The Combinatorics of the Longest-Chain Rule: Linear Consistency for Proof-of-Stake Blockchains

Erica Blum[1], Aggelos Kiayias[2,5], Cristopher Moore[3], Saad Quader[4], and Alexander Russell[4,5]

[1]University of Maryland, College Park
[2]University of Edinburgh
[3]Santa Fe Institute
[4]University of Connecticut
[5]IOHK

August 31, 2019

Abstract

Blockchain data structures maintained via the longest-chain rule have emerged as a powerful algorithmic tool for consensus algorithms. The technique—popularized by the Bitcoin protocol—has proven to be remarkably flexible and now supports consensus algorithms in a wide variety of settings. Despite such broad applicability and adoption, current analytic understanding of the technique is highly dependent on details of the protocol's leader election scheme. A particular challenge appears in the proof-of-stake setting, where existing analyses suffer from quadratic dependence on suffix length.

We describe an axiomatic theory of blockchain dynamics that permits rigorous reasoning about the longest-chain rule in quite general circumstances and establish bounds—optimal to within a constant—on the probability of a consistency violation. This settles a critical open question in the proof-of-stake setting where we achieve linear consistency for the first time.

Operationally, blockchain consensus protocols achieve consistency by instructing parties to remove a suffix of a certain length from their local blockchain. While the analysis of Bitcoin guarantees consistency with error 2^{-k} by removing $O(k)$ blocks, recent work on proof-of-stake (PoS) blockchains has suffered from quadratic dependence: (PoS) blockchain protocols, exemplified by Ouroboros (Crypto 2017), Ouroboros Praos (Eurocrypt 2018) and Sleepy Consensus (Asiacrypt 2017), can only establish that the length of this suffix should be $\Theta(k^2)$. This consistency guarantee is a fundamental design parameter for these systems, as the length of the suffix is a lower bound for the time required to wait for transactions to settle. Whether this gap is an intrinsic limitation of PoS—due to issues such as the "nothing-at-stake" problem—has been an urgent open question as deployed PoS blockchains further rely on consistency for protocol correctness: in particular, security of the protocol itself relies on this parameter. Our general theory directly improves the required suffix length from $\Theta(k^2)$ to $\Theta(k)$. Thus we show, for the first time, how PoS protocols can match proof-of-work blockchain protocols for exponentially decreasing consistency error.

Our analysis focuses on the articulation of a two-dimensional stochastic process that captures the features of interest, an exact recursive closed form for the critical functional of the process, and tail bounds established for associated generating functions that dominate the failure events. Finally, the analysis provides an explicit polynomial-time algorithm for exactly computing the exponentially-decaying error function which can directly inform practice.

1 Introduction

A blockchain is a data structure consisting of a collection of data blocks placed in linear order. It further requires that each block contains a collision-free hash of the previous block: thus blocks implicitly commit to the entire prefix of the blockchain preceding them. This elementary data structure has remarkable applications in distributed computing, and now appears as an essential component of consensus protocols in a wide variety of models and settings; this notably includes both the "permissionless" setting popularized by Bitcoin and the classic "permissioned" model.

Such consensus protocols call for players to collaboratively assemble a blockchain by repeatedly selecting players to add blocks. Specifically, the protocol determines a stochastic process resembling a lottery: each "leader" selected by the lottery is then responsible for broadcasting a new block. While the algorithmic details of this lottery depend heavily on the protocol, the outcome can be privately determined and provides the winning player a proof of leadership that can be publicly demonstrated. Assuming that the expected wait time for some player to win the lottery is constant, the blockchain experiences steady growth when players follow the protocol.

Network infelicities, adversarial behavior, or the possibility that two players simultaneously win the lottery can lead to disagreements among the players about the current blockchain. Thus protocols adopt a "chain selection rule" that determines how players should break ties among the various chains they observe on the network; ideally, the combination of the chain selection rule and the lottery should guarantee that the players' blockchains agree, perhaps with the exception of a short suffix. The emblematic chain selection strategy among such systems is the *longest-chain rule*, which calls for players to adopt the longest chain among various contenders.

The first blockchain protocol was the core of the sensational Bitcoin system [18]; it adopted a lottery mechanism based on a cryptographic puzzle [7, 1]—also known as *proof-of-work or PoW, for short*—and a chain selection rule favoring chains that represent more work. The system is particularly notable for its ability to survive in a permissionless setting—where players may freely join and depart—even when a portion of the players are actively attacking the protocol. Unfortunately, the proof-of-work mechanism makes quite striking energy demands: the system currently consumes as much electricity as a small country.[1] This motivated the blockchain community to exploring alternative lottery mechanisms, e.g., proof-of-stake (PoS) [3, 21, 13], proof of space [8, 20] and others [16]. The proof-of-stake mechanism is particularly attractive from the perspective of efficiency, as it makes no assumption of external computational resources.

The fundamental consistency property—critical in all these blockchain systems—is *common-prefix* (cf. [9]). It precisely captures the intuition described above: by trimming a k-block suffix from the chain held by any honest player the resulting blockchain is a prefix of the blockchain possessed by any honest party at any future point of the execution. A principal goal in the analysis of these systems is a to guarantee common prefix, for an appropriate value of k, even if some of the players collude to disrupt the protocol. Common prefix is typically only shown to hold with high probability $1 - \varepsilon$, where ε is an error term that is a function of k. The exact dependency of ε on k is critically important: it determines the length of the suffix that is to be removed from a blockchain in order to ensure that the remaining prefix will be retained at any future point of the execution. This directly imposes a lower bound on how long one has to wait for information in the blockchain (such as a payment transaction) to "settle." Additionally, many blockchain protocols internally rely on common prefix for correctness; thus the relationship between ε and k is critical to establishing the regime of correctness of the entire protocol.

A relatively straightforward lower bound for ε is $\varepsilon \geq \exp(-\alpha k)$ for some $\alpha > 0$. This lower bound applies when there is a coalition of adversarial players of constant fraction, the case of primary interest in practice. The result is easy to infer from the analysis of [18], where a strategy is demonstrated that violates common prefix with such probability (this is referred to as a "double-spending" attack in that paper). The tightness of this bound is an important open problem. For the special case of proof-of-work an upper bound of $\exp(-\Omega(k))$ was shown first in [9] and further verified in extended security models by [11, 24]. In the proof-of-stake setting, on the other hand, the tightness of the bound remains open. While recent proof-of-stake algorithms have been presented with rigorous analyses that rival proof-of-work in many regards, they suffer from a quadratic relationship between k and $\log(\varepsilon)$. For example, the Ouroboros protocols [13, 6, 2], as well as Snow White [4], provide an upper bound on ε of $\exp(-\Omega(\sqrt{k}))$; this should be compared with $\varepsilon = \exp(-\Theta(k))$ for proof-of-work. The significant gap from the known lower bound was attributed to a notable, general attack that distinguished PoS from PoW: Known as the *nothing-at-stake* problem, this refers to the ability of an adversarial coalition of players to strategically reuse a winning PoS lottery to extend multiple blockchains.

Our results. Our objective is to control the common-prefix error ε as tightly as possible while making minimal assumptions on the underlying blockchain protocol. We work in a general model formulated by a simple family of

[1] See e.g., `https://digiconomist.net/bitcoin-energy-consumption` where it is reported that Bitcoin annual energy consumption is on the order of at least 50 Twhr at the time of writing.

blockchain axioms. The axioms themselves are easy to interpret and few in number. This permits us to abstract many features of the underlying blockchain protocol (e.g., the details of the leader-election process, the cryptographic security of the relevant signature schemes and hash functions, and randomness generation), while still establishing results that are strong enough to directly incorporate into existing specific analyses.

Our most interesting finding is a quite tight theory of common prefix that depends *only on the schedule of participants certified to add a block*. Under common assumptions about this schedule, we achieve the optimal relationship $\varepsilon = \exp(-\Theta(k))$. This directly improves the common prefix guarantees (and settlement times) of existing proof-of-stake blockchains such as Snow White [4], Ouroboros [13], Ouroboros Praos [6], and Ouroboros Genesis [2]. Specifically, this improves the scaling in the exponent from $\sqrt{k}$ to k and establishes a tight characterization for $\varepsilon = \exp(-\Theta(k))$. (In fact, we even obtain reasonable control of the constants.) We remark that our assumptions about the schedule distribution can be weakened—without any effect on the final bounds—to apply to martingale-style distributions such as those that arise in the analysis of adaptive adversaries [6, 2].

Our new analysis offers an additional, but lower order, improvement for several of these blockchains. The existing analysis of, e.g., Ouroboros Praos [6], required a union bound to be taken over the entire lifetime of the protocol in order to rule out a common prefix violation at a particular point of time; thus such events were actually bounded above by a function of the form $T\exp(-\Omega(\sqrt{k}))$, where T is the lifetime of the protocol. While this event *does* depend on the entire dynamics of the protocol, we show how to avoid this pessimistic tail bound to achieve a "single shot" common prefix violation—at a particular time of interest—of form $\exp(-\Theta(k))$; this removes the dependence on T.

From a technical perspective, we contrast the structure of our proofs with existing techniques for the PoW case. The PoW results find a direct connection between common-prefix and the behavior of a biased, one-dimensional random walk. Interestingly, our results give a tight relationship between the general (e.g., PoS) case and a pair of *coupled* biased random walks. A major challenge in the analysis is to bound the behavior of this richer stochastic process. Our tools yield precise, explicit upper bounds on the probability of persistence violations that can be directly applied to tune the parameters of deployed PoS systems. See Appendix A where we record some concrete results of the general theory. The importance of these results in the practice of PoS blockchain systems cannot be understated: they provide, for the first time, concrete error bounds for settlement times for PoS blockchains that follow the longest chain rule.

Further analytic details. Our approach begins with the graph-theoretic framework of *forks* and *margin* developed for the analysis of the Ouroboros [13] protocol. We begin by generalizing the notion of margin to account for local, rather than global, features of a leader schedule, and provide an exact, recursive closed form for this new quantity. Curiously, this proof identifies an on-line—and optimal—adversary for PoS blockchain algorithms when divergence is measured in slots. This new adversary has the notable property of simultaneously achieving worst-case (slot) common-prefix violations over all starting points (see Section 5). We then study the stochastic process generated when the *characteristic string*—a Boolean string representing the outcome of the leader election scheme—is given by a family of i.i.d. Bernoulli random variables. In this case, we identify a generating function that bounds the tail events off interest, and analytically upper bound the growth of the function. We then show how to extend the analysis to the setting where the characteristic string is drawn from a martingale sequence. As it happens, this more general distribution arises naturally in the analyses of PoS protocols that survive adaptive adversaries; e.g., Ouroboros Genesis [2]. We obtain the pleasing result that the common prefix error probability in the martingale case is no more than that of the i.i.d. Bernoulli case.

Direct consequences. Our results establish consistency bounds in a quite general setting—see below: In particular, they directly imply $\exp(-\Theta(k))$ consistency for the Sleepy consensus (Snow White) [21], Ouroboros [13], Ouroboros Praos [6], and Ouroboros Genesis [2] blockchain protocols. (The Ouroboros Praos and Ouroboros Genesis analyses in fact directly relied on an earlier e-print version of the present article for their settlement estimates.)

Related work. Blockchain protocol analysis in the PoW-setting was initiated in [9] and further improved in [24, 11]. The established security bounds for consistency are linear in the security parameter. Sleepy consensus [21, Theorem 13] provides a consistency bound of the form $\exp(-\Omega(\sqrt{k}))$. Note that [21] is not a PoS protocol per se, but it is possible to turn it into one (as was demonstrated in [4]). The analysis of the Ouroboros blockchain [13] achieves $\exp(-\Omega(\sqrt{k}))$. We remark that the analyses of Ouroboros Praos [6] and Ouroboros Genesis [2] developed significant new machinery for handling other challenges (e.g., adaptive adversaries, partial synchrony), but directly referred to a preliminary version of this article to conclude their guarantees of $\exp(-\Omega(k))$.

Our results complement the recent results of [5], which also considers longest-chain PoS protocols. [5] focuses on identifying dynamics unique to longest-chain PoS protocols. In particular, they show that longest-chain PoS protocols that are *predictable* (i.e., for which some portion of the schedule of slot leaders is known ahead of time) are necessarily vulnerable to "predictable double-spends." The conventional defense against such attacks is to wait for the specified settlement time to elapse before accepting a transaction, which (until now) has resulted in slow confirmation times. As such, [5] raised the question of whether long confirmation times are a necessary evil in longest-chain PoS blockchains. As double-spending attacks imply a consistency violation, our results show that PoS protocols can safely decrease settlement times to asymptotically match PoW protocols without sacrificing security against double-spends.

Because we focus on the longest-chain rule, our analysis is not applicable to protocols like Algorand [15] which, in fact, offer settlement in expected constant time without invoking blockchain reorganisation or forks; however, Algorand lacks the ability to operate in the "sleepy" [21] or "dynamic availability" [2] setting. In our combinatorial analysis, synchronous operation is assumed against a rushing adversary; this is without loss of generality vis-a-vis the result of [6] where it was shown how to reduce the combinatorial analysis in the partially synchronous setting to the synchronous one. We note that a number of works have shown how to use a blockchain protocol to bootstrap a cryptographic protocol that can offer faster settlement time under stronger assumptions than honest majority, e.g., Hybrid Consensus [22] or Thunderella [23]; our results are orthogonal and synergistic to those since they can be used to improve the settlement time bounds of the blockchain protocol that operates as a fallback mechanism.

Outline. We begin in Section 2 by describing a simple general model for blockchain dynamics. Section 3 builds on this model to set down a number of basic definitions required for the proofs. The first part of the main proof is described in Section 5, which develops a "relative" version of the theory of margin from [13]; most details are then relegated to Section 7 in order to move quickly to the consistency estimates in Section 6. Finally, in Appendix A, we compute exact upper bounds on settlement probabilities for various values of k. We also outline a simple $O(k^3)$-time algorithm to compute these probabilities.

2 The blockchain axioms and the settlement security model

Typical blockchain consensus protocols call for each participant to maintain a *blockchain*; this is a data structure that organizes transactions and other protocol metadata into an ordered historical record of "blocks." A basic design goal of these systems is to guarantee that participants' blockchains always agree on a common prefix; the differing suffixes of these chains held by various participants roughly correspond to the possible future states of the system. Thus the major analytic challenge is to ensure that—despite evolving adversarial control of some of the participants—the portion of honest participants' blockchains that might pairwise disagree is confined to a short suffix. This analysis in turn supports the fundamental guarantee of *consistency* for these algorithms, which asserts that data appearing deep enough in the chain can be considered to be stable, or "settled."

We adopt a discrete notion of time organized into a sequence of *slots* $\{sl_0, sl_1, \ldots\}$ and assume all protocol participants have the luxury of synchronized clocks that report the current slot number. As discussed above, the protocols we consider rely on two algorithmic devices:

- A *leader election mechanism*, which randomly assigns to each time slot a set of "leaders" permitted to post a new block in that slot.
- The *longest-chain rule*, which calls for the leader(s) of each slot to add a block to the end of the longest blockchain she has yet observed, and broadcast this new chain to other participants.

The Bitcoin protocol uses a proof-of-work mechanism to carry out leader election, which can be modeled using a random oracle [9, 24, 11]. Proof-of-stake systems typically require more intricate leader election mechanisms; for example, the Ouroboros protocol [13] uses a full multi-party private computation to distribute clean randomness, while Snow White [4], Algorand [15], and Ouroboros Praos [6] use hashing and a family of values determined on-the-fly.

Despite these differences, all existing analyses show that the leader election mechanism suitably approximates an ideal distribution, which is also the approach we will adopt for our analysis.

2.1 The blockchain axioms and forks

To simplify our analysis, we assume a synchronous communication network in the presence of a *rushing* adversary: in particular, any message broadcast by an honest participant at the beginning of a particular slot is received by the adversary first, who may decide strategically and individually for each recipient in the network whether to inject additional messages and in what order all messages are to be delivered prior to the conclusion of the slot. (See §2.5 below for comments on this network assumption.)

Given this, the behavior of the protocol when carried out by a group of honest participants (who follow the protocol in the presence of an adversary who may only reorganize messages) is clear. Assuming that the system is initialized with a common "genesis block" corresponding to sl_0 and the leader election process in fact elects a single leader per slot, the players observe a common, linearly growing blockchain:

$$0 \to 1 \to 2 \to \cdots$$

Here node i represents the block broadcast by the leader of slot i and the arrows represent the direction of increasing time. (Note that the requirement of a single leader per slot is important in this simple picture; it is possible for a network adversary to induce divergent views between the players by taking advantage of slots where more than a single honest participant is elected a leader.)

The blockchain axioms: Informal discussion. The introduction of adversarial participants or multiple slot leaders complicates the family of possible blockchains that could emerge from this process. To explore this in the context of our protocols, we work with an abstract notion of a blockchain which (as informally suggested above) ignores all internal structure. We consider a fixed assignment of leaders to time slots, and assume that the blockchain uses a proof mechanism to ensure that any block labeled with slot sl_t was indeed produced by a leader of slot sl_t; this is guaranteed in practice by appropriate use of a secure digital signature scheme.

Specifically, we treat a *blockchain* as a sequence of abstract blocks, each labeled with a slot number, so that:

A1. The blockchain begins with a fixed "genesis" block, assigned to slot sl_0.

A2. The (slot) labels of the blocks are in strictly increasing order.

It is further convenient to introduce the structure of a directed graph on our presentation, where each block is treated as a vertex; in light of the first two axioms above, a blockchain is a path beginning with a special "genesis" vertex, labeled 0, followed by vertices with strictly increasing labels that indicate which slot is associated with the block. (See the example below.)

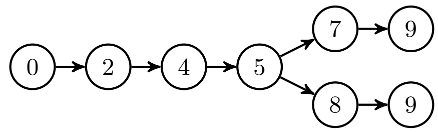
$$0 \to 2 \to 4 \to 5 \to 7 \to 9$$

The protocols of interest call for honest players to add a *single* block (to a single previous chain in its local state) during any slot. In particular:

A3. If a slot sl_t was assigned to a single honest player, then a single block is created—during the entire protocol—with the label sl_t.

Recall that blockchains are *immutable* in the sense that any block in the chain commits to the entire previous history of the chain; this is achieved in practice by including with each block a collision-free hash of the previous block. These properties imply that if a specific slot sl_t was assigned to a unique honest player, then any chain that includes the unique block from sl_t must also include that block's associated prefix in its entirety.

As we analyze the dynamics of blockchain algorithms, it is convenient to maintain an entire family of blockchains at once. As a matter of bookkeeping, when two blockchains agree on a common prefix, we can glue together the associated paths to reflect this, as indicated below.



When we glue together many chains to form such a diagram, we call it a "fork"—the precise definition appears below. Observe that while these two blockchains agree through the vertex (block) labeled 5, they contain (distinct) vertices labeled 9; this reflects two distinct blocks associated with slot 9.

Finally, in light of the fact that messages from honest players are delivered immediately, we note a direct consequence of the longest chain rule:

A4. If two honestly generated blocks B_1 and B_2 are labeled with slots sl_1 and sl_2 for which $sl_1 < sl_2$, then the length of the unique blockchain terminating at B_1 is strictly less than the length of the unique blockchain terminating at B_2.

Recall that the honest participant assigned to slot sl_2 will be aware of the blockchain terminating at B_1 that was broadcast by the honest player in slot sl_1 as a result of synchronicity; according to the longest-chain rule, it must have placed B_2 on a chain that was at least this long. In contrast, not all participants are necessarily aware of all blocks generated by dishonest players, and indeed dishonest players may often want to delay the delivery of an adversarial block to a participant or show one block to some participants and show a completely different block to others.

