An Analysis of Blockchain Consistency in Asynchronous Networks: Deriving a Neat Bound

Jun Zhao[1], Jing Tang[2], Zengxiang Li[3], Huaxiong Wang[4], Kwok-Yan Lam[1], Kaiping Xue[5]

[1]School of Computer Science and Engineering, Nanyang Technological University, Singapore
[2]Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore
[3]Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore
[4]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
[5]Department of Information Security, University of Science and Technology of China, China

[1]{junzhao, kwokyan.lam}@ntu.edu.sg, [2]isejtang@nus.edu.sg, [3]liz@ihpc.a-star.edu.sg, [4]hxwang@ntu.edu.sg, [5]kpxue@ustc.edu.cn

***Abstract*—Formal analyses of blockchain protocols have received much attention recently. Consistency results of Nakamoto's blockchain protocol are often expressed in a quantity c, which denotes the expected number of network delays before some block is mined. With μ (resp., ν) denoting the fraction of computational power controlled by benign miners (resp., the adversary), where $\mu+\nu=1$, we prove for the first time that to ensure the consistency property of Nakamoto's blockchain protocol in an asynchronous network, it suffices to have c to be just slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$. Such a result is both neater and stronger than existing ones. In the proof, we formulate novel Markov chains which characterize the numbers of mined blocks in different rounds.**

***Keywords*—Blockchain, consistency, asynchronous networks, Markov chains.**

I. Introduction

Nakamoto's blockchain protocol [1] supports the Bitcoin application and relies on the proof of work (POW). POW means that to create a block, a player needs to provide a solution of a cryptographic puzzle based on hash functions. Formal analyses of the protocol have received considerable interest recently [2]–[5].

Garay, Kiayas and Leonardos [2] propose the first formal modeling for Nakamoto's blockchain protocol. They also identify conditions which enable Nakamoto's protocol to achieve a common prefix-property, where honest players' blockchain views have a large common prefix.

The model of [2] assumes a *synchronous* network. Removing such a strong assumption, Pass, Seeman, and Shelat [3] consider an *asynchronous* network by allowing the adversary to adaptively and individually delay messages up to a delay limit Δ. We refer to this as the Δ-delay model.

One of the desired properties in a blockchain protocol is consistency. In this paper, we follow [3], [6] to define consistency as the property that for any positive integer T, with overwhelming probability in T, for any two rounds r and s with $r<s$, all but the last T blocks in the chain of any honest player i at round r is a prefix of the chain of any honest player j at round s. For an event to have an overwhelming probability in T, the probability of its complementary event should decay at least exponentially with respect to T.

Consistency results of Nakamoto's blockchain protocol are typically expressed in a quantity c defined as $\frac{1}{pn\Delta}$, where p denotes the hardness of the proof of work, n is the number of players, and Δ is the maximum delay of a message by the adversary (the notation will be summarized in Table I on Page 2). Roughly speaking, c means the expected number of network delays before some block is mined.

In this paper, we present a result for the consistency property of Nakamoto's blockchain protocol. Our consistency result is stronger than existing ones in the literature (e.g., the result of [3]). Under the Δ-delay model, with μ (resp., ν) denoting the fraction of computational power controlled by benign miners (resp., the adversary), where $\mu+\nu=1$ and $0<\nu<\mu$, we show that it suffices to achieve consistency for c denoting $\frac{1}{pn\Delta}$ to be just slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$. Our work is the first one in the literature to derive such a neat expression $\frac{2\mu}{\ln(\mu/\nu)}$. In Section II-A,, we will explain the superiority of our consistency result over existing results.

Contributions. Our contributions are as follows:

- ***(Major) Contribution 1 of proving Theorem 1:*** Our Theorem 1 to be presented on Page 5 gives the following neat condition to ensure the consistency property of Nakamoto's blockchain protocol: c denoting the expected number of network delays before some block is mined just needs to be slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$, where μ (resp., ν) denotes the fraction of computational power controlled by benign miners (resp., the adversary).
- ***(Secondary) Contribution 2 of fixing [6] and proving Theorem 2:*** We show an issue in the analysis of [6]: the probability that *only one* honest miner succeeds in solving a puzzle is computed as the probability that *at least one* honest miner succeeds in solving a puzzle in one round. Although [6] mentions "a single honest mined block", but its calculation actually uses "at least one honest mined block". After we fix the above issue and correct some minor notation typos of [6], the result of [6] will become the same as our Theorem 2 on Page 5 (We emphasize that Theorem 2 is our secondary contribution while Theorem 1 is our major contribution). Yet, in an effort to present an clearer explanation than that of [6], we formulate two novel Markov chains to prove Theorem 2. With a state of a round characterizing the number of mined blocks (e.g., no, one, or over one mined block), our first Markov chain models the transition of a variable denoting the suffix of the concatenation of the previous states and the current state. Our second Markov chain models the transition of a variable which denotes the concatenation of i) the suffix of previous states before the Δ to last state, ii) the previous Δ states, and iii) the current state.

Table I: Notation and their meanings.

Notation	Meanings
p	the hardness of the proof of work
n	the number of miners (either honest or corrupted), each with identical computing power
Δ	the maximum delay of a message by the adversary
c	$c := \frac{1}{pn\Delta}$. Roughly speaking, c means the expected number of Δ-delays before some block is mined.
μ	the fraction of computational power controlled by benign miners (i.e., the fraction of benign miners)
ν	the fraction of computational power controlled by the adversary (i.e., the fraction of corrupted miners)
α	α denotes the probability that *at least one* honest miner succeeds in solving a puzzle in one round. $\alpha = 1-(1-p)^{\mu n}$.
$\overline{\alpha}$	$\overline{\alpha}$ denotes the probability that *no* honest miner succeeds in solving a puzzle in one round. $\overline{\alpha} = (1-p)^{\mu n}$.
α_1	α_1 denotes the probability that *only one* honest miner succeeds in solving a puzzle in one round. $\alpha_1 = p\mu n \times (1-p)^{\mu n-1}$.
β	β denotes the expected number of blocks mined in each round by the adversary controlling ν fraction of computational power. $\beta := p\nu n$.

Organization of this paper. In Section II, we survey related studies. Section III explains the model for Nakamoto's blockchain protocol. Section IV presents our results for the consistency property of Nakamoto's blockchain protocol. In Sections V and VI, we discuss the proofs of Theorems 1 and 2, respectively. We conclude the paper in Section VII. Additional proof details are given in the Appendices of the online full version [7].

Notation. Table I lists the notation and their meanings.

II. Related Work

This section is organized as follows. In Section II-A, we elaborate the comparison between our results and related ones, where Figure 1 is plotted. Section II-B presents additional related studies.

A. Comparing our results and related ones

We compare our consistency results with [3], [6] and use Figure 1 to illustrate the comparison. Our Figure 1 adopts $n = 10^5$ and $\Delta = 10^{13}$ from Figure 1 of [3]. In Figure 1, all lines except the magenta line illustrate conditions used in different results to ensure the consistency property of Nakamoto's blockchain protocol. In particular, these red, brown, green, and blue lines plot the allowed maximum (or the limit superior) value for the fraction ν of computational power controlled by the adversary with respect to c, the expected number of network delays before some block is mined, in order to not break consistency according to the respective results. More details are as follows.

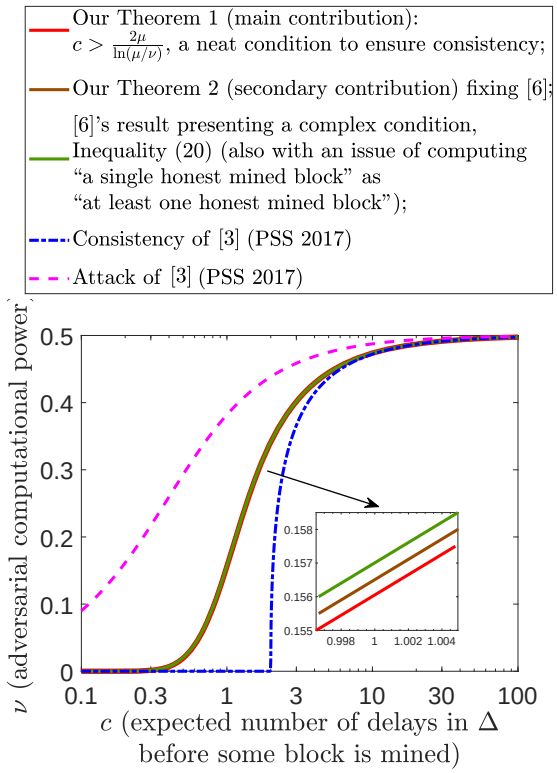

Fig. 1: A comparison of our consistency result with consistency of [6] by Kiffer, Rajaraman, and Shelat in ACM CCS 2018 as well as consistency and attack of [3] by Pass, Seeman, and Shelat (PSS) in Eurocrypt 2017. We adopt $n = 10^5$ and $\Delta = 10^{13}$ from Figure 1 of [3]. c denoting $\frac{1}{pn\Delta}$ roughly means the expected number of network delays before some block is mined. See Table I on the left-hand column for the meanings of the notation.

The red, brown, and green lines almost overlap in Figure 1. Hence, in the lower right corner of Figure 1, we also zoom some parts to show the (negligible) separation between the lines. The red line shows a neat condition on c to ensure consistency: $c > \frac{2\mu}{\ln(\mu/\nu)}$, given by our Theorem 1 (our main contribution) to be presented on Page 5. The brown line is from Theorem 2 (our secondary contribution) on Page 5. The green line shows Claim 1 on Page 5, which is Theorem 4.4 on Page 8 of [6] after we correct $\mu \cdot p$ to α in many places of [6] and perform some computations presented in Appendix A of the online full version [7] (see Table I for the notation's meanings). Yet, [6]'s result as well as its induced Claim 1 has a minor issue that the probability that *only one* honest miner succeeds in solving a puzzle is computed as the probability that *at least one* honest miner succeeds in solving a puzzle in one round, so we fix it to obtain Theorem 2 (our effort of obtaining Theorem 2 to fix [6] are based on new Markov chains, to present a more detailed explanation than [6]). In Claim 1 originating from [6]'s result, Inequality (20) as the condition on c for consistency is quite complex. In our Theorem 2 fixing [6], the condition on c for consistency is also quite complex after the expressions of α and α_1 are plugged in. In contrast, Theorem 1 as our main contribution presents the neat condition $c > \frac{2\mu}{\ln(\mu/\nu)}$ to ensure consistency. Since we obtain Theorem 1 based on Theorem 2, the neat condition of Theorem 1 is sufficient but not necessary to get the condition of Theorem 2. Yet, these two conditions are

almost the same since the red and brown lines almost overlap in Figure 1. This shows that we almost do not lose any tightness of the result in the move from Theorem 2 to Theorem 1 for seeking a neater condition.

Since the red, brown, and green lines almost overlap in Figure 1, we now focus on the red, blue, and magenta lines. As stated, the red line of Figure 1 shows our consistency result in Theorem 1 on Page 5. From the condition $c > \frac{2\mu}{\ln(\mu/\nu)} = \frac{2(1-\nu)}{\ln\frac{1-\nu}{\nu}}$, our maximal $\nu_{\max}$ can be solved numerically given c (strictly speaking, $\nu_{\max}$ cannot be achieved due to the strict inequality sign). This gives the red line.

The blue line of Figure 1 is from the consistency analysis of [3]. The consistency condition of [3] is $\alpha[1-(2\Delta+2)\alpha] > \beta$, where $\alpha := 1-(1-p)^{\mu n}$ and $\beta := \nu np$. Roughly speaking, $\alpha \approx \mu np$ and $2\Delta + 2 \approx 2\Delta$, so $\alpha[1-(2\Delta+2)\alpha] > \beta$ is approximately $1 - 2\Delta\mu np > \frac{\nu}{1-\nu}$, where we note $\mu = 1-\nu$. Then we further obtain $p < \frac{1-2\nu}{2(1-\nu)^2\Delta n}$ and hence $c := \frac{1}{pn\Delta} > \frac{2(1-\nu)^2}{1-2\nu}$. This implies $\nu < \frac{1}{2}(2 - c + \sqrt{c^2-2c})$, where $c > 2$. The blue line of Figure 1 shows this.

The magenta line of Figure 1 illustrates an attack of [3] which breaks consistency. Remark 8.5 of [3] presents an attack which works when $\frac{1}{c} > \frac{1}{\nu} - \frac{1}{1-\nu}$. This inequality means $\nu > \frac{2c+1-\sqrt{4c^2+1}}{2}$.

From Figure 1, the red line illustrating our consistency result is strictly above the blue line for consistency of [3]. Hence, our consistency result is much stronger than that of [3] in the sense that our result tolerates much more fraction of adversarial computational power. A future direction is to see whether it is possible to reduce the gap between the red line for our consistency result and the magenta line representing an attack on consistency from [3].

B. Additional related work

The essence of blockchain is a consensus protocol to achieve agreement among distributed nodes. The seminal blockchain protocol by Nakamoto [1] leads to the popular application of Bitcoin. Bitcoin is a cryptocurrency whose ledger is maintained by the public instead of trusted authorities.

Nakamoto's blockchain protocol is built on the proof of work (POW) [1]. When a node creates a block, the node should provide a solution of a cryptographic puzzle based on hash functions. Every node maintains its own chain and accepts the longest chain of the ones it receives from the network.

Recently, formal analyses of blockchain protocols have received considerable attention [2]–[5]. Three commonly analyzed properties are consistency, chain growth, and chain quality.

In [1], [2], *consistency* is defined as the property that with overwhelming probability in T, at any round, the chains of two honest players can differ only in the last T blocks. Pass, Seeman, and Shelat [3] identify that this definition is not sufficient for consensus, since it does not exclude a protocol which oscillates between different chains. Hence, they require an additional property, referred to as *future self-consistence*: with overwhelming probability in T, at any two rounds r and s, the chains of any honest player at r and s differs only in blocks within the last T blocks. The consistency notion used in [6] and our current paper combines the consistency definition of [1], [2] and future self-consistence of [3]. Specifically, by consistency, we mean that with overwhelming probability in T, for any two rounds r and s with $r < s$, all but the last T blocks in the chain of any honest player i at round r is a prefix of the chain of any honest player j at round s.

In addition to consistency analyzed by [1]–[3], [6], chain growth and chain quality for Nakamoto's blockchain protocol are also studied in the literature [3], [4], [8]–[10]. The chain growth is at least g if with overwhelming probability in T, the chain of honest players grew by at least T blocks in the last T/g rounds. The chain quality is at least q if with overwhelming probability in T, for any T consecutive blocks in any chain held by some honest player, the fraction of blocks contributed by honest players is at least q. In this paper, we analyze only consistency. A future direction is to investigate how to use our proof methods for the analyses of chain growth and chain quality.

After POW, blockchain protocols based on an alternative paradigm called the Proof of Stake (POS) have also been proposed [11]–[14]. POS typically consumes less computation power than POW. The ingenious Algorand protocol [15] combines POS and the classical practical Byzantine fault tolerance (PBFT) protocol of [16]. We refer interested readers to recent surveys [17], [18] for more details of POW, POS, and other types of blockchain protocols.

III. The Model for Nakamoto's Blockchain Protocol

As in many blockchain studies, we adopt the formalization of Garay, Kiayas and Leonardos [2] and Pass, Seeman, and Shelat [3] for Nakamoto's blockchain protocol. We will mostly follow the notation of [6], which presents a clear explanation of the formalization.

A blockchain is a pair of algorithms $(\Pi, \texttt{ext})$. The stateful algorithm Π maintains a local state variable $\mathcal{C}$ and also receives a security parameter κ as an input. The variable $\mathcal{C}$ is commonly referred to as the chain, since it contains a set of blocks. A block is an abstract record containing a message. The algorithm $\texttt{ext}(\kappa, \mathcal{C})$ outputs an ordered sequence of messages.

The execution of a blockchain protocol $(\Pi, \texttt{ext})$ is directed by an environment $Z(1^\kappa)$. It activates each of n players as either *honest* or *corrupt*. For simplicity, all n players are assumed to have identical computing power. Each honest player has a current view of the blockchain and aims to build blocks at the end of the chain. Each corrupted player is controlled by an adversary $\mathcal{A}$. We assume that at any point, $\mathcal{A}$ can corrupt an honest party or uncorrupt a corrupted player, but the fraction of corrupted players is at most ν. For ease of analysis, we can just consider the worst case where $\mathcal{A}$ controls ν fraction of corrupted players at each round.

We consider the network to be asynchronous, and allow the adversary $\mathcal{A}$ to have the following capabilities:

① $\mathcal{A}$ can delay and/or reorder all messages up to a delay of Δ rounds, but $\mathcal{A}$ cannot modify messages sent by honest players.

