

Nonlinear Blockchain Scalability: a Game-Theoretic Perspective

Lin Chen

Department of Computer Science, Texas Tech University,
TX, USA
chenlin198662@gmail.com

Zhimin Gao

Department of Computer Science, Auburn University at
Montgomery, AL, USA
mtion@msn.com

Lei Xu

Department of Computer Science, University of Texas Rio
Grande Valley, TX, USA
xuleimath@gmail.com

Keshav Kasichainula, Weidong Shi

Department of Computer Science, University of Houston,
TX, USA
kkasichainula@uh.edu,
larryshi@ymail.com

ABSTRACT

Recent advances in the blockchain research have been made in two important directions. One is refined resilience analysis utilizing game theory to study the consequences of selfish behaviors of users (miners), and the other is the extension from a linear (chain) structure to a non-linear (graphical) structure for performance improvements, such as IOTA and Graphcoin. The first question that comes to people's minds is what improvements that a blockchain system would see by leveraging these new advances. In this paper, we consider three major metrics for a blockchain system: full verification, scalability, and finality-duration. We establish a formal framework and prove that no blockchain system can achieve full verification, high scalability, and low finality-duration simultaneously. We observe that classical blockchain systems like Bitcoin achieves full verification and low finality-duration, Harmony and Ethereum 2.0 achieve low finality-duration and high scalability. As a complementary, we design a non-linear blockchain system that achieves full verification and scalability. We also establish, for the first time, the trade-off between scalability and finality-duration.

KEYWORDS

Scalability; finality-duration; Blockchain; Nash equilibrium

ACM Reference Format:

Lin Chen, Lei Xu, Zhimin Gao, and Keshav Kasichainula, Weidong Shi. 2020. Nonlinear Blockchain Scalability: a Game-Theoretic Perspective. In *Proceedings of ACM Conference (Conference '17)*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

Blockchain technology provides a mechanism for untrusted parties to carry out transactions without a fully trusted central system. The

basic idea behind this is that, instead of having trust in a centralized system or any other specific participant, each participant chooses to trust the majority of the participants and accepts the outcome achieved through consensus among them.

One major reason that hinders the adoption of blockchain is scalability [22]. For example, Bitcoin network can only process less than 10 transactions per second on average [6], while typical payment systems like Visa can process thousands of transactions per second.

Recently, a variety of approaches are proposed to address the scalability issue. Most of them follow the general framework of divide and conquer, e.g., Zilliqa [19], Harmony [18], and Ethereum 2.0 [17], and use a sharding scheme that allows transactions to be processed by a subgroup of nodes (a sharding committee). A sharding scheme usually has a critical issue in terms of resilience, as the correctness of each transaction now solely depends on a subgroup of voters. Consequently, if common consensus protocols like Proof-of-Work (PoW) or Byzantine Fault Tolerance (BFT) is used within subgroups, then the fraction of malicious nodes within every subgroup cannot exceed $1/2$ or $1/3$, which is a significantly stronger assumption than that of a standard blockchain system, which only requires that the fraction of malicious nodes does not exceed $1/2$ or $1/3$ of all nodes. We remark that both Harmony and Ethereum 2.0 claim that if subgroups are generated in a perfect randomized way, then the percentage of honest nodes within each subgroup is almost the same as their percentage in the whole group of nodes; however, this requires a perfect distributed random number generation as a separate procedure, which brings an additional assumption on the security of this additional procedure.

If we aim to guarantee that every transaction is correctly executed by only relying on the standard assumption that the majority of the nodes are honest, then it is natural to require every transaction to be verified by all the nodes. Consequently, it is a natural question of whether scalability is achievable at all, as it appears that any divide and conquer based solution would inevitably reduce the total number of verifications received by a transaction. A "non-linear" blockchain structure recently introduced by IOTA [13] and Graphcoin [2] brings the hope. The basic idea is to allow blocks to be connected as a directed acyclic graph (DAG) instead of a chain. Such a non-linear structure implements a divide and conquers approach implicitly by allowing multiple blocks to be appended simultaneously, as a general

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference '17, July 2017, Washington, DC, USA
© 2020 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/Y/YY/MM...\$15.00
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

graph can be extended in multiple directions. Meanwhile, if we treat different growing directions as soft forks or branches, then they have the possibility (depending on system parameters) to “merge” again in the future (see the following figure 2, where the sequences of blocks the following block A and B meet at block C). Therefore, if a node who tries to append a new block is required to verify a few previous blocks, there is a possibility that a block may still be verified by all the nodes, albeit the delay of such verification.

On a high level, there are three crucial metrics involved in a general blockchain system: full verification, scalability, and finality-duration. In a nutshell, full verification requires every transaction to be verified by essentially all the nodes (which thus ensures resilience under the standard assumption that majority of the nodes follow the protocol); scalability means the system throughput, or the total number of transactions executed per unit of time, is proportional to the total number of participating nodes; and finality-duration means the delay in reaching consensus on the correctness of the execution of each transaction. We give a precise definition later.

Classical blockchain systems like Bitcoin achieves full verification and low finality-duration, but not scalability. This is because Bitcoin requires every block, and hence the transactions within a block, to be verified by all the nodes; meanwhile, it has a constant finality-duration because every block is finalized after a constant number of blocks are appended afterward. However, it does not scale, as the increase in the number of nodes does not allow the system to handle more transactions per unit of time, which has been pointed out in many prior papers. On the other hand, blockchain systems like Harmony [18] and Ethereum 2.0 [17] achieve constant finality-duration and scalability, but not full verification.

In this paper, we provide a view on the relationships between full verification, low finality-duration, and scalability. We also consider the potential trade-off among these three metrics. Our detailed contribution is summarized as follows:

Our contributions. Based on a counting argument, we show that it is impossible to achieve full verification, low finality-duration and scalability simultaneously.

Given our impossibility result, and the fact that: (i) Bitcoin achieves full verification and low (constant) finality-duration, but not scalability; and (ii) Harmony and Ethereum 2.0 achieves low (constant) finality-duration and scalability, but not full verification, It is natural to ask whether there exists a blockchain system that satisfies both full verification and scalability. We prove that, interestingly, a modified version of a non-linear blockchain system, which is in line with a well-known blockchain system – IOTA, can achieve both full verification and scalability at the cost of high (non-constant) finality-duration. We further employ a game-theoretical analysis to characterize the trade-off between scalability and finality-duration for such a non-linear blockchain system. Informally speaking, the followings hold simultaneously for the non-linear blockchain system:

- (i) $O(s)$ new blocks are generated per unit of time on average;
- (i) after $O(s \log s)$ units of time, with a very high probability, each block will be verified by all users in the systems.

Here s is a system parameter that can be set suitably at the genesis block. When $s = 1$, the non-linear system degenerates to a linear system with a fixed block generation rate that is independent of the

nodes in the system, while the delay is a constant. This coincides with the classical Bitcoin system. On the other hand, s can be as high as $O(m)$ where m is the number of nodes in the system. In this case, the system is fully scalable, albeit that only a sufficiently long delay ($O(s \log s)$) can ensure full verification. However, if we set s to be $O(m)$ to enforce scalability and meanwhile enforce the delay to be some constant instead of $O(s \log s)$, then full verification cannot be guaranteed.

