Blockchains from Non-Idealized Hash Functions

Juan A. Garay
Texas A&M University
garay@cse.tamu.edu

Aggelos Kiayias*
University of Edinburgh
& IOHK
akiayias@inf.ed.ac.uk

Giorgos Panagiotakos
University of Edinburgh
giorgos.pan@ed.ac.uk

April 3, 2020

Abstract

The formalization of concrete, non-idealized hash function properties sufficient to prove the security of Bitcoin and related protocols has been elusive, as all previous security analyses of blockchain protocols have been performed in the random oracle model. In this paper we identify three such properties, and then construct a blockchain protocol whose security can be reduced to them in the standard model assuming a common reference string (CRS).

The three properties are: *collision resistance*, *computational randomness extraction* and *iterated hardness*. While the first two properties have been extensively studied, iterated hardness has been empirically stress-tested since the rise of Bitcoin; in fact, as we demonstrate in this paper, any attack against it (assuming the other two properties hold) results in an attack against Bitcoin.

In addition, iterated hardness puts forth a new class of search problems which we term *iterated search problems* (ISP). ISPs enable the concise and modular specification of blockchain protocols, and may be of independent interest.

*Research partly supported by EU Project No.780477, PRIVILEDGE.

Contents

1 Introduction

Blockchain protocols, introduced by Nakamoto [44], are seen as a prominent application of the "proof of work" (PoW) concept to the area of consensus protocol design. PoWs were initially introduced in the work of Dwork and Naor [27] as a spam protection mechanism, and subsequently found applications in other domains such as Sybil attack resilience [26] and denial of service protection [38, 4], prior to their application to the domain of distributed consensus hinted at early on by Aspnes *et al.* [3].

A PoW scheme is typified by a *proving* algorithm, that produces a solution given an input instance, as well as a *verification* algorithm that verifies the correctness of the witness with respect to the input. The fundamental property of a PoW scheme is that the proving algorithm allows for no significant shortcuts, i.e., it is hard to significantly make it more expedient, and hence any verified solution implies an investment of computational effort on behalf of the prover. Nevertheless, this "moderate hardness" property alone has been found to be insufficient for the utilization of PoWs in the context of various applications and other properties have been put forth to complement it. These include: (i) *amortization resistance*, which guarantees that the adversary cannot speed up the computation when solving multiple PoW instances together, and (ii) *fast verification*, which suggests a significant gap between the complexities of the proving and verification algorithms.

Despite the evolution of our understanding of the PoW primitive, as exemplified in recent works (e.g., [1, 6, 13, 33]), there has been no definitive analysis of the primitive in the context of blockchain protocol security in the standard model. Intuitively, PoWs are useful in the consensus setting because they make message passing (moderately) hard and hence generate stochastic opportunities for the parties running the protocol to unify their view of the current state of the system. This fundamentally relies on an assumption about the aggregate computational power of the honest parties, but not on their actual number, in relation to the computational power of the parties that may deviate from the protocol (the "Byzantine" parties)—a hallmark of the peer-to-peer setting Bitcoin is designed for. Despite the fact that the Bitcoin blockchain has been analyzed formally [30, 47, 31, 5], the required PoW properties have not been fully identified and most of the existing analysis has been carried out in the random oracle (RO) model [10]. The same is true for a wide variety of other protocols in the space, including [2, 39, 32].

We stress that despite the fact that the RO model has been widely used in the security analysis of practical protocols and primitives, it has also received significant criticism. For example, Canetti *et al.* [20] showed that there exist implementations of signatures and encryption schemes that are secure in the RO model but insecure for any implementation of the RO in the standard model; Nielsen [45] proved that efficient non-committing encryption has no instantiation in the standard model but a straightforward implementation in the RO model, while Goldwasser and Kalai [37] showed that the Fiat-Shamir heuristic [29] does not necessarily imply a secure digital signature, which is in contrast with the result by Pointcheval and Stern [48] in the RO model.

It follows that it is critical to discover security arguments for blockchain protocols that do not rely on the RO model. Note that we are looking for arguments as opposed to proofs since it is easy to observe that some computational assumption would still be needed for deriving the security of a blockchain protocol (recall that blockchain security cannot be inferred information theoretically as it fundamentally requires at minimum the collision resistance of the underlying hash function). In fact, the formalization of non-idealized, concrete hash function assumptions sufficient to prove security of Bitcoin and related protocols has been identified as a "fascinating open question" in [18].

Following the above, the main question that motivates the present work is the following:

> *Is it possible to prove the security of blockchain protocols in the standard model under non-idealized assumptions about the underlying hash function?*

Our results. In this paper we answer the above question in the positive, by identifying three properties of a hash function family $\{H_k(\cdot)\}_k$ and then constructing a blockchain protocol whose security can be reduced to these properties (together with NIZKs; see below) in the standard model.

The first property is *collision resistance*. Specifically, it should be hard for an adversary given a random key k, to find two distinct messages m, m' for which it holds $H_k(m) = H_k(m')$. This property is useful in the blockchain context, since intuitively collision resistance ensures that the hash-chain maintained by the parties ensures the chronologically correct encoding of information.

The second property of the underlying hash function family is that it should be *computational randomness extracting* (CRE). Specifically, there is a way to isolate a finite subset of the domain of the hash function family so that for any given key k, the function H_k is a (weak) computational randomness extractor. This property is useful in a few different ways in blockchain security. Firstly, it will help for symmetry breaking, making sure that parties work concurrently on independent instances of the underlying problem. Secondly, it will ensure that the problem instances generated by honest parties (in the form of new blocks), will be sufficiently unpredictable in the eyes of the adversary. Regarding the plausibility of a CRE hash, note that pseudorandom functions (PRFs) are known to imply weak computational randomness extractors [22], and assuming that a hash function implies a PRF is a fairly standard assumption [7, 25, 41].

The third property asks for the *iterative hardness* of the underlying hash function as multiple pre-images with near zero hashes are stringed together in the form of a chain. This assumption is implicit in the context of the Bitcoin protocol. In fact, as we show, an attack against iterative hardness would result in an attack against the protocol (assuming a CRE hash). This implies that there is (monetary) incentive to break this assumption, which coupled with the the fact that no significant attacks have been demonstrated in the context of the Bitcoin protocol, establishes iterated hardness of the underlying hash (in this case SHA256) as a plausible assumption.

Armed with the above, we show a novel blockchain protocol whose security can be reduced to the collision resistance, computational randomness extraction and iterative hardness of the underlying hash function. Our design adopts Bitcoin's hash-based blockchain structure, as well as the longest-chain selection rule. However, contrary to previous analyses of this type of protocols [30, 47, 5] in the random-oracle model, iterative hardness provides no guarantee that blocks are "non-malleable," in the sense that it may be easy to mine multiple blocks on the same height of the chain once you have mined the first one. Our solution is to instead construct a PoW that *is* malleable, and leverage it to show a reduction that breaks the underlying iterated hardness assumption given a common-prefix attack to the blockchain protocol. In order to achieve this, we also have to hide the block witnesses by taking advantage of NIZK proofs with efficient simulation, thus managing to efficiently extract a sequence of iterated witnesses despite the fact that the attacker may not produce consecutive blocks.

To describe and analyze the protocol modularly, we put forth a new class of search problems, which we call *iterated search problems* (ISP). Taking advantage of ISPs one can produce concise and modular specifications of blockchain protocols, as evidenced by the description of our protocol (Section 4.2); as such, ISPs can be of independent interest.

In a nutshell, an ISP instance is defined by a problem statement set X, a witness set W and a relation R that determines when a witness satisfies the problem statement. The ISP is also equipped with a *successor* algorithm S that given a statement x and a witness w, can produce a successor problem statement x'; a *solving* algorithm M which given an initial problem statement x can find a sequence of witnesses; and a *verification* algorithm V that takes a problem statement x and witness w and outpus 1 if $(x, w) \in R$, and 0 otherwise. Each witness corresponds to the next statement defined by algorithm S on input the previous statement and witness, starting from x. The *iterated hardness* property of the ISP asks that if the solving algorithm takes t steps to solve k instances iteratively, no alternative algorithm can substantially speed this process up and produce k iterative solutions with non-negligible probability.

We perform our analysis in the static setting with synchronous rounds as in [30], and prove that our protocol can thwart adversaries and environments that roughly take less than half the computational steps the honest parties collectively are allowed per round. It is straightforward to extend our results to the Δ-synchronous setting of [47]. Further, we leave as an open question the extension of our results to the dynamic setting of [31], as well as matching the 50% threshold on adversarial computational power of the Bitcoin blockchain which can be shown in the RO model.

Related work. A related but distinct notion of hardness, *sequential* (i.e., non-parallelizable) iterated hardness. This notion has been considered as early as [49], mainly in the domains of timed-release cryptography [15] and protocol fairness [34], and recently formalized in [14] under the term *iterated sequential functions* (ISF) in the context of Verifiable Delay Functions (VDFs). In addition, a number of candidate hard problems have been proposed, including squaring a group element of composite-modulus groups [49], hashing, and computing the modular square root of an element on a prime order group [42]. Nevertheless, we observe that if we base the Bitcoin protocol on an ISF (or VDF for that matter) it will be *insecure*. The fundamental issue is that it does not allow for parallelization, which is crucial for proving the security of any (Bitcoin-like) blockchain protocol. Indeed, an attacker with a single processor whose sequential speed is slightly faster than that of honest parties, can outperform potentially hundreds of them and mine longer chains first.

Another notion related to iterative hardness is the notion of "correlation intractability" (CI) [18]. The difference is that while CI only bounds the success probability in solving a single challenge, ISP fundamentally requires multiple instances. Further, while CI talks about any sparse relation, the iterative hardness definition talks about a specific non-sparse relation.

Finally, another related work focusing on sufficient conditions for the consensus problem in the permissionless setting is [33], which introduced the concept of "signatures of work" (SoW) as the basic underlying assumption. The only known implementation of SoWs however is in the RO model, hence it is unknown (and an interesting open question) whether SoWs can be realized under non-idealized hash function assumptions like the ones we consider here.

Organization of the paper. The basic computational model, definitions and cryptographic building blocks used by our constructions are presented in Section 2. The three hash-function properties that the security of our blockchain protocol is going to be based on are presented in Section 3. Section 4 is dedicated to the new blockchain protocol realizing a transaction ledger. The presentation of the protocol is modularized by first presenting ISP and an ISP-based blockchain protocol in Sections 4.1 and 4.2, respectively, and then presenting a security notion for ISPs and showing it sufficient to prove our protocol secure (Sections 4.3 and 4.4). The construction of a secure ISP based on the hash-function properties presented in Section 3 is presented in Section 4.5. Finally, the necessity of iterated hardness to prove the Bitcoin protocol secure is presented in Section 5.

2 Preliminaries

In this section we present basic notation and definitions that we will use in the rest of the paper.

For $k \in \mathbb{N}^+$, $[k]$ denotes the set $\{1, \ldots, k\}$. For strings x, z, $x||z$ is the concatenation of x and z, and $|x|$ denotes the length of x. We denote sequences by $(a_i)_{i\in I}$, where I is the index set which will always be countable. For a set X, $x \leftarrow X$ denotes sampling a uniform element from X. For a distribution $\mathcal{U}$ over a set X, $x \leftarrow \mathcal{U}$ denotes sampling an element of X according to $\mathcal{U}$. By $\mathcal{U}_m$ we denote the uniform distribution over $\{0,1\}^m$. For random variable X, we denote by $H_\infty(X)$ the min-entropy of X. We denote the statistical distance between two random variables X, Z with range U by $\Delta[X, Y]$, i.e., $\Delta[X, Z] = \frac{1}{2}\sum_{v\in U} |\Pr[X = v] - \Pr[Z = v]|$. A *random variable ensemble* $(X_i)_{i\in I}$, is a sequence of random variables indexed by I. By $(X_i)_i \approx (Y_i)_i$ (resp. $\stackrel{c}{\approx}$) we denote that two ensembles are statistical (resp. computational) indistinguishable. We let λ denote the security parameter.

Protocol execution and security model. In this paper we will follow a more concrete approach [8, 11, 34, 12] to security evaluation. We will use functions t, ϵ, whose range is $\mathbb{N}, \mathbb{R}$, respectively, and have possibly many different arguments, to denote concrete bounds on the running time (number of steps) and probability of adversarial success of an algorithm in some given computational model, respectively. When we speak about running time this will include the execution time plus the length of the code (cf. [12]; note also that we will be considering uniform machines). We will always assume that t is a polynomial in the security parameter λ, although we will sometimes omit this dependency for brevity.

Instead of using interactive Turing machines (ITMs) as the underlying model of distributed computation, we will use (interactive) RAMs. The reason is that we need a model where subroutine access and simulation do not incur a significant overhead. ITMs are not suitable for this purpose, since one needs to account for the additional steps to go back-and-forth all the way to the place where the subroutine is stored. A similar choice was made by Garay *et al.* [34]; refer to [34] for details on using interactive RAMs in a UC-like framework. Given a RAM M, we will denote by $\mathsf{Steps}_M(1^\lambda, x)$ the random variable that corresponds to the number of steps taken by M given input 1^λ and x. We will say that M is t-bounded if it holds that $\Pr[\mathsf{Steps}_M(1^\lambda, x) \leq t(\lambda)] = 1$.

Finally, we remark that in our analyses there will be asymptotic terms of the form $\mathsf{negl}(\lambda)$ and concrete terms; throughout the paper, we will assume that λ is large enough to render the asymptotic terms insignificant compared to the concrete terms.

The Bitcoin backbone model. In this section, we give an overview of the security model that we are going to use throughout this work, introduced in [33]. This model is a variant of the synchronous model presented in [30] for the analysis of the Bitcoin backbone protocol, extended to accommodate a standard-model analysis of PoW-based blockchain protocols. In turn the model of [30] is based on Canetti's formulation of "real world" execution for multi-party cryptographic protocols [16, 17].

An execution of some protocol Π is defined with respect to an "environment" program $\mathcal{Z}$, a "control" program C, and an "adversary" program $\mathcal{A}$. At a high level, $\mathcal{Z}$ is responsible for providing inputs to and obtaining outputs from different instances of Π, C is responsible for supervising the spawning and communication of all these programs, and $\mathcal{A}$ aims to disrupt the goals set by the protocol. The programs in question can be thought of as "interactive RAMs" communicating through registers in a well-defined manner.

We consider executions where the set of of parties $\{P_1, ..., P_n\}$ running Π is fixed and hardcoded to C. Moreover, we consider a "hybrid" model of computation [19], where the adversary $\mathcal{A}$ as well as all parties in the execution can access a number of "ideal" functionalities as subroutines; the functionalities are also modeled as RAMs and are presented later in detail. Initially $\mathcal{Z}$ is activated. $\mathcal{Z}$ can make special requests that result in the spawning of different parties and $\mathcal{A}$. In turn, $\mathcal{A}$ can corrupt different parties by sending messages of the form (Corrupt, P_i) to C, with the limitation that the total number of parties corrupted should be at most t; t is a parameter of the execution. We assume an active static[1] adversary.

We are working on the *synchronous setting* of computation, where the current round is known to all parties, and messages sent at one round are received at the beginning of the next one. The influence of the adversary in the network is going to be actively malicious following standard cryptographic practice. While we assume the adversary to be rushing and communication not to be authenticated, messages sent by honest parties are guaranteed to reach their destination.

All the above concerns are captured by the diffusion functionality $\mathcal{F}_{\mathsf{diff}}$. The functionality maintains a Receive string defined for each party P_i. A party is allowed at any moment to fetch the messages

[1] We note that all our results also hold against an adaptive adversary, if we allow erasures.

sent to it at the previous round that are contained in its personal Receive string. Moreover, when the functionality receives an instruction to diffuse a message m from party P_i, it marks the party as complete for the current round, and forwards the message to the adversary; note that m is allowed to be empty. At any moment, the adversary $\mathcal{A}$ is allowed to specify the contents of the Receive string for each party P_i. The adversary has to specify when it is complete for the current round. When all parties are complete for the current round, the functionality inspects the contents of all Receive tapes and includes any messages that were diffused by the parties in the current round but not contributed by the adversary to the Receive tapes. The variable round is then incremented. In the protocol description, we will use DIFFUSE as the message transmission command.

In addition, we assume the existence of a *common reference string* (CRS) functionality that samples the CRS in a trusted manner from a known efficiently samplable distribution, and is available for all parties to fetch at the start of the execution. Note, that from our modeling it is implicit that the adversary and the honest parties get access to the CRS at the same round.

Based on the above, we denote by $\{\textsc{view}^{P,t,n}_{\Pi,\mathcal{A},\mathcal{Z}}(z)\}_{z\in\{0,1\}^*}$ the random variable ensemble that corresponds to the view of party P at the end of an execution where $\mathcal{Z}$ takes z as input. We will consider stand-alone executions, hence z will always be of the form 1^λ, for $\lambda \in \mathbb{N}$. For simplicity, to denote this random variable ensemble we will use $\textsc{view}^{P,t,n}_{\Pi,\mathcal{A},\mathcal{Z}}$. By $\textsc{view}^{t,n}_{\Pi,\mathcal{A},\mathcal{Z}}$ we denote the concatenation of the views of all parties. The probability space where these variables are defined depends on the coins of all honest parties, $\mathcal{A}$, $\mathcal{Z}$ and the CRS generation procedure.

Furthermore, we are going to define a predicate on executions and prove our properties in disjunction with this predicate, i.e., either the property holds or the execution is not good.