Characteristic strings, forks, and the formal axioms. Note that with the axioms we have discussed above, whether or not a particular fork diagram (such as the one just above) corresponds to a valid execution of the protocol depends on how the slots have been awarded to the parties by the leader election mechanism. We introduce the notion of a "characteristic" string as a convenient means of representing information about slot leaders in a given execution.

Definition 1 (Characteristic string). *Let $sl_1, \ldots, sl_n$ be a sequence of slots. A* characteristic string *w is an element of $\{0,1\}^n$ defined for a particular execution of a blockchain protocol so that*

$$w_t = \begin{cases} 0 & \textit{if } sl_t \textit{ was assigned to a single honest participant,} \\ 1 & \textit{otherwise.} \end{cases}$$

With this discussion behind us, we set down the formal object we use to reflect the various blockchains adopted by honest players during the execution of a blockchain protocol. This definition formalizes the blockchains axioms discussed above.

Definition 2 (Fork; [13]). *Let $w \in \{0,1\}^n$ and let $H = \{i \mid w_i = 0\}$. A* fork *for the string w consists of a directed and rooted tree $F = (V, E)$ and a labeling $\ell : V \to \{0, 1, \ldots, n\}$. We insist that each edge of F is directed away from the root vertex and further require that*

(F1.) the root vertex r has label $\ell(r) = 0$;

(F2.) the labels of vertices along any directed path are strictly increasing;

(F3.) each index $i \in H$ is the label for exactly one vertex of F;

(F4.) for any vertices $i, j \in H$, if $i < j$, then the depth of vertex i in F is strictly less than the depth of vertex j in F.

If F is a fork for the characteristic string w, we write $F \vdash w$. Note that the conditions (F1.)–(F4.) are direct analogues of the axioms A1–A4 above. See Fig. 1 for an example fork. A final notational convention: If $F \vdash x$ and $\hat{F} \vdash w$, we say that F is a *prefix* of $\hat{F}$, written $F \sqsubseteq \hat{F}$, if the string $x \in \{0,1\}^\ell$ is a prefix of the string $w \in \{0,1\}^{\ell+m}$ and F appears as a consistently-labeled subgraph of $\hat{F}$. (Specifically, each path of F appears, with identical labels, in $\hat{F}$.)

Let w be a characteristic string. The directed paths in the fork $F \vdash w$ originating from the root are called *tines*; these are abstract representations of blockchains. (Note that a tine might not terminate at a leaf of the fork.) We naturally extend the label function ℓ for tines: i.e., $\ell(t) \triangleq \ell(v)$ where the tine t terminates at vertex v. The length of a tine t is denoted by $\mathrm{length}(t)$.

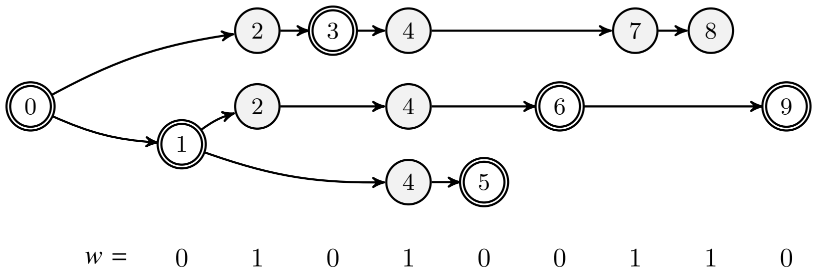


Figure 1: A fork F for the characteristic string $w = 010100110$; vertices appear with their labels and honest vertices are highlighted with double borders. Note that the depths of the (honest) vertices associated with the honest indices of w are strictly increasing. Note, also, that this fork has two disjoint paths of maximum depth.

Viable tines. The longest-chain rule dictates that honest players build on chains that are at least as long as all previously broadcast honest chains. It is convenient to distinguish such tines in the analysis: specifically, a tine t of F is called *viable* if its length is at least the depth of any honest vertex v for which $\ell(v) \leq \ell(t)$. A tine t is *viable at slot s* if the portion of t appearing over slots $0, \ldots, s$ has length at least that of any honest vertices labeled from this set. (As noted, the properties (F3.) and (F4.) together imply that an honest observer at slot s will only adopt a viable tine.) The *honest depth* function $\mathbf{d} : H \to [n]$ gives the depth of the (unique) vertex associated with an honest slot; by (F4.), $\mathbf{d}(\cdot)$ is strictly increasing.

2.2 Settlement and the common prefix property

We are now ready to explore the power of an adversary in this setting who has corrupted a (perhaps evolving) coalition of the players. We focus on the possibility that such an adversary can blatantly confound consistency of the honest player's blockchains. In particular, we consider the possibility that, at some time t, the adversary conspires to produce two blockchains of maximum length that diverge prior to a previous slot $s < t$; in this case honest players adopting the longest-chain rule may clearly disagree about the history of the blockchain after slot s. We call such a circumstance a *settlement violation*.

To reflect this in our abstract language, let $F \vdash w$ be a fork corresponding to an execution with characteristic string w. Such a settlement violation induces two viable tines t_1, t_2 with the same length that diverge prior to a particular slot of interest. We record this below.

Definition 3 (Settlement with parameters $s, k \in \mathbb{N}$)**.** *Let $w \in \{0,1\}^n$ be a characteristic string. We say that a slot s is* not k-settled *in a fork $F \vdash w_1 \ldots w_t$ if $s + k \leq t \leq n$ and F contains two tines t_1, t_2 of maximum length that "diverge prior to s," i.e., they either contain different vertices labeled with s, or one contains a vertex labeled with s while the other does not. Note that such tines are viable by definition. Otherwise,* slot s is k-settled in F. *We say that a slot s is* k-settled *(for the characteristic string w) if it is k-settled in every fork $F \vdash w_1, \ldots w_t$, for each $t \geq s + k$.*

Common prefix. Settlement violations are a convenient and intuitive proxy for the notion of common prefix discussed in the introduction. Indeed, as we show in Section 4, the two notions are equivalent, so we have the luxury of discussing settlement violations which have the advantage of a more ready interpretation. Concretely, we will simultaneously upper bound—using the same analytic techniques—the probability of settlement violations and common prefix violations.

Recall that the common prefix property with parameter k asserts that for any slot index s, if an honest observer at slot $s + k$ adopts a blockchain C, the prefix $C[0 : s]$ will be present in every honestly-held blockchain at or after slot $s + k$. (Here, $C[0 : s]$ denotes the prefix of the blockchain C containing only the blocks issued from slots $0, 1, \ldots, s$.)

We translate this property into the framework of forks. Consider a tine t of a fork $F \vdash w$. The *trimmed* tine $t^{\lceil k}$ is defined as the portion of t labeled with slots $\{0, \ldots, \ell(t) - k\}$. For two tines, we use the notation $t_1 \preceq t_2$ to indicate that the tine t_1 is a prefix of tine t_2.

Definition 4 (Common Prefix Property with parameter $k \in \mathbb{N}$)**.** *Let w be a characteristic string. A fork $F \vdash w$ satisfies k-$\mathrm{CP}^{\mathsf{slot}}$ if, for all pairs (t_1, t_2) of viable tines F for which $\ell(t_1) \le \ell(t_2)$, we have $t_1^{\lceil k} \preceq t_2$. Otherwise, we say that the tine-pair (t_1, t_2) is a witness to a k-$\mathrm{CP}^{\mathsf{slot}}$ violation. Finally, w* satisfies *k-$\mathrm{CP}^{\mathsf{slot}}$ if every fork $F \vdash w$ satisfies k-$\mathrm{CP}^{\mathsf{slot}}$.*

If a string w does not possess the k-$\mathrm{CP}^{\mathsf{slot}}$ property, we say that *w violates* k-$\mathrm{CP}^{\mathsf{slot}}$. Observe that we defined the common prefix property in terms of deleting any blocks associated with the *last k trailing slots* from a local blockchain C. Traditionally (cf. [10]), this property has been defined in terms of deleting a suffix of (block-)length k from C. We denote the block-deletion-based version of the common prefix property as the k-CP property. Note, however, that a k-CP violation immediately implies a k-$\mathrm{CP}^{\mathsf{slot}}$ violation, so bounding the probability of a k-$\mathrm{CP}^{\mathsf{slot}}$ violation is sufficient to rule out both events.

2.3 Adversarial attacks on settlement time; the settlement game

To clarify the relationship between forks and the chains at play in a canonical blockchain protocol, we define a game-based model below that explicitly describes the relationship between forks and executions. By design, the probability that the adversary wins this game is at most the probability that a slot s is not k-settled. We remark that while we focus on settlement violations for clarity, one could equally well have designed the game around common prefix violations.

Consider the $(\mathcal{D}, T; s, k)$-*settlement game*, played between an adversary $\mathcal{A}$ and a challenger C with a leader election mechanism modeled by an ideal distribution $\mathcal{D}$. Intuitively, the game should reflect the ability of the adversary to achieve a settlement violation; that is, to present two maximally-long viable blockchains to a future honest observer, thus forcing them to choose between two alternate histories which disagree on slot s. The challenger plays the role(s) of the honest players during the protocol.

Note that in typical PoS settings the distribution $\mathcal{D}$ is determined by the combined stake held by the adversarial players, the leader election mechanism, and the dynamics of the protocol. The most common case (as seen in Snow White [21] and Ouroboros [13]) guarantees that the characteristic string $w = w_1 \dots w_T$ is drawn from an i.i.d. distribution for which $\Pr[w_i = 1] \le (1-\epsilon)/2$; here the constant $(1-\epsilon)/2$ is directly related to the stake held by the adversary. Settings involving adaptive adversaries (e.g., Ouroboros Praos [6] and Ouroboros Genesis [2]) yield the weaker martingale-type guarantee that $\Pr[w_i = 1 \mid w_1, \dots, w_{i-1}] \le (1-\epsilon)/2$.

The $(\mathcal{D}, T; s, k)$-settlement game

1. A characteristic string $w \in \{0,1\}^T$ is drawn from $\mathcal{D}$. (This reflects the results of the leader election mechanism.)
2. Let $A_0 \vdash \varepsilon$ denote the initial fork for the empty string ε consisting of a single node corresponding to the genesis block.
3. For each slot $t = 1, \dots, T$ in increasing order:
   (a) If $w_t = 0$, this is an honest slot. In this case, the challenger is given the fork $A_{t-1} \vdash w_1 \dots w_{t-1}$ and must determine a new fork $F_t \vdash w_1 \dots w_t$ by adding a single vertex (labeled with t) to the end of a longest path in A_{t-1}. (If there are ties, the adversary may choose which path the challenger adopts.)
   (b) If $w_t = 1$, this is an adversarial slot. $\mathcal{A}$ may set $F_t \vdash w_1 \dots w_t$ to be an arbitrary fork for which $A_{t-1} \sqsubseteq F_t$.
   (c) (Adversarial augmentation.) $\mathcal{A}$ determines an arbitrary fork $A_t \vdash w_1 \dots, w_t$ for which $F_t \sqsubseteq A_t$.

   Recall that $F \sqsubseteq F'$ indicates that F' contains, as a consistently-labeled subgraph, the fork F.

$\mathcal{A}$ wins the settlement game if slot s is not k-settled in some fork A_t (with $t \ge s + k$).

Definition 5. *Let $\mathcal{D}$ be a distribution on $\{0,1\}^T$. Then define the (s,k)*-settlement insecurity *of $\mathcal{D}$ to be*

$$\mathbf{S}^{s,k}[\mathcal{D}] \triangleq \max_{\mathcal{A}} \Pr[\mathcal{A} \textit{ wins the } (\mathcal{D}, T; s, k)\textit{-settlement game}],$$

this maximum taken over all adversaries $\mathcal{A}$.

Remarks. Observe that the adversarial augmentation step permits the adversary to "suddenly" inject new paths in the fork between two honest players at adjacent slots; this corresponds to circumstances when the adversary chooses to deliver a new blockchain to an honest participant which may consist of an earlier honest chain with some adversarial blocks appended to the end. Observe, additionally, that the behavior of the challenger in the game is entirely deterministic, as it simply plays according to the longest-chain rule (even permitting the adversary to break ties). Thus the result of the game is entirely determined by the characteristic string w drawn from $\mathcal{D}$ and the choices of the adversary $\mathcal{A}$. We record the following immediate conclusion:

Lemma 1. *Let $s, k, T \in \mathbb{N}$. Let $\mathcal{D}$ be a distribution on $\{0,1\}^T$. Then*

$$\mathbf{S}^{s,k}[\mathcal{D}] \leq \Pr_{w \sim \mathcal{D}}[\textit{slot } s \textit{ is not } k\textit{-settled for } w] .$$

In the subsequent sections, we will develop some further notation and tools to analyze this event. We will investigate two different families of distributions, those with i.i.d. coordinates and those with martingale-type conditioning guarantees. For $T \in \mathbb{N}$ and $\epsilon \in (0,1)$, let $B_\epsilon = (B_1, \ldots, B_n)$ denote the random variable taking values in $\{0,1\}^T$ so that the B_i are independent and $\Pr[B_i = 1] = (1-\epsilon)/2$; we let $\mathcal{B}_\epsilon$ denote the distribution on $\{0,1\}^T$ associated with B_ϵ. When ϵ can be inferred from context, we simply write B and $\mathcal{B}$.

We also study a more general family of distributions, defined next.

Definition 6 (ϵ-martingale condition)**.** *Let $W = (W_1, \ldots, W_n)$ be a random variable taking values in $\{0,1\}^n$. We say that W satisfies the* ϵ-martingale condition *if for each $t \in \{1, \ldots, n\}$,*

$$\mathbb{E}[W_t \mid W_1, \cdots, W_{t-1}] \leq (1-\epsilon)/2 .$$

Equivalently, $\Pr[W_t = 1 \mid W_1, \ldots, W_{t-1}] \leq (1-\epsilon)/2$. The conditioning on the variables $W_1, \cdots, W_{t-1}$ is arbitrary in both cases; as a consequence, $\Pr[W_t = 1] \leq (1-\epsilon)/2$. As a matter of notation, we let $\mathcal{W}$ denote the distribution associated with the random variable W. We use the term "ϵ-martingale condition" to qualify both a random variable and its distribution.

There are settings, such as Genesis [2], where this martingale-type conditioning is important. Note that $\mathcal{B}_\epsilon$ satisfies the ϵ-martingale condition. Now we are ready to state our main theorem.

Theorem 1 (Main theorem)**.** *Let $\epsilon \in (0,1), s, k, T \in \mathbb{N}$. Let $\mathcal{W}$ and $\mathcal{B}_\epsilon$ be two distributions on $\{0,1\}^T$ where $\mathcal{B}_\epsilon$ is defined above and $\mathcal{W}$ satisfies the ϵ-martingale condition. Then*

$$\mathbf{S}^{s,k}[\mathcal{W}] \leq \mathbf{S}^{s,k}[\mathcal{B}_\epsilon] \leq \exp\bigl(-\Omega(\epsilon^3(1-O(\epsilon))k)\bigr) .$$

(Here, the asymptotic notation hides constants that do not depend on ϵ or k.)

By techniques similar to the ones used to prove this result, we obtain the following theorem pertaining directly to k-$\mathrm{CP}^{\mathsf{slot}}$ (and k-CP).

Theorem 2 (Main theorem; k-CP version)**.** *Let $\epsilon \in (0,1)$ and $T \in \mathbb{N}$. Let $w \in \{0,1\}^T$ be a random variable satisfying the ϵ-martingale condition. Then*

$$\Pr[w \textit{ violates } k\text{-CP}] \leq \Pr[w \textit{ violates } k\text{-CP}^{\mathsf{slot}}] \leq T \cdot \exp\bigl(-\Omega(\epsilon^3(1-O(\epsilon))k)\bigr) .$$

The proofs of these theorems are presented in Section 6.5. Additionally, we provide a $O(k^3)$-time algorithm for computing an explicit upper bound on these probabilities; cf. Appendix A.

2.4 Survey of the proofs of the main theorems

A central object in our combinatorial analysis is an "x-balanced fork" for a characteristic string $w = xy$. This fork contains two distinct, (viable,) maximum-length tines that are disjoint over y; See Definition 9 for details. A settlement violation for the slot $|x| + 1$ implies an x-balanced fork for the string xy; see Observation 1. In particular, for any distribution on characteristic strings in $\{0,1\}^n$ and $s + k \le n$,

$$\Pr_w[\text{slot } s \text{ is not } k\text{-settled}] \le \Pr_w \begin{bmatrix} \text{there is a decomposition } w = xyz \text{ and} \\ \text{a fork } F \vdash xy, \text{ where } |x| = s-1 \text{ and} \\ |y| \ge k+1, \text{ so that } F \text{ is } x\text{-balanced} \end{bmatrix}.$$

(This is a variant of Lemma 5 from Section 6.5.)

As promised above, common prefix violations can be handled the same way: we likewise establish (see Section 4; Theorem 3) that a common prefix violation implies that there exists a balanced fork for some prefix of w. Specifically, for any distribution of characteristic strings,

$$\Pr_w[w \text{ violates } k\text{-CP}^{\mathsf{slot}}] \le \Pr_w \begin{bmatrix} \text{there is a decomposition } w = xyz \text{ and a} \\ \text{fork } F \vdash xy, \text{ where } |y| \ge k+1, \text{ so that} \\ F \text{ is } x\text{-balanced} \end{bmatrix}. \tag{1}$$

Next, in Section 5, we define a combinatorial quantity called the "relative margin," written $\mu_x(y)$. In fact, for an arbitrary decomposition of the characteristic string $w = xy$, the event "there is an x-balanced fork for xy" is equivalent to the event "the relative margin $\mu_x(y)$ is non-negative;" this is Fact 1. In Lemma 3, we develop an exact recursive presentation for $\mu_x(y)$; hence we can bound the probability of a common prefix violation (or a settlement violation) by reasoning about the non-negativity of the relative margin and, in particular, without reasoning directly about forks.

In Section 6, we prove two bounds for the probability

$$\Pr_{xy}[\mu_x(y) \ge 0].$$

These bounds correspond to the distributions $\mathcal{B}_\epsilon$ and $\mathcal{W}$, respectively. For characteristic strings with distribution $\mathcal{B}_\epsilon$, we identify a random variable which stochastically dominates $\mu_x(y)$ and is amenable to exact analysis via generating functions; this yields the bound

$$\Pr_{w=xy}[\mu_x(y) \ge 0] \le \exp(-\Omega(|y|)).$$

Notice that this bound does not depend on the length of x. Next, in Lemma 4, we observe that the distribution $\mathcal{B}_\epsilon$ "stochastically dominates" the distribution $\mathcal{W}$. This observation has important repurcussions; in particular, the above inequality holds when w has the distribution $\mathcal{W}$ as well. See Section 6 for details.

It immediately follows that an (s,k)-settlement violation (or a k-CP$^{\mathsf{slot}}$ violation) is a rare event for distributions of interest. The multiplicative factor T in Theorem 2 comes from a union bound taken over all prefixes of w.

2.5 Comments on the model

Analysis in the Δ-synchronous setting. The security game above most naturally models a blockchain protocol over a synchronous network with immediate delivery (because each "honest" play of the challenger always builds on a fork that contains the fork generated by previous honest plays). However, the model can be easily adapted to protocols in the Δ-synchronous model adopted by the Snow White and Ouroboros Praos protocols and analyses. In particular, David et al. [6] developed a "Δ-reduction" mapping on the space of characteristic strings that permits analyses of forks (and the related statistics of interest, cf. §3) in the Δ-synchronous setting by direct appeal to the synchronous setting.

