The effect of network delays on Distributed Ledgers based on Direct Acyclic Graphs: A mathematical model

Navdeep Kumar, Alexandre Reiffers-Masson, Isabel Amigo, Santiago Ruano Rincon

The effect of network delays on Distributed Ledgers based on Direct Acyclic Graphs: A mathematical model

Navdeep Kumar[a,b], Alexandre Reiffers-Masson[a], Isabel Amigo[a], Santiago Ruano Rincón[a]

[a] *IMT Atlantique, firstname.lastname@imt-atlantique.fr*
[b] *Technion, Israel Institute of Technology, navdeepkumar@campus.technion.ac.il*

Abstract

We present a new stochastic model for the evolution of Directed Acyclic Graphs (DAG)-based distributed ledgers (DL), under the presence of heterogeneous delay. This model is used to analyse the performance metrics of the DL, showing in particular that the number of unapproved messages does not diverge to infinity, even under the presence of delay. We propose an analysis based on conveniently defined sets, as well as an alternative drift-based analysis. The former allows to get a bound on the number of unapproved messages, while the latter, through a simpler analysis, allows to probe the existence of such bound. For particular scenarios, we are able to derive the expected value of the drift of unapproved messages, through a Markov process-based approach. State-of-the-art mathematical models trying to capture the impact of delays on the performance of such DLs rely on some particular simplifications. In contrast, through our model, we are able to analytically derive similar performance guarantees, in a more realistic set-up. In particular, we focus on IOTA foundation's tangle, while our results can be extended to other DAG-based distributed ledgers. We compare our results to results obtained in a real testbed, showing good accordance between them.

Keywords: Distributed ledgers, DAG-based DLTs, Stochastic Process

1. Introduction

Distributed ledgers (DL), such as blockchains, constitute a huge innovation in the field of distributed systems. Such technologies are one of the available solutions to make secure and decentralized data storage possible. Well-known use cases of distributed ledgers are cryptocurrencies, as for instance the famous Bitcoin and Ethereum. In this paper, we study Directed Acyclic Graphs (DAG)-based DL, a specific type of DL that overcomes some of the drawbacks of blockchain-based DLs, and we model how its performance depends on the delay.

In blockchain technology, transactions are packed into blocks, which are then sequentially linked one to another. Typically, blocks are added to the ledger at a regular average interval (a.k.a. block interval). Block interval, along with the block side, is determined in such a way that blocks are propagated to most nodes in the overlay network before a new block is generated, reducing the probability of fork. This clearly limits transaction throughput, and in turn limits the throughput of the system, as has already been reported (see e.g. [6]). In addition, some blockchain implementations have very CPU-consuming consensus mechanisms, such as the Bitcoin's *mining*.

Solutions addressing the shortcomings of classical blockchain have been proposed. Several of them are still based on blockchain (e.g. lightning [18] and Algorand [14], see [24] for a recent survey on blockchain scalability solutions) and others are based on a different underlying structure: the so-called Directed-Acyclic Graph (DAG)-based DLs.

In a DAG-based DL, messages (including transactions) are represented as vertices of a graph, and directed edges from one vertex to another represent validation of the former by the latter. It is to note that messages or transactions are called also blocks, interchangeably. New messages are attached to the ledger by validating a number of already included messages. Such DLs, while still a young technology, present an appealing advantage compared to blockchains, as they make it easier (and thus faster) to attach new messages to the ledger, since they can be attached to any existent vertex in the graph. Different examples of DAG-based DLs exist, such as IOTA [19] and Obyte (formerly Byteball) [5]. Among them, IOTA, and its DAG-based structure called *the Tangle*, is one of the most active in terms of research as well as in terms of development. In this paper, we focus on the IOTA's Tangle as a subject of study,

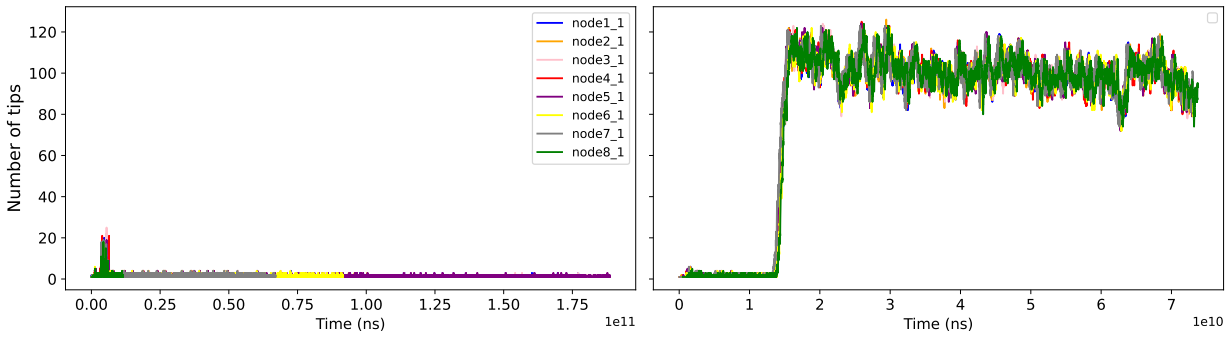


Figure 1: The impact of network delay on the Tangle regarding the number of concurrent unapproved messages (tips). A 8-node network without delay (left) versus the same network where a 800ms emulated delay is applied at around one minute(right).

though the extension of most results to other DAG-based DLs is straightforward. Indeed, the same set-based method could be applied by taking into account the specificity of the DL (for instance, number of messages to be validated or tip selection algorithm).

Altogether, Tangle-like DLs are promising solutions for supporting Internet-of-Things (IoT) applications such as micropayments in the case of IOTA, with its zero fee transactions. While advantages with respect to classical blockchain-based DLs are undeniable, DAG-based DLs also present some drawbacks. Nowadays one of the principal concerns is the lack of completely distributed solutions. In this vein, a research field, recently receiving a lot of attention, is the decentralization of the IOTA network, the so-called Coordicide project [20]. The Coordicide project addresses aspects such as consensus and admission control, however, there are still unanswered questions regarding the performance of such a DL.

In particular, under a fully distributed system, where nodes keep a local copy of the DL with a non-negligible delay between every pair of nodes, the state of the Tangle may not be the same for every node. Moreover, the performance of the register might be significantly impacted, resulting in an increased number of unapproved messages, with respect to the no-delay case. Empirical data obtained from a testbed supports such assertion, as shown in Figure 1. Indeed, we see that in a 8-node network, without delay between nodes, the mean average number of tips (unapproved messages, waiting for validation) is below 10. However, the moment we add a communication delay, 800 ms in this case, we see that the number of mean average tips increases up to 100. This simple example, coming from our testbed, illustrates the importance of understanding the impacts of communication delay in the DL's performance. We develop these observations through experiments in Section 5.

The evolution of the state of the Tangle is stochastic by nature, as a result of: (1) the random issuance of new messages; (2) the approval process, through which each new message approves randomly k unapproved messages and get attached to these selected messages. In addition, the system is distributed, meaning that due to delay between nodes, each node might see the system at a different state. Such elements make the system a complex one, and the task of finding a suitable mathematical model is a challenge. Indeed, several attempts in the literature have tried to assess the properties of DAG-based DLs, and in particular of the Tangle, such as the stability (in the number of unapproved messages) and the validation time of messages. These key metrics are revealing the health of the ledger, see for instance [19, 21, 2, 22], and Section 1.2 for a complete description of the related work. However, to the best of our knowledge, previous work has not yet provided a formal setup and mathematical insight about the dynamics of such ledgers in a realistic environment. This paper aims to take one more step towards filling such gap, by focusing on capturing the impact of heterogeneous delays in DAG-based DLs. We focus on deterministic delays, while we believe that results can easily be extended to random delays.

We hereafter refer to messages instead of transactions, since the DLs usage is not restricted to monetary transactions. As aforementioned, in order to attach a new message to the ledger, a number (two in the case of IOTA) of unapproved messages (the so-called *tips*) must be validated. We shall indistinguishably refer to unapproved messages or tips. Finally, due to the distributed nature of the system each node's copy of the register at a given instant might vary from the copy of any other node. We shall refer to the *view* of a node as the messages seen as tips at a given instant

by this node. Indeed, delay can lead a node to believe that a message is a tip even if it has already been approved by another node. Also, it might not be informed of the existence of a new message.

1.1. Contributions

We consider the case where a delay exists between every pair of nodes in the DL. The delay is coming from the underlying peer-to-peer network as well as from the time required to issue, process, and validate messages. The delay is constant, bounded, and potentially different between every pair of nodes. For this scenario:

- We provide a stochastic process model capturing the evolution of the number of tips (and the number of messages seen as tips by each node, called views in this paper) under the presence of heterogeneous delay.
- We deduce an upper bound on the expected number of tips and the expected number of messages seen by each node as tips. While the proposed bounds are loose, they make it possible to conclude that even under heterogeneous delays, these health metrics do not diverge to infinity. The model is based on the study of the evolution of specific stochastic sets of messages (Section 3). We also provide, an alternate analysis based on the asymptotic negative drift property of our stochastic process (see Section 7.3), proving that the expected number of tips is bounded over time.
- For the one-node case with validation delay, and for the multiple-node case with same delay value for all pairs of nodes, we derive closed forms for the expected drift of the number of tips and of the expected drift of the number of messages seen as tips for each node. In addition, for these two scenarios (Section 4):
  - We prove that the evolution of the number of tips (resp. views) can be captured by a Markov chain.
  - We show that the defined Markov chain is ergodic, aperiodic and irreducible.
  - Finally, we derive an upper bound on the tail of the stationary distribution of the number tips.
- Finally, the results are shown to be in good agreement with experimental results obtained from our running testbed (Section 5).

1.2. Related work

Several works have attempted to model performance metrics of DLs, most of them focusing on Blockchain-like DLs (see e.g. [8] and references there in). To the best of our knowledge, formal mathematical models for DAG-based DLs are rather limited in the literature. We hereafter review the state of the art related to performance evaluation of DAG-based DLs.

The first mathematical model of DAG-based ledgers has been introduced in [19] by S. Popov. The first assumption made in this paper is that the arrival of new messages in the ledger is captured by a unique Poisson process. The second assumption is about the delay observed by nodes regarding the state of the distributed ledger. The author assumes that one central node/user maintains the record of the ledger and the rest of the nodes view the state of the ledger (and send transactions to it) by requesting it to the central node. The communication delay is coming from the central server and the other nodes. It is assumed to be the same for every node. Finally, the analysis of the stochastic process used to captured the evolution of the distributed ledger is based on the conjecture that the number of unapproved transactions, at the stationary rate, is not deviating excessively from its average value. The author justifies this conjecture through simulation with homogeneous delays. It can be observed, in our simulations, that when the delays are heterogeneous, this assumption is not satisfied anymore.

This initial work has reinforced the interest on finding mathematical models to capture the performances of DAG-based ledgers. Among them, a lot of works have built on it, such as simulation-based works [12, 15] which try to understand the performances of DAG-based ledgers by simulating the extended versions of the model proposed in [19]. The impact of non-Poisson arrival rate has been also studied, from a mathematical point of view in [13]. Note that all these works are assuming the existence of one central node which maintains the record of the ledger. Moreover, the different theoretical analysis are always under the assumption that the number of unapproved messages, at the stationary rate, is not deviating excessively from its average value. The paper [17] by Penzkofer et al. builds on [19] to incorporate different delays in the system. They assume that multiple different classes of messages exist

(data messages and value messages), resulting in different processing times implying different reception delays. The authors again assume one central node/user maintains the record of the ledger. The different delays are coming from the different classes of messages and not from the fact that nodes are maintaining local copies of distributed ledger, and update their own copy by asking the neighbors the missing information. Moreover, to derive the mathematical results, they assume the above mentioned conjecture to be true.

A new type of theoretical analysis, based on fluid limit approximations, has been proposed in [10, 11]. In [10], the authors prove that the evolution of unapproved messages can be captured, under heavy traffic, by delayed differential equations. This model is even extended to a case where the strategy of selection of unapproved messages is not uniform. In this case, the authors in [11] show that the stochastic process is converging to a partial differential equation, still under a heavy traffic scenario. In this line of work, it is assumed that only one central node/user maintains the record of the ledger. From a more practical perspective, a recent work [23] has proposed a simulator framework for a multi-agent, DAG-based distributed ledgers with delays. In this work, no mathematical formulation is suggested. Even more recently, in [22], a mathematical model based on Markov Processes capturing the growth of such DAG-based DLs was proposed. They do not consider, however, heterogeneous delays among nodes.

We finally would like to mention the different works that try to improve the algorithm behind the selection of unapproved messages, regarding security [3, 1], fairness [4] and by incorporating strategic agents [21]. Such works are built on the previous mentioned mathematical models.

Our paper is extending the model proposed in [2] by Q. Bramas. In this paper, the author models the evolution of the number of unapproved transactions using a discrete time Markov chain, with countable state-space. In this paper, no delay is assumed (except the one produced by the discretization of the time) and only one node (or symmetric nodes) is present in the system. The author has been able to prove that the Markov chain has a positive stationary distribution. One main challenge in incorporating the heterogeneous delays is the fact that it is not possible to prove, as in [2], that the stochastic process is a Markov chain.

Of all those generalizations, to the best of our knowledge, no one has considered to model and to derive stability properties of the evolution of the number of unapproved messages when the nodes are maintaining local copies of distributed ledger and implying that heterogeneous delays are present in the system. Moreover a lot of the cited work rely on a conjecture that the number of unapproved transactions, at the stationary rate, are not deviating excessively from its average value, which has only been proved using simulation in homogeneous scenario.

