A Parallel Repetition Theorem for Any Interactive Argument

Iftach Haitner*

April 2, 2009

Abstract

The question whether or not parallel repetition reduces the soundness error is a fundamental question in the theory of protocols. While parallel repetition reduces (at an exponential rate) the error in interactive proofs and (at a weak exponential rate) in special cases of interactive arguments (e.g., 3-message protocols — Bellare, Impagliazzo and Naor [FOCS '97], and constant-round public-coin protocols — Pass and Venkitasubramaniam [STOC '07]), Bellare et al. gave example of interactive arguments for which parallel repetition does not reduce the soundness error at *all*.

We show that by slightly modifying *any* interactive argument, in a way that preserves its completeness and only slightly deteriorates its soundness, we get a protocol for which parallel repetition does reduce the error at a weak exponential rate. In this modified version, the verifier flips at the beginning of each round an $(1 - \frac{1}{4m}, \frac{1}{4m})$ biased coin (i.e., 1 is tossed with probability $1/4m$), where m is the round complexity of the (original) protocol. If the coin is one, the verifier halts the interaction and accepts, otherwise it sends the same message that the original verifier would. At the end of the protocol (if reached), the verifier accepts if and only if the original verifier would.

1 Introduction

In an interactive proof, a prover P is trying to convince the verifier V in the validity of some statement. Typically, P has some advantage over V, such as additional computational resources or some extra information (e.g., an NP witness that validates the claim). The two basic properties we would like such protocols to have are *completeness* and *soundness*. The completeness means that P convinces V to accept *valid* statements, and the soundness means that no cheating prover (of a certain class) can convince V to accept *invalid* statements. More generally, (P, V) has soundness $1 - \delta$ with respect to a given class of algorithms, if no malicious P^* from this class can convince V to accept an invalid statement with probability greater than δ. The bound δ is typically called the *soundness error* of the protocol.[1]

The basic distinction one may make about the soundness of a given protocol, is whether it holds unconditionally (i.e., even an all-powerful prover cannot break the soundness) or that it only holds against computationally bounded provers. Protocols with unconditional soundness are called *interactive proofs*, whereas protocols with the weaker type of soundness are called *interactive arguments*. In this work, the class of computationally bounded provers is the class of algorithms with polynomial running time.

A common paradigm for constructing protocols with low soundness error (i.e., the probability that the verifier accepts a false statement) is to start with construction a protocol with noticeable soundness error, and then manipulate the original protocol in a certain way that decreases its soundness error while keeping its completeness (i.e., the probability that the verifier accepts a true statement) high. The most natural way that comes to mind is to use repetition. Namely, to repeat the protocol many times (with independent randomness), where the verifier accepts only if the verifiers (of the original protocol) accept in all executions. The above repetition can be done essentially in two different ways: sequentially (known as *sequential*

*Microsoft Research, New England Campus. E-mail: iftach@microsoft.com.

[1] Additional properties (e.g., zero-knowledge) are often also required from the protocol, but in this work we only focus on the above properties.

ISSN 1433-8092

repetition), where the $(i+1)$ execution of the protocol is only started after the i^{th} execution is finished, or in parallel (known as *parallel repetition*), where all the executions are done simultaneously.

Sequential repetition is known to reduce the soundness error at an exponential rate in most computational models (cf., [DP98]). Unfortunately, sequential repetition has the undesired effect of increasing the round complexity. Parallel repetition on the other hand, does preserve the round complexity, and for the case of interactive proofs, it also reduces the soundness error at an exponential rate [Gol99]. Unfortunately, as shown by Bellare, Impagliazzo and Naor [BIN97], in the case of interactive arguments parallel repetition might not reduce the soundness error at *all*.

Let us be more precise about the latter statement. Parallel repetition does reduce the soundness error in the case of 3-message protocol [BIN97, CHS05, IJK06] and in the case of public-coin verifiers [PV07, HPPW08] (see Section 1.3 for more details). On the negative side, for any $k \in \mathbb{N}$ an 8-message protocol with soundness error $\frac{1}{2}$ whose k-parallel repetition soundness remains $\frac{1}{2}$, was presented by Bellare et al. [BIN97]. Recently, Pietrzak and Wikström [PW07] gave an example of a single protocol for which the above phenomena holds for all polynomial k simultaneously.[2] Moreover, both results extend to 4-message protocols, assuming that the soundness proof of the resulting protocol is "black box".

1.1 Our Result

We present a simple method for transforming any efficient interactive argument whose soundness error is bounded away from 1, into an efficient interactive argument with the same number of rounds and negligible soundness error. Given an m-round interactive protocol (P, V), we define the ***random-termination variant of*** V, denoted by $\widetilde{\mathrm{V}}$, as follows: through the interaction with P algorithm $\widetilde{\mathrm{V}}$ acts exactly as V does, but with the following additional step: at the beginning of all but the first round, $\widetilde{\mathrm{V}}$ tosses an $(1-1/4m, 1/4m)$ biased coin (i.e., 1 is tossed with probability $1/4m$). If the outcome of the coin is 1, then $\widetilde{\mathrm{V}}$ accepts the interaction (i.e., outputs 1) and halts. At the end of the interaction (if reached), $\widetilde{\mathrm{V}}$ accepts if and only if V does (upon receiving the same messages from P).

Note that the completeness of $(\mathrm{P}, \widetilde{\mathrm{V}})$ — the random-termination variant of (P, V), is at least as high as the completeness of (P, V), where the soundness error of $(\mathrm{P}, \widetilde{\mathrm{V}})$ is at most $1 - \frac{3}{4}\alpha$, given that the soundness error of (P, V) is at most $1 - \alpha$. Our main contribution is stated in the following theorem.

Theorem 1.1 (informal)**.** *Parallel repetition of the random-termination variant of* any *interactive argument, reduces the soundness error at a weak exponential rate.*[3]

We stress that our result holds with respect to any interactive protocol that can be cast as an interactive argument. For instance, our result yields a round-preserving binding amplification for computationally binding commitment schemes.[4] Our result also extends to the more general threshold case, where the prover in the k-fold repetition is only required to make $t < k$ of the verifiers accept.

