NIPoPoWs under Velvet Fork

1 Introduction

Since the release of Bitcoin about a decade ago, the interest in cryptocurrencies has increased tremendously, while a number of other "altcoins" have been constructed in the meantime. Given that cryptocurrencies are starting to be considered a generally accepted means of payment and are used for everyday transactions, the issue of efficiently handling cryptocurrencies by light clients, such as smartphones, has become of great importance.

In this work, we consider the problem of optimizing light clients, or "SPV clients" as described in the original Bitcoin paper[5]. As blockchains are ever growing, the main setback for efficient light client applications is the processing of data amount linear to the size of the blockchain, e.g. for synchronization purposes.

Our work is based on the construction of Non-Interactive Proofs of Proof of Work[1] that achieves SPV proofs of polylogarithmic portion of the blockchain size. The NIPoPoWs construction suggests a protocol update, that could be possibly implemented by a soft or a hard fork. Given the reluctancy of the Bitcoin community to proceed to such forks, we consider the case of a velvet fork[1][2], where it suffices only a portion of the total players to be updated.

Under this scope, our contributions come as follows:

- We revise the security proof for NIPoPoWs suffix proof protocol and compute a concrete value for the security parameter m
- We describe an attack, that we name *Chainsewing Attack*, against NIPoPoWs suffix proof construction, which is used for the light client's synchronization
- We suggest a patch to the NIPoPoWs protocol that eliminates the *Chainsewing Attack* and prove its security

2 Model Definition and Notation

3 NIPoPoWs under Soft or Hard Fork

3.1 Suffix Proofs

NIPoPoWs suffix proofs are used to prove predicates that pertain to the suffix of the blockchain. For example, this is the case of light client synchronization to the longest valid chain. [...]

3.1.1 Security of Suffix Proofs

In this section we provide the full security proof for the NIPoPoWs suffix proof protocol[1]. Apart from the proof itself (Theorem 2), we describe the definitions and lemmas being used. We try to give intuition for arguments and conclusions in each step.

Assume t adversarial out of n total parties, each with q PoW random oracle queries per round. We define $p = \frac{T}{2^\kappa}$ the probability of a successful Random Oracle query. We will call a query to the RO *μ-successful* if the RO returns a value h such that $h \leq 2^{-\mu}T$.

We define the boolean random variables X_r^μ, Y_r^μ, Z_r^μ. Fix some round r, query index j and adversarial party index k (out of t). If at round r an honest party obtains a PoW with $id < 2^{-\mu}T$, set $X_r^\mu = 1$, otherwise $X_r^\mu = 0$. If at round r exactly one honest party obtains $id < 2^{-\mu}T$, set $Y_r^\mu = 1$, otherwise $Y_r^\mu = 0$. If at round r the j-th query of the k-th corrupted party is μ-successful, set $Z_{rjk}^\mu = 1$, otherwise $Z_{rjk}^\mu = 0$. Let $Z_r^\mu = \sum_{k=1}^{t}\sum_{j=1}^{q} Z_{rjk}^\mu$. For a set of rounds S, let $X^\mu(S) = \sum_{r\in S} X_r^\mu$ and similarly define Y_S^μ, Z_S^μ.

Definition 1 (Typical Execution). *An execution of the protocol is (ϵ, η)-typical if:*

Block counts don't deviate. *For all $\mu \geq 0$ and any set S of consecutive rounds with $|S| \geq 2^\mu \eta k$, we have:*

- $(1-\epsilon)E[X^\mu(S)] < X^\mu(S) < (1+\epsilon)E[X^\mu(S)]$ and $(1-\epsilon)E[Y^\mu(S)] < Y^\mu(S)$
- $Z^\mu(S) < (1+\epsilon)E[Z^\mu(S)]$

Round count doesn't deviate. *Let S be a set of consecutive rounds such that $Z^\mu(S) \geq k$ for some security parameter k. Then $|S| \geq (1-\epsilon)2^\mu \frac{k}{pqt}$ with overwhelming probability.*

Chain regularity. *No insertions, no copies and no predictions [4] have occurred.*

Theorem 1 (Typicality). *Executions are (ϵ, η)-typical with overwhelming probability in k.*

Proof. **Block counts and regularity.** We refer to [4] for the full proof.

Round count. First, observe that for a specific round r we have $Z_{rjk} \sim Bern(p)$, so for the μ-level superblocks $Z_{rjk}^\mu \sim Bern(2^{-\mu}p)$ and these are jointly independent. Therefore, since for $|S|$ rounds we have $tq|S|$ adversarial RO queries, we have that $Z_S^\mu \sim \text{Bin}(tq|S|, 2^{-\mu}p)$. So $tq|S| \sim \text{NB}(Z_S^\mu, 2^{-\mu}p)$. Negative Binomial distribution is defined as $\text{NB}(r, p')$ and expresses the number of trials in a sequence of independent and identically distributed Bernoulli trials before a specified (r) number of successes

occurs. The expected total number of trials of a negative binomial distribution with parameters (r, p') is r/p'. To see this, imagine an experiment simulating the negative binomial performed many times, that is a set of trials is performed until r successes occur. Consider you perform n experiments of total N trials. Now we would expect $Np' = nr$, so $N/n = r/p'$. See that N/n is just the average number of trials per experiment. So we have $E[tq|S|] = \frac{Z_S^\mu}{2^{-\mu}p} \Rightarrow E[|S|] = 2^\mu \frac{Z_S^\mu}{tqp}$. So if $Z^\mu(S) \geq k$ then $E[|S|] \geq 2^\mu \frac{k}{tqp}$. Applying a tail bound to the negative binomial distribution, we obtain that $\Pr[|S| < (1-\epsilon)E(|S|)] \in \Omega(\epsilon^2 m)$.

Lemma 1. *Suppose S is a set of consecutive rounds $r_1 ... r_2$ and C_B is a chain adopted by an honest party at round r_2 of a typical execution. Let $C_S^B = \{\ b \in C_B : b$ was generated during $S\}$. Let $\mu_A, \mu_B \in \mathbb{N}$. Suppose $C_S^B \uparrow^{\mu_B}$ is good. Suppose C'_A is a μ_A-superchain containing only adversarially generated blocks generated during S and suppose that $|C'_A| \geq k$. Then $2^{\mu_A}|C'_A| < 2^{\mu_B}|C_S^B \uparrow^{\mu_B}|$.*

Proof. From $|C'_A| \geq k$ we have that $|Z_S^{\mu_A}| \geq k$. Applying Theorem 1, we conclude that $|S| \geq (1-\epsilon')2^{\mu_A}\frac{|C'_A|}{pqt}$. Applying the Chain Growth theorem [4] we obtain $|C_B^S| \geq (1-\epsilon)f|S|$. But from the goodness of $C_B^S \uparrow^{\mu_B}$, we know that $|C_B^S \uparrow^{\mu_B}| \geq (1-\delta)2^{-\mu_B}|C_B^S|$. So we have $|C_B^S \uparrow^{\mu_B}| \geq (1-\delta)2^{-\mu_B}(1-\epsilon)f|S|$ and follows that $|C_B^S \uparrow^{\mu_B}| \geq (1-\delta)2^{-\mu_B}(1-\epsilon)f(1-\epsilon')2^{\mu_A}\frac{|C'_A|}{pqt}$. Consequently we have that $2^{\mu_A}|C'_A| \leq \frac{pqt}{(1-\delta)(1-\epsilon)(1-\epsilon')f}2^{\mu_B}|C_B^S \uparrow^{\mu_B}|$.