② $\mathcal{A}$ fully controls all corrupted players; i.e., $\mathcal{A}$ reads all their inputs/messages and sets their outputs/messages to be sent.

Strategies taken by the adversary $\mathcal{A}$ can be letting all corrupted players work on the same block or different ones.

All players have access to a random function $H : \{0,1\}^* \to \{0,1\}^\kappa$ through the following two oracles. First, $\mathtt{H}(x)$ simply outputs $H(x)$. Second, the verification oracle $\mathtt{H.ver}(x, y)$ outputs 1 if and only if $H(x) = y$ and 0 otherwise. How $\mathtt{H}$ and $\mathtt{H.ver}$ can be accessed is specified as follows:

- In each round, the players, as well as the adversary $\mathcal{A}$, make *any* number of queries to $\mathtt{H.ver}$.
- In each round, each honest player can make only a single query $\mathtt{H}$ and the queries made by honest players are *parallel* so that even if they manage to mine several blocks, their longest chain can increase by at most 1. In contrast, the adversary $\mathcal{A}$ controlling q players can make q *sequential* queries to $\mathtt{H}$.

The above model captures that we account for only the effort of *finding* a solution to a "proof of work", and consider that *checking* the validity of a solution is negligible. A "proof of work" given the block h_{-1} and message m is to find a string η such that $\mathtt{H}(h_{-1}, \eta, m) \leq D_p$, where the blockchain protocol sets D_p such that the probability of finding η to satisfy the above relation is p. This quantity p is referred to as the hardness of the proof of work.

Given the above, we now describe an execution of a blockchain protocol. At the beginning, the environment $Z(1^\kappa)$ instantiate n players, which have identical computing power. The protocol proceeds in rounds as follows. At each round, each player i does the following:

- i receives blocks created by other players and includes the blocks in its chain based on the protocol Π;
- i can make at most one query to the oracle $\mathtt{H}$ and creates a block with probability p; and
- i receives some message from $Z(1^\kappa)$ and includes the message in the block that i tries to publish, where the message contains transactions to be included in the blockchain.

As already noted, ν denotes the fraction of corrupted players controlled by the adversary. With μ being the fraction of honest players, we have

$$\mu + \nu = 1. \tag{1}$$

Throughout the paper, we enforce

$$0 < \nu < \frac{1}{2} < \mu, \tag{2}$$

and the trivial condition

$$n \geq 4. \tag{3}$$

From Eq. (1), Inequality (2) simply means the following two conditions:

i) the fraction of computational power controlled by benign miners is greater than that controlled by the adversary; and
ii) the adversary controls non-zero fraction of computational power.

With n, p, μ, and ν introduced above, we now define α, $\overline{\alpha}$, and α_1, which will be used in our theorems to be presented in Section IV. All these notation are given in Table I on Page 2. The meanings of α, $\overline{\alpha}$, and α_1 are as follows:

$$\alpha: \text{ the probability that } \textit{at least one} \text{ honest miner, succeeds in solving a puzzle in one round,} \tag{4}$$

$$\overline{\alpha}: \text{ the probability that } \textit{no} \text{ honest miner, succeeds in solving a puzzle in one round,} \tag{5}$$

$$\alpha_1: \text{ the probability that } \textit{only one} \text{ honest miner, succeeds in solving a puzzle in one round.} \tag{6}$$

Next, we derive the expressions of α, $\overline{\alpha}$, and α_1. Since each honest node mines a block independently with probability p in a round, X denoting the number of blocks mined by the μn honest nodes in each round follows $\text{binom}(\mu n, p)$, which denotes a binomial distribution with μn being the number of trials and p being the success probability for each trial. Hence, we have

$$\alpha = \mathbb{P}[X > 0] = 1 - (1-p)^{\mu n}, \tag{7}$$

$$\overline{\alpha} = \mathbb{P}[X = 0] = 1 - \alpha = (1-p)^{\mu n}, \tag{8}$$

$$\alpha_1 = \mathbb{P}[X = 1] = p\mu n \times (1-p)^{\mu n - 1}. \tag{9}$$

Let β be the expected number of blocks mined in each round by the adversary controlling νn miners. Then it holds that

$$\beta = p\nu n. \tag{10}$$

IV. Our Results for the Consistency Property of Nakamoto's Blockchain Protocol

Our results for the consistency property of Nakamoto's blockchain protocol are presented as Theorems 1 and 2 below.

From [3], [6], blockchain consistency is defined as follows.

Definition 1 (Blockchain consistency)**.** *Nakamoto's blockchain protocol satisfies consistency if for any positive integer T, with at least $1 - O(1) \cdot \exp(-\Omega(T))$ probability, for any two rounds r and s with $r < s$, all but the last T blocks in the chain of any honest player i at round r is a prefix of the chain of any honest player j at round s.*

The asymptotic notation in this paper such as $O(\cdot)$ and $\Omega(\cdot)$ is standard[1]; see Footnote 1. The term[2] $O(1) \cdot \exp(-\Omega(T))$ above decays at least exponentially with respect to T. Intuitively, the above consistency notion implies that there is at least $1 - O(1) \cdot \exp(-\Omega(T))$ probability for the event that honest players agree on the current chain, except for T "unconfirmed" blocks at the end of the chain.

Based on Definition 1, Lemma 1 below presents a sufficient condition for consistency which we will use to prove our theorems.

Lemma 1 (Blockchain consistency)**.** *Nakamoto's blockchain protocol satisfies consistency if for any positive integer T, in a window of T slots, there is at least $1 - O(1) \cdot \exp(-\Omega(T))$ probability for the event that the number of convergence*

[1] Given two positive sequences f_T and g_T indexed by T, we have

- $f_T = O(g_T)$ means that there exist positive constants c_1 and T_1 such that $f_T \leq c_1 g_T$ for all $t \geq T_1$.
- $f_T = \Omega(g_T)$ means that there exist positive constants c_2 and T_2 such that $f_T \geq c_2 g_T$ for all $t \geq T_2$.

[2] Actually $1 - O(1) \cdot \exp(-\Omega(T))$ can be simplified as $1 - \exp(-\Omega(T))$ since $O(1) \cdot \exp(-\Omega(T)) = \exp(\ln O(1) - \Omega(T))$ and $\ln O(1) - \Omega(T)$ can also be written $-\Omega(T)$.

opportunities is greater than the number of blocks mined by the adversary, where a ***convergence opportunity*** *is an event which results in all honest players to agree on a single longest chain.*

Our main contribution on the consistency of Nakamoto's blockchain protocol is given as Theorem 1 below.

Theorem 1. *Nakamoto's blockchain protocol satisfies consistency when there exist constants ϵ_1 and ϵ_2 satisfying $0 < \epsilon_1 < 1$ and $\epsilon_2 > 0$ such that c denoting $\frac{1}{pn\Delta}$ satisfies*

$$c \geq \max\left\{\left(\frac{2\mu}{\ln\frac{\mu}{\nu}} + \frac{1}{\Delta}\right)\frac{1+\epsilon_2}{1-\epsilon_1}, \frac{(\ln\frac{\mu}{\nu}+1)\mu}{\epsilon_1\Delta\ln\frac{\mu}{\nu}}\right\}. \quad (11)$$

To better understand Inequality (11), we present the following result, which we will use in Remark 1 to show that Inequality (11) specifies c to be just slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$.

If there exist positive constants δ_1 and δ_2 satisfying $\delta_1 + \delta_2 < 1$ such that

$$\frac{1}{1+\exp(\Delta^{\delta_1})} \leq \nu \leq \frac{1}{1+\exp\left(\frac{1}{\Delta^{\delta_2}-1}\right)}, \quad (12)$$

we can write Inequality (11) as

$$c \geq \frac{2\mu}{\ln(\mu/\nu)} \cdot (1+\epsilon_2) \cdot \frac{1+\Delta^{\delta_1-1}}{1-\Delta^{\delta_1+\delta_2-1}}. \quad (13)$$

In Remark 1, we will explain that under Inequality (12), the condition on c as Inequality (13) enforces

$$c \text{ to be just slightly greater than } \frac{2\mu}{\ln(\mu/\nu)}.$$

Remark 1. *We now explain that Inequality (13) enforces c to be just slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$ for ν satisfying Inequality (12), which will be shown to cover almost all $\nu \in (0, \frac{1}{2})$. Here we consider $\Delta = 10^{13}$ which is used in Figure 1 of Pass* et al. *[3], a seminal work on the consistency property of Nakamoto's blockchain protocol, but our discussions readily apply to other values of Δ. We consider two sets of δ_1 and δ_2 values which cover slightly different ranges of ν.*

- *For $\Delta = 10^{13}$ of [3], we let $\delta_1 = \frac{1}{6}$ and $\delta_2 = \frac{1}{2}$ so that Inequalities (12) and (13) become*

  $$10^{-63} \leq \nu \leq 0.5 - 10^{-7}, \quad (14)$$

  and

  $$c \geq \frac{2\mu}{\ln(\mu/\nu)} \cdot (1+\epsilon_2) \cdot \left(1 + 5\times 10^{-5}\right). \quad (15)$$

  Inequalities (14) and (15) mean that c just needs to be slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$ for $10^{-63} \leq \nu \leq 0.5 - 10^{-7}$, since the positive constant ϵ_2 in Inequality (15) can be arbitrarily small.
- *Inequality (14) in the above case considers $10^{-63} \leq \nu \leq 0.5 - 10^{-7}$. Below we increase the upper bound for ν from $0.5-10^{-7}$ in Inequality (14) to $0.5-10^{-9}$ in Inequality (16) by increasing δ_2 from $\frac{1}{2}$ above to $\frac{2}{3}$ here. After increasing δ_2, to ensure that the term $\frac{1+\Delta^{\delta_1-1}}{1-\Delta^{\delta_1+\delta_2-1}}$ in Inequality (13) is still just slightly greater than 1, we slightly decrease δ_1 from $\frac{1}{6}$ above to $\frac{1}{8}$ here, which increases the lower bound for ν from 10^{-63} in Inequality (14) to 10^{-18} in Inequality (16). Specifically, for $\Delta = 10^{13}$ of [3], we let $\delta_1 = \frac{1}{8}$ and $\delta_2 = \frac{2}{3}$ so that Inequalities (12) and (13) become*

  $$10^{-18} \leq \nu \leq 0.5 - 10^{-9}, \quad (16)$$

  and

  $$c \geq \frac{2\mu}{\ln(\mu/\nu)} \cdot (1+\epsilon_2) \cdot \left(1 + 2\times 10^{-3}\right). \quad (17)$$

  Inequalities (16) and (17) mean that c just needs to be slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$ for $10^{-18} \leq \nu \leq 0.5 - 10^{-9}$, since the positive constant ϵ_2 in Inequality (15) can be arbitrarily small.

The proof of Theorem 1 will be explained in Section V. Below, we discuss the novelty of Theorem 1.

Novelty of our Theorem 1. *The analysis and results of our Theorem 1 are both novel. Moreover, with Inequality (14) considering $10^{-63} \leq \nu \leq 0.5 - 10^{-7}$ and Inequality (16) considering $10^{-18} \leq \nu \leq 0.5 - 10^{-9}$, we summarize Inequalities (14)–(17) to know that*

to ensure the consistency property of Nakamoto's blockchain protocol, c denoting $\frac{1}{pn\Delta}$ just needs to be slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$ for most $\nu \in (0, \frac{1}{2})$.

Our paper is the first one in the literature to derive such a neat expression $\frac{2\mu}{\ln(\mu/\nu)}$.

Our secondary contribution is the following Theorem 2, which fixes an issue of [6] (details later). We also use Theorem 2 to prove Theorem 1 above.

Theorem 2. *Nakamoto's blockchain protocol satisfies consistency if there exists a positive constant δ_1 such that*

$$\overline{\alpha}^{2\Delta}\alpha_1 \geq (1+\delta_1)\beta, \text{ for } \beta := p\nu n, \quad (18)$$

where $\overline{\alpha}$ (resp., α_1) denotes the probability that no (resp., only one) honest miner succeeds in solving a puzzle in one round, and is given by Eq. (8) (resp., Eq. (9)), while β denotes the expected number of blocks mined in each round by the adversary controlling νn miners.

The proof of Theorem 2 will be explained in Section VI. Below, we discuss the novelty of Theorem 2.

Novelty of our Theorem 2. Our Theorem 2 is also novel in the sense its result as Inequality (18) has not been presented in any related work. Although a recent study by Kiffer *et al.* [6] also adopts a Markov-chain based approach that our Theorem 2 uses, our Theorem 2 differentiates from [6] in the following aspects as we will discuss:

❶ First, [6] does not use the following two Markov chains which we propose for the first time and use to prove our Theorem 2:
  ① a Markov chain which models the transition of a variable denoting the suffix of the concatenation of the previous states and the current state,
  ② a Markov chain modeling the transition of a variable which denotes the concatenation of i) the suffix of previous states before the Δ to last state, ii) the previous Δ states, and iii) the current state.

❷ Second, the analysis of [6] has minor errors. In [6], the computations of ℓ_{11} and ℓ_{10} (defined on Page 7 of [6]) are incorrect. Specifically, $\frac{1}{\mu p}$ therein should be $\frac{1}{\alpha}$ (i.e., $\frac{1}{1-(1-p)^{\mu n}}$).

❸ Third, even after we correct $\mu \cdot p$ to α in many places of [6] and perform some computations to obtain Claim 1 below from Theorem 4.4 on Page 8 of [6], the result of [6] (and hence Claim 1) still has a minor issue. In [6],

to compute the convergence opportunities, one subevent is that *at least one* honest miner succeeds in solving a puzzle in one round (which happens with probability α in Eq. (7)), while the correct subevent should be that *only one* honest miner succeeds in solving a puzzle in one round (which happens with probability α_1 in Eq. (9)). Although [6] mentions "a single honest mined block", but its calculation actually uses "at least one honest mined block" (this leads to no appearance of α_1 in [6]'s consistency condition). Our Theorem 2 fixes the above issue of [6] (as noted in "❶" above, we also introduce novel Markov chains to present a clearer proof).

Here we give the reason why we present a detailed proof for Theorem 2 instead of just replacing α with α_1 in Inequality (19), a condition based on the analysis of [6]. We find the proof [6] not intuitive to understand. For instance, Page 6 of [6] uses the Markov chain $\circlearrowleft S_0 \rightleftarrows S_1 \circlearrowright$ to analyze consistency, with S_0 denoting the "messy" state where honest mined blocks occur in less than Δ rounds from one another, and S_1 denoting the state where quiet periods between honest mined blocks is at least Δ rounds. Taking the transition $S_1 \to S_1$ as an example, it happens after a honest mined block followed by a quiet period of at least Δ rounds. As the occurrence of the transition $S_1 \to S_1$ needs multiple rounds, the Markov chain $\circlearrowleft S_0 \rightleftarrows S_1 \circlearrowright$ of [6] cannot characterize the states in the middle of the transition. Due to this, we present a proof of Theorem 2 from scratch using more detailed Markov chains. We emphasize again that the detailed proof of Theorem 2 involving the novel Markov chains is our secondary contribution while Theorem 1 presenting a neat condition to ensure consistency is our major contribution. After obtaining an inequality (Inequality (78) in [7]) to ensure consistency, [6] does not analyze the inequality to provide a more understandable bound for c as our Theorem 1 does. Our proof of moving from Theorem 2 to Theorem 1 in Section V is quite involved.

We now state Claim 1 on Page 5, which is Theorem 4.4 on Page 8 of [6] after we correct $\mu \cdot p$ to α in many places of [6] and perform some computations presented in Appendix A of the online full version [7].

Claim 1 (Theorem 4.4 on Page 8 of [6] after we correct $\mu \cdot p$ to α in many places of [6] and perform some computations)**.** *Nakamoto's blockchain protocol satisfies consistency if there exists a positive constant δ_3 such that*

$$\overline{\alpha}^{2\Delta}\alpha \geq (1+\delta_3)\beta. \tag{19}$$

From the expressions of α and $\overline{\alpha}$ in Eq. (7) and Eq. (8) as well as $\beta = p\nu n$ and $c = \frac{1}{pn\Delta}$, Inequality (19) means the following complex condition involving c:

$$\left(1-\tfrac{1}{cn\Delta}\right)^{2\mu n\Delta}\left(1-\left(1-\tfrac{1}{cn\Delta}\right)^{\mu n}\right) \geq (1+\delta_3)\tfrac{\nu}{c\Delta}. \tag{20}$$

How we rewrite Theorem 4.4 as Claim 1 is presented in Appendix A of the online full version [7]. We present the result as the claim due to the issue mentioned in "❸" above.