We remark that the big- O notation in our statements hides a constant which is roughly the average time for a block to be generated, that is, we measure the delay in terms of the number of blocks; therefore, Bitcoin is considered as low finality-duration as the delay is constant blocks. Our result does not conflict with prior researches that complain about the “high” finality-duration of Bitcoin because of the long time it takes to generate a single block. The research that tries to decrease such a block generation time is parallel to this paper. For example, if a lighter version of PoW can be used in the existing Bitcoin system, then it can also be used directly in our non-linear blockchain system, while our impossibility result, as well as the trade-off between finality-duration and scalability, remain the same.

2 RELATED WORKS

The study of e-cash systems dates back to 1983 [3, 15]. However, all such systems require a centrally or quasi-centrally controlling authority. A well-known exception, Bitcoin, was introduced by Nakamoto [10] in 2008, which uses a public ledger known as a *blockchain* to record transactions carried out between users. Following this line of research, various alternative blockchain-based transaction systems are proposed [9, 16, 20], further improving the performance and security of Bitcoin as well as extending the system to deal with applications beyond transactions (e.g., smart contracts). We refer the readers to several surveys on blockchain systems [1, 4, 7, 8, 14, 21, 23]. In particular, [21] provides a comprehensive introduction to the bitcoin network, [4, 7, 14] focus on the security and privacy results on blockchain, [1] focuses on the applications of blockchain. The most relevant survey to this paper is [8], which summarizes recent results on game-theoretical studies of blockchain. However, most of the existing game theoretical research primarily focuses on the traditional linear blockchain system, only a very recent paper of Popov et al. [12] gives the first game-theoretical analysis of IOTA. Their result, however, does not establish the trade-off between scalability and finality-duration.

2.1 Classical and Non-linear Blockchain

Chain-structured blockchain. Most of the existing blockchain systems, e.g., Bitcoin, Ethereum, Hyperledger, follow the classical structure where blocks form a chain as illustrated by Fig 1.

Non-linear (graph-structured) blockchain. Popov introduced the concept of *tangle* [13] which allows a blockchain to adopt a directed acyclic graph (DAG) architecture. We summarize the abstract model of a non-linear blockchain in Section 3. We briefly review IOTA, which is the most well-known non-linear blockchain system so far. On a high level, IOTA allows each transaction be an individual node linked in the distributed ledger. We may interpret a transaction as a block in such a system. In the tangle, each user needs to select

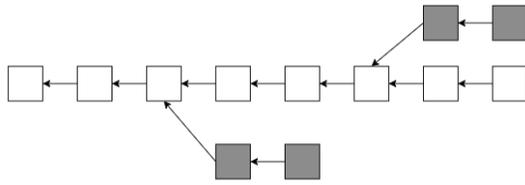


Figure 1: Chain-structured blockchain. White squares form the main chain, and gray squares form the side chains that are discarded eventually.

one transaction from the pool as well as two previous blocks (transactions) in the system. The user verifies these two transactions and mines a new block referring to them. Then this new block (transaction) is broadcasted to the tangle network. Figure 2 gives a simple example of a non-linear blockchain.

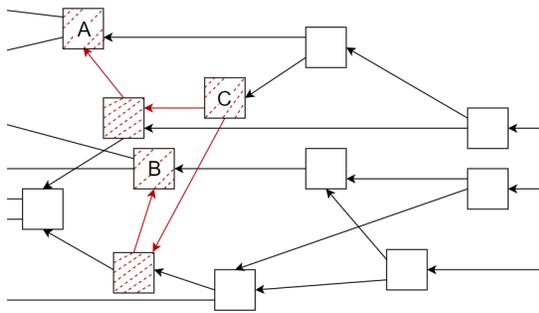


Figure 2: A non-linear blockchain. White squares are verified transactions/blocks.

3 THE ABSTRACT MODEL

We describe an abstract model of a non-linear blockchain which is general enough to incorporate existing well-known non-linear blockchain systems like IOTA and Graphcoin.

A non-linear blockchain NLB is defined by a quadruple

$$NLB : (\mathcal{A}, C, \mathcal{R}, \mathcal{E}), \text{ where}$$

- \mathcal{A} defines the rules of building and adding a new block to the blockchain. Since we are considering non-linear blockchain, \mathcal{A} allows multiple blocks to be added simultaneously.
- C defines the way to check a block, including validity verification, such as whether the block has the correct format and whether transactions included in the block are valid, and whether the block is finalized.
- \mathcal{R} defines the way how the award is assigned to a user who adds a new block to the DAG. A NLB needs to encourage users to participate in the construction of the blockchain by giving rewards to those who add new blocks.
- \mathcal{E} defines the rules to eliminate conflict blocks. Similar to a linear blockchain, it is possible that multiple participants have different local copies of the blockchain, and \mathcal{E} determines which version should be kept.

Next, we provide formal definitions of the three metrics of a blockchain system we mentioned earlier.

DEFINITION 1 (FULL VERIFICATION). A blockchain system satisfies the property of full verification if every block is verified by all the nodes in the system before it is finalized.

If a blockchain system satisfies full verification, then resilience follows directly from standard assumptions on the percentage of honest nodes among all nodes, e.g., if the blockchain uses PoW or BFT as the consensus protocol, then it is resilient once the majority or 2/3 of nodes follow the protocol.

DEFINITION 2 (SCALABILITY). The throughput of a blockchain system is the number of blocks n_b that can be added to the system in a fixed time. A blockchain system scales with the number of nodes m in the system if $n_b \rightarrow \infty$ when $m \rightarrow \infty$. Particularly, a blockchain system fully scales with the number of nodes m if $n_b = \Omega(m)$.

It should be clear that the definition of scalability or fully scalability does not depend on the length of the time period chosen for throughput. It captures the possibility of speeding up blockchain generation with more participating nodes; consequently, classical blockchain systems like Bitcoin does not scale at all.

DEFINITION 3 (FINALITY-DURATION). The finality-duration of a blockchain system is the time difference between the time point when a block is appended and the time point when a block receives full verification.

We say the finality-duration of a blockchain system is low (or constant) if the finality-duration is independent of the nodes in the system; consequently, classical blockchain systems like Bitcoin has a low finality-duration because after a fixed number of blocks are appended, all the nodes start following the main chain, thus blocks on the main chain will receive full verification.

4 IMPOSSIBILITY RESULT

THEOREM 1. There does not exist a blockchain system that simultaneously satisfies (i) scalability; (ii) low finality-duration; and (iii) full verification.

The proof follows from a counting argument on the total number of verifications. See full version.

5 ACHIEVING FULL VERIFICATION WITH TRADE-OFF BETWEEN SCALABILITY AND FINALITY-DURATION

As we have mentioned before, classical blockchain systems like Bitcoin achieve full verification and low finality-duration at the cost of scalability, Harmony and Ethereum 2.0 achieve scalability and low finality-duration at the cost of full verification. In this section, we complement the characterization by constructing a non-linear blockchain system, which is a modified version of IOTA, and show that it achieves full verification and scalability at the cost of finality-duration. We will further characterize the trade-off between the scalability and finality-duration, allowing the system to balance these two parameters for different applications.

5.1 Non-linear Blockchain (NLB) Construction

We first propose a concrete construction of NLB that achieves both security and scalability under the agent model. Without loss of generality, we assume that each block only includes one transaction. In the following, we broaden the terms and use them interchangeably. We first define some concepts.