So, according to the above equation we have that $2^{\mu_A}|C'_A| < 2^{\mu_B}|C_S^B \uparrow^{\mu_B}|$ considering that honest majority assumption holds, specifically considering that $\frac{pqt}{f} \approx \frac{t}{n-t} \leq 1$.

Definition 2 (Adequate level of honest proof). *Let π be an honestly generated proof constructed upon some adopted chain C and let $b \in \pi$. Then μ' is defined as $\mu' = \max\{\mu : |\pi\{b:\} \uparrow^\mu| \geq \max(m+1, (1-\delta)2^{-\mu}|\pi\{b:\} \uparrow^{\mu\downarrow}|)\}$. We call μ' the adequate level of proof π with respect to block b with security parameters δ and m. Note that the adequate level of a proof is a function of both the proof π and the chosen block b.*

Intuitively, adequate is the level μ' of a proof π for a block b if there are at least m blocks after b in π under the condition that there is good chain quality for this level, meaning that there are at least so many blocks at this level as expected considering the number of 0-level blocks.
NOTE: adequate level is mostly useful for Claim 1a of the Security Proof (Theorem 2).

Lemma 2. *Let π be some honest proof generated with security parameters δ, m. Let C be the underlying chain, $b \in C$ be any block and μ' be the adequate level of the proof with respect to b and the same security parameters.*
Then $C\{b:\} \uparrow^{\mu'} = \pi\{b:\} \uparrow^{\mu'}$.

Proof. $\pi\{b:\} \uparrow^{\mu'} \subseteq C\{b:\} \uparrow^{\mu'}$ is trivial. For the converse, we have that in the iteration of the *Prove for loop*[1] with $\mu = \mu^*$, the block stored in variable B precedes b in C.

Note that the Prover's for loop iterates over all levels in the interlink structure, and places in the proof all of the blocks that are of the corresponding level and succeed B in C.

Suppose $\mu = \mu^*$ is the first *for* iteration during which the property is violated. This cannot be the first iteration since $B = C[0]$ and Genesis precedes all blocks. By induction hypothesis we see that during the iteration $\mu = mu^* + 1$, B preceded b. From the definition of μ' we know that μ' is the highest level for which $|\pi\{b:\} \uparrow^{\mu}| \geq \max(m, (1-\delta)2^{-\mu}|\pi\{b:\} \uparrow^{\mu}\downarrow|)$.

Hence, this property cannot hold for $\mu^* > \mu$ and therefore $|\pi\{b:\} \uparrow^{\mu}| < m$ or $\neg\text{local-good}_{\delta}(\pi\{b:\} \uparrow \mu^*[1:], C, \mu^*)$.

In case local-good is violated, variable B remains unmodified and the induction step holds. If local-good is not violated, then $|\pi\{b:\} \uparrow^{\mu^*} [1:]| < m$ and so $\pi \uparrow^{\mu^*} [-m]$, which is the updated value of B at the end of μ^* iteration, precedes b.

Lemma 3. *Suppose the verifier evaluates $\pi_A \geq \pi_B$ in a protocol interaction where B is honest and assume during the comparison that the compared level of the honest party is μ_B. Let $b = LCA(\pi_A, \pi_B)$ and let μ'_B be the adequate level of π_B with respect to b. Then $\mu'_B \geq \mu_B$.*

Proof. Because μ_B is the compared level of the honest party, from the definition of the $\geq_m$ operator, we have $2^{\mu_B}|\pi\{b:\} \uparrow^{\mu_B}| > 2^{\mu'_B}|\pi\{b:\} \uparrow^{\mu'_B}|$. This is true, otherwise the Verifier would have chosen level μ'_B as level of comparison. The proof is by contradiction. Suppose $\mu'_B < \mu_B$. By definition, μ'_B is the maximum level such that $|\pi_B\{b:\} \uparrow^{\mu} [1:]| \geq max(m, (1-\delta)2^{-\mu}|\pi_B\{b:\} \uparrow^{\mu} [1:] \downarrow|)$, therefore μ_B does not satisfy this condition. But we know that $|\pi_B\{b:\} \uparrow^{\mu} [1:]| > m$ because μ_B was selected by the Verifier. Therefore $2^{\mu_B}|\pi\{b:\} \uparrow^{\mu_B}| < (1-\delta)|C\{b:\}|$.
But also μ'_B satisfies goodness, so $2^{\mu'_B}|\pi\{b:\} \uparrow^{\mu'_B}| > (1-\delta)|C\{b:\}|$.
From the last two equations we obtain $2^{\mu_B}|\pi\{b:\} \uparrow^{\mu_B}| < 2^{\mu'_B}|\pi\{b:\} \uparrow^{\mu'_B}|$ which contradicts the initial equation.

@To Be Discussed: would the verifier ever choose a non-adequate level for proof comparison?
Intuitively the above Lemma says: the comparison level chosen by the

Verifier can be no other than the adequate level in respect to block b ($LCA(\pi_A, \pi_B)$), since any other choice would be a level of non-good quality, because of the definition of the adequate level. A level of non-good quality would contain less PoW than that of the adequate level for the range of interest $C\{b:\}$.

Theorem 2. (Security) *Assuming honest majority, the non-interactive proofs-of-proof-of-work construction for computable κ-stable monotonic suffix-sensitive predicates is secure with overwhelming probability in κ.*

Proof. By contradiction. Let Q be a $\kappa-$stable monotonic suffix-sensitive chain predicate. Assume NIPoPoWs on Q is insecure. Then, during an execution at some round r_3, $Q(C)$ is defined and the verifier V disagrees with some honest participant. Assume the execution is typical. V communicates with adversary A and honest prover B. The verifier receives proofs π_A, π_B. Because B is honest, π_B is a proof constructed based on underlying blockchain C_B (with $\pi_B \subseteq C_B$), which B has adopted during round r_3 at which π_B was generated. Furthermore, π_A was generated at round $r'_3 \leq r_3$.

The verifier outputs $\neg Q(C_B)$. Thus it is necessary that $\pi_A \geq \pi_B$. We will show that this is a negligible event.

Let $b = LCA(\pi_A, \pi_B)$. Let $b^\star$ be the most recently honestly generated block in C_B preceding b. Note that $b^\star$ necessarily exists because Genesis is honestly generated. Let the levels of comparison decided by the verifier be μ_A and μ_B respectively. Let μ'_B be the adequate level of proof π_B with respect to block b. Call $\alpha_A = \pi_A \uparrow^{\mu_A} \{b:\}$, $\alpha'_B = \pi_B \uparrow^{\mu'_B} \{b:\}$.