Definition 1. Let $(t_{\mathcal{A}}, \theta)$-good be a predicate defined on executions in the hybrid setting described above. Then E *is* $(t_{\mathcal{A}}, \theta)$*-good*, where E is one such execution, if

- the total number of steps taken by $\mathcal{A}$ and $\mathcal{Z}$ per round is no more than $t_{\mathcal{A}}$;[2]
- the adversary sends at most θ messages per round.

Definition 2. Given a predicate Q and bounds $t_{\mathcal{A}}, \theta, t, n \in \mathbb{N}$, with $t < n$, we say that protocol Π satisfies property Q for n parties assuming the number of corruptions is bounded by t, provided that for all PPT $\mathcal{Z}, \mathcal{A}$, the probability that $Q(\textsc{view}^{t,n}_{\Pi,\mathcal{A},\mathcal{Z}})$ is false and the execution is $(t_{\mathcal{A}}, \theta)$-good is negligible in λ.

Cryptographic primitives and building blocks. We will make use of the following cryptographic primitives: Cryptographic hash functions, (computational) randomness extractors [46, 22], robust non-interactive zero-knowledge (NIZK) [50], and iterated sequential functions [14].

Randomness extractors. We make use of the notion of *weak* computational randomness extractors, as formalized in [22].

Definition 3. An extractor is a family of functions $\mathsf{Ext} = \{\mathsf{Ext}_\lambda : \{0,1\}^{n(\lambda)} \times \{0,1\}^{d(\lambda)} \to \{0,1\}^{m(\lambda)}\}_{\lambda\in\mathbb{N}}$, where $n(\cdot), d(\cdot)$ and $m(\cdot)$ are polynomials. The extractor is called weak $k(\cdot)$-computational if Ext_λ is PPT, and for all efficiently samplable probability ensembles $\{X_\lambda\}_\lambda$ with min-entropy $k(\lambda)$:

$$(\mathsf{Ext}_\lambda(X_\lambda, U_{d(\lambda)}))_{\lambda\in\mathbb{N}} \overset{c}{\approx} (U_{m(\lambda)})_{\lambda\in\mathbb{N}}$$

where computational indistinguishability is defined w.r.t. a non-uniform distinguisher.

Robust non-interactive zero-knowledge. We make use of the following composable notion of non-interactive zero-knowledge, introduced in [50].

[2] The adversary cannot use the running time of honest parties that it has corrupted; it is activated instead of them during their turn.

Definition 4. Given an NP relation R, let $L = \{x : \exists w \text{ s.t. } R(x,w) = 1\}$. $\Pi = (q, P, V, S = (S_1, S_2), E))$ is a *robust NIZK argument* for L, if $P, V, S, E \in PPT$ and $q(\cdot)$ is a polynomial such that the following conditions hold:

1. Completeness. For all $x \in L$ of length λ, all w such that $R(x,w) = 1$, and all $\Omega \in \{0,1\}^{q(\lambda)}$, $\mathsf{V}(\Omega, x, \mathsf{P}(\Omega, w, x))] = 1$.

2. Multi-Theorem Zero-knowledge. For all PPT adversaries $\mathcal{A}$, we have that $\textsc{Real}(\lambda) \approx \textsc{Sim}(\lambda)$, where
$$\textsc{Real}(\lambda) = \{\Omega \leftarrow \{0,1\}^{q(\lambda)}; out \leftarrow \mathcal{A}^{\mathsf{P}(\Omega,\cdot,\cdot)}(\Omega); \text{Output } out\},$$
$$\textsc{Sim}(\lambda) = \{(\Omega, tk) \leftarrow \mathsf{S}_1(1^\lambda); out \leftarrow \mathcal{A}^{\mathsf{S}_2'(\Omega,\cdot,\cdot,tk)}(\Omega); \text{Output } out\},$$
and $\mathsf{S}_2'(\Omega, x, w, tk) \stackrel{def}{=} \mathsf{S}_2(\Omega, x, tk)$ if $(x,w) \in R$, and outputs `failure` if $(x,w) \notin R$.

3. Extractability. There exists a PPT algorithm E such that, for all PPT $\mathcal{A}$,
$$\Pr\begin{bmatrix}(\Omega, tk) \leftarrow \mathsf{S}_1(1^\lambda); (x,\pi) \leftarrow \mathcal{A}^{\mathsf{S}_2(\Omega,\cdot,tk)}(\Omega); w \leftarrow \mathsf{E}(\Omega,(x,\pi),tk) : \\ R(x,w) \neq 1 \wedge (x,\pi) \notin \mathcal{Q} \wedge \mathsf{V}(\Omega, x, \pi) = 1\end{bmatrix} \leq \mathsf{negl}(\lambda)$$
where $\mathcal{Q}$ contains the successful pairs (x_i, π_i) that $\mathcal{A}$ has queried to S_2.

As in [28], we also require that the proof system supports labels. That is, algorithms $\mathsf{P}, \mathsf{V}, \mathsf{S}, \mathsf{E}$ take as input a label ϕ, and the completeness, zero-knowledge and extractability properties are updated accordingly. This can be achieved by adding the label ϕ to the statement x. In particular, we write $\mathsf{P}^\phi(\Omega, x, w)$ and $\mathsf{V}^\phi(\Omega, x, \pi)$ for the prover and the verifier, and $\mathsf{S}_2^\phi(\Omega, x, tk)$ and $\mathsf{E}^\phi(\Omega, (x,\pi), tk)$ for the simulator and the extractor.

Theorem 5 ([50])**.** *Assuming trapdoor permutations and a dense cryptosystem exist, robust NIZK arguments exist for all languages in $\mathcal{NP}$.*

Iterated sequential functions. We recite the hardness definition introduced in [14]:

Definition 6. $f : X \to Y$ is a (t, ϵ)*-sequential function* for $\lambda = O(\log(|X|))$, if the following conditions hold:

1. There exists an algorithm that for all $x \in X$ evaluates f in parallel time t using $\mathsf{poly}(\log(t), \lambda)$ processors.

2. For all $\mathcal{A}$ that run in parallel time strictly less than $(1-\epsilon)\cdot t$ with $\mathsf{poly}(t, \lambda)$ processors:
$$\Pr[y_{\mathcal{A}} = f(x) | y_{\mathcal{A}} \leftarrow \mathcal{A}(\lambda, x), x \leftarrow X] < \mathsf{negl}(\lambda).$$

Definition 7. Let $g : X \to X$ be a function which satisfies (t, ϵ)-sequentiality. A function $f : \mathbb{N} \times X \to X$ defined as $f(k, x) = g^{(k)}(x) = g \circ g \circ \ldots \circ g$ is called an *iterated sequential function* (with round function g), if for all $k = 2^{o(\lambda)}$, the function $h : X \to X$ such that $h(x) = f(k, x)$ is (kt, ϵ)-sequential.

Robust public transaction ledgers. Our work is concerned with necessary and sufficient conditions to implement a public transaction ledger. Next, we give the transaction ledger definition introduced in [30], with Liveness slightly strengthened, as in [47].

A *public transaction ledger* is defined with respect to a set of valid ledgers $\mathcal{L}$ and a set of valid transactions $\mathcal{T}$, each one possessing an efficient membership test. A ledger $\mathbf{x} \in \mathcal{L}$ is a vector of sequences of transactions $\mathrm{tx} \in \mathcal{T}$. Ledgers correspond to chains of blocks in the Bitcoin protocol. It is possible for the adversary to create two transactions that are conflicting; valid ledgers must not contain conflicting transactions. Moreover, it is assumed that in the protocol execution there also exists an oracle Txgen that generates valid transactions, and is unambiguous, i.e., the adversary cannot create transactions that come in 'conflict' with the transactions generated by the oracle. A transaction is called *neutral* if there does not exist any transactions that come in conflict with it.

Definition 8. A protocol Π implements a *robust public transaction ledger* if it organizes the ledger as a chain of blocks of transactions and it satisfies the following two properties:

- **Consistency** *(parameterized by the "depth" parameter $k \in \mathbb{N}$):* If in a certain round an honest player reports a ledger that contains a transaction tx in a block more than k blocks away from the end of the ledger, where $k \in \mathbb{N}$ is the "depth" parameter (such transactions are called *stable*), then tx will be reported as stable and in the same position in the ledger by any honest player from this round on.
- **Liveness** *(parameterized by $k, u \in \mathbb{N}$—the "depth" and "wait time" parameters, resp.):* For every u consecutive rounds, there exists a round and an honest party, such that the transactions given as input to that party at this round that are either (i) issued by Txgen or (ii) neutral, will be reported by all honest parties as stable at the end of this round interval.

3 Hash Functions Properties for Blockchain Security

In this section we describe the three falsifiable hash function properties that the security of our protocol is going to be based on. Two of these properties, namely, collision resistance [23] and weak computational randomness extraction [22], have been extensively studied in the hash function literature. The third one is new, and has to do with the moderate hardness of computing *sequences* of small hashes. We proceed to discuss each of the properties in detail.

We start with collision resistance. Most known blockchain protocols make use of a collision-resistant hash function in order to establish basic structural properties, e.g., that the adversary cannot create a blockchain that contains a cycle. That is exactly the way we are going to use this property here. We will use the following security definition [36].[3]

Definition 9. Let $\mathcal{H} = \{\{H_k : \{0,1\}^* \to \{0,1\}^\lambda\}_{k \in K(\lambda)}\}_{\lambda \in \mathbb{N}}$ be a hash-function family, and $\mathcal{A}$ be a PPT adversary. Then $\mathcal{H}$ is **collision resistant** if and only if for any $\lambda \in \mathbb{N}$ and corresponding $\{H_k\}_{k \in K}$ in $\mathcal{H}$,

$$\Pr_{k \leftarrow K}[(m, m') \leftarrow \mathcal{A}(1^\lambda, k) : (m \neq m') \wedge (H_k(m) = H_k(m'))] \leq \mathsf{negl}(\lambda).$$

Our second security assumption has to do with the existence of a fixed-length-input hash function family that is a *weak* computational randomness extractor. As explained in [22], this assumption is weaker than assuming a fixed-length-input pseudorandom function family (FI-PRF), a common assumption in the hash function literature [7, 25, 41]. We adapt the definition of a weak computational randomness extractor given in Appendix **??** to the context of a hash function family.

Definition 10. Let $\mathcal{H} = \{\{H_k : \{0,1\}^{d\lambda} \to \{0,1\}^\lambda\}_{k \in K(\lambda)}\}_{\lambda \in \mathbb{N}}$, for some $d \in \mathbb{N}$, $d > 1$, be a fixed-length input hash-function family. $\mathcal{H}$ is a **computational randomness extracting** (CRE) hash function family if for some $c \in \mathbb{N}^+$, $c < d$, the function family $E = \{E_\lambda : \{0,1\}^{(c+1)\lambda} \times \{0,1\}^{(d-c-1)\lambda} \to \{0,1\}^\lambda\}_\lambda$, where $E_\lambda(x, i) \stackrel{\text{def}}{=} H_k(x||i)$, is a weak $(c\lambda)$-computational extractor (Definition 3), for any $k \in K(\lambda)$.

This property will be useful in our protocol for two reasons. First, to ensure that the distributions of blocks generated by honest parties are identical and independent. Second, to establish that the blocks generated by honest parties, and which the adversary has the choice to mine on, look sufficiently random and hence the moderate hardness of the underlying problem is preserved.

[3] Throughout our exposition for simplicity we will assume that $\mathcal{H}$ takes one step to be evaluated. We note that our results can be generalized to the case where $\mathcal{H}$ takes more time.

Our third assumption about hash functions has to do with the hardness of finding sequences of small hashes in the hash-based PoW construction proposed for Bitcoin. In more detail, given the hash x of some block, computing a valid PoW for this construction consists of finding witnesses w_1, w_2 such that $H_k(H_k(x||w_1)||w_2) < T$. In turn, our hardness property requires that any adversary should take a number of steps proportional to the number of PoWs computed, when these PoWs form a sequence starting from a uniformly random string x. The property is parameterized by t, the number of steps the adversary takes to generate each PoW on average.

Definition 11. Let $\mathcal{H} = \{\{H_k : \{0,1\}^{d\lambda} \to \{0,1\}^\lambda\}_{k \in K(\lambda)}\}_{\lambda \in \mathbb{N}}$, for some $d \in \mathbb{N}$, $d > 1$, be a fixed-length input hash-function family, and let T be some hardness parameter. $\mathcal{H}$ is **t-iteratively hard** iff there exists a polynomial $k_0(\cdot)$, such that for any PPT RAM $(\mathcal{A}_1, \mathcal{A}_2)$, $\lambda \in \mathbb{N}$, and $k \geq k_0(\lambda)$, it holds that:

$$\Pr_{\substack{\sigma \leftarrow K(\lambda); \\ x_0 \leftarrow [0,T]}} \left[\begin{array}{l} st \leftarrow \mathcal{A}_1(1^\lambda, \sigma); (w_i, w_i')_{i \in [k]} \leftarrow \mathcal{A}_2(1^\lambda, st, x_0) : \\ \forall i \in [k] : x_i := H_\sigma(H_\sigma(x_{i-1}||w_i)||w_i') < T \\ \wedge\, \mathsf{Steps}_{\mathcal{A}_2}(st, x_0) < k \cdot t \end{array} \right] \leq \mathsf{negl}(\lambda)$$

Our choice to base the security of our protocol on the iterated hardness of Bitcoin's PoW construction is not accidental. Empirical evidence about the security of Bitcoin, provided by the fact that no attacks have been publicly disclosed in the last ten years that this construction has been actively used in Bitcoin, and the fact that any attack on iterated hardness implies an attack on Bitcoin, as we show in Appendix 5, also constitute empirical evidence in favor of iterated hardness. Note that this would not necessarily be the case if we based security on a stronger hardness property that was not necessary to prove Bitcoin secure, as it would then be possible that an attack against the property is known and the adversary does not have any incentive to reveal it, as it does not affect the security of the protocol in any way.

We note that to prove the security of our protocol both properties in Definitions 10 and 11 should hold for the same hash function and for suitable parameters [4], which we discuss in the next section; collision resistance may hold for a different hash function. Finally, in our protocol analysis we will also make use of a number of other standard assumptions, such as the existence of a NIZK-PoK scheme or that the honest parties control the majority of the computational power. The theorem we prove is as follows:

Theorem 43 (Informal) *Assume the existence of collision-resistant hash functions, a hash function family that is CRE and iteratively hard for appropriate parameters, a one-way trapdoor permutation and a dense cryptosystem (for the NIZK), and that $t_{\mathcal{A}}$ is (roughly) less than half the total running time of honest parties per round. Then there exists a protocol that implements a robust public transaction ledger.*

Finally, Gentry and Wichs [35] define as falsifiable the cryptographic assumptions that can be expressed as a game between an efficient challenger and an adversary. All cryptographic assumptions of Theorem 43 are falsifiable in this sense, with two caveats: First, due to the concrete security approach our work takes, the challenger should take as input the number of steps of the adversary. Second, in the computational randomness extraction property we quantify over all keys of the hash, which is not immediate to express in the framework of [35]. We could instead first choose they key randomly, and then have the adversary choose the source distribution of the extractor. To simplify our presentation we adopt the former version of the definition. However, we note that the proof techniques we use can be adapted to handle the latter.

[4] Intuitively, the adversary should not be able to compute small hashes much faster than the rate at which honest parties generate blocks that is guaranteed by the computational extractor property.

4 Blockchains from Non-Idealized Hash Functions

In this section we present and prove secure a protocol that implements a transaction ledger and is based on a hash function that satisfies the properties described in Section 3. We modularize our presentation and analysis by first introducing the concept of *iterated search problems* (ISP) in Section 4.1, and an ISP-based blockchain protocol in Section 4.2. Then, in Section 4.3, we introduce a "blockchain friendly" ISP security definition, that show in Section 4.4 to be sufficient to prove our protocol secure. Finally, in Section 4.5 we construct a secure ISP based on the hash properties defined in Section 3, which in combination with our protocol can be shown to satisfy Theorem 43.

As always the choice of modularizing the protocol analysis has multiple benefits. In particular, it firstly allows us to formally capture all required properties that the moderately hard problem that our protocol is built on should satisfy for the analysis to go through. We hope that this will motivate building other constructions in the future. Secondly, it makes it easier to take advantage of previous efforts to analyze relevant protocols [30, 47, 33]. While we adapt some of the proof techniques presented there, a contribution of our work is that the ISP notion which we built on is considerably weaker and can be instantiated in the standard model from fairly simple assumptions.

4.1 Iterated Search Problems

In this section we introduce a class of problems inspired by Bitcoin's underlying computational problem. The straightforward properties that this class should have, are the ability to find a witness for a problem statement and to verify that the witness is correct, matching Bitcoin's block mining and block verification procedures, respectively. In addition, the notion models the ability to generate a new problem statement from a valid statement/witness pair. This captures the fact that in Bitcoin the problem that a miner solves depends on a previous block (i.e., a statement/witness pair). This concept has appeared before in the study of iterated *sequential* functions [14], whose name we draw from. Syntactically, the key difference here is that in each iteration we are not evaluating a function, but instead we are solving a search problem with possibly many witnesses. Moreover, as we already commented in Section 1 iterated sequential functions are *not* the correct abstractions for Bitcoin's underlying computational problem, as they allow for an attack against the protocol. We proceed to give a formal definition of ISPs.

Definition 12 (Iterated Search Problem)**.** An *iterated search problem* (ISP) $\mathcal{I}$ specifies a collection $(I_\lambda)_{\lambda\in\mathbb{N}}$ of distributions.[5] For every value of the security parameter $\lambda \geq 0$, I_λ is a probability distribution of *instance descriptions*. An instance description Λ specifies

1. finite, non-empty sets X, W, and
2. a binary relation $R \subset X \times W$.