V. Proof of Theorem 1 Given Theorem 2

We decompose Inequality (11) of Theorem 1 into Inequalities (21) and (22), to present Theorem 3 below.

Theorem 3. *Consistency of Nakamoto's blockchain protocol holds in a window of T rounds with probability at least $1 - O(1)\cdot\exp(-\Omega(T))$, when there exist constants ϵ_1 and ϵ_2 satisfying $0 < \epsilon_1 < 1$ and $\epsilon_2 > 0$ such that we have*

$$pn \leq \frac{\epsilon_1 \ln\frac{\mu}{\nu}}{(\ln\frac{\mu}{\nu}+1)\mu}, \tag{21}$$

and c denoting $\frac{1}{pn\Delta}$ satisfies

$$c \geq \left[\frac{2\mu}{\ln(\mu/\nu)} + \frac{1}{\Delta}\right]\frac{1+\epsilon_2}{1-\epsilon_1}. \tag{22}$$

Since c denotes $\frac{1}{pn\Delta}$, it is straightforward to show that a combination of Inequalities (21) and (22) is the same as Inequality (11), which is a condition of Theorem 1.

Below we present the proof of Theorem 3 using Theorem 2. In Appendix E of the online full version [7], we use Theorem 3 to show Theorem 1.

A. Proof of Theorem 3 using Theorem 2

To prove Theorem 3 based on Theorem 2, we will show that given Inequality (21), Inequality (22) implies Inequality (18). To this end, we analyze Inequality (18) through a series of transformations. Before stating the transformations, we note that in the rest of the paper, "$\Longleftarrow$", "$\Longrightarrow$", and "$\Longleftrightarrow$" represent "is implied by", "implies", and "is equivalent to", respectively. To prove Theorem 3, we will convert Inequality (18) in a number of steps and obtain the following results, where we will explain soon how to set δ_1 and δ_5.

Nakamoto's blockchain protocol satisfies consistency

$$\xLeftarrow{\text{Theorem 2}} \left\{\overline{\alpha}^{2\Delta}\alpha_1 \geq (1+\delta_1)p\nu n\right\} \tag{23}$$

$$\xLeftarrow{\text{Lemma 2}} \left\{\overline{\alpha} \geq \left(\frac{1+\delta_1}{1-p\mu n}\cdot\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right\} \tag{24}$$

$$\xLeftarrow{\text{Lemma 3}} \left\{\overline{\alpha} \geq \left(1+\frac{\delta_5}{2\Delta}\right)\cdot\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right\} \tag{25}$$

$$\xLeftarrow{\text{Lemma 4}} \left\{c \geq \frac{1}{n\Delta\left\{1-\left[\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]^{1/(\mu n)}\right\}}\right\} \tag{26}$$

$$\xLeftarrow{\text{Lemma 5}} \left\{c \geq \frac{\mu}{\Delta\left[1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]}\right\} \tag{27}$$

$$\xLeftarrow{\text{Lemma 6}} \left\{c \geq \frac{\mu}{\Delta\left[1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]}\cdot\left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right)\right\} \tag{28}$$

$$\xLeftarrow{\text{Lemma 7}} \left\{c \geq \left[\frac{2\mu}{\ln(\mu/\nu)}+\frac{\mu}{\Delta}\right]\cdot\left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right)\right\} \tag{29}$$

$$\xLeftarrow{\text{Lemma 8}} \left\{c \geq \left[\frac{2\mu}{\ln(\mu/\nu)}+\frac{1}{\Delta}\right]\cdot\frac{1+\epsilon_2}{1-\epsilon_1}\right\}$$

(i.e., Inequality (22) of Theorem 3)). (30)

The statements of Lemmas 2–8 used above are deferred to the end of this subsection for clarity, while their proofs will be presented in the Appendicies of the online full version [7].

Lemmas 2–8 also involve extra conditions on pn, δ_1, and δ_5, which are not explicitly stated in (23)–(30). We will

show on Page 8 that these conditions on pn are implied by Inequality (21) of Theorem 3. To satisfy conditions on δ_1 and δ_5 in Lemmas 2–8 for proving Theorem 3 (the conditions will be discussed in detail on Page 8), we will set δ_5 and δ_1 as follows:

$$\delta_5 = \frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{\epsilon_1+\epsilon_2+(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)}, \text{ and} \tag{31}$$

$$\delta_1 = (1+\delta_5)\cdot\left(1-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{\ln\frac{\mu}{\nu}+1}\right)-1 \text{ with the above } \delta_5. \tag{32}$$

We note that δ_5 and δ_1 in Eq. (31) and Eq. (32) are both positive for $0<\epsilon_1<1$ and $\epsilon_2>0$. The details are given in Appendix F of the online full version [7].

Below, we give *intuitive* explanations for a) how we obtain the condition on pn in Inequality (21) of Theorem 3, and b) why we set δ_5 and δ_1 according to Eq. (31) and (32) in order to have (23)–(30) get through. The explanations are just intuitive since some steps come from necessity arguments while some other steps result from sufficiency arguments. On Page 8, we will formally explain that enforcing the condition on pn in Inequality (21) and setting constants δ_5 and δ_1 according to Eq. (31) and (32) will ensure that all conditions of Lemmas 2–8 are satisfied.

How do we obtain the condition on pn in Inequality (21) of Theorem 3?

In (28) and (29), we observe the expression $\ln\frac{\mu}{\nu}-\delta_5$, which requires δ_5 to be smaller than $\ln\frac{\mu}{\nu}$, as will become clear in Lemmas 6 and 7. From (25), we see that Lemma 3 is used to provide $\left(\frac{1+\delta_1}{1-p\mu n}\right)^{1/(2\Delta)} \le 1+\frac{\delta_5}{2\Delta}$. A necessary condition for this is $\left(\frac{1}{1-p\mu n}\right)^{1/(2\Delta)} < 1+\frac{\delta_5}{2\Delta}$, for which a sufficient condition is $\frac{1}{1-p\mu n} < 1+\delta_5$ since we know from the binomial series that $1+\delta_5 < \left(1+\frac{\delta_5}{2\Delta}\right)^{2\Delta}$. For $\delta_5 < \ln\frac{\mu}{\nu}$, this implies $\frac{1}{1-p\mu n} < 1+\ln\frac{\mu}{\nu}$, for which a sufficient condition is

$$pn \le \frac{\epsilon_1\ln\frac{\mu}{\nu}}{(\ln\frac{\mu}{\nu}+1)\mu} \text{ for a constant } 0<\epsilon_1<1. \tag{33}$$

This Inequality (33) is stronger than $pn<\frac{1}{\mu}$ used in Lemma 2. Hence, our condition on pn is just Inequality (33), which is exactly Inequality (21) of Theorem 3.

How do we set constants δ_5 and δ_1 according to Eq. (31) and (32) to have (23)–(30) get through?

As discussed above, from (25), we see that Lemma 3 is used to provide $\left(\frac{1+\delta_1}{1-p\mu n}\right)^{1/(2\Delta)} \le 1+\frac{\delta_5}{2\Delta}$, for which a sufficient condition is $\frac{1+\delta_1}{1-p\mu n} \le 1+\delta_5$ since we know from the binomial series that $1+\delta_5 < \left(1+\frac{\delta_5}{2\Delta}\right)^{2\Delta}$. We have just explained above the reasoning behind enforcing Inequality (21) of Theorem 3; i.e., $pn \le \frac{\epsilon_1\ln\frac{\mu}{\nu}}{(\ln\frac{\mu}{\nu}+1)\mu}$ for a positive constant $\epsilon_1<1$. Then a sufficient condition to ensure the existence of positive δ_1 satisfying $\frac{1+\delta_1}{1-p\mu n} \le 1+\delta_5$ discussed just above is $\frac{1}{1-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+\ln\frac{\mu}{\nu}}} < 1+\delta_5$, which gives $\delta_5 > \frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln\frac{\mu}{\nu}}$. For such δ_5, the expression $\left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right)$ appearing in (29) is greater than $\left[1+\frac{\frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln\frac{\mu}{\nu}}}{\ln\frac{\mu}{\nu}-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln\frac{\mu}{\nu}}}\right] = \left[1+\frac{\epsilon_1}{(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)}\right]$. Then we can select δ_5 such that $\left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right)$ equals $\left[1+\frac{\epsilon_1+\epsilon_2}{(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)}\right]$ for a positive constant ϵ_2. This gives δ_5 by Eq. (31).

Recalling $\frac{1+\delta_1}{1-p\mu n} \le 1+\delta_5$ discussed above to produce Lemma 3, we have $1+\delta_1 \le (1+\delta_5)\cdot(1-p\mu n)$, for which we know from Inequality (21) of Theorem 3 (i.e., $pn \le \frac{\epsilon_1\ln\frac{\mu}{\nu}}{(\ln\frac{\mu}{\nu}+1)\mu}$ for a positive constant $\epsilon_1<1$) that a sufficient condition is $1+\delta_1 \le (1+\delta_5)\cdot\left(1-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{\ln\frac{\mu}{\nu}+1}\right)$. Taking "$\le$" here as "$=$" for simplicity, we set δ_1 by Eq. (32).

We now state Lemmas 2–8, which are proved in the Appendicies of the online full version [7].

Lemma 2. *Under*

$$0 < p\mu n < 1, \tag{34}$$

if

$$\overline{\alpha} \ge \left(\frac{1+\delta_1}{1-p\mu n}\cdot\frac{\nu}{\mu}\right)^{1/(2\Delta)}, \tag{35}$$

then Inequality (18) of Theorem 2 follows; i.e., $\overline{\alpha}^{2\Delta}\alpha_1 \ge (1+\delta_1)p\nu n$.

Remark 2. *The above result shows (24) under (34), where (24) is*

$$\left\{\overline{\alpha}^{2\Delta}\alpha_1 \ge (1+\delta_1)p\nu n\right\} \xLeftarrow{\text{Lemma 2}} \text{Inequality (35)}. \tag{36}$$

Lemma 3. *If Inequality (21) of Theorem 3 holds; i.e., if there exists a positive constant $0<\epsilon_1<1$ such that Inequality (21) of Theorem 3 holds, then for*

$$\delta_5 > \frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln\frac{\mu}{\nu}}, \tag{37}$$

and δ_1 given by

$$\delta_1 = (1+\delta_5)\cdot\left(1-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{\ln\frac{\mu}{\nu}+1}\right)-1, \tag{38}$$

we have $\delta_5>0$, $\delta_1>0$, and

$$\left(\frac{1+\delta_1}{1-p\mu n}\right)^{1/(2\Delta)} \le 1+\frac{\delta_5}{2\Delta}. \tag{39}$$

Remark 3. *Inequality (39) means that under*

$$\overline{\alpha} \ge \left(1+\frac{\delta_5}{2\Delta}\right)\cdot\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}, \tag{40}$$

Inequality (35) of Lemma 2 follows. Thus, under (21) and (37), we have

$$\text{Inequality (35)} \xLeftarrow{\text{Lemma 3}} \text{Inequality (40)}. \tag{41}$$

Lemma 4. *Under*

$$0 < \delta_5 < \ln\frac{\mu}{\nu}, \tag{42}$$

if c denoting $\frac{1}{pn\Delta}$ satisfies

$$c \ge \frac{1}{n\Delta\left\{1-\left[\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]^{1/(\mu n)}\right\}}, \tag{43}$$

then we have Inequality (40). Note that under Inequality (42), the denominator in Inequality (43) is positive from Proposition 1 to be presented soon.

Remark 4. *From the above result, under Inequality (42), we have*

$$\text{Inequality (40)} \xLeftarrow{\text{Lemma 4}} \text{Inequality (43)}. \tag{44}$$

Remark 5. *For proving Theorem 3, we set δ_5 and δ_1 according to Eq. (31) and Eq. (32) with constants $0<\epsilon_1<1$ and $\epsilon_2>0$, so that (37) (38) and (42) of Lemmas 3 and 4 are satisfied, as explained below. First, the result that Eq. (31) implies (37) has been shown in (100). Second, Eq. (32) is the same as (38). Finally, for δ_5 in Eq. (31) with $0<\epsilon_1<1$, we have $\delta_5 = \frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{\epsilon_1+\epsilon_2+(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)} < \frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{\epsilon_1+\epsilon_2} = \ln\frac{\mu}{\nu}$, which gives (42).*

Proposition 1. *Under Inequality (42), we have*

$$1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)} > 0.$$

Lemma 5. *Under Inequality (42), we have*

$$\frac{\mu}{\Delta\left[1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]} \geq \frac{1}{n\Delta\left\{1-\left[\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]^{1/(\mu n)}\right\}}, \tag{45}$$

where the denominators in both sides of Inequality (45) are positive from Proposition 1 above.

Remark 6. *Inequality (45) means that if c denoting $\frac{1}{pn\Delta}$ satisfies*

$$c \geq \frac{\mu}{\Delta\left[1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]}, \tag{46}$$

then Inequality (43) of Lemma 4 follows. Thus, under Inequality (42), we have

$$\text{Inequality (43)} \xLeftarrow{\text{Lemma 5}} \text{Inequality (46)}. \tag{47}$$

Lemma 6. *Under Inequality (42), we have*

$$\frac{1}{1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}} \cdot \left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right) > \frac{1}{1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}}. \tag{48}$$

Remark 7. *Inequality (48) means that if c denoting $\frac{1}{pn\Delta}$ satisfies*

$$c \geq \frac{\mu}{\Delta\left[1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]} \cdot \left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right), \tag{49}$$

then Inequality (46) follows. Thus, under Inequality (42), we have

$$\text{Inequality (46)} \xLeftarrow{\text{Lemma 6}} \text{Inequality (49)}. \tag{50}$$

Lemma 7. *We have*

$$\frac{2}{\ln(\mu/\nu)} \leq \frac{1}{\Delta\left[1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]} \leq \frac{2}{\ln(\mu/\nu)} + \frac{1}{\Delta}. \tag{51}$$

Remark 8. *Inequality (51) means that if c denoting $\frac{1}{pn\Delta}$ satisfies*

$$c \geq \left[\frac{2\mu}{\ln(\mu/\nu)} + \frac{\mu}{\Delta}\right] \cdot \left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right), \tag{52}$$

where δ_5 satisfies $0 < \delta_5 < \ln\frac{\mu}{\nu}$ (i.e., Inequality (42)), then Inequality (49) follows. Thus, under Inequality (42), we have

$$\text{Inequality (49)} \xLeftarrow{\text{Lemma 7}} \text{Inequality (52)}. \tag{53}$$

Lemma 8. *For constants $0 < \epsilon_1 < 1$ and $\epsilon_2 > 0$, with δ_5 given by Eq. (31), we have*

$$1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5} < \frac{1+\epsilon_2}{1-\epsilon_1}. \tag{54}$$

Remark 9. *Inequality (54) means that under Inequality (31), if c denoting $\frac{1}{pn\Delta}$ satisfies Inequality (22) of Theorem 3 (i.e., $c \geq \left[\frac{2\mu}{\ln(\mu/\nu)} + \frac{\mu}{\Delta}\right] \cdot \frac{1+\epsilon_2}{1-\epsilon_1}$), then Inequality (52) follows. Thus, under Inequality (31), we have*

$$\text{Inequality (52)} \xLeftarrow{\text{Lemma 8}} \text{Inequality (22)}. \tag{55}$$

Putting All Things Together to Prove Theorem 3. The above results (36) (41) (44) (47) (50) (53) (55) are exactly (24) (25) (26) (27) (28) (29) (30) discussed earlier, which along with (23) implies the desired result of Theorem 3 that consistency of Nakamoto's blockchain protocol follows if Inequality (22) holds, under the assumption that we enforce all conditions of (36) (41) (44) (47) (50) (53) (55). Now we discuss these conditions:

- (36) needs the condition (34) that Lemma 2 requires,
- (41) needs the condition (21) and (37) that Lemma 3 requires, and
- (44) (resp. (47) (50) and (53)) needs the condition (42) that Lemma 4 (resp. Lemmas 5, 6, and 7) requires,
- (55) needs the condition (31) that Lemma 8 requires.

Hence, to complete proving Theorem 3, we just need to enforce (34) (37) (42) and (31) given Inequality (21) (i.e., $pn \leq \frac{\epsilon_1 \ln\frac{\mu}{\nu}}{(\ln\frac{\mu}{\nu}+1)\mu}$) with $0 < \epsilon_1 < 1$ and $\epsilon_2 > 0$ from Theorem 3. To this end, we have the following:

- We obtain Inequality (34) from Inequality (21) with $0 < \epsilon_1 < 1$, in view of $pn \leq \frac{\epsilon_1 \ln\frac{\mu}{\nu}}{(\ln\frac{\mu}{\nu}+1)\mu} < \frac{1}{\mu}$.
- After we define δ_5 according to (31), we obtain Inequality (37) in view of (100), and obtain Inequality (42) in view of $\delta_5 = \frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{(\epsilon_1+\epsilon_2)+(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)} < \frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{(\epsilon_1+\epsilon_2)} = \ln\frac{\mu}{\nu}$.