Note that we consider the parts of the proofs succeeding block b the decisive ones for the verifier's choice. This is to adversary's advantage, since the parts preceding this block demonstrate the proof-of-work contained in the common (sub)chain, thus the adversary could only include equal or less proof-of-work in her proof for this part of the chain.

We will now show three successive claims: First, α_A and $\alpha'_B\downarrow$ are mostly disjoint. Second, a_A contains mostly adversarially generated blocks. And third, the adversary is able to produce this a_A with negligible probability.

Let $\alpha_A = k_1 + k_2 + k_3$ and let k_1, k_2, k_3 be as defined in the following Claims.

Claim 1: $\alpha_A, \alpha'_B\downarrow$ are mostly disjoint. We show this by taking the two possible cases for the relation of μ_A, μ'_B.

Claim 1a: If $\mu'_B \leq \mu_A$ then they are completely disjoint. In such a case of inequality, every block in α_A would also be of lower level μ'_B. Applying Lemma 2 to $C\{b:\}\uparrow^{\mu'_B}$ we see that $C\{b:\}\uparrow^{\mu'_B} = \pi\{b:\}\uparrow^{\mu'_B}$. Subsequently, any block in $\pi_A \uparrow^{\mu_A} \{b:\}[1:]$ would also be included in proof α'_B, but $b = LCA(\pi_A, \pi_B)$ so there can be no succeeding block common in α_A, α'_B.

Claim 1b: If $\mu'_B > \mu_A$ then $|\alpha_A[1:] \cap \alpha'_B\downarrow[1:]| = k_1 \leq 2^{\mu'_B - \mu_A}$.

First observe that because the adversary is winning $2^{\mu_A}|\alpha_A| > 2^{\mu'_B}|\alpha'_B| \geq 2^{\mu'_B}m \Rightarrow |\alpha_A| > 2^{\mu'_B-\mu_A}m$. Let's call b_1 the first block in α'_B after block b. Suppose for contradiction that $k_1 > 2^{\mu'_B-\mu_A}$. Since $C_B^{\mu'_B}$ is of good chain quality, this would mean that block b_1, of level μ'_B is also of level μ_A. Since it is of level μ_A the adversary could include it in the proof but b_1 cannot exist in both α_A, α'_B since $\alpha_A \cap \alpha'_B = \emptyset$ by definition. In case that the adversary chooses not to include b_1 in the proof than she can include no other blocks of C_B in her proof, since it would not consist a valid chain.

From Claim 1a and Claim 1b, we conclude that there are $|\alpha_A| - k_1$ blocks after block b in α_A which do not exist in $\alpha_B \downarrow$. We now set $b_2 = LCA(C_B, \alpha_A)$. This makes b_2 the last block before the fork point at the 0-level chain included in the adversary's proof.

Intuition: in this case the common blocks of $\alpha_A, \alpha'_B \downarrow$ may only be blocks of level μ_A which precede the first μ'_B block appearing in α'_B. If this block of level μ'_B was common, it could also be included in α_A. If it is included this would be the LCA of α_A, α'_B. If it is not, then the adversary could no more include blocks from the common part of chain C_B in her proof since they no longer form a valid chain in α_A. The quantity $2^{\mu'_B-\mu_A}$ means: in the range between two consequent μ'_B-level blocks, we have $n = 2^{\mu'_B}$ 0-level blocks and, thus, $2^{-\mu_A}n = 2^{\mu'_B-\mu_A}$ blocks of μ_A-level.

Claim 2: At least k_3 superblocks of α_A are adversarially generated. We show this by showing that $\alpha_A[k_1 + k_2 + 1 :]$ contains no honestly generated blocks. Suppose for contradiction that the block $\alpha_A[i]$ for some $i \geq k_1+k_2+1$ was honestly generated. This means that an honest party adopted the chain $\alpha_A[: i-1] \downarrow$ at some round $r_2 \leq r_3$. Because of the way honest parties adopt chains, the superchain $\alpha_A[: i-1]$ has an underlying properly constructed 0-level anchored chain C_A such that $\alpha_A[: i-1] \subseteq C_A$. Let j be the index of block b_2 within α_A, $j_\downarrow$ be the index of block b_2 within C_A and $k_{2\downarrow} = |\alpha_A[j : j + k_2] \downarrow|$. See Figure 1 for a demonstration. Observe that $|C_A[: \{\alpha_A[i-1]\}]| \geq |C_A[: j_\downarrow + k_{2\downarrow}]|$, while $C_A[j_\downarrow : j_\downarrow + k_{2\downarrow}] \not\preceq C_B$ as proved in Claim 1. But C_A was adopted by an honest party at round r_2, which is prior to round r_3 during which C_B was adopted by an honest party B. This contradicts the Common Prefix[4] with parameter $k_{2\downarrow}$. It follows that with overwhelming probability in $k_{2\downarrow}$, the $k_3 = |\alpha_A| - k_2 - k_1$ last blocks of the adversarial proof have been adversarially generated.

Intuition: Because of Common Prefix on $k_{2\downarrow}$ parameter, where $k_{2\downarrow} = |\alpha_A[j : j + k_2] \downarrow|$, where $E[k_{2\downarrow}] = 2^{\mu_A}k_2$, there can be no honest party adopting C_A at any round $i \geq k_1 + k_2 + 1$.

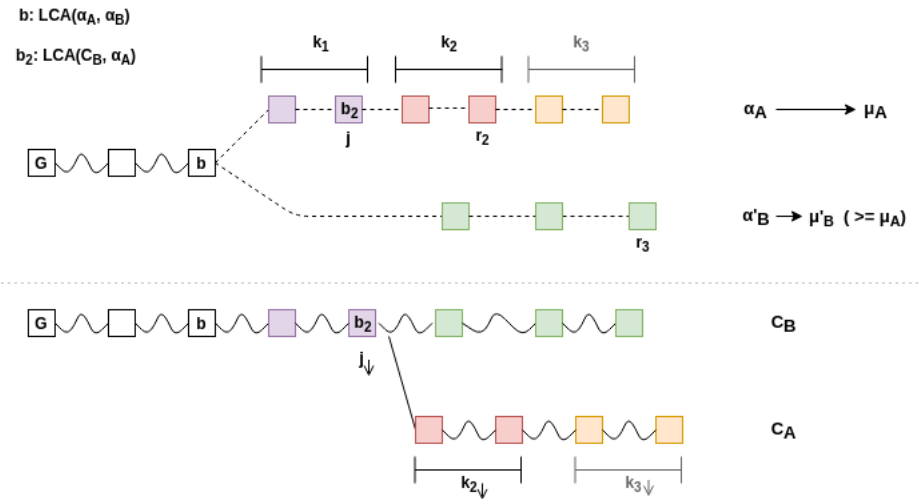

Figure 1: *Two competing proofs at different levels. At the bottom the corresponding 0-level chains are represented.*

From these two Claims we have that k_1 blocks in α_A are blocks of the common zero-level chain, while k_2 blocks are blocks after the fork point at the zero-level chain. Subsequently, k_2 is subject to the Common Prefix $\kappa-$parameter limitations as described in the Backbone paper[4].