We write $\Lambda[X, W, R]$ to indicate that the instance Λ specifies X, W and R as above.

An ISP also provides several algorithms. For this purpose, we require that the instance descriptions, as well as the elements of the sets X and W, can be uniquely encoded as bit strings of length polynomial in λ, and that both X and $(I_\lambda)_{\lambda\in\mathbb{N}}$ have polynomial-time samplers. The ISP algorithms are as follows, all parameterized by $\Lambda[X, W, R]$:

– *Verification* algorithm $V_\Lambda(x, w)$: A deterministic algorithm that takes as input a problem statement x and a witness w and outputs 1 if $(x, w) \in R$ and 0 otherwise.

[5] Here we follow the notation used in [21] to define *subset membership problems*. We remark that no other connection exists between the two papers.

- *Successor* algorithm $S_\Lambda(x, w)$: A deterministic algorithm that takes as input a problem statement[6] x and a valid witness w and outputs a new instance $x' \in X$.
- *Solving* algorithm $M_\Lambda(x, k)$: A probabilistic algorithm that takes as input a problem statement x and a number $k \in \mathbb{N}^+$ and outputs a sequence of k witnesses $(w_i)_{i \in [k]}$.

In the sequel, we will omit writing Λ as a parameter of V, S, M when it is clear from the context. In order to ease the presentation, we recursively extend the definitions of S and R to sequences of witnesses as follows: Let $S(x, \emptyset) := x$ and for any $k > 1$, $S(x, (w_i)_{i \in [k]}) := S(S(x, (w_i)_{i \in [k-1]}), w_k)$ and $(x, (w_i)_{i \in [k]}) \in R$ iff $\bigwedge_{i=1}^{k} (S(x, (w_j)_{j \in [i-1]}), w_i) \in R$. Further, we assume that M is *correct*, i.e., for $(w_i)_{i \in [k]} \leftarrow M(x, k)$, it holds that $(x, (w_i)_{i \in [k]}) \in R$.

Example. Next, we present as an example Bitcoin's underlying computational problem captured as an ISP.

Construction 1. Let T be a protocol parameter representing how hard it is to solve a problem instance.[7] Then:
- I_λ is the uniform distribution over functions $H : \{0,1\}^* \to \{0,1\}^\lambda$ in some family of hash functions $\mathcal{H}$, i.e., $\Lambda = \{H\}$;
- $X = \{x | x < T \wedge x \in \{0,1\}^\lambda\}$ and $W = \{0,1\}^* \times \{0,1\}^\lambda$;
- $R = \{(x, w) | H(H(x||m)||ctr) < T, \text{ for } w = m||ctr\}$;
- $V(x, w)$ checks whether $H(H(x||m)||ctr) < T$, for $w = m||ctr$;
- $S(x, w) = H(H(x||m)||ctr)$, and
- $M(x, 1)$ tests whether $V(x, (m, ctr))$ is true, for different (m, ctr) pairs, until it finds a solution. $M(x, k)$ is defined inductively, by running $M(x', 1)$ on the statement x' produced by $M(x, k-1)$. The output consists of all witnesses found.

4.2 Blockchain Protocol Description

Having established the basic syntax for ISPs, we proceed to describe our protocol. The main challenge we have to overcome is that while the protocol's security is going to be based on iterated hardness (Definition 11), it operates in a setting where the adversary can also take advantage of the work of honest parties. This includes, the adversary seeing the information leaked by the honestly produced blocks, as well as, honest parties directly working on the chain it is extending. In contrast, the iteratively hard experiment does not provide any guarantees about these cases; the adversary does not receive any externally computed witnesses.

Towards this end, blocks in our protocol, instead of exposing the relevant witness computed, contain a proof of knowledge (PoK) of such a valid witness through a non-interactive zero-knowledge (NIZK) proof. At first, the fact that we use NIZK proofs for a language that is moderately hard may seem counter-intuitive, due to the fact that a trivial simulator and extractor exist for the zero-knowledge and soundness properties, since computing a new witness for a given statement takes polynomial time. Instead, following our general approach, we make concrete assumptions regarding the efficiency of both the simulator and the extractor. Informally, we require that the time it takes to simulate a proof or extract a witness is a lot smaller than the time it takes for honest parties to compute a witness (see Assumption 2). Note that in practice this can be achieved by making the underlying problem hard enough, which however will affect the performance of the resulting ledger being implemented.

[6]We could formalize S more generally, to take as input a sequence of problem statements. However, for our exposition the current formulation suffices. Note, that a more general definition would be needed for the variable difficulty case [31], which we do not study here, where the next block's difficulty depends on the whole chain.

[7]For simplicity, in our exposition the hardness parameter for each ISP is fixed, and we do not capture it explicitly.

Finally, the protocol adopts the longest-chain selection rule of the Bitcoin protocol, which as we will see later allows it to operate even if the witnesses of the ISP are malleable. To make our analysis cleaner, the hash chain structure of blocks is decoupled from the underlying computational problem.

4.2.1 The protocol in detail.

Next, we are going to describe our new protocol. Our protocol, $\Pi_{\mathsf{PL}}^{\mathsf{new}}$, uses as building blocks three cryptographic primitives: An ISP $\mathcal{I} = (M, V, S)$, a collision-resistant hash function family $\mathcal{H}$, and a robust NIZK (Definition 4) protocol $\Pi_{\mathsf{NIZK}} = (q, \mathsf{P}, \mathsf{V}, \mathsf{S} = (\mathsf{S}_1, \mathsf{S}_2), \mathsf{E})$ for the language[8]

$$L = \{(\Lambda[X, W, R], x, x') | \exists w \in W : (x, w) \in R \wedge S(x, w) == x'\}$$

where $\Lambda[X, W, R]$ is an ISP instance of $\mathcal{I}$. Π_{NIZK} also supports labels, which we denote as a superscript on P and V. The initialization of these primitives happens through the CRS all parties share at the start of the execution, which contains: An instance description $\Lambda[X, W, R]$, a statement $x_{\mathtt{Gen}}$, the description of a hash function $H : \{0,1\}^* \to \{0,1\}^\lambda$ and the NIZK reference string Ω, each randomly sampled from $I_\lambda, X, \mathcal{H}, \{0,1\}^{q(\lambda)}$, respectively. Moreover, as in [30], our protocol is parameterized by the chain validation predicate $\mathrm{V}(\cdot)$, the chain reading function $\mathrm{R}(\cdot)$, and the input contribution function $\mathrm{I}(\cdot)$ to capture higher-level applications, e.g., Bitcoin.

Next, we introduce some notation used in the description of our protocol. We use the terms block and chain to refer to tuples of the form $\langle s, m, x, \pi \rangle \in \{0,1\}^\lambda \times \{0,1\}^* \times X \times \{0,1\}^{\mathsf{poly}(\lambda)}$, and sequences of such tuples, respectively. The rightmost (resp., leftmost) block of chain $\mathcal{C}$ is denoted by head$(\mathcal{C})$ (resp., tail$(\mathcal{C})$). Each block contains the hash of the previous block s, a message m, the next problem x to be solved, and a NIZK proof π. We denote by $B_{\mathtt{Gen}} = \langle 0^\lambda, 0^\lambda, x_{\mathtt{Gen}}, 0^\lambda \rangle$ a special block called the *genesis block*; note that $x_{\mathtt{Gen}}$ is part of the CRS. A chain $\mathcal{C} = (\langle s_i, m_i, x_i, \pi_i \rangle)_{i \in [k]}$ is valid if: (i) The first block of $\mathcal{C}$ is equal to $B_{\mathtt{Gen}}$; (ii) the contents of the chain $\mathbf{m}_{\mathcal{C}} = (m_1, \ldots, m_k)$ are valid according to the chain validation predicate V, i.e., $\mathrm{V}(\mathbf{m}_{\mathcal{C}})$ is true; (iii) $s_{i+1} = H(s_i || m_i || x_i || i)$[9] for all $i \in [k]$, and (iv) $\mathsf{V}^{s_{i+1}}((\Lambda, x_{i-1}, x_i), \pi_i)$ is true for all $i \in [k] \setminus \{1\}$; see Algorithm 1. We call $H(s_i || m_i || x_i || i)$ the *hash of block* B_i and denote it by $H(B_i)$, and define $H(\mathcal{C}) \triangleq H(\mathrm{head}(\mathcal{C}))$. We will consider two valid blocks or chains as equal, if all their parts match, except possibly for the NIZK proofs.

We proceed to describe the main function of the protocol, presented in Algorithm 4. At each round, each party chooses the longest valid chain among the ones it has received (Algorithm 2) and tries to extend it by computing a new witness. If it succeeds, it diffuses the new block to the network. In more detail, each party will run the solver M on the problem x defined in the last block $\langle s, m, x, \pi \rangle$ of the chosen chain $\mathcal{C}$. If it succeeds on finding a witness w, it will then compute a NIZK proof that it knows a witness w such that $(x, w) \in R$ and $S(x, w) = x'$, for some $x' \in X$. The proof should also have a label $H(H(head(\mathcal{C}))||m'||x'||(|\mathcal{C}| + 1))$, where m' is the output of the input contribution function $\mathrm{I}(\cdot)$, i.e., the message encoded in the block; see Algorithm 3. Then, the party diffuses the extended chain to the network. Finally, if the party is queried by the environment, it outputs $\mathrm{R}(\mathcal{C})$, where $\mathcal{C}$ is the chain selected by the party; the chain reading function $\mathrm{R}(\cdot)$ interprets $\mathcal{C}$ differently depending on the higher-level application running on top of the backbone protocol. We assume that all honest parties take the same number of steps t_H per round.

In order to turn the above protocol into a protocol realizing a public transaction ledger, we define functions $\mathrm{V}(\cdot), \mathrm{R}(\cdot), \mathrm{I}(\cdot)$ exactly as in [30]. For completeness we give these definitions in Table 1. We denote the new public ledger protocol by $\Pi_{\mathsf{PL}}^{\mathsf{new}}$.

[8] We assume that both V and S are efficiently computable. Hence, $L \in \mathcal{NP}$.

[9] We include a fixed length (λ-bit) encoding of the height of the block in the hash on purpose. This way, the contents of the hash chain form a suffix-free code [9], which in turn implies collision resistance. See Lemma 19.

Algorithm 1 The validate procedure, parameterized by $B_{\mathtt{Gen}}$, the hash function $H(\cdot)$, the *chain validation predicate* $V(\cdot)$, and the verification algorithm V of Π_{NIZK}. The input is $\mathcal{C}$.

function $\mathsf{validate}(\mathcal{C})$
  $b \leftarrow \mathrm{V}(\mathbf{m}_{\mathcal{C}}) \wedge (\mathrm{tail}(\mathcal{C}) = B_{\mathtt{Gen}})$ ▷ $\mathbf{m}_{\mathcal{C}}$ describes the contents of chain $\mathcal{C}$.
  if $b = \mathrm{True}$ **then** ▷ The chain is non-empty and meaningful w.r.t. $V(\cdot)$
    $s' \leftarrow H(B_{\mathtt{Gen}})$ ▷ Compute the hash of the genesis block.
    $x' \leftarrow x_{\mathtt{Gen}}$
    $\mathcal{C} \leftarrow \mathcal{C}^{1\rceil}$ ▷ Remove the genesis from $\mathcal{C}$
    while $(\mathcal{C} \neq \epsilon \wedge b = \mathrm{True})$ **do**
      $\langle s, m, x, \pi\rangle \leftarrow \mathrm{tail}(\mathcal{C})$
      $s'' \leftarrow H(\mathrm{tail}(\mathcal{C}))$
      if $(s = s' \wedge \mathsf{V}^{s''}(\Omega, (\Lambda, x', x), \pi))$ **then**
        $s' \leftarrow s''$ ▷ Retain hash value
        $x' \leftarrow x$
        $\mathcal{C} \leftarrow \mathcal{C}^{1\rceil}$ ▷ Remove the tail from $\mathcal{C}$
      else
        $b \leftarrow \mathrm{False}$
  return (b)

Algorithm 2 The function that finds the "best" chain, parameterized by function $\max(\cdot)$. The input is $\{\mathcal{C}_1, \ldots, \mathcal{C}_k\}$.

function $\mathsf{maxvalid}(\mathcal{C}_1, \ldots, \mathcal{C}_k)$
  $temp \leftarrow \varepsilon$
  for $i = 1$ to k **do**
    if $\mathsf{validate}(\mathcal{C}_i)$ **then**
      $temp \leftarrow \max(\mathcal{C}, temp)$
  return $temp$

4.3 ISP Security Properties

Next, we present a set of ISP properties sufficient to prove our protocol secure. Later in Section 4.5 we show how to instantiate them.

In the same spirit as in Boneh *et al.* [14]'s definition of an iterated sequential function (cf. Definition 7), we can define the notion of a *hard* iterated search problem. Our definition is parameterized by t, δ and k_0, all functions of λ which we omit for brevity. Unlike the former definition, we take in account the *total* number of steps instead of only the sequential ones, and we require the error probability to be negligible after at least k_0 witnesses have been found instead of one. In that sense, our notion relaxes the strict convergence criterion of [14]. Finally, note that the adversary is allowed some precomputation time.

Definition 13. An ISP $\mathcal{I} = (V, M, S)$ is (t, δ, k_0)-*hard* iff it holds that

- For $\lambda \in \mathbb{N}$ and for all polynomially large $k \geq k_0$:

$$\Pr_{\substack{\Lambda[X,W,R]\leftarrow I_\lambda;\\ x\leftarrow X}}\left[\begin{array}{l}(w_i)_{i\in[k]} \leftarrow M(x,k) : (x,(w_i)_i) \in R\\ \wedge\ \mathsf{Steps}_M(x,k) \leq k\cdot t\end{array}\right] \geq 1 - \mathsf{negl}(\lambda), \text{ and}$$

Algorithm 3 The *proof of work* function is parameterized by the hash function $H(\cdot)$, and the proving algorithm P of Π_{NIZK}. The input is $(m', \mathcal{C})$.

function $\mathsf{pow}(m', \mathcal{C})$
  $\langle s, m, x, \pi \rangle \leftarrow \mathrm{head}(\mathcal{C})$
  $w \leftarrow M(x)$ ▷ Run the honest solving algorithm of the ISP.
  if $w \neq \bot$ **then**
    $x' \leftarrow S(x, w)$ ▷ Compute the next problem to be solved.
    $s' \leftarrow H(s||m||x|||\mathcal{C}|)$ ▷ Compute the hash of the last block.
    $s'' \leftarrow H(s'||m'||x'|||\mathcal{C}| + 1)$ ▷ Compute the hash of the new block.
    $\pi' \leftarrow \mathsf{P}^{s''}(\Omega, (\Lambda, x, x'), w)$ ▷ Compute the NIZK proof.
    $B \leftarrow \langle s', m', x', \pi' \rangle$
  $\mathcal{C} \leftarrow \mathcal{C}B$ ▷ Extend chain
  return $\mathcal{C}$

Algorithm 4 The Bitcoin backbone protocol, parameterized by the *input contribution function* $\mathrm{I}(\cdot)$ and the *chain reading function* $\mathrm{R}(\cdot)$.

$\mathcal{C} \leftarrow B_{\mathsf{Gen}}$ ▷ Initialize $\mathcal{C}$ to the genesis block.
$st \leftarrow \varepsilon$
$round \leftarrow 0$
while TRUE **do**
  $\tilde{\mathcal{C}} \leftarrow \mathsf{maxvalid}(\mathcal{C}$, any chain $\mathcal{C}'$ found in RECEIVE())
  $\langle st, m \rangle \leftarrow \mathrm{I}(st, \tilde{\mathcal{C}}, round, \mathrm{INPUT}(), \mathrm{RECEIVE}())$ ▷ Determine the m-value.
  $\mathcal{C}_{\mathsf{new}} \leftarrow \mathsf{pow}(m, \tilde{\mathcal{C}})$
  if $\mathcal{C} \neq \mathcal{C}_{\mathsf{new}}$ **then**
    $\mathcal{C} \leftarrow \mathcal{C}_{\mathsf{new}}$
    DIFFUSE($\mathcal{C}$)
  $round \leftarrow round + 1$
  if INPUT() contains READ **then**
    write $\mathrm{R}(\mathbf{m}_{\mathcal{C}})$ to OUTPUT()

– for any PPT RAM $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, $\lambda \in \mathbb{N}$, and all polynomially large $k \geq k_0$, it holds that

$$\Pr_{\substack{\Lambda[X,W,R] \leftarrow I_\lambda; \\ x \leftarrow X}} \left[\begin{array}{l} st \leftarrow \mathcal{A}_1(1^\lambda, \Lambda); (w_i)_{i \in [k]} \leftarrow \mathcal{A}_2(1^\lambda, st, x) : \\ (x, (w_i)_i) \in R \wedge \mathit{Steps}_{\mathcal{A}_2}(st, x) < (1 - \delta)k \cdot t \end{array} \right] \leq \mathsf{negl}(\lambda).$$

The next property, has to do with establishing an upper bound t on the the running time of the verification algorithm V. Intuitively, the product $\theta \cdot t$ should be a lot smaller than the number of steps $t_{\mathcal{H}}$ per round available to honest parties, to avoid resource depletion attacks.

Definition 14. An ISP $\mathcal{I} = (V, M, S)$ is *t-verifiable* iff algorithm V takes time at most t (on all inputs).

In general, attacking an honest solver amounts to finding a certain set of inputs over which the honest solving algorithm fails to produce witnesses sufficiently fast. In order to combat this attack, we introduce the following property: We say that an ISP $\mathcal{I}$ is (t, α)-*successful* when the probability that M[10] computes a witness in t steps is at least α.