Summarizing the above, we have shown Theorem 3 using (36) (41) (44) (47) (50) (53) (55), which hold respectively after we prove Lemmas 2–8 in Appendices H–O of the online full version [7]. In Appendix E of [7], we use Theorem 3 to show Theorem 1. ■

VI. Proof of Theorem 2

We use $A(t_0, t_0+T-1)$ to denote the number of blocks mined the adversary in the T rounds from round t_0 to t_0+T-1, and use $C(t_0, t_0+T-1)$ to denote the number of times that $HN^{\geq\Delta}||H_1N^{\Delta}$ is visited (i.e., the number of convergence opportunities) in the T rounds from round t_0 to t_0+T-1. Then we will show in Section VI-A that Inequality (18) of Theorem 2 is the same as

$$\mathbb{E}\left[C(t_0, t_0+T-1)\right] \geq (1+\delta_1)\cdot\mathbb{E}\left[A(t_0, t_0+T-1)\right]. \tag{56}$$

Here we discuss the intuition of requiring Inequality (56), which then gives Inequality (18). First, we will prove that the probability of $C(t_0, t_0+T-1)$ being a constant factor smaller than its expectation $\mathbb{E}\left[C(t_0, t_0+T-1)\right]$ is exponentially small in T. Formally, for any positive constant $\delta_2 < 1$, we have

$$\begin{aligned}&\mathbb{P}\left[C(t_0, t_0+T-1) \leq (1-\delta_2)\cdot\mathbb{E}\left[C(t_0, t_0+T-1)\right]\right]\\ &\leq O(1)\cdot\exp\left(-\Omega\left(T\right)\right).\end{aligned} \tag{57}$$

Second, we will prove that the probability of $A(t_0, t_0+T-1)$ being a constant factor greater than its expectation $\mathbb{E}\left[A(t_0, t_0+T-1)\right]$ is exponentially small in T. Formally, for any positive constant δ_4, we have

$$\begin{aligned}&\mathbb{P}\left[A(t_0, t_0+T-1) \geq (1+\delta_4)\cdot\mathbb{E}\left[A(t_0, t_0+T-1)\right]\right]\\ &\leq O(1)\cdot\exp\left(-\Omega\left(T\right)\right).\end{aligned} \tag{58}$$

In (57) and (58), the term $O(1)$ is with respect to T.

Via a union bound to combine Inequalities (57) and (58), $(C(t_0, t_0+T-1) \leq (1-\delta_2)\cdot\mathbb{E}\left[C(t_0, t_0+T-1)\right]) \vee (A(t_0, t_0+T-1) \geq (1+\delta_4)\cdot\mathbb{E}\left[A(t_0, t_0+T-1)\right])$ happens with probability no greater than the result of summing the bounds in the right hand side (RHS) of Inequalities (57) and (58), which can also be written as $O(1)\cdot\exp\left(-\Omega\left(T\right)\right)$. Then we have (at least) $1-O(1)\cdot\exp\left(-\Omega\left(T\right)\right)$ probability for the above union event's complement, $(C(t_0, t_0+T-1) > (1-\delta_2)\cdot\mathbb{E}\left[C(t_0, t_0+T-1)\right]) \wedge (A(t_0, t_0+T-1) < (1+\delta_4)\cdot\mathbb{E}\left[A(t_0, t_0+T-1)\right])$, implying that $C(t_0, t_0+T-1) - A(t_0, t_0+T-1)$ is greater than

$$(1-\delta_2)\cdot\mathbb{E}\left[C(t_0, t_0+T-1)\right] - (1+\delta_4)\cdot\mathbb{E}\left[A(t_0, t_0+T-1)\right]. \tag{59}$$

From Inequality (56), we bound the term in (59) by

$$(59) \geq \left[(1-\delta_2)\cdot(1+\delta_1) - (1+\delta_4)\right]\cdot\mathbb{E}\left[A(t_0, t_0+T-1)\right]. \tag{60}$$

Then to obtain the desired result that $C(t_0, t_0+T-1) - A(t_0, t_0+T-1)$ is $\Omega(T)$ with probability $1-O(1)\cdot\exp\left(-\Omega\left(T\right)\right)$, we select positive constants $\delta_2 < 1$ and δ_4 such that the term in (60) is $\Omega(T)$. It will become clear from Eq. (65) that $A(t_0, t_0+T-1)$ can be

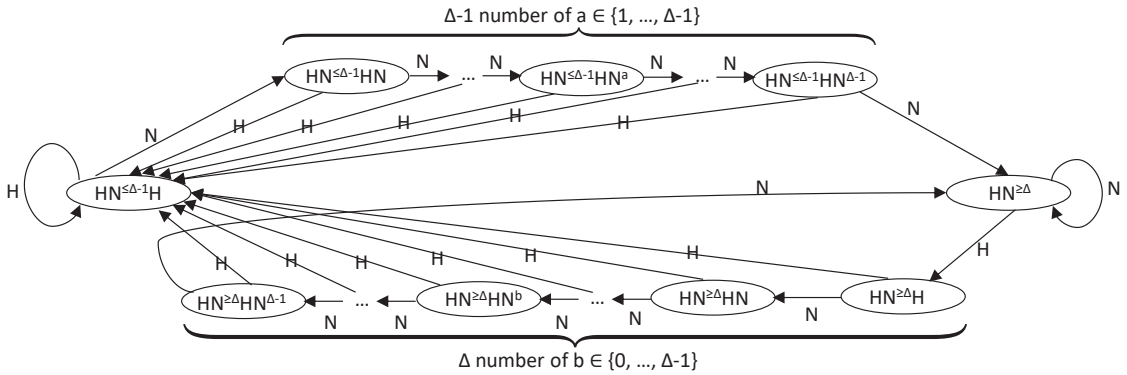

Fig. 2: The suffix-of-previous-and-current-states Markov chain $\mathcal{C}_{\boldsymbol{F}}$, which models the transition of a variable denoting the suffix of the concatenation of the previous states and the current state.

written as $\Omega(T)$, so we select positive constants $\delta_2 < 1$ and δ_4 such that $[(1-\delta_2)\cdot(1+\delta_1)-(1+\delta_4)]$ appearing in (60) is a positive constant. To this end, we set

$$\delta_2 := 1-(1+\delta_1)^{-1/3}, \quad \delta_4 := (1+\delta_1)^{1/3}-1, \tag{61}$$

so that Inequality (60) becomes

$$(59) \geq \left[(1+\delta_1)^{2/3}-(1+\delta_1)^{1/3}\right]\cdot\mathbb{E}\left[A(t_0,t_0+T-1)\right]. \tag{62}$$

Summarizing the above, we have

If Inequalities (56) (57) and (58) hold, then

$C(t_0,t_0+T-1)-A(t_0,t_0+T-1)$ is greater than

$\left[(1+\delta_1)^{2/3}-(1+\delta_1)^{1/3}\right]\cdot\mathbb{E}\left[A(t_0,t_0+T-1)\right]$

with probability $1-O(1)\cdot\exp\left(-\Omega\left(T\right)\right)$. (63)

In the rest of this section, we will first prove in Section VI-A that Inequality (18) of Theorem 2 is the same as Inequality (56). It will also become clear in Section VI-A that $A(t_0,t_0+T-1)$ can be written as $\Omega(T)$. We prove Inequalities (57) and (58) in Appendices C and D of the online full version [7]. In Section VI-B, we combine the results of Appendices C and D of [7] with (63) to complete the proof of to Theorem 2.

A. Proving that Inequality (18) is the same as Inequality (56)

Inequality (18) of Theorem 2 is $\overline{\alpha}^{2\Delta}\alpha_1 \geq (1+\delta_1)p\nu n$. To show Inequality (56), we will explain

$$\mathbb{E}\left[C(t_0,t_0+T-1)\right] = T\overline{\alpha}^{2\Delta}\alpha_1, \tag{64}$$

and

$$\mathbb{E}\left[A(t_0,t_0+T-1)\right] = Tp\nu n, \tag{65}$$

We first show Eq. (65). Since the adversary controls νn nodes and each node mines a block independently with probability p in each round, the number of blocks mined (by νn nodes controlled) by the adversary in each round follows binom$(\nu n, p)$, which denotes a binomial distribution with νn being the number of trials and p being the success probability for each trial. Then $A(t_0,t_0+T-1)$ denoting the number of blocks mined the adversary in the T rounds from round t_0 to t_0+T-1 is the sum of T indepdendent random variables, each of which obeys binom$(\nu n,p)$. Hence, $A(t_0,t_0+T-1)$ follows binom$(T\nu n,p)$. Then Eq. (65) clearly follows.

We now present the proof of Eq. (64.

In each round, one of the following events will happen: i) H, which means that at least one block is mined by the benign (i.e., honest) nodes, and ii) N, which means that no block is mined by the benign nodes. By a round's state, we refer to whether H or N happens, and we know from the definitions of α and $\overline{\alpha}$ in Eq. (7) and Eq. (8) that H (resp., N) happens with probability α (resp., $\overline{\alpha}$). Then we define State-Set to characterize the possible values that a round's state can take:

$$\text{State-Set} := \{H,\ N\}. \tag{66}$$

Let S_t be the random variable representing the state at round t. We will use $s_t \in$ State-Set as an instantiation of S_t.

We consider a Markov chain $\mathcal{C}_{\boldsymbol{F}}$ for the suffix of all the states in all rounds up to round t, where "$\boldsymbol{F}$" means suffix. We will explain that Figure 2 can represent this Markov chain. To avoid confusion, we use "$\boldsymbol{F}$" instead of "$\boldsymbol{S}$" since the symbol S is used to represent the state at a round. We will call $\mathcal{C}_{\boldsymbol{F}}$ as the suffix-of-previous-and-current-states Markov chain. At round t, let random variable $\boldsymbol{F}_t$ represent the suffix of the states in all rounds up to round t; i.e., $\boldsymbol{F}_t$ represents the vertex visited at round t in the Markov chain $\mathcal{C}_{\boldsymbol{F}}$.

After at least two H have happened by round t (which holds for sufficiently large t), we will explain below that we can characterize all possible $\boldsymbol{F}_t$ by the following $2\Delta+1$ values which form the Suffix-Set:

$$\text{Suffix-Set} := \left\{ \begin{matrix} HN^{\leq\Delta-1}H,\ HN^{\leq\Delta-1}HN^{a}, \\ HN^{\geq\Delta},\ HN^{\geq\Delta}HN^{b} \end{matrix} \middle|\ \begin{matrix} a\in\{1,\ldots,\Delta-1\}, \\ b\in\{0,\ldots,\Delta-1\} \end{matrix} \right\}. \tag{67}$$

In (67), the term $N^{\leq\Delta-1}$ means a series of N which has at most $\Delta-1$ number of consecutive N; i.e., zero N (i.e., null), one N, …, or $\Delta-1$ number of N. Similarly, $N^{\geq\Delta}$ means a series of N which has at least Δ number of consecutive N, while N^a (resp., N^b) means a (resp., b) number of consecutive N. Supposing $\Delta = 3$ for the purpose of giving an example (practical Δ is much larger) and the states from round 1 to round 10 are H,N,H,H,N,N,H,N,N,N, then the corresponding $\boldsymbol{F}_7,\boldsymbol{F}_8,\boldsymbol{F}_9$, and $\boldsymbol{F}_{10}$ (i.e., $\boldsymbol{F}_t$ at time $t=7,8,9$, and 10) are $HN^{\leq\Delta-1}H$, $HN^{\leq\Delta-1}HN^a$ with $a=1$, $HN^{\leq\Delta-1}HN^a$ with $a=2$, and $HN^{\geq\Delta}$ ($HN^{\geq 3}$ covers HN^3), respectively.

To see why we can characterize all possible $\boldsymbol{F}_t$ by (67), we discuss the following cases, where we recall that S_i represents the state at round i:

- If S_t is H and S_{t-1} is H, then we can set $\boldsymbol{F}_t$ as $HN^{\leq\Delta-1}H$ which covers HH when "$N^{\leq\Delta-1}$" becomes 0 number of N (i.e., null);
- If S_t is H and S_{t-1} is N, as we consider that at least two H have happened by round t (which holds for sufficiently large t), suppose the previous H closest to round t happens at round $t-c$ for some $c>0$. In other words, S_{t-c} and S_t are H while S_i for each $i\in\{t-c+1,\ldots,t-1\}$ is N so that the series $S_{t-c}\ldots S_t$ can be written as $HN^{c-1}H$. Then if $c-1\leq\Delta-1$, we can set $\boldsymbol{F}_t$ as $HN^{\leq\Delta-1}H$; if $c-1\geq\Delta$, we can set $\boldsymbol{F}_t$ as $HN^{\geq\Delta}HN^b$ which covers $HN^{\geq\Delta}H$ when b takes 0.

- If S_t is N, as we consider that at least two H have happened by round t (which holds for sufficiently large t), suppose the H closest to and before round t happens at round $t-d$ for some $d>0$. In other words, S_{t-d} is H while S_i for each $i \in \{t-d+1,\ldots,t\}$ is N so that the series $S_{t-d}\ldots S_t$ can be written as HN^d. Then we have two subcases:
  - If $d \geq \Delta$, we can set $\boldsymbol{F}_t$ as $HN^{\geq\Delta}$.
  - If $d \leq \Delta-1$, given that $S_{t-d}\ldots S_t$ is HN^d, we further discuss the states before round $t-d$. Again, since we consider that at least two H have happened by round t (which holds for sufficiently large t), suppose the previous H closest to round $t-d$ happens at round $t-f$ for some $f>d$. In other words, S_{t-f} is H while S_i for each $i \in \{t-f+1,\ldots,t-d-1\}$ is N so that the series $S_{t-f}\ldots S_t$ can be written as $HN^{f-d-1}HN^d$. Recalling this subcase discusses $d \leq \Delta-1$, if $f-d-1 \leq \Delta-1$, we can set $\boldsymbol{F}_t$ as $HN^{\leq\Delta-1}HN^a$ for $a=d \in \{1,\ldots,\Delta-1\}$; if $f-d-1 \geq \Delta$, we can set $\boldsymbol{F}_t$ as $HN^{\geq\Delta}HN^b$ for $b=d \in \{1,\ldots,\Delta-1\}$.

In Figure 2 on Page 9, we plot the transition of $\boldsymbol{F}_t$ in the suffix-of-previous-and-current-states Markov chain $\mathcal{C}_{\boldsymbol{F}}$, which is *time-homogeneous*, *irreducible*, and *ergodic*. In particular, from [6], [19], time-homogeneous means that the transition does not depend on the time; irreducible means getting to any state from any other state has non-zero probability; and ergodic means that each state has a positive mean recurrence time and is aperiodic (i.e., the period is 1).

As illustrated in Figure 2, the transition rules in the Markov chain $\mathcal{C}_{\boldsymbol{F}}$ are as follows:

① First, for any $a \in \{1,\ldots,\Delta-1\}$, the event that $\boldsymbol{F}_t$ at time t is $HN^{\leq\Delta-1}HN^a$ can only result from that $\boldsymbol{F}_{t-1}$ at time $t-1$ is $HN^{\leq\Delta-1}HN^{a-1}$, which by a recursive argument can only result from that $\boldsymbol{F}_{t-a}$ at time $t-a$ is $HN^{\leq\Delta-1}H$. Moreover, moving from $\boldsymbol{F}_{t-a}=HN^{\leq\Delta-1}H$ to $\boldsymbol{F}_t=HN^{\leq\Delta-1}HN^a$ requires that S_i for each $i \in \{t-a+1,\ldots,t\}$ is N.

② Second, for any $b \in \{0,\ldots,\Delta-1\}$, the event that $\boldsymbol{F}_t$ at time t is $HN^{\geq\Delta}HN^b$ can only result from that $\boldsymbol{F}_{t-1}$ at time $t-1$ is $HN^{\geq\Delta}HN^{b-1}$, which by a recursive argument can only result from that $\boldsymbol{F}_{t-b}$ at time $t-b$ is $HN^{\geq\Delta}H$, and also $\boldsymbol{F}_{t-b-1}$ at time $t-b-1$ is $HN^{\geq\Delta}$. Moreover, moving from $\boldsymbol{F}_{t-b-1}=HN^{\geq\Delta}$ to $\boldsymbol{F}_t=HN^{\geq\Delta}HN^b$ requires that S_{t-a} is H, and S_i for each $i \in \{t-a+1,\ldots,t\}$ is N.