2. Model

2.1. Mathematical Formulation

Let $\mathcal{I} := \{1, 2, .., I\}$ be the set of nodes participating to the DL. We consider that the state of the DL is evolving on a discrete time $n \in \mathbb{N}_+$. We denote by C_n^i the set of new messages sent into the system by node i at time n. Moreover, let V_n^i be the set of messages viewed as unapproved messages (tips) by node i at time n. We also denote by $D_n^i \subseteq V_n^i$ the set of messages approved by node i at instant n. We assume that $V_0^i = \{0\}, C_0^i = D_0^i = \emptyset$, for all $i \in \mathcal{I}$ and that if a set is negatively indexed then it is the null set. The evolution of the sets C_n^i, D_n^i and V_n^i are captured by the following equations:

$$C_n^i = \left\{s_n^i + 1, s_n^i + 2, \cdots, s_n^i + r_i\right\}, \text{ with } s_n^i := (n-1)(\sum_{j\in\mathcal{I}} r_j) + \sum_{j=1}^{i-1} r_j, \ \forall n \geq 1, \tag{1}$$

$$D_n^i = \bigcup_{k=1}^{2r_i} \{c_k\}, \text{ with } c_k \text{ sampled uniformly (with replacement) from } V_n^i, \ \forall n \geq 1, \tag{2}$$

$$V_n^i = \left(\bigsqcup_{j\in\mathcal{I}} C_{n-1-d^{j,i}}^j\right) \bigsqcup \left(V_{n-1}^i - \bigcup_{j\in\mathcal{I}} D_{n-1-d^{j,i}}^j\right), \ \forall n \geq 1, \tag{3}$$

where:

- for all $i \in \mathcal{I}$ and for all $n \in \mathbb{N}_+$, $r_i \geq 1$ is the number of new messages sent by node i;

- $d^{i,j} \in \left\{0, \ldots, d^*\right\}$, with $d^* \in \mathbb{N}$, is the delay for node j to observe the new messages sent by node i. $d^{i,j}$ is also capturing the delay of node j to observe the messages approved by node i.

In this model, we have chosen to provide a unique identifier to every message sent to the distributed ledger, which explains (1). We assume that every node i, is creating r_i new messages, at every instant n. The interpretation of the evolutions of D_n^i and V_n^i are more complex. At every instant n, in a distributed ledger based on a DAG architecture, we assume that every new message sent from node i validates two messages chosen uniformly, independently and with replacement, in V_n^i. Therefore it is possible that a message validates twice the same message in V_n^i. It is also possible that a message in V_n^i is approved by more than one message from C_n^i. This set of approved messages is therefore equal to D_n^i, according to (2). The messages in C_n^i are tips (new messages in the system at time n) and are thus waiting to be approved. Node j is informed of the existence of this new set of unapproved messages after $d^{j,i}$ units of time starting from n. Moreover, node j is also informed of the fact that messages in D_n^i should not be approved after $d^{j,i}$ units of time starting from n. The last two points complete the explanation of (3), where we have used notation $\bigsqcup$ to indicate union of disjoint sets, and $\bigcup$ for union of possibly not disjoint sets.

For convenience, we also define C_n (resp. D_n) to be the set of messages generated (resp. approved) by all nodes at time step n, i.e. $C_n := \bigsqcup_{i \in \mathcal{I}} C_n^i$ ($D_n := \bigsqcup_{i \in \mathcal{I}} D_n^i$). Please note that such sets do not correspond to the state seen by all the nodes (i.e. there is no such global view in the distributed system), but they are convenient for deriving theoretical results. Let Y_n be the set of messages which are actually tips (i.e. transaction that have not received any validation) at time n. It can be defined as

$$Y_n := \bigcup_{t=0}^{n} C_t - \bigcup_{t=0}^{n} D_t, \qquad n \geq 0. \tag{4}$$

It can be computed recursively as

$$Y_n = C_n \sqcup (Y_{n-1} - D_n).$$

We are interested in the cardinality of the tips set at each time step, noted as X_n and defined as $X_n := |Y_n|$.For some part of our work, we will require the following assumption about the system's delays.

Assumption A: the system's delays satisfies the triangle inequality i.e.: $d^{i,j} + d^{j,k} \geq d^{i,k}$, $\forall i, j, k \in \mathcal{I}$.

This assumption is satisfied if the nodes in the system perform a 'full gossip' algorithm to propagate the state of the distributed ledger. In a non-full gossip set-up, information from a node reaches another node through only one possible way, that is, via the direct link between the two nodes. But in a full gossip set-up, there are multiple ways for information to go between any two nodes. To be precise, let $\tau^{i,j}$ be the delay in the direct route from node i to j. We define effective delay between any node i and j to be, for the full gossip set-up, to be equal to:

$$d^{i,j} = \min\{\sum_{k=1}^{n} \tau^{i_k, i_{k+1}} | i_1 = i, i_{n+1} = j\}, \tag{5}$$

and in the non-full gossip set-up, $d^{i,j} = \tau^{i,j}$. It is easy to see from the above definition that system's (effective) delay satisfies the triangle inequality in the full gossip set-up.

2.2. Model's discussion

In this section we highlight the strengths and limitations of our model. First of all, let us recall that the purpose of the paper is to study the impact of delay on DAG-based DLs. As pointed out in Section 1.2, when delay is considered in a limited way (one central node/user maintains the record of the ledger), the Tangle is stable (i.e. number of tips and validation time remains bounded. See, for instance [21]). In the remaining of the paper we prove that, under a pretty realistic setting, non-divergence to infinity of the system can still be guaranteed. In that sense, one of the main richness of our model is to consider delay between any pair of nodes, such delay can vary from one node to the other. On the other hand, we consider constant values in our model, and we have not taken into account the effect of resolving the proof-of-work. While making those delay random values would provide our model with more realism, that is left for future work. However, please note that the constant values can be set at the worst case of delay between two nodes.

Our model is based on a discrete time, where different tasks (messages sent, message validation) are performed in a synchronised way. Still, the model can be enriched by considering asynchronous events. However, as we shall show in Section 5, our theoretical results match quite well experimental results obtained in a real testbed. We conclude that the impact of this model assumption on the results is minor.

Regarding tips selection algorithm, we have assumed that tips are selected in a uniform way and with replacement. While, it is well understood that the IOTA protocol cannot force a node to use a particular tip selection algorithm, it is a reasonable assumption that in the absence of malicious actors, nodes behaves well (i.e. respect protocols and conventions). In addition, for such scenarios, it has been shown in the literature that uniform random tip selection minimizes the time until first approval of transactions, with respect to other algorithms such as a random walk [12]. Other proposals exist in the literature, as the one in [20], where tips are weighted before uniform selection, which ultimately is claimed to incentive node participation in ledger's maintenance tasks. When it comes to modelling the behaviour of the system, it is common to assume a simple solution such as in [2] and in the present work. Uniform random selection with replacement is also a lightweight solution for nodes, making it a reasonable assumption for real implementations.

Message issuing dynamics is also a relevant model parameter. While it is classical in the literature to consider random time between messages, in particular exponential distributed interarrival time between messages, our model focuses on the amount of messages sent at each time interval, and considers constant values. While it could be interesting to extend results to rate varying as a function of time, our assumption is classical in implemented solutions. As with delay parameter, these constant values can be seen as worst case.

Finally, throughout the paper we focus on the evolution and number of tips (unapproved messages) as a performance metric. In particular, we focus on its boundedness in expectation. Tips can be measured directly from any simulation or implementation, and have been related to the average approval time of a transaction [13]. While more complex measurements could be defined, such as the final time (time when a transaction is consider definitively attached to the tangle) there is no consensus in the literature on how a node can determine such metric in a completely distributed way.

3. General case: finite bound over the expectation of the cardinality of the views set and the tips set

This section is dedicated to our first main result. We derive a finite upper bound on the average number of tips and a finite upper bound on the average number of messages viewed by a node as tips (cardinality of the view sets), for every node, at every instant. Such results show that DAG-based distributed ledgers do not diverge, even in the presence of delays. These analytical results are valid for the general case, that is, when the system has more than one node. The section is organized as follows: we first introduce the recursions on the tips sets and the view sets, and we then compute the upper bound of such quantities. Numerical validation and comparison with state-of-the-art results are addressed in Section 5. The proof of most of the results are presented in Appendix 7.2. The proofs derived in this section are mainly based on the recursion shown in proposition 1, and are not based on classical tools from stochastic processes. It is interesting to note, that an alternate upper bound can also be derived using stochastic processes tools, using assumptions such as bounded increments and negative drift. In addition, for such proof Assumption A is not needed. The reader is referred to Appendix 7.3 for such results.

Let us now recall that C_n^i (resp. D_n^i) is the set of messages created (resp. approved) by node i at instant n. As before, we use $C_n = \sqcup_{i\in\mathcal{I}} C_n^i$, $D_n = \cup_{i\in\mathcal{I}} D_n^i$, $x_n = \mathbb{E}[X_n]$, $v_n^i = \mathbb{E}[|V_n^i|]$ as shorthand notations. Note that the above expectations are only conditioned with respect to the initial condition of the system ($V_0^i = \{0\}$, $C_0^i = D_0^i = \emptyset$).

Proposition 1. *We assume that $V_0^i = \{0\}, C_0^i = D_0^i = \emptyset$, for all $i \in \mathcal{I}$. If assumption A is satisfied, for every i and every n, we have the different recursions:*

$$v_n^i \leq \sum_{j\in\mathcal{I}} \sum_{m=d^{j,i}+1}^{n} r_j \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} (1 - \frac{1}{v_{n-m+l}^k + a + 1})^{2r_k}$$

and

$$x_n \leq \sum_{m=0}^{n} \sum_{j\in\mathcal{I}} r_j \prod_{l=1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} (1 - \frac{1}{v_{n-m+l}^i + a + 1})^{2r_i},$$

where $a = \max_i r_i + 1$.

Proof. The proof is available in the Appendix 7.2. □

We now show the existence of an upper bound on the iterates introduced in proposition 1. Recall $d^* = \max_{i,j} d^{i,j}$ is the maximum delay and $r = \sum_i r_i$ is total rate of the system.

Theorem 1. *If assumption A is satisfied and if we assume that* $V_0^i = \{0\}, C_0^i = D_0^i = \emptyset$*, for all* $i \in \mathcal{I}$ *then the* $X_n, |V_n^i|$ *are bounded in expectation for every n and i. That is,*

$$x_n \leq 3rd^* + \frac{3r}{2} + a + 1 + \frac{r}{4d^*}, \qquad v_n^i \leq B, \qquad \forall i, n,$$

where $B = 4rd^* + 2r + a + 1$*, and* $a = \max_i r_i + 1$.

Proof. We first prove the bound on v_k^i and then on x_k. The proof for the bound of v_k^i is based on an induction argument. *View iterates bound:* As mentioned above, we prove that $v_k^i \leq B$ by using an induction argument. The initialization is trivial. We assume that $v_n^i \leq B$ for all $i \in \mathcal{I}$ and $k \leq n$. We need to prove $v_{n+1}^i \leq 2rd^* + \frac{(B+a+1)^2}{2(B+a+1-r)} \leq B$, for all i. Let $b := (1 - \frac{1}{B+a+1})^2$. Using the recursion in proposition 1,

$$\begin{aligned}
v_n^i &\leq \sum_{j\in\mathcal{I}} \sum_{m=d^{j,i}+1}^{n} r_j \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} (1 - \frac{1}{v_{n-m+l}^k + a + 1})^{2r_k} \\
&\leq \sum_{j\in\mathcal{I}} \sum_{m=d^{j,i}+1}^{n} r_j \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k}, \qquad \text{(using assumption } v_k^i \leq B, \forall i, k \leq n) \\
&\leq \sum_{j\in\mathcal{I}} \sum_{m=1}^{n} r_j \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k}
\end{aligned}$$

The last term can be rewritten as:

$$\sum_{j\in\mathcal{I}} r_j [\sum_{m=1}^{2d^*} \underbrace{\prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k}}_{\leq 1} + \sum_{m=2d^*+1}^{n} \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k}],$$

implying

$$\begin{aligned}
v_n^i \leq & 2rd^* + \sum_{j\in\mathcal{I}} r_j \sum_{m=2d^*+1}^{n} \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k} \\
= & 2rd^* + \sum_{j\in\mathcal{I}} r_j \sum_{m=2d^*+1}^{n} \underbrace{\prod_{l=1}^{d^*} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k}}_{\leq 1} \prod_{l=d^*+1}^{m-d^*-1} \underbrace{\prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}}}_{=\mathcal{I}} b^{r_k} \\
& \times \underbrace{\prod_{l=m-d^*}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} b^{r_k}}_{\leq 1} \leq 2rd^* + \sum_{j\in\mathcal{I}} r_j \sum_{m=2d^*+1}^{n} \prod_{l=d^*+1}^{m-d^*-1} b^r \\
\leq & 2rd^* + r \sum_{m=2d^*+1}^{\infty} b^{r(m-2d^*-1)} \leq [2rd^* + \frac{(B+a+1)^2}{2(B+a+1-r)}], \qquad \text{(using lemma 3).}
\end{aligned}$$

We can conclude the first part of the proof using Lemma 4.