[10]For brevity, we use $M(x)$ instead of $M(x, 1)$ in this section.

Content validation predicate $\mathrm{V}(\cdot)$	$\mathrm{V}(\cdot)$ is true if its input $\langle m_1, \ldots, m_\ell \rangle$ is a valid ledger, i.e., it is in $\mathcal{L}$.
Chain reading function $\mathrm{R}(\cdot)$	$\mathrm{R}(\cdot)$ returns the contents of the chain if they constitute a valid ledger, otherwise it is undefined.
Input contribution function $\mathrm{I}(\cdot)$	$\mathrm{I}(\cdot)$ returns the largest subsequence of transactions in the input and receive registers that constitute a valid ledger, with respect to the contents of the chain $\mathcal{C}$ the party already has, preceded by a neutral random transaction.

Table 1: *The instantiation of functions* $\mathrm{V}(\cdot), \mathrm{R}(\cdot), \mathrm{I}(\cdot)$ *for protocol* $\Pi^{\mathsf{new}}_{\mathsf{PL}}(\mathcal{I})$.

Definition 15. An ISP $\mathcal{I} = (V, M, S)$ is (t, α)-*successful* iff for $\lambda \in \mathbb{N}$, $\Lambda[X, W, R] \in I_\lambda$, and for all $x \in X$ it holds that: $\Pr[\mathsf{Steps}_M(x) < t] \geq \alpha$.

The iterated hardness property (Defintion 13) does not give any guarantees regarding composition. For blockchain protocols, however, this is necessary as many parties concurrently try to solve the same ISP. To address this issue, we introduce the next property that ensures that learning how long it took for a witness to be computed or what the next problem defined by such witness is, does not leak any information that could help the adversary find a witness himself. More formally, there exists an efficient simulator whose output is computationally indistinguishable from the distribution of the time it takes to compute a witness w for some statement x and the next statement $S(x, w)$. Note that, crucially, the simulator does not depend on the instance description Λ or the problem statement x, and that we consider a non-uniform distinguisher.

Definition 16. An ISP $\mathcal{I} = (V, M, S)$ is *t-next-problem simulatable* iff there exists a t-bounded RAM Ψ such that for any PPT RAM D, any $\lambda \in \mathbb{N}$, any $z \in \{0,1\}^{poly(\lambda)}$, any $\Lambda[X, W, R] \in I_\lambda$, and any $x \in X$, it holds that

$$|\Pr[D(1^\lambda, z, \Lambda, x, (S(x, M(x)), \mathsf{Steps}_M(x))) = 1] - \Pr[D(1^\lambda, z, \Lambda, x, \Psi(1^\lambda)) = 1]| \leq \mathsf{negl}(\lambda).$$

The next property has to do with a party's ability to "cheaply" compute witnesses for a statement, if it already knows one. This will be important to ensure that even if the adversary has external help to produce some of the witnesses needed by the hard ISP experiment, as is the case for blockchain protocols, still the overall process remains hard with respect to the number of consecutive blocks the adversary actually produced. We call this ISP property *witness malleability*.

Definition 17. An ISP $\mathcal{I} = (V, M, S)$ is *t-witness malleable* iff there exists a t-bounded RAM Φ such that for any PPT RAM D, any $\lambda \in \mathbb{N}$, any $z \in \{0,1\}^{poly(\lambda)}$, any $\Lambda[X, W, R] \in I_\lambda$, and any $(x, w) \in R$, it holds that $(x, \Phi(x, w)) \in R$, and

$$|\Pr[D(1^\lambda, z, \Lambda, x, w, S(x, \Phi(x, w))) = 1] - \Pr[D(1^\lambda, z, \Lambda, x, w, S(x, M(x))) = 1]| \leq \mathsf{negl}(\lambda).$$

Finally, we call an ISP that satisfies all the above properties *secure*.

Definition 18. An ISP $\mathcal{I} = (V, M, S)$ is $(t_{\mathsf{ver}}, t_{\mathsf{succ}}, \alpha, t_{\mathsf{nps}}, t_{\mathsf{mal}}, t_{\mathsf{hard}}, \delta_{\mathsf{hard}}, k_{\mathsf{hard}})$-*secure* iff it is t_{ver}-verifiable, $(t_{\mathsf{succ}}, \alpha)$-successful, t_{nps}-next-problem simulatable, t_{mal}-witness malleable, and $(t_{\mathsf{hard}}, \delta_{\mathsf{hard}}, k_{\mathsf{hard}})$-hard.

An ISP scheme with trivial parameters is of limited use in a distributed environment; for example, if $\delta_{\mathsf{hard}} \ll 1$ or $t_{\mathsf{hard}} \ll t_{\mathsf{ver}}$. Hence, next we describe the parameters' ranges that make for a non-trivial secure ISP. First off, and ignoring negligible terms, one can show that $\alpha \leq \frac{t_{\mathsf{succ}}}{(1-\delta_{\mathsf{hard}})t_{\mathsf{hard}}}$ (see Lemma 25).

On the other hand, the successful property always holds for $\alpha = 0$. Therefore, for a non-trivial ISP scheme it should hold that α is close to $\frac{t_{\mathsf{succ}}}{(1-\delta_{\mathsf{hard}})t_{\mathsf{hard}}}$. To avoid denial of service attacks, $\theta \cdot t_{\mathsf{ver}}$ must be sufficiently small compared to t_{hard}, the running time of the solving algorithm M. Furthermore, t_{mal} should be a lot smaller than t_{hard}, otherwise M can be used as a trivial simulator. We note, that the security of the protocol that we presented earlier relies on the fact that a secure ISP scheme with favorable parameters exists, mainly reflected in Assumption 2 (Section 4.4).

4.4 Security of the ISP-based Blockchain Protocol

In this subsection we prove that $\Pi^{\mathsf{new}}_{\mathsf{PL}}$ implements a robust public transaction ledger (cf. Definition 8), assuming the underlying ISP $\mathcal{I}$ is secure.

Security Proof of the ISP-based Protocol. We proceed to the main part of the protocol analysis. The first assumption we are going to make is that the underlying ISP $\mathcal{I}$ is secure, and that the runtimes of the procedures of the NIZK system are upper bounded.

Assumption 1 (ISP Assumption). For parameters $t_{\mathsf{ver}}, t'_{\mathcal{H}}, \alpha, t_{\mathsf{nps}}, t_{\mathsf{mal}}, t_{\mathsf{hard}}, \delta_{\mathsf{hard}}, k_{\mathsf{hard}}, t_{\mathsf{P}}, t_{\mathsf{V}}, t_{\mathsf{S}}$, and t_{E} we assume that:

- ISP $\mathcal{I}$ is $(t_{\mathsf{ver}}, t'_{\mathcal{H}}, \alpha, t_{\mathsf{nps}}, t_{\mathsf{mal}}, t_{\mathsf{hard}}, \delta_{\mathsf{hard}}, k_{\mathsf{hard}})$-secure;[11]
- running the prover (resp., verifier, simulator, extractor) of Π_{NIZK} takes t_{P} (resp. $t_{\mathsf{V}}, t_{\mathsf{S}}, t_{\mathsf{E}}$) steps.

Next, we introduce some additional notation necessary to formalize our second assumption that has to do with the computational power of the honest parties and the adversary. For brevity, and to better connect our analysis to previous work [30, 47, 33], we denote by $\beta = ((1 - \delta_{\mathsf{hard}}) \cdot t_{\mathsf{hard}})^{-1}$, the upper bound on the rate at which the adversary can compute witnesses in the iterated hardness game. We introduce two variables, $t'_{\mathcal{H}}$ and $t'_{\mathcal{A}}$, that have to do with the effectiveness of honest parties and the adversary in producing witnesses for $\mathcal{I}$. $t'_{\mathcal{H}}$ is a lower bound on the number of steps each honest party takes per round running M. It holds that in any round at least $n - t$ parties will run M for at least $t'_{\mathcal{H}}$ steps. $t'_{\mathcal{A}}$ denotes the maximum time needed by a RAM machine to simulate the adversary, the environment and the honest parties in one round of the protocol execution, without taking into account calls made to M by the latter, and with the addition of one invocation of the NIZK extractor. They amount to:

$$t'_{\mathcal{A}} = t_{\mathcal{A}} + \theta \cdot t_{\mathsf{V}} + t_{\mathsf{E}} + n(t_{\mathsf{bb}} + t_{\mathsf{nps}} + t_{\mathsf{mal}} + t_{\mathsf{S}}) \quad \text{and} \quad t'_{\mathcal{H}} = t_{\mathcal{H}} - t_{\mathsf{bb}} - \theta t_{\mathsf{V}} - t_{\mathsf{P}},$$

where t_{bb} (bb for backbone) is an upper bound on the number of steps needed to run the code of an honest party in one round besides the calls to $M, \mathsf{P}, \mathsf{V}$.

We are now ready to state our main computational assumption regarding the honest parties and the adversary. Besides assuming that the total number of steps the honest parties take per round exceed those of the adversary, and that the total block generation rate is bounded, we have to additionally assume that the efficiency of the solving algorithm M used by honest parties is comparable to that of the adversary; i.e, as explained earlier, α should be comparable to $\beta t'_{\mathcal{H}}$, otherwise the adversary will be able to compute long chains of blocks fast and break the security of the protocol. The observation we just made, corresponds to the first condition in our formalization, which we present next. To avoid confusion, we cast most of our analysis based on the δ parameter. Furthermore, note that under optimal conditions – i.e., δ_{ISP} close to 0 and $t_{\mathsf{P}}, t_{\mathsf{V}}, t_{\mathsf{E}}, t_{\mathsf{S}}, t_{\mathsf{nps}}, t_{\mathsf{mal}}$ a lot smaller than $t_{\mathcal{H}}$ – our assumption allows for an adversary that controls up to 1/3 of the total computational power available (vs. 1/2 in the RO model).

Assumption 2. There exist $\delta_{\mathsf{ISP}}, \delta_{\mathsf{Steps}}$ and $\delta \in (0, 1)$, such that for sufficiently large $\lambda \in \mathbb{N}$:

[11] $t'_{\mathcal{H}}$ is related to our model and we formally define it in the next paragraph.

λ :	security parameter
n :	number of parties
t :	number of parties corrupted
$t_{\mathcal{H}}$:	number of steps per round per honest party
$t_{\mathcal{A}}$:	total number of adversarial steps per round
$t'_{\mathcal{H}}$:	lower bound on number of steps running M per round per honest party
$t'_{\mathcal{A}}$:	round simulation cost, excluding honest calls to M
θ :	upper bound on the number of messages sent by the adversary per round
β :	upper bound on the rate at which the adversary computes witnesses per step
α :	probability that M outputs a witness after $t'_{\mathcal{H}}$ steps
f :	probability that at least one party computes a block in a round
γ :	probability that exactly one party computes a block in a round
δ :	upper bound on the total block generation rate
δ_{Steps} :	honest parties' advantage on number of steps
δ_{ISP} :	adversary's advantage on ISP witnesses computation rate
k_{hard} :	convergence parameter of ISP hardness

Table 2: *The parameters in our analysis:* $\lambda, n, t, t_{\mathcal{H}}, t_{\mathcal{A}}, t'_{\mathcal{H}}, t'_{\mathcal{A}}, \theta, k_{hard}$ *are in* $\mathbb{N}$, $\alpha, f, \gamma, \beta, \delta, \delta_{\mathsf{Steps}}, \delta_{\mathsf{ISP}}$ *are in* $(0, 1)$.

- $\alpha \geq (1 - \delta_{\mathsf{ISP}})\beta t'_{\mathcal{H}} > \mathsf{negl}(\lambda)$ (ISP generation gap)
- $(n - t)t'_{\mathcal{H}}(1 - \delta_{\mathsf{Steps}}) \geq 2 \cdot t'_{\mathcal{A}}$ (steps gap)
- $\frac{\delta_{\mathsf{Steps}} - \delta_{\mathsf{ISP}}}{2} \geq \delta > \beta(t'_{\mathcal{A}} + nt_{\mathcal{H}})$ (bounded block generation rate)

Next, we focus on structural properties of blockchains in our protocol. We follow a similar approach to [33] based on a collisions resistant hash function. Observe that the hash structure of any blockchain in our protocol is similar to the Merkle-Damgard transform [24], defined as:

$$\mathrm{MD}(IV, (x_i)_{i \in [m]}) : z = IV; \text{ for } i = 1 \text{ to } m \text{ do } z = H(z || x_i); \text{ return } z,$$

where H is the hash function described in the CRS, and IV is set to $B_{\mathtt{Gen}}$. Based on this observation, as in [33], we can show that no efficient adversary can find distinct chains with the same hash value, as this would result to finding a collision on the underlying hash function.

Lemma 19. *Let $\mathcal{H}$ be a collision-resistant hash function family. The probability that any PPT RAM $\mathcal{A}$, given $B_{\mathtt{Gen}}$, can find two distinct valid chains $\mathcal{C}_1, \mathcal{C}_2$ such that $H(\mathcal{C}_1) = H(\mathcal{C}_2)$, is negligible in λ.*

Proof. To show that the adversary cannot find distinct chains with the same hash, we are going to take advantage of the following property of the MD transform: For any non-empty valid chain $\mathcal{C} = B_1, \ldots, B_k$, where $B_i = \langle s_i, m_i, x_i, \pi_i \rangle$, it holds that for any $j \in [k]$, $H(\mathrm{head}(\mathcal{C})) = \mathrm{MD}(H(B_j), ((m_i || x_i || i))_{i \in \{j+1, \ldots, k\}})$. Let $\mathcal{C}_1 = B_{\mathtt{Gen}}, B_1, \ldots, B_{|\mathcal{C}_1|}$, $\mathcal{C}_2 = B_{\mathtt{Gen}}, B'_1, \ldots, B'_{|\mathcal{C}_2|}$, $z = ((m_i || x_i || i))_{i \in [|\mathcal{C}_1|]}$ and $z' = ((m'_i || x'_i || i))_{i \in [|\mathcal{C}_2|]}$. For the sake of contradiction, assume that the lemma does not hold and there exists an adversary $\mathcal{A}$ that can find valid chains $\mathcal{C}_1$, $\mathcal{C}_2$ such that $H(\mathcal{C}_1) = H(\mathcal{C}_2)$, with non-negligible probability. By our observation above, this implies that $\mathrm{MD}(H(B_{\mathtt{Gen}}), z) = \mathrm{MD}(H(B_{\mathtt{Gen}}), z')$.

We will construct an adversary $\mathcal{A}'$ that breaks the collision resistance of H with non-negligible probability. We have two cases. In the first case, $|\mathcal{C}_1| \neq |\mathcal{C}_2|$. Then, since the height of the chain is part of the hash of blocks $B_{|\mathcal{C}_1|}, B'_{|\mathcal{C}_2|}$, and $H(B_{|\mathcal{C}_1|}) = H(\mathrm{head}(\mathcal{C}_1)) = H(\mathrm{head}(\mathcal{C}_2)) = H(B_{|\mathcal{C}_2|})$, it follows that a collision in H has been found. In the second case, where $|\mathcal{C}_1| = |\mathcal{C}_2|$, following the classical inductive argument for the MD transform, it can be shown that either $\mathcal{C}_1$ and $\mathcal{C}_2$ are equal, or there exists $\ell \in [|\mathcal{C}_1|]$, such that $\mathrm{MD}(H(B_{\texttt{Gen}}), ((m_i||x_i||i))_{i\in[\ell]}) = \mathrm{MD}(H(B_{\texttt{Gen}}), ((m'_i||x'_i||i))_{i\in[\ell]})$ and $(m_\ell, x_\ell) \neq (m'_\ell, x'_\ell)$, which implies again that a collision has been found for H. Since the probability that a collision can be found for H is negligible, the lemma follows.

The corollary follows easily by observing that insertions and copies imply there exist distinct chains that end on the same block. This in turn implies that the hash of the two chains is the same, which violates Lemma 19. Hence, the corollary follows with overwhelming probability in λ. □

Lemma 19 implies that insertion and copy properties[12] of [30], that have to do with the way blocks are connected, do not occur with overwhelming probability in λ.

Definition 20. An *insertion* occurs when, given a chain $\mathcal{C}$ with two consecutive blocks B and B_0, a block B^* created after B_0 is such that B, B^*, B_0 form three consecutive blocks of a valid chain. A *copy* occurs if the same block exists in two different positions.

Corollary 21. *Let $\mathcal{H}$ be a collision-resistant hash function family. Then, for any PPT $\mathcal{A}, \mathcal{Z}$ no insertions or copies occur in $\mathrm{VIEW}^{t,n}_{\Pi^{\mathsf{new}}_{\mathsf{PL}},\mathcal{A},\mathcal{Z}}$ with probability $1 - \mathsf{negl}(\lambda)$.*

We proceed to the main part of the analysis. First, we introduce some useful notation. For each round j, we define the Boolean random variables X_j and Y_j as follows. Let $X_j = 1$ if and only if j was a *successful round*, i.e., at least one honest party computed a witness at round j, and let $Y_j = 1$ if and only if j was a *uniquely successful round*, i.e., *exactly one honest party* computed a witness at round j. With respect to a set of rounds R, let $X(R) = \sum_{j\in R} X_j$ and define $Y(R)$ similarly.

Moreover, with respect to some block B computed by an honest party P at some round r, let $Z^P_r(R)$ denote the maximum number of distinct blocks diffused by the adversary during R that have B as their ancestor and lie on the same chain; note that honest parties compute at most one block per round. If P is corrupted or did not compute any block at r, let $Z^P_r(R) = 0$. We extend the definition of random variable $X(r)$ to $X^P_r(R)$ similarly.