Public leader schedules. One attractive feature of this model is that it gives the adversary full information about the future schedule of leaders. The analysis of some protocols indeed demand this (e.g., Ouroboros, Snow White). Other protocols—especially those designed to offer security against adaptive adversaries (Praos, Genesis)—in fact contrive to keep the leader schedule private. Of course, as our analysis is in the more difficult "full information" model, it applies to all of these systems.

Bootstrapping multi-phase algorithms; stake shift. We remark that several existing proof-of-stake blockchain protocols proceed in phases, each of which is obligated to generate the randomness (for leader election, say) for the next phase based on the current stake distribution. The blockchain security properties of each phase are then individually analyzed—assuming clean randomness—which yields a recursive security argument; in this context the game outlined above precisely reflects the single phase analysis.

3 Definitions

We rely on the elementary framework of forks and margin from Kiayias et al. [13]. We restate and briefly discuss the pertinent definitions below. With these basic notions behind us, we then define a new "relative" notion of margin, which will allow us to significantly improve the efficacy of these tools for reasoning about settlement times. In particular, these tools will allow us to reason about the possibility that an adversary can produce two alternate histories of the blockchain that diverge prior to a particular block.

Recall that for a given execution of the protocol, we record the result of the leader election process via a *characteristic string* $w \in \{0,1\}^T$, defined such that $w_i = 0$ when a unique and honest party is assigned to slot i and $w_i = 1$ otherwise. A vertex of a fork is said to be *honest* if it is labeled with an index i such that $w_i = 0$.

Definition 7 (Tines, length and height)**.** *Let $F \vdash w$ be a fork for a characteristic string. A* tine *of F is a directed path starting from the root. For any tine t we define its* length *to be the number of edges in the path, and for any vertex v we define its* depth *to be the length of the unique tine that ends at v. The* height *of a fork (as usual for a tree) is the length of the longest tine, denoted* $\mathrm{height}(F)$.

Definition 8 (The $\sim_x$ relations)**.** *For two tines t_1 and t_2 of a fork F, it is convenient to define an equivalence relation that reflects whether t_1 and t_2 share an edge of F. If so, we write $t_1 \sim t_2$; otherwise we write $t_1 \nsim t_2$. We additionally extend this notation to reflect whether the tines share an edge over a particular suffix of w: for $w = xy$ we define $t_1 \sim_x t_2$ if t_1 and t_2 share an edge that terminates at some node labeled with an index in y; otherwise, we write $t_1 \nsim_x t_2$ (observe that in this case the paths share no vertex labeled by a slot associated with y). We sometimes call such pairs of tines* disjoint *(or, if $t_1 \nsim_x t_2$ for a string $w = xy$,* disjoint over *y). Note that $\sim$ and $\sim_\varepsilon$ are the same relation.*

Informally, $t_1 \sim_x t_2$ indicates that when we restrict our view of history to only blocks *after* the prefix x, t_1 and t_2 share an edge (and thus agree on at least one block after that point).

The basic structure we use to use to reason about settlement times is that of a "balanced fork."

Definition 9 (Balanced fork; cf. "flat" in [13])**.** *A fork F is* balanced *if it contains a pair of tines t_1 and t_2 for which $t_1 \nsim t_2$ and* $\mathrm{length}(t_1) = \mathrm{length}(t_2) = \mathrm{height}(F)$. *We define a relative notion of balance as follows: a fork $F \vdash xy$ is* x-balanced *if it contains a pair of tines t_1 and t_2 for which $t_1 \nsim_x t_2$ and* $\mathrm{length}(t_1) = \mathrm{length}(t_2) = \mathrm{height}(F)$.

Thus, balanced forks contain two completely disjoint, maximum-length tines, while x-balanced forks contain two maximum-length tines that may share edges in x but must be disjoint over the rest of the string. See Figures 2 and 3 for examples of balanced forks.

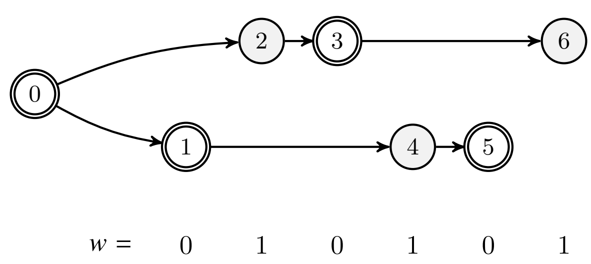


Figure 2: A balanced fork

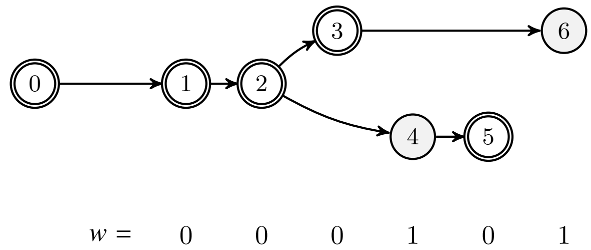


Figure 3: An x-balanced fork, where $x = 00$

Balanced forks and settlement time. A fundamental questions arising naturally in typical blockchain settings is how to determine *settlement time*, the delay after which the contents of a particular block of a blockchain can be considered stable. The existence of a balanced fork is a precise indicator for "settlement violations" in this sense. Specifically, consider a characteristic string xy and a transaction appearing in a block associated with the first slot of y (that is, slot $|x| + 1$). One clear violation of settlement at this point of the execution is the existence of two chains—each of maximum length—which diverge *prior to* y; in particular, this indicates that there is an x-balanced fork F for xy. Let us record this observation below.

Observation 1. *Let $s, k \in \mathbb{N}$ be given and let w be a characteristic string. Slot s is not k-settled for the characteristic string w if there exist a decomposition $w = xyz$, where $|x| = s - 1$ and $|y| \geq k + 1$, and an x-balanced fork for xy.*

In fact, every k-$\mathrm{CP}^{\mathsf{slot}}$ violation produces a balanced fork as well; see Theorem 3 in Section 4. In particular, to provide a rigorous k-slot settlement guarantee—which is to say that the transaction can be considered settled once k slots have gone by—it suffices to show that with overwhelming probability in choice of the characteristic string determined by the leader election process (of a full execution of the protocol), no such forks are possible. Specifically, if the protocol runs for a total of T time steps yielding the characteristics string $w = xy$ (where $w \in \{0, 1\}^T$ and the transaction of interest appears in slot $|x| + 1$ as above) then it suffices to ensure that there is no x-balanced fork for $x\hat{y}$, where $\hat{y}$ is a prefix of y of length at least $k + 1$; see Corollary 1 in Section 6. Note that for systems adopting the longest chain rule, this condition must necessarily involve the *entire future dynamics* of the blockchain. We remark that our analysis below will in fact let us take $T = \infty$.

Definition 10 (Closed fork)**.** *A fork F is* closed *if every leaf is honest. For convenience, we say the trivial fork is closed.*

Closed forks have two nice properties that make them especially useful in reasoning about the view of honest parties. First, a closed fork must have a unique longest tine (since honest parties are aware of honest blocks, and honest parties always add blocks to the longest seen chain). Second, recalling our description of the settlement game, closed forks intuitively capture decision points for the adversary. The adversary can potentially show many tines to many honest parties, but once an honest node has been placed on top of a tine, any adversarial blocks beneath it are part of the public record and are visible to all honest parties. For these reasons, we will often find it easier to reason about closed forks than arbitrary forks (without loss of generality).

The next few definitions are the start of a general toolkit for reasoning about an adversary's capacity to build highly diverging paths in forks, based on the underlying characteristic string.

Definition 11 (Gap, reserve, and reach)**.** *For a closed fork $F \vdash w$ and its unique longest tine $\hat{t}$, we define the* gap *of a tine t to be* $\mathrm{gap}(t) = \mathrm{length}(\hat{t}) - \mathrm{length}(t)$*. Furthermore, we define the* reserve *of t, denoted* $\mathrm{reserve}(t)$*, to be the number of adversarial indices in w that appear after the terminating vertex of t. More precisely, if v is the last vertex of t, then*

$$\mathrm{reserve}(t) = |\{\, i \mid w_i = 1 \text{ and } i > \ell(v)\}|\,.$$

These quantities together define the reach *of a tine:* $\mathrm{reach}(t) = \mathrm{reserve}(t) - \mathrm{gap}(t)$*.*

The notion of reach can be intuitively understood as a measurement of the resources available to our adversary in the settlement game. Reserve tracks the number of slots in which the adversary has the right to issue new blocks. When reserve exceeds gap (or equivalently, when reach is nonnegative), it indicates that a tine with nonnegative reach could be extended using a sequence of dishonest blocks until it is as long as (or longer than) the longest tine. Such a tine could be offered to an honest player who would prefer it over, e.g., the current longest tine in the fork. In contrast, a tine with negative reach is too far behind to be useful to the adversary at that time.

Definition 12 (Maximum reach)**.** *For a closed fork $F \vdash w$, we define $\rho(F)$ to be the largest reach attained by any tine of F, i.e.,*

$$\rho(F) = \max_{t} \operatorname{reach}(t)\,.$$

Note that $\operatorname{reach}(F)$ *is never negative (as the longest tine of any fork always has reach at least 0). We overload this notation to denote the maximum reach over all forks for a given characteristic string:*

$$\rho(w) = \max_{\substack{F \vdash w \\ F \text{ closed}}} \left[\max_{t} \operatorname{reach}(t)\right].$$

Definition 13 (Margin)**.** *The* margin *of a fork $F \vdash w$, denoted $\mu(F)$, is defined as*

$$\mu(F) = \max_{t_1 \nsim t_2}\bigl(\min\{\operatorname{reach}(t_1), \operatorname{reach}(t_2)\}\bigr)\,, \tag{2}$$

where this maximum is extended over all pairs of disjoint tines of F; thus margin reflects the "second best" reach obtained over all disjoint tines. In order to study splits in the chain over particular portions of a string, we generalize this to define a "relative" notion of margin: If $w = xy$ for two strings x and y and, as above, $F \vdash w$, we define

$$\mu_x(F) = \max_{t_1 \nsim_x t_2}\bigl(\min\{\operatorname{reach}(t_1), \operatorname{reach}(t_2)\}\bigr)\,.$$

Note that $\mu_\varepsilon(F) = \mu(F)$.

For convenience, we once again overload this notation to denote the margin of a string. $\mu(w)$ refers to the maximum value of $\mu(F)$ over all possible closed forks F for a characteristic string w:

$$\mu(w) = \max_{\substack{F \vdash w, \\ F \text{ closed}}} \mu(F)\,.$$

Likewise, if $w = xy$ for two strings x and y we define

$$\mu_x(y) = \max_{\substack{F \vdash w, \\ F \text{ closed}}} \mu_x(F)\,.$$

Note that, at least informally, "second-best" tines are of natural interest to an adversary intent on the construction of an x-balanced fork, which involves two (partially disjoint) long tines.

Balanced forks and relative margin. Kiayias et al. [13] showed that a balanced fork can be constructed for a given characteristic string w if and only if there exists some closed $F \vdash w$ such that $\mu(F) \geq 0$. We record a relative version of this theorem below, which will ultimately allow us to extend the analysis of [13] to more general class of disagreement and settlement failures.

Fact 1. *Let $xy \in \{0,1\}^n$ be a characteristic string. Then there is an x-balanced fork $F \vdash xy$ if and only if $\mu_x(y) \geq 0$.*

Proof. The proof is immediate from the definitions. We sketch the details for completeness.

Suppose F is an x-balanced fork for xy. Then F must contain a pair of tines t_1 and t_2 for which $t_1 \nsim_x t_2$ and $\operatorname{length}(t_1) = \operatorname{length}(t_2) = \operatorname{height}(F)$. We observe that (1) $\operatorname{gap}(t_i) = 0$ for both t_1 and t_2, and (2) reserve is always a nonnegative quantity. Together with the definition of reach, these two facts immediately imply $\operatorname{reach}(t_i) \geq 0$. Because t_1 and t_2 are edge-disjoint over y and $\min\{\operatorname{reach}(t_1), \operatorname{reach}(t_2)\} \geq 0$, we conclude that $\mu_x(y) \geq 0$, as desired.

Suppose $\mu_x(y) \geq 0$. Then there is some closed fork F for xy such that $\mu_x(F) \geq 0$. By the definition of relative margin, we know that F has two tines t_1, t_2 such that $t_1 \nsim_x t_2$ and $\operatorname{reach}(t_i) \geq 0$. Recall that we define reach by $\operatorname{reach}(t) = \operatorname{reserve}(t) - \operatorname{gap}(t)$, and so in this case it follows that $\operatorname{reserve}(t_i) - \operatorname{gap}(t_i) \geq 0$. Thus, an x-balanced fork $F' \vdash xy$ can be constructed from F by appending a path of $\operatorname{gap}(t_i)$ adversarial vertices to each t_i. $\square$

Indeed, we can define the "forkability" of a characteristic string in terms of its margin.

Definition 14 (Forkable strings)**.** *A charactersitic string w is* forkable *if its margin is non-negative, i.e., $\mu(w) \geq 0$. Equivalently, w is forkable if there is a balanced fork for w.*

Although this definition is not necessary for our presentation, it helps make our statements compatible with the existing literature.

4 Common prefix violation and balanced forks

We will show that for any characteristic string, a common prefix violation implies the existence of a balanced fork on a prefix of that string. This allows us to bound consistency errors by reasoning about balanced forks. In particular, inequality (1) is a direct consequence of the theorem below.

Theorem 3. *Let $k, T \in \mathbb{N}$. Let $w \in \{0,1\}^T$ be a characteristic string which violates k-$\mathrm{CP}^{\mathsf{slot}}$. Then there exist a decomposition $w = xyz$ and a fork $\hat{F} \vdash xy$, where $|y| \geq k+1$, so that $\hat{F}$ is x-balanced.*

Proof. Recall that $\ell(t)$ is the slot index of the last vertex of tine t. Define $A \triangleq \cup_{F \vdash w} A_F$ where, for a given fork $F \vdash w$, define

$$A_F \triangleq \left\{ (\tau_1, \tau_2) \;:\; \begin{array}{l} \tau_1, \tau_2 \text{ are two viable tines in the fork } F, \\ \ell(\tau_1) \leq \ell(\tau_2), \text{ and the pair } (\tau_1,\tau_2) \text{ is a wit-} \\ \text{ness to a } k\text{-}\mathrm{CP}^{\mathsf{slot}} \text{ violation} \end{array} \right\} .$$

Define the *slot divergence* of two tines as $\mathrm{div}_{\mathsf{slot}}(\tau_1, \tau_2) \triangleq \ell(\tau_1) - \ell(\tau_1 \cap \tau_2)$ where $\tau_1 \cap \tau_2$ denotes the common prefix of the tines τ_1 and τ_2. Recalling the definition of a k-$\mathrm{CP}^{\mathsf{slot}}$ violation, it is clear that

$$\mathrm{div}_{\mathsf{slot}}(\tau_1, \tau_2) \geq k+1 \quad \text{for all } (\tau_1, \tau_2) \in A\,. \tag{3}$$

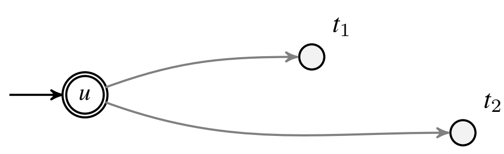
Notice that there must be a tine-pair $(t_1, t_2) \in A$ which satisfies the following two conditions:

$$\mathrm{div}_{\mathsf{slot}}(t_1, t_2) \text{ is maximum over } A\,, \text{and} \tag{4}$$

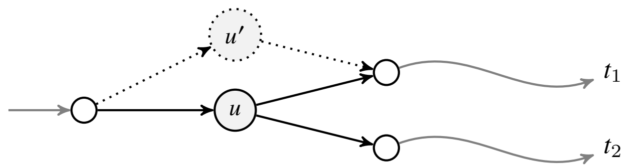
$$|\ell(t_2) - \ell(t_1)| \text{ is minimum among all tine-pairs in } A \text{ for which (4) holds.} \tag{5}$$

The tines t_1, t_2 will play a special role in our proof; let F be a fork containing these tines.

The prefix x, fork F_x, and vertex u. Let u denote the last vertex on the tine $t_1 \cap t_2$, as shown in the diagram below, and let $\alpha \triangleq \ell(u) = \ell(t_1 \cap t_2)$. Let $x \triangleq w_1, \ldots, w_\alpha$ and let F_x be the fork-prefix of F supported on x. We will argue that u must be honest and, in addition, that F_x must contain a unique longest tine t_u terminating at the vertex u. We will also identify a substring $y, |y| \geq k+1$ such that w can be written as $w = xyz$. Then we will construct a balanced fork $\tilde{F}_y \vdash y$ by modifying the subgraph of F supported on y. We will finish the proof by constructing an x-balanced fork by suitably appending $\tilde{F}_y$ to F_x.



u **must be an honest vertex.** We observe, first of all, that the vertex u cannot be adversarial: otherwise it is easy to construct an alternative fork $F' \vdash w$ and a pair of tines in F' that violate (4). Specifically, construct F' from F by adding a new (adversarial) vertex u' to F for which $\ell(u') = \ell(u)$, adding an edge to u' from the vertex preceding u, and replacing the edge of t_1 following u with one from u'; then the other relevant properties of the fork are maintained, but the slot divergence of the resulting tines has increased by at least one. (See the diagram below.)



F_x has a unique, longest (and honest) tine t_u. A similar argument implies that the fork F_x has a unique vertex of depth $\mathrm{depth}(u)$: namely, u itself. In the presence of another vertex u' (of F_x) with depth $\mathrm{depth}(u)$, "redirecting" t_1 through u' (as in the argument above) would likewise result in a fork with a larger slot divergence. To see this, notice that $\ell(u')$ must be strictly less than $\ell(u)$ since $\ell(u)$ is an honest slot (which means u is the only vertex at that slot). Thus $\ell(\cdot)$ would indeed be increasing along this new tine (resulting from redirecting t_1). As α is the last index of the string x, this additionally implies that F_x has no vertices of depth exceeding $\mathrm{depth}(u)$. Let $t_u \in F_x$ be the tine with $\ell(t_u) = \alpha$.

$$\text{The honest tine } t_u \text{ is the unique longest tine in } F_x\,. \tag{6}$$

Identifying y. Let β denote the smallest honest index of w for which $\beta \geq \ell(t_2)$, with the convention that if there is no such index we define $\beta = T + 1$. Observe that $\beta - 1 \geq \ell(t_1)$. (If $\ell(t_2)$ is an honest slot then $\beta = \ell(t_2)$ but $\ell(t_1) < \ell(t_2)$. The case $\ell(t_1) = \ell(t_2)$ is possible if $\ell(t_2)$ is an adversarial slot; but then $\beta > \ell(t_2)$.) These indices, α and β, distinguish the substrings $y = w_{\alpha+1} \dots w_{\beta-1}$ and $z = w_\beta \dots w_T$; we will focus on y in the remainder of the proof. Since the function $\ell(\cdot)$ is strictly increasing along any tine, observe that

$$|y| = \beta - \alpha - 1 \geq \ell(t_1) - \ell(u) \geq k + 1\,.$$

Hence y has the desired length and it suffices to establish that it is forkable. We can extract from F a balanced fork (for y) in two steps: (i.) we subject the fork F to some minor restructuring to ensure that all "long" tines pass through u; (ii.) we construct a flat fork by treating the vertex u as the root of a portion of the subtree of F labeled with the indices of y. At the conclusion of the construction, the segments of the two tines t_1 and t_2 will yield the required "long, disjoint, equal-length" tines satisfying the definition of a balanced fork.