Tips iterates bound: From Proposition 1, using the fact that $v_k^i \leq B$ we have

$$x_n \leq \sum_{m=0}^{n} \sum_{j\in\mathcal{I}} r_j \prod_{l=1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} (1 - \frac{1}{v^i_{n-m+l} + a + 1})^{2r_i} \leq \sum_{m=0}^{n} \sum_{j\in\mathcal{I}} r_j \prod_{l=1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} b^{r_i},$$

$$= \sum_{j\in\mathcal{I}} r_j [\sum_{m=0}^{d*-1} \underbrace{\prod_{l=1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} b^{r_i}}_{\leq 1} + \sum_{m=d*}^{n} \prod_{l=1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} b^{r_i}]$$

$$\leq \sum_{j\in\mathcal{I}} r_j [d^* + \sum_{m=d*}^{n} \prod_{l=1}^{d*} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} b^{r_i} \prod_{l=d*+1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} b^{r_i}]$$

The fact that $\prod_{l=1}^{d*} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} b^{r_i} < 1$ is then implying:

$$x_n \leq \sum_{j\in\mathcal{I}} r_j [d^* + \sum_{m=d*}^{n} \prod_{l=d*+1}^{m} \prod_{i\in\underbrace{\{i\in\mathcal{I}|l>d^{j,i}\}}_{=\mathcal{I}}} b^{r_i}]$$

$$= r[d^* + \sum_{m=d*}^{n} \prod_{l=d*+1}^{m} b^r] \qquad \leq r[d^* + \sum_{m=d*}^{\infty} \prod_{l=d*+1}^{m} b^r]$$

$$\leq r[d^* + \frac{1}{1-b^r}] \leq r[d^* + \frac{(B+a+1)^2}{2r(B+1-r)}] \qquad \text{(using lemma 3)}$$

$$= rd^* + \frac{(B+a+1)^2}{2(B+a+1-r)} \leq 4rd^* + \frac{5r}{2} + 1 + \frac{r}{12d^*} \qquad \text{(using Lemma 5).}$$

The last inequality concludes our proof. □

Lemma 1. *For all i and for all $n \geq \max_{i,j} d^{i,j} + 2$ the view set and tips are lower bounded as follows:*

$$|V_n^i| \geq \sum_j r_j = r, \qquad \forall i, \qquad\qquad X_n \geq \sum_i r_i (\min_j d^{ij} + 1).$$

Proof. From the definition (3), we have

$$V_n^i = (\sqcup_{j\in\mathcal{I}} C^j_{n-1-d^{j,i}}) \bigsqcup (V^i_{n-1} - \cup_{j\in\mathcal{I}} D^j_{n-1-d^{j,i}}) \implies \sqcup_{j\in\mathcal{I}} C^j_{n-1-d^{j,i}} \subset V_n^i \implies \sum_j |C^j_{n-1-d^{j,i}}| = \sum_j r_j \leq |V_n^i|.$$

$n \geq \max_{i,j} d^{i,j} + 2$ ensures that the indices $n - 1 - d^{i,i}$ in the above equations are positive, which concludes the proof for V_n^i. Regarding the tips lower bound. We can observe that

$$Y_n = \bigcup_{t=0}^{n} C_t - \bigcup_{t=0}^{n} D_t, \qquad n \geq 0.$$

Notice that if $c \in C_m^i$ then $c \notin D_k^j, \forall k \leq m + d^{i,j} + 1$, implying that $c \notin D_k, \forall k \leq m + \min_j d^{i,j} + 1$. Therefore $\forall c \in C_m^i$, $\forall m \geq n - \min_j d^{i,j} + 1$, $c \in Y_n$ and the result follows. □

As the proof suggest, the above bounds are loose. For the interested reader, we point out the weaknesses in the proof. In the views recursion (3): we split summation over m into two terms (i.e $\sum_{m=0}^{2d^*} \cdot + \sum_{m=2d^*+1}^{n} \cdot$). We can improve the bound on the first term, if we analyze terms inside the sum node by node. For the sake of simplicity, we have assumed that the distributed ledger, at instant 0, only contains one tip ($V_0^i = \{0\}, C_0^i = D_0^i = \emptyset$). However, this condition is not as restrictive as it seems and the above results hold for any initial condition. That is, iterates remains bounded in expectation for all initial conditions. Of course, the exact bound will change. Note also that if we start with any tips less than B, then above results and their proofs hold without any change. We illustrate this different remarks in Section 5.

4. One-node case: stationary probability distribution of the number of tips

The mathematical analysis of the previous section was about deriving properties of the average evolution of X_n, without using tools from stochastic processes such as Markov chain or the theory of martingales. In this section, we demonstrate that it is possible, when there is only one node in the distributed ledger, to cast the evolution of X_n as the evolution of a specific Markov chain with a countable state space. We first derive a general formula for the expected number of tips, at time n with respect to the appropriate σ-field. Then we show that for the one node case, using the Foster Theorem, the studied Markov chain is ergodic, admits a unique stationary probability distribution and that the stationary probability distribution has an exponential tail. Note that in the one-node case, there are no delays related to the underlying peer-to-peer network, delays account mostly for validation delay.

First, for every $S \in \mathcal{P}(\mathcal{I})$, where $\mathcal{P}(\mathcal{I})$ is the power set of $\mathcal{I}$, we define the disjoint partition $E_n(S)$ of Y_{n-1} as:

$$E_n(S) := \left\{ s \in Y_{n-1} | s \in V_n^i, \forall i \in S, s \notin V_n^j, \forall j \notin S, i, j \in \mathcal{I} \right\}. \tag{6}$$

$E_n(S)$ can be interpreted as the set of messages in Y_{n-1} that are only visible to nodes in S. We can observe that $E_n(\cdot)$ satisfies the following recursion relations:

$$E_n(S) = \left(\bigcap_{i \in S} V_n^i \right) \bigcap (Y_{n-1}) - \bigcup_{S \subsetneq S'} E_n(S'),\ E_n(\mathcal{I}) = \left(\bigcap_{i \in \mathcal{I}} V_n^i \right) \bigcap (Y_{n-1}) \text{ and } E_n(\emptyset) = Y_{n-1} - \bigcup_{i=1}^{I} V_n^i.$$

For every n, let $\mathcal{F}_n = \sigma(C_k^i, D_k^i, k \leq n-1, i \in \mathcal{I})$ be the σ-field which encodes the information available at time n. In the next proposition, we characterize the expected number of tips with respect to $\mathcal{F}_n$.

Proposition 2. *If the distributed ledger is only having a single node $\mathcal{I} = \{1\}$, with a sending rate r and a delay d then the expected number of tips at instant n, with respect to $\mathcal{F}_n$ is equal to:*

$$\mathbb{E}[X_n | \mathcal{F}_n] = (d+1)r + (X_{n-1} - dr)(1 - \frac{1}{X_{n-d-1}})^{2r}. \tag{7}$$

Moreover, for $|\mathcal{I}| \geq 1$, the expected number of tips at instant n, with respect to $\mathcal{F}_n$ is equal to:

$$\mathbb{E}[X_n | \mathcal{F}_n] = \sum_{i=1}^{I} r_i + \sum_{S \in \mathcal{P}(\mathcal{I})} |E_n(S)| \prod_{i \in S} (1 - \frac{1}{|V_n^i|})^{2r_i}. \tag{8}$$

Proof. The first part of the proof is dedicated to the proof of (8). Then the second part is simply an application of (8) to the one node case.

Proof of (8)*:* We start by rewriting (4) as $Y_n = C_n \bigsqcup (Y_{n-1} - D_n)$. From the definition of (1), we have $X_n = |C_n| + |Y_{n-1} - D_n| = \sum_{i=1}^{I} r_i + |Y_{n-1} - D_n|$. Let s be a message in Y_{n-1} , the probability of $s \in Y_n$ is the same as the probability of s not belonging to any D_n^i. We get,

$$X_n = \sum_{i=1}^{I} r_i + \sum_{s \in Y_{n-1}} \mathbb{1}(s \notin \bigcup_{i=1}^{I} D_n^i).$$

Recall the creation of D_n^i, all nodes act independently, therefore

$$X_n = \sum_{i=1}^{I} r_i + \sum_{s \in Y_{n-1}} \prod_{i \in \mathcal{I}} \mathbb{1}(s \notin D_n^i).$$

Let $\mathcal{F}_n = \sigma(C_k^i, D_k^i | k \leq n-1, i \in \mathcal{I})$ be the σ-algebra which encodes the information available at time n. Using the linearity of the expectation, we get

$$\mathbb{E}[X_n | \mathcal{F}_n] = \sum_{i=1}^{I} r_i + \sum_{s \in Y_{n-1}} \prod_{i \in \mathcal{I}} \mathbb{P}(s \notin D_n^i).$$

Moreover, using the disjoint partition $\{E_n(S)\}$ of Y_{n-1}, we obtain:

$$\mathbb{E}[X_n|\mathcal{F}_n] = \sum_{i=1}^{I} r_i + \sum_{S\in\mathcal{P}(\mathcal{I})} \sum_{s\in E_n(S)} \prod_{i\in\mathcal{I}} \mathbb{P}(s \notin D_n^i).$$

Observe that $s \in E_n(S)$ can't belong to D_n^i for $i \notin S$ because s is only viewed by nodes in S (in the sense that $s \notin V_n^i, \forall i \notin S, s \in E_n(S)$). If a node can't view a tip, it can't attach its messages to it. If $s \in V_n^i$, then the event $s \notin D_n^i$ has the same probability of non-selection of a particular tip, uniformly, and $2r_i$ (independently and with replacement) in V_n^i. Therefore we have:

$$\mathbb{P}(s \notin D_n^i) = \begin{cases} 1 & if \quad s \notin V_n^i \\ (1 - \frac{1}{|V_n^i|})^{2r_i} & if \quad s \in V_n^i \end{cases}$$

Recall that if $s \in E_n(S)$ implies $s \in V_n^i, \forall i \in S$, and $s \notin V_n^j, \forall j \notin S$. We have

$$\mathbb{E}[X_n|\mathcal{F}_n] = \sum_{i=1}^{I} r_i + \sum_{S\in\mathcal{P}(\mathcal{I})} \sum_{s\in E_n(S)} \prod_{i\in S} (1 - \frac{1}{|V_n^i|})^{2r_i}.$$

The last equality concludes the first part of the proof.

Proof of (7)*:* Recall from (1), (4) and (3) that

$$V_n = \bigcup_{t=0}^{n-1-d} C_t - \bigcup_{t=0}^{n-1-d} D_t,\ Y_n = \bigcup_{t=0}^{n} C_t - \bigcup_{t=0}^{n} D_t,\ C_n = \{(n-1)r+1, (n-1)r+2, ..., nr\}.$$

The relation $V_n = Y_{n-1-d}$ is direct from the previous definitions. In this case there are only two subsets of $\mathcal{I}$ namely $\emptyset$ and $\{1\}$. Let $E_n(\emptyset), E_n(\{1\})$ be the disjoint partition of Y_{n-1} according to (6). $\mathcal{F}_n = \sigma(X_k|k \le n)$ is the sigma algebra which contains all the information till time n. Equation (8) gives us

$$\mathbb{E}[X_n|\mathcal{F}_n] = r + \sum_{S\in\mathcal{P}(\mathcal{I})} |E_n(S)| \prod_{i\in S} (1 - \frac{1}{|V_n|})^{2r}.$$

From above discussion, we get

$$\mathbb{E}[X_n|\mathcal{F}_n] = r + |E_n(\emptyset)| \prod_{i\in\emptyset} (1 - \frac{1}{|V_n|})^{2r} + |E_n(\{1\})| \prod_{i\in\{1\}} (1 - \frac{1}{|V_n|})^{2r}.$$

Using the convention that $\prod$ over null set is 1, it simplifies to

$$\mathbb{E}[X_n|\mathcal{F}_n] = r + |E_n(\emptyset)| + |E_n(\{1\})|(1 - \frac{1}{|V_n|})^{2r}.$$

We now compute $|E_n(\emptyset)|$ and $|E_n(\{1\})|$. Recall that $E_n(\emptyset)$ is the set of tips in Y_{n-1} which node 1 has not seen at time n, in other words:

$$E_n(\emptyset) = Y_{n-1} - V_n = (\bigcup_{t=0}^{n-1} C_t - \bigcup_{t=0}^{n-1} D_t) - (\bigcup_{t=0}^{n-1-d} C_t - \bigcup_{t=0}^{n-1-d} D_t) = \bigcup_{t=n-d}^{n-1} C_t \implies |E_n(\emptyset)| = \sum_{t=n-d}^{n-1} |C_t| = dr.$$

Since, $|E_n(\emptyset)|$ and $|E_n(\{1\})|$ are disjoint partition sets of Y_{n-1},

$$|E_n(\{1\})| = |Y_{n-1}| - |E_n(\emptyset)| = |X_{n-1}| - dr.$$

Hence the (7) has been proved. □

In the case of one node, we have a stochastic process $\{X_n, n \geq 0\}$ which satisfies the recursion (7). Observe that X_n depends on X_{n-1} and X_{n-d-1}. To convert the above stochastic process into a standard Markov chain, we consider a new state space $Z_n := (X_n, X_{n-1}, ..., X_{n-d})^T$. We define the following set:

$$\mathcal{A} := \{(x_0, \cdots, x_d) | \mathbb{P}\left((X_n = x_0, \cdots, X_{n-d} = x_d) | X_0 = r(d+1)\right) > 0, \text{for some } n \geq 0\},$$

where as a convention, negatively indexed variable $X_m, m < 0$ shall denote X_0. We restrict the Markov chain to the set $\mathcal{A}$ to make sure that all the states are visited infinitely often. We also fix $X_0 = r(d+1)$ because, if we start from $x < r(d+1)$ tips, the distributed ledger never returns to x tips again. From structure of the problem we know that our Markov chain makes bounded jump i.e $|X_{n+1} - X_n| \leq r, \forall n$. It is then clear from the bounded jumps and the lower bound of the Markov chain $\{X_n, n \geq 0\}$ that

$$\mathcal{A} \subseteq \{(x_0, \cdots, x_d) | x_i \geq (d+1)r, |x_i - x_{i+1}| \leq r, \forall i\}.$$

Let $q : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$ be the transition kernel of Markov chain $\{Z_n, n \geq 0\}$, defined as

$$q(Z_{n+1} = \hat{z} | Z_n = z) = p(\hat{z}^1 | z^1, z^{n+d+1}) \delta(z^1 = \hat{z}^2) \delta(z^2 = \hat{z}^3) \cdots \delta(z^{n-d-1} = \hat{z}^{n-d}),$$

where $z, \hat{z} \in \mathcal{A}$ and a^k denotes the k^{th} coordinate of $a \in \mathcal{A}$. The second main result of our paper is the following theorem.