③ Third, the event that $\boldsymbol{F}_t$ at time t is $HN^{\leq\Delta-1}H$ can result from the combination of the following two events: i) $\boldsymbol{F}_{t-1}$ at time $t-1$ is $HN^{\leq\Delta-1}H$ or $HN^{\leq\Delta-1}HN^a$ for $a \in \{1,\ldots,\Delta-1\}$ or $HN^{\geq\Delta}HN^b$ for $b \in \{0,\ldots,\Delta-1\}$; and ii) S_t is H.

④ Fourth, the event that $\boldsymbol{F}_t$ at time t is $HN^{\geq\Delta}$ can result from the combination of the following two events: i) $\boldsymbol{F}_{t-1}$ at time $t-1$ is $HN^{\geq\Delta}$ or $HN^{\leq\Delta-1}HN^{\Delta-1}$ or $HN^{\geq\Delta}HN^{\Delta-1}$; and ii) S_t is N.

In Appendix B of the online full version [7], we derive the stationary distribution of the suffix-of-previous-and-current-states Markov chain $\mathcal{C}_{\boldsymbol{F}}$ as follows:

$$\begin{cases} \pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}H)=\alpha\cdot(1-\overline{\alpha}^{\Delta}), & \text{(68a)}\\ \pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}HN^a)=\alpha\cdot(1-\overline{\alpha}^{\Delta})\cdot\overline{\alpha}^a, \quad \forall a\in\{1,\ldots,\Delta-1\}, & \text{(68b)}\\ \pi_{\boldsymbol{F}}(HN^{\geq\Delta})=\overline{\alpha}^{\Delta}, & \text{(68c)}\\ \pi_{\boldsymbol{F}}(HN^{\geq\Delta}HN^b)=\alpha\cdot\overline{\alpha}^{\Delta+b}, \quad \forall b\in\{0,\ldots,\Delta-1\}. & \text{(68d)} \end{cases}$$

We now use the suffix-of-previous-and-current-states Markov chain $\mathcal{C}_{\boldsymbol{F}}$ to construct another Markov chain. For notational purpose, we let $\boldsymbol{P}$ stand for $S_{t-\Delta}\ldots S_t$, which are states in the previous Δ rounds and the state in the current round t. Then we consider a Markov chain to represent the transition of $\boldsymbol{F}_{t-\Delta-1}S_{t-\Delta}\ldots S_t$, and denote this Markov chain by $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$, where "||" intuitively means concatenation. The random variable $\boldsymbol{F}_{t-\Delta-1}$ represents the suffix of the states in all rounds up to round $t-\Delta-1$, so $\boldsymbol{F}_{t-\Delta-1}S_{t-\Delta}\ldots S_t$ means the concatenation of i) the suffix of previous states before the Δ to last state, ii) the previous Δ states, and iii) the current state. We can see that the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ is *time-homogeneous*, *irreducible*, and *ergodic*.

As it will become clear, to analyze S_t of $\boldsymbol{F}_{t-\Delta-1}S_{t-\Delta}\ldots S_t$, knowing whether S_t is H or N is not enough, and we need to know the exact number of blocks mined by the honest nodes at round t in the case of S_t being H (i.e., when at least one block is mined by the honest nodes at round t). To this end, we let H_h be the event that the honest nodes mine h number of block at round t. Then the values that S_t can take is given by the following set:

$$\text{Detailed-State-Set} := \{H_h,\ N \mid 1\leq h\leq \mu n\}. \quad (69)$$

Clearly, the H state comprises all H_h states for $1\leq h\leq\mu n$.

Below we analyze the stationary distribution of the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$. For $\boldsymbol{f}\in$ Suffix-Set, $s^{(1)}\in$ Detailed-State-Set, …, $s^{(\Delta+1)}\in$ Detailed-State-Set, we let $\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)})$ be the stationary probability of vertex $\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)}$, where Suffix-Set and Detailed-State-Set are given by Eq. (69) and (67); i.e.,

$$\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)}) = \lim_{t\to\infty}\mathbb{P}\left[\boldsymbol{F}_{t-\Delta-1}S_{t-\Delta}\ldots S_t=\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)}\right]. \quad (70)$$

Since $\mathbb{P}\left[\boldsymbol{F}_{t-\Delta-1}S_{t-\Delta}\ldots S_t=\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)}\right]$ equals $\mathbb{P}\left[\boldsymbol{F}_{t-\Delta-1}=\boldsymbol{f}\right]\prod_{i=1}^{\Delta+1}\mathbb{P}\left[S_{t-\Delta-1+i}=s^{(i)}\right]$, we obtain from Eq. (86) and (70) that

$$\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)})=\pi_{\boldsymbol{F}}(\boldsymbol{f})\prod_{i=1}^{\Delta+1}\mathbb{P}\left[s^{(i)}\right]. \quad (71)$$

We can also prove Eq. (71) by analyzing the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ directly. A proof is deferred to Appendix P of the online full version [7].

From Eq. (71), we can compute the stationary distribution $\pi_{\boldsymbol{F}||\boldsymbol{P}}$ of the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ using expressions of $\pi_{\boldsymbol{F}}$ in Eq. (68a)–(68d) and the following Eq. (72):

$$\mathbb{P}\left[s^{(i)}\right]=\begin{cases}\binom{\mu n}{h}p^h(1-p)^{\mu n-h}, & \text{if } s^{(i)}=H_h,\\ & \text{for each } h \text{ satisfying } 1\leq h\leq\mu n,\\ \overline{\alpha}, \text{ for } \overline{\alpha}=(1-p)^{\mu n}, & \text{if } s^{(i)}=N.\end{cases} \quad (72)$$

Eq. (72) follows from the result that since each honest node mines a block independently with probability p in a round, the number of blocks mined by the μn honest nodes in each round follows $\text{binom}(\mu n,p)$, which denotes a binomial distribution with μn being the number of trials and p being the success probability for each trial.

We now explain that when we have $\left(\boldsymbol{f}=HN^{\geq\Delta}\right)\wedge\left(s^{(1)}=H_1\right)\wedge\left(s^{(2)}=\ldots=s^{(\Delta+1)}=N\right)$, the $\boldsymbol{F}||\boldsymbol{P}$ state $\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)}$, which we write as $HN^{\geq\Delta}||H_1N^{\Delta}$ for notational simplicity, represents a convergence opportunity. Specifically, the pattern of $HN^{\geq\Delta}||H_1N^{\Delta}$ means the following consecutive events:

i) a benign node mines a block in a round,

ii) at least Δ rounds pass in which no benign node mines a block, which means that at the end of the Δ rounds, all

benign nodes know all benign blocks and hence agree on the maximum length of the chain (they may not agree on the same chain),

iii) a benign node mines a block $\mathcal{B}$ in a new round and thus extends a chain by one more block than the longest chain of the previous round, and

iv) Δ rounds pass in which no benign node mines a block. Thus, at the end, all honest miners know the new block $\mathcal{B}$ and agree on the single longest chain as the one having $\mathcal{B}$.

Then we compute the stationary probability of the $\boldsymbol{F}||\boldsymbol{P}$ state $\boldsymbol{f}s^{(1)}\dots s^{(\Delta+1)}$ state being $HN^{\geq\Delta}||H_1N^{\Delta}$ as follows by using Eq. (71):

$$\pi_{\boldsymbol{F}||\boldsymbol{P}}(HN^{\geq\Delta}||H_1N^{\Delta}) = \pi_{\boldsymbol{F}}(HN^{\geq\Delta})\mathbb{P}\left[H_1\right]\left(\mathbb{P}\left[N\right]\right)^{\Delta}. \tag{73}$$

From Eq. (72), it holds that

$$\mathbb{P}\left[H_1\right] = \alpha_1 \text{ for } \alpha_1 := p\mu n \times (1-p)^{\mu n - 1}. \tag{74}$$

From Eq. (68c) and Eq. (74), we obtain

$$\pi_{\boldsymbol{F}||\boldsymbol{P}}(HN^{\geq\Delta}||H_1N^{\Delta}) = \overline{\alpha}^{\Delta}\cdot\alpha_1\cdot\overline{\alpha}^{\Delta} = \overline{\alpha}^{2\Delta}\alpha_1. \tag{75}$$

We define f_t as the indicator function that the visited vertex at time t is the state $HN^{\geq\Delta}||H_1N^{\Delta}$. For the T-step random walk on the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ in the T rounds from round t_0 to t_0+T-1, let the visited vertices be $V_{t_0},\dots,V_{t_0+T-1}$. Then from Eq. (75), we have that for $t\in\{t_0,\dots,t_0+T-1\}$:

- $f_t(V_t)$ equals 1 if V_t is the state $HN^{\geq\Delta}||H_1N^{\Delta}$, which happens with probability $\overline{\alpha}^{2\Delta}\alpha_1$;
- $f_t(V_t)$ equals 0 if V_t is not the state $HN^{\geq\Delta}||H_1N^{\Delta}$, which happens with probability $1-\overline{\alpha}^{2\Delta}\alpha_1$.

Then the expectation of the binary variable $f_t(V_t)$ is

$$\mathbb{E}\left[f_t(V_t)\right] = \overline{\alpha}^{2\Delta}\alpha_1. \tag{76}$$

With $C(t_0,t_0+T-1)$ being the number of times that $HN^{\geq\Delta}||H_1N^{\Delta}$ is visited (i.e., the number of convergence opportunities) from round t_0 to t_0+T-1, we have

$$C(t_0,t_0+T-1) = \textstyle\sum_{t=t_0}^{t_0+T-1} f_t(V_t). \tag{77}$$

From the above, the random variables $f_t(V_t)|_{t=t_0}^{t_0+T-1}$ are identically distributed, but are not independent. Since the linearity of expectation holds regardless of whether the random variables are independent, we use Eq. (76) to obtain

$$\mathbb{E}\left[C(t_0,t_0+T-1)\right] = \textstyle\sum_{t=t_0}^{t_0+T-1}\mathbb{E}\left[f_t(V_t)\right] = T\overline{\alpha}^{2\Delta}\alpha_1;$$

i.e., Eq. (64) is proved. Using Eq. (64) and Eq. (65) which we have both shown, we know that Inequality (18) is the same as Inequality (56).

B. Putting things together to prove Theorem 2

We have proved in Section VI-A that Inequality (18) as a condition of Theorem 2 is the same as Inequality (56). Also, Eq. (65) in Section VI-A shows that $A(t_0,t_0+T-1)$ can be written as $\Omega(T)$. We prove Inequalities (57) and (58) in Appendices C and D of the online full version [7] (where [20]–[23] are cited). Then we combine (63) and (56) (57) (58) with $A(t_0,t_0+T-1)=\Omega(T)$ to complete proving Theorem 2. ■

VII. Conclusion

In this paper, we analyze the consistency of Nakamoto's blockchain protocol. Let μ (resp., ν) be the fraction of computational power controlled by benign miners (resp., the adversary), where $\mu+\nu=1$. With c denoting the expected number of network delays before some block is mined, we prove for the first time that to ensure the consistency property of Nakamoto's blockchain protocol in an asynchronous network, it suffices to have c to be just slightly greater than $\frac{2\mu}{\ln(\mu/\nu)}$. This expression is both neater and stronger than existing ones. In the proof, we formulate novel Markov chains which characterize the numbers of mined blocks in different rounds.

References

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] J. Garay, A. Kiayias, and N. Leonardos, "The bitcoin backbone protocol: Analysis and applications," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*, 2015, pp. 281–310.

[3] R. Pass, L. Seeman, and A. Shelat, "Analysis of the blockchain protocol in asynchronous networks," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*, 2017, pp. 643–673.

[4] R. Pass and E. Shi, "Fruitchains: A fair blockchain," in *ACM Symposium on Principles of Distributed Computing (PODC)*, 2017, pp. 315–324.

[5] E. Shi, "Analysis of deterministic longest-chain protocols," *IACR Cryptology ePrint Archive*, vol. 2018, p. 1079, 2018.

[6] L. Kiffer, R. Rajaraman, and A. Shelat, "A better method to analyze blockchain consistency," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2018, pp. 729–744.

[7] J. Zhao, J. Tang, Z. Li, H. Wang, K.-Y. Lam, and K. Xue, "An analysis of blockchain consistency in asynchronous networks: Deriving a neat bound," 2020, full version of this paper. Available online at http://www.ntu.edu.sg/home/junzhao/BlockchainConsistency.pdf.

[8] A. Kiayias and G. Panagiotakos, "Speed-security tradeoffs in blockchain protocols," *IACR Cryptology ePrint Archive*, vol. 2015, p. 1019, 2015.

[9] J. Garay, A. Kiayias, and N. Leonardos, "The bitcoin backbone protocol with chains of variable difficulty," in *Annual International Cryptology Conference (CRYPTO)*, 2017, pp. 291–323.

[10] R. Zhang and B. Preneel, "Lay down the common metrics: Evaluating proof-of-work consensus protocols' security," in *IEEE Symposium on Security and Privacy (SP)*, 2019.

[11] A. Kiayias, A. Russell, B. David, and R. Oliynykov, "Ouroboros: A provably secure proof-of-stake blockchain protocol," in *Annual International Cryptology Conference (CRYPTO)*, 2017, pp. 357–388.

[12] B. David, P. Gaži, A. Kiayias, and A. Russell, "Ouroboros Praos: An adaptively-secure, semi-synchronous proof-of-stake blockchain," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*, 2018, pp. 66–98.

[13] C. Badertscher, P. Gaži, A. Kiayias, A. Russell, and V. Zikas, "Ouroboros genesis: Composable proof-of-stake blockchains with dynamic availability," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2018, pp. 913–930.

[14] W. Li, S. Andreina, J.-M. Bohli, and G. Karame, "Securing proof-of-stake blockchain protocols," in *Data Privacy Management, Cryptocurrencies and Blockchain Technology*. Springer, 2017, pp. 297–315.

[15] Y. Gilad, R. Hemo, S. Micali, G. Vlachos, and N. Zeldovich, "Algorand: Scaling byzantine agreements for cryptocurrencies," in *Proceedings of the 26th Symposium on Operating Systems Principles (SOSP)*, 2017, pp. 51–68.

[16] M. Castro, B. Liskov *et al.*, "Practical byzantine fault tolerance," in *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, vol. 99, no. 1999, 1999, pp. 173–186.

[17] W. Wang, D. T. Hoang, P. Hu, Z. Xiong, D. Niyato, P. Wang, Y. Wen, and D. I. Kim, "A survey on consensus mechanisms and mining strategy management in blockchain networks," *IEEE Access*, vol. 7, pp. 22 328–22 370, 2019.

[18] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," *Future Generation Computer Systems*, vol. 8, p. 274, 2017.

[19] D. A. Levin and Y. Peres, *Markov chains and mixing times*. American Mathematical Soc., 2017, vol. 107.

[20] K.-M. Chung, H. Lam, Z. Liu, and M. Mitzenmacher, "Chernoff–Hoeffding bounds for Markov chains: Generalized and simplified," *arXiv preprint arXiv:1201.0559*, 2012.

[21] R. Arratia and L. Gordon, "Tutorial on large deviations for the binomial distribution," *Bulletin of mathematical biology*, vol. 51, no. 1, pp. 125–131, 1989.

[22] J. Zhao, O. Yağan, and V. Gligor, "k-connectivity in random key graphs with unreliable links," *IEEE Transactions on Information Theory*, vol. 61, no. 7, pp. 3810–3836, 2015.

[23] A. E. Taylor, "L'hospital's rule," *The American Mathematical Monthly*, vol. 59, no. 1, pp. 20–24, 1952.

APPENDIX

A. Proof of Claim 1

We first present Claim 2 on Page 5, which is Theorem 4.4 on Page 8 of [6] after we correct $\mu \cdot p$ to α in many places of [6]. Then we perform some computations to show Claim 1 using Claim 2.