DEFINITION 4 (BLOCK DISTANCE, DESCENDANT, AND ANCESTOR). *Given two blocks A and B , we define the distance between the two blocks as the length of the shortest directly path from A to B , which is denoted as $d(A, B)$. If there is no such a directed path, we define $d(A, B) = \infty$. If $0 < d(A, B) < \infty$, we say A is a descendant of B , and B is an ancestor of A . For a block B and each $1 \leq k \leq \ell$, let $Anc(B, k) = \{X | d(X, B) = k\}$ and $Des(B, k) = \{d(B, X) = k\}$, where ℓ is a given parameter.*

The new NLB is constructed as follows:

- \mathcal{A} . The new NLB assumes that there is a pool of new transactions from which a user can select one to construct a new block, which refers to two previous blocks¹. The user then does lightweight mining to fix this information in the newly constructed block. Suppose that the newly built block is B , the user also verifies blocks in $Anc(B, k)$, $1 \leq k \leq \ell$, where ℓ is a pre-defined system parameter that determines how many previous blocks the producer of a new block should verify.
- \mathcal{C} . To check a block B , the algorithm first checks whether the block format is correct, including the verification of the mining outcome. The algorithm also checks whether B is finalized or not, which is determined by

$$num_{Anc} \leftarrow |\cup_{k=1}^{\ell} Anc(B, k)|.$$

If num_{Anc} is larger than the system pre-defined threshold, B is finalized.

- \mathcal{R} . Each block has a reward value and the system imposes an upper bound on the maximal reward offered by a transaction, so that the largest and smallest reward among transactions (and blocks) can differ by a factor at most Δ . The producer of the new block also receives rewards from previous blocks. Specifically, each block is associated with a uniform verification cost vrf , which is divided into ℓ parts such that $vrf_1 < vrf_2 < \dots < vrf_{\ell}$ and $\sum_{k=1}^{\ell} vrf_k = vrf$. For each $1 \leq k \leq \ell$, the producer of block B gets reward $vrf_k / |Des(X, k)|$ for each $X \in Anc(B, k)$. This means that the verification reward of vrf_k from block X is evenly distributed among all its descendants in $Des(B, k)$. Note that the reward will be only collected when the new block is finalized.
- \mathcal{E} . The constructed NLB adopts the largest-weighted-descendants principle (LWD) to eliminate disagreement. For each block B , let $Des(B) = \cup_{h \geq 1} \{X | d(B, X) = h\}$. If there are two conflicting blocks B, B' and $|Des(B)| > |Des(B')|$, then B prevails, that is, users will abandon B' together with all its descendants in the sense that a new block will not refer to any of these blocks.

¹Our analysis in this paper also works if a new block refers to any fixed constant (greater than or equal to 2) number of blocks. For ease of presentation, we take this number to be 2 throughout this paper.

5.2 Scalability and Finality-duration Analysis

We first give a high-level summarization of the workflow of the proposed NLB system. Transactions are generated over time and form a pool. Each transaction is associated with a distinct transaction reward and a fixed verification reward vrf . Each time, a miner will select one transaction from the pool and append a block, which refers to two previous blocks. Here the miner needs to decide two things: (i) which transaction to include, and (ii) which two previous blocks to refer to. As we assume that miners are rational players, they will strategically make their decisions to maximize their profits, and this section is devoted to analyzing the scalability and security of the system under an arbitrary Nash equilibrium.

We formalize the problem as follows. Let the pool consist of n transactions, with the transaction reward being p_1, p_2, \dots, p_n . Let m be the number of miners, with computational powers being u_1, u_2, \dots, u_m . As we mentioned, each miner will mine a new block by including one transaction from the pool. If multiple miners say, miners in the subset of S , all choose the same transaction, then they compete, and only one of them will succeed, and the probability that some miner $i \in S$ succeeds is $\frac{u_i}{\sum_{h \in S} u_h}$. If, however, all miners choose different transactions, then each of them can append a new block.

In the following section, we will analyze the scalability and finality-duration of the constructed NLB separately.

5.2.1 Scalability. For scalability, we are interested in how many different transactions from the pool can be selected by the miners simultaneously. Note that the more different transactions are chosen, the higher scalability is. By considering the situation where miners choose transactions simultaneously, we are considering the worst-case because if miners are selecting transactions at different times, later ones may be able to avoid conflicts with earlier ones. Let n be the number of available transactions and m be the number of miners, the following Theorem 2 implies that the system is scalable even in the worst case such that when there are sufficiently many transactions, the throughput will be $O(m/\Delta)$ where Δ is a system parameter part of \mathcal{R} . By controlling Δ , we can control the scalability of the system. In particular, when we set Δ to be a constant, the system becomes fully scale with the number of nodes.

The following of this subsection is devoted to proving Theorem 2.

THEOREM 2. *With probability at least $1 - \max\{e^{-\Theta(m)}, e^{-\Theta(n)}\}$, the number of blocks mined by m miners in an arbitrary Nash equilibrium is at least $\min\{c_1 m / \Delta, c_2 n\}$ for some universal constants c_1, c_2 .*

Notice that a Nash equilibrium always exists by allowing mixed strategies [11]. Towards the proof, we introduce some notations. For simplicity, let all the transaction rewards be $p_1 \geq p_2 \geq \dots \geq p_n$. By the design of our system we require that $p_1/p_n \leq \Delta$. Note that the strategy of a miner is to select one transaction. We consider the general mixed strategy of a miner where he/she can specify a probability for each transaction.

Consider an arbitrary Nash equilibrium and let $\pi^{(i)} = (\pi_1^{(i)}, \pi_2^{(i)}, \dots, \pi_n^{(i)})$ be the strategy of miner i in the equilibrium, where $\pi_j^{(i)}$ is the probability that he chooses transaction j . It is obvious that $\sum_{j=1}^n \pi_j^{(i)} = 1$ for any $1 \leq i \leq m$. Let $X_j^{(i)}$ be the 0-1 random

variable that indicates whether miner i chooses transaction j . Then $X_j^{(i)} = 1$ with probability $\pi_j^{(i)}$ and $X_j^{(i)} = 0$ with probability $1 - \pi_j^{(i)}$.

Consider the above Nash equilibrium. Intuitively, if only a small number of transactions are selected, then miners must have devoted their probabilities to a few transactions. Therefore, to show that a sufficient number of distinct transactions are selected in expectation by miners, we need to show that the miners are distributing their probabilities in a fair way among transactions, as is implied by the following lemma.

LEMMA 1. *If there exists some transaction j_1 such that $\sum_{i=1}^m \pi_{j_1}^{(i)} \geq 12\Delta$, then for every transaction j it holds that $\sum_{i=1}^m \pi_j^{(i)} \geq 1/2$.*

PROOF. Suppose on the contrary that the lemma is not true, that is, there exists some transactions j_1 and j_2 such that $\sum_{i=1}^m \pi_{j_1}^{(i)} \geq 12\Delta$ and $\sum_{i=1}^m \pi_{j_2}^{(i)} < 1/2$. Consider the set of miners that choose transaction j_1 with positive probability. For simplicity, let these miners be miner $1, 2, \dots, k$ such that $u_1 \geq u_2 \geq \dots \geq u_k$. We show in the following that miner k can change his strategy to get a strictly higher profit, contradicting the fact that this is a Nash equilibrium, and consequently, the lemma is proved. More precisely, we argue that player k can get strictly larger profit (in expectation) by increasing his probability of choosing transaction j_2 and meanwhile decreasing his probability of choosing j_1 .