Theorem 2. *The Markov chain $\{Z_n\}$ satisfies the following properties:*

1. *Let $f : \mathcal{A} \to \mathcal{R}$ be a non-negative Lyanpunov function, defined as the projection on first coordinate, i.e. $f(Z) := Z^1$. Let $A := \{Z \in \mathcal{A} | Z^{d+1} < 4dr + 2r\}$, where Z^{d+1} denotes the $(d+1)$-th component of vector Z, be a finite subset of $\mathcal{A}$ and $\mathcal{F}_n := \sigma(Z_k, k \leq n) = \sigma(X_k, k \leq n)$ be a sigma algebra containing all information until time n. There exists $\epsilon > 0$ such that for any n,*

$$\mathbb{E}[f(Z_{n+1}) - f(Z_n) | \mathcal{F}_n] < -\epsilon, \qquad \forall Z_n \notin A,$$

and

$$\mathbb{E}[f(Z_{n+1}) | \mathcal{F}_n] < \infty, \quad \forall Z_n \in A, \qquad \sup |f(Z_{n+1}) - f(Z_n)| \leq r.$$

2. *The Markov chain $\{Z_n, n \geq 0\}$ is irreducible and aperiodic.*
3. *The Markov chain $\{Z_n, n \geq 0\}$ is ergodic. This is implying that the Markov Chain admits a unique stationary probability distribution, denoted α.*
4. *There exists $C_3, \beta_3 > 0$ such that $\alpha(z) \leq C_3 e^{-\beta_3 f(z)}$, $\forall z \in \mathcal{A}$.*

Proof. We prove each property separately.

(1) Let $Z_n \notin A$ then

$$\begin{aligned}
\mathbb{E}[f(Z_{n+1}) - f(Z_n) | \mathcal{F}_n] &= \mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] \overset{\text{from (7)}}{=} (d+1)r + (X_n - dr)(1 - \frac{1}{X_{n-d}})^{2r} - X_n \\
\overset{inclusion-exclusion}{\leq} & (d+1)r + (X_n - dr)[1 - \frac{2r}{X_{n-d}} + \frac{r(2r-1)}{(X_{n-d})^2}] - X_n = r[1 - \frac{2(X_n - dr)}{X_{n-d}} + \frac{(2r-1)(X_n - dr)}{(X_{n-d})^2}] \\
\leq & r[1 - \frac{2(X_{n-d} - 2dr)}{X_{n-d}} + \frac{(2r-1)(X_n - dr)}{(X_{n-d})^2}] \qquad (\text{from } X_n \geq X_{n-d} - dr) \\
= & r[-1 + \frac{4dr}{X_{n-d}} + \frac{(2r-1)}{(X_{n-d})} \frac{(X_n - dr)}{(X_{n-d})}] \\
\leq & r[-1 + \frac{4dr}{X_{n-d}} + \frac{2r-1}{X_{n-d}}] \qquad (\text{from } X_n \leq X_{n-d} + dr) \\
\leq & r[-1 + \frac{4dr + 2r - 1}{4dr + 2r}] \qquad (Z_n \notin A \text{ therefore } Z_n^{d+1} = X_{n-d} \geq 4dr + 2r) \\
= & \frac{-r}{4dr + 2r}
\end{aligned}$$

By taking ϵ equal $\frac{r}{4dr+2r}$ we obtain the first part. The second part is trivial from the fact that our Markov chain has bounded jump.

(2) We first prove the irreducibility of the Markov chain and then we focus on the aperiodicity.
Irreducibility: By definition, all the states of $\mathcal{A}$ can be reached from the target state $z' = ((d+1)r, \cdots, (d+1)r)) \in \mathcal{A}$. To prove the irreducibility of the Markov chain $\{Z_n, n \geq 0\}$, it is enough to prove that it is possible to reach a target z' from any state $z \in \mathcal{A}$, that is, there exists $n(z) \geq 0, z \in \mathcal{A}$

$$\mathbb{P}(Z_{n(z)} = z' | Z_0 = z) > 0, \qquad \forall z \in \mathcal{A}.$$

Here, we illustrate a way to reach the target state z' from any state $z \in \mathcal{A}$. The key ideas are : a) r(d+1) is the minimum of tips possible in the system. And if the system maintains it for consecutive $d+1$ time, then the Markov chain reaches state z'. b) Now, no matter in which state system is, if it starts decreasing tips then the state z' shall be reached. c) To do that, D_{n+1} must choose its parents from $Y_n \cap V_{n+1}$ (maximum distinct parents as possible). Indeed, by a simple manipulation of sets, we get

$$Y_{n+1} = (\sqcup_{k=n-d+1}^{n+1} C_k) \sqcup (Y_n \cap V_{n+1} - D_{n+1}), \qquad Y_{n+1} = (\sqcup_{k=n-d+1}^{n} C_k) \sqcup (Y_n \cap V_{n+1})$$

where $Y_n \cap V_{n+1} = \cup_{k=0}^{n-d} C_k - \cup_{k=0}^{n} D_k$. Observe that $|Y_n| = |(\cup_{k=n-d+1}^{n} C_k)| + |(Y_n \cap V_{n+1})| = rd + |(Y_n \cap V_{n+1})| \geq r(d+1)$, implies $|(Y_n \cap V_{n+1})| \geq r$. And $|Y_{n+1}| = |(\sqcup_{k=n-d+1}^{n+1} C_k)| + |(Y_n \cap V_{n+1}| - |(Y_n \cap V_{n+1} \cap D_{n+1})| = r(d+1) + |(Y_n \cap V_{n+1}| - |(Y_n \cap V_{n+1} \cap D_{n+1})|$. Recall, D_{n+1} is chooses $2r$ messeages from V_{n+1}. So, we allow D_{n+1} to choose as many $(\min(2r, |(Y_n \cap V_{n+1}|))$ distinct parents possible from $Y_n \cap V_{n+1}$. So, $|Y_{n+1}| < |Y_n|$ if $|Y_n| > r(d+1)$, and $|Y_{n+1}| = |Y_n|$ if $|Y_n| = r(d+1)$. This concludes the proof. Hence, the irreducibility is established.
Aperiodicity: We turn our attention to the proof of aperiodicity. Let $z' \in \mathcal{A}$ be the initial state as defined above. Let $Z_n = z'$, this implies that $Y_n = \cup_{k=n}^{n-d} C_k$. Note that $V_{n+1} = C_{n-d}$, therefore there is a non-zero probability that $D_{n+1} = C_{n-d}$. Therefore we have $\mathbb{P}(Z_{n+1} = z' | Z_n = z') > 0$. We had already established that z' is reachable from every state in $\mathcal{A}$. Combining these two facts, we immediately get that every state in $\mathcal{A}$ must have a period of 1.

(3) We have already proved the all the conditions from Theorem 2.2.3 in [9], therefore the ergodicity is established. In our case, the ergodicity means that every state is positive recurrent (by definition) which implies the existence of a unique stationary probability distribution π (Theorem 5.5.12 in [7]).

(4) This type of result is not surprising for this type of systems. Much stronger results exist for much general systems, for example see [9], especially chapter 7. It can be proved using techniques from [9]. □

We end this section with a example extending the one-node case to the multiple-node case, under specific delay conditions.

Example 1. *(Multiple nodes with equal delay) In a I-node system, if the delay is constant, that is $D^{i,j} = d, \forall i, j$, then the system behaves like a one-node system with rate $\sum_i r_i$ and delay d.*

Proof. In this setting, $V_n^i = Y_{n-1-d}$ for all $i \in \mathcal{I}$, $E_n^{\emptyset} = d \sum_i r_i$, $E_n^{\mathcal{I}} = Y_{n-1} - d \sum_i r_i$, and $E_n^{\mathcal{I}} = \emptyset \quad \forall S \in \mathcal{P}(\mathcal{I}), S \neq \emptyset, \mathcal{I}$. Using these facts in lemma 8, we get the following recursion:

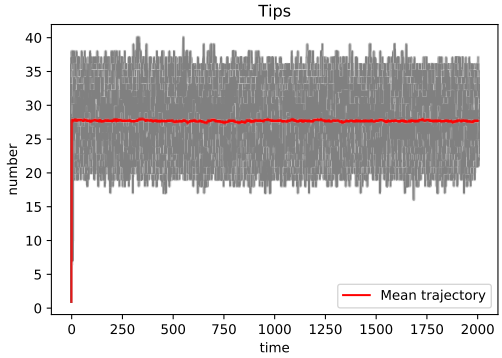
$$\mathbb{E}[X_n | \mathcal{F}_n] = (d+1) \sum_i r_i + (X_{n-1} - d \sum_i r_i)(1 - \frac{1}{X_{n-d-1}})^{2 \sum_i r_i},$$

which is the same as the recursion for one node with rate $\sum_i r_i$ and delay d system.

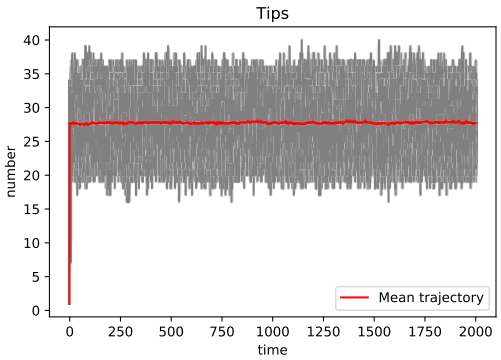
□

5. Empirical validation

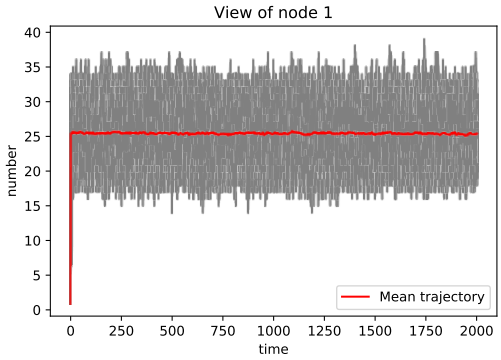
To validate our mathematical model we rely on two approaches: (1) the simulation of the model and (2) an experimental testbed of IOTA Tangle nodes where we apply network delays. In this section, we describe both approaches and we compare their results regarding the amount of concurrent tips (tipsCount) through time. Also, we compare our work with existing results in the literature.



(a) Tips count (X_n) as a function of time, with initial conditions at zero values. Parameters of Monte Carlo simulations: $r = 1$, $I = 3$, $d^{i,j} = 8$ and $d^{i,i} = 1$. 500 simulations



(b) Tips count (X_n) as a function of time. Each run starts from random initial conditions.



(c) View count for node 1 ($|V_n^1|$) as a function of time. Each run starts from random initial conditions.

Figure 2: Monte Carlo simulations: Parameters: $r = 1$, $I = 3$,v$d^{i,j} = 8$ and $d^{i,i} = 1$. 500 simulations and different initial conditions. The system is shown to be stable, and the effect of initial conditions negligible.

5.1. Stochastic model simulation

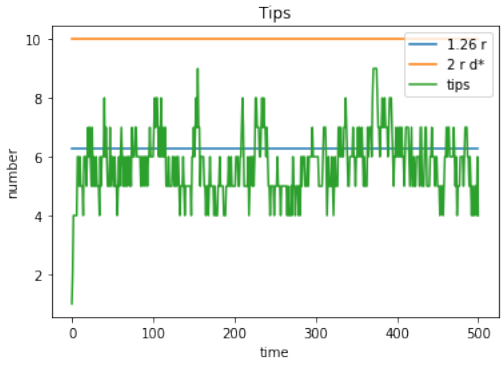
We have implemented the stochastic process captured by (1), (2) and (3). Our code was developed in Python and regular data manipulation tools, and it is available at [Removed for blind-review]. We now summarize some numerical findings, for illustrative purposes, as well as for providing more insight on the dynamics of the simulation system.

We first evaluate results with a 3-node network ($I = 3$), with delay equal to eight units of time between different nodes ($d^{i,j} = 8\ \forall i, j \in \mathcal{I},\ i \neq j$) and equal to one unit within the same node ($d^{i,i} = 1\ \forall\, i \in \mathcal{I}$). We have considered different initial conditions, and run each simulation scenario 500 times.

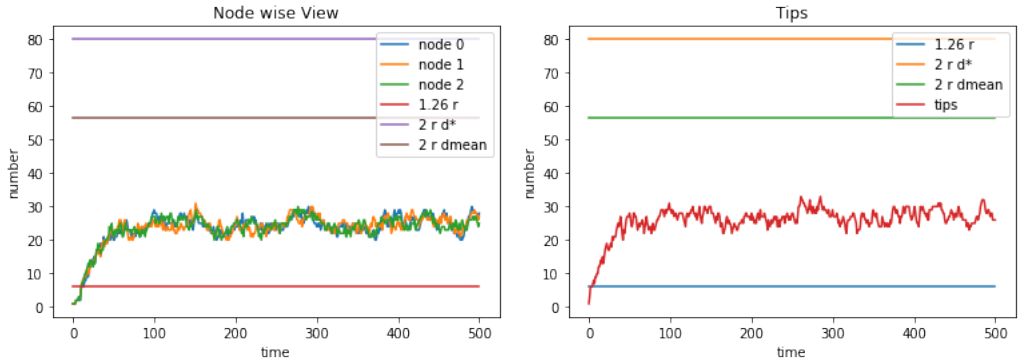
Figure 2a considers initial conditions at zero values ($V_0^i = \{0\}$, $C_0^i = D_0^i = \emptyset$). We can first observe that, as claimed by our results, the number of tips remains bounded. Moreover, the trajectories do not deviate too much form the mean value.