1.2 Our Technique

Let (P, V) be an interactive argument with soundness error ε and let $(\mathrm{P}^{(k)}, \mathrm{V}^{(k)})$ be its k^{th} parallel repetition. We show that if (P, V) is a random-termination variant of some protocol, then any efficient algorithm $\mathrm{P}^{(k)^*}$

[2] Both negative results hold under common cryptographic assumptions.

[3] We are using a rather relaxed interpretation of weak exponential rate. Namely, $\varepsilon_k \leq \max\{\mathrm{neg}, \varepsilon^{\frac{k}{\mathrm{poly}(m,(1/1-\varepsilon))}}\}$, where m is the round complexity of the protocol, and ε and ε_k are the soundness error of the original protocol and its k-parallel repetition respectively. See Corollary 3.2 for the exact statement.

[4] Given a weakly binding commitment (S, R), consider the protocol (P, V) where P and V play the role of S and R in a random commit stage of (S, R) respectively. Following the commit stage, P sends two strings to the V, and V outputs "1" iff both strings are valid decommitments to different values. The weakly binding property of (S, R) yields that the soundness error of (P, V) is noticeably far from one. Thus, Theorem 1.1 yields that the parallel repetition of the random-termination variant of (P, V), has negligible soundness error. It follows that the parallel repetition of the random-termination variant of (S, R) is strongly binding.

that breaks the soundness of $(\mathrm{P}^{(k)}, \mathrm{V}^{(k)})$ with "too high" probability ε_k, implies an efficient algorithm P^* that breaks the soundness of (P, V) with probability higher than ε.

Verifiers with "test mode". We start by considering a toy protocol that enjoys an extra property that will enable us to derive P^* from ${\mathrm{P}^{(k)}}^*$, and then show how to emulate this useful property in any random-termination protocol. We say that a verifier V has a ***test mode*** if in at any given time through the execution of the protocol, it agrees (upon request) to move to the following test mode: in this mode the verifier agrees to interact in a random continuation of the protocol, where in such random continuations, the verifier samples uniformly at random the values of the random coins that it *did not flip* prior to moving to the test mode. When the verifier is told that the test mode is ended, it sets its state back to its value before moving to the test mode (and when the "real" interaction continues, it flips the rest of its random coins independently of the values it flipped in the test mode). There is no limit on the number of times the verifier agrees to move to a test mode.

Let us now present an efficient implementation of P^*, assuming that V has a test mode. In order to interact with V, algorithm P^* emulates a random execution of $({\mathrm{P}^{(k)}}^*, \mathrm{V}^{(k)})$, where the "real" V plays the role of the i^{*th} V, for i^* that is chosen at random from $[k]$, and P^* emulates the execution of the other $k-1$ verifiers and of ${\mathrm{P}^{(k)}}^*$. In the j^{th} round (for $j \in [m]$), P^* acts as follows: upon receiving the j^{th} message from V, it samples at random the value of rc_j — the random coins flipped by the emulated verifiers in order to send their j^{th} messages, and evaluates their "quality" — the probability that ${\mathrm{P}^{(k)}}^*$ makes $\mathrm{V}^{(k)}$ accept conditioned on rc_j. In order to do so, P^* samples many random continuations of the protocol, and measures the fraction of accepting ones (i.e., where all the verifiers accept).[5] Sampling such random continuations requires the ability to sample a random continuation of each of the k verifiers. This sampling is always easy for the emulated verifiers (as P^* knows their random coins), and is also possible for a real verifier with a test mode. If this success probability is higher than some threshold (e.g., larger than $\beta_j = \varepsilon_k \cdot (1 - (j/4m))$), P^* let $a^j = (a_1^j, \ldots, a_k^j)$ as the messages that ${\mathrm{P}^{(k)}}^*$ sends in the j^{th} round upon receiving the messages induced by rc_j, and sends $a_{i^*}^j$ back to the real V. In addition, P^* sets the state of the emulated verifiers and ${\mathrm{P}^{(k)}}^*$ according to rc_j. If the above sampling was unsuccessful (the accepting rate was below the threshold), P^* samples new value for rc_j and evaluates its quality as above. Algorithm P^* aborts after n/ε_k unsuccessful attempts to sample good value for rc_j. Note that if P^* does not abort in any of the rounds, then V accepts.

For proving that P^* breaks the soundness of $(\mathrm{P}^*, \mathrm{V})$ with high probability (i.e., higher than ε), we show that the following condition holds with high probability in the end of the emulation's j^{th} round: conditioned on the random coins sampled by P^* and V in the first j rounds of the emulated execution of $({\mathrm{P}^{(k)}}^*, \mathrm{V}^{(k)})$ (where in case of P^*, it stands for the value of $(rc_1, \ldots, rc_j)$), the probability that ${\mathrm{P}^{(k)}}^*$ makes $\mathrm{V}^{(k)}$ accept in a random continuation of $({\mathrm{P}^{(k)}}^*, \mathrm{V}^{(k)})$ is very close to β_j. (In particular, this condition for $j = m$ yields that V has accepted in the real execution of $(\mathrm{P}^*, \mathrm{V})$). The proof goes by induction on j. Assuming that the conditional success portability of ${\mathrm{P}^{(k)}}^*$ in the end of the j^{th} round is β_j, it suffices to show that with high enough probability ${\mathrm{P}^{(k)}}^*$'s conditional success probability is noticeably larger than β_{j+1}, when conditioning also on the actual random coins flipped by V in the beginning of the $j+1$ round. While in the worst case the latter probability might be arbitrarily small, using a result of Raz [Raz98, Claim 5.1] one can show that with high probability over the choice of i^* (where this probability is a function of β_j and k), it holds that the latter probability is high.[6]

Verifiers with no test mode. While in some special cases, such as public-coin and 3-message protocols, implementing the test mode without the verifier's help is easy (indeed, these are exactly the cases where parallel repetition theorems are known), for general protocols implementing such a test mode might be

[5] We assume for the sake of this presentation that V's decision whether to accept or not is only a function of the protocol's transcript, and not of the transcript and its random coins. In the actual proof, we handle the more general case using the "soft decision" approach taken by [BIN97, CHS05, IJK06, HPPW08].

[6] We remark that a conclusion of [Raz98, Claim 5.1] was used in similar settings by [IJK06, HPPW08].

infeasible.[7] Our main technical contribution is an efficient approximation of the test mode for any random-termination protocol.