An important part of our analysis will be to establish lower and upper bounds for these random variables. First, in Lemma 23 we will show that the rate at which the adversary produces witnesses is upper bounded by $\beta \cdot t'_{\mathcal{A}}$. Then, in Lemma 25 we prove that the expected rate of successful and uniquely successful rounds is lower bounded by f and γ, respectively, both defined bellow:

$$f = 1 - (1-\alpha)^{n-t} \text{ and } \gamma = (n-t)\cdot\alpha\cdot(1-\beta t_{\mathcal{H}})^{n-1}$$

Finally, for our analysis to go through, γ should be twice as big as $\beta \cdot t'_{\mathcal{A}}$. As we demonstrate next, this follows from the fact that in Assumption 2 the honest parties take at least double the steps the adversary takes per round.

Lemma 22. *Assume an ISP that complies with Assumptions 1 and 2. It holds that $\gamma \geq 2(1+\delta)\beta t'_{\mathcal{A}}$.*

Proof. For γ it holds that:

$$\begin{aligned}\gamma =&(n-t)\cdot\alpha\cdot(1-\beta t_{\mathcal{H}})^{n-1} \geq (n-t)\cdot\alpha\cdot(1-\beta t_{\mathcal{H}} n)\\ \geq&(n-t)\cdot(1-\delta_{\mathsf{ISP}})\cdot\beta t'_{\mathcal{H}}\cdot(1-\delta) \geq \frac{(1-\delta_{\mathsf{ISP}})(1-\delta)}{(1-\delta_{\mathsf{Steps}})}\cdot 2\cdot\beta t'_{\mathcal{A}} \geq 2(1+\delta)\beta t'_{\mathcal{A}}\end{aligned}$$

where we have first used Bernouli's inequality, and then the three conditions from Assumption 2. The last inequality follows from the fact that $\frac{\delta_{\mathsf{Steps}}-\delta_{\mathsf{ISP}}}{2} \geq \delta$. □

[12]A third property, called "prediction," also introduced in [30], is not needed in our proof as it is captured by the fact that the ISP is hard even in the presence of adversarial precomputation.

As promised, we prove next that the adversary cannot mine blocks extending a single chain, with rate and probability better than that of breaking the iterative hardness property.

Lemma 23. *For any set of consecutive rounds R, where $|R| \geq k_{\mathsf{hard}}/\beta t'_{\mathcal{A}}$, for any party P, and any round $i \in R$, the probability that $Z_i^P(R) \geq \beta t'_{\mathcal{A}}|R|$ is* $\mathsf{negl}(\lambda)$.

Proof. W.l.o.g., let i be the first round of $R = \{i' | i \leq i' < i+s\}$, and let E be the event where in $\text{VIEW}^{t,n}_{\Pi^{\mathsf{new}}_{\mathsf{PL}},\mathcal{A},\mathcal{Z}}$ party P at round i mined a block B, and the adversary mined at least $\beta t'_{\mathcal{A}} s$ blocks until round $i+s$ that extend B and are part of a single chain. For the sake of contradiction, assume that the lemma does not hold, and thus $\Pr[E]$ is non-negligible. Using $\mathcal{A}$, we will construct an adversary $\mathcal{A}' = (\mathcal{A}'_1, \mathcal{A}'_2)$ that breaks the iterative hardness (Definition 13) of $\mathcal{I}$ with non-negligible probability.

$\mathcal{A}'$ is going to run internally $\mathcal{A}$ and $\mathcal{Z}$, while at the same time simulating the work honest parties do using the NIZK proof simulator. Moreover, $\mathcal{A}'$ is also going to use the witness malleability property, to trick $\mathcal{A}$ to produce blocks in a sequence, instead of interleaved adversarial and (simulated) honest blocks. Finally, using the NIZK extractor, $\mathcal{A}'$ is going to extract the witnesses from the adversarial blocks, and win the iterative hardness game. By a hybrid argument, we will show that the view of $\mathcal{A}, \mathcal{Z}$ is indistinguishable both in the real and the simulated run, and thus the probability that E happens will be the same in both cases.

Next, we describe the behavior of $\mathcal{A}'$ in more detail. We are going to describe the two stages of $\mathcal{A}'$ separately, i.e. before and after obtaining x. First, $\mathcal{A}'_1(\Lambda)$ sets $(\Lambda, x_{\mathsf{Gen}}, H, \Omega)$ as the common input for $\mathcal{A}$ and $\mathcal{Z}$, where Ω has been generated using S_1 and the rest of the inputs using the default samplers, and stores the NIZK trapdoor tk. Then, it perfectly simulates honest parties up to round $i-1$ and at the same time runs $\mathcal{A}$ and $\mathcal{Z}$ in a black-box way. Finally, it outputs the contents of the registers of $\mathcal{A}$ and $\mathcal{Z}$ and the NIZK trapdoor tk, as variable st. It can do all this, since in the iterated hardness experiment it has polynomial time on λ on his disposal. Note, that up until this point in the eyes of $\mathcal{A}$ and $\mathcal{Z}$ the simulated execution is perfectly indistinguishable compared to the real one.

For the second stage, $\mathcal{A}'_2(st, x)$, is first going to use st to reset $\mathcal{A}$ and $\mathcal{Z}$ to the same state that they were. We assume that this can be done efficiently, e.g., by having $\mathcal{A}$ and $\mathcal{Z}$ read from the registers where st is stored whenever they perform some operation on their registers. It will also continue to simulate honest parties, this time in a more efficient way.

$\mathcal{A}'_2$ takes as input a problem statement x sampled from X, as in Definition 13. It should somehow introduce x to the simulated protocol execution, without the adversary noticing any difference that could help him distinguish from the real execution. Let $B_0 = \langle s_0, m_0, x_0, \pi_0 \rangle$ be the head of chain $\mathcal{C}$ that party P is extending at round i, and m_1 the block input it produced for this round using the input contribution function $\mathrm{I}(\cdot)$. $\mathcal{A}'_2$ is first going to run M on input x for the amount of steps available to P. If it is successful and produces some witnesses w, it will diffuse the following block:

$$B_1 = \langle H(B_0), m_1, S(x,w), \mathsf{S}_2^{H(H(B_0),m_1,S(x,w),|\mathcal{C}|+1)}(\Omega, (\Lambda, x_0, S(x,w)), tk) \rangle$$

where the last component is a simulated NIZK proof for the statement $(x_0, S(x,w))$. Note, that $\mathcal{A}'_2$ does not know any witness for this statement, and it is possible that no such witness exists. Later, we will argue that the output of the simulator on this input should be indistinguishable from the output on the statement $(x_0, S(x_0, M(x_0)))$. Also, note that due to the next-problem simulatability property, $\mathcal{A}_2$ will not be able to tell the difference of P running M on x_0 or x at this round.

$\mathcal{A}'_2$ will follow a more complex strategy to simulate the rest of the honest parties invocations. For each honest party, it will run the next-problem simulator $\Psi(1^\lambda)$ and check if the numbers of steps output is less than the number of steps available on this invocation. If they are not, $\mathcal{A}'_2$ will proceed by just updating the state of this party for the round. Otherwise, it will simulate its behavior when being successful, as follows: Let block $B^* = \langle s^*, m^*, x^*, \pi^* \rangle$ be the head for the chain $\mathcal{C}^*$ the honest

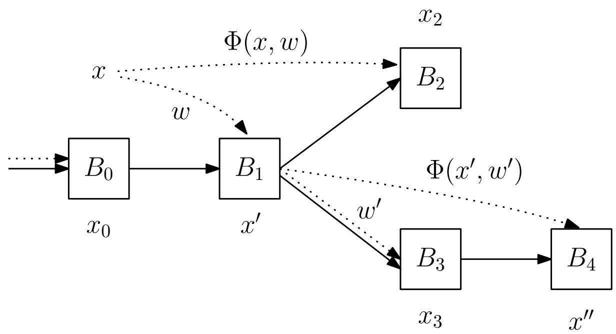

Figure 1: *A possible scenario according to Lemma 23. The blocks have been generated in order B_0, B_1, B_2, B_3, B_4, with B_3 being the only adversarial block. The cases where a valid witness is either known or can be extracted, and a NIZK proof has either been computed or simulated for the depicted transitions, correspond to the dotted and normal arrows, respectively.*

party was trying to extend with message m'' in this round. Let $B_j = \langle s_j, m_j, x_j, \pi_j \rangle$ be the adversarial block that descends B_1 and maximizes the number of adversarial blocks between itself and B_1. Let $B' = \langle s', m', x', \pi' \rangle$ be the parent of B_j. If no such adversarial block exists, assume that $B_j = B_1$ and $B' = \langle \emptyset, x, w, \emptyset \rangle$. $\mathcal{A}_2'$ first runs the NIZK extractor $\mathsf{E}^{H(B_j)}(\Omega, ((\Lambda, x', x_j), \pi_j), tk)$ to obtain a witness w' for x'. Then, it runs $\Phi(x', w')$ and obtains a new witness w'' for x'; let $x'' = S(x', w'')$. Finally, it is going to make $\mathcal{A}_2$ believe that the block it has computed extends B^*, instead of B', by simulating a NIZK proof as follows: $\pi'' = \mathsf{S}_2^{H(H(B^*), m'', x'', |\mathcal{C}^*|+1)}(\Omega, (\Lambda, x^*, S(x', w'')), tk)$. The new block that $\mathcal{A}_2'$ is going to diffuse is $\langle H(B^*), m'', x'', \pi'' \rangle$. We point to Figure 1 for an example of the procedure described above. If $\mathcal{A}_2'$ was not successful when it run $M(x)$ to extend B_0, it is going to simulate honest parties work as follows: to extend block $\hat{B} = \langle \hat{s}, \hat{m}, \hat{x}, \hat{\pi} \rangle$, it will first use Ψ to see if it succeeds and if yes generate the next problem statement $\hat{x}'$, and then use S as above to generate a NIZK proof $\hat{\pi}'$ for block $\langle H(\hat{B}), \hat{m}', \hat{x}', \hat{\pi}' \rangle$.

In the following claim we argue that the view H_{sim} of the adversary in the simulated run we just described is computationally indistinguishable from its view H_0 in $\mathrm{VIEW}^{t,n}_{\Pi^{\mathsf{new}}_{\mathsf{PL}}, \mathcal{A}, \mathcal{Z}}$.

Claim 1. *$H_{sim} \overset{c}{\approx} H_0$.*

Proof. We start by describing a sequence of hybrids:

- Hybrid H_0: The view of the adversary in $\mathrm{VIEW}^{t,n}_{\Pi^{\mathsf{new}}_{\mathsf{PL}}, \mathcal{A}, \mathcal{Z}}$.
- Hybrid H_1: Same as H_0, with the only difference being replacing honest parties' calls to P by calls to S_2, and Ω being generated by S_1.
- Hybrid $H_{1,i}$ to $H_{n,s}$: In hybrid $H_{u,v}$, we replace the next statement and the NIZK in the block produced by party u at round v if successful, with a possibly wrong statement and proof computed as described in the proof above.

By the zero knowledge property of the NIZK proof system (Definition 4) it easily follows that H_0 is indistinguishable from H_1; H_0 corresponds to the real execution, while H_1 to the simulated one.

Next, we will argue that $H_{u-1,v}$ is indistinguishable from $H_{u,v}$, for some $u \in [n], v \in R$ (let $H_{0,i} = H_1$), by contradiction . Assume $H_{u-1,v} \overset{c}{\not\approx} H_{u,v}$. There are two cases.

In the first case, $u = P$, $v = i$. The difference between the two executions, is that in $H_{P,i}$, instead of running $M(x_0)$ and computing $S(x_0, M(x_0))$, $M(\mathcal{X})$ is run and the next problem computed is $S(\mathcal{X}, M(\mathcal{X}))$, where by $\mathcal{X}$ we denote the uniform distribution over X. Assuming that the two hybrids are distinguishable, by an averaging argument there exists a PPT distinguisher

D, some auxiliary register z, and $\Lambda[X, W, \hat{R}] \in \mathcal{I}_\lambda, x_0, x_1 \in X, t_0 \leq t_\mathcal{H}$ such that D distinguishes $(\Lambda, x_0, (S(x_0, M(x_0)), \mathsf{Steps}_M(x_0) < t_0))$ from $(\Lambda, x_0, (S(x_1, M(x_1)), \mathsf{Steps}_M(x_1) < t_0))$; z will be equal to the state of an execution where P has t_0 steps to extend problem x_0. This is a contradiction, since by the next-problem simulatability property it follows that:[13]

$$(\Lambda, x_0, (S(x_0, M(x_0)), \mathsf{Steps}_M(x_0))) \overset{c}{\approx} (\Lambda, x_0, \Psi(1^\lambda)) \overset{c}{\approx} (\Lambda, x_0, (S(x_1, M(x_1)), \mathsf{Steps}_M(x_1)))$$

where the last part follows from the fact that $\Psi_0(1^\lambda)$ and x_1 do not depend x_0.

For the second case, assume that either $u \neq P$ or $v \neq i$. W.l.o.g., assume that u is successful. Similarly, by an averaging argument we can show that there exists a PPT distinguisher D, some auxiliary register z, and $\Lambda[X, W, \hat{R}] \in \mathcal{I}_\lambda, x_0, x_1 \in X, w_1 \in W, t_0 \leq t_\mathcal{H}$ such that D distinguishes $(\Lambda, x_0, (S(x_0, M(x_0)), \mathsf{Steps}_M(x_0) < t_0))$ from $(\Lambda, x_0, (S(x_1, \Phi(x_1, w_1)), \Psi_1(1^\lambda) < t_0))$ and $(x_1, w_1) \in \hat{R}$, where $\Psi_1(1^\lambda)$ is the steps component of $\Psi(1^\lambda)$. We arrive to a contradiction due to the witness malleability property:

$$\begin{aligned}(\Lambda, x_0, &(S(x_0, M(x_0)), \mathsf{Steps}_M(x_0))) \overset{c}{\approx} (\Lambda, x_0, \Psi(1^\lambda)) \\ &\overset{c}{\approx} (\Lambda, x_0, (S(x_1, M(x_1)), \mathsf{Steps}_M(x_1))) \overset{c}{\approx} (\Lambda, x_0, (S(x_1, \Phi(x_1, w_1)), \Psi_1(1^\lambda)))\end{aligned}$$

If the second and third distributions are distinguishable, then we can construct a distinguisher for the next-problem simulatable property as before, while if the third and the fourth distributions are distinguishable, we can construct a distinguisher for the witness malleability property.

The claim follows by the fact that $H_{n,s}$ is the same as H_{sim}. ⊣

Since $\mathcal{A}$ and $\mathcal{Z}$ cannot distinguish between the real execution and the one we described above, E will occur with non-negligible probability in H_{sim}, i.e. $\mathcal{A}$ will compute at least $\beta t'_\mathcal{A} s$ blocks starting from round i and up to round $i+s$ that descend B_1 and lie on the same chain. By the way honest blocks are constructed, $\mathcal{A}'_2$ knows the witnesses of the honest blocks in this chain, and using the NIZK extractor it can extract the witnesses of the adversarial ones. Now, note that each adversarial block includes a witness to the problem statement defined by the previous block, while at the same time each subsequent honest block defines a problem statement that lies in a sequence starting from x and followed by at least as many witnesses as on the previous block. It follows that $\mathcal{A}'_2$ can extract a sequence of valid witnesses of length at least $\beta t'_\mathcal{A} s + 1$, where the plus one comes from the witness computed by P at round i, and win in the iterative hardness game with non-negligible probability, since it takes at most

$$t_\mathcal{H} + s \cdot (t_\mathcal{A} + \theta \cdot t_\mathsf{V} + t_\mathsf{E}) + s \cdot n(t_\mathsf{bb} + t_{nps} + t_{mal} + t_\mathsf{S}) \leq s \cdot t'_\mathcal{A} + t_\mathcal{H}$$

steps. Hence, $\mathcal{A}'_2$ has computed $\beta t'_\mathcal{A} s + 1 \geq \beta(s \cdot t'_\mathcal{A} + t_\mathcal{H}) \geq k_\mathsf{hard}$ blocks in $s \cdot t'_\mathcal{A} + t_\mathcal{H} = (1 - \delta_\mathsf{hard}) t_\mathsf{hard} \cdot \beta(s \cdot t'_\mathcal{A} + t_\mathcal{H})$ steps with non-negligible probability. This is a contradiction to our initial assumption that $\mathcal{I}$ is a $(t_\mathsf{hard}, \delta_\mathsf{hard}, k_\mathsf{hard})$-hard ISP. □

Note that we can do exactly the same reduction without simulating honest parties' work. Then, the total running time of the second stage of $\mathcal{A}'$ is $s \cdot (t'_\mathcal{A} + n t'_\mathcal{H})$-bounded. Hence, we can derive the following bound on the longest chain that can be produced by both honest and malicious parties during a certain number of rounds.

Corollary 24. *For any set of consecutive rounds R, where $|R| \geq k_\mathsf{hard}/\beta(t'_\mathcal{A} + n t'_\mathcal{H})$, for any party P, and any round $i \in R$, the probability that $Z_i^P(R) + X_i^P(R) \geq \beta(t'_\mathcal{A} + n t'_\mathcal{H}) \cdot |R|$ is $\mathsf{negl}(\lambda)$.*

[13]For brevity, we abuse notation here and use the computational indistinguishability relation to random variables, instead of random variable ensembles. The related random variable ensembles can be easily deduced.

