1}=0\,|\,X_i=1\} \\ \leq 1-p_{i+1}. \end{cases} \quad (54)$$

The rest of the proof is completed in [15]. The proof follows. ■

APPENDIX C
PROOF OF THEOREM 4

Proof: Claim 1: Since $\{X_i\}$ is homogeneous, the matrix P of $\{X_i\}$'s one-step transition probabilities is fixed and defined by (20)–(23) for $\alpha_i = \alpha$ and $p_i = p, \forall i$. We prove (38)[11] by mathematical induction.

Base Case: Using (20)–(23) and (38)

$$P \triangleq \{p_{jk}\} = \Big\{\Pr\{X_{i+1} = k \mid X_i = j\}\Big\} = \begin{bmatrix} 1-(1-\alpha)p & (1-\alpha)p \\ (1-\alpha)(1-p) & \alpha(1-p)+p \end{bmatrix} = P^{(n)}\,|_{n=1} \quad (55)$$

where $j, k \in \{0,1\}; i \in \{1,2,\ldots\}$. So, (38) holds for $n = 1$.

Inductive Hypothesis: Suppose (38) holds for $n = q-1$, then

$$P^{(q)} \triangleq \left\{p_{jk}^{(q)}\right\} = P^{(q-1)}P = \begin{bmatrix} 1-\left(1-\alpha^{(q-1)}\right)p & \left(1-\alpha^{(q-1)}\right)p \\ \left(1-\alpha^{(q-1)}\right)(1-p) & \alpha^{(q-1)}(1-p)+p \end{bmatrix} P = \begin{bmatrix} 1-(1-\alpha^q)p & (1-\alpha^q)p \\ (1-\alpha^q)(1-p) & \alpha^q(1-p)+p \end{bmatrix}. \quad (56)$$

So, (38) holds for $n = q$, and thus also holds $\forall n \in \{0,1,\ldots\}$.

Claim 2: Equation (40) follows directly from (38). States j ($j \in \{0,1\}$) are ergodic *iff* j is positive recurrent and aperiodic. Clearly, j is aperiodic (period $d = 1$). Since $\{X_i\}$ is a finite-state Markov chain, we only need prove that j is recurrent, which is true because from (38), $0 < p < 1$, and $\alpha \in [0,1]$, we obtain the following:

$$\begin{cases} \lim_{n\to\infty} \sum_{r=1}^{n} p_{00}^{(r)} = \sum_{n=1}^{\infty}(1-p) + \sum_{n=1}^{\infty} p\alpha^n = \infty \\ \lim_{n\to\infty} \sum_{r=1}^{n} p_{11}^{(r)} = \sum_{n=1}^{\infty}(1-p)\alpha^n + \sum_{n=1}^{\infty} p = \infty. \end{cases} \quad (57)$$

Thus, $\{X_i\}$'s states are ergodic. Equation (57) gives the second part of (39). Since states j ($j \in \{0,1\}$) are aperiodic, we have $\limsup_{n\to\infty} p_{jj}^{(n)} = \lim_{n\to\infty} p_{jj}^{(n)}$. By (40), we get $\limsup_{n\to\infty} p_{jj}^{(n)} = \lim_{n\to\infty} p_{jj}^{(n)} > 0$, which proves the first part of (39).

Claim 3: Since $\alpha \in [0,1)$, i.e., $\alpha \neq 1$, implying $\{X_i\}$ is not perfectly dependent, $\{X_i\}$ is irreducible. So, the Markov chain $\{X_i\}$ is ergodic due to Claim 2. Thus, $\{X_i\}$ has the unique limiting/equilibrium state probabilities determined by (38): $\pi_0 = \lim_{n\to\infty} p_{j0}^{(n)} = 1-p = \Pr\{X_i = 0\}; \pi_1 = \lim_{n\to\infty} p_{j1}^{(n)} = p = \Pr\{X_i = 1\}$ where $\alpha \in [0,1)$ and $\forall j \in \{0,1\}$. So, (41) follows.