Let V be a random-termination verifier and assume without loss of generality that it chooses all but its decision bits (the bits uses for deciding whether or not to terminate the executions) before the interaction starts. In order to approximate the test mode in the random-termination verifier (i.e., sample a random continuation conditioned on its already flipped random coins), we sample the random coins conditioned on the event that the verifier's decision bit in the current round is one. Sampling in this case is very easy, since the verifier sends no further messages.

The obvious problem with the above approach is that the additional conditioning might effect the expected success probability of ${\rm P}^{(k)^*}$. If the latter happens, ${\rm P}^*$ might choose a bad value for rc and we have no guarantee for the success probability of ${\rm P}^*$ in this case. Our main technical contribution is showing that the latter does not happen for most values of $i^* \in [k]$. Thus, for most choices of i^* it holds that ${\rm P}^*$ breaks the soundness of $({\rm P}^*, {\rm V})$ with high probability, and therefore it breaks the soundness with high probability for a random $i^* \in [k]$.

1.3 Related Work

Babai and Moran [BM88] showed that parallel repetition reduces the soundness error of Arthur-Merlin protocols, where Goldreich [Gol99] showed that the same holds with respect to interactive proofs. Parallel repetition is also known to reduce the error in the important case of two-prover interactive proofs [Raz98] (in all the above cases the soundness error reduces at exponential rate).

Bellare, Impagliazzo and Naor [BIN97] showed that parallel repetition of 3-message interactive arguments reduces the soundness error at weak exponential rate. For two-message protocols, Canetti et al. [CHS05] gave a proof with better parameters, and Impagliazzo et al. [IJK06] showed that the same holds with respect to the threshold case. Finally for public-coin protocols, Pass and Venkitasubramaniam [PV07] gave a parallel repetition theorem for constant-round protocols, where recently Håstad et al. [HPPW08] extended this result to a polynomial number of rounds. The result of [HPPW08] extends to the case where it is feasible to sample a random continuation of the verifier's actions, given a partial transcript of the protocol (i.e., the verifier has an implicit "test mode", see Section 1.2). All the latter protocols reduce the soundness error at a weak exponential rate. Recently, Haitner et al. [HRVW] showed a round-preserving computational binding amplification of a specific "m-phase" computational binding commitment. The random-termination protocol used in this work, is inspired by their construction. Finally, the phenomena that by changing the verifier to send less information in a single execution (thus increasing the soundness error), we reduce the soundness error when repeating the protocol in parallel, also happens in the work of Feige and Kilian [FK94] in the context of two-prover protocols.

1.4 Paper Organization

We present the notations and formal definitions used in this paper in Section 2, where we also state (and prove) our main technical lemma that bounds the number of variables that significantly effect the expectation of a given function. Our main result is formally stated and proved in Section 3.

2 Preliminaries

For $\alpha, \beta > 0$, we let $(\alpha \pm \beta) := [\alpha - \beta, \alpha + \beta]$. We use calligraphic letters to denote sets, capital letters for random variable, and lower case letters for values. We use superscripts to denote tuples, e.g., $X^n := (X_1, \ldots, X_n)$ and $x^n := (x_1, \ldots x_n)$, and denote by X^n_{-i} the tuple contains all but the i^{th} entry of the tuple,

[7] Since the transcript of the protocol does not always fully determine the random coins used to generate it, there is sufficient information for implementing the test mode. A different approach would be to sample a random continuation of the protocol conditioned on the *transcript*. This approach can be used for proving parallel a repetition theorem for interactive *proofs*, but such a sampling might be hard in the general case (it is equivalent to finding a random preimage of an arbitrary function).

e.g., $X^n_{-i} := (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$. Given a set $\mathcal{X}$, we let $1_{\mathcal{X}}(x) = 1$ if $x \in \mathcal{X}$ and zero otherwise. For a random variable X taking values in a finite set $\mathcal{U}$, we write $x \leftarrow \mathcal{U}$ to indicate that x is selected according to the uniform distribution over $\mathcal{U}$. We adopt the convention that when the same random variable occurs several times in an expression, all occurrences refer to a single sample. For example, $\Pr[f(X) = X]$ is defined to be the probability that when $x \leftarrow X$, we have $f(x) = x$. If a distribution P_{XY} over $\mathcal{X} \times \mathcal{Y}$ is given, we write P_X and P_Y to denote the marginal distributions, e.g., $\mathsf{P}_X(x) := \sum_{y \in \mathcal{Y}} \mathsf{P}_{XY}(x, y)$. The conditional distribution $\mathsf{P}_{Y|X=x}$ is $\mathsf{P}_{Y|X=x}(y) = \mathsf{P}_{XY}(x, y)/\mathrm{P}_X(x)$.

For $k \in \mathbb{N}$ and $p \in (0, 1]$, we let U^p_k be the distribution induced on $\{0, 1\}^k$ by independently setting each of the bits to 1 with probability p. For $i \in [k]$ and $b \in \{0, 1\}$, let the distribution $U^p_{k,i=b}$ be the distribution U^p_k conditioned that i^{th} bit is b. Given a set $\mathcal{S}$ let $1_{\mathcal{S}}(x) = 1$ if $x \in \mathcal{S}$ and 0 otherwise. The statistical difference of two distributions P^1_X and P^2_X defined over a set $\mathcal{X}$, is equal to $\left\|\mathsf{P}^1_X - \mathsf{P}^2_X\right\| = \frac{1}{2} \sum_{x \in \mathcal{X}} \left|\mathsf{P}^1_X(x) - \mathsf{P}^2_X(x)\right| = \max_{\mathcal{X}' \subseteq \mathcal{X}} \left\{\mathsf{P}^1_X(\mathcal{X}') - \mathsf{P}^2_X(\mathcal{X}')\right\}$. When bounding the statistical difference of two distributions, we often use the following proposition.