Claim 2 (Theorem 4.4 of [6] after we correct $\mu \cdot p$ to α in many places of [6])**.** *With some notation defined below, Nakamoto's blockchain protocol satisfies consistency if there exists a positive constant δ_3 such that*

$$\frac{P_\Delta^2}{\sum_{i,j\in\{0,1\}} P_{ij}\pi_i\ell_{ij}} \geq (1+\delta_3)\beta, \text{ for } \beta := p\nu n, \tag{78}$$

where

1): *β denotes the expected number of blocks mined in each round by the adversary controlling νn miners;*
2): *P_Δ (defined on Page 6 of [6]) denotes the probability of Δ silent rounds (after we correct $\mu \cdot p$ to α on Page 6 of [6], it holds that $P_\Delta = (1-\alpha)^\Delta = \overline{\alpha}^\Delta$ from the definitions of α and $\overline{\alpha}$ in Eq. (7) and Eq. (8));*
3): *π_0 and π_1 (denoting stationary probabilities of states S_0 and S_1 in the Markov chain $\circlearrowleft S_0 \rightleftarrows S_1 \circlearrowright$ on Page 6 of [6]) are given as follows:*
   3a): *π_0 denotes the stationary probability of a "messy" state S_0 where honest mined blocks occur in less than Δ rounds from one another ($\pi_0 = 1 - P_\Delta$ from Page 7 of [6]);*
   3b): *π_1 denotes the stationary probability of the state S_1 where quiet periods between honest mined blocks is at least Δ rounds ($\pi_1 = P_\Delta$ from Page 7 of [6]);*
4): *P_{ij} for $i,j \in \{0,1\}$ denotes the probability of event e_{ij}, which represents the transition from state S_i to state S_j; more specifically,*
   4a): *P_{00} denotes the probability of e_{00}, meaning one quiet period of less than Δ rounds, followed by a round with at least one block mined by honest players[3] ($P_{00} = 1 - P_\Delta$ from Page 7 of [6]);*
   4b): *P_{01} denotes the probability of e_{01}, meaning one quiet period that is at least Δ rounds ($P_{01} = P_\Delta$ from Page 7 of [6]);*
   4c): *P_{11} denotes the probability of e_{11}, meaning a single honest mined block[3], followed by a quiet period of at least Δ rounds ($P_{11} = P_\Delta$ from Page 7 of [6]);*
   4d): *P_{10} denotes the probability of e_{10}, meaning a round with at least one block mined by honest players, followed by one quiet period of less than Δ rounds, followed by a round with at least one block mined by honest players ($P_{10} = 1 - P_\Delta$ from Page 7 of [6]);*
5): *ℓ_{ij} for $i,j \in \{0,1\}$ denoting the expected time spent on the edge $S_i \to S_j$ in the Markov chain $\circlearrowleft S_0 \rightleftarrows S_1 \circlearrowright$ on Page 6 of [6]:*
   ◦ *With $p_{i|\leq\Delta}$ denoting $\mathbb{P}[\text{hit at time } i \,|\, \text{silience lasted} \leq \Delta]$ (after we correct $\mu \cdot p$ to α on Page 7 of [6], it holds that $p_{i|\leq\Delta} = \frac{(1-\alpha)^{i-1}\alpha}{\sum_{j=1}^{\Delta}(1-\alpha)^{j-1}\alpha}$), the expressions of ℓ_{00}, ℓ_{01}, ℓ_{11}, and ℓ_{10} are as follows after we correct $\mu \cdot p$ to α on Page 7 of [6]: $\ell_{00} = \sum_{i=1}^{\Delta} i p_{i|\leq\Delta}$, $\ell_{01} = \Delta$, $\ell_{11} = \frac{1}{\alpha} + \Delta$, and $\ell_{10} = \frac{1}{\alpha} + \sum_{i=1}^{\Delta} i p_{i|\leq\Delta}$.*

[3]Note the phrase "a round with at least one block mined by honest players" in the definitions of e_{00}, e_{11}, and e_{10} (and hence P_{00}, P_{11}, and P_{10}) of Theorem 2. On Page 6 of [6], actually the phrase "single honest mined block" is used. However, for e_{ij} to exactly mean the transition from state S_i to state S_j for $i,j \in \{0,1\}$ (defined in "3a)" and "3b)" of the list in Theorem 2), there is no reason for requiring "single honest mined block".

We now use Claim 2 to show Claim 1. First, we use the expression of $p_{i|\leq\Delta}$ to compute ℓ_{00}:

$$\begin{aligned}
\ell_{00} &= \sum_{i=1}^{\Delta} i p_{i|\leq\Delta} \\
&= \sum_{i=1}^{\Delta} \frac{i(1-\alpha)^{i-1}\alpha}{\sum_{j=1}^{\Delta}(1-\alpha)^{j-1}\alpha} \\
&= \frac{1}{\alpha} - \frac{\Delta(1-\alpha)^\Delta}{1-(1-\alpha)^\Delta} \\
&= \frac{1}{\alpha} - \frac{\Delta P_\Delta}{1-P_\Delta}.
\end{aligned} \tag{79}$$

Then we calculate the left-hand side in Inequality (78):

$$\begin{aligned}
&\frac{P_\Delta^2}{\sum_{i,j\in\{0,1\}} P_{ij}\pi_i\ell_{ij}} \\
&= \frac{P_\Delta^2}{P_{00}\pi_0\ell_{00} + P_{01}\pi_0\ell_{01} + P_{11}\pi_1\ell_{11} + P_{10}\pi_1\ell_{10}} \\
&= \frac{P_\Delta^2}{\left[\begin{array}{l}(1-P_\Delta)(1-P_\Delta)(\frac{1}{\alpha} - \frac{\Delta P_\Delta}{1-P_\Delta}) + P_\Delta(1-P_\Delta)\Delta \\ +P_\Delta P_\Delta(\frac{1}{\alpha}+\Delta) + (1-P_\Delta)P_\Delta(\frac{2}{\alpha} - \frac{\Delta P_\Delta}{1-P_\Delta})\end{array}\right]} \\
&= P_\Delta^2\alpha \\
&= \overline{\alpha}^{2\Delta}\alpha.
\end{aligned} \tag{80}$$

Hence, Claim 1 follows from Claim 2.

B. Deriving the stationary distribution of the suffix-of-previous-and-current-states Markov chain $\mathcal{C}_{\boldsymbol{F}}$

We now derive the stationary distribution of the suffix-of-previous-and-current-states Markov chain $\mathcal{C}_{\boldsymbol{F}}$. To this end, we first analyze the state transition in $\mathcal{C}_{\boldsymbol{F}}$.

Let s_t be $\boldsymbol{f}_t$'s state in round t. We define a function $\text{suffix}(\cdot)$ such that $\left(\boldsymbol{F}_{t-1} = \boldsymbol{f}_{t-1}\right) \wedge (S_t = s_t)$ produces $\boldsymbol{F}_t = \text{suffix}(\boldsymbol{f}_{t-1}||s_t)$. Then we have

$$\begin{aligned}
&\mathbb{P}[\boldsymbol{F}_t = \boldsymbol{f}_t] \\
&= \sum_{\substack{\boldsymbol{f}_{t-1}\in\text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-1}||s_t)=\boldsymbol{f}_t}} \mathbb{P}\left[\left(\boldsymbol{F}_{t-1} = \boldsymbol{f}_{t-1}\right) \wedge (S_t = s_t)\right] \\
&= \sum_{\substack{\boldsymbol{f}_{t-1}\in\text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-1}||s_t)=\boldsymbol{f}_t}} \left(\mathbb{P}\left[\boldsymbol{F}_{t-1} = \boldsymbol{f}_{t-1}\right]\mathbb{P}[S_t = s_t]\right),
\end{aligned} \tag{81}$$

where the last step uses the independence between $\left(\boldsymbol{F}_{t-1} = \boldsymbol{f}_{t-1}\right)$ and $(S_t = s_t)$.

Based on Eq. (81), we now set $\boldsymbol{f}_t$ as each vertex of Markov chain $\mathcal{C}_{\boldsymbol{F}}$ to obtain the specific transition rules.

Case of $\boldsymbol{f}_t$ in Eq. (81) being $HN^{\leq\Delta-1}HN^a$. We obtain from Eq. (81) and the above result ① that for any $a \in \{1,\ldots,\Delta-1\}$,

$$\begin{aligned}
&\mathbb{P}\left[\boldsymbol{F}_t = HN^{\leq\Delta-1}HN^a\right] \\
&= \mathbb{P}\left[\boldsymbol{F}_{t-a} = HN^{\leq\Delta-1}H\right] \prod_{i=t-a+1}^{t} \mathbb{P}[S_i = N] \\
&= \mathbb{P}\left[\boldsymbol{F}_{t-a} = HN^{\leq\Delta-1}H\right] \cdot \overline{\alpha}^a,
\end{aligned} \tag{82}$$

where the last step uses $\mathbb{P}[S_i = N] = \overline{\alpha}$.

Case of $\boldsymbol{f}_t$ in Eq. (81) being $HN^{\geq\Delta}HN^b$. We obtain from Eq. (81) and the above result ② that for any $b \in$

$\{0,\ldots,\Delta-1\}$,

$$\begin{aligned}
&\mathbb{P}\left[\boldsymbol{F}_t = HN^{\geq\Delta}HN^b\right]\\
&= \mathbb{P}\left[\boldsymbol{F}_{t-b-1} = HN^{\geq\Delta}\right]\mathbb{P}\left[S_{t-b}=H\right]\prod_{i=t-b+1}^{t}\mathbb{P}\left[S_i=N\right]\\
&= \mathbb{P}\left[\boldsymbol{F}_{t-b-1} = HN^{\geq\Delta}\right]\cdot\alpha\cdot\overline{\alpha}^b, \qquad (83)
\end{aligned}$$

where the last step uses $\mathbb{P}\left[S_{t-b}=H\right]=\alpha$ and $\mathbb{P}\left[S_i=N\right]=\overline{\alpha}$.

Case of $\boldsymbol{f}_t$ in Eq. (81) being $HN^{\leq\Delta-1}H$. We obtain from Eq. (81) and the above result ③ that

$$\begin{aligned}
&\mathbb{P}\left[\boldsymbol{F}_t = HN^{\leq\Delta-1}H\right]\\
&= \mathbb{P}\left[S_t=H\right]\cdot\Bigg(\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\leq\Delta-1}H\right]\\
&\qquad + \sum_{a=1}^{\Delta-1}\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\leq\Delta-1}HN^a\right]\\
&\qquad + \sum_{b=0}^{\Delta-1}\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\geq\Delta}HN^b\right]\Bigg)\\
&= \alpha\cdot\Bigg(\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\leq\Delta-1}H\right]\\
&\qquad + \sum_{a=1}^{\Delta-1}\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\leq\Delta-1}HN^a\right]\\
&\qquad + \sum_{b=0}^{\Delta-1}\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\geq\Delta}HN^b\right]\Bigg), \qquad (84)
\end{aligned}$$

where the last step uses $\mathbb{P}\left[S_t=H\right]=\alpha$.

Case of $\boldsymbol{f}_t$ in Eq. (81) being $HN^{\geq\Delta}$. We obtain from Eq. (81) and the above result ④ that

$$\begin{aligned}
&\mathbb{P}\left[\boldsymbol{F}_t = HN^{\geq\Delta}\right]\\
&= \mathbb{P}\left[S_t=N\right]\cdot\Bigg(\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\geq\Delta}\right]\\
&\qquad + \mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\leq\Delta-1}HN^{\Delta-1}\right]\\
&\qquad + \mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\geq\Delta}HN^{\Delta-1}\right]\Bigg)\\
&= \overline{\alpha}\cdot\Bigg(\mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\geq\Delta}\right]\\
&\qquad + \mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\leq\Delta-1}HN^{\Delta-1}\right]\\
&\qquad + \mathbb{P}\left[\boldsymbol{F}_{t-1} = HN^{\geq\Delta}HN^{\Delta-1}\right]\Bigg), \qquad (85)
\end{aligned}$$

where the last step uses $\mathbb{P}\left[S_t=N\right]=\overline{\alpha}$.

Below we analyze the stationary distribution of the Markov chain $\mathcal{C}_{\boldsymbol{F}}$. For $\boldsymbol{f}\in$ Suffix-Set, we let $\pi_{\boldsymbol{F}}(\boldsymbol{f})$ be the stationary probability of vertex $\boldsymbol{f}$, where Suffix-Set is given by Eq. (67); i.e.,

$$\pi_{\boldsymbol{F}}(\boldsymbol{f}) = \lim_{t\to\infty}\mathbb{P}\left[\boldsymbol{F}_t=\boldsymbol{f}\right]. \qquad (86)$$

Summarizing Eq. (82)–(85), to derive Markov chain $\mathcal{C}_{\boldsymbol{F}}$'s stationary distribution denoted by $\pi_{\boldsymbol{F}}$, we obtain

$$\begin{cases}
\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}HN^a) = \pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}H)\cdot\overline{\alpha}^a, & (87a)\\
\qquad\qquad \forall a\in\{1,\ldots,\Delta-1\},\\
\pi_{\boldsymbol{F}}(HN^{\geq\Delta}HN^b) = \pi_{\boldsymbol{F}}(HN^{\geq\Delta})\cdot\alpha\cdot\overline{\alpha}^b, & (87b)\\
\qquad\qquad \forall b\in\{0,\ldots,\Delta-1\},\\
\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}H) = \alpha\cdot\Bigg(\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}H)\\
\quad + \sum_{a=1}^{\Delta-1}\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}HN^a) + \sum_{b=0}^{\Delta-1}\pi_{\boldsymbol{F}}(HN^{\geq\Delta}HN^b)\Bigg), & (87c)\\
\pi_{\boldsymbol{F}}(HN^{\geq\Delta}) = \overline{\alpha}\cdot\Bigg(\pi_{\boldsymbol{F}}(HN^{\geq\Delta})\\
\quad + \pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}HN^{\Delta-1}) + \pi_{\boldsymbol{F}}(HN^{\geq\Delta}HN^{\Delta-1})\Bigg), & (87d)\\
\begin{bmatrix}\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}H) + \sum_{a=1}^{\Delta-1}\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}HN^a)\\ + \pi_{\boldsymbol{F}}(HN^{\geq\Delta}) + \sum_{b=0}^{\Delta-1}\pi_{\boldsymbol{F}}(HN^{\geq\Delta}HN^b)\end{bmatrix} = 1, & (87e)
\end{cases}$$

where Eq. (87a)–(87d) are from Eq. (82)–(85), respectively, and Eq. (87e) simply means that the stationary probabilities of all the states sum to 1.

From Eq. (87a)–(87e), we derive that

$$\begin{cases}
\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}H) = \alpha\cdot(1-\overline{\alpha}^{\Delta}), & (88a)\\
\pi_{\boldsymbol{F}}(HN^{\leq\Delta-1}HN^a) = \alpha\cdot(1-\overline{\alpha}^{\Delta})\cdot\overline{\alpha}^a, & (88b)\\
\qquad\qquad \forall a\in\{1,\ldots,\Delta-1\},\\
\pi_{\boldsymbol{F}}(HN^{\geq\Delta}) = \overline{\alpha}^{\Delta}, & (88c)\\
\pi_{\boldsymbol{F}}(HN^{\geq\Delta}HN^b) = \alpha\cdot\overline{\alpha}^{\Delta+b}, & (88d)\\
\qquad\qquad \forall b\in\{0,\ldots,\Delta-1\}.
\end{cases}$$

C. Proving Inequality (57)

Recall from the previous subsection that the T-step random walk on the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ in the T rounds from round t_0 to t_0+T-1 visits vertices $V_{t_0},\ldots,V_{t_0+T-1}$. Let ϕ be the initial distribution of the random walk; i.e., ϕ represents the distribution at round t_0. Also recall that the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ is *time-homogeneous*, *irreducible*, and *ergodic*. Let $\tau(\epsilon,\alpha,\Delta)$ be the ϵ-mixing time of $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$, for $0<\epsilon\leq 1/8$. With $f_t(V_t)$ and $C(t_0,t_0+T-1)$ defined above, we use Theorem 3.1 of Reference [20] on the Chernoff–Hoeffding bounds for Markov chains to obtain the existence of a positive constant c independent of T,n,p,μ,Δ such that

$$\begin{aligned}
&\mathbb{P}\left[C(t_0,t_0+T-1)\leq(1-\delta_2)\cdot\mathbb{E}\left[C(t_0,t_0+T-1)\right]\right]\\
&\leq c\|\phi\|_{\pi}\exp\left(-\frac{{\delta_2}^2T\overline{\alpha}^{2\Delta}\alpha_1}{72\tau(\epsilon,\alpha,\Delta)}\right),\ \text{for constant } 0<\delta_2<1, \qquad (89)
\end{aligned}$$

where $\|\phi\|_{\pi}$, denoting the π-norm of the vector ϕ, is given by

$$\|\phi\|_{\pi} := \sqrt{\sum_{(\boldsymbol{f}||\boldsymbol{p})\in\text{Domain}(\boldsymbol{F}||\boldsymbol{P})}\frac{(\phi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}||\boldsymbol{p}))^2}{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}||\boldsymbol{p}))}},$$

where $\text{Domain}(\boldsymbol{F}||\boldsymbol{P}) :=$ Suffix-Set$\times$(Detailed-State-Set)$^{\Delta+1}$ since Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ represents the transition of $\boldsymbol{F}_{t-\Delta-1}S_{t-\Delta}\ldots S_t$. The term $T\overline{\alpha}^{2\Delta}\alpha_1$ in Inequality (89) comes from Eq. (64). We can also use Theorem 3.1 of Reference [20] to compute a bound for the tail probability $\mathbb{P}\left[C(t_0,t_0+T-1)\geq(1+\delta_2)\cdot\mathbb{E}\left[C(t_0,t_0+T-1)\right]\right]$. We do not present the result here since it is not needed.