Claim 3: Adversary A is able to produce α_A that wins against α_B with negligible probability.

Let b' be the latest honestly generated block in a_A, or $b' = b^*$ if no such block exists in a_A. Let r_1 be the round when b' was generated. Consider the set S of consecutive rounds $r_1..r_3$. Every block in $\alpha_A[-k_3:]$ has been adversarially generated during S and $|\alpha_A[-k_3:]| = |\alpha_A\{b':\}| = k_3$. C_B is a chain adopted by an honest party at round r_3 and filtering the blocks by the rounds during which they were generated to obtain C_B^S, we see that if b'' is the most recently generated block in α_B in a round $r \leq r_1$, then $C_B^S = C_B\{b'':\}$. But $C_B^S \uparrow^{\mu'_B}$ is good with respect to C_B^S. Applying Lemma 1, we obtain that with overwhelming probability $2^{\mu_A}|\alpha_A\{b':\}| < 2^{\mu'_B}|C_B^S \uparrow^{\mu'_B}|$, which is equal to

$$2^{\mu_A}|\alpha_A\{b':\}| < 2^{\mu'_B}|\alpha'_B\{b'':\}| \tag{1}$$

since α'_B contains all the μ'_B-level blocks in C_B^S.

In order to complete the proof, let us now consider $\alpha_A^{k_1}$, $\alpha_A^{k_2}$, $\alpha_A^{k_3}$ the parts of α_A where the k_1, k_2, k_3 blocks reside and $\alpha_B^{k_1}$, $\alpha_B^{k_2}$, $\alpha_B^{k_3}$ the parts of α_B containing blocks generated in the corresponding round sets as illustrated

in Figure 2.

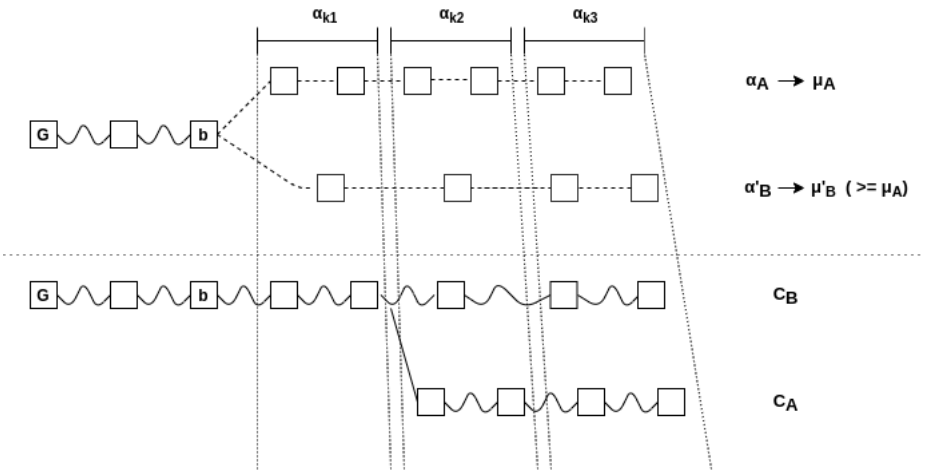

Figure 2: *The three round sets in two competing proofs at different levels. The vertical dashed lines denote the area of interest, across proofs and chains, corresponding to each round set. At the bottom the corresponding 0-level chains are represented.*

Subsequently to the above Claims we have that:
Because of the common underlying chain in the first round set:

$$2^{\mu_A}|\alpha_A^{k_1}| \leq 2^{\mu'_B}|\alpha'^{k_1}_B| \tag{2}$$

Because of the adoption by an honest party of chain C_B at a later round r_3, we have for the second round set:

$$2^{\mu_A}|\alpha_A^{k_2}| \leq 2^{\mu'_B}|\alpha'^{k_2}_B| \tag{3}$$

Because of Equation (1), we have for the third round set:

$$2^{\mu_A}|\alpha_A^{k_3}| < 2^{\mu'_B}|\alpha'^{k_3}_B| \tag{4}$$

So we have

$$2^{\mu_A}(|\alpha_A^{k_1}| + |\alpha_A^{k_2}| + |\alpha_A^{k_3}|) < 2^{\mu'_B}(|\alpha'^{k_1}_B| + |\alpha'^{k_2}_B| + |\alpha'^{k_3}_B|)$$

and finally

$$2^{\mu_A}|\alpha_A| < 2^{\mu'_B}|\alpha'_B| \tag{5}$$

Therefore we have proven that $2^{\mu'_B}|\pi_B \uparrow^{\mu'_B}| > 2^{\mu_A}|\pi_A^{\mu_A}|$. From the definition of μ_B, we know that $2^{\mu_B}|\pi_B \uparrow^{\mu_B}| > 2^{\mu'_B}|\pi_B \uparrow^{\mu'_B}|$ because it was chosen μ_B as level of comparison by the Verifier. So we conclude that $2^{\mu_B}|\pi_B \uparrow^{\mu_B}| > 2^{\mu_A}|\pi_A \uparrow^{\mu_A}|$.

□

It remains to compute the security parameter m that guarantee that all the above hold true in every implementation. It suffices to compute a security parameter value for each set of rounds k_1, k_2, k_3, so that the proof equations 2, 3, 4 hold and then sum these values to obtain parameter m.

In the first set of rounds, for the first k_1 blocks in α_A, we only need 1 block included in α_B for the part of the proof described in Equation 2. In the second set of rounds we need $2^{-\mu_B}\kappa$ blocks for the part of the proof described in Equation 3, just as it directly results from the Common Prefix property. In order to make m independent of any specific level it suffices to consider the upper bound of κ blocks for this set of rounds. In the last set of rounds we need at least κ adversarially generated blocks in $\alpha_A^{k_3}$ so that Lemma 1 is applicable. Since we assume honest majority, obliging to at least κ blocks for this set of rounds suffices to guarantee for Equation 4.

So, we finally conclude to the following upper bound for the value of the security parameter:

$$m = 2\kappa + 1 \tag{6}$$

3.2 Infix Proofs

3.3 Succinctness

4 NIPoPoWs under Velvet Fork

4.1 Suffix Proofs

4.1.1 The Chainsewing Attack

We will now describe an explicit attack against the NIPoPoW suffix proof construction under a velvet fork. Note that since the protocol is implemented under a velvet fork, any adversarial block that is mined in the proper way except containing false interlink data structure will be accepted as valid. A false interlink may contain invalid pointers, for example pointers to superblocks of a fork chain, as shown in Figure 3. Taking advantage of this fact, an adversary maintaining a fork chain could produce suffix proofs that claim blocks of the chain adopted by an honest player as her own. The attack is described in detail in the following.