Honest indices in xy have low depths. The minimality assumption (5) implies that any honest index h for which $h < \beta$ has depth no more than $\min(\mathrm{length}(t_1), \mathrm{length}(t_2))$: specifically,

$$h < \beta \quad \Longrightarrow \quad \mathbf{d}(h) \leq \min(\mathrm{length}(t_1), \mathrm{length}(t_2))\,. \tag{7}$$

To see this, consider an honest index $h, h < \beta$ and a tine t_h for which $\ell(t_h) = h$. Recall that t_1 and t_2 are viable and that $h < \ell(t_2)$. (If $\ell(t_2)$ is honest, it is obvious. Otherwise, $h < \ell(t_2) < \beta$ since $\ell(t_2)$ is adversarial.) As t_2 is viable, it follows immediately that $\mathbf{d}(h) = \mathrm{length}(t_h) \leq \mathrm{length}(t_2)$. Similarly, if $h \leq \ell(t_1)$ then $\mathbf{d}(h) \leq \mathrm{length}(t_1)$ since t_1 is viable as well. The remaining case, i.e., when $\ell(t_1) < h < \ell(t_2)$, can be ruled out by the argument below.

There is no honest index between $\ell(t_1)$ and $\ell(t_2)$. We claim that

$$\text{There is no honest index } h \text{ satisfying } \ell(t_1) < h < \ell(t_2)\,. \tag{8}$$

The claim above is trivially true if $\ell(t_1) = \ell(t_2)$. Otherwise, suppose (toward a contradiction) that h is an honest index satisfying $\ell(t_1) < h < \ell(t_2)$. Let t_h be the (honest) tine at slot h. The tine-pair (t_1, t_h) may or may not be in A. We will show that both cases lead to contradictions.

- If (t_1, t_h) is in A and $\ell(t_1 \cap t_h) \leq \ell(u)$, $\mathrm{div}_{\mathsf{slot}}(t_1, t_h)$ is at least $\mathrm{div}_{\mathsf{slot}}(t_1, t_2)$. In fact, due to (4), this inequality must be an equality. However, the assumption $\ell(t_1) < h < \ell(t_2)$ contradicts (5).
- If (t_1, t_h) is in A and $\ell(t_1 \cap t_h) > \ell(u)$, it follows that $\mathrm{div}_{\mathsf{slot}}(t_h, t_2) > \mathrm{div}_{\mathsf{slot}}(t_1, t_2)$. As the latter quantity is at least $k + 1$, (t_h, t_2) must be in A. The preceding inequality, however, contradicts (4).
- If $(t_1, t_h) \notin A$, $\mathrm{div}_{\mathsf{slot}}(t_1, t_h)$ is at most k. As $\mathrm{div}_{\mathsf{slot}}(t_1, t_2)$ is at least $k + 1$, t_h and t_1 must share a vertex after slot $\ell(u)$. Since $\ell(t_1) < h < \ell(t_2)$ by assumption, $\mathrm{div}_{\mathsf{slot}}(t_h, t_2) > \mathrm{div}_{\mathsf{slot}}(t_1, t_2) \geq k + 1$ and, as a result, $(t_h, t_2) \in A$. However, the preceding strict inequality violates condition (4).

A fork $F^{\triangleright u \triangleleft}$ where all long tines go through u. In light of the remarks above, we observe that the fork F may be "pinched" at u to yield an essentially identical fork $F^{\triangleright u \triangleleft} \vdash w$ with the exception that all tines of length exceeding $\mathrm{depth}(u)$ pass through the vertex u. Specifically, the fork $F^{\triangleright u \triangleleft} \vdash w$ is defined to be the graph obtained from F by changing every edge of F directed towards a vertex of depth $\mathrm{depth}(u) + 1$ so that it originates from u. To see that the resulting tree is a well-defined fork, it suffices to check that $\ell(\cdot)$ is still increasing along all tines of $F^{\triangleright u \triangleleft}$. For this purpose, consider the effect of this pinching on an individual tine t terminating at a particular vertex v—it is replaced with a tine $t^{\triangleright u \triangleleft}$ defined so that:

- If $\mathrm{length}(t) \leq \mathrm{depth}(u)$, the tine t is unchanged: $t^{\triangleright u \triangleleft} = t$.
- Otherwise, $\mathrm{length}(t) > \mathrm{depth}(u)$ and t has a vertex v of depth $\mathrm{depth}(u) + 1$; note that $\ell(v) > \ell(u)$ because F_x contains no vertices of depth exceeding $\mathrm{depth}(u)$. Then $t^{\triangleright u \triangleleft}$ is defined to be the path given by the tine terminating at u, a (new) edge from u to v, and the suffix of t beginning at z. (As $\ell(v) > \ell(u)$ this has the increasing label property.)

Thus the tree $F^{\triangleright u \triangleleft}$ is a legal fork on the same vertex set; note that the depths of vertices in F and $F^{\triangleright u \triangleleft}$ are identical.

Constructing a shallow fork $F_y \vdash y$. By excising the tree rooted at u from this pinched fork $F^{\triangleright u \triangleleft}$, we may extract a fork for the string $w_{\alpha+1} \dots w_T$. Specifically, consider the induced subgraph $F^{u \triangleleft}$ of $F^{\triangleright u \triangleleft}$ given by the vertices $\{u\} \cup \{v \mid \mathrm{depth}(v) > \mathrm{depth}(u)\}$. By treating u as a root vertex and suitably defining the labels $\ell^{u \triangleleft}$ of $F^{u \triangleleft}$ so that $\ell^{u \triangleleft}(v) = \ell(v) - \ell(u)$, this subgraph has the defining properties of a fork for $w_{\alpha+1} \dots w_T$. In particular, considering that α is honest it follows that each honest index $h > \alpha$ has depth $\mathbf{d}(h) > \mathrm{length}(u)$ and hence h labels a vertex in $F^{u \triangleleft}$. For a tine t of $F^{\triangleright u \triangleleft}$, we let $t^{u \triangleleft}$ denote the suffix of this tine beginning at u, which forms a tine in $F^{u \triangleleft}$. (If $\mathrm{length}(t) \leq \mathrm{depth}(u)$, we define $t^{u \triangleleft}$ to consist solely of the vertex u.) Note that ${t_1}^{u \triangleleft}$ and ${t_2}^{u \triangleleft}$ share no edges in the fork $F^{u \triangleleft}$.

Finally, let F_y denote the subtree obtained from $F^{u \triangleleft}$ as the union of all tines $t^{u \triangleleft}$ of $F^{u \triangleleft}$ so that all labels of $t^{u \triangleleft}$ are drawn from y (as it appears as a prefix of $w_{\alpha+1} \dots w_T$), and

$$\mathrm{length}(t^{u \triangleleft}) \leq \max_{\substack{h \leq |y| \\ h \text{ honest}}} \mathbf{d}(h)\,. \tag{9}$$

It is immediate that $F_y \vdash y$.

Two longest viable tines in F_y. Consider the tines ${t_1}^{u \triangleleft}$ and ${t_2}^{u \triangleleft}$. As mentioned above, they share no edges in $F^{u \triangleleft}$ and hence the prefixes $\check{t}_1$ and $\check{t}_2$ (of ${t_1}^{u \triangleleft}$ and ${t_2}^{u \triangleleft}$) appearing in F_y share no edges. We wish to show that these prefixes have the maximal length in F_y, making F_y balanced, as desired. Let h be the largest honest index in y. Since the lengths of the tines in F_y are at most $\mathbf{d}(h)$, it suffices to show that the lengths of $\check{t}_i, i \in \{1, 2\}$ is at least $\mathbf{d}(h)$.

This is immediate for the tine $\check{t}_1$ since all labels of ${t_1}^{u \triangleleft}$ are drawn from y and, considering (7), its depth is at least that of all relevant honest vertices. As for $\check{t}_2$, observe that if $\ell(t_2)$ is not honest then $\beta > \ell(t_2)$ so that, as with $\check{t}_1$, the tine $\check{t}_2$ is labeled by y so that the same argument, relying on (7), ensures that the $\mathrm{length}(\check{t}_2)$ is at least the depth of all relevant honest vertices. If $\ell(t_2)$ is honest, $\beta = \ell(t_2)$, and the terminal vertex of ${t_2}^{u \triangleleft}$ does not appear in F_y (as $\ell({t_2}^{u \triangleleft})$ falls outside y). In this case, however, $\mathrm{length}({t_2}^{u \triangleleft}) > \mathbf{d}(h)$ for any honest index h of y. It follows that $\mathrm{length}(\check{t}_2)$, which equals $\mathrm{length}({t_2}^{u \triangleleft}) - 1$, is at least the depth of any honest index of y, as desired. Thus we have proved

$$\check{t}_1 \text{ and } \check{t}_2 \text{ are two maximally long viable tines in } F_y \vdash y\,. \tag{10}$$

Constructing a flat fork $\tilde{F}_y \vdash y$. Let us identify the fork prefix $\tilde{F}_y \sqsubseteq F_y$ which is either identical to F_y or differs from F_y in only one of the tines $\check{t}_1, \check{t}_2$. In particular, if $\mathrm{length}(\check{t}_1) = \mathrm{length}(\check{t}_2)$, we set $\tilde{F}_y = F_y$. Otherwise, let $\check{t}_a$ be the longer of the two tines $\check{t}_1, \check{t}_2$; let $\check{t}_b$ be the shorter one. We modify F_y by deleting some trailing adversarial nodes from $\check{t}_a$ until it has the same length as $\check{t}_b$; we set $\tilde{F}_y$ as the resulting fork and, in addition, set $\tilde{t}_b = \check{t}_b$ and $\tilde{t}_a$ as the tine after trimming $\check{t}_a$.

We claim that $\tilde{F}_y$ is balanced. The claim is obvious if $\mathrm{length}(\check{t}_1) = \mathrm{length}(\check{t}_2)$. Otherwise, thanks to (10), it remains to show that the longer tine, $\check{t}_a$, has sufficiently many trailing adversarial nodes which, if deleted, yields $\mathrm{length}(\tilde{t}_1) = \mathrm{length}(\tilde{t}_2)$. To that end, let h_i be the index of the last honest vertex on $\check{t}_i \in F_y, i \in \{1, 2\}$.

Suppose $\operatorname{length}(\check{t}_2) > \operatorname{length}(\check{t}_1)$. By (8), we also have $\operatorname{length}(\check{t}_1) \geq \mathbf{d}(h_2)$ and hence we can trim some of the trailing adversarial nodes from $\check{t}_2$ to get the tine $\tilde{t}_2$ whose length is the same as that of $\check{t}_1$. Otherwise, suppose $\operatorname{length}(\check{t}_1) > \operatorname{length}(\check{t}_2)$. Since t_2 is a viable tine in F, we also have $\operatorname{length}(\check{t}_2) \geq \mathbf{d}(h_1)$. Thus we can trim some of the trailing adversarial nodes from $\check{t}_1$ to have a tine $\tilde{t}_1$ whose length is the same as that of $\check{t}_2$. In any case, the quantity $\min(\operatorname{length}(\tilde{t}_1), \operatorname{length}(\tilde{t}_2))$ remains the same as $\min(\operatorname{length}(\check{t}_1), \operatorname{length}(\check{t}_2))$. Thus the fork $\tilde{F}_y$ has at least two tines, $\tilde{t}_1$ and $\tilde{t}_2$, that achieve the maximum length of all tines in $\tilde{F}_y$; hence $\tilde{F}_y$ is balanced.

An x-balanced fork $\hat{F} \sqsubseteq F$. Let us identify the root of the fork $\tilde{F}_y$ with the vertex u of F_x and let $\hat{F}$ be the resulting graph (after "gluing" the root of $\tilde{F}_y$ to u). By (6), it is easy to see that the fork $\hat{F} \sqsubseteq F$ is indeed a valid fork on the string xy. Moreover, $\hat{F}$ is x-balanced since $\tilde{F}_y$ is balanced. The claim in Theorem 3 follows immediately since $|y| \geq k + 1$. □

5 A simple recursive formulation of relative margin

A significant finding of Kiayias et al. [13] is that the margin of a characteristic string $\mu(w)$—the maximum value of a quantity taken over a (typically) exponentially-large family of forks—can be given a simple, mutually recursive formulation with the associated quantity of reach $\rho(w)$. Specifically, they prove the following lemma.

Lemma 2 ([13, Lemma 4.19])**.** *$\rho(\varepsilon) = 0$ where ε is the empty string, and, for all nonempty strings $w \in \{0,1\}^*$,*

$$\rho(w1) = \rho(w) + 1\,, \qquad \text{and} \qquad \rho(w0) = \begin{cases} 0 & \text{if } \rho(w) = 0, \\ \rho(w) - 1 & \text{otherwise.} \end{cases} \tag{11}$$

Furthermore, margin satisfies the mutually recursive relationship $\mu(\varepsilon) = 0$ and for all $w \in \{0,1\}^*$,

$$\mu(w1) = \mu(w) + 1\,, \qquad \text{and} \qquad \mu(w0) = \begin{cases} 0 & \text{if } \rho(w) > \mu(w) = 0, \\ \mu(w) - 1 & \text{if } \rho(w) = 0, \\ \mu(w) - 1 & \text{otherwise.} \end{cases} \tag{12}$$

Additionally, there exists a closed fork $F \vdash w$ such that $\rho(F) = \rho(w)$ and $\mu(F) = \mu(w)$. (It is convenient to separate the case $\rho(w) = 0$ from the other case which also yields $\mu(w) - 1$ in the proof, so we reflect that in the statement of the theorem.)

We prove an analogous recursive statement for relative margin, recorded below.

Lemma 3 (Relative margin)**.** *Given a fixed string $x \in \{0,1\}^*$, $\mu_x(\varepsilon) = \rho(x)$ where ε is the empty string, and, for all nonempty strings $w = xy \in \{0,1\}^*$,*

$$\mu_x(y1) = \mu_x(y) + 1\,, \qquad \text{and} \qquad \mu_x(y0) = \begin{cases} 0 & \text{if } \rho(xy) > \mu_x(y) = 0\,, \\ \mu_x(y) - 1 & \text{if } \rho(xy) = 0\,, \\ \mu_x(y) - 1 & \text{otherwise.} \end{cases} \tag{13}$$

Additionally, there exists a closed fork $F \vdash xy$ such that $\rho(F) = \rho(xy)$ and $\mu_x(F) = \mu_x(y)$.

We delay the proof of Lemma 3 to Section 7, preferring to immediately focus on the application to settlement times in Section 6.

Discussion. The proof of Lemma 3 shares many technical similarities with the proof of Lemma 2 given by Kiayias et al. [13]. However, there is an important respect in which the proofs differ. Each of the proofs requires the definition of a particular adversary (which, in effect, constructs a fork achieving the worst case reach and margin guaranteed by the lemma). The adversary constructed by [13] can create a balanced fork for w whenever $\mu(w) \geq 0$ (i.e., w is "forkable"). However, the adversary only focuses on the problem of producing disjoint tines over the *entire string w* (consistent with the definition of $\mu(\cdot)$). The "online adversary," developed during the proof of Lemma 3, uses a more sophisticated logic for extending chains (tines) of the fork, which allows it to *simultaneously maximize relative margin over all prefixes of the string*. This remarkable property is important for the settlement proofs in Section 6.

6 General settlement guarantees and proof of main theorems

With the recursive formulation for relative margin in hand, we study the stochastic process that arises when the characteristic string w is chosen from a distribution satisfying the ϵ-martingale condition. Let us write $w = xy$ (where the decomposition is arbitrary) and let E be the event that the relative margin $\mu_x(y)$ is non-negative. As Fact 1 and Observation 1 point out, this event has a direct bearing on the settlement violation on w.

In this section, we prove two bounds on the probability of the event E. The first bound corresponds to the distribution $\mathcal{B}_\epsilon$ whereas the second bound applies to any distribution that satisfies the ϵ-martingale condition. (Recall that the distribution $\mathcal{B}_\epsilon$, mentioned in Theorem 1, satisfies the ϵ-martingale condition with equality.) Our exposition in this section culminates in the proofs of our main theorems.

We start with the following theorem which is a direct consequences of these bounds; see Section 6.1 for a proof.

Theorem 4. *Let $T, k \in \mathbb{N}$. Let $w \in \{0,1\}^T$ be a random variable satisfying the ϵ-martingale condition. Consider the decomposition $w = xy, |y| = k$. Then*

$$\Pr_{w=xy}[\text{there is an } x\text{-balanced fork for } xy] = \Pr_{w=xy}[\mu_x(y) \geq 0] \leq \exp(-\Omega(k))\,.$$

(The asymptotic notation hides constants that depend only on ϵ.)

Notice how the final bound does not depend on $|x|$. Indeed, as we show in Lemma 4, the reach of a Boolean string x drawn from the distribution $\mathcal{B}_\epsilon$ converges to a fixed exponential distribution as $|x| \to \infty$. This limiting distribution "stochastically dominates" any distribution that satisfies the ϵ-martingale condition; see Section 6.2. The following corollary is immediate.

Corollary 1. *Let $T, s, k \in \mathbb{N}$. Let $w \in \{0,1\}^T$ be a random variable satisfying the ϵ-martingale condition. Then*

$$\Pr_w\begin{bmatrix}\text{there is a decomposition } w = xyz, \text{ where } |x| = \\ s-1 \text{ and } |y| \geq k, \text{ so that } \mu_x(y) \geq 0\end{bmatrix} \leq O(1)\cdot\exp(-\Omega(k))\,. \tag{14}$$

Proof. Notice that Theorem 4 works for *any* prefix x of the characteristic string $w = xy$. Thus we can fix the prefix x with length $s-1$ and sum the bound in Theorem 4 over all suffixes y with length at least k. This would give an upper bound to the left-hand side of our claim, the bound being $\sum_{t\geq k}\exp(-\Omega(t)) = O(1)\cdot\exp(-\Omega(k))$. □

We obtain another imporant corollary by setting $|x| = 0$ and $|y| = n$ in Theorem 4.

Corollary 2. *Let $w \in \{0,1\}^n$ be a random variable satisfying the ϵ-martingale condition. Then*

$$\Pr[w \text{ is forkable}] = \Pr[\mu(w) \geq 0] \leq \exp(-\Omega(n))\,.$$

Thus *forkable strings are rare* where "forkable" is defined in Definition 14. This result significantly strengthens the $\exp(-\Omega(\sqrt{n}))$ bound obtained in Theorem 4.13 of [13]. The improvement comes in two respects: first, Corollary 1 improves the exponent from $\sqrt{n}$ to n, and second, the characteristic string is allowed to be drawn from any distribution satisfying the ϵ-martingale condition. For comparison, the characteristic string in Theorem 4.13 of [13] has the distribution $\mathcal{B}_\epsilon$, i.e., the bits were i.i.d. Bernoulli random variables with expectation $(1-\epsilon)/2$.

6.1 Two bounds for non-negative relative margin

The main ingredients to proving Theorem 4 are two bounds on the event that for a characteristic string xy, the relative margin $\mu_x(y)$ is non-negative.

Bound 1. *Let $x \in \{0,1\}^m$ and $y \in \{0,1\}^k$ be independent random variables, each chosen according to $\mathcal{B}_\epsilon$. Then*

$$\Pr[\mu_x(y) \geq 0] \leq \exp(-\epsilon^3(1-O(\epsilon))k/2)\,.$$

Bound 2. *Let $x \in \{0,1\}^m$ and $y \in \{0,1\}^k$ be random variables (jointly) satisfying the ϵ-martingale condition with respect to the ordering $x_1, \ldots, x_m, y_1, \ldots, y_k$. Let $x' \in \{0,1\}^m$ and $y' \in \{0,1\}^k$ be independent random variables, each chosen independently according to $\mathcal{B}_\epsilon$. Then*

$$\Pr[\mu_x(y) \geq 0] \leq \Pr[\mu_{x'}(y') \geq 0] .$$

As a result,

$$\Pr[\mu_x(y) \geq 0] \leq \exp(-\epsilon^3(1 - O(\epsilon))k/2) .$$

Proof of Theorem 4. The equality is Fact 1 and the inequality is Bound 2. □

6.2 Stochastic dominance among distributions

The statement of Bound 2 stems from the fact that the distribution $\mathcal{B}_\epsilon$ stochastically dominates any distribution satisfying the ϵ-martingale condition; we define the notion of dominance below. For notational convenience, we denote the probability distribution associated with a random variable using uppercase script letters; for example, the distribution of a random variable R is denoted by $\mathcal{R}$. This usage should be clear from the context.