Proposition 2.1. *Let P^1_X and P^2_X be two distributions over $\mathcal{X}$ and let $\mathcal{X}' \subseteq \mathcal{X}$. Then*

$$\left\|\mathsf{P}^1_X - \mathsf{P}^2_X\right\| \le \frac{1}{2}\mathsf{P}^1_X(\mathcal{X}') + \sum_{x \in \mathcal{X} \setminus \mathcal{X}'} \left|\mathsf{P}^1_X(x) - \mathsf{P}^2_X(x)\right| \ .$$

Proof. Let $\delta := \sum_{x \in \mathcal{X} \setminus \mathcal{X}'} \left|\mathsf{P}^1_X(x) - \mathsf{P}^2_X(x)\right|$, and note that $\mathsf{P}^2_X(\mathcal{X}') \le \mathsf{P}^1_X(\mathcal{X}') + \delta$. Hence

$$\begin{aligned} \left\|\mathsf{P}^1_X - \mathsf{P}^2_X\right\| &= \frac{1}{2} \sum_{x \in \mathcal{X}} \left|\mathsf{P}^1_X(x) - \mathsf{P}^2_X(x)\right| \\ &\le \frac{1}{2} \cdot \max\left\{\mathsf{P}^1_X(\mathcal{X}'), \mathsf{P}^2_X(\mathcal{X}')\right\} + \frac{1}{2}\delta \le \frac{1}{2}\mathsf{P}^1_X(\mathcal{X}') + \delta \ . \end{aligned}$$

□

2.1 Interactive Arguments

An interactive argument for a language $L \subseteq \{0, 1\}^*$, is an interactive protocol between the prover P and the verifier V. The parties get as common input a security parameter 1^n and an element $x \in \{0, 1\}^*$, and the prover might get an additional private input (e.g., witness). We assume for simplicity that V speaks first, where each round of the protocol consists of exchange of two message, from V to P and back. The protocol has completeness $\gamma(n)$, if for every $x \in L$, there exits $w \in \{0, 1\}^*$ such that $\Pr(\mathrm{P}(w), \mathrm{V})(1^n, x) \neq 1] \le \gamma(n)$. The protocol has soundness error $\delta(n)$, if for every $x \notin L$, and for every efficient P^*, it holds that $\Pr(\mathrm{P}^*, \mathrm{V})(1^n, x) = 1] \le \delta(n)$, where the prover is allowed to use an arbitrary private input. In this work we focus on the soundness of the protocol.

2.2 Random-termination Verifiers

Definition 2.2. *[random-termination verifiers] Let V be a verifier of an m-round protocol. The* random-termination variant of V, *denoted as $\widetilde{\mathrm{V}}$, acts exactly as V does, but with the following additional step: in the beginning in all but the first round, $\widetilde{\mathrm{V}}$ tosses an $(1 - 1/4m, 1/4m)$ biased coin (i.e., 1 is tossed with probability $1/4m$), if the outcome of the coin is 1, then $\widetilde{\mathrm{V}}$ accepts (i.e., outputs 1) and halts. Given a protocol (P, V), we call $(\mathrm{P}, \widetilde{\mathrm{V}})$ the random-termination variant of (P, V).*

2.3 Large-Effect Variables

Our soundness analysis for the parallel repetition of random-terminating protocols gives rise to the following question: let f be a real function from $\{0, 1\}^n$ to $[0, 1]$, we are interested in a bound of the number of indices in $[n]$ for which conditioning that the i^{th} bit input of f is 1, significantly changes the expectation of f with respect to a given input distribution.[8] Bounds for the number of variables "large-effect" were given in several other papers (e.g., [HKM06, GRYY08]), but in different settings that than the ones discussed in this paper. For the input distribution U_n^p (i.e., each input bit is independently set to 1 with probability p), we give the following answer.

Lemma 2.3. *Let $f : \{0,1\}^n \mapsto [0,1]$, let $b \in \{0,1\}$ and let p and μ be in $(0, \frac{1}{2}]$. Let $\delta := \mathsf{E}_{U_n^p}[f]$ and let $\mathrm{LargeEfct}_{f,p,\mu,b} := \left\{ i \in [n] \colon \mathsf{E}_{U^p_{n,i=b}}[f] \notin (1 \pm \mu)\delta \right\}$. Then,*

1. *For any function f and $i \in [n]$ it holds that $\mathsf{E}_{U^p_{n,i=0}}[f] \leq \frac{1}{1-p} \cdot \mathsf{E}_{U_n^p}[f]$ and $\mathsf{E}_{U^p_{n,i=1}}[f] \leq \frac{1}{p} \cdot \mathsf{E}_{U_n^p}[f]$.*
2. *For any Boolean function f it holds that $\left|\mathrm{LargeEfct}_{f,p,\mu,b}\right| \in O(\frac{-\log\delta}{p^2\mu^2})$,*
3. *for any function f it holds that $\left|\mathrm{LargeEfct}_{f,p,\mu,b}\right| \in O(\frac{\log\delta\cdot\log(\delta\mu p)}{p^2\mu^2})$.*

Proof. The proof of the first item is immediate by a Markov bound. We prove the other two items for $b = 1$, and the proof for $b = 0$ would immediately follow via a straight forward average argument. In the following we let $\mathrm{LargeEfct}_{f,p,\mu} := \mathrm{LargeEfct}_{f,p,\mu,1}$.

Proof of 2. For $i \in [n]$ and $x \in \{0,1\}^n$, let $r_i(x) := \sqrt{(1-p)/p}$ if $x_i = 1$, and $-\sqrt{p/(1-p)}$ otherwise. Note that for every $i \in [n]$, it holds that $(\mathsf{E}_{U_n^p}[r_i \cdot f])^2 = (1-p)p \cdot (\mathsf{E}_{U^p_{n,i=1}}[f] - \mathsf{E}_{U^p_{n,i=0}}[f])^2 \geq \frac{p}{2}(\mathsf{E}_{U^p_{n,i=1}}[f] - \mathsf{E}_{U^p_{n,i=0}}[f])$. Also note that for every $i \in \mathrm{LargeEfct}_{f,p,\mu}$, it holds that $\left|\mathsf{E}_{U^p_{n,i=1}}[f] - \mathsf{E}_{U^p_{n,i=0}}[f]\right| \geq \mu\delta$. Therefore,

$$\sum_{i \in \mathrm{LargeEfct}_{f,p,\mu}} (\mathsf{E}_{U_n^p}[r_i \cdot f])^2 \geq \left|\mathrm{LargeEfct}_{f,p,\mu}\right| \cdot p\delta^2\mu^2 \tag{1}$$

We conclude the proof using the following proposition, which is a generalization of [Tal96, Proposition 2.2] for the case of $p \neq \frac{1}{2}$.