Assume that chain C_B was adopted by an honest player B and chain C_A, a fork of C_B at some block, maintained by an adversary A. Assume the adversary wants to produce a suffix proof in order to attack an honest light client to have him adopt chain C_A. In order to achieve this, the adversary needs to include a greater amount of PoW in her suffix proof, π_A, in comparison to the honest player's proof, π_B, so as to achieve $\pi_A \geq_m \pi_B$. For this she produces some blocks in chains C_A and C_B containing false interlink

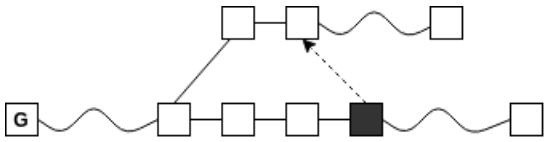

Figure 3: *Example of false interlink structure of an adversarial block, coloured black, in an honest player's chain. The dashed arrow is a pointer to a fork chain superblock included in the interlink.*

pointers which will allow for claiming blocks of chain C_B as of chain C_A in her suffix proof.

The general setting of this attack is represented in Figure 4. The dashed arrows represent interlink pointers of some level μ_A. Starting from the most recently mined block in the adversary's fork chain and following the interlink pointers a chain is formed which consists the adversary's suffix proof. Blocks of both chains are included in this proof and a verifier could not distinguish the false interlink pointers forming this chain proof and, as a result, would consider it a valid proof.

As the generic attack scheme may seem a bit complicated we will now describe a more specific attack case. Consider that the adversary acts as described below. Assume that the adversary chooses to attack at some level μ_A. As shown in Figure 5 she first generates a superblock b' in her fork chain C_A and a superblock a' in the honest chain C_B which are connected via an invalid interlink pointer from a' to b'. As argued earlier, block a' will be accepted as valid in the honest chain C_B despite the false pointers in the interlink data structure. After that the adversary may mine on chain C_A or C_B, or not mine at all. At some point she produces a block a in C_A containing an interlink pointer to a block b of the honest player's chain C_B. Because of the way blocks are generated by updated honest miners there will be successive interlink pointers leading from block b to block a'. Thus following the interlink pointers a chain is formulated which connects C_A blocks a and b' and contains an arbitrarily large part of the honest player's chain C_B.

At this point the adversary will produce a suffix proof for chain C_A containing the subchain $C\{ab\} \cup C\{b : a'\} \cup C\{a' : b'\}$. Notice that following the interlink pointers constructed in such a way, a light client perceives $C\{ab\} \cup C\{b : a'\} \cup C\{a' : b'\}$ as a valid chain.

In this attack the adversary uses false interlink pointers to "sew" portions of the chain adopted by an honest player to her own fork. This remark justifies the name given.

Note that in order to make this attack successful, the adversary has to produce only a few superblocks which let her arrogate an arbitrary large

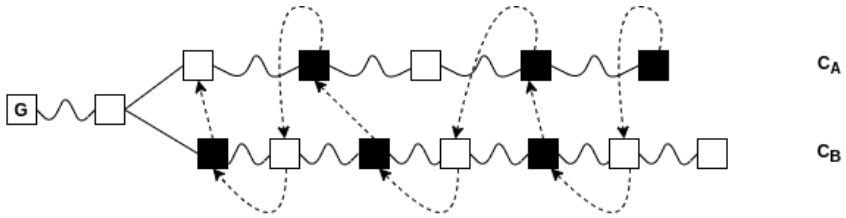

Figure 4: *Generic Chainsewing Attack. C_B is the chain of an honest player and C_A the adversary's chain. Blocks generated by the adversary are colored black. Dashed arrows represent interlink pointers included in the adversary's suffix proof. Wavy lines imply one or more blocks.*

number of blocks of an honest player's chain, while she can mine for her own fork chain. Thus intuitively we expect this attack to succeed with overwhelming probability.

@TODO

Needs to be proven. It is not obvious that the attacker will succeed in high probability, since the most important adversarially generated blocks , a and a', set a limit to the adversarial blocks produced in parallel to the honest blocks of subchain $C\{ab\} \cup C\{b : a'\} \cup C\{a' : b'\}$ and can take part in the suffix proof.

4.2 Protocol Update

In order to eliminate the Chainsewing Attack we propose an update to the NIPoPoWs protocol under velvet fork. The core problem is that in her suffix proof the adversary is able to claim not only blocks of the fork chain, which are in majority adversarially generated due to the Common Prefix property,

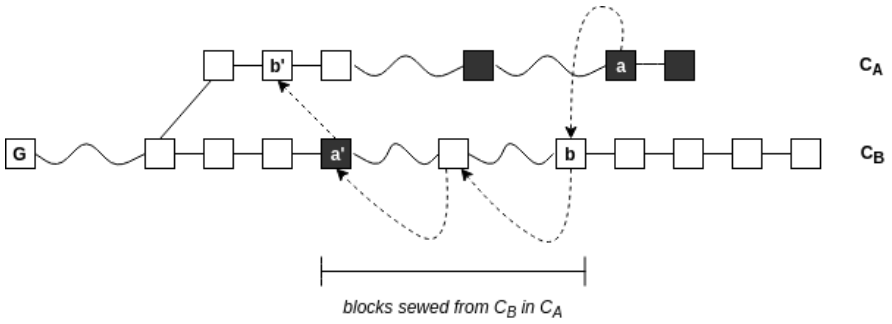

Figure 5: *Chainsewing Attack. The chain at the bottom represents the chain of an honest player, C_B, while the above one is the adversarial fork, C_A. Blocks generated by the adversary are colored black. Dashed arrows represent interlink pointers included in the suffix proof by the adversary. Wavy lines imply one or more blocks. Firm lines imply the previousId relationship between two sequential blocks.*

but also an arbitrarily long part of the chain adopted by an honest player. Since blocks containing false interlink pointers are accepted as valid, the verifier cannot distinguish blocks that actually belong in a chain from blocks that only seem to belong in the same chain because they are pointed to via a false interlink pointer.

These facts make it possible to provide a secure solution for the velvet fork conditions only under the assumption of adversary of 1/3 of the total hashing power. This claim is described and discussed in the following.

Impossibility of a secure protocol for (1/2)-bounded adversary

During our study on the problem we failed to prove the security of several protocol constructions under $(\frac{1}{2})$-bounded adversary. We finally concluded that such a secure NIPoPoWs construction is impossible under velvet fork conditions. Though it is hard to provide a typical proof for this claim, as one should consider any possible construction, in this section we will try to argue for it.

Claim: *Assume t adversarial out of total n parties. There is no construction for NIPoPoWs suffix proofs under velvet fork conditions, which is secure for every adversary t, such that $\frac{t}{n-t} < \delta$.*

Discussion. As explained earlier, since the adversary may use the interlink structure so as to include pointers to arbitrary blocks, she may construct her own chain history utilizing the false pointers included in the blocks she

herself generates. Such an example is given in Figure 4.

Let us consider a construction p which allows for NIPoPoWs suffix proofs under velvet fork and is secure for the underlined conditions. In order p to have a chance for being secure, it should be possible to challenge the submitted proofs, so that an honest player can contest against an inconsistent proof. Note that the same power to challenge any submitted proof is also provided to the adversary.