Definition 15 (Stochastic dominance)**.** *Let X and Y be random variables taking values in some set Ω endowed with a partial order $\leq$. We say that X* stochastically dominates *Y, written $Y \preceq X$, if $\mathcal{X}(A) \geq \mathcal{Y}(A)$ for all* monotone sets *$A \subseteq \Omega$, where a set $A \subseteq \Omega$ is called monotone if, for all $x \leq y$, $x \in A$ implies $y \in A$. As a special case, when $\Omega = \mathbb{R}$, $Y \preceq X$ if $\Pr[X \geq \Lambda] \geq \Pr[Y \geq \Lambda]$ for every $\Lambda \in \mathbb{R}$. We extend this notion to probability distributions in the natural way.*

Observe that if $Y \preceq X$ and Z is independent of both X and Y, then $Z + Y \preceq Z + X$. In addition, for any non-decreasing function u defined on Ω, $Y \preceq X$ implies $u(Y) \preceq u(X)$.

Let $m \in \mathbb{N}$ and suppose $W = (W_1, \ldots, W_m) \in \{0,1\}^m$ satisfies the ϵ-martingale condition. It turns out that $\rho(W)$ is stochastically dominated by the distribution of $\rho(B_1, \ldots, B_m)$, where each $B_i \in \{0,1\}$ is an independent Bernoulli random variable with parameter $(1-\epsilon)/2$. In addition, $\rho(B_1, \ldots, B_m)$ is stochastically dominated by its limiting (stationary) distribution where we take $m \to \infty$.

Lemma 4. *Suppose $W = (W_1, \ldots, W_n) \in \{0,1\}^n$ satisfies the ϵ-martingale condition. Let $\epsilon \in (0,1)$ and $B = (B_1, \ldots, B_n) \in \{0,1\}^n$ where each B_i is independent with expectation $(1-\epsilon)/2$. Let $R_\infty \in \{0, 1, \ldots\}$ be a random variable whose distribution $\mathcal{R}_\infty$ is defined as*

$$\mathcal{R}_\infty(k) = \Pr[R_\infty = k] \triangleq \left(\frac{2\epsilon}{1+\epsilon}\right) \cdot \left(\frac{1-\epsilon}{1+\epsilon}\right)^k \qquad \text{for } k = 0, 1, 2, \ldots \, . \tag{15}$$

Then

$$\rho(W) \preceq \rho(B) \preceq R_\infty .$$

Proof.

***B* stochastically dominates *W*.** As a matter of notation, for any fixed values $w_1, \ldots, w_k \in \{0,1\}^k$, let

$$\theta[w_1, \ldots, w_k] = \Pr[W_{k+1} = 1 \mid W_i = w_i, \text{for } i \leq k] \leq (1-\epsilon)/2$$

and $\theta[\varepsilon] = \Pr[W_1 = 1]$ where ε is the empty string. Then consider n uniform and independent real numbers $(A_1, \ldots, A_n)$, each taking a value in the unit interval $[0,1]$; we use these random variables to construct a monotone coupling between W and B. Specifically, define $\beta : [0,1]^n \to \{0,1\}^n$ by the rule $\beta(\alpha_1, \ldots, \alpha_n) = (b_1, \ldots, b_n)$ where

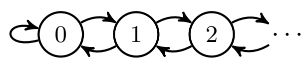
$$b_t = \begin{cases} 1 & \text{if } \alpha_t \leq (1-\epsilon)/2, \\ 0 & \text{if } \alpha_t > (1-\epsilon)/2, \end{cases}$$

and define $B = (B_1, \ldots, B_n) = \beta(A_1, \ldots, A_n)$; these B_is are independent zero-one Bernoulli random variables with expectation $(1-\epsilon)/2$. Likewise define the function $\omega : [0,1]^n \to \{0,1\}^n$ so that $\omega(\alpha_1, \ldots, \alpha_n) = (w_1, \ldots, w_n)$ where each w_t is assigned by the iterative rule

$$w_{t+1} = \begin{cases} 1 & \text{if } \alpha \leq \theta[w_1, \ldots, w_t], \\ 0 & \text{if } \alpha > \theta[w_1, \ldots, w_t], \end{cases}$$

and observe that the probability law of $\omega(A_1, \ldots, A_n)$ is precisely that of $W = (W_1, \ldots, W_n)$. For convenience, we simply identify the random variable W with $\omega(A_1, \ldots, A_n)$. Note that for any $\alpha = (\alpha_1, \ldots, \alpha_n)$ and for each i, the ith coordinates of $\beta(\alpha)$ and $\omega(\alpha)$ satisfy $\omega(\alpha)_i \leq \beta(\alpha)_i$ (which is to say that $W_i \leq B_i$). It follows immediately that $\rho(\omega(\alpha)) \leq \rho(\beta(\alpha))$ with probability 1 and hence $\rho(W) \leq \rho(B)$. See [14, Lemma 22.5].

R_∞ stochastically dominates $\rho(B)$. To complete the proof, we now establish that $\rho(B) \leq R_\infty$. We remark that the random variables $\rho(B)$ (and R_∞) have an immediate interpretation in terms of the Markov chain corresponding to a biased random walk on $\mathbb{Z}$ with a "reflecting boundary" at -1. Specifically, consider the Markov chain on $\{0, 1, \ldots\}$ given by the transition diagram



where edges pointing right have probability $(1-\epsilon)/2$ and edges pointing left—including the loop at 0—have probability $(1+\epsilon)/2$. Examining the recursive description of $\rho(w)$, it is easy to confirm that the random variable $\rho(B_1, \ldots, B_n)$ is precisely given by the result of evolving the Markov chain above for n steps with all probability initially placed at 0. It is further easy to confirm that the distribution given by (15) above is stationary for this chain.

To establish stochastic dominance, it's convenient to work with the underlying distributions and consider walks of varying lengths: let $\mathcal{R}_n : \mathbb{Z} \to \mathbb{R}$ denote the probability distribution given by $\rho(B_1, \ldots, B_n)$; likewise define $\mathcal{R}_\infty$. For a distribution $\mathcal{R}$ on $\mathbb{Z}$, we define $[\mathcal{R}]_0$ to denote the probability distribution obtained by shifting all probability mass on negative numbers to zero; that is, for $x \in \mathbb{Z}$,

$$[\mathcal{R}]_0(x) = \begin{cases} \mathcal{R}(x) & \text{if } x > 0, \\ \sum_{t \leq 0} \mathcal{R}(t) & \text{if } x = 0, \\ 0 & \text{if } x < 0. \end{cases}$$

We observe that if $A \leq C$ then $[A]_0 \leq [C]_0$ for any distributions A and C on $\mathbb{Z}$. It will also be convenient to introduce the shift operators: for a distribution $\mathcal{R} : \mathbb{Z} \to \mathbb{R}$ and an integer k, we define $S^k\mathcal{R}$ to be the distribution given by the rule $S^k\mathcal{R}(x) = \mathcal{R}(x-k)$. With these operators in place, we may write

$$\mathcal{R}_t = \left(\frac{1-\epsilon}{2}\right) S^1\mathcal{R}_{t-1} + \left(\frac{1+\epsilon}{2}\right) \left[S^{-1}\mathcal{R}_{t-1}\right]_0,$$

with the understanding that $\mathcal{R}_0$ is the distribution placing unit probability at 0. The proof now proceeds by induction. It is clear that $\mathcal{R}_0 \leq \mathcal{R}_\infty$. Assuming that $\mathcal{R}_n \leq \mathcal{R}_\infty$, we note that for any k

$$S^k\mathcal{R}_n \leq S^k\mathcal{R}_\infty \qquad \text{and, additionally, that} \qquad [S^{-1}\mathcal{R}_n]_0 \leq [S^{-1}\mathcal{R}_\infty]_0 .$$

Finally, it is clear that stochastic dominance respects convex combinations, in the sense that if $A_1 \leq C_1$ and $A_2 \leq C_2$ then $\lambda A_1 + (1-\lambda)A_2 \leq \lambda C_1 + (1-\lambda)C_2$ (for $0 \leq \lambda \leq 1$). We conclude that

$$\mathcal{R}_{t+1} = \left(\frac{1-\epsilon}{2}\right) S^1\mathcal{R}_t + \left(\frac{1+\epsilon}{2}\right) \left[S^{-1}\mathcal{R}_t\right]_0 \leq \left(\frac{1-\epsilon}{2}\right) S^1\mathcal{R}_\infty + \left(\frac{1+\epsilon}{2}\right) \left[S^{-1}\mathcal{R}_\infty\right]_0 .$$

By inspection, the right-hand side equals $\mathcal{R}_\infty$, as desired. Hence $\rho(B) \leq R_\infty$.

□

Remark. In fact, the random variable $\rho(B)$ actually converges to R_∞ as $n \to \infty$. This can be seen, for example, by solving for the stationary distribution of the Markov chain in the proof above. However, we will only require the dominance for our exposition. Importantly, since $\mu_x(\varepsilon) = \rho(x)$, and $\Pr[\mu_x(y) \geq 0]$ increases monotonically with an increase in $\Pr[\mu_x(\varepsilon) \geq r]$ for any $r \geq 0$, it suffices to take $|x| \to \infty$ when reasoning about an uppor bound on $\Pr[\mu_x(y) \geq 0]$.

6.3 Proof of Bound 1

Anticipating the proof, we make a few remarks about generating functions and stochastic dominance. We reserve the term *generating function* to refer to an "ordinary" generating function which represents a sequence $a_0, a_1, \ldots$ of non-negative real numbers by the formal power series $\mathsf{A}(Z) = \sum_{t=0}^{\infty} a_t Z^t$. When $\mathsf{A}(1) = \sum_t a_t = 1$ we say that the generating function is a *probability generating function*; in this case, the generating function A can naturally be associated with the integer-valued random variable A for which $\Pr[A = k] = a_k$. If the probability generating functions A and B are associated with the random variables A and B, it is easy to check that $\mathsf{A} \cdot \mathsf{B}$ is the generating function associated with the convolution $A + B$ (where A and B are assumed to be independent). Translating the notion of stochastic dominance to the setting with generating functions, we say that the generating function A *stochastically dominates* B if $\sum_{t \leq T} a_t \leq \sum_{t \leq T} b_t$ for all $T \geq 0$; we write $\mathsf{B} \preceq \mathsf{A}$ to denote this state of affairs. If $\mathsf{B}_1 \preceq \mathsf{A}_1$ and $\mathsf{B}_2 \preceq \mathsf{A}_2$ then $\mathsf{B}_1 \cdot \mathsf{B}_2 \preceq \mathsf{A}_1 \cdot \mathsf{A}_2$ and $\alpha \mathsf{B}_1 + \beta \mathsf{B}_2 \preceq \alpha \mathsf{A}_1 + \beta \mathsf{A}_2$ (for any $\alpha, \beta \geq 0$). Moreover, if $\mathsf{B} \preceq \mathsf{A}$ then it can be checked that $\mathsf{B}(\mathsf{C}) \preceq \mathsf{A}(\mathsf{C})$ for any probability generating function $\mathsf{C}(Z)$, where we write $\mathsf{A}(\mathsf{C})$ to denote the composition $\mathsf{A}(\mathsf{C}(Z))$.

Finally, we remark that if $\mathsf{A}(Z)$ is a generating function which converges as a function of a complex Z for $|Z| < R$ for some non-negative R, R is called the *radius of convergence* of A. It follows from [25, Theorem 2.19] that $\lim_{k \to \infty} a_k R^k = 0$ and $|a_k| = O(R^{-k})$. In addition, if A is a probability generating function associated with the random variable A then it follows that $\Pr[A \geq T] = O(R^{-T})$.

We define $p = (1 - \epsilon)/2$ and $q = 1 - p$ and as in the proof of Bound 2, consider the independent $\{0, 1\}$-valued random variables $w_1, w_2, \ldots$ where $\Pr[w_t = 1] = p$. We also define the associated $\{\pm 1\}$-valued random variables $W_t = (-1)^{1 + w_t}$.

Although our actual interest is in the random variable $\mu_x(y)$ from (13) on a characteristic string $w = xy$, we begin by analyzing the case when $|x| = 0$.

Case 1: x is the empty string. In this case, the random variable $\mu_x(y)$ is identical to $\mu(w)$ from (12) with $w = y$. Our strategy is to study the probability generating function

$$\mathsf{L}(Z) = \sum_{t=0}^{\infty} \ell_t Z^t$$

where $\ell_t = \Pr[t \text{ is the last time } \mu_t = 0]$. Controlling the decay of the coefficients ℓ_t suffices to give a bound on the probability that $w_1 \ldots w_k$ is forkable because

$$\Pr[w_1 \ldots w_k \text{ is forkable}] \leq 1 - \sum_{t=0}^{k-1} \ell_t = \sum_{t=k}^{\infty} \ell_t \,.$$

It seems challenging to give a closed-form algebraic expression for the generating function L; our approach is to develop a closed-form expression for a probability generating function $\hat{\mathsf{L}} = \sum_t \hat{\ell}_t Z^t$ which stochastically dominates L and apply the analytic properties of this closed form to bound the partial sums $\sum_{t \geq k} \hat{\ell}_k$. Observe that if $\mathsf{L} \preceq \hat{\mathsf{L}}$ then the series $\hat{\mathsf{L}}$ gives rise to an upper bound on the probability that $w_1 \ldots w_k$ is forkable as $\sum_{t=k}^{\infty} \ell_t \leq \sum_{t=k}^{\infty} \hat{\ell}_t$.

The coupled random variables ρ_t and μ_t are Markovian in the sense that values (ρ_s, μ_s) for $s \geq t$ are entirely determined by (ρ_t, μ_t) and the subsequent values $W_{t+1}, \ldots$ of the underlying variables W_i. We organize the sequence $(\rho_0, \mu_0), (\rho_1, \mu_1), \ldots$ into "epochs" punctuated by those times t for which $\rho_t = \mu_t = 0$. With this in mind, we define $\mathsf{M}(Z) = \sum m_t Z^t$ to be the generating function for the first completion of such an epoch, corresponding to the least $t > 0$ for which $\rho_t = \mu_t = 0$. As we discuss below, $\mathsf{M}(Z)$ is not a probability generating function, but rather $\mathsf{M}(1) = 1 - \epsilon$. It follows that

$$\mathsf{L}(Z) = \epsilon(1 + \mathsf{M}(Z) + \mathsf{M}(Z)^2 + \cdots) = \frac{\epsilon}{1 - \mathsf{M}(Z)} \,. \tag{16}$$

Below we develop an analytic expression for a generating function $\hat{\mathsf{M}}$ for which $\mathsf{M} \preceq \hat{\mathsf{M}}$ and define $\hat{\mathsf{L}} = \epsilon/(1 - \hat{\mathsf{M}}(Z))$. We then proceed as outlined above, noting that $\mathsf{L} \preceq \hat{\mathsf{L}}$ and using the asymptotics of $\hat{\mathsf{L}}$ to upper bound the probability that a string is forkable.

In preparation for defining $\hat{\mathsf{M}}$, we set down two elementary generating functions for the "descent" and "ascent" stopping times. Treating the random variables $W_1, \ldots$ as defining a (negatively) biased random walk, define D to be the generating function for the *descent stopping time* of the walk; this is the first time the random walk, starting at 0, visits -1. The natural recursive formulation of the descent time yields a simple algebraic equation for the descent generating function, $\mathsf{D}(Z) = qZ + pZ\mathsf{D}(Z)^2$, and from this we may conclude

$$\mathsf{D}(Z) = \frac{1 - \sqrt{1 - 4pqZ^2}}{2pZ} .$$

We likewise consider the generating function $\mathsf{A}(Z)$ for the *ascent stopping time*, associated with the first time the walk, starting at 0, visits 1: we have $\mathsf{A}(Z) = pZ + qZ\mathsf{A}(Z)^2$ and

$$\mathsf{A}(Z) = \frac{1 - \sqrt{1 - 4pqZ^2}}{2qZ} .$$

Note that while D is a probability generating function, the generating function A is not: according to the classical "gambler's ruin" analysis [12], the probability that a negatively-biased random walk starting at 0 ever rises to 1 is exactly p/q; thus $\mathsf{A}(1) = p/q$.

Returning to the generating function M above, we note that an epoch can have one of two "shapes": in the first case, the epoch is given by a walk for which $W_1 = 1$ followed by a descent (so that ρ returns to zero); in the second case, the epoch is given by a walk for which $W_1 = -1$, followed by an ascent (so that μ returns to zero), followed by the eventual return of ρ to 0. Considering that when $\rho_t > 0$ it will return to zero in the future almost surely, it follows that the probability that such a biased random walk will complete an epoch is $p + q(p/q) = 2p = 1 - \epsilon$, as mentioned in the discussion of (16) above. One technical difficulty arising in a complete analysis of M concerns the second case discussed above: while the distribution of the smallest $t > 0$ for which $\mu_t = 0$ is proportional to A above, the distribution of the smallest subsequent time t' for which $\rho_{t'} = 0$ depends on the value t. More specifically, the distribution of the return time depends on the value of ρ_t. Considering that $\rho_t \leq t$, however, this conditional distribution (of the return time of ρ to zero conditioned on t) is stochastically dominated by D^t, the time to descend t steps. This yields the following generating function $\hat{\mathsf{M}}$ which, as described, stochastically dominates M:

$$\hat{\mathsf{M}}(Z) = pZ \cdot \mathsf{D}(Z) + qZ \cdot \mathsf{D}(Z) \cdot \mathsf{A}(Z \cdot \mathsf{D}(Z)) .$$

It remains to establish a bound on the radius of convergence of $\hat{\mathsf{L}}$. Recall that if the radius of convergence of $\hat{\mathsf{L}}$ is $\exp(\delta)$ it follows that $\Pr[w_1 \ldots w_k \text{ is forkable}] = O(\exp(-\delta k))$. A sufficient condition for convergence of $\hat{\mathsf{L}}(z) = \epsilon/(1 - \hat{\mathsf{M}}(z))$ at z is that that all generating functions appearing in the definition of $\hat{\mathsf{M}}$ converge at z and that the resulting value $\hat{\mathsf{M}}(z) < 1$.