Next, we prove lower bounds on the rate of successful and uniquely successful rounds. In our proof we are going to take advantage of the next-problem simulatable property of $\mathcal{I}$ and the zero-knowledge property of the robust NIZK we are using. The main idea is to first use these two properties and similar arguments as in Lemma 23 to construct an "ideal" execution where: (i) honest parties' behavior is efficiently simulated using Ψ, and (ii) is computationally indistinguishable from the "real" execution. Then, since the outputs of different invocations of the runtime simulator $\Psi(1^\lambda)$ are independent, it will be much easier to establish lower bounds for $X(\cdot)$ and $Y(\cdot)$ in the ideal execution. Finally, due to the the fact that the two executions are computationally indistinguishable, and the execution properties we examine can be efficiently checked, it will follow that the same bounds should also hold for the real execution with negligible difference in probability.

Lemma 25. *For any set of consecutive rounds R, with $|R| \geq \lambda/\gamma\delta^2$, the following two events occur with negligible probability in λ:*

- *the number of uniquely successful rounds in R is less or equal to $(1-\frac{\delta}{4})\gamma \cdot |R|$;*
- *the number of successful rounds in R is less or equal to $(1-\frac{\delta}{4})f \cdot |R|$.*

Proof. W.l.o.g., let $R = \{1, \dots, s\}$. Our proof strategy will be to first prove the results of the lemma in an "ideal" execution where honest parties behavior is simulated using the simulators Ψ and S of the next-problem simulatable property and the NIZK proof system, similarly to Lemma 23; Ψ is used to determine whether a party is successful and the next problem statement, and S is used to generate the required NIZK. Then, using similar arguments as in Lemma 23, we can show that the view of the adversary in the real execution and its view on the ideal one are computationally indistinguishable, and thus the results should also hold for the real execution with negligible difference in probability. We denote by $E_{\mathtt{ideal}}(1^\lambda)$ (resp. $E_{\mathtt{real}}(1^\lambda)$) the ideal (resp. real) execution.

Next, we analyze the probability of successful and uniquely successful rounds occurring in the ideal execution. We start by deriving lower and upper bounds for $\Psi(1^\lambda)$. First, from the Successful property it follows that $\Pr[\Psi(1^\lambda) \leq t'_{\mathcal{H}}] \geq \alpha - \mathsf{negl}(\lambda)$. Otherwise, we can construct a distinguisher for $(\Lambda, x, \Psi(1^\lambda))$ and $(\Lambda, x, \mathsf{Steps}_M(x))$, for any Λ, x, by checking whether the input to the distinguisher is smaller than $t'_{\mathcal{H}}$. This violates the next-problem simulatable property. Similarly, we can upper bound $\Pr[\Psi(1^\lambda) \leq t_{\mathcal{H}}]$.

Claim 2. $\Pr[\Psi(1^\lambda) \leq t_{\mathcal{H}}] \leq t_{\mathcal{H}}\beta + \mathsf{negl}(\lambda)$.

Proof. For the sake of contradiction, assume that the difference $\Pr[\Psi(1^\lambda) \leq t_{\mathcal{H}}] - t_{\mathcal{H}}\beta$ is non-negligible. First, we will argue that there exists an $x \in X$, such that $\Pr[\mathsf{Steps}_M(x) \leq t_{\mathcal{H}}] - t_{\mathcal{H}}\beta$ is negligible. For the sake of contradiction, assume that for all $x \in X$, $\Pr[\mathsf{Steps}_M(x) \leq t_{\mathcal{H}}] - t_{\mathcal{H}}\beta$ is non-negligible. By the iterated hardness property, we have that for $k \geq k_{\mathsf{hard}}$, any k/β-bounded adversary will compute k or more witnesses with negligible probability in λ (assume we pick a k that is a multiple of β). This implies that the expected number of blocks any such adversary computes is at most $k + \mathsf{negl}(\lambda)$. Let an adversary that is based on M work as follows: on some initial input x, it runs M for at most $t_{\mathcal{H}}$ steps. If, it succeeds on producing a witness, it computes the next problem, and runs M again with the new input. If not, it runs M on the initial input. By our assumption and the linearity of expectation, the expected number of blocks our adversary will mine on k/β steps, is greater than $(\beta t_{\mathcal{H}} + \epsilon)\frac{k}{\beta t_{\mathcal{H}}} \geq k(1 + \frac{\epsilon}{\beta t_{\mathcal{H}}})$, where ϵ is a non-negligible function. This is a contradiction. Hence, there exists an $x_0 \in X$, such that $\Pr[\mathsf{Steps}_M(x_0)] - t_{\mathcal{H}}\beta$ is negligible. This in turn implies that we can construct a distinguisher for $\Psi(1^\lambda)$ and $\mathsf{Steps}_M(x_0)$, by checking whether the input of the distinguisher is less or equal to $t_{\mathcal{H}}$. This is a contradiction to the next-problem simulatable property. Therefore, the claim follows. ⊣

We proceed to analyze the probability of successful and uniquely successful rounds occurring in $E_{\texttt{ideal}}(1^\lambda)$. Let random variables $\hat{X}(\cdot), \hat{Y}(\cdot)$ for $E_{\texttt{ideal}}(1^\lambda)$ be defined similarly to $X(\cdot), Y(\cdot)$ for $E_{\texttt{real}}(1^\lambda)$. We define an additional two random variables $X'(\cdot), Y'(\cdot)$ that will be helpful in our analysis. Let $L_{i,j}$ be equal to the second output of $\Psi(1^\lambda)$ (number of steps) for its invocation in the simulation of honest party P_j at round i. Then, $X'(\{i\})$ is equal to 1, where i is some round of the execution, if there exists some party, among the first $n-t$ honest parties that are activated at round i, such that $L_{i,j} \le t'_{\mathcal{H}}$. Note that $X'(\{i\}) = 1$ implies that $\hat{X}(\{i\}) = 1$. Further, if we define $X'(R) = \sum_{i \in S} X'(\{i\})$, it follows that $\hat{X}(R) \ge X'(R)$. Next, let $Y'(\{i\})$ be equal to 1, if there exists a unique party, among the first $n-t$ honest parties that are activated at round i, such that $L_{i,j} \le t'_{\mathcal{H}}$, and for all other $n-1$ parties it holds that $L_{i,j} > t_{\mathcal{H}}$. Again, it holds that $\hat{Y}(R) \ge Y'(R)$. Note that due to the way $X'(\cdot), Y'(\cdot)$ are defined, and the fact that different invocations of $\Psi(1^\lambda)$ are independent and hence $\{L_{i,j}\}_{(i,j) \in [s] \times [n]}$ are i.i.d random variables, these two random variables do *not* depend on the behavior of the adversary and thus are easier to analyze. We proceed to analyze their expectations.

Claim 3. *For any $i \in R$, it holds that $\mathbb{E}[Y'(\{i\})] \ge \gamma$ and $\mathbb{E}[X'(\{i\})] \ge f$.*

Proof of Claim.

$$\begin{aligned}
\mathbb{E}[Y'(\{i\})] =& \Pr[Y'(\{i\}) = 1] \\
=& \Pr[\bigvee_{j \in [n-t]} (L_{i,j} \le t'_{\mathcal{H}} \wedge \bigwedge_{m \in [n] \setminus \{j\}} L_{i,m} > t_{\mathcal{H}})] \\
=& \sum_{j \in [n-t]} \Pr[L_{i,j} \le t'_{\mathcal{H}}] \cdot \prod_{m \in [n] \setminus \{j\}} \Pr[L_{i,m} > t_{\mathcal{H}}] \\
\ge& (n-t) \cdot \alpha \cdot (1 - \beta t_{\mathcal{H}})^{n-1} = \gamma
\end{aligned}$$

where the second equality follows from the fact that the events are mutually exclusive, and the last inequality follows from the bounds established earlier about $L_{i,j}$.

$$\begin{aligned}
\mathbb{E}[X'(\{i\})] =& \Pr[X'(\{i\}) = 1] = \Pr[\bigvee_{j \in [n-t]} L_{i,j} \le t'_{\mathcal{H}}] \\
=& 1 - \Pr[\bigwedge_{j \in [n-t]} L_{i,j} > t'_{\mathcal{H}}] \\
=& 1 - \prod_{j \in [n-t]} \Pr[L_{i,j} > t'_{\mathcal{H}}] \\
\ge& 1 - (1-\alpha)^{n-t} = f
\end{aligned}$$

where the last inequality follows as before. $\dashv$

By the linearity of expectation we have that $\mathbb{E}[Y'(R)] \ge \gamma |R|$ and $\mathbb{E}[X'(R)] \ge f|R|$. Moreover, it easy to see that $\{Y'(\{i\})\}_{i \in R}$ are independent, and thus we can apply the Chernoff Bound:

$$\Pr[Y'(R) \le (1 - \frac{\delta}{4}) \gamma |R|] \le \Pr[Y'(R) \le (1 - \frac{\delta}{4}) \mathbb{E}[Y'(R)]] \le e^{-\Omega(\delta^2 \gamma |R|)}$$

Similarly, we can show that $\Pr[X'(R) \le (1 - \frac{\delta}{4}) f|R|] \le e^{-\Omega(\delta^2 f |R|)}$. Since $\hat{X}(R) \ge X'(R)$ and $\hat{Y}(R) \ge Y'(R)$, the same bounds hold for $\hat{X}(R), \hat{Y}(R)$.

Since the conditions of the above two events of the ideal execution can be checked in polynomial time, it follows that they should also hold for the real execution with negligible difference in probability. Otherwise, a distinguisher would be able to use them to efficiently distinguish between the two executions. The lemma follows.

□

Following the strategy of [30], we are now ready to define the set of *typical executions* for this setting.

Definition 26 (Typical execution). An execution is *typical* if and only if $\lambda \geq 9/\delta$ and for any set R of consecutive rounds with $|R| \geq \max\{4k_{\mathsf{hard}}, \lambda\}/\gamma\delta^2$, the following hold:

1. $Y(R) > (1 - \frac{\delta}{4})\gamma|R|$ and $X(R) > (1 - \frac{\delta}{4})f|R|$;
2. for any party P, any round $i \in R$: $Z_i^P(R) < \frac{\gamma}{2(1+\delta)} \cdot |R|$ and $Z_i^P(R) + X_i^P(R) < \beta(t'_{\mathcal{A}} + nt'_{\mathcal{H}}) \cdot |R|$; and
3. no insertions and no copies occurred.

Theorem 27. *An execution is typical with probability* $1 - \mathsf{negl}(\lambda)$.

Proof. In order for an execution to not be typical, one of the three items of Definition 26 must not hold with non-negligible probability for some big enough set of rounds. Point 3 is implied by Corollary 21. For a specific set of rounds R, where $|R| \geq \frac{2\lambda}{\gamma\delta^2}$, item 1 is implied by Lemma 25 with overwhelming probability in λ.

Regarding item 2, by an application of Lemma 23 for $t'_{\mathcal{A}} = \frac{\gamma}{2(1+\delta)\beta}$, it follows that $Z_i^P(R) < \frac{\gamma}{2(1+\delta)} \cdot |R|$ with probability $\mathsf{negl}(\beta \cdot t'_{\mathcal{A}} \cdot |R|)$, where $\frac{k_{\mathsf{hard}}}{\beta t'_{\mathcal{A}}} \leq \frac{4k_{\mathsf{hard}}}{\gamma\delta^2} \leq |R|$. Note, that due to Assumptions 1 and 2 and Lemma 22, necessarily $t'_{\mathcal{A}} \leq \frac{\gamma}{2(1+\delta)\beta}$. Similarly, Corollary 24 implies that $Z_i^P(R) + X_i^P(R) < \beta(t'_{\mathcal{A}} + nt'_{\mathcal{H}})|R|$ with overwhelming probability in λ, since $\frac{k_{\mathsf{hard}}}{\beta(t'_{\mathcal{A}} + nt'_{\mathcal{H}})} \leq \frac{4k_{\mathsf{hard}}}{\gamma} \leq |R|$. Hence, item 2 also follows with overwhelming probability in λ.

Finally, we can bound the probability that an execution is not typical by applying the union bound on the negation of these events over all sets of consecutive rounds of sufficiently large size, where the probability of each event occurring is negligible in λ. □

Having established that typical rounds happen with overwhelming probability, the rest of the proof follows closely that of [30]. The only difference is that to prove the corresponding common-prefix lemma, although we can match blocks mined in uniquely successful rounds to adversarial blocks in one of the two chains that constitute the fork, the typicality of the execution only provides a bound on the maximum number of blocks in a single chain. Hence, only half of the blocks matched must outnumber the uniquely successful rounds in this interval, which is also the reason why our proof only works with an adversary controlling up to 1/3 of the parties. Next, we describe the changes one has to do after proving the typical execution theorem with respect to the analysis of [30], in order to prove the security of the protocol in our model. We only give brief proof sketches of lemmas and theorems from [30] that are exactly the same for our own setting.

Security properties of the blockchain. First, we describe a number of desired basic properties for the blockchain introduced in [30, 40, 47]. At a high level, the first property, called *common prefix*, has to do with the existence, as well as persistence in time, of a common prefix of blocks among the chains of honest players. Here we will consider a stronger variant of the property, presented in [47], which allows for the black-box proof of application-level properties (such as the *persistence* of transactions entered in a public transaction ledger built on top of the Bitcoin backbone). We will use $\mathcal{C} \preceq \mathcal{C}'$ to denote that some chain $\mathcal{C}$ is a prefix of some other chain $\mathcal{C}'$, and $\mathcal{C}^{\lceil k}$ to denote the chain resulting from removing the last k blocks of $\mathcal{C}$. We will call a block honest, if it was diffused for the first time in the execution by some honest party, and adversarial otherwise.

Definition 28 ((Strong) Common Prefix). The *strong common prefix* property Q_{cp} with parameter $k \in \mathbb{N}$ states that the chains $\mathcal{C}_1, \mathcal{C}_2$ reported by two, not necessarily distinct honest parties P_1, P_2, at rounds r_1, r_2 in $\text{VIEW}^{t,n}_{\Pi,\mathcal{A},\mathcal{Z}}$, with $r_1 \leq r_2$, satisfy $\mathcal{C}_1^{\lceil k} \preceq \mathcal{C}_2$.

The next property relates to the proportion of honest blocks in any portion of some honest player's chain.

Definition 29 (Chain Quality). The *chain quality* property Q_{cq} with parameters $\mu \in \mathbb{R}$ and $k \in \mathbb{N}$ states that for any honest party P with chain $\mathcal{C}$ in $\text{VIEW}^{t,n}_{\Pi,\mathcal{A},\mathcal{Z}}$, it holds that for any k consecutive blocks of $\mathcal{C}$ the ratio of adversarial blocks is at most μ.

Further, in the derivations in [30] an important lemma was established relating to the rate at which the chains of honest players were increasing as the Bitcoin backbone protocol was run. This was explicitly considered in [40] as a property under the name *chain growth.*

Definition 30 (Chain Growth). The *chain growth* property Q_{cg} with parameters $\tau \in \mathbb{R}$ (the "chain speed" coefficient) and $s, r_0 \in \mathbb{N}$ states that for any round $r > r_0$, where honest party P has chain $\mathcal{C}_1$ at round r and chain $\mathcal{C}_2$ at round $r+s$ in $\text{VIEW}^{t,n}_{\Pi,\mathcal{A},\mathcal{Z}}$, it holds that $|\mathcal{C}_2| - |\mathcal{C}_1| \geq \tau \cdot s$.

Security Analysis. Next, we proceed to the security analysis. We first prove that the rate at which the adversary generates blocks in any big enough round interval, is at most half the rate of uniquely successful rounds. This relation is going to be at the center of the security proof we are going to develop next.