Proposition 2 below provides an upper bound for $\|\phi\|_{\pi}$.

Proposition 2. *We have* $\|\phi\|_\pi \leq \frac{1}{\sqrt{\min \pi_{\boldsymbol{F}||\boldsymbol{P}}}}$, *where* $\min \pi_{\boldsymbol{F}||\boldsymbol{P}}$ *denotes the minimal value among* $\pi_{\boldsymbol{F}||\boldsymbol{P}}$ *and is given by* $\alpha \cdot \overline{\alpha}^{\Delta-1} \cdot \min\left\{1-\overline{\alpha}^{\Delta}, \overline{\alpha}^{\Delta}\right\} \cdot \left(\min\left\{p^{\mu n}, (1-p)^{\mu n}\right\}\right)^{\Delta+1}$.

We prove Proposition 2 in Appendix G.

From Proposition 2, $\|\phi\|_\pi$ is upper bounded by a term that depends on α and Δ (note that when α is given, $\overline{\alpha} := 1-\alpha$ is also given). Also, $\tau(\epsilon, \alpha, \Delta)$ denoting the ϵ-mixing time of the Markov chain $\mathcal{C}_{\boldsymbol{F}||\boldsymbol{P}}$ is clearly a non-increasing function of ϵ given α and Δ. In view of $0 < \epsilon \leq 1/8$, we can select ϵ as $1/8$ so that the bound in the right hand side of Inequality (89) is maximized. Then $\tau(1/8, \alpha, \Delta)$ depends on only α and Δ. Recall from Eq. (7) that α depends on n, p, μ. Hence, given n, p, μ, Δ, we use Inequality (89) to obtain the desired result (57) that $\mathbb{P}\left[C(t_0, t_0+T-1) \leq (1-\delta_2) \cdot \mathbb{E}\left[C(t_0, t_0+T-1)\right]\right]$ is upper bounded by $O(1) \cdot \exp\left(-\Omega\left(T\right)\right)$, where $O(1)$ is with respect to T.

D. Proving Inequality (58)

As already explained in Section VI-A, $A(t_0, t_0+T-1)$ follows the binomial distribution $\text{binom}(T\nu n, p)$. From [21], for a positive constant δ_4, with $D\left((1+\delta_4)p||p\right)$ denoting the relative entropy between a Bernoulli distribution of parameter $(1+\delta_4)p$ and a Bernoulli distribution of parameter p; i.e., defining

$$D\left((1+\delta_4)p||p\right) := (1+\delta_4)p\ln(1+\delta_4) + [1-(1+\delta_4)p]\ln\frac{1-(1+\delta_4)p}{1-p}, \tag{90}$$

we have

$$\begin{aligned}&\mathbb{P}\left[A(t_0, t_0+T-1) \geq (1+\delta_4) \cdot \mathbb{E}\left[A(t_0, t_0+T-1)\right]\right] \\ &\leq \exp\left(-T\nu n \cdot D\left((1+\delta_4)p||p\right)\right).\end{aligned} \tag{91}$$

Thus, given n, p, ν, we obtain the desired result (58) that $\mathbb{P}\left[A(t_0, t_0+T-1) \geq (1+\delta_4) \cdot \mathbb{E}\left[A(t_0, t_0+T-1)\right]\right]$ is upper bounded by $O(1) \cdot \exp\left(-\Omega\left(T\right)\right)$, where $O(1)$ is with respect to T.

E. Using Theorem 3 to prove Theorem 1

For c denoting $\frac{1}{pn\Delta}$, it is straightforward to show that a combination of Inequalities (21) and (22) in Theorem 3 is the same as Inequality (11) of Theorem 1. Hence, given Theorem 3, we know that if Inequality (11) holds, then the consistency of Nakamoto's blockchain protocol holds in a window of T rounds with probability at least $1 - O(1) \cdot \exp\left(-\Omega\left(T\right)\right)$.

To complete the proof of Theorem 1, next we show that under Inequality (12), we can write Inequality (11) as Inequality (13).

From $\mu = 1-\nu$ and the condition $\nu \geq \frac{1}{1+\exp(\Delta^{\delta_1})}$ of Inequality (12), we have

$$\ln\frac{\mu}{\nu} = \ln\frac{1-\nu}{\nu} \leq \ln\frac{1-\frac{1}{1+\exp(\Delta^{\delta_1})}}{\frac{1}{1+\exp(\Delta^{\delta_1})}} = \Delta^{\delta_1}. \tag{92}$$

From $\mu = 1-\nu$ and the condition $\nu \leq \frac{1}{1+\exp\left(\frac{1}{\Delta^{\delta_2}-1}\right)}$ of Inequality (12), we have

$$\ln\frac{\mu}{\nu} = \ln\frac{1-\nu}{\nu} \geq \ln\frac{1-\frac{1}{1+\exp\left(\frac{1}{\Delta^{\delta_2}-1}\right)}}{\frac{1}{1+\exp\left(\frac{1}{\Delta^{\delta_2}-1}\right)}} = \frac{1}{\Delta^{\delta_2}-1}, \tag{93}$$

which implies

$$\frac{\ln\frac{\mu}{\nu}+1}{\Delta\ln\frac{\mu}{\nu}} = \frac{1}{\Delta}\left(1+\frac{1}{\ln\frac{\mu}{\nu}}\right) \leq \Delta^{\delta_2-1}. \tag{94}$$

Here we set ϵ_1 by

$$\epsilon_1 := \Delta^{\delta_1+\delta_2-1}. \tag{95}$$

From (92) (94) (95) and the condition $\delta_1 + \delta_2 < 1$, letting ϵ_1 be $\Delta^{\delta_1+\delta_2-1}$, we obtain

$$\frac{2\mu}{\ln\frac{\mu}{\nu}} \geq \frac{2\mu}{\Delta^{\delta_1}} = \frac{2\mu}{\epsilon_1} \cdot \Delta^{\delta_2-1} > \frac{2(\ln\frac{\mu}{\nu}+1)\mu}{\epsilon_1\Delta\ln\frac{\mu}{\nu}}, \tag{96}$$

which means that Inequality (11) (i.e., $c \geq \max\left\{\left(\frac{2\mu}{\ln\frac{\mu}{\nu}}+\frac{1}{\Delta}\right)\frac{1+\epsilon_2}{1-\epsilon_1}, \frac{(\ln\frac{\mu}{\nu}+1)\mu}{\epsilon_1\Delta\ln\frac{\mu}{\nu}}\right\}$) becomes

$$c \geq \left[\frac{2\mu}{\ln(\mu/\nu)} + \frac{1}{\Delta}\right]\frac{1+\epsilon_2}{1-\epsilon_1}. \tag{97}$$

From (92) and $\mu > \frac{1}{2}$, we get

$$\frac{1}{\Delta} = \Delta^{-\delta_1} \cdot \Delta^{\delta_1-1} < \frac{2\mu}{\Delta^{\delta_1}} \cdot \Delta^{\delta_1-1} \leq \frac{2\mu}{\ln\frac{\mu}{\nu}} \cdot \Delta^{\delta_1-1}, \tag{98}$$

which means that a sufficient condition for (97) is

$$\begin{aligned}c &\geq \left[\frac{2\mu}{\ln(\mu/\nu)} + \frac{2\mu}{\ln\frac{\mu}{\nu}} \cdot \Delta^{\delta_1-1}\right]\frac{1+\epsilon_2}{1-\epsilon_1} \\ &= \frac{2\mu}{\ln(\mu/\nu)} \cdot (1+\epsilon_2) \cdot \frac{1+\Delta^{\delta_1-1}}{1-\Delta^{\delta_1+\delta_2-1}},\end{aligned} \tag{99}$$

where the last step uses (95).

The above result (99) gives Inequality (13). Hence, we have completed proving Theorem 1. ■

F. Explaining that δ_5 and δ_1 in Eq. (31) and Eq. (32) are both positive for $0 < \epsilon_1 < 1$ and $\epsilon_2 > 0$

Clearly, $\delta_5 > 0$ since the nominator and denominator of Eq. (31) are both positive. In addition, given

$$\begin{aligned}\delta_5 &> \frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{(\epsilon_1+\epsilon_2) + \frac{\epsilon_1+\epsilon_2}{\epsilon_1} \cdot (1-\epsilon_1) \cdot (\ln\frac{\mu}{\nu}+1)} \\ &= \frac{\epsilon_1\ln\frac{\mu}{\nu}}{\epsilon_1 + (1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)} = \frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln\frac{\mu}{\nu}},\end{aligned} \tag{100}$$

we have

$$\begin{aligned}\delta_1 &= (1+\delta_5)\cdot\left(1-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{\ln\frac{\mu}{\nu}+1}\right) - 1 \\ &> \left[1+\frac{\epsilon_1\ln\frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln\frac{\mu}{\nu}}\right]\cdot\left(1-\frac{\epsilon_1\ln\frac{\mu}{\nu}}{\ln\frac{\mu}{\nu}+1}\right) - 1 = 0.\end{aligned} \tag{101}$$

G. Proof of Proposition 2

The π-norm of ϕ is

$$\begin{aligned}\|\phi\|_\pi &= \sqrt{\sum_{(\boldsymbol{f}||\boldsymbol{p}) \in \text{Domain}(\boldsymbol{F}||\boldsymbol{P})} \frac{(\phi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}||\boldsymbol{p}))^2}{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}||\boldsymbol{p}))}} \\ &\leq \sqrt{\sum_{(\boldsymbol{f}||\boldsymbol{p}) \in \text{Domain}(\boldsymbol{F}||\boldsymbol{P})} \frac{\phi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}||\boldsymbol{p})}{\min\pi_{\boldsymbol{F}||\boldsymbol{P}}}} \\ &= \frac{1}{\sqrt{\min\pi_{\boldsymbol{F}||\boldsymbol{P}}}},\end{aligned} \tag{102}$$

where $\min\pi_{\boldsymbol{F}||\boldsymbol{P}}$ denotes the minimal value among $\pi_{\boldsymbol{F}||\boldsymbol{P}}$.

Recall from Eq. (71) that

$$\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}s^{(1)}\ldots s^{(\Delta+1)}) = \pi_{\boldsymbol{F}}(\boldsymbol{f})\prod_{i=1}^{\Delta+1}\mathbb{P}\left[s^{(i)}\right]. \tag{103}$$

For $s^{(i)} \in$ Detailed-State-Set for Detailed-State-Set in Eq. (69), we have

$$\min_{s^{(i)} \in \text{Detailed-State-Set}} \mathbb{P}\left[s^{(i)}\right] = \begin{cases} p^{\mu n}, & \text{if } p \leq \frac{1}{2}, \\ (1-p)^{\mu n}, & \text{if } p > \frac{1}{2}, \end{cases}$$

so that we can write

$$\min_{s^{(i)} \in \text{Detailed-State-Set}} \mathbb{P}\left[s^{(i)}\right] = \min\left\{p^{\mu n}, (1-p)^{\mu n}\right\}. \quad (104)$$

Then Eq. (103) implies that

$$\min \pi_{\boldsymbol{F}||\boldsymbol{P}} = (\min \pi_{\boldsymbol{F}}) \cdot \left(\min\left\{p^{\mu n}, (1-p)^{\mu n}\right\}\right)^{\Delta+1}, \quad (105)$$

where the minimal value among $\pi_{\boldsymbol{F}}$ is

$$\begin{aligned} \min \pi_{\boldsymbol{F}} &= \min\left\{\alpha \cdot (1-\overline{\alpha}^{\Delta}) \cdot \overline{\alpha}^{\Delta-1}, \alpha \cdot \overline{\alpha}^{2\Delta-1}\right\} \\ &= \alpha \cdot \overline{\alpha}^{\Delta-1} \cdot \min\left\{1-\overline{\alpha}^{\Delta}, \overline{\alpha}^{\Delta}\right\}. \end{aligned} \quad (106)$$

Combining (102) (105) (106), we complete proving Proposition 2. ■

H. Proof of Lemma 2

Recall the expression of α_1 in Inequality (74); i.e., $\alpha_1 = p\mu n \times (1-p)^{\mu n - 1}$. Then given the condition $0 < p\mu n < 1$ and the result $\mu n - 1 > \frac{1}{2}n - 1 \geq 1$ from $\mu > \frac{1}{2}$ and $n \geq 4$, we use Fact 2 on Page 20 of [22] to obtain

$$\begin{aligned} \alpha_1 &= p\mu n \cdot (1-p)^{\mu n - 1} \geq p\mu n \cdot [1 - p \cdot (\mu n - 1)] \\ &\geq p\mu n \cdot (1 - p\mu n). \end{aligned} \quad (107)$$

Then Inequality (107) induces

$$\begin{aligned} &\left\{p\mu n \cdot (1-p\mu n)\, \overline{\alpha}^{2\Delta} \geq (1+\delta_1) p\nu n\right\} \\ &\Longrightarrow \left\{\overline{\alpha}^{2\Delta} \alpha_1 \geq (1+\delta_1) p\nu n\right\}. \end{aligned} \quad (108)$$

The statement $p\mu n\,(1-p\mu n)\,\overline{\alpha}^{2\Delta} \geq (1+\delta_1)p\nu n$ is equivalent to $\overline{\alpha} \geq \left(\frac{1+\delta_1}{1-p\mu n} \cdot \frac{\nu}{\mu}\right)^{1/(2\Delta)}$; i.e., Inequality (35). This along with Inequality (108) implies the desired result. ■

I. Proof of Lemma 3

Proof of $\delta_5 > 0$: Given $\delta_5 > \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln \frac{\mu}{\nu}}$ with $0 < \epsilon_1 < 1$ and $\ln \frac{\mu}{\nu} > 0$ from $0 < \nu < \mu$, we have $\delta_5 > 0$.

Proof of $\delta_1 > 0$: Given $\delta_5 > \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{1+(1-\epsilon_1)\ln \frac{\mu}{\nu}}$ and $\delta_1 = (1+\delta_5) \cdot \left(1 - \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{\ln \frac{\mu}{\nu} + 1}\right) - 1$, we have

$$\delta_1 > \left(1 + \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{1 + (1-\epsilon_1)\ln \frac{\mu}{\nu}}\right) \cdot \left(1 - \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{\ln \frac{\mu}{\nu} + 1}\right) - 1 = 0.$$

Proof of $\left(\frac{1+\delta_1}{1-p\mu n}\right)^{1/(2\Delta)} < 1 + \frac{\delta_5}{2\Delta}$: Using the conditions $pn \leq \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{(\ln \frac{\mu}{\nu}+1)\mu}$ and $\delta_1 = (1+\delta_5) \cdot \left(1 - \frac{\epsilon_1 \ln \frac{\mu}{\nu}}{\ln \frac{\mu}{\nu}+1}\right) - 1$, we have $1 + \delta_1 \leq (1+\delta_5) \cdot (1 - p\mu n)$, which means $\frac{1+\delta_1}{1-p\mu n} \leq 1 + \delta_5$. Moreover, we have $1 + \delta_5 < \left(1 + \frac{\delta_5}{2\Delta}\right)^{2\Delta}$ from the binomial series. Summarizing the above results, we obtain $\left(\frac{1+\delta_1}{1-p\mu n}\right)^{1/(2\Delta)} < 1 + \frac{\delta_5}{2\Delta}$. ■

J. Proof of Lemma 4

Recalling $\overline{\alpha} = (1-p)^{\mu n}$ from Eq. (8), we have

$$\begin{aligned} &\left\{\overline{\alpha} \geq \left(1 + \frac{\delta_5}{2\Delta}\right) \cdot \left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right\} \\ &\Longleftrightarrow \left\{(1-p)^{\mu n} \geq \left(1 + \frac{\delta_5}{2\Delta}\right) \cdot \left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right\} \\ &\Longleftrightarrow \left\{p \leq 1 - \left[\left(1 + \frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]^{1/(\mu n)}\right\} \\ &\Longleftrightarrow \left\{c := \frac{1}{pn\Delta} \geq \frac{1}{n\Delta\left\{1 - \left[\left(1 + \frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]^{1/(\mu n)}\right\}}\right\}. \end{aligned}$$

■

K. Proof of Proposition 1

Our goal is to prove

$$1 - \left(1 + \frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)} > 0, \quad (109)$$

given the condition $0 < \delta_5 < \ln \frac{\mu}{\nu}$.