We then analyze the impact of the initial conditions by randomly setting them at each run. Figure 2b and Figure 2c show the number of tips and views, respectively. We can observe how the effect of the initial conditions vanishes out as the system converges.

Figure 3 shows simulation results for a 3-node network compared against results in the literature for different delay values. We first consider a scenario with a delay between different nodes equal to one unit, and no delay within each node itself. This makes it possible to compare the results with the proposals made by Popov in [21] and by Bramas in [2]. Indeed, Bramas considers no delay other than that one imposed by time discretization (so one unit), finding $\mathbb{E}[X_n] \sim 1.26r$. While Popov's continuous time model considers an homogeneous constant value for delay between any node in the network (say d^*), finding that $\mathbb{E}[X_n] = 2rd^*$. Figure 3a shows results obtained for our model



(a) Tips count for delay equal to one for every node. ($d_{i,j} = 1 \,\forall\, i, j \in \mathcal{I}$). Our results agree with the literature's.



(b) Views (left) and Tips (right) count for delay equal to 8 between every node ($d_{i,j} = 8 \,\forall\, i, j \in \mathcal{I}$) and equal to 1 within same node ($d^{i,i} = 1, \forall\, i \in \mathcal{I}$). The model not taking into account delay [2] underestimates the number of tips, while the model considering the same maximum communication delay for all nodes [21] overestimates it.

Figure 3: Comparison with State-of-the-Art results for different values of delay. As expected, results match when delay is equal to one unit.

and for such models of the literature. We can observe that our model perfectly matches the model in [2], which is not surprising since our model is an extension of theirs. Naturally, the continuous time model proposed in [21] provides a slightly higher bound as the formula differs from Bramas' just in a multiplicative constant greater than 1. The difference with Popov's model can, however, be intrinsic to the different settings (continuous vs discrete time). Figure 3b considers a case with delay between different nodes equal to 8 units. We observe that Bramas' model underestimates the number of tips, while Popov's model provides a higher bound.

5.2. Comparison with testbed results

To validate our model against a Tangle ledger, we use a network of virtualised Goshimmer nodes (the in-development decentralised version of IOTA) relying on docker containers. Also, we set up network delay on each node using a network emulator for docker containers (`https://alexei-led.github.io/post/pumba_docker_netem/`). This network emulator is based on the Linux stack of tools, such as tc. This means the applied delay concerns the outbound traffic on each node's network interface. Our network of Goshimmer nodes is composed of up to eight similar containers, each of them has exclusively assigned two CPU cores of a Intel Xeon @ 2.4Ghz CPU. The data and code to analyse these tests can be found at the same URL listed above in Section 5.1.

The differences between the model implementation and the tangle testbed are the following: (1) The implemented simulation does not execute any proof-of-work to attach new messages. (2) In the simulation there is neither message handling nor related overhead produced by network packet processing in the operating system. This means that, apart from the applied delay, the time needed to attach a new message in the simulated DL is negligible. Also, (3) for convenience, we have run the simulation code on personal workstations instead of on the same machine hosting the tangle testbed, implying that the computing resources are different.

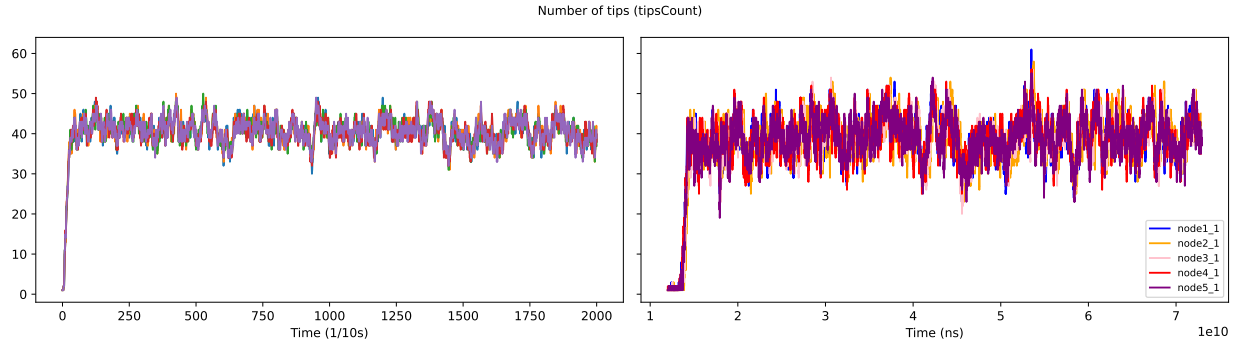


Figure 4: Number of concurrent tips versus time, according to Model simulation (left) and Tangle testbed (right)

nb_of_nodes	pow	mpm	delay_time	median TB	variance TB	median simulation	variance simulation
8	10	600	100	6.0	3.500	9.0	1.236
8	10	600	200	21.0	10.095	37.0	5.810
8	10	600	400	33.0	21.211	63.0	13.611
8	10	600	800	101.0	55.822	110.0	21.449

Table 1: Median and variance of tip counts produced by the tangle testbed (TB) and our model simulation

For comparison's sake, we run the simulated model and testbed experiments using comparable parameters. We consider a setup of five nodes and 500ms node-to-node delay. In order to have similar number of messages being attached to the DL, we set the rate, in the simulation, to 600 messages per minute (MPM), and no Proof-of-Work difficulty (powd), while in the testbed we set rate to 800 MPM and powd to 10 (in a range from 0 to 22 with increasing difficulty. Figure 4 shows the resulting tip count through time, as seen by each node, in the simulation (left) and in the testbed (right). Results show that, despite the aforementioned differences between the implemented model and the testbed, we have been able to find a testbed setup that produces results quite similar to simulation results.

We then run different scenarios namely different delay values across the pair of nodes. To control the number of variables in the experiment, we have arbitrarily fixed rate to 600 MPM, and powd to 10. We consider 8 nodes.Same values were used to run the simulation of the implemented model. Results are shown in Table 1. Once again we observe that results of simulation and testbed quite match.

6. Conclusion

In this paper, we proposed a new mathematical model for capturing the behavior of DAG-based DL under the presence of heterogeneous delays between nodes. The first main result, derived through two alternative methods, (manipulating stochastic sets of messages and a drift-based analysis) proved the existence of an upper bound on the expected number of tips and on the expected number of messages seen as tips by each node (named *views* in this paper). Moreover, we were able to deduce an upper bound on such quantities. The second main result is regarding the one-node case with validation delay, and for the multiple-node case with same delay value for all pairs of nodes. We proved that the evolution of the number of tips (resp. views) can be captured by a Markov chain with a countable state space. Moreover, we showed that the chain is ergodic, aperiodic and irreducible, implying in this case that the Markov Chain admits a unique stationary probability distribution and this distribution has an exponential tail. We have finally made extensive simulations to verify our results and we also observe that analytical results are in very good agreement with experimental results obtained from our running testbed. The natural extensions of this work are to incorporate random delays and study different regimes (such as the mean-field regime).

References

[1] Attias, V., Bramas, Q.: How to choose its parents in the tangle. In: International Conference on Networked Systems. pp. 275–280. Springer (2019)

[2] Bramas, Q.: The Stability and the Security of the Tangle (Apr 2018), `https://hal.archives-ouvertes.fr/hal-01716111`, working paper or preprint

[3] Bramas, Q.: Efficient and secure tsa for the tangle. In: International Conference on Networked Systems. pp. 161–166. Springer (2021)

[4] Bu, G., Gürcan, Ö., Potop-Butucaru, M.: G-iota: Fair and confidence aware tangle. In: IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). pp. 644–649. IEEE (2019)

[5] Churyumov, A.: Byteball: A decentralized system for storage and transfer of value. URL https://byteball. org/Byteball. pdf (2016)

[6] Croman, K., Decker, C., Eyal, I., Gencer, A.E., Juels, A., Kosba, A., Miller, A., Saxena, P., Shi, E., Gün Sirer, E., Song, D., Wattenhofer, R.: On scaling decentralized blockchains. In: Clark, J., Meiklejohn, S., Ryan, P.Y., Wallach, D., Brenner, M., Rohloff, K. (eds.) Financial Cryptography and Data Security. p. 106–125. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)

[7] Durrett, R.: Probability: Theory and Examples. Thomson (2005)

[8] Fan, C., Ghaemi, S., Khazaei, H., Musilek, P.: Performance evaluation of blockchain systems: A systematic survey. IEEE Access **8**, 126927–126950 (2020). https://doi.org/10.1109/ACCESS.2020.3006078

[9] Fayolle, G., Malyshev, V.A., Menshikov, M.V.: Topics in the Constructive Theory of Countable Markov Chains. Cambridge University Press (1995). https://doi.org/10.1017/CBO9780511984020

[10] Ferraro, P., King, C., Shorten, R.: Distributed ledger technology for smart cities, the sharing economy, and social compliance. IEEE Access **6**, 62728–62746 (2018)

[11] Ferraro, P., King, C., Shorten, R.: Iota-based directed acyclic graphs without orphans. arXiv preprint arXiv:1901.07302 (2018)

[12] Kusmierz, B., Sanders, W., Penzkofer, A., Capossele, A., Gal, A.: Properties of the tangle for uniform random and random walk tip selection. In: 2019 IEEE International Conference on Blockchain (Blockchain). pp. 228–236. IEEE (2019)

[13] Li, Y., Cao, B., Peng, M., Zhang, L., Zhang, L., Feng, D., Yu, J.: Direct acyclic graph-based ledger for internet of things: performance and security analysis. IEEE/ACM Transactions on Networking **28**(4), 1643–1656 (2020)

[14] Micali, S.: ALGORAND: the efficient and democratic ledger. CoRR **abs/1607.01341** (2016), `http://arxiv.org/abs/1607.01341`

[15] Park, S., Oh, S., Kim, H.: Performance analysis of dag-based cryptocurrency. In: 2019 IEEE International Conference on Communications workshops (ICC workshops). pp. 1–6. IEEE (2019)

[16] Pemantle, R., Rosenthal, J.S.: Moment conditions for a sequence with negative drift to be uniformly bounded in lr. Stochastic Processes and their Applications **82**(1), 143–155 (1999). https://doi.org/https://doi.org/10.1016/S0304-4149(99)00012-5, `https://www.sciencedirect.com/science/article/pii/S0304414999000125`

[17] Penzkofer, A., Saa, O., Dziubałtowska, D.: Impact of delay classes on the data structure in iota. In: Data Privacy Management, Cryptocurrencies and Blockchain Technology, pp. 289–300. Springer (2021)

[18] Poon, J., Dryja, T.: The bitcoin lightning network:, `https://lightning.network/lightning-network-paper.pdf`

[19] Popov, S.: The tangle (2018)

[20] Popov, S., Moog, H., Camargo, D., Capossele, A., Dimitrov, V., Gal, A., Greve, A., Kusmierz, B., Mueller, S., Penzkofer, A., et al.: The coordicide (2020), *https://files.iota.org/papers/20200120_Coordicide_WP.pdf*

[21] Popov, S., Saa, O., Finardi, P.: Equilibria in the tangle. Computers & Industrial Engineering **136**, 160–172 (2019)

[22] Song, X.S., Li, Q.L., Chang, Y.X., Zhan, C.: A markov process theory for network growth processes of dag-based blockchain systems (2022). https://doi.org/10.48550/ARXIV.2209.01458, `https://arxiv.org/abs/2209.01458`

[23] Zander, M., Waite, T., Harz, D.: Dagsim: Simulation of dag-based distributed ledger protocols. ACM SIGMETRICS Performance Evaluation Review **46**(3), 118–121 (2019)

[24] Zhou, Q., Huang, H., Zheng, Z., Bian, J.: Solutions to scalability of blockchain: A survey. IEEE Access **8**, 16440–16455 (2020). https://doi.org/10.1109/ACCESS.2020.2967218

7. Appendix

7.1. Technical lemmas

We present the technical lemmas used in this paper along with the different proofs.

Lemma 2. *For every $n > 1$ and $0 \leq x \leq 1$, following holds:*

$$(1-x)^n \geq \sum_{k=0}^{m} (-1)^k C_k^n x^k, \qquad \textit{if } m \textit{ is odd},$$

$$(1-x)^n \leq \sum_{k=0}^{m} (-1)^k C_k^n x^k, \qquad \textit{if } m \textit{ is even}.$$

Proof. Consider $A_1, A_2, \cdots, A_n$ be independent events such that

$$\mathbb{P}(A_k) = x, \quad \forall k.$$

Let $S_m := \sum_{1\leq k_1<k_2<\cdots<k_m\leq n} \mathbb{P}(A_{k_1}\cap\cdots\cap A_{k_m}) = C_m^n x^m$, and $B := \mathbb{P}(\cup_k A_k) = 1-(1-x)^n$. From Bonferroni inequalities, we have the following:

$$\mathbb{P}(\cup_k A_k) \leq \sum_{k=1}^{m}(-1)^{k-1}S_k, \qquad \text{if } m \text{ is odd,}$$
$$\mathbb{P}(\cup_k A_k) \geq \sum_{k=1}^{m}(-1)^{k-1}S_k, \qquad \text{if } m \text{ is even.}$$

implying that:

$$(1-x)^n \geq 1+\sum_{k=1}^{m}(-1)^k C_k^n x^k = \sum_{k=0}^{m}(-1)^k C_k^n x^k, \qquad \text{if } m \text{ is odd,}$$
$$(1-x)^n \leq 1+\sum_{k=1}^{m}(-1)^k C_k^n x^k = \sum_{k=0}^{m}(-1)^k C_k^n x^k, \qquad \text{if } m \text{ is even.}$$

□

Lemma 3. *The following inequality holds for all $x > 0$, and all positive integer r,* $\dfrac{1}{1-(1-\frac{1}{x})^{2r}} \leq \dfrac{x^2}{2r(x-r)}$.