The expected profit that miner k can get from transaction j_1 and j_2 using his current strategy is equal to

$$p_{j_1} \mathbb{E}[\Gamma_1] + p_{j_2} \mathbb{E}[\Gamma_2],$$

where for $h = 1, 2$, we have

$$\Gamma_h = \begin{cases} 0, & \text{if } \sum_{i=1}^n u_i X_{j_h}^{(i)} = 0 \\ \frac{u_k X_{j_h}^{(k)}}{\sum_{i=1}^m u_i X_{j_h}^{(i)}}, & \text{Otherwise.} \end{cases}$$

If k changes his strategy by choosing j_1 with the probability of 0 and choosing j_2 with the probability of $\pi_{j_1}^{(k)} + \pi_{j_2}^{(k)}$, then the expected profit he can get from j_1 and j_2 is equal to $p_{j_2} \mathbb{E}[\tilde{\Gamma}_2]$, where

$$\tilde{\Gamma}_2 = \begin{cases} 0, & \text{if } \sum_{i \neq k} u_i X_{j_2}^{(i)} + u_k \tilde{X}_{j_2}^{(k)} = 0 \\ \frac{u_k \tilde{X}_{j_2}^{(k)}}{\sum_{i \neq k} u_i X_{j_2}^{(i)} + u_k \tilde{X}_{j_2}^{(k)}}, & \text{Otherwise.} \end{cases}$$

and $\tilde{X}_{j_2}^{(k)}$ is the 0-1 random variable that takes the value 1 with the probability of $\pi_{j_1}^{(k)} + \pi_{j_2}^{(k)}$.

In the following we show that

$$p_{j_2} \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2] > p_{j_1} \mathbb{E}[\Gamma_1],$$

which implies the correctness of the lemma. We prove the following two claims.

$$\text{CLAIM 1. } \mathbb{E}[\Gamma_1] < \frac{\pi_{j_1}^{(k)}}{2\Delta}.$$

$$\text{CLAIM 2. } \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2] \geq \pi_{j_1}^{(k)}/2.$$

Given the two claims and the fact that $p_{j_2} \leq \Delta p_{j_1}$, $p_{j_2} \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2] > p_{j_1} \mathbb{E}[\Gamma_1]$ follows and the lemma is proved. The proof of the two claims is quite involved. Please refer to the full version for details. \square

Lemma 1 shows that: Either no transaction has received a total amount of probability that is larger than 12Δ , or every transaction receives a total amount of probability at least $1/2$. Note that the two cases are not mutually exclusive. Nevertheless, we show in the following that in both cases, miners will select sufficiently many transactions with very high probability. The proofs of the following lemmas are mathematically involved. Please refer to the full version for details.

LEMMA 2. *Let $k, 1/\delta \in \mathbb{N}_+$ such that $k\delta = \zeta \in (0, 1)$. Let a_1, a_2, \dots, a_n be n numbers such that $a_i \in [0, \delta]$, $\sum_{i=1}^n a_i = 1$. Let $\mathcal{A} = \{S \mid S \subseteq \{1, 2, \dots, n\}, |S| = k\}$. Then we have*

$$\sum_{S \in \mathcal{A}} \left(\sum_{j \in S} a_j \right)^m \leq e^k \zeta^{m-k}.$$

LEMMA 3. *If $\sum_{i=1}^m \pi_j^{(i)} < 12\Delta$ holds for every transaction j , and $\Delta \leq m/12$, then the probability that only $k \leq \frac{m}{12e\Delta} = \Theta(m/\Delta)$ different transactions are selected by m miners is at most $(1/e)^{\Theta(m)}$.*

Note that if $\Delta > m/12$, $\frac{m}{12e\Delta} < 1$. As miners complete at least 1 transaction, the lemma is trivially true. Now we consider the other case and have the following.

LEMMA 4. *If $\sum_{i=1}^m \pi_j^{(i)} \geq 1/2$ holds for every transaction j , then the probability that no more than n/e^2 transactions are selected is at most $(1/e)^{\Theta(n)}$.*

Given Lemma 3 and Lemma 4, Theorem 2 follows directly.

5.2.2 Finality-duration. For ease of presentation, let $s = \min\{c_1 m/\Delta, c_2 n\}$ (recall Theorem 2). We will characterize finality-duration in terms of s .

Recall that a miner needs to make two decisions: (i) which transaction to include in the new block, and (ii) which two previous blocks to refer to. The two decisions are independent. In the previous subsection we have discussed (i), and in this subsection we focus on (ii), as this affects how the DAG grows.

It should be clear that since the verification reward of a transaction (block) is evenly distributed among miners who append a block of the same distance to it, a miner always prefers a block with no descendants. At any particular time t , we call a block without descendants as a leaf at t , and denote by L_t the set of leaves. We are interested in the size of L_t . Notice that in the classical blockchain system, $|L_t|$ is 1 since it is a chain. However, in a non-linear model, $|L_t|$ is not necessarily 1. In principle, $|L_t|$ could grow arbitrarily large, but what we will show in this section is that, $|L_t|$ is always bounded when miners are taking their equilibrium strategies. In this case, although we are considering a non-linear model, it is "almost linear", as implied by Theorem 3. Based on this result, we further leverage the techniques from random walk to prove that, for every block, after a delay of $O(s \log s)$ units of time, all blocks will be its descendant (Theorem 4), consequently, if we set $\ell = \Theta(s \log s)$ in our design, every block will be verified by all the users, and security follows.

As we mentioned before, each new block will refer to two leaves in L_t . As every block offers the same total amount of verification reward, every leaf appears the same to the miners (unless they are in conflict with previous blocks and then miners will be biased based on the LWD rule). Therefore, a new block will randomly select

two leaves to refer to. Assuming leaves are not conflicting with previous blocks, we show that $|L_t|$ will be $O(s)$ in the long run with an extremely high probability.

First, it is easy to see that if $|L_t| \leq s$, then $L_{t+1} \geq s$ as the s new blocks will be leaves at $t + 1$. The following lemma shows that if $|L_t|$ is sufficiently large, then with very high probability it will reduce to $O(s)$ after enough time.

LEMMA 5. *Let ϵ be an arbitrary small constant. If $|L_t| \geq 1/\epsilon^3$ and $|L_t| \geq 4s$, then with sufficiently high probability (at least $1 - O(\epsilon)$), $|L_{t+1}| = |L_t| - X + s \leq |L_t| - \frac{(1-3\epsilon)s}{2}$, i.e., L_t decreases by at least $\Omega(s)$.*

PROOF. Consider an arbitrary L_t . For any $\tau_i \in L_t$ and any new block b , the probability that b refers to τ_i is $2/|L_t|$, hence, the probability that none of the s new blocks refer to τ_i is $(1 - 2/|L_t|)^s$, i.e., the probability that an arbitrary τ_i has descendant(s) at $t + 1$ is $\theta = 1 - (1 - 2/|L_t|)^s$.