Lemma 31. *Assume a typical execution. For any set of consecutive rounds $R = \{i, \ldots, j\}$, where $|R| \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta^2}$, and for any party P and round $r \in R$, it holds that $(1-\frac{\delta}{4})\frac{Y(R\setminus\{i\})}{2} > Z_r^P(R)$.*

Proof. It holds that:

$$(1-\frac{\delta}{4})\frac{Y(R\setminus\{i\})}{2} > (1-\frac{\delta}{4})^2\frac{\gamma}{2}(|R|-1) \geq \frac{\gamma|R|}{2(1+\delta)} > Z_r^P(R)$$

where, the first and last inequalities follow from the assumption that the execution is typical, while the middle one follows from the fact that $|R| \geq 9/\delta \geq (1-\frac{1}{(1-\delta/4)^2(1+\delta)})^{-1}$. ☐

Lemma 32. *(Chain-Growth Lemma). Suppose that at round r an honest party has a chain of length ℓ. Then, by round $s \geq r$, every honest party has adopted a chain of length at least $\ell + X(\{r, \ldots, s-1\})$.*

Proof. The main idea of the proof of this lemma is that, after each successful round at least one honest party will have received a chain that is at least one block longer than the chain it had, and all parties pick only chains that are longer than the ones they had. ☐

Theorem 33 (Chain Growth). *In a typical execution the chain-growth property holds with parameters $\tau = (1-\frac{\delta}{4})f$ and $s \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta^2}$.*

Proof. Let R be any set of at least s consecutive rounds. Then, since the execution is typical: $X(R) \geq (1-\frac{\delta}{4})f \cdot |R| \geq \tau \cdot |R|$. By Lemma 32, each honest player's chain will have grown by that amount of blocks at the end of this round interval. Hence, the chain growth property follows. ☐

Lemma 34. *Assume a typical execution. Let B be some honest or the genesis block. Any sequence of $k \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$ consecutive blocks in some chain $\mathcal{C}$, where the first block in the sequence directly descends B, have been computed in at least k/δ rounds, starting from the round that B was computed.*

Proof. W.l.o.g, let B be an honest block computed by party P at round i. First, note that due to Assumption 2, it holds that $\beta(t'_{\mathcal{A}} + nt'_{\mathcal{H}}) < \delta$. For the sake of contradiction, assume that for some $k \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$, there is a set of rounds R', such that $|R'| < k/\delta$ and at least k blocks that descend block B have been computed during R'. This implies that there is a set of rounds R, where $|R| \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta^2}$, such that $X_i^P(R) + Z_i^P(R) \geq k \geq |R|\delta > |R|\beta(t'_{\mathcal{A}} + nt'_{\mathcal{H}})$. This contradicts the typicality of the execution, and the lemma follows. □

Lemma 35 (Common Prefix Lemma). *Assume a typical execution and consider two chains $\mathcal{C}_1$ and $\mathcal{C}_2$ such that $len(\mathcal{C}_2) \geq len(\mathcal{C}_1)$. If $\mathcal{C}_1$ is adopted by an honest party at round r, and $\mathcal{C}_2$ is either adopted by an honest party or diffused at round r, then $\mathcal{C}_1^{\lceil k} \leq \mathcal{C}_2$ and $\mathcal{C}_2^{\lceil k} \leq \mathcal{C}_1$, for $k \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$.*

Proof. The proof in [30] shows that for every block mined at a uniquely successful round, there exists an adversarial block in one of the two chains. This in turn implies that one of the two chain has a number of adversarial blocks that is at least as big as half the number of uniquely successful rounds. Using the previous lemma the proof proceeds as in [30], reaching a contradiction with Lemma 31. Note, that all adversarial blocks in the matching between uniquely successful rounds and adversarial blocks are descendants of the last honest block in the common prefix of $\mathcal{C}_1$ and $\mathcal{C}_2$. □

Theorem 36 (Common Prefix). *In a typical execution the common-prefix property holds with parameter $k \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$.*

Proof. The main idea of the proof is that if there exists a deep enough fork between two chains, then the previously proved lemma cannot hold. Hence, the theorem follows. □

Theorem 37 (Chain Quality). *In a typical execution the chain-quality property holds with parameter $\mu < 1 - \delta/4$ and $\ell \geq \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$.*

Proof. The main idea of the proof is the following: a large enough number of consecutive blocks will have been mined in a set rounds that satisfies the properties of Definition 26. Hence, the number of blocks that belong to the adversary will be upper bounded, and all other blocks will have been mined by honest parties. □

Finally, the Consistency and Liveness properties follow from the three basic properties, albeit with different parameters than in [30].

Lemma 38 (Consistency). *It holds that $\Pi_{\mathsf{PL}}^{\mathsf{new}}$ with $k = \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$ satisfies Persistence with overwhelming probability in λ.*

Proof. The main idea is that if consistency is violated, then the common-prefix property will also be violated. Hence, if the execution is typical the lemma follows. □

Lemma 39 (Liveness). *It holds that $\Pi_{\mathsf{PL}}^{\mathsf{new}}$ with $k = \frac{\max\{4k_{\mathsf{hard}},\lambda\}}{\gamma\delta}$ and $u = \frac{2k}{(1-\frac{\delta}{4})f}$ rounds satisfies Liveness with overwhelming probability in λ.*

Proof. The main idea here is that after u rounds at least $2k$ successful rounds will have occurred in a typical execution. Thus, by the chain growth lemma the chain of each honest party will have grown by $2k$ blocks, and by the chain quality property at least one of these blocks that is deep enough in the chain is honest. □

Next, we state our theorem. Note that both Consistency and Liveness depend on the convergence parameter k_{hard} of $\mathcal{I}$.

Theorem 40. *Assuming the existence of a collision-resistant hash function family, a one-way trapdoor permutation and a dense cryptosystem (for the NIZK), and a secure ISP problem $\mathcal{I}$ that comply with Assumptions 1 and 2, protocol $\Pi_{\mathsf{PL}}^{\mathsf{new}}$ implements a robust public transaction ledger with parameters $k = \max\{4k_{\mathsf{hard}}, \lambda\}/\gamma\delta$ and $u = 2k/(1-\frac{\delta}{4})f$, except with negligible probability in λ.*

4.5 Realizing ISPs from Non-Idealized Hash Functions

Next, we present a secure ISP problem assuming the existence of a hash function that satisfies both the computational extraction and iterated hardness properties presented in Section 3.

Construction 2. Let $\mathcal{H}$ be a hash function family as in Definitions 10 and 11. Let $T \in \{0,1\}^\lambda$ be a hardness parameter. An instance of a secure ISP is as follows:

- I_λ is the uniform distribution over $K(\lambda)$, i.e., $\Lambda = \{k\}$;
- $X = \{0,1\}^\lambda, W = \{0,1\}^{2(d-1)\lambda}$;
- $R = \{(x,w) | H_k(x||w_1) < T$ for $w = w_1||w_2\}$;
- $M(x,1)$ iteratively samples w_1 from $\mathcal{U}_{(d-1)\lambda}$, and tests whether $H_k(x||w_1) < T$, until it finds a solution. It then samples a uniformly random w_2 from $\mathcal{U}_{(d-1)\lambda}$, and outputs $w_1||w_2$.
- $S(x,w) = H_k(H_k(x||w_1)||w_2)$, for $w = w_1||w_2$.

Construction 2 is similar to Bitcoin's ISP construction (see Section 4.1, Construction 1), with the following differences:

1. In our construction $H_k(x||w_1)$ is required to be smaller than the hardness parameter T, while in Bitcoin $H_k(H_k(x||w_1)||w_2)$ is expected to be small, where w_1 is the hash of some message. This change allows a party who already knows a witness (w_1, w_2) for some statement, to produce a new one by changing w_2 arbitrarily.
2. Each time M tests a new possible witness, w_1 is sampled randomly, instead of just being increased by one, as in Bitcoin. This will help us later on to argue that each test succeeds with probability proportional to T.

Obviously, if used in "native" Bitcoin this construction is totally insecure, as by the time an honest party publishes a block, anyone can compute another valid block with minimal effort. However, it is good enough for our new protocol, where the witnesses are not exposed, and thus only a party who knows a witness can generate new witnesses for free. Next, we argue the security of the construction.

Assuming $\mathcal{H}$ is a computational randomness extractor is sufficient for the security properties that make up a secure ISP, besides hardness, to be satisfied. First, the fact that $H_k(x||w_1)$ is computationally indistinguishable from uniform, for any $x \in X$, implies that the runtime and the output of M are computationally indistinguishable from a process that sampled repeatedly a uniform value from $\{0,1\}^\lambda$ until it finds one that is smaller than T. This in turn implies that the runtime distribution of M is indistinguishable from the geometric distribution with parameter $T/2^\lambda$, and thus the successful ISP property is satisfied. Further, since w_2 is also chosen uniformly at random, we can show that a simulator that samples a random value from $\mathcal{U}_\lambda$ and the geometric distribution, satisfies the next-problem simulatability property. Finally, by resampling a new w_2 uniformly at random, an admissible witness is produced, and the witness malleability property follows. Thus, we are able to state the following lemma.

Lemma 41. *If $\mathcal{H}$ is a CRE hash family (Definition 10), then Construction 2 is $O(\lambda)$-next-problem simulatable, $O(\lambda)$-witness malleable, and $(t, \mathcal{C}_{T/2^\lambda}(O(t)))$-successful for any $t \in poly(\lambda)$, where $\mathcal{C}_{T/2^\lambda}$ is the cumulative geometric distribution with parameter $T/2^\lambda$.*

Proof. We start by showing that even if the adversary chooses the problem statement x maliciously, hashing it once together with a uniformly random string, will result in a string that is computationally indistinguishable from a uniformly sampled string. Fix some $\lambda \in \mathbb{N}$, $k \in K(\lambda)$ and $x \in X$. Let random variable Z be equal to $H_k(x||\mathcal{U}_{c\lambda}||\mathcal{U}_{(d-c-1)\lambda})$. By our assumption that $E(x,i) \stackrel{def}{=} H_k(x||i)$ is a $(c\lambda)$-computational extractor, and since $x||\mathcal{U}_{c\lambda}$ has $c\lambda$ bits of min-entropy, it follows that[14] $Z \stackrel{c}{\approx} \mathcal{U}_\lambda$. Moreover, x is fixed, hence $(x, Z) \stackrel{c}{\approx} (x,\mathcal{U}_\lambda)$. Assume instead that x is sampled from some efficiently samplable distribution $\hat{\mathcal{X}}$, as it will be the case in an actual execution, and let $\hat{Z} = H_k(\hat{\mathcal{X}}||\mathcal{U}_{c\lambda}||\mathcal{U}_{(d-c-1)\lambda})$. For any PPT distinguisher D, sufficiently large $\lambda \in \mathbb{N}$, all $z \in \{0,1\}^{poly(\lambda)}$ it holds that:

$$
\begin{aligned}
&|\Pr[D(1^\lambda, z, \hat{\mathcal{X}}, \hat{Z}) = 1] - \Pr[D(1^\lambda, z, \hat{\mathcal{X}}, \mathcal{U}_\lambda) = 1]| \\
&= |\sum_{x' \in X} \Pr[x' = \hat{\mathcal{X}}]\big(\Pr[D(1^\lambda, z, x', H_k(x'||\mathcal{U}_{(d-1)\lambda})) = 1] - \Pr[D(1^\lambda, z, x', \mathcal{U}_\lambda) = 1]\big)| \\
&\leq \sum_{x' \in X} \Pr[x' = \hat{\mathcal{X}}] \cdot |\Pr[D(1^\lambda, z, x', H_k(x'||\mathcal{U}_{(d-1)\lambda})) = 1] - \Pr[D(1^\lambda, z, x', \mathcal{U}_\lambda) = 1]| \\
&\leq \sum_{x' \in X} \Pr[x' = \hat{\mathcal{X}}] \cdot \mathsf{negl}(\lambda) \leq \mathsf{negl}(\lambda)
\end{aligned}
$$

where the last inequality follows from the fact that $(x, Z) \stackrel{c}{\approx} (x,\mathcal{U}_\lambda)$ for any $x \in X$. Hence, $(\hat{\mathcal{X}}, \hat{Z}) \stackrel{c}{\approx} (\hat{\mathcal{X}},\mathcal{U}_\lambda)$.

We next argue about the distribution of the running time of M. Algorithm M on input x iteratively samples a uniformly random $w_1||w_1'$ from $\mathcal{U}_{c\lambda} \times \mathcal{U}_{(d-c-1)\lambda}$, and tests whether $H_k(x||w_1||w_1') < T$, until it finds a solution. For a moment, assume that M instead tested whether a value sampled from $\mathcal{U}_\lambda$ is smaller than T. Then, its running time would be distributed according to the geometric distribution $\mathcal{G}_p$ with parameter $p = T/2^\lambda$. Since $Z \stackrel{c}{\approx} \mathcal{U}_\lambda$, we can use a hybrid argument to show that the distribution of $\mathsf{Steps}_M(x)$ is computationally indistinguishable from $c_1 \cdot \mathcal{G}_{T/2^\lambda} + c_2$, where c_1 is a constant related to the cost of sampling a random value for each test and evaluating H, and c_2 to the cost of sampling w_2. The hybrid argument proceeds by replacing a computation of $H_k(x||\mathcal{U}_{(d-1)\lambda}) < T$ at some step of M, with $\mathcal{U}_\lambda < T$. If between any two hybrids the distributions of the runtime of the respective modified M is not computationally indistinguishable, then we can easily construct a distinguisher for $H_k(x||\mathcal{U}_{(d-1)\lambda})$ and $\mathcal{U}_\lambda$, which contradicts our previous analysis. Hence, $\mathsf{Steps}_M(x)$ should be computationally indistinguishable from $c_1 \cdot \mathcal{G}_{T/2^\lambda} + c_2$. It follows that M must be $(t, \mathcal{C}_{T/2^\lambda}(O(t)))$-successful, for any $t \in poly(\lambda)$.

Next, note that M, after finding a small hash, hashes again the result with a fresh randomly sampled string w_2. Using the same hybrid argument as in the previous paragraph we can show that $(x, M(x), \mathsf{Steps}_M(x)) \stackrel{c}{\approx} (x, H_k(\mathcal{W}||\mathcal{U}_\lambda), c_1 \cdot \mathcal{G}_{T/2^\lambda} + c_2)$, where $\mathcal{W}$ is the uniform distribution over the hash images that are smaller than T. By our previous analysis it follows that $(x, H_k(\mathcal{W}||\mathcal{U}_\lambda), c_1 \cdot \mathcal{G}_{T/2^\lambda} + c_2) \stackrel{c}{\approx} (x, \mathcal{U}_\lambda, c_1 \cdot \mathcal{G}_{T/2^\lambda} + c_2)$. By the transitivity of computational indistinguishability it follows that the simulator Ψ that outputs a randomly sampled pair from $\mathcal{U}_\lambda$ and $c_1 \cdot \mathcal{G}_{T/2^\lambda} + c_2$ satisfies the next-problem simulatability property. Note, that using the inverse transform technique, we can sample from the geometric distribution (truncated to 2^λ) in $O(\lambda)$ steps.

Finally, the witness malleability property holds for $\Phi(x, (w_1, w_2))$ that outputs the witness (w_1, w_2'), where w_2' is sampled uniformly at random. Again, $S(x, \Phi(x, (w_1, w_2)))$ will be indistinguishable from $\mathcal{U}_\lambda$. The lemma follows.

□

[14] We abuse the notation and use the $\stackrel{c}{\approx}$ relation with random variables, instead of random variable ensembles. The relevant ensembles can be easily deduced.

Regarding the hard-ISP property, we are going to take advantage of the iterative hardness of Bitcoin's ISP construction and the fact that Construction 2 is closely related to it. The main idea is that if there exists an attacker against our construction, then we can use it to break the iterative hardness property (Definition 11). In more detail, given as input a statement x, the iterated hardness attacker runs the attacker of our construction with input $H(x||w)$, where w is sampled at random. It is easy to see that if $((w_1, w'_1), \ldots, (w_m, w'_m))$ are the witnesses it is going to produce, then $((w, w_1), (w'_1, w_2), \ldots, (w'_{m-1}, w_m))$ are valid witnesses for Construction 1, and also against the iterative hardness property. The following lemma highlights this relation.

Lemma 42. *Assume Construction 2 is based on a hash family $\mathcal{H}$ that is CRE and t-iteratively hard. Then, for some polynomial $k_0(\cdot)$, any $\sigma \in (0,1)$ and $t' = \frac{2^\lambda}{(1-\sigma)T}$, Construction 2 is $(t', 1-t'/t, k_0)$-hard.*

Proof. Due to the fact that $\mathcal{H}$ is CRE, we can show that our construction satisfies the first item of Definition 13 by a simple application of the Chernoff bound as in Lemma 25. Each witness test succeeds with probability $T/2^\lambda$, and thus for any $\sigma \in (0,1)$ it holds that for any $k \geq \lambda$, $M(x,k)$ will finish in less than kt' steps with probability negligible in λ.

Regarding the second item, as mentioned Construction 2 is a mirror image of Construction 1, in the sense that the first hash, instead of the second, is required to be smaller than T, and the other one can have an arbitrary value. Hence, we will reduce the hardness of our construction to the iterative hardness property. Let k_0 be equal to the maximum between the polynomial described in Definition 11 and λ. For the sake of contradiction, assume that Construction 2 is not $(t', 1-t'/t, k_0)$-hard. It has to be the case that there exists an attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ that for infinitely many λ and some $m \geq k_0$ breaks the hardness of Construction 2. Using $\mathcal{A}$, we are going to construct an attacker $\mathcal{A}'$ that breaks the iterative hardness property.

Let $\mathcal{A}'$ work as follows: First, $\mathcal{A}'_1$ runs $\mathcal{A}_1$ and forwards variable st to $\mathcal{A}'_2$. Then, $\mathcal{A}'_2$ on input st and a randomly sampled problem statement x, runs $\mathcal{A}_2$ on input $H(x||w)$, where w is sampled at random. If $\mathcal{A}_2$ succeeds, it outputs witnesses $((w_1, w'_1), \ldots, (w_m, w'_m))$. Then, $\mathcal{A}'_2$ outputs $((w, w_1), \ldots, (w'_{m-1}, w_m))$. Note, that in that case (w, w_1) is a witness for x, for the game $\mathcal{A}'$ is playing, since $H(H(x||w)||w_1) < T$. Moreover, it should hold that $H(H(H(H(x||w)||w_1)||w'_1)||w_2) < T$. In turn, this implies that $((w, w_1), (w'_1, w_2))$ is a valid sequence of witnesses for $\mathcal{A}'$. Similarly, it follows that $((w, w_1), \ldots, (w'_{m-1}, w_m))$ is a valid sequence of m witnesses for the game $\mathcal{A}'$ is playing. Hence, whenever $\mathcal{A}$ wins, $\mathcal{A}'$ also wins.

We proceed to analyze the winning probability of $\mathcal{A}'$. We have already argued that whenever $\mathcal{A}$ wins, $\mathcal{A}'$ also wins. Moreover, we have assumed that $\mathcal{A}$ succeeds in producing $m \geq k$ witnesses with non-negligible probability. Due to the randomness extraction property of $\mathcal{H}$, the distribution of $H(x,w)$ will be computationally indistinguishable from the uniform distribution over $\{0,1\}^\lambda$. Hence, the probability that $\mathcal{A}$ wins is negligibly close to the probability that it wins on a uniformly random input, and thus $\mathcal{A}'$ wins also with non-negligible probability. This is a contradiction, and the lemma follows.