Proposition 2.4. *For any $p \in (0,1]$ and any subset $A \subseteq \{0,1\}^n$ it holds that*

$$\sum_{i \in [n]} (\mathsf{E}_{U_n^p}[r_i \cdot 1_A])^2 \in O\left(\frac{1}{p} \cdot (\mathsf{E}_{U_n^p}[1_A])^2) \cdot \log\frac{1}{\mathsf{E}_{U_n^p}[1_A]}\right) \quad .$$

Letting $A := f^{-1}(1)$, Proposition 2.4 and Equation 1 yield that

$$\left|\mathrm{LargeEfct}_{f,p,\mu}\right| \in \frac{1}{p^2\delta^2\mu^2} O(\frac{1}{p} \cdot \delta^2 \cdot \log\frac{1}{\delta}) \in O(\frac{1}{p^2\mu^2} \cdot \log\frac{1}{\delta}) \quad .$$

Proof. (of Proposition 2.4) The proof follows the same lines as the original proof, with the following changes. We replace [Tal96, Proposition 2.1] with the following proposition.

Proposition 2.5. *For any real numbers $\{\alpha_i\}_{i\in[n]}$ and $t \geq 0$ it holds that*

$$\mathsf{Pr}\left[\left|\sum_{i\in[n]} \alpha_i \cdot r_i(U_n^p)\right| \geq t\right] \leq 2 \cdot e^{-\frac{p \cdot t^2}{2\sum_{i\in[n]}\alpha_i}}$$

[8] We remark that unless f is monotone, the above notion is not the same as the influence of a variable defined as $\mathsf{Pr}[f(X_1,\ldots,X_i,\ldots,X_n) \neq f(X_1,\ldots,\overline{X_i},\ldots,X_n)]$. In particular, it easy to come up with an example of function and distribution for which the influence of a variable is maximal (i.e., 1), but has small effect (i.e., conditioning that the variable is 1, does changes the expectation of the function).

Proof. The proof is immediate by Hoeffding bound (note that $\mathsf{E}_{U_n^p}[\sum_{i\in[n]}\alpha_i\cdot r_i]=0$, and that for every $i\in[n]$ and $x\in\{0,1\}^n$ it holds that $\alpha_i\cdot r_i(x)\in[-|\alpha_i|\cdot\sqrt{2p},|\alpha_i|\cdot\sqrt{1/p}])$. □

For any real numbers $\{\alpha_i\}_{i\in[n]}$ such that $\sum_{i\in[n]}\alpha_i^2=1$, let $f:=\sum_{i\in[n]}\alpha_i\cdot r_i$. As in [Tal96, Proposition 2.1], it follows that for any $t_0\geq 1$ it holds that $\mathsf{E}_{U_n^p}[f\cdot 1_A]\leq\mathsf{E}_{U_n^p}[1_A]\cdot t_0+2e^{-p\cdot t_0^2/2}$. Taking $t_0=\sqrt{1/p}\cdot\sqrt{2\log\frac{e}{\mathsf{E}_{U_n^p}[1_A]}}\geq 1$, we have that $\mathsf{E}_{U_n^p}[f\cdot 1_A]\in O\left(\sqrt{1/p}\cdot\mathsf{E}_{U_n^p}[1_A]\sqrt{\log\frac{1}{\mathsf{E}_{U_n^p}[1_A]}}\right)$, which concludes the proof. □

Proof of 3. Let $\gamma:=\delta\mu^2p/4$. For each $j\in\mathbb{N}$, let f_j be the function defined by the j^{th} bit of f (i.e., $f_j(x):=f(x)_j$) and let $\delta_j:=\mathsf{E}_{U_n^p}[f_j]$. Let $\mathcal{J}:=\{0\leq j\leq-\log\gamma\colon\delta_j\geq\gamma\}$, let $f_{\mathcal{J}}:=\sum_{j\in\mathcal{J}}2^{-j}\cdot f_j$ and $f_{\overline{\mathcal{J}}}:=\sum_{j\in\mathbb{N}\setminus\mathcal{J}}2^{-j}\cdot f_j$. Let $\mathrm{Bad}:=\bigcup_{j\in\mathcal{J}}\mathrm{LargeEfct}_{f,p,(\mu/2)}$, the linearity of expectation yields that $\mathrm{LargeEfct}_{f_{\mathcal{J}},p,(\mu/2)}\subseteq\mathrm{Bad}$. Hence, for every $i\notin\mathrm{Bad}$ it holds that $(1-\mu/2)\mathsf{E}_{U_n^p}[f_{\mathcal{J}}]\leq\mathsf{E}_{U_{n,i=1}^p}[f_{\mathcal{J}}]\leq\mathsf{E}_{U_{n,i=1}^p}[f]$ and thus

$$(1-\mu)\mathsf{E}_{U_n^p}[f]\leq(1-\mu)(1+\mu^2)\mathsf{E}_{U_n^p}[f_{\mathcal{J}}]\leq(1-\mu)(1+\mu^2)\frac{\mathsf{E}_{U_n^p}[f]}{1-\mu/2}<\mathsf{E}_{U_{n,i=1}^p}[f] \tag{2}$$

Since $\mathsf{E}_{U_n^p}[f_{\overline{\mathcal{J}}}]<2\gamma$, item 1. yields that $\mathsf{E}_{U_n^p}[f_{\overline{\mathcal{J}}}]<2\gamma/p$. Thus for every $i\notin\mathrm{Bad}$ it holds that

$$\begin{aligned}\mathsf{E}_{U_{n,i=1}^p}[f]&=\mathsf{E}_{U_{n,i=1}^p}[f_{\mathcal{J}}]+\mathsf{E}_{U_{n,i=1}^p}[f_{\overline{\mathcal{J}}}]\\&\leq\ \mathsf{E}_{U_{n,i=1}^p}[f_{\mathcal{J}}]+2\gamma/p\\&\leq\ (1+\mu/2)\mathsf{E}_{U_n^p}[f_{\mathcal{J}}]+\delta\mu^2/2\leq(1+\mu)\mathsf{E}_{U_n^p}[f]\ .\end{aligned}$$

Applying 2., we conclude that $|\mathrm{LargeEfct}|\leq|\mathcal{J}|\cdot O(\frac{-\log\delta}{p^2\mu^2})\in O(\frac{\log\delta\cdot\log(\delta\mu p)}{p^2\mu^2})$.