Now assume an adversary of $\frac{n}{3} < t < \frac{n}{2}$. Then, as explained in the Backbone and Selfish Mining papers [4][3] it is possible for the adversary to maintain the longest valid chain of less than 50% chain quality in favor of adversarially generated blocks. This is illustrated in Figure 6.

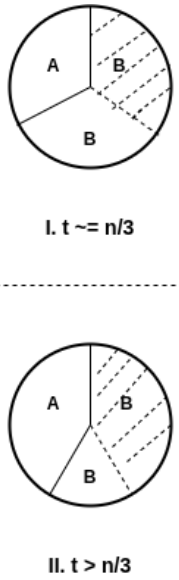

Figure 6: *Pie chart of adversarially and honestly generated blocks distribution during a round set S. Part* ***A*** *stands for blocks mined by the adversary while* ***B*** *for blocks mined by honest parties. Lined out parts denote honestly mined blocks that were defeated by adversarially mined ones in the same round due to selfish mining. The rest denote blocks placed in the chain adopted by honest parties.* ***I.*** *With $t = n/3$, 50% of the total blocks are adversarially generated in the worst case scenario.* ***II.*** *With $t > n/3$, more than half of the total blocks are adversarially generated in the worst case scenario.*

Observe that the property that is violated in the described attack is the *prevId* relation between sequential blocks. However since the main purpose of our work is to achieve succinctness, the *prevId* relations could not be utilized in a viable manner in order to contest adversarial proofs, as it would

require proofs of length linear to that of the whole chain. We thus consider that we should mostly rely on the information given by the interlink data structure as far as the validity of any proof is concerned so as to keep our protocol construction efficient.

Now assume, to honest parties' favor, that it is both possible and efficient to inspect the whole chain history that each block is commited to via its interlink pointers. Keep in mind that an adversary can keep a consistent chain considering the interlink pointers by using only her own mined blocks and, at the same time, these blocks may overwhelm the honestly generated ones. In our protocol p it should be decided under what policy an honest party generates blocks and constructs suffix proofs. A decision should be made for block generation:

1. interlink data are neutral as for adversarially and honestly generated blocks or

2. interlink data point out inconsistent blocks, meaning blocks with incorrect interlinks

3. interlink data exclude inconsistent blocks from being part of the valid chain formed by the super-pointers

Now let's examine the above choices.
In the first case the arguments that an honest player could provide against an inconsistent proof could only be based on the 0-level pointers. Such a contesting proof should at least provide the 0-level subchain of length equal to the following: starting from a block included in the suffix proof but is not a block of the chain until the closest block included in the suffix proof and is a valid block of the chain. This means that contesting proofs are expected to be of length 2^{μ} where $\mu = log(|C|)$, and subsequently are of complexity $\mathcal{O}(|C|)$ which ruins the succinctness of our protocol.

In the second case an honest party could utilize this extra information provided in the interlink to prove an inconsistent proof wrong. However, keep in mind that the adversary could make such claims too. So, a claim of inconsistency for a specific block included should be followed by a proof of its inconsistency. Whatever the information considering the incorrect blocks in the interlink may be, the adversary could make some of it in her own generated blocks in order to contest an honest proof too. So we fall to the previous case turning to the 0-level in order to proof the true chain history, where the contesting proofs are unacceptable efficiency-wise.

In the third case we demand from the honest players to validate the interlink structures of the blocks in the chain and exclude the inconsistent ones from the chain history that the interlinks provide. In this way we could possibly provide contesting proofs using information only from the interlinks thus keeping our proofs succinct. This solution would result to

two different chain histories considering the interlink pointers: the one of the honest parties that includes only blocks with valid interlinks, and the one of the adversary which may form her own history of invalid blocks but could not include honest blocks if at least one invalid block participates. This solution seems to bring us to a notable trade-off point. On the one hand we manage to uncouple from the 0-level blocks. On the other hand, we compromise that honest players cannot use valid adversarially generated blocks in their proof, while the adversary cannot include honestly generated blocks in her proofs if inconsistent blocks are also included. This solution could not work properly for $(\frac{1}{2})$-bounded adversary though. The reason lies in the bad chain quality as pointed out earlier and shown in Figure 6. If the adversary could create his own chain history including inconsistent blocks from two different chains, then she could easily construct a suffix proof that wins over an honest one because the majority of the blocks included in the chain could be adversarial with high probability.

We conclude that such a protocol could not exist for $(\frac{1}{2})$-bounded adversary.

Solution for (1/3)-bounded adversary

The vulnerability that makes this attack possible is the acceptance of blocks containing false interlink pointers. Since we operate under a velvet fork we cannot eliminate such blocks, but we need, however, to restrict the adversary from being able to claim portion of another chain as part of her own fork chain. The key observation on the Chainsewing Attack is that the adversary needs at least one adversarially generated block in an honest player's chain (block a'), in order to create a path of superpointers connecting blocks of two, or more, diverse chains. The final proof chain will make use of this crossing point and may contain both honest or adversarial blocks. In case the proof contains only adversarial blocks, the attack cannot harm security for an adversary of less than 1/3 hashing power. This fact will be clear in the security proof section. In short, as argued in the previous section an adversary of 1/3 of the total hashing power may contribute at most 50% of the total blocks in the longest valid chain. An attacker of more than 1/3 hashing power could dominate as of total hashing power expressed in mined blocks in the chain, while the inverse would be true for an attacker of less hashing power than this threshold.

So in order to be successful, the attacker needs to also "sew" honestly generated blocks. Thus there will be at least one honest block in the superblock path connecting blocks a and a', which points to an adversarial block containing false interlink or, by induction, pointing to a block containing false interlink.

The idea is to ban all blocks generated by honest players from participating in this superblock path. In this way the adversary could not misuse

hashing power of the honest players and the sewed blocks could only be adversarially generated, thus the attack would never succeed for an adversary of less than 1/3 of the total hashing power.

We describe a protocol patch that operates as follows. The NIPoPoW protocol under velvet fork works as usual but each miner constructs a block's interlink excluding the blocks with false interlink (except the pointers of level 0). In this way, blocks containing false interlink pointers are integrated in the chain but are not taken into consideration when updating the interlink structure for the next block to be mined. No honest block could now point to an adversarial superblock that may act as the passing point to the fork chain in an adversarial suffix proof. Thus, after this protocol update the adversary is only able to inject adversarially generated blocks from the chain adopted by an honest party to her own fork. At the same time, adversarially generated blocks cannot participate in an honestly generated suffix proof, as these blocks do not form a valid chain along with honestly mined blocks. Consequently, as far as the hashing power included in a suffix proof is concerned, these blocks can be conceived as belonging in the adversary's fork chain. Figure 7 illustrates this remark.
The protocol patch we suggest can be summarized as follows:

Protocol Patch for NIPoPows under velvet fork. In order to make NIPoPoWs safe under velvet fork conditions we suggest:

1. Strengthen the Honest Majority Assumption so that $\frac{t}{n-t} \leq \frac{1-\delta}{2} \cdot (n-t)$
2. The NIPoPoW protocol under velvet fork works as usual but each miner constructs a block's interlink without considering the blocks with false interlink except for pointers of level $\mu = 0$.