Clearly, Inequality (109) holds once we show

$$1 - \left(1 + \frac{1}{2\Delta}\ln \frac{\mu}{\nu}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)} > 0. \quad (110)$$

After defining $f(x) := x^{1/(2\Delta)} - \frac{1}{2\Delta}\ln x - 1$ for $x \geq 1$, the term $1 - \left(1 + \frac{1}{2\Delta}\ln \frac{\mu}{\nu}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}$ in Inequality (110) becomes $f\left(\frac{\mu}{\nu}\right) \cdot \left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}$, so Inequality (110) holds once we prove $f\left(\frac{\mu}{\nu}\right) > 0$. To this end, we derive $f'(x) := \frac{1}{2x\Delta}\left(x^{1/(2\Delta)} - 1\right) > 0$ for $x > 1$, so that $f(x)$ is a strictly increasing function for $x \geq 1$. Then given $f(1) = 0$, we obtain $f(x) > 0$ for $x > 1$ and thus $f\left(\frac{\mu}{\nu}\right) > 0$ given $\frac{\mu}{\nu} > 1$. The result $f\left(\frac{\mu}{\nu}\right) > 0$ means $\left(\frac{\mu}{\nu}\right)^{1/(2\Delta)} - \frac{1}{2\Delta}\ln \frac{\mu}{\nu} - 1 > 0$, which implies Inequality (110) and thus Inequality (109). ■

L. Proof of Lemma 5

First, we know from Proposition 1 that the denominators in both sides of Inequality (45) of Lemma 5 are positive.

With A defined by

$$A := 1 - \left(1 + \frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}, \quad (111)$$

we know $A > 0$ from Proposition 1. Also, clearly $A < 1$. With $0 < A < 1$ and $\mu n > \frac{n}{2} \geq 2$ from $\mu > \frac{1}{2}$ and $n \geq 4$, we use Fact 2 on Page 20 of [22] to obtain $(1 - \frac{A}{\mu n})^{\mu n} \geq 1 - \frac{A}{\mu n} \cdot \mu n = 1 - A > 0$, which implies $(1-A)^{1/(\mu n)} \leq 1 - \frac{A}{\mu n}$. Hence,

$$\frac{\mu}{A\Delta} = \frac{1}{n\Delta[1 - (1 - A/(\mu n))]} \geq \frac{1}{n\Delta[1 - (1-A)^{1/(\mu n)}]}. \quad (112)$$

We plug Eq. (111) (i.e., the expression of A) into (112) and complete proving Lemma 5. ■

M. Proof of Lemma 6

We evaluate $\frac{1}{1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}}$ appearing in the desired result. We have

$$\frac{1}{1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}} = \frac{\left(\frac{\mu}{\nu}\right)^{1/(2\Delta)}}{\left(\frac{\mu}{\nu}\right)^{1/(2\Delta)}-\left(1+\frac{\delta_5}{2\Delta}\right)}$$
$$= \left[1+\frac{\frac{\delta_5}{2\Delta}}{\left(\frac{\mu}{\nu}\right)^{1/(2\Delta)}-\left(1+\frac{\delta_5}{2\Delta}\right)}\right]\cdot\frac{1}{1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}}. \quad (113)$$

We further bound the term $\left(\frac{\mu}{\nu}\right)^{1/(2\Delta)}-\left(1+\frac{\delta_5}{2\Delta}\right)$ in Eq. (113):

$$\left(\frac{\mu}{\nu}\right)^{1/(2\Delta)}-\left(1+\frac{\delta_5}{2\Delta}\right)$$
$$= \exp\left(\frac{1}{2\Delta}\ln\frac{\mu}{\nu}\right)-\left(1+\frac{\delta_5}{2\Delta}\right)$$
$$> 1+\frac{1}{2\Delta}\ln\frac{\mu}{\nu}-\left(1+\frac{\delta_5}{2\Delta}\right)$$
$$= \frac{\ln\frac{\mu}{\nu}-\delta_5}{2\Delta}, \quad (114)$$

where the step of "$>$" uses $\exp(x) > 1+x$ for $x>0$ as well as $\ln\frac{\mu}{\nu}>0$ from $0<\nu<\mu$.

Applying Inequality (114) to Eq. (113), we obtain

$$\frac{1}{1-\left(1+\frac{\delta_5}{2\Delta}\right)\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}}$$
$$< \left(1+\frac{\frac{\delta_5}{2\Delta}}{\frac{\ln\frac{\mu}{\nu}-\delta_5}{2\Delta}}\right)\cdot\frac{1}{1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}}$$
$$= \left(1+\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}\right)\cdot\frac{1}{1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}}. \quad (115)$$

■

N. Proof of Lemma 7

We define

$$\lambda := \frac{\nu}{\mu} \quad (116)$$

and for $0<x\le 1$,

$$f(x) := \frac{x}{1-\lambda^x}. \quad (117)$$

Clearly, $0<\lambda<1$ follows from $0<\nu<\mu$. Then the derivative of $f(x)$ is

$$f'(x) = \frac{1-\lambda^x - x\cdot(-\ln\lambda)\lambda^x}{(1-\lambda^x)^2} = \frac{g(x)}{(1-\lambda^x)^2}, \quad (118)$$

where we define $g(x)$ as

$$g(x) := 1-(1-x\ln\lambda)\lambda^x. \quad (119)$$

To analyze the sign of $f'(x)$ in Eq. (118), we discuss the sign of $g(x)$ in Eq. (119). Hence, we compute the derivative of $g(x)$ as $g'(x) = (\ln\lambda)^2\lambda^x x > 0$ for $0<x\le 1$, given $0<\lambda<1$. Hence, $g(x)$ strictly increases as x increases for $0<x\le 1$, implying $g(x)>g(0)=0$ for $0<x\le 1$. Using this in Eq. (118), we have $f'(x)>0$ for $0<x\le 1$, so that $f(x)$ strictly increases as x increases for $0<x\le 1$. Then for any $\epsilon_4\in(0,\frac{1}{2\Delta})$, we have

$$f\left(\frac{1}{2\Delta}\right) > f(\epsilon_4), \quad (120)$$

and

$$f\left(\frac{1}{2\Delta}\right)-f(\epsilon_4) \le \left(\frac{1}{2\Delta}-\epsilon_4\right)\cdot\max_{x\in[\epsilon_4,\frac{1}{2\Delta}]} f'(x). \quad (121)$$

Letting $\epsilon_4\to 0$ in Inequality (120), we obtain

$$f\left(\frac{1}{2\Delta}\right) \ge \lim_{\epsilon_4\to 0} f(\epsilon_4) = \lim_{\epsilon_4\to 0}\frac{\epsilon_4}{1-\lambda^{\epsilon_4}}. \quad (122)$$

To compute $\lim_{\epsilon_4\to 0}\frac{\epsilon_4}{1-\lambda^{\epsilon_4}}$ of (122), we note that the nominator and denominator both converge to 0 as $\epsilon_4\to 0$, and are also both differentiable for $\epsilon_4>0$, so we use L'Hospital's rule (see [23]) to obtain

$$\lim_{\epsilon_4\to 0} f(\epsilon_4) = \lim_{\epsilon_4\to 0}\frac{\epsilon_4}{1-\lambda^{\epsilon_4}} = \lim_{\epsilon_4\to 0}\frac{1}{-\lambda^{\epsilon_4}\cdot\ln\lambda} = \frac{1}{\ln(1/\lambda)}, \quad (123)$$

which together with (122) means

$$f\left(\frac{1}{2\Delta}\right) \ge \frac{1}{\ln(1/\lambda)}. \quad (124)$$

To analyze Inequality (121), we now check the monotonicity of $f'(x)$. To this end, the second-order derivatives of $f(x)$ is

$$f''(x) = \frac{h(x)}{(1-\lambda^x)^3}. \quad (125)$$

where we define $h(x)$ as

$$h(x) := [x\ln\lambda(1+\lambda^x)+2(1-\lambda^x)]\lambda^x\ln\lambda. \quad (126)$$

To analyze the sign of $f''(x)$ in Eq. (125), we discuss the sign of $h(x)$ in Eq. (126). Hence, we compute the derivative of $h(x)$ as $h'(x) = \ln\lambda[1-(1-x\ln\lambda)\lambda^x] = \lambda g(x)$. Given $g(x)>0$ for $0<x\le 1$, we have $h'(x)>0$ for $0<x\le 1$. Hence, $h(x)$ strictly increases as x increases for $0<x\le 1$, implying $h(x)>h(0)=0$ for $0<x\le 1$. Using this in Eq. (125), we obtain $f''(x)>0$ for $0<x\le 1$, so that $f'(x)$ strictly increases as x increases for $0<x\le 1$. Thus, we know for any ϵ_4 satisfying $0<\epsilon_4<\frac{1}{2\Delta}\le\frac{1}{2}<1$ from $\Delta\ge 1$ that

$$\max_{x\in[\epsilon_4,\frac{1}{2\Delta}]} f'(x) < f'(1) = \frac{1-[1+\ln(1/\lambda)]\lambda}{(1-\lambda)^2}. \quad (127)$$

Now we bound the nominator $1-[1+\ln(1/\lambda)]\lambda$ in (127). For a lower bound, we use $\ln(1/\lambda)\le\lambda^{-1}-1$ given $0<\lambda<1$ to obtain $1-[1+\ln(1/\lambda)]\lambda\ge 0$. For an upper bound, we use $\ln(1/\lambda)\ge 1-\lambda$ given $0<\lambda<1$ to obtain $1-[1+\ln(1/\lambda)]\lambda\le 1-(2-\lambda)\lambda=(1-\lambda)^2$. These two bounds imply that $f'(1)$ in (127) satisfies $0\le f'(1)\le 1$. Then (127) gives

$$\max_{x\in[\epsilon_4,\frac{1}{2\Delta}]} f'(x) < 1.$$

which is used in Inequality (121) to induce

$$f\left(\frac{1}{2\Delta}\right) \le f(\epsilon_4)+\frac{1}{2\Delta}-\epsilon_4. \quad (128)$$

Letting $\epsilon_4\to 0$ in Inequality (128), we obtain

$$f\left(\frac{1}{2\Delta}\right) \le \lim_{\epsilon_4\to 0} f(\epsilon_4)+\lim_{\epsilon_4\to 0}\left(\frac{1}{2\Delta}-\epsilon_4\right)$$
$$= \frac{1}{\ln(1/\lambda)}+\frac{1}{2\Delta}, \quad (129)$$

where the last step uses Inequality (123).

Given $\lambda=\frac{\nu}{\mu}$ and $f\left(\frac{1}{2\Delta}\right)=\frac{1}{2\Delta\left[1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]}$ from Eq. (116) and Eq. (117), the combination of Inequalities (124) and (129) gives the desired result

$$\frac{2}{\ln(\mu/\nu)} \le \frac{1}{\Delta\left[1-\left(\frac{\nu}{\mu}\right)^{1/(2\Delta)}\right]} \le \frac{2}{\ln(\mu/\nu)}+\frac{1}{\Delta}.$$

■

O. Proof of Lemma 8

For $\delta_5=\frac{(\epsilon_1+\epsilon_2)\ln\frac{\mu}{\nu}}{(\epsilon_1+\epsilon_2)+(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)}$, the term $\frac{\delta_5}{\ln\frac{\mu}{\nu}-\delta_5}$ equals $\frac{(\epsilon_1+\epsilon_2)}{(1-\epsilon_1)\cdot(\ln\frac{\mu}{\nu}+1)}$. Given $0<\nu<\mu$, we have $\ln\frac{\mu}{\nu}>0$, which

with $0 < \epsilon_1 < 1$ and $\epsilon_2 > 0$ gives

$$1 + \frac{\epsilon_1 + \epsilon_2}{(1-\epsilon_1) \cdot (\ln \frac{\mu}{\nu} + 1)} < 1 + \frac{\epsilon_1 + \epsilon_2}{1 - \epsilon_1} = \frac{1 + \epsilon_2}{1 - \epsilon_1}.$$

Thus, Lemma 8 is proved. ■

P. Proof of Eq. (71)

$$\begin{aligned}
&\mathbb{P}\left[\boldsymbol{F}_{t-\Delta-1} S_{t-\Delta} \dots S_t = \boldsymbol{f}_{t-\Delta-1} s_{t-\Delta} \dots s_t\right] \\
&= \sum_{\substack{\boldsymbol{f}_{t-\Delta-2} \in \text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-\Delta-2} || s_{t-\Delta-1}) = \boldsymbol{f}_{t-\Delta-1}}} \\
&\quad \mathbb{P}\left[\begin{array}{l}\left(\boldsymbol{F}_{t-\Delta-2} S_{t-\Delta-1} \dots S_{t-1} = \boldsymbol{f}_{t-\Delta-2} s_{t-\Delta-1} \dots s_{t-1}\right) \\ \wedge \left(S_t = s_t\right)\end{array}\right] \\
&= \mathbb{P}\left[S_t = s_t\right] \sum_{\substack{\boldsymbol{f}_{t-\Delta-2} \in \text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-\Delta-2} || s_{t-\Delta-1}) = \boldsymbol{f}_{t-\Delta-1}}} \\
&\quad \mathbb{P}\left[\boldsymbol{F}_{t-\Delta-2} S_{t-\Delta-1} \dots S_{t-1} = \boldsymbol{f}_{t-\Delta-2} s_{t-\Delta-1} \dots s_{t-1}\right].
\end{aligned} \tag{130}$$

$$\begin{aligned}
&\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-1} s_{t-\Delta} \dots s_t) \\
&= \mathbb{P}\left[S_t = s_t\right] \times \\
&\quad \sum_{\substack{\boldsymbol{f}_{t-\Delta-2} \in \text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-\Delta-2} || s_{t-\Delta-1}) = \boldsymbol{f}_{t-\Delta-1}}} \pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-2} s_{t-\Delta-1} \dots s_{t-1}).
\end{aligned} \tag{131}$$

$$\begin{aligned}
&\frac{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-1} s_{t-\Delta} \dots s_t)}{\prod_{i=t-\Delta}^{t} \mathbb{P}\left[S_i = s_i\right]} \\
&= \sum_{\substack{\boldsymbol{f}_{t-\Delta-2} \in \text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-\Delta-2} || s_{t-\Delta-1}) = \boldsymbol{f}_{t-\Delta-1}}} \\
&\quad \left(\frac{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-2} s_{t-\Delta-1} \dots s_{t-1})}{\prod_{i=t-\Delta-1}^{t-1} \mathbb{P}\left[S_i = s_i\right]} \cdot \mathbb{P}\left[S_{t-\Delta-1} = s_{t-\Delta-1}\right]\right).
\end{aligned} \tag{132}$$

Using Eq. (81) and replacing t therein by $t - \Delta - 1$, we have

$$\begin{aligned}
&\mathbb{P}\left[\boldsymbol{F}_{t-\Delta-1} = \boldsymbol{f}_{t-\Delta-1}\right] \\
&= \sum_{\substack{\boldsymbol{f}_{t-\Delta-2} \in \text{Suffix-Set:} \\ \text{suffix}(\boldsymbol{f}_{t-\Delta-2} || s_{t-\Delta-1}) = \boldsymbol{f}_{t-\Delta-1}}} \\
&\quad \left(\mathbb{P}\left[\boldsymbol{F}_{t-\Delta-2} = \boldsymbol{f}_{t-\Delta-2}\right] \mathbb{P}\left[S_{t-\Delta-1} = s_{t-\Delta-1}\right]\right).
\end{aligned} \tag{133}$$

From Eq. (132) and (133), the transition from $\frac{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-2} s_{t-\Delta-1} \dots s_{t-1})}{\prod_{i=t-\Delta-1}^{t-1} \mathbb{P}[S_i = s_i]}$ to $\frac{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-1} s_{t-\Delta} \dots s_t)}{\prod_{i=t-\Delta}^{t} \mathbb{P}[S_i = s_i]}$ has the same rule as the transition from $\pi_{\boldsymbol{F}}(\boldsymbol{f}_{t-\Delta-2})$ to $\pi_{\boldsymbol{F}}(\boldsymbol{f}_{t-\Delta-1})$, so we can conclude $\frac{\pi_{\boldsymbol{F}||\boldsymbol{P}}(\boldsymbol{f}_{t-\Delta-1} s_{t-\Delta} \dots s_t)}{\prod_{i=t-\Delta}^{t} \mathbb{P}[S_i = s_i]} = \pi_{\boldsymbol{F}}(\boldsymbol{f}_{t-\Delta-1})$, which is exactly the desired result Eq. (71). ■