□

We generalize the above Lemma for non-Boolean variable as follows.

Corollary 2.6. *Let $f:(\{0,1\}^t)^n\mapsto[0,1]$, $b\in\{0,1\}$ and let p and μ be in $(0,\frac{1}{2}]$. Let $\delta:=\mathsf{E}_{(U_t)^n}[f]$ and let $\mathrm{LargeEfct}^t_{f,p,\mu}:=\left\{i\in[n]\colon\mathsf{Pr}_{a\leftarrow U_t}[\mathsf{E}_{(U_t)^n|(U_t)^n_i=a}[f]\notin(1\pm\mu)\cdot\delta]>p\right\}$. Then,*

1. *For any Boolean function f it holds that $\left|\mathrm{LargeEfct}^t_{f,p,\mu,b}\right|\in O(\frac{-\log\delta}{p^2\mu^2})$,*

2. *for any function f it holds that $\left|\mathrm{LargeEfct}^t_{f,p,\mu,b}\right|\in O(\frac{\log\delta\cdot\log(\delta\mu p)}{p^2\mu^2})$.*

Proof. Again we prove for $\mathrm{LargeEfct}^t_{f,p,\mu}:=\mathrm{LargeEfct}^t_{f,p,\mu,1}$. We assume without lost of generality that $p\cdot 2^n\in\mathbb{N}$ (otherwise we would consider the smallest $p'>p$ for which the latter holds), and assume for simplicity that for each $i\in\mathrm{LargeEfct}^t_{f,p,\mu}$ it holds that $\mathsf{Pr}_{a\leftarrow U_t}[\mathsf{E}_{(U_t)^n|(U_t)^n_i=a}[f]<(1-\mu)\cdot\delta]>p$. For each $i\in[n]$, let L_i be the set of the lightest $p\cdot 2^n$ elements inside $\{0,1\}^n$ according to the weight function $w(a):=\mathsf{E}_{(U_t)^n|(U_t)^n_i=a}[f]$. We define $h\colon\{0,1\}^n\mapsto[0,1]$ as $h(x):=\mathsf{E}[f(X^1_{x_1},\ldots,X^n_{x_n})]$, where X^i_1 is uniformly distributed over L_i and X^i_0 is uniformly distributed over $\{0,1\}^n\setminus L_i$. Note that $\mathsf{E}_{U_n^p}[h]=\delta$, and that for every $i\in\mathrm{LargeEfct}^t_{f,p,\mu}$ it holds that $\mathsf{E}_{U_{n,i=1}^p}[h]<(1-\mu)\cdot\delta$. Thus, we are in the settings of Lemma 2.3, and the proof of the corollary follows. □

3 Parallel Repetition Theorem for Random-termination Protocols

In this section we present the main contribution if this paper. The following theorem relates the soundness of the k^{th} parallel repetition of the random-termination variant of a protocol to the soundness of the original (non random-termination) verifier. In Corollary 3.2, we use this result to show that the soundness error of the k parallel repetition decays at a weak exponential rate.

Theorem 3.1. *[restatement of Theorem 1.1] Let* (P,V) *be an* m*-round protocol and let* $\widetilde{\mathrm{V}}$ *the random-termination variant of* V. *There exists an oracle-aided algorithm* P^* *such that the following holds: let* $x \in \{0,1\}^*$, $n \geq 2\cdot\log m$, $n^8m^{12} \leq k \in \mathrm{poly}(n)$ *and* $t \in [k]$. *Then for any algorithm* ${\mathrm{P}^{(k)}}^*$ *for which* $\varepsilon_k(n) := \mathsf{Pr}[$*at least* t *verifiers accept in*$({\mathrm{P}^{(k)}}^*, \widetilde{\mathrm{V}}^{(k)})(1^n,x)] > 2^{-n/4}$, *it holds that*

$$\mathsf{Pr}[(\mathrm{P}^{*{\mathrm{P}^{(k)}}^*}(t),\mathrm{V})(1^n,x)=1] > \frac{t}{k} - O(m\cdot k^{-\frac{1}{10}}) \ .$$

The running time of P^* *is bounded by* $O(k\cdot T_{{\mathrm{P}^{(k)}}^*}/\varepsilon_k^3)$, *where* $T_{{\mathrm{P}^{(k)}}^*}$ *is an upper bound on the running time of* $({\mathrm{P}^{(k)}}^*, \widetilde{\mathrm{V}}^{(k)})(1^n,x)$.

Theorem 1.1 yields that following useful corollary.

Corollary 3.2. *Let* (P,V), n, m, k, t, ${\mathrm{P}^{(k)}}^*$ *and* ε_k *be as in Theorem 3.1. Let* $\varepsilon > 0$ *be a bound on the soundness error of* (P,V) *(against* PPT *'s). Assuming that* ${\mathrm{P}^{(k)}}^*$ *is a* PPT *and that* $\delta := \frac{t}{k} - \varepsilon$ *is noticeable (e.g.,* $\exists p \in \mathrm{poly}$ *such that* $\delta > \frac{1}{p(n)}$*), then* $\varepsilon_k \leq \max\{\mathrm{neg}(n), \varepsilon^{(\frac{\delta}{m})^{10}\cdot k}\}$, *where* neg *stands for any negligible function.*

Proof. Assumes that $\varepsilon_k > \varepsilon^{(\frac{\delta}{m})^{10}\cdot k} > \mathrm{neg}(n)$. Let $k' := C\cdot\frac{m}{\delta})^{10}$, for $C > 0$, be a multiple of k. Consider the adversary ${\mathrm{P}^{(k')}}^*$ for interacting with $\mathrm{V}^{(k')}$ that is the k'/k "parallel repetition" of ${\mathrm{P}^{(k)}}^*$. Namely, ${\mathrm{P}^{(k')}}^*$ partitions the verifiers into groups of size k and acts as ${\mathrm{P}^{(k)}}^*$ "against" each of this groups. A direct calculation shows that $\varepsilon_{k'} = \mathsf{Pr}[\text{at least } \frac{t}{k} \text{ fraction of the verifiers accept in}({\mathrm{P}^{(k')}}^*, \mathrm{V}^{(k')})(1^n,x)] \in \Omega(\varepsilon)$. Thus, Theorem 1.1 yields that (for the proper choice of C) there exists an efficient algorithm P^* that runs it time $\mathrm{poly}(k', 1/\varepsilon, T_{{\mathrm{P}^{(k')}}^*})$ and breaks the soundness of (P,V) with probability better than ε. The fact that $\varepsilon_k > \mathrm{neg}$, implies via a standard hybrid argument that $\varepsilon > \mathrm{neg}$. Hence, P^* contradicts the soundness guarantee of (P,V). □