Let a random variable X_i denote whether the event that τ_i has descendant(s) at $t + 1$, then $\Pr(X_i = 1) = \theta$ and $\Pr(X_i = 0) = 1 - \theta$. Denote by $X = \sum_{i=1}^{|L_t|} X_i$ the total number of leaves in L_t that has descendant(s) at $t + 1$.

According to Chernoff bound, with the probability of at least $1 - 2e^{-\epsilon^2 |\mathbb{E}X|/3}$, $(1 - \epsilon)\mathbb{E}X \leq X \leq (1 + \epsilon)\mathbb{E}X$. Now we estimate $\mathbb{E}X$. It is easy to verify that

$$2s - \frac{2s(s-1)}{|L_t|} \leq \mathbb{E}X = |L_t| - |L_t|(1 - \frac{2}{|L_t|})^s \leq 2s.$$

Hence, if $|L_t| \geq 1/\epsilon^3$ and $|L_t| \geq 4s$, then with sufficiently high probability (at least $1 - 2e^{-1/(3\epsilon)} = 1 - O(\epsilon)$), $X \geq (1 - \epsilon)\mathbb{E}X \geq 3/2 \cdot (1 - \epsilon)s$, consequently,

$$|L_{t+1}| = |L_t| - X + s \leq |L_t| - \frac{(1-3\epsilon)s}{2}. \quad \square$$

The above lemma shows that if $|L_t|$ is large, then with high probability $|L_t|$ shall decrease, however, what we are interested in is the probability that $|L_t| \leq O(s)$ for all $t \geq 0$. Towards this, we need to cast the problem as a *random walk*. Lemma 5 shows that with the probability of $(1 - O(\epsilon))^3 = 1 - O(\epsilon)$, $|L_t|$ can decrease by $\frac{3(1-3\epsilon)s}{2} \geq s$, while with probability of at most $O(\epsilon)$, $|L_t|$ can increase by at most s . This can be interpreted as a random walk which walks right (increase) by s steps with the probability of $1 - O(\epsilon)$, and walks left (decrease) by s steps with the probability of $O(\epsilon)$. The following lemma is proved for a general random walk.

LEMMA 6 ([5], PP.272). *Consider a random walk starting at $RW_0 = 0$, $\Pr(RW_{i+1} - RW_i = s) = p$, $\Pr(RW_{i+1} - RW_i = -s) = q$ where $p + q = 1$ and $s \in \mathbb{Z}_{>0}$. If $p > q$, then*

$$\lim_{n \rightarrow \infty} \Pr(RW_i \geq 0, \forall 1 \leq i \leq n) = \frac{p - q}{p}.$$

If $p < q$, the above limit is 0.

Now we are ready to prove the following theorem.

THEOREM 3. *Let ϵ be a small constant such that $s > 1/\epsilon^3$. With very high probability (at least $1 - O(\epsilon)$), $|L_t| \leq 5s$ for all $t \geq 0$.*

PROOF. Recall that $|L_0| = 0$. Let t^* be the smallest time where $|L_{t^*}| \geq 4s$, then $|L_{t^*}| \leq 5s$. Now we take t^* as a starting time, $|L_{t^*}|$

as a starting point and take the random walk interpretation. Using Lemma 6, we have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr(|L_t| \leq |L_{t^*}|, \forall 1 \leq t \leq n) &\leq \frac{1 - O(\epsilon) - O(\epsilon)}{1 - O(\epsilon)} \\ &= 1 - O(\epsilon). \end{aligned}$$

Therefore, the probability that $|L_t|$ is bounded by $5s$ for all $t \geq 0$ is at least $1 - O(\epsilon)$. \square

LEMMA 7. *Let ϵ be a small constant such that $s > 1/\epsilon^3$. For any transaction at t that is not in conflict with prior transactions, with sufficiently high probability (at least $1 - O(\epsilon)$) every block appended at or after $t + O(s \log s)$ will be its descendant.*

Given the above lemma, if we set ℓ , the verification depth to be $\ell \geq O(s \log s)$, then any transaction at t will be verified by all the users after $O(s \log s)$ units of time with high probability. The following theorem is thus true.

THEOREM 4. *If $s > 1/\epsilon^3$ and $\ell \geq O(s \log s)$, then with probability at least $1 - O(\epsilon)$, any transaction at t will be verified by all the users after $O(s \log s)$ units of time.*

Remark. Recall that the value of $s = \min\{c_1 m / \Delta, c_2 n\}$ can be adjusted by setting different values of the system parameter Δ . Theorem 4 shows the trade-off between the scalability and finality-duration.

6 CONCLUSION

We provide the first systematic analysis on blockchain systems with respect to three major parameters, full verification, scalability, and finality-duration. We establish an impossibility result showing no blockchain system can simultaneously achieve the three properties. We complement the existing blockchain systems by establishing the first NLB that achieves both full verification and scalability. We also reveal, for the first time, the trade-off between scalability and finality-duration in NLB. It is not clear whether a better trade-off exists or not.

REFERENCES

- [1] Muhammad Salek Ali, Massimo Vecchio, Miguel Pincheira, Koustabh Dolui, Fabio Antonelli, and Mubashir Husain Rehmani. Applications of blockchains in the internet of things: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 2018.
- [2] Xavier Boyen, Christopher Carr, and Thomas Haines. Blockchain-free cryptocurrencies. 2017.
- [3] David Chaum. Blind signatures for untraceable payments. In *Advances in cryptography*, pages 199–203. Springer, 1983.
- [4] Mauro Conti, E Sandeep Kumar, Chhagan Lal, and Sushmita Ruj. A survey on security and privacy issues of bitcoin. *IEEE Communications Surveys & Tutorials*, 20(4):3416–3452, 2018.
- [5] William Feller. *An introduction to probability theory and its applications: volume 1*, volume 3. John Wiley & Sons New York, 1968.
- [6] BF França. Homomorphic mini-blockchain scheme, 2015.
- [7] Merve Can Kus Khalilov and Albert Levi. A survey on anonymity and privacy in bitcoin-like digital cash systems. *IEEE Communications Surveys & Tutorials*, 20(3):2543–2585, 2018.
- [8] Ziyao Liu, Nguyen Cong Luong, Wenbo Wang, Dusit Niyato, Ping Wang, Ying-Chang Liang, and Dong In Kim. A survey on applications of game theory in blockchain. *arXiv preprint arXiv:1902.10865*, 2019.
- [9] Ian Miers, Christina Garman, Matthew Green, and Aviel D Rubin. Zerocoin: Anonymous distributed e-cash from bitcoin. In *Security and Privacy (SP), 2013 IEEE Symposium on*, pages 397–411. IEEE, 2013.
- [10] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008.
- [11] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.
- [12] Serguei Popov, Olivia Saa, and Paulo Finardi. Equilibria in the tangle. *Computers & Industrial Engineering*, 136:160–172, 2019.
- [13] Serguei Popov. The tangle. *cit. on*, page 131, 2016.
- [14] Tara Salman, Maede Zolanvari, Aiman Erbad, Raj Jain, and Mohammed Samaka. Security services using blockchains: A state of the art survey. *IEEE Communications Surveys & Tutorials*, 21(1):858–880, 2018.
- [15] Tomas Sander and Amnon Ta-Shma. Auditable, anonymous electronic cash. In *Annual International Cryptology Conference*, pages 555–572. Springer, 1999.
- [16] Eli Ben Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. Zerocash: Decentralized anonymous payments from bitcoin. In *2014 IEEE Symposium on Security and Privacy*, pages 459–474. IEEE, 2014.
- [17] The Ethereum Team. On sharding blockchains, 2019.
- [18] The Harmony Team. Harmony - technical whitepaper, 2018.
- [19] The Zilliqa Team. The zilliqa technical whitepaper, 2017.
- [20] Nikolaos Petros Triantafyllidis and TNO Oskar van Deventer. Developing an ethereum blockchain application. 2016.
- [21] Florian Tschorsch and Björn Scheuermann. Bitcoin and beyond: A technical survey on decentralized digital currencies. *IEEE Communications Surveys & Tutorials*, 18(3):2084–2123, 2016.
- [22] Marko Vukolić. The quest for scalable blockchain fabric: Proof-of-work vs. bft replication. In *International Workshop on Open Problems in Network Security*, pages 112–125. Springer, 2015.
- [23] Ruizhe Yang, F Richard Yu, Pengbo Si, Zhaoxin Yang, and Yanhua Zhang. Integrated blockchain and edge computing systems: A survey, some research issues and challenges. *IEEE Communications Surveys & Tutorials*, 2019.