The generating function $\mathsf{D}(z)$ (and $\mathsf{A}(z)$) converges when the discriminant $1 - 4pqz^2$ is positive; equivalently $|z| < 1/\sqrt{1-\epsilon^2}$ or $|z| < 1 + \epsilon^2/2 + O(\epsilon^4)$. Considering $\hat{\mathsf{M}}$, it remains to determine when the second term, $qzD(z)\mathsf{A}(z\mathsf{D}(z))$, converges; this is likewise determined by positivity of the discriminant, which is to say that

$$1 - (1 - \epsilon^2)\left(\frac{1 - \sqrt{1 - (1 - \epsilon^2)z^2}}{1 - \epsilon}\right)^2 > 0 .$$

Equivalently,

$$|z| < \sqrt{\frac{1}{1+\epsilon}\left(\frac{2}{\sqrt{1-\epsilon^2}} - \frac{1}{1+\epsilon}\right)} = 1 + \epsilon^3/2 + O(\epsilon^4) .$$

Note that when the series $pz \cdot \mathsf{D}(z)$ converges, it converges to a value less than $1/2$; the same is true of $qz \cdot \mathsf{A}(z)$. It follows that for $|z| = 1 + \epsilon^3/2 + O(\epsilon^4)$, $|\hat{\mathsf{M}}(z)| < 1$ and $\hat{\mathsf{L}}(z)$ converges, as desired. We conclude that

$$\Pr[w_1 \ldots w_k \text{ is forkable}] = \exp(-\epsilon^3(1 + O(\epsilon))k/2) . \tag{17}$$

Case 2: x is non-empty. The relative margin before y begins is $\mu_x(\varepsilon)$. Recalling that $\mu_x(\varepsilon) = \rho(x)$ and conditioning on the event that $\rho(x) = r$, let us define the random variables $\{\tilde{\mu}_t\}$ for $t = 0, 1, 2, \cdots$ as follows: $\tilde{\mu}_0 = \rho(x)$ and

$$\Pr[\tilde{\mu}_t = s] = \Pr[\mu_x(y) = s \mid, \rho(x) = r \text{ and } |y| = t] .$$

If the $\tilde{\mu}$ random walk makes the rth descent at some time $t < n$, then $\tilde{\mu}_t = 0$ and the remainder of the walk is identical to an $(k-t)$-step μ random walk which we have already analyzed. Hence we investigate the probability generating function

$$\mathsf{B}_r(Z) = \mathsf{D}(Z)^r \mathsf{L}(Z) \quad \text{with coefficients} \quad b_t^{(r)} := \Pr[t \text{ is the last time } \tilde{\mu}_t = 0 \mid \tilde{\mu}_0 = r]$$

where $t = 0, 1, 2, \cdots$. Our interest lies in the quantity

$$b_t := \Pr[t \text{ is the last time } \tilde{\mu}_t = 0] = \sum_{r \geq 0} b_t^{(r)} \mathcal{R}_m(r) ,$$

where the *reach distribution* $\mathcal{R}_m : \mathbb{Z} \to [0, 1]$ associated with the random variable $\rho(x), |x| = m$ is defined as

$$\mathcal{R}_m(r) = \Pr_{x : |x| = m}[\rho(x) = r] . \tag{18}$$

Let $\mathsf{R}_m(Z)$ be the probability generating function for the distribution $\mathcal{R}_m$. Using Lemma 4 and Definition 15, we deduce that $\mathsf{R}_m \preceq \mathsf{R}_\infty$ for every $m \geq 0$ since $\mathcal{R}_m \preceq \mathcal{R}_\infty$. In addition, it is easy to check from (15) that the probability generating function for $\mathcal{R}_\infty$ is in fact $\mathsf{R}_\infty(Z) = (1-\beta)/(1-\beta Z)$ where $\beta := (1-\epsilon)/(1+\epsilon)$. Thus the generating function corresponding to the probabilities $\{b_t\}_{t=0}^{\infty}$ is

$$\begin{aligned}
\mathsf{B}(Z) &= \sum_{t=0}^{\infty} b_t Z^t = \sum_{r=0}^{\infty} \mathcal{R}_m(r) \sum_{t=0}^{\infty} b_t^{(r)} Z^t = \sum_{r=0}^{\infty} \mathcal{R}_m(r) \mathsf{B}_r(Z) \\
&= \mathsf{L}(Z) \sum_{r=0}^{\infty} \mathcal{R}_m(r) \mathsf{D}(Z)^r = \mathsf{L}(Z)\, \mathsf{R}_m(\mathsf{D}(Z)) \preceq \hat{\mathsf{L}}(Z)\, \mathsf{R}_\infty(\mathsf{D}(Z)) \\
&= \frac{(1-\beta)\, \hat{\mathsf{L}}(Z)}{1 - \beta \mathsf{D}(Z)} .
\end{aligned} \tag{19}$$

The dominance notation above follows because $\mathsf{L} \preceq \hat{\mathsf{L}}$ and $\mathsf{R}_m \preceq \mathsf{R}_\infty$.

For $\mathsf{B}(Z)$ to converge, we need to check that $\mathsf{D}(Z)$ should never converge to $1/\beta$. One can easily check that the radius of convergence of $\mathsf{D}(Z)$—which is $\sqrt{1-\epsilon^2}$—is strictly less than $1/\beta$ when $\epsilon > 0$. We conclude that $\mathsf{B}(Z)$ converges if both $\mathsf{D}(Z)$ and $\mathsf{L}(Z)$ converge. The radius of convergence of $\mathsf{B}(Z)$ would be the smaller of the radii of convergence of $\mathsf{D}(Z)$ and $\mathsf{L}(Z)$. We already know from the previous analysis that $\hat{\mathsf{L}}(Z)$ has the smaller radius of the two; therefore, the bound in (17) applies to the relative margin $\mu_x(y)$ for $|x| \geq 0$. ◻

6.4 Proof of Bound 2

Let $\epsilon \in (0, 1)$, $W \in \{0,1\}^m, W' \in \{0,1\}^k$ where both $(W_1, \ldots, W_n)$ and $(W'_1, \ldots, W'_n)$ satisfy the ϵ-martingale condition. Let $B \in \{0,1\}^m, B' \in \{0,1\}^k$ where the components of B, B' are independent with expectation $(1-\epsilon)/2$. By Lemma 4,

$$W \preceq B \quad \text{and} \quad W' \preceq B' . \tag{$*$}$$

Let us define the partial order $\leq$ on Boolean strings $\{0,1\}^k, k \in \mathbb{N}$ as follows: $a \leq b$ if and only if for all $i \in [k]$, $a_i = 1$ implies $b_i = 1$. Let $\mu : \{0,1\}^k \to \mathbb{Z}$ be the margin function from Lemma 3. Observe that for Boolean strings a, a', b, b' with $|a| = |a'|$ and $|b| = |b'|$, (i.) $b \leq b'$ implies $\mu_a(b) \leq \mu_a(b')$ and (ii.) $a \leq a'$ implies $\mu_a(b) \leq \mu_{a'}(b)$. That is,

$$\mu_a(b) \text{ is non-decreasing in both } a \text{ and } b . \tag{$\dagger$}$$

Using ($*$) and ($\dagger$), it follows that $\mu_W(W') \preceq \mu_B(B')$. Writing $x = W$ and $y = W'$, we have

$$\Pr[\mu_x(y) \geq 0] = \Pr[\mu_W(W') \geq 0] \leq \Pr[\mu_B(B') \geq 0]$$

where the inequality comes from the definition of stochastic dominance. A bound on the right-hand side is obtained in Bound 1. □

In Appendix B, we present a weaker bound on $\Pr[\mu_x(y) \geq 0]$ where the sequence $x_1, \ldots, x_m, y_1, \ldots, y_k$ satisfies ϵ-martingale conditions. The proof directly uses the properties of the martingale and Azuma's inequality but it does not use a stochastic dominance argument. Although it gives a bound of $3\exp\left(-\epsilon^4(1-O(\epsilon))k/64\right)$, the reader might find the proof of independent interest.

6.5 Proof of main theorems

Proof of Theorem 1. Let us start with the following observation. It allows us to formulate the (s,k)-settlement insecurity of a distribution $\mathcal{D}$ directly in terms of the relative margin.

Lemma 5. *Let $s, k, T \in \mathbb{N}$. Let $\mathcal{D}$ be any distribution on $\{0,1\}^T$. Then*

$$\mathbf{S}^{s,k}[\mathcal{D}] \leq \Pr_{w\sim\mathcal{D}}\left[\begin{matrix}\textit{there is a decomposition } w = xyz, \textit{ where}\\ |x| = s-1 \textit{ and } |y| \geq k+1\textit{, so that } \mu_x(y) \geq 0\end{matrix}\right].$$

Proof. Lemma 1 implies that $\mathbf{S}^{s,k}[\mathcal{D}]$ is no more than the probability that slot s is not k-settled for the characteristic string w. By Observation 1, this probability, in turn, is no more than the probability that there exists an x-balanced fork $F \vdash xy$ where we write $w = xyz, |x| = s-1, |y| \geq k+1, |z| \geq 0$. Finally, Fact 1 states that for any characteristic string xy, the two events "exists an x-balanced fork $F \vdash xy$" and "$\mu_x(y)$ is non-negative" have the same measure. Hence the claim follows. □

If the distribution $\mathcal{D}$ in the lemma above satisfies the ϵ-martingale condition, the probability in this lemma is no more than the probability in the left-hand side of Corollary 1. Finally, by retracing the proof of Corollary 1 using the explicit probability from Bound 2, we see that the bound in Corollary 1 is $O(1)\cdot\exp\bigl(-\Omega(\epsilon^3(1-O(\epsilon))k)\bigr)$. Since $\mathcal{B}_\epsilon$ satisfies the ϵ-martingale condition, we conclude that $\mathbf{S}^{s,k}[\mathcal{B}_\epsilon]$ is no more than this quantity as well.

For any player playing the settlement game, the set of strings on which the player wins is monotone with respect to the partial order $\leq$ defined in Section 6.4. To see why, note that if the adversary wins with a specific string w, he can certainly win with any string w' where $w \leq w'$. As $\mathcal{B}_\epsilon$ stochastically dominates $\mathcal{W}$, it follows that $\mathbf{S}^{s,k}[\mathcal{W}] \leq \mathbf{S}^{s,k}[\mathcal{B}_\epsilon]$. □

Proof of Theorem 2 For the first inequality, observe that if w violates k-CP, it must violate k-CP$^{\mathsf{slot}}$ as well. It remains to prove the second inequality. Let $\mathcal{D}$ be any distribution on $\{0,1\}^T$. We can apply Fact 1 on the statement of Theorem 3 to deduce that

$$\Pr_{w\sim\mathcal{D}}[w \text{ violates } k\text{-CP}^{\mathsf{slot}}] \leq \Pr_{w\sim\mathcal{D}}\left[\begin{matrix}\text{there is a decomposition } w = xyz,\\ \text{where } |y| \geq k \text{, so that } \mu_x(y) \geq 0\end{matrix}\right].$$

By using a union bound over $|x|$, the above probability is at most

$$\sum_{s=1}^{T-k+1} \Pr_{w}\left[\begin{matrix}\text{there is a decomposition } w = xyz, \text{ where}\\ |x| = s-1 \text{ and } |y| \geq k\text{, so that } \mu_x(y) \geq 0\end{matrix}\right].$$

Since w satisfies the ϵ-martingale condition, we can upper bound the probability inside the sum using Corollary 1. As we have seen in the proof of Theorem 1, the bound in Corollary 1 is $O(1)\cdot\exp\bigl(-\Omega(\epsilon^3(1-O(\epsilon))k)\bigr)$. It follows that the sum above is at most $T\exp\bigl(-\Omega(\epsilon^3(1-O(\epsilon))k)\bigr)$. □

It remains to prove the recursive formulation of the relative margin from Section 5; we tackle it in the next section.

7 Relative margin proofs and algorithm

As promised, we now step back to consider the recursive definition of relative margin in detail (and prove Lemma 3). Our analysis proceeds by cases. In each case, we prove an upper bound on relative margin in slot sl_{i+1} given the state of

Algorithm $\mathcal{A}$

1. Fix $w \in \{0,1\}^n$, and set $F_0 \vdash \varepsilon$ to the trivial fork (consisting of a single node corresponding to the genesis block).
2. For each slot $sl_i = \{sl_1, \ldots, sl_n\}$ in increasing order:
   (a) Set $S \leftarrow \{t \in F_i : \text{reach}(t) = 0\}$; $\hat{t} \leftarrow \arg\max_{t \in F_i} \text{length}(t)$; $t_1 \leftarrow \arg\max_{t \in F_i} \text{reach}(t)$
   (b) If $w_i = 0$:
      - If S is nonempty: select the tine $t \in S$ that diverges from t_1 earliest in w. Determine F_{i+1} from F_i by appending to t a chain of $\text{gap}(t)$ dishonest nodes followed by one honest node.
      - If S is empty: determine F_{i+1} from F_i by appending one honest node to $\hat{t}$.
   (c) Else if $w_i = 1$: set F_{i+1} equal to F_i.

the system in slot sl_i. Furthermore, we give a simple strategy for the settlement game that achieves this upper bound in each slot.

The analysis of [13] defines an adversary that maximizes margin by processing a characteristic string. In particular, the authors show that she that will always succeed in building a balanced fork, if one exists for that characteristic string. This result naturally suggests the possibility of an adversary who maximizes relative margin and builds x-balanced forks. An adversary that is able to maximize relative margin would also be able to optimally position herself to cause k-settlement violations (cf. our discussion of x-balanced forks, relative margin, and settlement violations in Section 4).

In this section, we will define a new adversary $\mathcal{A}$ who seeks to maximize $\mu_x(y)$. Furthermore, we will prove that this new adversary is able to simultaneously maximize $\mu_x(y)$ for *all* possible decompositions of a characteristic string w into components $w = xy$.

In order to strengthen our results (and match the setting used for the settlement game), we assume that $\mathcal{A}$ is given w in its entirety ahead of time; however, we note that $\mathcal{A}$ does not need this information in order to maximize relative margin. In fact, $\mathcal{A}$'s decision-making at each time slot depends only on whether the current slot is honest or dishonest, the current relative margin, and the current fork.

Notation. We again adopt the notation used in [13], with a few additions: let t_1 and t_2 be the disjoint tines of F for which $\rho(F) = \text{reach}(t_1)$ and $\mu(F) = \text{reach}(t_2)$, and let $\hat{t}$ be the longest tine of F. Finally, let S represent the set of tines t of F such that $\text{reach}(t) = 0$. (We will sometimes refer to such tines as *critical tines*.)

Overview of the algorithm. In order to maximize relative margin, the adversary uses the following reasoning (formally described in Figure 7). During each slot, she learns whether the next slot is honest or dishonest, and decides what (if any) changes to make to the fork F. If the next token of the characteristic string is a 1, she makes no changes to F and simply bides her time. If the next token of the characteristic string is a 0, she looks for tines with reach precisely 0. Any such tine (1) is close enough in length to the longest tine that she can expend her reserve to catch up, and (2) may fall too far behind if she waits any longer, so she will choose to add new blocks to some critical tine $t \in S$. (If there is more than one tine with reach 0, she selects $t \in S$ that branches from t_1 earliest in the fork.) She appends $\text{gap}(t)$ dishonest nodes to t before adding an honest node to the end. If there are no tines with reach exactly 0, she simply appends an honest node to the longest tine, $\hat{t}$. (Note that this description is consistent with the behavior of an adversary and challenger in the settlement game, but for the sake of brevity the adversary plays both roles here.) We will prove that this strategy maximizes relative margin below.

Recall from Section 2 the definition of fork prefixes (denoted $\sqsubseteq$):

Definition 16 (Fork prefixes)**.** *Given $w \in \{0,1\}^*$ and some prefix $x \subseteq w$, we say the fork $F \vdash x$ is a* prefix *of $F' \vdash w$ if F is a consistently labeled subgraph of F' (i.e., all vertices and edges of F also appear in F', and the label of any vertex appearing in both F and F' is identical). We denote this relationship by $F \sqsubseteq F'$.*

Pairs of forks $F_{i-1} \vdash w_1 \dots w_{i-1}$ and $F_i \vdash w_1 \dots w_i$ (such that $F_{i-1} \sqsubseteq F_i$) are particularly relevant to our analysis, because we are interested in how the fork changes with each new slot. Suppose $w_i = 0$, i.e., slot sl_i is honest, and consider the closed prefixes of F_{i-1} and F_i. (Recall that a closed fork is derived by "pruning" any adversarial nodes from the end of tines until all leaves are honest.) The graphs of these closed prefixes differ by either a single honest node, or a chain of dishonest nodes—an "adversarial augmentation," in the language of the settlement game—followed by a single honest node. We call the tine of F_i that contains all the new nodes an *extension*, since it is the tine that the adversary has chosen to extend and present to the challenger during that slot. In other words, an extension consists of the "new" honest node added by the challenger, as well as any adversarial nodes beneath the new honest node that did not appear in the closed fork from the previous slot. When considering closed forks, we observe that there is a *unique* extension associated with each honest slot.

We can immediately derive two useful results related to extensions. As in [13], we use the notation $\mathrm{reach}_F(t)$ (or $\mathrm{reserve}_F(t)$, or $\mathrm{gap}_F(t)$, etc.) to indicate the reach (or reserve, or gap) of the tine t in the context of a particular fork F. (This notation is especially useful when discussing properties of t as it appeared in two different forks.)

Claim 1 (Reach of extended tines)**.** *Consider a closed fork $F \vdash w$ and some closed fork $F' \vdash w0$ such that $F \sqsubseteq F'$. If a tine t of F' is an extension, then $\mathrm{reach}_{F'}(t) = 0$.*

Proof. We have assumed that t is an extension, so its terminal vertex must be the new honest node. By definition, $\mathrm{reach}_{F'}(t) = \mathrm{reserve}_{F'}(t) - \mathrm{gap}_{F'}(t)$. Honest players will only place nodes at a depth strictly greater than all other honest nodes, so we infer that t is the longest tine of F', and so $\mathrm{gap}_{F'}(t) = 0$. Moreover, we observe that there are no 1s occurring after this point in the characteristic string, and so $\mathrm{reserve}_{F'}(t) = 0$. Plugging these values into our definition of reach we see that $\mathrm{reach}_{F'}(t) = 0 - 0 = 0$. □

Claim 2 (Reach of non-extended tines)**.** *Consider a closed fork $F \vdash w$ and some closed fork $F' \vdash w0$ such that $F \sqsubseteq F'$. If a tine t of F' did not arise from extension, i.e., it existed in F, then $\mathrm{reach}_{F'}(t) < \mathrm{reach}_F(t)$.*

Proof. Definitionally, we know that $\mathrm{reach}_{F'}(t) = \mathrm{reserve}_{F'}(t) - \mathrm{gap}_{F'}(t)$. From F to F', the length of the longest tine increases, and the length of t does not change, so we observe that $\mathrm{gap}_{F'}(t) > \mathrm{gap}_F(t)$. The reserve of t does not change, because there are no new 1s in the characteristic string. Therefore,

$$\mathrm{reach}_{F'}(t) = \mathrm{reserve}_{F'}(t) - \mathrm{gap}_{F'}(t) < \mathrm{reserve}_F(t) - \mathrm{gap}_F(t) = \mathrm{reach}_F(t)\,.$$ □

Now we are ready to proceed with our proof of Lemma 3. The structure of the proof closely follows the analogous proof for the recursive definition of margin given in Lemma 4.19 of [13]; however, it incorporates the definition and analysis of the new adversary.

Proof of Lemma 3. Let F be a fork for the characteristic string xy. In the base case, where $y = \varepsilon$, we observe that any two tines of F are disjoint over y. Moreover, even a single tine t_1 is disjoint with itself over ε! Therefore, the relative margin $\mu_x(\varepsilon)$ must be greater than or equal to the reach of the tine t that achieves reach$(t) = \rho(x)$. The relative margin must also be less than or equal to $\rho(x)$, because that is, by definition, the maximum reach over all tines in all forks $F \vdash w$. Putting these facts together, we have $\mu_x(\varepsilon) = \rho(x)$.

Moving beyond the base case, we will consider a pair of closed forks $F \vdash xy$ and $F' \vdash xyc$ such that $F \sqsubseteq F'$, $x, y \in \{0,1\}^*$, y is nonempty, and $c \in \{0,1\}$.

Suppose the next slot is dishonest ($c = 1$). Then F' must necessarily be equal to F, because we are dealing with closed forks and have not introduced any new honest nodes. The reach of each tine increases by 1 from F to F' because the gap has not changed and reserve has increased by one. Therefore, $\mu_x(y1) = \mu_x(y) + 1$, as desired.