□

Due to Theorem 40 and the previous two lemmas, we can implement a ledger assuming the existence of a robust NIZK, a hash family that is collision-resistant, another hash function family that is both CRE and iteratively hard for appropriate parameters, and that the adversary controls less than a third of the total computational power. The following theorem holds.[15]

[15]For simplicity, we assume that the cost in computational steps of evaluating H, and the hidden constant in the successful property of Lemma 41 are both 1. The theorem can be easily generalized for arbitrary costs.

Theorem 43. *Assuming the existence of collision-resistant hash functions, a hash function family that is CRE and t_{hard}-iteratively hard, a one-way trapdoor permutation and a dense cryptosystem (for the NIZK), and that for some $\delta_{\mathsf{Steps}} \in (0,1)$, sufficiently large $\lambda \in \mathbb{N}$, and T equal to $\lfloor 2^\lambda \cdot \min\{\frac{\ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1})}{t'_{\mathcal{H}}}, \frac{\delta_{\mathsf{Steps}}/4}{(t'_{\mathcal{A}}+nt'_{\mathcal{H}})(1+\delta_{\mathsf{Steps}}/2)}\}\rfloor$ it holds that :*
- $t_{\mathsf{hard}} \geq (1+\delta_{\mathsf{Steps}}/2)^{-1} \cdot \frac{2^\lambda}{T}$*; and*
- $2 \cdot t'_{\mathcal{A}} \leq (1-\delta_{\mathsf{Steps}}) \cdot (n-t)t'_{\mathcal{H}}$

protocol $\Pi_{\mathsf{PL}}^{\mathsf{new}}$ based on Construction 2 implements a robust public transaction ledger, except with negligible probability in λ.

Proof. We have already proved in Lemmas 41 and 42 that Construction 2 satisfies all properties described in Assumption 1. It remains to argue about Assumption 2.

Set $\delta = \delta_{\mathsf{Steps}}/4$ and $\delta_{\mathsf{ISP}} = \delta_{\mathsf{Steps}}/2$. It follows that $\frac{\delta_{\mathsf{Steps}}-\delta_{\mathsf{ISP}}}{2} \geq \delta$. Let $p_h = T/2^\lambda$. It holds that:

$$\begin{aligned}\alpha \geq (1-\delta_{\mathsf{ISP}})\beta t'_{\mathcal{H}} &\Leftrightarrow 1-(1-p_h)^{t'_{\mathcal{H}}} \geq (1-\delta_{\mathsf{ISP}})t_{\mathsf{hard}}^{-1}t'_{\mathcal{H}}\\ &\Leftarrow (1-p_h)^{t'_{\mathcal{H}}} + (1-\delta_{\mathsf{ISP}})(1+\delta_{\mathsf{Steps}}/2)\cdot p_h t'_{\mathcal{H}} \leq 1\\ &\Leftarrow e^{-p_h t'_{\mathcal{H}}} + (1-\delta_{\mathsf{Steps}}^2/4)\cdot p_h t'_{\mathcal{H}} \leq 1\end{aligned}$$

Now, let $f(u) = e^{-u} + (1-\delta_{\mathsf{Steps}}^2/4)u$. It holds that $f(0)=1$, $\frac{df}{du} = -e^{-u} + (1-\delta_{\mathsf{Steps}}^2/4)$, $\frac{df}{du}(\ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1}) = 0$, and $\frac{df}{du}$ is strictly increasing as u grows. Since $(1-\delta_{\mathsf{Steps}}^2/4) \in (0,1)$, it follows that $\ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1}) > 0$, which further implies that f is decreasing in $(0, \ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1})]$, and thus for any positive $u < \ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1})$ it follows that $f(u) < 1$. Hence, for $T/2^\lambda = p_h \leq \frac{\ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1})}{t'_{\mathcal{H}}}$, it follows that $f(T/2^\lambda \cdot t'_{\mathcal{H}}) < 1$, which implies that $\alpha(h) \geq (1-\delta_{\mathsf{ISP}})\beta t'_{\mathcal{H}}$, for any $t'_{\mathcal{H}}$.

Hence, if we set T less than $2^\lambda \cdot \min\{\frac{\ln((1-\delta_{\mathsf{Steps}}^2/4)^{-1})}{t'_{\mathcal{H}}}, \frac{\delta_{\mathsf{Steps}}/4}{(t'_{\mathcal{A}}+nt'_{\mathcal{H}})(1+\delta_{\mathsf{Steps}}/2)}\}$, the preconditions of Theorem 40 for Construction 2 are satisfied. Note, that we assume that λ is large enough so that (i) $\lambda > 9/\delta$, (ii) there exists a T that satisfies the inequality above and $T/2^\lambda$ is non-negligible. ☐

5 Iterated Hardness is Necessary

In this section we demonstrate that an attack against iterated hardness implies an attack against the Bitcoin protocol, assuming the underlying hash function is CRE (Definition 10). We phrase our attack, against an abstraction of the Bitcoin protocol which appeared in [30], and from which it is easy to extract a version of the protocol for our model, where honest parties run the PoW generation algorithm (presented in Construction 1) for $t_{\mathcal{H}}$ steps, while the adversary takes $t_{\mathcal{A}}$ steps per round.

Theorem 44. *Let $n, t, t_{\mathcal{H}}, t_{\mathcal{A}}$ such that $t_{\mathcal{A}} = c\cdot(n-t)t_{\mathcal{H}}$, for some $c \in (0,1)$. If $\mathcal{H}$ is CRE, and the Bitcoin protocol satisfies Consistency with parameter k, then $\mathcal{H}$ is $\frac{c}{2}\cdot\frac{(n-t)t_{\mathcal{H}}}{(1-T/2^\lambda)^{(n-t)t_{\mathcal{H}}}}$-iteratively hard, for any polynomial k.*

Proof. Let $t' = \frac{c}{2}\cdot\frac{(n-t)t_{\mathcal{H}}}{(1-T/2^\lambda)^{(n-t)t_{\mathcal{H}}}}$. For the sake of contradiction, assume that the theorem does not hold. Then, for any polynomially large k_0, there exists an adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ and some $k' \geq k_0$, such that $\mathcal{A}_2$ can compute k' witnesses in less than $k't'$ steps (see Definition 11) with non-negligible probability. For our case take $k_0 = k+2$. We are going to construct an adversary $\mathcal{A}'$ that breaks the Consistency of the Bitcoin protocol.

We start by describing our attacker. $\mathcal{A}'$ gets as input the CRS, and runs $\mathcal{A}_1$ with input the hash key. When $\mathcal{A}_1$ finishes with output st, $\mathcal{A}'$ waits for a new honest block to be mined, say at height l of some chain. When this happens, it invokes $\mathcal{A}_2$ with input the hash of this block. Upon $\mathcal{A}_2$ finishing, it waits until an honest party computes a block at height $l+k+1$, sends the chain created by $\mathcal{A}_2$ to a different honest party, and halts.

We will argue that $\mathcal{A}'$ breaks Consistency. We make two simplifying assumptions in our analysis. First, we assume that in the PoW construction of the Bitcoin protocol (Construction 1), honest parties sample every time a uniformly random counter, instead of increasing it. Otherwise we have to assume that $\mathcal{H}$ is a PRF instead of a CRE for the analysis to go through. Secondly, we assume that the witnesses output by $\mathcal{A}_2$ constitute valid blocks. Otherwise, we have to rephrase the iterative hardness assumption to capture Bitcoin transaction semantics.

We proceed with the analysis. Similarly to Lemma 25, honest parties mine chains of length at most $(1+\delta)f \cdot s$ in s rounds, for any large enough s, and $\delta \in (0,1)$. Moreover, due to the CRE property, the hash of the block provided as input to $\mathcal{A}_2$ is going to be computationally indistinguishable from a uniformly sampled element from $[0,T]$; the hash of the blocks of hte protocol has to be less or equal to T. Hence, $\mathcal{A}_2$ is going to compute k' witnesses in less than $k't'$ steps (less than $k't'/t_{\mathcal{A}}$ rounds), with non-negligible probability. Putting everything together, at the round $\mathcal{A}_2$ finishes, honest parties will have computed at most:

$$(1+\delta)f \cdot k't'/t_{\mathcal{A}} \leq \frac{(1+\delta)(1-T/2^{\lambda})^{(n-t)t_{\mathcal{H}}}}{t_{\mathcal{A}}} \cdot \frac{c \cdot (n-t)t_{\mathcal{H}}}{2(1-T/2^{\lambda})^{(n-t)t_{\mathcal{H}}}} k' < k'$$

blocks, for $\delta \in (0,1)$ and with overwhelming probability. Hence, an honest party will later compute a block at height $l+k+1$, at which point $\mathcal{A}'$ will send its longer chain to another honest party, and break Consistency, since their blocks at position $l+1$ will differ with overwhelming probability. $\square$

As expected, as the computational power of the adversary decreases, the iteratively hard hash function needs to be less secure.

References

[1] J. Alwen and B. Tackmann. Moderately hard functions: Definition, instantiations, and applications. In *TCC 2017*, pages 493–526, 2017.

[2] M. Andrychowicz and S. Dziembowski. Pow-based distributed cryptography with no trusted setup. In R. Gennaro and M. Robshaw, editors, *CRYPTO 2015*, volume 9216 of *Lecture Notes in Computer Science*, pages 379–399. Springer, 2015.

[3] J. Aspnes, C. Jackson, and A. Krishnamurthy. Exposing computationally-challenged Byzantine impostors. Technical Report YALEU/DCS/TR-1332, Yale University Department of Computer Science, July 2005.

[4] A. Back. Hashcash–a denial of service counter-measure, 2002.

[5] C. Badertscher, U. Maurer, D. Tschudi, and V. Zikas. Bitcoin as a transaction ledger: A composable treatment. In *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part I*, pages 324–356, 2017.

[6] M. Ball, A. Rosen, M. Sabin, and P. N. Vasudevan. Proofs of work from worst-case assumptions. In *CRYPTO 2018*, pages 789–819, 2018.

[7] M. Bellare, R. Canetti, and H. Krawczyk. Pseudorandom functions revisited: The cascade construction and its concrete security. In *37th Annual Symposium on Foundations of Computer Science, FOCS '96, Burlington, Vermont, USA, 14-16 October, 1996*, pages 514–523, 1996.

[8] M. Bellare, A. Desai, E. Jokipii, and P. Rogaway. A concrete security treatment of symmetric encryption. In *FOCS '97*, pages 394–403, 1997.

[9] M. Bellare, J. Jaeger, and J. Len. Better than advertised: Improved collision-resistance guarantees for md-based hash functions. CCS '17, pages 891–906, New York, NY, USA, 2017. ACM.

[10] M. Bellare and P. Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *CCS '93*.

[11] M. Bellare and P. Rogaway. The exact security of digital signatures - how to sign with RSA and rabin. In Maurer [43], pages 399–416.

[12] D. J. Bernstein and T. Lange. Non-uniform cracks in the concrete: The power of free precomputation. In *ASIACRYPT 2013*, pages 321–340, 2013.

[13] N. Bitansky, S. Goldwasser, A. Jain, O. Paneth, V. Vaikuntanathan, and B. Waters. Time-lock puzzles from randomized encodings. In M. Sudan, editor, *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, Cambridge, MA, USA, January 14-16, 2016*, pages 345–356. ACM, 2016.

[14] D. Boneh, J. Bonneau, B. Bünz, and B. Fisch. Verifiable delay functions. In *CRYPTO 2018*, pages 757–788, 2018.

[15] D. Boneh and M. Naor. Timed commitments. In *Advances in Cryptology - CRYPTO 2000, 20th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2000, Proceedings*, pages 236–254, 2000.

[16] R. Canetti. Security and composition of multiparty cryptographic protocols. *J. Cryptology*, 13(1):143–202, 2000.

[17] R. Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 136–145. IEEE Computer Society, 2001.

[18] R. Canetti, Y. Chen, L. Reyzin, and R. D. Rothblum. Fiat-shamir and correlation intractability from strong kdm-secure encryption. In *EUROCRYPT 2018*.

[19] R. Canetti and M. Fischlin. Universally composable commitments. In *CRYPTO 2001*, pages 19–40, 2001.

[20] R. Canetti, O. Goldreich, and S. Halevi. The random oracle methodology, revisited. *J. ACM*, 51(4):557–594, July 2004.

[21] R. Cramer and V. Shoup. Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In L. R. Knudsen, editor, *Advances in Cryptology — EUROCRYPT 2002*, Berlin, Heidelberg, 2002.

[22] D. Dachman-Soled, R. Gennaro, H. Krawczyk, and T. Malkin. Computational extractors and pseudorandomness. In R. Cramer, editor, *Theory of Cryptography*, pages 383–403, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[23] I. Damgård. Collision free hash functions and public key signature schemes. In *Advances in Cryptology - EUROCRYPT '87, Workshop on the Theory and Application of of Cryptographic Techniques, Amsterdam, The Netherlands, April 13-15, 1987, Proceedings*, pages 203–216, 1987.

[24] I. Damgård. A design principle for hash functions. In *CRYPTO '89*, 1989.

[25] Y. Dodis, R. Gennaro, J. Håstad, H. Krawczyk, and T. Rabin. Randomness extraction and key derivation using the cbc, cascade and hmac modes. In M. Franklin, editor, *CRYPTO 2004*.

[26] J. R. Douceur. The sybil attack. In P. Druschel, M. F. Kaashoek, and A. I. T. Rowstron, editors, *IPTPS 2002*, volume 2429 of *Lecture Notes in Computer Science*. Springer, 2002.

[27] C. Dwork and M. Naor. Pricing via processing or combatting junk mail. CRYPTO '92, pages 139–147, London, UK, UK, 1993. Springer-Verlag.

[28] S. Faust, P. Mukherjee, J. B. Nielsen, and D. Venturi. Continuous non-malleable codes. In *TCC 2014*, pages 465–488, 2014.

[29] A. Fiat and A. Shamir. How to prove yourself: Practical solutions to identification and signature problems. In A. M. Odlyzko, editor, *Advances in Cryptology - CRYPTO '86*, volume 263 of *Lecture Notes in Computer Science*. Springer, 1986.

[30] J. A. Garay, A. Kiayias, and N. Leonardos. The bitcoin backbone protocol: Analysis and applications. In *Advances in Cryptology - EUROCRYPT 2015*, pages 281–310.

[31] J. A. Garay, A. Kiayias, and N. Leonardos. The bitcoin backbone protocol with chains of variable difficulty. In *CRYPTO 2017*, pages 291–323, 2017.

[32] J. A. Garay, A. Kiayias, N. Leonardos, and G. Panagiotakos. Bootstrapping the blockchain, with applications to consensus and fast PKI setup. In *PKC 2018*.

[33] J. A. Garay, A. Kiayias, and G. Panagiotakos. Consensus from signatures of work. Cryptology ePrint Archive, Report 2017/775. To appear in CT-RSA 2020.

[34] J. A. Garay, P. MacKenzie, M. Prabhakaran, and K. Yang. Resource fairness and composability of cryptographic protocols. *Journal of cryptology*, 24, 2011.

[35] C. Gentry and D. Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In L. Fortnow and S. P. Vadhan, editors, *STOC 2011*. ACM, 2011.

[36] O. Goldreich. *Foundations of Cryptography: Volume 1*. Cambridge University Press, New York, NY, USA, 2006.

[37] S. Goldwasser and Y. T. Kalai. On the (in)security of the fiat-shamir paradigm. In *Foundations of Computer Science (FOCS 2003)*. IEEE Computer Society, 2003.

[38] A. Juels and J. G. Brainard. Client puzzles: A cryptographic countermeasure against connection depletion attacks. In *NDSS 1999*. The Internet Society, 1999.

[39] J. Katz, A. Miller, and E. Shi. Pseudonymous secure computation from time-lock puzzles. *IACR Cryptology ePrint Archive*, 2014:857, 2014.

[40] A. Kiayias and G. Panagiotakos. Speed-security tradeoffs in blockchain protocols. Technical report, IACR: Cryptology ePrint Archive, 2015.

[41] H. Krawczyk. Cryptographic extraction and key derivation: The hkdf scheme. In T. Rabin, editor, *CRYPTO 2010*.

[42] A. K. Lenstra and B. Wesolowski. A random zoo: sloth, unicorn, and trx. Cryptology ePrint Archive, Report 2015/366, 2015. `https://eprint.iacr.org/2015/366`.

[43] U. M. Maurer, editor. *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, volume 1070 of *Lecture Notes in Computer Science*. Springer, 1996.

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

[45] J. B. Nielsen. Separating random oracle proofs from complexity theoretic proofs: The non-committing encryption case. In M. Yung, editor, *Advances in Cryptology - CRYPTO 2002*, volume 2442 of *Lecture Notes in Computer Science*. Springer, 2002.

[46] N. Nisan and D. Zuckerman. Randomness is linear in space. *J. Comput. Syst. Sci.*, 52(1):43–52, Feb. 1996.

[47] R. Pass, L. Seeman, and A. Shelat. Analysis of the blockchain protocol in asynchronous networks. In J. Coron and J. B. Nielsen, editors, *EUROCRYPT 2017*, volume 10211 of *Lecture Notes in Computer Science*, pages 643–673, 2017.

[48] D. Pointcheval and J. Stern. Security proofs for signature schemes. In Maurer [43].

[49] R. L. Rivest, A. Shamir, and D. A. Wagner. Time-lock puzzles and timed-release crypto. Technical report, Cambridge, MA, USA, 1996.

[50] A. D. Santis, G. D. Crescenzo, R. Ostrovsky, G. Persiano, and A. Sahai. Robust non-interactive zero knowledge. In *Advances in Cryptology - CRYPTO 2001*, 2001.