Proof. (of Theorem 1.1) We say that $\widetilde{\mathrm{V}}^{(k)}$ accepts if at least t of the $\widetilde{\mathrm{V}}$'s do. We omit 1^n and x from our notations whenever their values is clear from the context. We assume without lost of generality that ${\mathrm{P}^{(k)}}^*$ is deterministic, as handling randomized ${\mathrm{P}^{(k)}}^*$ would only increase the running time of our adversary by $O(n/\varepsilon_k^2)$, while reducing its success probability by $O(2^{-n})$.

Let $\mathrm{len} \in \mathbb{N}$ be a bound on the number of random coins used by V in any interaction (with respect to security parameter 1^n). We assume without lost generality that the partial view of $\widetilde{\mathrm{V}}^{(k)}$ in an interaction with ${\mathrm{P}^{(k)}}^*$ is of the form $\mathrm{view}(r^k, \mathcal{S}_1,\ldots,\mathcal{S}_\ell)$, where $r^k \in \{0,1\}^{k\cdot\mathrm{len}}$ denotes the random coins of the k embedded V's inside $\widetilde{\mathrm{V}}^{(k)}$, and $\mathcal{S}_j$ (for $j \in [\ell]$) denotes the indices of those verifier that decided to halt on the beginning of the j^{th} round. Since ${\mathrm{P}^{(k)}}^*$ is deterministic, we omit its messages from $\widetilde{\mathrm{V}}^{(k)}$'s view. We let $\mathcal{S}_j(\mathrm{view})$ be the value of the entry '$\mathcal{S}_j$' in view, let $\mathcal{S}_{>j}(\mathrm{view}) := [k] \setminus (\bigcup_{j'=1}^{j}\mathcal{S}_{j'}(\mathrm{view}))$, let $\mathit{round}(\mathrm{view} = (r^k, \mathcal{S}_1,\ldots,\mathcal{S}_\ell)) := \ell+1$ and let $\mathit{round}(\perp) := 1$.

We start by considering an algorithm $\widehat{\mathrm{P}}$ that given V's random coins as input, makes V accept with high probability. We then complete the proof by showing how to implement $\widehat{\mathrm{P}}$ efficiently, without no access to these random coins. We stress the heart of our proof lies in the implementation of $\widehat{\mathrm{P}}$ (given below), where the shift to an algorithm without access to V's random coins is rather standard. Algorithm $\widehat{\mathrm{P}}$ follows rather closely the intuition given in Section 1.2, but with the following main differences:

1. In all but for choosing the last random coins of the emulated verifiers, $\widehat{\mathrm{P}}$ uses a soft threshold for evaluating the "quality" of these random coins. Namely, the probability that $\widehat{\mathrm{P}}$ return a given value for these random coins, decays relatively to the distance of the induced success probability of ${\mathrm{P}^{(k)}}^*$ below a certain threshold (and not set to zero as described in the introduction). The use of such soft threshold is needed, as our estimation of the above quality is not accurate enough.

2. When choosing the last random coins, $\widehat{\mathrm{P}}$ uses a soft threshold for deciding whether the number of accepting verifiers in a given execution of $({\mathrm{P}^{(k)}}^*, \widetilde{\mathrm{V}}^{(k)})$ is "high enough". This change will later allows us to handle the case where V's random coins are unknown. We note that similar approach was taken by [BIN97, CHS05, IJK06, HPPW08].

We define Algorithm $\widehat{\mathrm{P}}$ is as follows.

Algorithm 3.3. $\widehat{\mathrm{P}}$.

Oracle: ${\mathrm{P}^{(k)}}^*$

Input: *A string* $r \in \{0,1\}^{\mathrm{len}}$.

Operations:

```
1. Choose $i^* \in [k]$ uniformly at random and set $\mathrm{view} = \perp$.
2. For $j = 1$ to $m$:
    (a) Set $\mathrm{view} = (\mathrm{view}, \mathrm{GetNextRC}^{{\mathrm{P}^{(k)}}^*}(\mathrm{view}, i^*, r))$.
    (b) Send $a^j_{i^*}$ to $\mathrm{V}$, where $a^j$ is the message that ${\mathrm{P}^{(k)}}^*$ sends to $\widetilde{\mathrm{V}}$ in the $j^{th}$ round of view.
```

Algorithm 3.4. GetNextRC.

Oracle: ${\mathrm{P}^{(k)}}^*$

Input: $\widetilde{\mathrm{V}}^{(k)}$*'s view* — view, *an index* $i^* \in [k] \cup \perp$ *and a string* $r \in \{0,1\}^{\mathrm{len}} \cup \perp$.

Operations:

```
1. Set $round = round(\mathrm{view})$, $\mu = k^{-\frac{1}{10}}$ and $p = 1/4m$.
2. Do the following for $16mn/\varepsilon_k$ times:
        //Sample next random coins
    (a) If $round = 1$,
        i. Choose $r^k$ uniformly at random from $\{0,1\}^{k \cdot \mathrm{len}}$ conditioned that $r^k_{i^*} = r$.
        ii. Set $rc = r^k$.
    (b) Otherwise,
        i. Chooses $\mathcal{S}_{round} \subseteq \mathcal{S}_{>round-1}(\mathrm{view})$ conditioned that $i^* \notin \mathcal{S}_{round}$, where each $i \in \mathcal{S}_{>round-1}$ is (independently) chosen to be in $\mathcal{S}_{round}$ with probability $p$.
        ii. Set $rc = \mathcal{S}_{round}$.
        //Evaluate the random coins
```