Claim 4: Let $\vec{p}(n) \triangleq \big[p_0(n)\ p_1(n)\big] = \big[p_0(1)\ p_1(1)\big]P^{(n-1)}$ denote the vector of state probabilities at link n (≥ 1). Since $\alpha = 1, P^{(n)} = I$ (unit matrix), $\forall n$ by (38). Thus, we get $\vec{\pi} \triangleq \lim_{n\to\infty} \vec{p}(n) = \big[p_0(1)\ p_1(1)\big]$, completing the proof. ■

ACKNOWLEDGMENT

The authors thank S. H. Low and D. E. Lapsley for their preprints and the reviewers of IEEE INFOCOM 2001 and this TRANSACTIONS for their comments that improved the presentation of this paper.

[11]It is trivial to prove that (38) also holds for $n = 0$.

REFERENCES

[1] X. Zhang and K. G. Shin, "Delay analysis of feedback-synchronization signaling for multicast flow control," *IEEE/ACM Trans. Networking*, vol. 11, pp. 436–450, June 2003.
[2] X. Zhang, K. G. Shin, D. Saha, and D. Kandlur, "Scalable flow control for multicast ABR services in ATM networks," *IEEE/ACM Trans. Networking*, vol. 10, pp. 67–85, Feb. 2002.
[3] L. Roberts, "Point-to-multipoint ABR operation," ATM Forum, Contribution 95-0834, Aug. 1995.
[4] H. Saito, K. Kawashima, H. Kitazume, A. Koike, M. Ishizuka, and A. Abe, "Performance issues in public ABR service," *IEEE Commun. Mag.*, vol. 11, pp. 40–48, Nov. 1996.
[5] K.-Y. Siu and H.-Y. Tzeng, "On max-min fair congestion control for mulltiicast ABR services in ATM," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 545–556, Apr. 1997.
[6] Y.-Z. Cho and M.-Y. Lee, "An efficient rate-based algorithm for point-to-multipoint ABR service," in *Proc. IEEE GLOBECOM*, Nov. 1997.
[7] S. Fahmy, R. Jain, R. Goyal, B. Vandalor, and S. Kalyanaraman, "Feedbackback consolidation algorithms for ABR point-to-mulipoint connections in ATM networks," in *Proc. IEEE INFOCOM*, Apr. 1998.
[8] S. H. Low and D. Lapsley, "Optimization flow control—I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, pp. 861–874, Dec. 1999.
[9] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Trans. Networking*, vol. 1, pp. 397–413, Aug. 1993.
[10] D. E. Comer and D. L. Stevens, *Internetworking with TCP/IP, Volume II, Design, Implementation, and Internals*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[11] S. Sathaye, "ATM Forum Traffic Management Specifications Version 4.0," ATM Forum, Aug. 1995. Contribution 95-0013R7.1.
[12] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[13] P. P. Mishra and H. R. Kanakia, "A hop-by-hop rate-based congestion control scheme," in *Proc. ACM SIGCOMM*, Aug. 1992.
[14] S. McCanne and S. Floyd, ns-Network Simulator, Apr. 2002.
[15] X. Zhang and K. G. Shin, "Markov-chain modeling for multicast signaling delay analysis," Networking and Information Systems Labs., Dept. Electr. Eng., Texas A&M Univ., College Station, Tech. Rep. [Online.] Available: http://ece.tamu.edu/~xizhang/papers/mcast_mc.pdf, July 2003.
[16] S. Keshav, *An Engineering Approach to Computer Networking, ATM Networks, the Internet, and the Telephone Network*. Reading, MA: Addison Wesley, 1997.

Xi Zhang (S'89–SM'98) received the Ph.D. degree in electrical engineering and computer science (Electrical Engineering—Systems) from The University of Michigan, Ann Arbor.

He is currently an Assistant Professor and the Founding Director of the Networking and Information Systems Laboratory, Department of Electrical Engineering, Texas A&M University, College Station. He has published over 50 technical papers.

Dr. Zhang received the U.S. National Science Foundation CAREER Award in 2004 for his research in the areas of mobile wireless and multicast networking and systems.

Kang G. Shin (S'75–M'78–SM'83–F'92) received the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, in 1978.

He is the Kevin and Nancy O'Connor Chair Professor of Computer Science and Founding Director of the Real-Time Computing Laboratory, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor. He has supervised 49 Ph.D. theses, has published more than 600 technical papers, and has written the book *Real-Time Systems* (New York, NY: McGraw-Hill, 1997).

Dr. Shin is a Fellow of the Association for Computing Machinery (ACM).