Proof. We have the following succession of inequalities:

$$\begin{aligned}\frac{1}{1-(1-\frac{1}{x})^{2r}} &\leq \frac{1}{1-(1-\frac{2r}{x}+\frac{r(2r-1)}{x^2})}, \qquad \text{from } (1-a)^n \leq 1-na+\frac{n(n-1)a^2}{2}) \text{ using Lemma 2}\\ =&\frac{x^2}{r(2x-2r+1)}, \qquad \leq \frac{x^2}{2r(x-r)}.\end{aligned}$$

□

Lemma 4. *If $B = 2s+2r+a+1$ then $s+0.5(B+a+1)^2/(B+a+1-r) \leq B$.*

Proof. First let us study the inequality $t+\frac{x^2}{2(x-r)} \leq x$. Note that if $x > r$ then we have:

$$t+\frac{x^2}{2(x-r)} \leq x \Leftrightarrow x^2-2(r+t)x+2rt \geq 0.$$

This implies, $x \leq x_-$ or $x \geq x_+$, with $x_\pm = (r+t)\pm\sqrt{r_2+t^2}$. The largest root is $x_+ < 2(r+t)$. So, the previous inequality is satisfied for all $x \geq 2(r+t)$. Secondly, remember that we want to solve $s+\frac{(B+a+1)^2}{2(B+a+1-r)} \leq B$ which is equivalent to $s+a+1+\frac{(B+a+1)^2}{2(B+a+1-r)} \leq B+a+1$. By simply using the previous result with $x = B+a+1$ and $t = s+a+1$, we obtain the desired result. □

Lemma 5. *For $B = 4rd^*+2r+a+1$ the following holds,*

$$rd^*+\frac{(B+a+1)^2}{2(B+a+1-r)} \leq 3rd^*+\frac{3r}{2}+a+1+\frac{r}{4d^*}.$$

Proof. We simply need to find an upper bound on $\frac{(B+a+1)^2}{(B+a+1-r)}$. We have

$$\begin{aligned}\frac{(B+a+1)^2}{(B+a+1-r)} =&\frac{(B+a+1-r+r)^2}{(B+a+1-r)} = \frac{(B+a+1-r)^2+r_2+2r(B+a+1-r))}{(B+a+1-r)}\\ =&(B+a+1-r)+2r+\frac{r_2}{(B+a+1-r)} = B+a+1+r+\frac{r_2}{(B+a+1-r)}\\ =&4rd^*+3r+2a+2+\frac{r_2}{(4rd^*+2r+a+1+a+1-r)}, \qquad \text{(plugging the value of B)}\\ \leq&4rd^*+3r+2a+2+\frac{r}{4d^*}.\end{aligned}$$

We have all the elements to derive the upper bound and conclude the proof. □

Lemma 6. *For all $j,k \in \mathcal{I}$ and $\tau \in \mathbb{N}_+$, we define*

$$f_\tau^{j,k}(v) := \begin{cases} (1-\frac{1}{v})^{2r_k} & \text{if } \tau > d^{j,k}, \\ 1 & \text{if } \tau \leq d^{j,k}. \end{cases}$$

If $v \geq \max_k r_k + 1$, then the function $f_\tau^{j,k}(v)$ is strictly increasing and concave in v for $\tau > d^{j,k}$,.

Proof. Observe that it is enough to show that the $g(v) := (1-\frac{1}{v})^{2r'}$ is strictly increasing and concave for $v \geq r'+1$. Note that $g'(v) = \frac{2r'}{v^2}(1-\frac{1}{v})^{2r'-1} > 0$, implying that the function $g(v)$ is strictly increasing in $v > 1$ and thus in $v \geq \max_k r_k + 1$. Moreover, we have that

$$g''(v) = -\frac{4r}{v^3}(1-\frac{1}{v})^{2r'-1} + \frac{2r'(2r'-1)}{v^4}(1-\frac{1}{v})^{2r'-2} = \frac{4r}{v^4}(1-\frac{1}{v})^{2r'-2}[r'+0.5-v] \leq 0,$$

where the last inequality implies that $g(v)$ is concave in $v \geq \max_k r_k + 1$. □

7.2. Tips and views recursion

This section is dedicated to the proof of Proposition 1. First, we introduce some definitions. $H_n^i := \{(k,l) | l < n - d^{k,i}, i,k \in \mathcal{I}, n,l \in \mathbb{N}\}$ is the set of tuples (k,l) such that D_l^k and C_l^k are seen by node i at time n. We also define the set $A_n^m := \{(i,l) | i \in \mathcal{I}, l \in \mathbb{N}, m+1 \leq l \leq n-1\}$. Before starting the proof of Proposition 1, we first need to prove the following lemma.

Lemma 7. *If assumption A is satisfied then:*

1. *For all $i,j \in \mathcal{I}$, $m,l \in \mathbb{N}$ we have*

$$\mathbb{E}[|V_l^i| | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j] \leq \mathbb{E}[|V_l^i|] + 1 \tag{9}$$

2. *If there exist the following finite sequence $(k_t, s_t) \in H_{s_{t-1}}^{k_{t-1}}$, $\forall 1 \leq t \leq T$, then*

$$(k_t, s_t) \in H_{s_0}^{k_0}, \forall 1 \leq t \leq T, \text{ and } H_{s'}^{k'} \subseteq H_s^k, \forall (k',s') \in H_s^k. \tag{10}$$

3. *For all $l > m$,*

$$\mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H_l^k} D_s^u, c \in C_m^j) = \mathbb{P}(c' \in V_l^k | c \notin \cup_{(u,s)\in A_l^m} D_s^u, c \in C_m^j). \tag{11}$$

4. *For all H such that $H_l^k \cap A_l^m \subseteq H \cap A_l^m$, the following holds*

$$\mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H_l^k} D_s^u, c \in C_m^j) = \mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H} D_s^u, c \in C_m^j). \tag{12}$$

5. *If $(k,l) \in H_n^i$, then*

$$\mathbb{E}[|V_l^k| | c \notin \cup_{s=m+1}^{l-1} \cup_{u\in\{u | s<n-d_{u,i}\}} D_s^u, c \in C_m^j] \leq \mathbb{E}[|V_l^k|] + 1, \ c \in C_m^j. \tag{13}$$

Proof. We prove the different points successively.

(1) Using the linearity of the expectation, we first expand the expectation as:

$$\mathbb{E}[|V_l^i|] = \sum_{c'} \mathbb{P}(c' \in V_l^i) = \mathbb{P}(c \in V_l^i) + \sum_{c' \neq c} \underbrace{\mathbb{P}(c' \in V_l^i)}_{term_1},$$

and the conditional expectation as:

$$\begin{aligned} \mathbb{E}[|V_l^i| | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j] &= \sum_{c'} \mathbb{P}(c' \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j) \\ &= \mathbb{P}(c \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j) + \sum_{c' \neq c} \underbrace{\mathbb{P}(c' \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j)}_{term_2}. \end{aligned}$$

We now focus on the different individual terms and prove that:

$$\mathbb{P}(c' \in V_l^i) \geq \mathbb{P}(c' \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j), \qquad \forall c' \neq c.$$

The right-hand side of the previous equation is equal to the probability of survival of the message $c' \neq c$ in the view set of node i at time l, conditioned on the event that the message c has not been approved until time $l-1$ by any node. Note that all the other messages in V_l^i except c, have slightly more (or same but not less) chances of being approved, when no information is given about the status of c, compared to the case where it is known that the message c has not been approved yet. Indeed, in the last case, the number of available tips is increasing and the incoming messages looking for tips to attach is remaining the same. This observation leads to the desired inequality. We now can finish the proof of *(1)* by observing that:

$$\begin{aligned}
\mathbb{E}[|V_l^i| | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j] =& \mathbb{P}(c \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j) + \sum_{c'\neq c} \underbrace{\mathbb{P}(c' \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j)}_{term_2} \\
\leq& \mathbb{P}(c \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j) + \sum_{c'\neq c} \underbrace{\mathbb{P}(c' \in V_l^i)}_{term_1} \\
=& \underbrace{\mathbb{P}(c \in V_l^i | c \notin \cup_{i\in\mathcal{I}} \cup_{k=m+1}^{l-1} D_k^i, c \in C_m^j) - \mathbb{P}(c \in V_l^i)}_{\leq 1} + \mathbb{E}[|V_l^i|] \\
\leq& \mathbb{E}[|V_l^i|] + 1 = v_l^i + 1.
\end{aligned}$$

(2) From the definition of $H_{s_{t-1}}^{k_{t-1}}$, we have $s_t < s_{t-1} - d^{k_t,k_{t-1}}$. By using a telescoping sum argument, $s_t < s_{t-1} - d^{k_t,k_{t-1}}$ implies that $s_T < s_0 - \sum_{t=1}^{T} d^{k_t,k_{t-1}}$. Using Assumption A, we can then deduce that $s_T < s_0 - d^{k_T,k_0}$. We can conclude that $(k_T, s_T) \in H_{s_0}^{k_0}$. The second statement is direct from the first part.

(3) An event (C_s^k, D_s^k) affects the view V_l^i of node i at time l only if, there exist a finite sequence such that $(k_t, s_t) \in H_{s_{t-1}}^{k_{t-1}}$, $\forall 1 \leq t \leq T$, where $(k_T, s_t) = (k, s)$ and $(k_0, s_0) = (i, l)$. Indeed, If (C_s^k, D_s^k) has to affect V_l^i directly then it has to be in its view, that is $(k, s) \in H_l^i$. Or (C_s^k, D_s^k) can affect a node k_1 at time s_1 directly (i.e $(k, s) \in H_{s_1}^{k_1}$), given that (k_1, s_1) is in the view of node i at time l (i,e $(k, s) \in H_l^i$). This can go on till time 0, hence a finite sequence suffices. Finally, equation (10) says that all the tuples (k, s) (node k at time s) that can affect V_l^i (directly or indirectly) lye inside H_l^i. Hence if $(k, s) \notin H_l^i$ then it is irrelevant for the view V_l^i. That is, event '$c' \in V_l^i$ given $c \notin \cup_{(u,s)\in H_l^k} D_s^u$' is independent of the event '$c \notin D_s^k$ given $c \notin \cup_{(u,s)\in H_l^k} D_s^u$' for all $(k, s) \notin H_l^i$. And we know that, $\mathbb{P}(A|B) = \mathbb{P}(A)$, if the events A and B are independent. So, we have

$$\mathbb{P}(c' \in V_l^i | c \notin \cup_{(u,s)\in H_l^k} D_s^u, c \notin \cup_{(u,s)\in A} D_s^u, c \in C_m^j) = \mathbb{P}(c' \in V_l^i | c \notin \cup_{(u,s)\in H_l^k} D_s^u, c \in C_m^j), \qquad \forall A \cap H_l^i = \emptyset.$$

In addition if a message is created at time m (i.e $c \in C_m^j$), then $c \notin D_s$ for all $s \leq m$. This happens with probability 1, and conditioning on such events doesn't change anything, that is $\mathbb{P}(A|B) = \mathbb{P}(A)$ if $\mathbb{P}(B) = 1$. We have

$$\mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H_l^k} D_s^u, c \in C_m^j) = \mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H_l^k \cap A_l^m} D_s^u, c \in C_m^j)$$

as $\mathbb{P}(c \notin \cup_{(u,s)\in H_l^k \cap (A_l^m)^c} D_s^u | c \in C_m^j) = 1$, where A^c complement of A.

(4) Note that H has more information than $H_l^k \cap A_l^m$, however the extra information that doesn't affect V_l^k directly nor indirectly, therefore using (11) we have

$$\mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H_l^k} D_s^u, c \in C_m^j) = \mathbb{P}(c' \in V_l^k \big| c \notin \cup_{(u,s)\in H} D_s^u, c \in C_m^j).$$

(5) We have $(k, l) \in H_n^i$ therefore $H_l^k \subseteq H_n^i$ from (10). This implies

$$H_l^k \cap A_l^m \subseteq H_n^i \cap A_l^m. \tag{14}$$

Moreover, we have

$$\begin{aligned}
\mathbb{E}[|V_l^k \| c \notin \cup_{s=m+1}^{l-1} \cup_{u\in\{u|s<n-d_{ui}\}} D_s^u, c \in C_m^j] &= \mathbb{E}[|V_l^k \| c \notin \cup_{(u,s)\in H_n^i \cap A_l^m} D_s^u, c \in C_m^j], \qquad \text{(by definition)} \\
&= \mathbb{E}[|V_l^k \| c \notin \cup_{(u,s)\in H_l^k \cap A_l^m} D_s^u, c \in C_m^j], \qquad \text{(using (14) and (12))} \\
&= \mathbb{E}[|V_l^k \| c \notin \cup_{(u,s)\in A_l^m} D_s^u, c \in C_m^j], \qquad \text{(from (11))} \\
&\leq v_l^k + 1. \qquad \text{(from (9))}
\end{aligned}$$

The last inequality concludes our proof. □

We have all the elements to prove Proposition 1. We restate the proposition below to help the reader.