The following two Lemmas come as immediate results from the suggested protocol update.

Lemma 4. *A velvet suffix proof constructed by an honest player cannot contain any block with incorrect interlink.*

Lemma 5. *Consider C_B, C_A the 0-level chains adopted respectively by an honest player and the adversary at some round r'. Let π_A be a velvet suffix proof constructed by the adversary. If there exist blocks with incorrect interlinks in π_A, consider block b_A which was produced at some round r_A and is the first block in π_A that has incorrect interlink. Then $\forall r : r \geq r_A$ no honest blocks generated at round r can be included in π_A.*

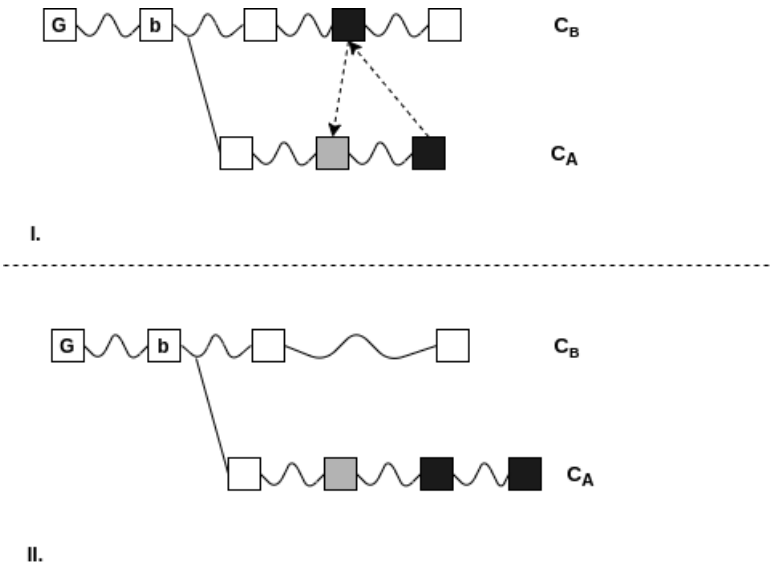

Figure 7: *The adversarial fork chain C_A and chain C_B of an honest player. Blocks generated by the adversary are colored black. Dashed arrows represent interlink pointers. Wavy lines imply one or more blocks. When an adversarially generated block is sewed from C_B into the adversary's suffix proof the verifier conceives C_A as longer and C_B as shorter.* ***I:*** *The real picture of the chains.* ***II:*** *Equivalent picture from the verifier's perspective considering the hashing power included in the corresponding suffix proof of each chain.*

4.2.1 Security of Suffix Proofs

Our security proof is based on the security proof of NIPoPoWs Suffix Security Proof under a soft or hard fork. Consequently, most of the Definitions and Lemmas presented in Section 3.1.1 will be referenced and applied for the security proof under velvet fork conditions. Keep in mind that we operate under the condition of (1/3)-bounded adversary.

The notion of Chain Quality is decisive for the operation of NIPoPoWs under a velvet fork. Before stepping into the security proof we provide the formal definition of this notion and the basic Theorem where its basic metric is computed.

Definition 3 (Chain Quality Property)[4].The chain quality property Q_{cq} with parameters $\mu \in \mathbb{R}$ and $l \in \mathbb{N}$ states that for any honest party P with chain C it holds that for any l consecutive blocks of C the ration of honest blocks is at least μ.

Theorem 3 (Chain Quality)[4]. *In a typical execution the chain quality property holds with parameter* $\mu > 1 - (1 + \frac{\delta}{2}) \cdot \frac{t}{n-t} - \frac{\delta}{2}$.
Proof. See [4].

Theorem 4. (Security under velvet fork) *Assuming honest majority under velvet fork conditions such that* $\frac{t}{n-t} \leq \frac{1-\delta}{2}$, *the non-interactive proofs-of-proof-of-work construction for computable* κ*-stable monotonic suffix-sensitive predicates under velvet fork conditions is secure with overwhelming probability in* κ.

Proof. By contradiction. We follow the proof construction of Theorem 2 and extend it. Let Q be a $\kappa-$stable monotonic suffix-sensitive chain predicate. Assume NIPoPoWs under velvet fork on Q is insecure. Then, during an execution at some round r_3, $Q(C)$ is defined and the verifier V disagrees with some honest participant. Assume the execution is typical. V communicates with adversary A and honest prover B. The verifier receives proofs π_A, π_B. Because B is honest, π_B is a proof constructed based on underlying blockchain C_B (with $\pi_B \subseteq C_B$), which B has adopted during round r_3 at which π_B was generated. Consider C_A the chain containing at least some of the blocks in π_A, while the remaining blocks π_A must belong in C_B. The verifier outputs $\neg Q(C_B)$. Thus it is necessary that $\pi_A \geq \pi_B$. We show that $\pi_A \geq \pi_B$ is a negligible event. Let $b = LCA(\pi_A, \pi_B)$. Let the levels of comparison decided by the verifier be μ_A and μ_B respectively. Let μ'_B be the adequate level of proof π_B with respect to block b. Call $\alpha_A = \pi_A \uparrow^{\mu_A} \{b :\}$, $\alpha'_B = \pi_B \uparrow^{\mu'_B} \{b :\}$.

Our proof construction is based on the following scheme: we show that the competing suffix proofs can be conceived as consisting of three distinct parts. Each part denotes a specific round set and is called after the number of blocks existing in π_A for that round set. Part k_1 lies for the first part of the proofs between blocks $b = LCA(\pi_A, \pi_B)$ and $b_2 = LCA(C_A, C_B)$ meaning for the common 0-level part of α_A, α_B. Part k_2 lies for the second part of the proofs considering the rounds from block $b_2 = LCA(C_A, C_B)$ up until the Common Prefix is established at the 0-level chains for that fork point. The third and last part, k_3 lies for the rest blocks in the proofs.
The above are illustrated, among other, in Parts I, II of Figure 8.

We will now show three successive claims under velvet fork conditions: First, $\alpha_A \downarrow\uparrow^{\mu_A}$ and $\alpha'_B \downarrow$ are mostly disjoint. Second, a_A contains mostly adversarially generated blocks. And third, the adversary is able to produce this a_A with negligible probability.
Let $\alpha_A = k_1 + k_2 + k_3$ and let k_1, k_2, k_3 be as defined in the following Claims. Let round r_1 be the round when block b is generated and round r_2 when block $b_2 = LCA(\alpha_A, \alpha'_B \downarrow)$ is generated.