If instead, the next slot is honest ($c = 0$), there are more possibilities to consider. We will break this part of the proof into several cases based on the relative reach and margin of the fork. In each case, we will prove the lower bound by showing how the adversary $\mathcal{A}$ can achieve some value of $\mu_x(y)$, and then use a proof by contradiction to show that this value is also the upper bound.

Case 1: $\rho(xy) > 0$ **and** $\mu_x(y) = 0$**.** Let F be some fork for xy such that $\rho(F) = \rho(xy)$ and $\mu_x(F) = \mu_x(y)$. $\mathcal{A}$ will build on some tine t with $\mathrm{reach}(t) = 0$, and break ties by choosing to extend the tine that branches from t_1 as early as possible. In fact, in this case we are guaranteed that any tine she chooses will diverge from t_1 prior to the beginning of y: because $\mu_x(y) = 0$, we know that the tine t_2 associated with $\mu_x(y)$ is disjoint with t_1 over y and is in the set of critical tines. Based on our description of $\mathcal{A}$, she will either build on t_2, or on another tine that diverges from t_1 even earlier, and is also disjoint with t_1 over y. This shows that any such extension guarantees $\mu_x(y0)$ is at least 0, as the extension and t_1 form a pair of tines disjoint over $y0$.

Now we will show the corresponding upper bound. Let F' be a closed fork for the characteristic string $w = xy0$ such that $\rho(F') = \rho(xy0)$ and $\mu_x(F') = \mu_x(y0)$, and let $F \vdash xy$ be the unique closed fork such that $F \sqsubseteq F'$. Let t_1 and t_2 be the tines of F' that achieve $\rho(xy0)$ and $\mu_x(y0)$, respectively. Suppose (toward a contradiction) that $\mu_x(y0) > 0$. Then neither t_1 or t_2 is an extension because, as we proved in Claim 1, extensions have reach exactly 0. This means that t_1 and t_2 existed in F, and had strictly greater reach in F than they do presently in F' (by Claim 2). Because t_1 and t_2 have been implicitly assumed to be disjoint over $y0$, they must also be disjoint over y; therefore the margin of F must be at least $\min\{\mathrm{reach}_F(t_1), \mathrm{reach}_F(t_2)\}$. Following this line of reasoning, we have

$$\mu_x(y) \geq \min\{\mathrm{reach}_F(t_i)\} > \min\{\mathrm{reach}_{F'}(t_i)\} = \mu_x(y0) > 0.$$

This contradicts our original assumption for the case, which states that $\mu_x(y) = 0$. We can conclude that $\mu_x(y0) \leq 0$, as desired.

Case 2: $\rho(xy) = 0$**.** We will analyze this case with the help of subcases based on the contents of S, the set of critical tines. If $S = \{t_1\}$, $\mathcal{A}$ will extend t_1. The resulting extension has reach 0, so $\rho(xy0) \geq 0$. Additionally, t_2's reach decreases by 1, and the extension and t_2 are still disjoint over y, so $\mu_x(y0) \geq \mu_x(y) - 1$. If S contains both t_1 and t_2, $\mathcal{A}$ extends t_2, because it is totally disjoint from t_1 over y and has reach 0. The extension still has reach 0, so $\rho(xy0) \geq 0$. Furthermore, the reach of t_1 decreases by 1, and the extension and t_1 are disjoint over y, so $\mu_x(y0) \geq \rho(xy) - 1 \geq \mu_x(y) - 1$. Lastly, if S contains some critical tine s distinct from t_1 but S does not contain t_2, $\mathcal{A}$ will extend s. The resulting extension of s has reach 0, so $\rho(xy0) \geq 0$. Note that because t_2 is not in S, $\mathrm{reach}(t_2) < 0$. This implies that s (and its extension) must share an edge with t_1 somewhere over y, as otherwise we would achieve $\mu_x(y) = 0$. As a result, t_2 and the extension of s must be disjoint over y, and they have reach $\mu_x(y) - 1$ and 0 respectively, so they act as witnesses to prove that $\mu_x(y0) \geq \mu_x(y) - 1$.

Next, we want to prove the corresponding upper bound. Suppose $F' \vdash xy0$ is a closed fork such that $\rho(xy0) = \rho(F')$ and $\mu_x(y0) = \mu_x(F')$, and let $F \vdash xy$ be the unique closed fork such that $F \sqsubseteq F'$. Define t_1, t_2 to be a pair of tines disjoint over y in F' such that $\mathrm{reach}_{F'}(t_1) = \rho(F')$ and $\mathrm{reach}_{F'}(t_2) = \mu_x(F') = \mu_x(y0)$. First, it will be helpful to determine some facts about t_1. Specifically, we claim that t_1 must be an extension. Suppose t_1 is not an extension. The fact that t_1 achieves maximum reach implies that t_1 has non-negative reach, because the longest tine always achieves reach 0, so t_1 must do at least as well as the longest tine. Furthermore, Claim 2 states that all tines other than the extended tine see their reach decrease. Therefore, if t_1 was not extended, then t_1 as it appeared in F must have had strictly positive reach. This contradicts the central assumption of the case, i.e., that $\rho(xy) = 0$. Therefore, we conclude that t_1 arose from extension.

Having established that t_1 must arise from extension, we know that the tine prefix of t_1 that is present in F must have reach of at least 0. Additionally, we have assumed $\rho(xy) = 0$, so $\mathrm{reach}_F(t_1) \leq 0$. Together, these statements tell us that $\mathrm{reach}_F(t_1) = 0$. Restricting our view to F, we see that t_1 (as it appeared in F) and t_2 are disjoint over y, and so it must be true that $\min\{\mathrm{reach}_F(t_1), \mathrm{reach}_F(t_2)\} \leq \mu_x(y)$. Because $\mathrm{reach}_F(t_1) = 0$ and $\mathrm{reach}_F(t_2) \leq \rho(xy) = 0$, we can simplify that statement to $\mathrm{reach}_F(t_2) \leq \mu_x(y)$. Finally, because t_2 was not extended from F to F', Claim 2 tells us that $\mathrm{reach}_{F'}(t_2) < \mathrm{reach}_F(t_2)$. Taken together, these two inequalities show that $\mathrm{reach}_{F'}(t_2) < \mathrm{reach}_F(t_2) \leq \mu_x(y)$. Reach is always an integer, and so $\mu_x(y0) = \mathrm{reach}_{F'}(t_2) < \mu_x(y)$ implies $\mu_x(y0) = \mathrm{reach}_{F'}(t_2) \leq \mu_x(y) - 1$, as desired.

Case 3: $\rho(xy) > 0, \mu_x(y) \neq 0$**.** Suppose by induction that we have $F \vdash xy$ and tines t_1, t_2 such that $\rho(xy) = \rho(F) = \mathrm{reach}_F(t_1)$ and $\mu_x(y) = \mu_x(F) = \mathrm{reach}_F(t_2)$. $\mathcal{A}$ will minimally extend a tine s with reach 0, if one exists, or $\hat{t}$. As a result of this extension, we know that $\mathrm{reach}_{F'}(t_i) = \mathrm{reach}_F(t_i) - 1$. The witnesses t_1 and t_2 will still be disjoint over $y0$, so $\mu_x(y0) \geq \mu_x(y) - 1$.

Now we need to prove the corresponding upper bound. Let $F' \vdash xy0$ be a closed fork such that $\mu_x(y0) = \mu_x(F')$, and let $F \vdash xy$ be the unique closed fork such that $F \sqsubseteq F'$. Additionally, let t_1 and t_2 be tines disjoint over y such that $\mathrm{reach}_{F'}(t_1) = \rho(F')$ and $\mathrm{reach}_{F'}(t_2) = \mu_x(y0)$. We will break this case into sub-cases. In the first sub-case, suppose that neither t_1 nor t_2 arose from extension. Then $\min\{\mathrm{reach}_F(t_1), \mathrm{reach}_F(t_2)\} \leq \mu_x(y)$, because t_1 and t_2 existed in F and must be disjoint over y (by virtue of being disjoint over $y0$). Furthermore, our claim about reach of non-extended tines implies that $\mathrm{reach}_{F'}(t_i) < \mathrm{reach}_F(t_i)$ for $i \in \{1, 2\}$. Therefore,

$$\mu_x(y0) = \min\{\mathrm{reach}_{F'}(t_1), \mathrm{reach}_{F'}(t_2)\} < \min\{\mathrm{reach}_F(t_1), \mathrm{reach}_F(t_2)\} \leq \mu_x(y),$$

as desired. For the second sub-case, suppose either t_1 or t_2 arose from extension. It must be true that $\mathrm{reach}_{F'}(t_2) \leq 0$, because either t_2 is the extension (and therefore has reach exactly 0) or t_1 is the extension and we have $\mathrm{reach}_{F'}(t_2) = \mu_x(y0) \leq \rho(xy0) = \mathrm{reach}_{F'}(t_1) = 0$. Recall that we have assumed $\mu_x(y) \neq 0$. If $\mu_x(y) > 0$, we are done: certainly $\mu_x(y0) \leq 0 < \mu_x(y)$. If, however, $\mu_x(y) < 0$, there is more work to do. Suppose $\mu_x(y) < 0$. In this case, it is not possible for t_2 to have been the extension. To see why, consider the following: if t_2 arose from extension, then it must have had some precursor in F with non-negative reach. Additionally, by our claim about non-extended tines, we see that $\mathrm{reach}_F(t_1) > \mathrm{reach}_{F'} \geq 0$. Therefore, t_1 and the precursor to t_2 would be a pair of tines that achieve margin greater than or equal to 0. By contradiction, t_2 cannot have arisen from extension, so we do not need to worry about this case. The last remaining scenario is the one in which $\mu_x(y) < 0$ and t_1 arises from extension. In this scenario, t_2 cannot have been the extension (since there is only one!) so we can invoke our claim about reach of non-extended tines once again to see that $\mathrm{reach}_F(t_2) > \mathrm{reach}_{F'}(t_2)$. Using a now-familiar line of reasoning, note that t_2 and t_1 (prior to its extension) are a valid choice for a pair of tines achieving margin in F, and therefore $\mathrm{reach}_F(t_2) \leq \mu_x(y)$. We now have $\mu_x(y) \geq \mathrm{reach}_F(t_2) > \mathrm{reach}_{F'}(t_2) = \mu_x(y0)$. Because reach is always an integer, the value of $\mu_x(y0)$ must be less than or equal to $\mu_x(y) - 1$, as desired. □

Observe that the lower bounds are derived by showing that the adversary $\mathcal{A}$ is able to achieve that value of $\mu_x(y)$ in each case. Because that value matches the upper bound, we know that $\mathcal{A}$ maximizes $\mu_x(y)$.

Perhaps surprisingly, this strategy allows our adversary to maximize relative reach and margin over *all* possible decompositions $w = xy$. This is because her strategy is independent of any particular decomposition; she will always build pairs of viable tines that are edge-disjoint over as much of the string as possible, which is the best she can hope to do with respect to any decomposition.

This adversary is closely tied to our investigation of the $(\mathcal{D}, T; s, k)$-settlement game. In fact, by maximizing relative margin, she can play the $(\mathcal{D}, T; s, k)$-settlement game optimally (i.e., win whenever there exists a winning configuration for the given challenge characteristic string). For the settlement game to be winnable, there must exist some $\hat{y}$, a prefix of y, so that $|\hat{y}| \geq k + 1$ and $\mu_x(\hat{y}) \geq 0$. Recall from 1 that for a characteristic string xy, there is an x-balanced fork $F \vdash xy$ if and only if $\mu_x(y) \geq 0$. Because in each slot, our adversary builds a fork that achieves the maximum value of relative margin for all possible decompositions, then at slot $|x\hat{y}|$ she will have built a fork $F \vdash x\hat{y}$ such that $\mu_x(F) \geq 0$. As we argued in the proof of Fact1, the definition of relative margin tells us that F has two tines t_1 and t_2 with nonnegative reach that diverge prior to the start of $\hat{y}$. Consequently, she is able to append $\mathrm{gap}(t_i)$ adversarial vertices from our reserve to each t_i so that they become maximum length, thus winning the $(\mathcal{D}, T; s, k)$-settlement game.

References

[1] Adam Back. Hashcash. http://www.cypherspace.org/hashcash, 1997.

[2] Christian Badertscher, Peter Gazi, Aggelos Kiayias, Alexander Russell, and Vassilis Zikas. Ouroboros genesis: Composable proof-of-stake blockchains with dynamic availability. *IACR Cryptology ePrint Archive*, 2018:378, 2018.

[3] Iddo Bentov, Ariel Gabizon, and Alex Mizrahi. Cryptocurrencies without proof of work. *CoRR*, abs/1406.5694, 2014.

[4] Iddo Bentov, Rafael Pass, and Elaine Shi. Snow white: Provably secure proofs of stake. *IACR Cryptology ePrint Archive*, 2016:919, 2016.

[5] Jonah Brown-Cohen, Arvind Narayanan, Christos-Alexandros Psomas, and S. Matthew Weinberg. Formal barriers to longest-chain proof-of-stake protocols. *CoRR*, abs/1809.06528, 2018.

[6] Bernardo David, Peter Gazi, Aggelos Kiayias, and Alexander Russell. Ouroboros praos: An adaptively-secure, semi-synchronous proof-of-stake blockchain. In Nielsen and Rijmen [19], pages 66–98.

[7] Cynthia Dwork and Moni Naor. Pricing via processing or combatting junk mail. In Ernest F. Brickell, editor, *Advances in Cryptology – CRYPTO'92*, volume 740 of *Lecture Notes in Computer Science*, pages 139–147, Santa Barbara, CA, USA, August 16–20, 1993. Springer, Heidelberg, Germany.

[8] Stefan Dziembowski, Sebastian Faust, Vladimir Kolmogorov, and Krzysztof Pietrzak. Proofs of space. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *Advances in Cryptology – CRYPTO 2015, Part II*, volume 9216 of *Lecture Notes in Computer Science*, pages 585–605, Santa Barbara, CA, USA, August 16–20, 2015. Springer, Heidelberg, Germany.

[9] Juan A. Garay, Aggelos Kiayias, and Nikos Leonardos. The bitcoin backbone protocol: Analysis and applications. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part II*, volume 9057 of *Lecture Notes in Computer Science*, pages 281–310. Springer, 2015.

[10] Juan A. Garay, Aggelos Kiayias, and Nikos Leonardos. The bitcoin backbone protocol with chains of variable difficulty. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017, Part I*, volume 10401 of *Lecture Notes in Computer Science*, pages 291–323, Santa Barbara, CA, USA, August 20–24, 2017. Springer, Heidelberg, Germany.

[11] Juan A. Garay, Aggelos Kiayias, and Nikos Leonardos. The bitcoin backbone protocol with chains of variable difficulty. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part I*, volume 10401 of *Lecture Notes in Computer Science*, pages 291–323. Springer, 2017.

[12] Charles M. Grinstead and J. Laurie Snell. *Introduction to Probability*. American Mathematical Association, 1997.

[13] Aggelos Kiayias, Alexander Russell, Bernardo David, and Roman Oliynykov. Ouroboros: A provably secure proof-of-stake blockchain protocol. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference*, volume 10401 of *Lecture Notes in Computer Science*, pages 357–388. Springer, 2017.

[14] David A Levin, Yuval Peres, and Elizabeth L Wilmer. *Markov chains and mixing times*, volume 58. American Mathematical Society, 2009.

[15] Silvio Micali. ALGORAND: the efficient and democratic ledger. *CoRR*, abs/1607.01341, 2016.

[16] Tal Moran and Ilan Orlov. Proofs of space-time and rational proofs of storage. Cryptology ePrint Archive, Report 2016/035, 2016. `http://eprint.iacr.org/2016/035`.

[17] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, New York, NY, USA, 1995.

[18] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. http://bitcoin.org/bitcoin.pdf, 2008.

[19] Jesper Buus Nielsen and Vincent Rijmen, editors. *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *Lecture Notes in Computer Science*, 2018. Springer.

[20] Sunoo Park, Krzysztof Pietrzak, Albert Kwon, Joël Alwen, Georg Fuchsbauer, and Peter Gazi. Spacemint: A cryptocurrency based on proofs of space. *IACR Cryptology ePrint Archive*, 2015:528, 2015.

[21] Rafael Pass and Elaine Shi. The sleepy model of consensus. In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part II*, volume 10625 of *Lecture Notes in Computer Science*, pages 380–409. Springer, 2017.

[22] Rafael Pass and Elaine Shi. Hybrid consensus: Efficient consensus in the permissionless model. In Andréa W. Richa, editor, *31st International Symposium on Distributed Computing, DISC 2017, October 16-20, 2017, Vienna, Austria*, volume 91 of *LIPIcs*, pages 39:1–39:16. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

[23] Rafael Pass and Elaine Shi. Thunderella: Blockchains with optimistic instant confirmation. In Nielsen and Rijmen [19], pages 3–33.

[24] Rafael Pass, Lior Seeman, and Abhi Shelat. Analysis of the blockchain protocol in asynchronous networks. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part II*, volume 10211 of *Lecture Notes in Computer Science*, pages 643–673, 2017.

[25] Herbert S Wilf. *generatingfunctionology*. AK Peters/CRC Press, 3 edition, 2005.

A Exact settlement probabilities

Let $m, k \in \mathbb{N}$ and $\epsilon \in (0, 1]$. Let w be a characteristic string of length $T = m + k$ such that the bits of w are i.i.d. Bernoulli with expectation $\alpha = (1 - \epsilon)/2$. Write w as $w = xy$ where $|x| = m, |y| = k$. The recursive definition of relative margin (cf. Lemma 3) implies an algorithm for computing the probability $\Pr[\mu_x(y) \geq 0]$ in time $\operatorname{poly}(m, k)$. In typical circumstances, however, it is more interesting to establish an explicit upper bound on $\Pr[\mu_x(y) \geq 0]$ where $|x| \to \infty$; this corresponds to the case where the distribution of the initial reach $\rho(x)$ is the dominant distribution $\mathcal{R}_\infty$ in Lemma 4. Due to dominance, $\mathcal{R}_\infty(m)$ serves as an upper bound on $\rho(x)$ for any finite $m = |x|$. For this purpose, one can implicitly maintain a sequence of matrices (M_t) for $t = 0, 1, 2, \cdots, k$ such that $M_0(r, r) = \mathcal{R}_\infty(r)$ for all $0 \leq r \leq 2k$ and the invariant

$$M_t(r, s) = \Pr_{y \sim \mathcal{B}(t, \alpha)}[\rho(xy) = r \text{ and } \mu_x(y) = s]$$

is satisfied for every integer $t \in [1, k]$, $r \in [0, 2k]$, and $s \in [-2k, 2k]$. Here, $M(i, j)$ denotes the entry at the ith row and jth column of the matrix M. Observe that $M_t(r, s)$ can be computed solely from the neighboring cells of M_{t-1}, that is, from the values $M_{t-1}(r \pm 1, s \pm 1)$. Of course, only the transitions approved by the recursions in Lemma 2 and Lemma 3 should be considered.

Finally, one can compute $\Pr[\mu_x(y) \geq 0]$ by summing $M_k(r, s)$ for $r, s \geq 0$. Table 1 contains these probabilities where α ranges from 0.05 to 0.40 and k ranges from 50 to 1000. In addition, Figure 4 shows the base-10 logarithm of these probabilities. The points corresponding to a fixed α appear to form a straight line. This means the probability decays exponentially in k, or equivalently, that the exponent depends linearly on k, as stipulated by Bound 1.

B A forkability bound for strings satisfying the ϵ-martingale condition

Below we present a bound (Bound 3) on the probability that a characteristic string satisfying the ϵ-martingale condition has a non-negative relative margin. We remark that the bound below is weaker than Bound 2. Before we proceed, recall the following standard large deviation bound for supermartingales.