APPENDIX

A PROOF OF THEOREM 1

PROOF OF THEOREM 1. Suppose, on the contrary, that there exists such a blockchain. Then by definition, every block or transaction will receive verifications from all the nodes within a constant delay. Let c_0 be the constant delay. Consider an arbitrary node x and let τ_x be the fixed time it takes for node x to perform one verification. Let the throughput of the blockchain be n_b , then by definition of scalability we have $n_b = n_b(m) \rightarrow \infty$ when $m \rightarrow \infty$. Note that all the n_b blocks generated shall be verified by every node within the delay of c_0 , which means every node should perform n_b verifications within c_0 . However, node x can only perform c_0/τ_x verifications, which is a constant. Since $n_b \rightarrow \infty$, when m is sufficiently large, $n_b > c_0/\tau_x$. Therefore, it is impossible for an arbitrary node x to complete all the verifications. Hence, the three properties, scalability, low finality-duration and full verifications, cannot be satisfied simultaneously. \square

B PROOF OF CLAIM 1

PROOF OF CLAIM 1. Let $X = \sum_{i \neq k} X_{j_1}^{(i)}$ and $\mu = \mathbb{E}[X] \geq 5\Delta/\theta - \mathbb{E}[X_{j_1}^{(k)}] = 6\Delta/\theta - \pi_{j_1}^{(k)}$. For any $\delta \in (0, 1)$, we have

$$\begin{aligned} \mathbb{E}[\Gamma_1] &= \mathbb{E}[\Gamma_1 | X > (1-\delta)\mu] \Pr(X > (1-\delta)\mu) \\ &\quad + \mathbb{E}[\Gamma_1 | X \leq (1-\delta)\mu] \Pr(X \leq (1-\delta)\mu) \end{aligned}$$

Given that $\Gamma_1 \leq 1$, we know that

$$\begin{aligned} &\mathbb{E}[\Gamma_1 | X \leq (1-\delta)\mu] \\ &= \mathbb{E}[\Gamma_1 | X \leq (1-\delta)\mu, X_{j_1}^{(k)} = 1] \Pr(X_{j_1}^{(k)} = 1) \\ &\quad + \mathbb{E}[\Gamma_1 | X \leq (1-\delta)\mu, X_{j_1}^{(k)} = 0] \Pr(X_{j_1}^{(k)} = 0) \\ &= \pi_{j_1}^{(k)} \cdot \mathbb{E}[\Gamma_1 | X \leq (1-\delta)\mu, X_{j_1}^{(k)} = 1] \leq \pi_{j_1}^{(k)}. \end{aligned}$$

Meanwhile, by $u_1 \geq u_2 \geq \dots \geq u_k$, we have

$$\Gamma_1 \leq \frac{X_{j_1}^{(k)}}{\sum_{i=1}^k X_{j_1}^{(i)}}.$$

According to Chernoff bound, we know that

$$\Pr(X \leq (1-\delta)\mu) \leq e^{-\frac{\delta^2\mu}{2}}.$$

Therefore,

$$\begin{aligned} \mathbb{E}[\Gamma_1] &= \mathbb{E}[\Gamma_1 | X > (1-\delta)\mu] \Pr(X > (1-\delta)\mu) \\ &\quad + \mathbb{E}[\Gamma_1 | X \leq (1-\delta)\mu] \Pr(X \leq (1-\delta)\mu) \\ &\leq \mathbb{E}\left[\frac{X_{j_1}^{(k)}}{(1-\delta)\mu + X_{j_1}^{(k)}}\right] + e^{-\frac{\delta^2\mu}{2}} \cdot \pi_{j_1}^{(k)} \end{aligned}$$

Consider the function $\varphi(x) = \frac{x}{(1-\delta)\mu+x}$, it is easy to verify that $-\varphi(x)$ is convex in $x \in (0, +\infty)$, therefore, by Jensen's inequality we have

$$\begin{aligned} \frac{\mathbb{E}[X_{j_1}^{(k)}]}{(1-\delta)\mu + \mathbb{E}[X_{j_1}^{(k)}]} &= \varphi(\mathbb{E}[X_{j_1}^{(k)}]) \geq \mathbb{E}[\varphi(X)] \\ &= \mathbb{E}\left[\frac{X_{j_1}^{(k)}}{(1-\delta)\mu + X_{j_1}^{(k)}}\right]. \end{aligned}$$

Now consider the function xe^{-x} . It is easy to verify that the function decreases when $x \geq 1$, therefore $e^{-x} \leq \frac{1}{ex}$ for $x \geq 1$, hence, for $\frac{\delta^2\mu}{2} \geq 1$, we have

$$\begin{aligned} \mathbb{E}[\Gamma_1] &\leq \frac{\mathbb{E}[X_{j_1}^{(k)}]}{(1-\delta)\mu + \mathbb{E}[X_{j_1}^{(k)}]} + e^{-\frac{\delta^2\mu}{2}} \cdot \pi_{j_1}^{(k)} \\ &\leq \frac{\pi_{j_1}^{(k)}}{(1-\delta)\mu + \pi_{j_1}^{(k)}} + \frac{2}{e\delta^2\mu} \cdot \pi_{j_1}^{(k)}. \end{aligned}$$

Using the fact that $\mu \geq 12\Delta - \pi_{j_1}^{(k)} \geq 11\Delta$ and taking $\delta = 1/2$, we have

$$\mathbb{E}[\Gamma_1] \leq \frac{\pi_{j_1}^{(k)}}{6\Delta} + \frac{2\pi_{j_1}^{(k)}}{e \cdot 1/4 \cdot 11\Delta} < \frac{\pi_{j_1}^{(k)}}{2\Delta}. \quad \square$$