Proposition 3. *We assume that $V_0^i = \{0\}, C_0^i = D_0^i = \emptyset$, for all $i \in \mathcal{I}$. If assumption A is satisfied, for every i and every n, we have the difference inclusions:*

$$x_n \leq \sum_{m=0}^{n} \sum_{j\in\mathcal{I}} r_j \prod_{l=1}^{m} \prod_{i\in\{i\in\mathcal{I}|l>d^{j,i}\}} (1 - \frac{1}{v_{n-m+l}^i + a + 1})^{2r_i},$$

and

$$v_n^i \leq \sum_{j\in\mathcal{I}} \sum_{m=d^{j,i}+1}^{n} r_j \prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} (1 - \frac{1}{v_{n-m+l}^k + a + 1})^{2r_k},$$

where $a = \max_i r_i + 1$.

Proof. In this proof, we prove the recursion on x_n (tips recursion) and then v_n^i (views recursion).

Tips recursion: The following proof is divided in two steps. First, we write $\mathbb{E}[X_n]$ as a function of $f_{l-m}^{j,i}(v)$ (defined below and in Lemma 6) for all $j, i \in \mathcal{I}$, $l, m \in \mathbb{N}$. Then we setup the environment to use Jensen's inequality and conclude. A similar method is used to prove the recursion for v_n^i.

By definition we have

$$\mathbb{E}[X_n] = \sum_{m=0}^{n} \sum_{c\in C_m} \mathbb{P}(c \in Y_n) = \sum_{m=0}^{n} \sum_{c\in C_m} \mathbb{P}(c \notin \cup_{k=0}^{n} D_k) = \sum_{m=0}^{n} \sum_{c\in C_m} \mathbb{P}(c \notin \cup_{k=m+1}^{n} D_k),$$

where the last equality is coming from the fact that if $c \in C_m$ then $\mathbb{P}(c \notin D_k) = 1$ for all $k \leq m$. We can rewrite the previous equation as:

$$\mathbb{E}[X_n] = \sum_{m=0}^{n} \sum_{c\in C_m} \prod_{l=m+1}^{n} \mathbb{P}(c \notin D_l | c \notin \cup_{k=m+1}^{l-1} D_k),$$

using the observation that $\mathbb{P}(b \notin B_1 \cup B_2) = \mathbb{P}(b \notin B_1 | b \notin B_2)\mathbb{P}(b \notin B_2)$ and by assuming that $\cup_{k=n}^{m} A_k = \emptyset$ for any set A when $m < n$. At every instant, each node chooses the tips to validate independently from the choice of the other node. Therefore, if $c \notin \cup_{k=0}^{l-1} D_k$ then the events $c \in D_l^i, i \in \mathcal{I}$ are independent impying that

$$\mathbb{P}(c \notin D_l | c \notin \cup_{k=0}^{l-1} D_k) = \prod_{i\in\mathcal{I}} \mathbb{P}(c \notin D_l^i | c \notin \cup_{k=0}^{l-1} D_k) = \prod_{i\in\mathcal{I}} \mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k).$$

Therefore, we have

$$\mathbb{E}[X_n] = \sum_{m=0}^{n} \sum_{c\in C_m} \prod_{l=m+1}^{n} \prod_{i\in\mathcal{I}} \mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k) = \sum_{m=0}^{n} \sum_{j\in\mathcal{I}} \sum_{c\in C_m^j} \prod_{l=m+1}^{n} \prod_{i\in\mathcal{I}} \mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j),$$

where the last equality is coming from the definition of C_m. Now we are going to focus on each individual terms $\mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j)$. Using the law of total probability, we get

$$\begin{aligned}
&\mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j) \\
&= \sum_{v\in\mathbf{N}} \mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j, |\ V_l^i\ | = v) \mathbb{P}(|\ V_l^i\ | = v \mid c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j),
\end{aligned} \tag{15}$$

where V_l^i is the view set of node i at time l. The fact that node i, at instant l, only choose to validate $2r_i$ tips uniformly implies that we have:

$$\mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j, |V_l^i| = v) = \begin{cases} (1-\frac{1}{v})^{2r_i}, & \text{if } c \in V_l^i, \\ 1, & \text{if } c \notin V_l^i, \end{cases} = \begin{cases} (1-\frac{1}{v})^{2r_i}, & \text{if } l > m + d^{j,i}, \\ 1, & \text{if } l \leq m + d^{j,i}, \end{cases} =: f_{l-m}^{j,i}(v)$$

The last equality follows from the condition $c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j$. It states that message $c \in C_m^j$ has not been attached by any node until time $l-1$. Therefore, whether a node k is able to see the message c as a tip at time l depends on whether it has observed it. And that depends only on whether or not $l-m$ is greater than $d^{j,i}$. Coming back to (15) we obtain:

$$\mathbb{P}(c \notin D_l^i | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j) = \sum_{v \in \mathbb{N}} \mathbb{P}(|V_l^i| = v | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j) f_{l-m}^{j,i}(v).$$

Recall, from Lemma 6 that function $f_{l-m}^{j,i}(v)$ is concave in the range $v \geq a = \max_i r_i + 1$. Moreover function $f_{l-m}^{j,i}(v)$ is strictly increasing in $v \geq 1$ (as long as $l > m + d^{j,i}$). We have therefore the following sequence of inequalities:

$$\begin{aligned}
\mathbb{E}[X_n] &= \sum_{m=0}^{n} \sum_{j \in \mathcal{I}} \sum_{c \in C_m^j} \prod_{l=m+1}^{n} \prod_{i \in \mathcal{I}} \sum_{v \in \mathbb{N}} \mathbb{P}(|V_l^i| = v | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j) f_{l-m}^{j,i}(v), \\
&\leq \sum_{m=0}^{n} \sum_{j \in \mathcal{I}} \sum_{c \in C_m^j} \prod_{l=m+1}^{n} \prod_{i \in \mathcal{I}} \sum_{v \in \mathbb{N}} \mathbb{P}(|V_l^i| = v | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j \underbrace{f_{l-m}^{j,i}(v+a)}_{\text{focus}}, \qquad \text{(strictly increasing f, see Lemma 6)} \\
&\leq \sum_{m=0}^{n} \sum_{j \in \mathcal{I}} \sum_{c \in C_m^j} \prod_{l=m+1}^{n} \prod_{i \in \mathcal{I}} f_{l-m}^{j,i}(\underbrace{\mathbb{E}[|V_l^i| + a | c \notin \cup_{k=m+1}^{l-1} D_k, c \in C_m^j]}_{\leq \mathbb{E}[|V_l^i|] + a + 1, \quad \text{(see (9))}}), \qquad \text{(Jensen, see Lemma 6)} \\
&\leq \sum_{m=0}^{n} \sum_{j \in \mathcal{I}} \sum_{c \in C_m^j} \prod_{l=m+1}^{n} \prod_{i \in \mathcal{I}} f_{l-m}^{j,i}(v_l^i + a + 1), \qquad \text{(strictly increasing f, see Lemma 6)} \\
&= \sum_{m'=0}^{n} \sum_{j \in \mathcal{I}} r_j \prod_{l'=1}^{m'} \prod_{i \in \mathcal{I}} f_{l'}^{j,i}(v_{n-m'+l'}^i + a + 1), \qquad \text{(putting: } m' = n-m \text{ and } l' = l-m) \\
&= \sum_{m=0}^{n} \sum_{j \in \mathcal{I}} r_j \prod_{l=1}^{m} \prod_{i \in \mathcal{I}} f_l^{j,i}(v_{n-m+l} + a + 1), \qquad \text{(changing back name of variables)} \\
&= \sum_{m=0}^{n} \sum_{j \in \mathcal{I}} r_j \prod_{l=1}^{m} \prod_{i \in \{i \in \mathcal{I} | l > d^{j,i}\}} (1 - \frac{1}{v_{n-m+l}^i + a + 1})^{2r_i} \qquad \text{(putting definition of } f\text{)}
\end{aligned} \tag{16}$$

The last equation concludes the first part of the proof.

View recursion: By definition of the view V_n^i (see (3)), we have $|V_n^i| = \sum_{j \in \mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} \sum_{c \in C_m^j} \mathbb{1}(c \in V_n^i)$. Taking the expectation on both sides and using the following fact on conditional probability $\mathbb{P}(a \notin \cup_{k=0}^{n} A_k) = \prod_{l=0}^{n} \mathbb{P}(a \notin A_l | a \notin \cup_{k=0}^{l-1} A_k)$, we obtain:

$$\begin{aligned}
\mathbb{E}[|V_n^i|] &= \sum_{j \in \mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \mathbb{P}(c \in V_n^i | c \in C_m^j) = \sum_{j \in \mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \mathbb{P}(c \notin \cup_{k \in \mathcal{I}} \cup_{l < n-d^{k,i}} D_l^k | c \in C_m^j) \\
&= \sum_{j \in \mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \mathbb{P}(c \notin \cup_{l \leq n} \cup_{k:\, d^{k,i} < n-l} D_l^k | c \in C_m^j). \qquad \text{(interchanging unions)}
\end{aligned}$$

If $c \in C_m$ then $\mathbb{P}(c \notin D_l) = 1$ for all $l \leq m$. Indeed, if c is created at time m, it can only be approved after time m. Therefore, we get

$$\mathbb{E}[|V_n^i|] = \sum_{j \in \mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \mathbb{P}(c \notin \cup_{l=m+1}^{n} \cup_{d^{k,i} < n-l} D_l^k | c \in C_m^j).$$

Using again the same fact on conditional probability, we obtain the following equality:

$$\mathbb{E}[|V_n^i|] = \sum_{j\in\mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \prod_{l=m+1}^{n} \mathbb{P}(c \notin \cup_{d^{k,i}<n-l} D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{d^{k,i}<n-s} D_s^k, c \in C_m^j).$$

At a given time each node acts independently, therefore if a message $c \in C_m^j$ is a tip (i.e. $c \notin \cup_{s=m+1}^{l-1} \cup_{d^{k,i}<n-s} D_s^k$) at time $l-1$ then the events $c \in D_l^k, k \in \mathcal{I}$ are independent. That is, for any fixed $i \in \mathcal{I}, c \in C_m^j, m < l < n$, we have

$$\mathbb{P}(c \notin \cup_{d_{ti}<n-l} D_l^t | c \notin \cup_{s=m+1}^{l-1} \cup_{d^{k,i}<n-s} D_s^k) = \prod_{d_{ti}<n-l} \mathbb{P}(c \notin D_l^t | c \notin \cup_{s=m+1}^{l-1} \cup_{d^{k,i}<n-s} D_s^k).$$

So we have

$$\mathbb{E}[|V_n^i|] = \sum_{j\in\mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \prod_{l=m+1}^{n} \prod_{d^{k,i}<n-l} \mathbb{P}(c \notin D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{d_{ti}<n-s} D_s^t, c \in C_m^j)$$

Now we are going to focus on each individual term $\mathbb{P}(c \notin D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{d_{ti}<n-s} D_s^t, c \in C_m^j)$. Recall the following conditional probability fact $\mathbb{P}(a \notin A|B) = \sum_{c\in C} \mathbb{P}(a \notin A|B, X = c)\mathbb{P}(X = c|B)$ for all appropriate A, B, C, X. Using this fact, for all $i, k \in \mathcal{I}, c \in C_m^j, m < l < n$, we get

$$\mathbb{P}(c \notin D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t)$$

$$= \sum_{v\in\mathbf{N}} \mathbb{P}(c \notin D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t, |V_l^k| = v)\mathbb{P}(|V_l^k| = v | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t).$$

If node k at time l cannot see a message $c \in C_m^j$ as a tip then it can't attach to it, that is, if $c \notin V_l^k$ then $c \notin D_l^k$. On the other hand, if $c \in V_l^k$ then we have $\mathbb{P}(c \notin D_l^k) = (1 - \frac{1}{|V_l^k|})^{2r_k}$, due to the fact that node k at instant l chooses uniformly $2r_k$ messages to approve. To summarize, we have

$$\mathbb{P}(c \notin D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{r,s<n-d_{ri}} D_s^r, c \in C_m^j, |V_l^k| = v) = \begin{cases} (1-\frac{1}{v})^{2r_k} & \text{if } c \in V_l^k \\ 1 & \text{if } c \notin V_l^k. \end{cases}$$

$$= \begin{cases} (1-\frac{1}{v})^{2r_k} & \text{if } l > m + d^{j,k} \\ 1 & \text{if } l \leq m + d^{j,k} \end{cases} \qquad = \qquad f_{l-m}^{j,k}(v).$$

The last equality follows from the condition $c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t$. It says that message $c \in C_m^j$ has not been attached by any node until time $l-1$. Therefore, whether a node k is able to see the message c as a tip at time l depends on whether it has observed it. And that depends only on whether or not $l - m$ is greater than $d^{j,k}$. Putting it back everything, we get

$$\mathbb{P}(c \notin D_l^k | c \notin \cup_{s=m+1}^{l-1} \cup_{r,s<n-d_{ri}} D_s^r, c \in C_m^j) = \sum_{v\in\mathbf{N}} f_{l-m}^{j,k}(v)\mathbb{P}(|V_l^k| = v | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t, c \in C_m^j).$$