Theorem 5 (Azuma's inequality (Azuma; Hoeffding). See [17, 4.16] for a discussion)**.** *Let $X_0, \ldots, X_n$ be a sequence of real-valued random variables so that, for all t, $\mathbb{E}[X_{t+1} \mid X_0, \ldots, X_t] \leq X_t$ and $|X_{t+1} - X_t| \leq c$ for some constant c. Then $\Pr[X_n - X_0 \geq \Lambda] \leq \exp\left(-\Lambda^2/2nc^2\right)$ for every $\Lambda \geq 0$.*

Table 1: Exact probabilities $\Pr[\mu_x(y) \geq 0]$ where the bits of the characteristic string xy are i.i.d. Bernoulli with expectation α. Each row of the table corresponds to a different $k = |y|$.

k	α							
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
50	5.37E-15	1.16E-09	1.02E-06	8.68E-05	1.96E-03	1.86E-02	9.36E-02	2.92E-01
100	1.23E-28	5.10E-18	3.52E-12	2.28E-08	1.03E-05	8.00E-04	1.72E-02	1.37E-01
150	2.83E-42	2.24E-26	1.22E-17	6.05E-12	5.54E-08	3.57E-05	3.30E-03	6.74E-02
200	6.49E-56	9.82E-35	4.21E-23	1.61E-15	2.98E-10	1.60E-06	6.40E-04	3.36E-02
250	1.49E-69	4.31E-43	1.46E-28	4.27E-19	1.61E-12	7.21E-08	1.25E-04	1.69E-02
300	3.42E-83	1.89E-51	5.05E-34	1.14E-22	8.67E-15	3.25E-09	2.44E-05	8.52E-03
350	7.84E-97	8.29E-60	1.75E-39	3.02E-26	4.67E-17	1.46E-10	4.78E-06	4.31E-03
400	1.80E-110	3.64E-68	6.06E-45	8.02E-30	2.52E-19	6.59E-12	9.37E-07	2.18E-03
450	4.13E-124	1.60E-76	2.10E-50	2.13E-33	1.36E-21	2.97E-13	1.84E-07	1.11E-03
500	9.47E-138	7.00E-85	7.26E-56	5.67E-37	7.32E-24	1.34E-14	3.60E-08	5.62E-04
550	2.17E-151	3.07E-93	2.51E-61	1.51E-40	3.95E-26	6.02E-16	7.05E-09	2.86E-04
600	4.98E-165	1.35E-101	8.70E-67	4.00E-44	2.13E-28	2.71E-17	1.38E-09	1.45E-04
650	1.14E-178	5.91E-110	3.01E-72	1.06E-47	1.15E-30	1.22E-18	2.71E-10	7.37E-05
700	2.62E-192	2.59E-118	1.04E-77	2.83E-51	6.19E-33	5.51E-20	5.31E-11	3.75E-05
750	6.02E-206	1.14E-126	3.61E-83	7.52E-55	3.33E-35	2.48E-21	1.04E-11	1.91E-05
800	1.38E-219	4.99E-135	1.25E-88	2.00E-58	1.80E-37	1.12E-22	2.04E-12	9.69E-06
850	3.17E-233	2.19E-143	4.33E-94	5.31E-62	9.69E-40	5.04E-24	4.00E-13	4.93E-06
900	7.27E-247	9.61E-152	1.50E-99	1.41E-65	5.23E-42	2.27E-25	7.84E-14	2.50E-06
950	1.67E-260	4.22E-160	5.19E-105	3.75E-69	2.82E-44	1.02E-26	1.54E-14	1.27E-06
1000	3.83E-274	1.85E-168	1.80E-110	9.98E-73	1.52E-46	4.61E-28	3.01E-15	6.48E-07

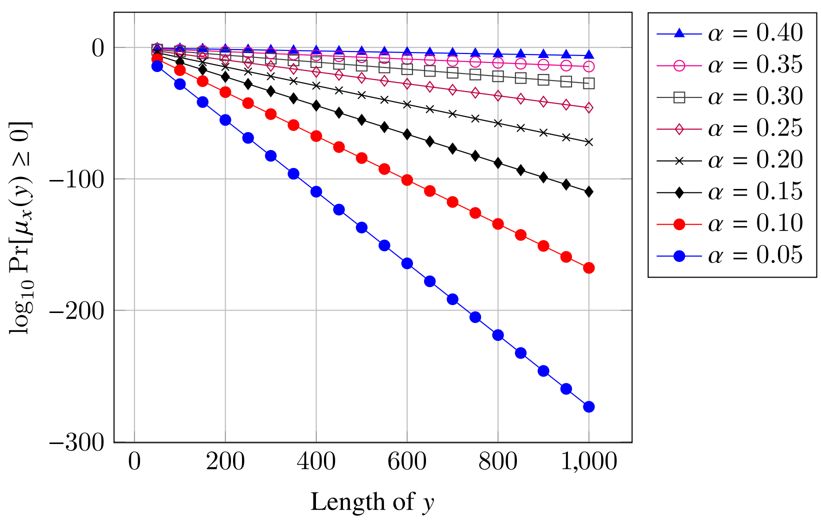


Figure 4: The probabilities from Table 1 drawn in the base-10 logarithmic scale.

Bound 3. *Let $x \in \{0,1\}^m$ and $y \in \{0,1\}^k$ be random variables, satisfying the ϵ-martingale condition (with respect to the ordering $x_1, \ldots, x_m, y_1, \ldots, y_k$). Then*

$$\Pr[\mu_x(y) \geq 0] \leq 3 \exp\left(-\epsilon^4(1 - O(\epsilon))k/64\right) .$$

Proof. Let $w_1, w_2, \ldots$ be random variables obeying the ϵ-martingale condition. Specifically, $\Pr[w_t = 1 \mid E] \leq (1-\epsilon)/2$

conditioned on any event E expressed in the variables $w_1, \ldots, w_{t-1}$. For convenience, define the associated $\{\pm 1\}$-valued random variables $W_t = (-1)^{1+w_t}$ and observe that $\mathbb{E}[W_t] \leq -\epsilon$.

If x is empty. Observe that in this case, the relative margin $\mu_x(y)$ reduces to the non-relative margin $\mu(y)$ from Lemma 2. Since the sequence $y_1, y_2, \ldots$ in the statement of the claim is identical to the sequence $w_1, w_2, \ldots$ defined above, we focus on the reach and margin of the latter sequence. Specifically, define $\rho_t = \rho(w_1 \ldots w_t)$ and $\mu_t = \mu(w_1 \ldots w_t)$ to be the two random variables from Lemma 2 acting on the string $w = w_1 \ldots w_t$. The analysis will rely on the ancillary random variables $\overline{\mu}_t = \min(0, \mu_t)$. Observe that $\Pr[w \text{ forkable}] = \Pr[\mu(w) \geq 0] = \Pr[\overline{\mu}_k = 0]$, so we may focus on the event that $\overline{\mu}_k = 0$. As an additional preparatory step, define the constant $\alpha = (1+\epsilon)/(2\epsilon) \geq 1$ and define the random variables $\Phi_t \in \mathbb{R}$ by the inner product

$$\Phi_t = (\rho_t, \overline{\mu}_t) \cdot \begin{pmatrix} 1 \\ \alpha \end{pmatrix} = \rho_t + \alpha \overline{\mu}_t \,.$$

The Φ_t will act as a "potential function" in the analysis: we will establish that $\Phi_k < 0$ with high probability and, considering that $\alpha \overline{\mu}_k \leq \rho_k + \alpha \overline{\mu}_k = \Phi_k$, this implies $\overline{\mu}_k < 0$, as desired.

Let $\Delta_t = \Phi_t - \Phi_{t-1}$; we claim that—conditioned on any fixed value (ρ, μ) for (ρ_t, μ_t)—the random variable $\Delta_{t+1} \in [-(1+\alpha), 1+\alpha]$ has expectation no more than $-\epsilon$. The analysis has four cases, depending on the various regimes of ρ and μ from Lemma 2. When $\rho > 0$ and $\mu < 0$, $\rho_{t+1} = \rho + W_{t+1}$ and $\overline{\mu}_{t+1} = \overline{\mu} + W_{t+1}$, where $\overline{\mu} = \max(0, \mu)$; then $\Delta_{t+1} = (1+\alpha) W_{t+1}$ and $\mathbb{E}[\Delta_{t+1}] \leq -(1+\alpha)\epsilon \leq -\epsilon$. When $\rho > 0$ and $\mu \geq 0$, $\rho_{t+1} = \rho + W_{t+1}$ but $\overline{\mu}_{t+1} = \overline{\mu}$ so that $\Delta_{t+1} = W_{t+1}$ and $\mathbb{E}[\Delta_{t+1}] \leq -\epsilon$. Similarly, when $\rho = 0$ and $\mu < 0$, $\overline{\mu}_{t+1} = \overline{\mu} + W_{t+1}$ while $\rho_{t+1} = \rho + \max(0, W_{t+1})$; we may compute

$$\mathbb{E}[\Delta_{t+1}] \leq \frac{1-\epsilon}{2}(1+\alpha) - \frac{1+\epsilon}{2}\alpha = \frac{1-\epsilon}{2} - \epsilon\alpha = \frac{1-\epsilon}{2} - \epsilon\left(\frac{1}{\epsilon} \cdot \frac{1+\epsilon}{2}\right) = -\epsilon \,.$$

Finally, when $\rho = \mu = 0$ exactly one of the two random variables ρ_{t+1} and $\overline{\mu}_{t+1}$ differs from zero: if $W_{t+1} = 1$ then $(\rho_{t+1}, \overline{\mu}_{t+1}) = (1, 0)$; likewise, if $W_{t+1} = -1$ then $(\rho_{t+1}, \overline{\mu}_{t+1}) = (0, -1)$. It follows that

$$\mathbb{E}[\Delta_{t+1}] \leq \frac{1-\epsilon}{2} - \frac{1+\epsilon}{2}\alpha \leq -\epsilon \,.$$

Thus $\mathbb{E}[\Phi_k] = \mathbb{E}\sum_{t=1}^{k} \Delta_t \leq -\epsilon k$. We wish to apply Azuma's inequality to conclude that $\Pr[\Phi_k \geq 0]$ is exponentially small. For this purpose, we transform the random variables Φ_t to a related supermartingale by shifting them: specifically, define $\tilde{\Phi}_t = \Phi_t + \epsilon t$ and $\tilde{\Delta}_t = \Delta_t + \epsilon$ so that $\tilde{\Phi}_t = \sum_i^t \tilde{\Delta}_t$. Then

$$\mathbb{E}[\tilde{\Phi}_{t+1} \mid \tilde{\Phi}_1, \ldots, \tilde{\Phi}_t] = \mathbb{E}[\tilde{\Phi}_{t+1} \mid W_1, \ldots, W_t] \leq \tilde{\Phi}_t \,, \qquad \tilde{\Delta}_t \in [-(1+\alpha)+\epsilon, 1+\alpha+\epsilon] \,,$$

and $\tilde{\Phi}_k = \Phi_k + \epsilon k$. It follows from Azuma's inequality that

$$\begin{aligned}
\Pr[w \text{ forkable}] = \Pr[\overline{\mu}_k = 0] \leq \Pr[\Phi_k \geq 0] &= \Pr[\tilde{\Phi}_k \geq \epsilon k] \\
&\leq \exp\left(-\frac{\epsilon^2 k^2}{2k(1+\alpha+\epsilon)^2}\right) = \exp\left(-\left(\frac{2\epsilon^2}{1+3\epsilon+2\epsilon^2}\right)^2 \cdot \frac{k}{2}\right) \\
&\leq \exp\left(-\frac{2\epsilon^4}{1+35\epsilon} \cdot k\right) \,.
\end{aligned} \tag{20}$$

If x is not empty. In this case, we go back to study the sequences x and y as in the statement of the claim. Recall the reach distribution (i.e., the distribution of the random variable $\rho(x)$) $\mathcal{R}_m : \mathbb{Z} \to [0,1]$ from (18). Since $x = (x_1, \ldots, x_m)$ satisfies the ϵ-martingale condition, Lemma 4 states that $\mathcal{R}_m \preceq \mathcal{R}_\infty$. We reserve the symbol $\mu_x^{(r)}$ for the relative margin random walk μ_x which starts at a non-negative initial position r. Thus $\rho(x) = \mu_x(\epsilon) = r$, and

$$\Pr[\mu_x(y) \geq 0] = \sum_{r \geq 0} \mathcal{R}_m(r) \Pr[\mu_x^{(r)}(y) \geq 0] \leq \sum_{r \geq 0} \mathcal{R}_\infty(r) \Pr[\mu_x^{(r)}(y) \geq 0] \tag{21}$$

since the sequence $(\Pr[\mu_x^{(r)}(y) \geq 0])_{r=0}^{\infty}$ is non-decreasing and $\mathcal{R}_m \leq \mathcal{R}_\infty$. Fix a "large enough" positive integer r^* whose value will be assigned later in the analysis. Let us define the following events:

- Event B_r: it occurs when $r \in [0, r^*]$ and the $\mu_x^{(r)}$ walk is strictly positive on every prefix of y with length at most $k/2$; and
- Event $\mathsf{C}_{r,s}$: it occurs when $r \in [0, r^*]$ and $\hat{y}$ is the smallest prefix of y of length $s \in [r, k/2]$ such that $\mu_x^{(r)}(\hat{y}) = 0$. We say that $\hat{y}$ is a witnesses to the event $\mathsf{C}_{r,s}$.

The right-hand side of (21) can be written as

$$\sum_{r>r^*} \mathcal{R}_\infty(r) \Pr[\mu_x^{(r)}(y) \geq 0] + \sum_{r \leq r^*} \mathcal{R}_\infty(r) \Pr[\mathsf{B}_r] \cdot \Pr\left[\mu_x^{(r)}(y) \geq 0 \mid \mathsf{B}_r\right]$$
$$+ \sum_{r \leq r^*} \mathcal{R}_\infty(r) \sum_{s=r}^{k/2} \Pr[\mathsf{C}_{r,s}] \cdot \Pr[\mu_x^{(r)}(y) \geq 0 \mid \mathsf{C}_{r,s}]\,.$$

We observe that the probabilities $\Pr[\mu_x^{(r)}(y) \geq 0]$ and $\Pr[\mu_x^{(r)}(y) \geq 0 \mid \mathsf{B}_r]$ are at most one. In addition, recall that for two non-negative sequences $(a_i), (b_i)$ of equal lengths, we have $\sum a_i b_i \leq \max b_i$ if $\sum a_i \leq 1$. Thus (21) can be simplified as

$$\begin{aligned}\Pr[\mu_x(y) \geq 0] &\leq \sum_{r>r^*} \mathcal{R}_\infty(r) + \sum_{r \leq r^*} \mathcal{R}_\infty(r) \Pr[\mathsf{B}_r] \\ &\quad + \sum_{r \leq r^*} \mathcal{R}_\infty(r) \max_{r \leq s \leq k/2} \Pr[\mu_x^{(r)}(y) \geq 0 \mid \mathsf{C}_{r,s}] \\ &\leq \sum_{r>r^*} \mathcal{R}_\infty(r) + \max_{r \leq r^*} \Pr[\mathsf{B}_r] + \max_{\substack{r \leq r^* \\ r \leq s \leq k/2}} \Pr[\mu_x^{(r)}(y) \geq 0 \mid \mathsf{C}_{r,s}]\,. \end{aligned} \tag{22}$$

The first term in (22) is the right-tail of the distribution $\mathcal{R}_\infty$. Using Lemma 4, this quantity is at most β^{r^*} where $\beta := (1-\epsilon)/(1+\epsilon)$. Furthermore, it can be easily checked that the above quantity is at most $\exp(-5\epsilon/3)$.

The second term in (22) concerns the event B_r and calls for more care. Define

$$S_k^{(r)} := \sum_{t=0}^{k} W_t$$

where $W_0 = r$ and the random variables W_t are defined at the outset of this proof for $t \geq 1$. We know that the $\mu_x^{(r)}$ walk starts with $\rho(x) = \mu(x) = r \geq 0$. Since B_r holds, both the margin $\mu_x(\hat{y})$ and the reach $\rho(x\hat{y})$ remain non-negative for all prefixes $\hat{y}$ of length $t = 1, 2, \cdots, k/2$. These two facts imply that the random variable $\mu_x^{(r)}(\hat{y})$ is identical to the sum $S_t^{(r)}$ for all prefixes $\hat{y}$ of length $t = 1, 2, \cdots, k/2$.

To be precise,

$$\Pr[\mathsf{B}_r] = \Pr[S_t^{(r)} \geq 0 \quad \text{for all } t \leq k/2]\,.$$

The latter probability is at most $\Pr[S_{k/2}^{(r)} \geq 0]$ because the event $S_{k/2}^{(r)} \geq 0$ does not constrain the intermediate sums $S_t^{(r)}$ for $t < k/2$. Since $\Pr[S_{k/2}^{(r)} \geq 0]$ increases monotonically in r, we conclude that the second term in (22) is at most $\Pr[S_{k/2}^{(r^*)} \geq 0]$. Now we are free to shift our focus from the relative margin walk to the sum of a martingale sequence.

For notational clarity, let us write $S := S_{k/2}^{(r^*)}$. Since the sequence (w_t) obeys the ϵ-martingale condition, $\mathbb{E}\,S$ is at most $M := r^* - k\epsilon/2$. Let us set $r^* = W_0 = k\epsilon/4$. Then $\mathbb{E}\,S$ is at most $-k\epsilon/4$ and Azuma's inequality gives us

$$\Pr[S \geq 0] = \Pr[(S - \mathbb{E}\,S) \geq k\epsilon/4] \leq \exp\left(-\frac{(k\epsilon/4)^2}{2(k/2)\cdot 2^2}\right) = \exp\left(-\frac{k\epsilon^2}{64}\right)\,.$$

This is an upper bound on the second term in (22).

The third term in (22) concerns the event $\mathsf{C}_{r,s}$ and it can be bounded using our existing analysis of the $|x| = 0$ case. Specifically, suppose $y = \hat{y}z$ where $\hat{y}$ is a witness to the event $\mathsf{C}_{r,s}$. Since the $\mu_x^{(r)}$ walk remains non-negative over the entire string $\hat{y}$, it follows that $\rho(x\hat{y}) = \mu(x\hat{y}) = 0$ and as a consequence, the $\mu_{x\hat{y}}$ walk on z is identical to the μ walk on z. Our analysis in the $|x| = 0$ case suggests that $\Pr[\mu(z) \geq 0]$ is at most $A(k-s,\epsilon)$ where $|z| = k - s$ and $A(k,\epsilon)$ is the bound in (20). Since $A(\cdot,\epsilon)$ decreases monotonically in the first argument, $A(k-s,\epsilon)$ is at most $A(k/2,\epsilon)$. However, since the last quantity is independent of r, the third term in (22) is at most $A(k/2,\epsilon) = \exp\left(-k\epsilon^4/(1+35\epsilon)\right)$.

Returning to (22) and using $r^* = k\epsilon/4$, we get

$$\Pr[\mu_x(y) \geq 0] \leq \exp\left(-\frac{5\epsilon}{3}\cdot\frac{k\epsilon}{4}\right) + \exp\left(-\frac{2\epsilon^4}{1+35\epsilon}\cdot\frac{n}{2}\right) + \exp\left(-\frac{k\epsilon^2}{64}\right) .$$

It is easy to check that the above quantity is at most $3\exp\left(-k\epsilon^4/(64+35\epsilon)\right) = 3\exp\left(-\epsilon^4(1-O(\epsilon))k/64\right)$.

$\square$