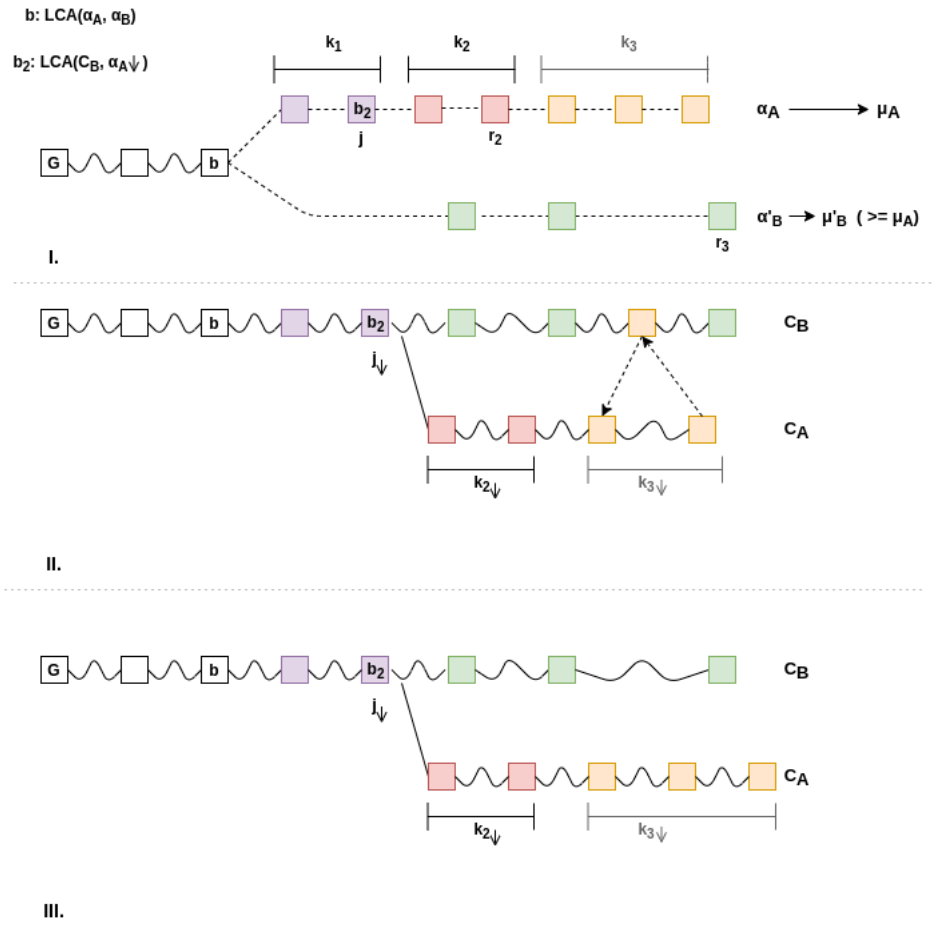

Figure 8: *Wavy lines imply one or more blocks. Dashed lines and arrows imply interlink pointers to superblocks.* ***I****: the three round sets in two competing proofs at different levels,* ***II****: the corresponding 0-level chains,* ***III****: blocks participating in chains C_B, C_A as conceived by the verifier's perspective.*

Claim 1: As for honestly generated blocks, α_A and $\alpha'_B \downarrow$ are mostly disjoint. Following the proof of Theorem 2 we conclude that $|\alpha_A \downarrow\uparrow^{\mu_A} [1 :] \cap \alpha'_B \downarrow [1 :]| \leq k_1 = 2^{\mu'_B - \mu_A}$. In order to see this under the velvet fork conditions consider that the adversary behaves honestly for blocks in her proof between b and b_2, where Claim1 of Theorem 2 applies directly. In the opposite case, the adversary includes a block with false interlink after block b and before block b_2 and because of Lemma 5 no more honestly generated blocks can be included in α_A and we can immediately proceed to Claim 3 of this proof.

So we conclude that there are at least $|\alpha_A| - k_1$ blocks after block b in α_A

which are not honestly generated blocks existing in $\alpha'_B\downarrow$. In other words, there are $|\alpha_A| - k_1$ blocks after block b in α_A, which are either adversarially generated existing in $\alpha_B\downarrow$ either don't belong in $\alpha_B\downarrow$. This makes b_2 the last block before the 0-level fork point included in the adversary's proof.

Claim 2. At least k_3 superblocks of α_A are adversarially generated. Just as the proof of Theorem 2 and using a similar notation, because of the Common Prefix property on parameter $k_{2\downarrow}$, $\alpha_A[k_1 + k_2 :]$ could contain no honestly generated blocks. In order to see this for the velvet fork conditions let's again consider the case that the adversary behaves honestly for the first $(k_1 + k_2)$ blocks of her proof in which case Claim 2 of Theorem 2 is immediately applied. In the opposite case, consider that the adversary includes in her proof a block with invalid interlink at some earlier point. Again, because of Lemma 5 no more honestly generated blocks can be included in α_A and we can proceed to Claim 3 of this proof.

Claim 3. Adversary can submit a suffix proof such that $\alpha_A >= \alpha_B$ with negligible probability. The last k_3 blocks included in α_A may belong either in C_A either in C_B but are all adversarially generated. In the worst case all k_3 blocks are sewed from C_B. This is the worst case scenario since each adversarially generated block in C_B may have dropped one honest block out of the chain because of selfish mining. Considering this scenario, because of the strengthened Honest Majority Assumption for (1/3)-bounded adversary, Theorem 3 for Chain Quality guarantees that the majority of the blocks in C_B was computed by honest parties, thus the honestly generated blocks in C_B for the same round set sum to more amount of hashing power.

From all the above Claims we have that:
In the first round set, because of the common underlying chain:

$$2^{\mu_A}|\alpha_A^{k_1}| \leq 2^{\mu'_B}|\alpha'^{k_1}_B| \tag{7}$$

Because of the adoption by an honest party of chain C_B at a later round r_3, we have for the second round set:

$$2^{\mu_A}|\alpha_A^{k_2}| \leq 2^{\mu'_B}|\alpha'^{k_2}_B| \tag{8}$$

In the third round set, because of good Chain Quality under the strengthened Honest Majority Assumption and Theorem 3 we have:

$$2^{\mu_A}|\alpha_A^{k_3}| < 2^{\mu'_B}|\alpha'^{k_3}_B| \tag{9}$$

Consequently we have:

$$2^{\mu_A}|\alpha_A| < 2^{\mu'_B}|\alpha'_B| \tag{10}$$

4.3 Infix Proofs

References

[1] Kiayias A., Miller A., and Zindros D. Non-interactive proofs of proof-of-work. *IACR Cryptology ePrint Archive*, 2017.

[2] Zamyatin A., Stifter N., Judmayer A., Schindler P., Weippl E., and Knottenbelt W.J. A wild velvet fork appears! inclusive blockchain protocol changes in practice. *Zohar A. et al. (eds) Financial Cryptography and Data Security. FC 2018*, 2019.

[3] Ittay Eyal and Emin Gun Sirer. Majority is not enough: Bitcoin mining is vulnerable, 2013.

[4] Juan Garay, Aggelos Kiayias, and Nikos Leonardos. The bitcoin backbone protocol: Analysis and applications. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 281–310, 2015.

[5] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2009.