We go back to v_n^i. We have following succession of inequalities:

$$
\begin{aligned}
v_n^i &= \sum_{j\in\mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \prod_{l=m+1}^{n} \prod_{d^{k,i}<n-l} \sum_{v\in\mathbf{N}} f_{l-m}^{j,k}(v)\mathbb{P}(|V_l^k| = v | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t, c \in C_m^j) \\
&\leq \sum_{j\in\mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \prod_{l=m+1}^{n} \prod_{d^{k,i}<n-l} \sum_{v\in\mathbf{N}} f_{l-m}^{j,k}(v+a)\mathbb{P}(|V_l^k| = v | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t, c \in C_m^j), \qquad \text{(strictly increasing f, see Lemma 6)} \\
&\quad \sum_{j\in\mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \prod_{l=m+1}^{n} \prod_{d^{k,i}<n-l} \sum_{v\in\mathbf{N}} f_{l-m}^{j,k}(\underbrace{\mathbb{E}[|V_l^k| + a | c \notin \cup_{s=m+1}^{l-1} \cup_{s<n-d_{ti}} D_s^t, c \in C_m^j]}_{\leq v_l^k+a+1, \quad \text{(see (13))}}) \qquad \text{(Jensen, see Lemma 6)} \\
&\leq \sum_{j\in\mathcal{I}} \sum_{m=0}^{n-d^{j,i}-1} r_j \prod_{l=m+1}^{n} \prod_{d^{k,i}<n-l} \underbrace{\sum_{v\in\mathbf{N}} f_{l-m}^{j,k}(v_l^k + a + 1)}, \qquad \text{(strictly increasing f, see Lemma 6)} \\
&= \sum_{j\in\mathcal{I}} \sum_{m'=d^{j,i}+1}^{n} r_j \prod_{l'=1}^{m'} \prod_{d^{k,i}<m'-l'} f_{l'}^{j,k}(v_{n-m'+l'}^k + a + 1), \qquad \text{(by putting } m' = n-m, \text{ and } l' = l-m). \\
&= \sum_{j\in\mathcal{I}} \sum_{m=d^{j,i}+1}^{n} r_j \underbrace{\prod_{l=1}^{m} \prod_{d^{k,i}<m-l} f_l^{j,k}(v_{n-m+l}^k + a + 1)}_{\text{Term}}, \\
&= \sum_{j\in\mathcal{I}} \sum_{m=d^{j,i}+1}^{n} r_j \underbrace{\prod_{l=1}^{m} \prod_{k\in\{k|m-d^{k,i}>l>d^{j,k}\}} (1 - \frac{1}{v_{n-m+l}^k + a + 1})^{2r_k}}_{\text{Term}}, \qquad \text{(by definition of } f).
\end{aligned}
$$

The last equation concludes the second part of the proof. □

7.3. Deriving an upper bound for tips and views from bounded increments and negative drifts.

This section is dedicated to provide a new proof regarding the existence of a finite bound over the expectation of the cardinal of the tips set. Such results show that DAG-based distributed ledgers do not diverge, as long as the delays is bounded. Recall that C_n^i (resp. D_n^i) is the set of messages created (resp. approved) by node i at instant n. As before, we use $C_n = \sqcup_{i\in\mathcal{I}} C_n^i$, $D_n = \bigcup_{i\in\mathcal{I}} D_n^i$, $x_n = \mathbb{E}[X_n]$, $v_n^i = \mathbb{E}[|V_n^i|]$ as shorthand notations. Note that the above expectations are only conditioned with respect to the initial condition of the system ($V_0^i = \{0\}$, $C_0^i = D_0^i = \emptyset$). Let us define the set A_n as

$$A_n := (\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{t=0}^{n} D_t). \tag{17}$$

This set is the set of messages (tips) that were created till time $(n - d^*)$ and have never (until time n) been attached to the DL, i.e. the set of messages that have been created until time $(n - d^*)$ that are tips at time n.

Lemma 8. *A_n has the following useful properties:*

1. *(Common tips) For every $i \in \mathcal{I}$, and $\forall n \in \mathbb{N}$, $Y_n \supseteq A_n \subseteq V_{n+1}^i$.*
2. *(Tips bound) At every instant n, the number of tips X_n is bounded as follows:*

$$|A_n| \leq X_n \leq |A_n| + rd^*, \qquad \forall n \in \mathbb{N}.$$

3. *(View bound) The cardinality of the views set is bounded as follows:*

$$|A_n| \leq |V_{n+1}^i| \leq |A_n| + 3rd^*, \qquad \forall i \in \mathcal{I}, \forall n \in \mathbb{N}.$$

4. *(View bounded by tips) Views are upper bounded by tips as follows:*

$$|V_{n+1}^i| \leq X_n + 2rd^* \qquad \forall i \in \mathcal{I},\ \forall n \in \mathbb{N}.$$

5. *The number of tips is lower bounded by r, that is*

$$X_n \geq r, \qquad \forall n \in \mathbb{N}.$$

Proof. We prove each point separately.

1. From the definition of the view set V_n^i in (3), we have:

$$\begin{aligned}
V_{n+1}^i =& \Big(\bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} C_t^j - \bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} D_t^j \Big), \\
\supseteq& \Big(\bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} C_t^j - \bigcup_{t=0}^{n} D_t \Big), \qquad (\text{ as } D' \supseteq D \implies C - D \supseteq C - D') \\
\supseteq& \Big(\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{t=0}^{n} D_t \Big), \qquad (\text{ as } C \supseteq C' \implies C - D \supseteq C' - D) \\
=& A_n.
\end{aligned}$$

We focus on the evolution of the tips. By definition, we have

$$Y_n = \Big(\bigcup_{t=0}^{n} C_t - \bigcup_{t=0}^{n} D_t \Big) \supseteq \Big(\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{t=0}^{n} D_t \Big) = A_n.$$

2. By definition, we have

$$\begin{aligned}
Y_n =& \Big(\bigcup_{t=0}^{n} C_t - \bigcup_{t=0}^{n} D_t \Big) && (18) \\
\subseteq& \Big(\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{t=0}^{n} D_t \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} C_t \Big), \qquad (\text{as } (C \cup C') - D \subseteq (C - D) \cup C') && (19) \\
=& A_n \cup \Big(\bigcup_{t=n-d^*+1}^{n} C_t \Big) && (20) \\
\implies X_n \leq& \Big| A_n \cup \Big(\bigcup_{t=n-d^*+1}^{n} C_t \Big) \Big| \leq \Big| A_n \Big| + \Big| \bigcup_{t=n-d^*+1}^{n} C_t \Big| = \Big| A_n \Big| + rd^*. && (21)
\end{aligned}$$

3. We move to the bound on the view sets.

$$\begin{aligned}
V_{n+1}^i =& \Big(\bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} C_t^j - \bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} D_t^j \Big), \\
\subseteq& \Big(\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} D_t^j \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} C_t \Big), \qquad (\text{as } (C \cup C') - D \subseteq (C - D) \cup C') \\
\subseteq& \Big(\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{t=0}^{n} D_t \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} C_t \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} D_t \Big), \qquad (\text{as } C - D \subseteq (C - (D \cup D')) \cup D'). \\
\implies |V_{n+1}^i| \leq& \Big| \Big(\bigcup_{t=0}^{n-d^*} C_t - \bigcup_{t=0}^{n} D_t \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} C_t \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} D_t \Big) \Big| \\
\leq& |A_n| + rd^* + 2rd^*.
\end{aligned}$$

The lower bound is direct from the first point.

4. Again by definition, we have

$$\begin{aligned}
V^i_{n+1} &= \Big(\bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} C^j_t - \bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} D^j_t \Big) \subseteq \Big(\bigcup_{t=0}^{n} C_t - \bigcup_{j\in\mathcal{I}} \bigcup_{t=0}^{n-d^{ji}} D^j_t \Big), \\
&\subseteq \Big(\bigcup_{t=0}^{n} C_t - \bigcup_{t=0}^{n} D_t \Big) \cup \Big(\bigcup_{t=n-d^*+1}^{n} D_t \Big). \\
&= Y_n \cup \Big(\bigcup_{t=n-d^*+1}^{n} D_t \Big) \Longrightarrow \big| V^i_{n+1} \big| \leq X_n + \Big| \bigcup_{t=n-d^*+1}^{n} D_t \Big| \\
&\leq X_n + 2rd^*.
\end{aligned}$$

□

We use the above properties to obtain a bound on the drift of the stochastic process.

Lemma 9. *(Drift) The number of tips X_n has a uniform negative drift outside some set. Precisely, the drift is bounded by*

$$\mathbb{E}[X_{n+1} - X_n \mid X_n = x] \leq r - \frac{2r(x - rd^*)}{x + 2rd^*} \Big(1 - \frac{r}{x + 2rd^*} \Big).$$

And as X_n becomes large, the drift tends to $-r$, that is

$$\lim_{x\to\infty} \mathbb{E}[X_{n+1} - X_n \mid X_n = x] \leq \lim_{x\to\infty} [r - \frac{2r(x - rd^*)}{x + 2rd^*} \Big(1 - \frac{r}{x + 2rd^*} \Big)] = -r.$$

Also, defining $a = 7(rd^* + r)$, we have

$$\mathbb{E}[X_{n+1} - X_n \mid X_n = x, \forall x > a] \leq -r/7.$$

Proof. Let $a \in A_n$ (note this implies $a \in Y_n$ and ,$a \in V^i_{n+1}$, $\forall i$) be a tip that is visible to all nodes. Remember that A_n is the set of messages that have been created until time $(n - d^*)$ that are tips at time n. The probability that node i doesn't attach its new messages to a at time $n + 1$ is given by,

$$\mathbb{P}\ (\ a \notin D^i_{n+1} \mid a \in A_n\ ) = (1 - \frac{1}{|V^i_{n+1}|})^{2r_i}.$$

The probability that no node attaches their new messages to it is given by

$$\mathbb{P}\ (\ a \notin D_{n+1} \mid a \in A_n\ ) = \prod_{i\in\mathcal{I}} (1 - \frac{1}{|V^i_{n+1}|})^{2r_i}.$$

The expected number of tips in A_n that no longer remain as tips at time $n + 1$ is given by,

$$\begin{aligned}
\mathbb{E}\ [\ |A_n \cap D_{n+1}|\ |\ X_n = x\ ] &= |A_n| - |A_n| \mathbb{P}\ (\ a \notin D_{n+1} \mid a \in A_n\ ) = |A_n| - |A_n| \prod_{i\in\mathcal{I}} (1 - \frac{1}{|V^i_{n+1}|})^{2r_i} \\
&\geq |A_n| - |A_n| \prod_{i\in\mathcal{I}} (1 - \frac{1}{x + 2rd^*})^{2r_i}, \qquad \text{(from Lemma 8)} \\
&= |A_n| - |A_n| (1 - \frac{1}{x + 2rd^*})^{2r}, \\
&\geq |A_n| - |A_n| \Big(1 - \frac{2r}{x + 2rd^*} + \frac{2r(2r-1)}{2(X_n + 2rd^*)^2} \Big), \qquad \text{(inclusion exclusion } m = 2) \\
&= |A_n| \Big(\frac{2r}{x + 2rd^*} - \frac{r(2r-1)}{(x + 2rd^*)^2} \Big), \\
&\geq \frac{2r|A_n|}{x + 2rd^*} \Big(\underbrace{1 - \frac{r}{x + 2rd^*}}_{\geq 0 \text{ as } x \geq r} \Big) \\
&\geq \frac{2r(x - rd^*)}{x + 2rd^*} \Big(1 - \frac{r}{x + 2rd^*} \Big), \qquad \text{(from lemma 8, } x \leq |A_n| + rd^*).
\end{aligned}$$

We are in position to prove the first part of the theorem. From the definition, we have $Y_{n+1} = C_{n+1} \sqcup (Y_n - D_{n+1})$ which implies that $X_{n+1} = r + X_n - |Y_n \cap D_{n+1}|$. We then have our first claim:

$$\begin{aligned}
\mathbb{E}[X_{n+1} \mid X_n = x] &\leq r + x - \mathbb{E}[|A_n \cap D_{n+1}| \mid X_n = x], \qquad \text{(from Property 8, } A_n \subseteq Y_n\text{)} \\
&\leq r + x - \frac{2r(x - rd^*)}{x + 2rd^*} \Big(1 - \frac{r}{x + 2rd^*} \Big), \qquad \text{(putting the value of } |A_n \cap D_{n+1}| \text{ from above)}
\end{aligned}$$

The second claim about the limit comes directly from the drift equation.

We conclude the proof by proving the third claim. We consider $x \geq a$.

$$\begin{aligned}
&\mathbb{E}[X_{n+1} - X_n \mid X_n = x,\ x \geq a] \leq r - \underbrace{\frac{2r(x - rd^*)}{x + 2rd^*}}_{:=T1(x)} \underbrace{\Big(1 - \frac{r}{x + 2rd^*} \Big)}_{:=T2(x)} \\
&\leq r - \frac{2r(a - rd^*)}{a + 2rd^*} \Big(1 - \frac{r}{a + 2rd^*} \Big), \qquad \text{(as both the function } T1(x) \text{ and } T2(x) \text{ are increasing in } x \geq 0\text{)} \\
&= r - \frac{2r(6rd^* + 7r)}{9rd^* + 7r} \Big(1 - \frac{r}{9rd^* + 7r} \Big), \\
&\leq r - \frac{2r(6rd^* + 7r)}{9rd^* + 7r} \Big(1 - \frac{r}{7r} \Big), \\
&\leq r - \frac{2r(6rd^*)}{9rd^*} \Big(1 - \frac{r}{7r} \Big) = -\frac{r}{7},
\end{aligned}$$

□

We turn our attention to the proof of the existence of an upper bound on the expected number of tips. We refer the reader to theorem 1 of [16]. Note that we have a bounded jump in our process, by construction, that is $\mathbb{E}[|X_{n+1} - X_n|^p \mid X_n, \cdots, X_0] \leq (2r)^p$ for all $p > 0$. Moreover, it is clear from the above lemma that we have uniform negative drift after some threshold a. So, Theorem 1 of [16] implies that the expected number of tips is bounded, that is, there exist some constant c such that $\forall n$ we have $\mathbb{E}[X_n] \leq c$.