C PROOF OF CLAIM 2

PROOF OF CLAIM 2. Let $Y = \sum_{i \neq k} X_{j_2}^{(i)}$. Then

$$\begin{aligned} \Pr(Y = 0) &= \Pr(X_{j_2}^{(i)} = 0, \forall i \neq k) = \prod_{i \neq k} (1 - \pi_{j_2}^{(i)}) \\ &\geq 1 - \sum_{i \neq k} \pi_{j_2}^{(i)} \\ &\geq 1 - (1/2 - \pi_{j_2}^{(k)}) \\ &= 1/2 + \pi_{j_2}^{(k)}. \end{aligned}$$

Note that $X_{j_2}^{(k)}$ and Y are independent, hence we have the following,

$$\begin{aligned} &\mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2] \\ &\geq \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0] \Pr(Y = 0, \tilde{X}_{j_2}^{(k)} > 0) \\ &\geq (\pi_{j_1}^{(k)} + \pi_{j_2}^{(k)})(1/2 + \pi_{j_2}^{(k)}) \cdot \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0]. \end{aligned}$$

Further notice that $\tilde{X}_{j_2}^{(k)}$ and $X_{j_2}^{(k)}$ are independent, thus

$$\begin{aligned} &\mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0] \\ &= \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0, X_{j_2}^{(k)} = 0] \\ &\quad \cdot \Pr(X_{j_2}^{(k)} = 0 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0) \\ &\quad + \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0, X_{j_2}^{(k)} > 0] \\ &\quad \cdot \Pr(X_{j_2}^{(k)} > 0 | Y = 0, \tilde{X}_{j_2}^{(k)} > 0) \\ &= \mathbb{E}\left[\frac{u_k \tilde{X}_{j_2}^{(k)}}{Y + u_k \tilde{X}_{j_2}^{(k)}} \middle| Y = 0, \tilde{X}_{j_2}^{(k)} > 0, X_{j_2}^{(k)} = 0\right] \cdot \Pr(X_{j_2}^{(k)} = 0) \\ &\quad + \mathbb{E}\left[\frac{u_k \tilde{X}_{j_2}^{(k)}}{Y + u_k \tilde{X}_{j_2}^{(k)}} - \frac{u_k X_{j_2}^{(k)}}{Y + u_k X_{j_2}^{(k)}} \middle| Y = 0, \tilde{X}_{j_2}^{(k)} > 0, X_{j_2}^{(k)} > 0\right] \\ &\quad \cdot \Pr(X_{j_2}^{(k)} > 0) \\ &= 1 - \pi_{j_2}^{(k)} \end{aligned}$$

Now we have

$$\mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2] \geq (\pi_{j_1}^{(k)} + \pi_{j_2}^{(k)})(1/2 + \pi_{j_2}^{(k)})(1 - \pi_{j_2}^{(k)})$$

For simplicity, let $\tau = \pi_{j_1}^{(k)} + \pi_{j_2}^{(k)} \in [0, 1]$ and $x = \pi_{j_2}^{(k)}$, then we have

$$f(x) = \mathbb{E}[\tilde{\Gamma}_2 - \Gamma_2] - 1/2 \cdot \pi_{j_1}^{(k)} = \tau(1/2 + x)(1 - x) - 1/2 \cdot (\tau - x).$$

Notice that f is a quadratic function in x whose quadratic term has negative coefficient, therefore if $f(0) \geq 0$ and $f(1) \geq 0$, then for any $x \in [0, 1]$ we have $f(x) \geq 0$ and the claim is proved. It is easy to see that $f(0) = 0$, and $f(1) = 1/2 \cdot (1 - \tau) \geq 0$, hence, the claim is proved. \square

D PROOF OF LEMMA 2

PROOF. Without loss of generality we assume that $a_1 \geq a_2 \geq \dots \geq a_n$. We define

$$f(a_1, a_2, \dots, a_n) = \sum_{S \in \mathcal{A}} \left(\sum_{j \in S} a_j \right)^m.$$

For arbitrary $i < j$ and some small $\epsilon \geq 0$ such that $a_{i-1} \geq a_i + \epsilon$, $a_j - \epsilon \geq a_{j+1}$, we prove in the following that

$$f(\dots, a_{i-1}, a_i + \epsilon, a_{i+1}, \dots, a_{j-1}, a_j - \epsilon, a_{j+1}, \dots) \geq f(a_1, a_2, \dots, a_n).$$

For simplicity we write $f(a_1, a_2, \dots, a_n) = f_{i,j}(a_i, a_j)$. We have

$$\begin{aligned} & f_{i,j}(a_i + \epsilon, a_j - \epsilon) \\ &= \sum_{\substack{S: S \in \mathcal{A} \\ i, j \notin S}} \left(\sum_{h \in S} a_h \right)^m + \sum_{\substack{S: S \in \mathcal{A} \\ i \in S, j \notin S}} (a_i + \epsilon + \sum_{h \in S, h \neq i} a_h)^m \\ &+ \sum_{\substack{S: S \in \mathcal{A} \\ i \notin S, j \in S}} (a_j - \epsilon + \sum_{h \in S, h \neq j} a_h)^m + \sum_{\substack{S: S \in \mathcal{A} \\ i, j \in S}} \left(\sum_{h \in S} a_h \right)^m \end{aligned}$$

Consider the function $g(x) = x^m$. Due to its convexity, for $x + \epsilon > x > y > y - \epsilon$, we have

$$g(x + \epsilon) + g(y - \epsilon) \geq g(x) + g(y).$$

Hence, let $\mathcal{A}' = \{S | S \subseteq \{1, 2, \dots, n\} \setminus \{i, j\}, |S| = k - 1\}$, then for any $S' \in \mathcal{A}'$, we have

$$\begin{aligned} & \left(\sum_{h \in S'} a_h + a_i + \epsilon \right)^m + \left(\sum_{h \in S'} a_h + a_j - \epsilon \right)^m \\ & \geq \left(\sum_{h \in S'} a_h + a_i \right)^m + \left(\sum_{h \in S'} a_h + a_j \right)^m, \end{aligned}$$

therefore, $f_{i,j}(a_i + \epsilon, a_j - \epsilon) \geq f_{i,j}(a_i, a_j)$. Now we can iteratively change a_1, a_2, \dots, a_n into a'_1, a'_2, \dots, a'_n such that $a'_i = \delta$ for $1 \leq i \leq 1/\delta$, and $a'_i = 0$ otherwise, and get

$$f(a_1, a_2, \dots, a_n) \leq f(a'_1, a'_2, \dots, a'_n).$$

It is not difficult to compute that

$$f(a'_1, a'_2, \dots, a'_n) = \binom{1/\delta}{k} (k\delta)^m \leq \frac{(1/\delta)^k}{k!} \cdot \zeta^m \leq \left(\frac{e}{\zeta}\right)^k \zeta^m,$$

where for the last inequality we make use of the Stirling's approximation that $k! \geq (k/e)^k$. Hence, the lemma is proved. \square

E PROOF OF LEMMA 4

PROOF OF LEMMA 4. Consider the event that at most θn transactions are selected by m miners for some $\theta \in (0, 1)$. Note that in this case $m = \sum_{j=1}^n \sum_{i=1}^m \pi_j^{(i)} \geq n/2$, hence $n \leq 2m$. Again we define $\mathcal{A} = \{S | S \subseteq \{1, 2, \dots, n\}, |S| = (1 - \theta)n\}$ as the superset of all the subsets of cardinality $(1 - \theta)n$. The probability that miner i does not select any transaction in some $S \in \mathcal{A}$ is $1 - \sum_{j \in S} \pi_j^{(i)}$. Given that miners select transactions independently, the probability that all miners do not select transactions in S is

$$\begin{aligned} \prod_{i=1}^m \left(1 - \sum_{j \in S} \pi_j^{(i)} \right) & \leq \left(\frac{m - \sum_{j \in S} \sum_{i=1}^m \pi_j^{(i)}}{m} \right)^m \\ & \leq \left(\frac{m - \frac{(1-\theta)n}{2}}{m} \right)^m \leq e^{-\frac{(1-\theta)n}{2}}, \end{aligned}$$

where the first inequality follows by inequality of arithmetic and geometric means, the second inequality follows by the fact that $\sum_{i=1}^m \pi_j^{(i)} \geq 1/2$, and the third inequality follows by $(1 - 1/x)^x \leq e^{-1}$ for $x \geq 1$. Taking the summation over all possible $S \in \mathcal{A}$, the probability that at most θn transactions are selected is at most

$$\begin{aligned} \sum_{S \in \mathcal{A}} \prod_{i=1}^m \left(1 - \sum_{j \in S} \pi_j^{(i)} \right) & \leq \binom{n}{\theta n} e^{-\frac{(1-\theta)n}{2}} \\ & \leq \frac{n^{\theta n}}{(\theta n)!} \cdot e^{-\frac{(1-\theta)n}{2}} \\ & \leq \left(\frac{e}{\theta}\right)^{\theta n} \cdot e^{-\frac{(1-\theta)n}{2}}. \end{aligned}$$

Taking $\theta = e^{-2}$, simple calculation shows that the right side of the inequality above is e^{cn} for $c \leq -0.02$, and the lemma is proved. \square

F PROOF OF LEMMA 3

PROOF OF LEMMA 3. Consider the event that at most k different transactions are selected by m miners. Let $\mathcal{A} = \{S | S \subseteq \{1, 2, \dots, n\}, |S| = k\}$ be the superset of all subsets of cardinality k . For any $S \in \mathcal{A}$, let σ_S be any mapping that maps $\{1, 2, \dots, m\}$ to S , and Ω_S be the set of all such mappings. Consider the event that miner i selects transaction $\sigma_S(i)$, we know that the event happens with the probability $\Pr(\sigma_S) = \prod_{i=1}^m \pi_{\sigma_S(i)}^{(i)}$. Taking summation over all possible mappings, the event that only transactions in S are selected is at most

$$\sum_{\sigma_S \in \Omega_S} \Pr(\sigma_S) = \sum_{\sigma_S \in \Omega_S} \prod_{i=1}^m \pi_{\sigma_S(i)}^{(i)} = \prod_{i=1}^m \sum_{j \in S} \pi_j^{(i)}.$$

The last equality follows as when we expand $\prod_{i=1}^m \sum_{j \in S} \pi_j^{(i)}$, any summand corresponds to $\prod_{i=1}^m \pi_{\sigma(i)}^{(i)}$ for some mapping σ . Using the inequality of arithmetic and geometric means, we have

$$\sum_{\sigma_S \in \Omega_S} \Pr(\sigma_S) = \prod_{i=1}^m \sum_{j \in S} \pi_j^{(i)} \leq \left(\frac{\sum_{i=1}^m \sum_{j \in S} \pi_j^{(i)}}{m} \right)^m = \left(\sum_{j \in S} \frac{\sum_{i=1}^m \pi_j^{(i)}}{m} \right)^m.$$

Taking summation over all possible $S \in \mathcal{A}$, the event that at most k transactions are selected is at most

$$\sum_{S \in \mathcal{A}} \sum_{\sigma_S \in \Omega_S} \Pr(\sigma_S) \leq \sum_{S \in \mathcal{A}} \left(\sum_{j \in S} \frac{\sum_{i=1}^m \pi_j^{(i)}}{m} \right)^m.$$

Now let $a_j = \frac{\sum_{i=1}^m \pi_j^{(i)}}{m}$. It is easy to see that $\sum_{j=1}^n a_j = 1$. Furthermore, since $\sum_{i=1}^m \pi_j^{(i)} < 12\Delta$, we have $a_j \leq 12\Delta/m$. We apply

Lemma 2 by taking $\delta = \frac{1}{\lceil m/(12\Delta) \rceil} \leq 12\Delta/(m+12\Delta)$, $\zeta = k\delta \leq e^{-1}$, we know that

$$\sum_{S \in \mathcal{A}} \sum_{\sigma_S \in \Omega_S} \Pr(\sigma_S) \leq e^k \zeta^{m-k} \leq e^{-m/2}.$$

Hence, the lemma is proved. \square

G PROOF OF LEMMA 7

PROOF OF LEMMA 7. According to Theorem 3, we focus on the event that $|L_t| \leq 5s$ for all $t \geq 0$, which happens with $1 - O(\epsilon)$ probability.

For $h \geq t$, let Ψ_h be the subset of blocks in L_h which has a directed path from some fixed block $\tau_0 \in L_t$, which is a random subset. Let $\psi_h = \mathbb{E}(|\Psi_h|)$. Consider L_{h+1} . For any block $\tau_i \in L_{h+1}$, let X_i be a binary random variable indicating whether τ_i refers to some block in Ψ_h , and hence admits a directed path from τ_0 . Then we know

$$\begin{aligned} \Pr(X_i = 1) &= \frac{\binom{|\Psi_h|}{2} + |\Psi_h|(|L_h| - |\Psi_h|)}{\binom{|L_h|}{2}} \\ &= \frac{|\Psi_h|(2|L_h| - |\Psi_h| - 1)}{|L_h|(|L_h| - 1)}. \end{aligned}$$

We consider $|\Psi_{h+1}|$. It is obvious that if $|\Psi_h| = |L_h|$, then every block in L_{h+1} refers to some block in Ψ_h and thus admits a directed path from τ_0 , hence, $|L_{h+1}| = |\Psi_{h+1}|$, and similarly we have $|L_{h+j}| = |\Psi_{h+j}|$ for all $j \geq 1$.

Otherwise, we assume $1 \leq |\Psi_h| \leq |L_h| - 1$. Then $2|L_h| - |\Psi_h| - 1 \geq |L_h|$, and we have

$$\mathbb{E}(X_i) = \mathbb{E}\left(\frac{|\Psi_h|(2|L_h| - |\Psi_h| - 1)}{|L_h|(|L_h| - 1)}\right) \geq \frac{\psi_h}{|L_h| - 1}.$$

Note that $|\Psi_{h+1}| = \sum_i X_i$. It is easy to calculate that

$$\psi_{h+1} = \mathbb{E}(|\Psi_{h+1}|) \geq \psi_h \left(1 + \frac{1}{|L_h| - 1}\right).$$

This means, starting from $\psi_t = 1$, for each ψ_h where $h \geq t$, either $\psi_h = |L_h|$ and thus $\psi_{h'} = |L_{h'}$ for all $h' \geq h$, or $\psi_{h+1} \geq \left(1 + \frac{1}{|L_h| - 1}\right) \psi_h$. Since $|L_h| \leq 5s$, it is easy to see that ψ_h increases to a value sufficiently close to $|L_h| \leq 5s$ when $h \geq t + O(s \log s)$, and the theorem is proved. \square