(c) *If round* $< m$*,*

  i. *Sample independently* $8 \cdot n/\varepsilon_k^2\mu^2$ *different values for* $\text{view}' = (\text{view}, rc, \mathcal{S}_{round+1}, \ldots, \mathcal{S}_m)$ *conditioned that* $i^* \in \mathcal{S}_{round+1}(\text{view}')$*, where each* $i \in \mathcal{S}_{>\ell}(\text{view}')$ *(for* $\ell \in \{round, \ldots, m-1\}$*) is chosen to be in* $\mathcal{S}_{\ell+1}$ *with probability* p*.*

  ii. *Let* α *be the fraction of accepting* view$'$*'s and let* $\beta = \varepsilon_k(1 - \frac{(2+3\cdot round)}{8m})$*. If* $\alpha \geq \beta$*, return rc with probability* α*, otherwise return rc with probability* $\beta \cdot (\frac{\alpha}{\beta})^{8nm}$*.*

(d) *Otherwise, return rc with probability* $\min\left\{1, 2^{\mu\cdot(|\mathcal{T}|-t)}\right\}$*, where* $\mathcal{T} = \left\{i \in [k] \colon \widetilde{\mathrm{V}}_i \text{ accepts in } (\text{view}, rc)\right\}$*.*

3. *Abort the execution.*

We assume that once the execution of $(\widehat{\mathrm{P}}(r), \mathrm{V}(r))$ ends, $\widehat{\mathrm{P}}$ outputs (view, i^*), and let $\mathsf{P}^0_{\text{View},I^*}$ be the output distribution of $\widehat{\mathrm{P}}$ induced by an execution of $(\widehat{\mathrm{P}}(U_{\text{len}}), \mathrm{V}(U_{\text{len}}))$. We also assume that if GetNextRC is called with $i^* = \perp$, it does the sampling of Lines 2.(a).i, 2.(b).i and 2.(c).i *without* the conditioning on $\imath^*$. We are interested in the probability over $\mathsf{P}^0_{\text{View},I^*}$ that $\widetilde{\mathrm{V}}_{i^*}$ accepts in View. For lower bounding this probability we introduce the following family of experiments $\left\{\text{Exp}^\ell\right\}_{\ell\in[m]}$.

Experiment 3.5. Exp^ℓ.

Definition:

1. *Set* view $=\perp$.

2. *For* $j = 1$ *to* ℓ *do*

   set view $= (\text{view}, \text{GetNextRC}^{\mathrm{P}^{(k)^*}}(\text{view}, \perp, \perp))$.

3. *Select uniformly at random* $i^* \in \mathcal{S}_{>\ell}(\text{view})$.

4. *For* $j = \ell + 1$ *to* m *do*

   set view $= (\text{view}, \text{GetNextRC}^{\mathrm{P}^{(k)^*}}(\text{view}, i^*, \perp))$.

5. *Output* (view, i^*).

Let $\mathsf{P}^\ell_{\text{View},I^*}$ be the output distribution of Exp^ℓ. The proof of the theorem follows by the next two claims.

Claim 3.6. *It holds that* $\left\|\mathsf{P}^0_{\text{View},I^*} - \mathsf{P}^m_{\text{View},I^*}\right\| \in O(m \cdot \mu)$*.*

Claim 3.7. $\mathsf{P}^m_{\text{View},I^*}(\widetilde{\mathrm{V}}_{I^*}$ *accepts in* View$) \geq \frac{t}{k} - O(\mu)$*.*

Before proving the above claims, let us use the above claims for proving Theorem 3.1. Claim 3.6 and Claim 3.7 yield that $\widehat{\mathrm{P}}$ makes V accept with probability $\frac{t}{k} - O(m\mu)$, so it is left to show how to implement $\widehat{\mathrm{P}}$ without knowing the verifier random coins. Since $\widehat{\mathrm{P}}$ calls GetNextRC(view, i^*, r) only after receiving the first *round*(view)$+1$ messages from V, and this knowledge suffices for the computation of GetNextRC(view, i^*, r), we only need to know r for identifying the set $\mathcal{T}$ in the m^{th} call to GetNextRC.

Let GetNextRC$'$ be a variant of GetNextRC that in case *round* $= m$, sets $\mathcal{T} = \{i \in [k] \backslash \{i^*\} : \widetilde{\mathrm{V}}_i \text{ accepts}\}$ (i.e., the decision of the i^{*th} verifier is ignored). Let P^* be that variant of $\widehat{\mathrm{P}}$ that uses GetNextRC$'$ instead of GetNextRC and let $\mathsf{P}^{\text{Real}}_{\text{View},I^*}$ be the distribution induced by a random execution of $(\mathrm{P}^*, \mathrm{V}(U_{\text{len}}))$ on (i^*, view). For any fixed values of view $= (r^k, \ldots, \mathcal{S}_{m-1})$ and i^*, the“soft threshold” decision used by GetNextRC yields that probability that a single loop of GetNextRC$(\text{view}, i^*, \cdot)$ returns a value rc, is at most $1/(1-\mu)$

larger (and never smaller) than the probability of a single loop of $\mathrm{GetNextRC}'(\mathrm{view}, i^*, \cdot)$ to return rc. It follows that $\left\|\mathsf{P}^0_{\mathrm{View},I^*} - \mathsf{P}^{\mathrm{Real}}_{\mathrm{View},I^*}\right\| \in O(\mu)$ and hence $\left\|\mathsf{P}^{\mathrm{Real}}_{\mathrm{View},I^*} - \mathsf{P}^m_{\mathrm{View},I^*}\right\| \in O(m\mu)$, and we conclude that $\Pr[(\hat{\mathrm{P}}, \mathrm{V}(U_{\mathrm{len}})) = 1)] > \frac{t}{k} - O(m \cdot \mu)$. □

We start with giving some notations and make several observations about algorithm GetNextRC.

Definition 3.8. *Let* view *be a partial view of* $\widetilde{\mathrm{V}}^{(k)}$ *in an execution of* $({\mathrm{P}^{(k)}}^*, \widetilde{\mathrm{V}}^{(k)})$*, let* $i^* \in [k] \cup \{\perp\}$ *and let* $r \in \{0,1\}^{\mathrm{len}} \cup \{\perp\}$*. We let* $\delta_{\mathrm{view},i^*,r}(rc)$ *be the probability that* $\mathrm{GetNextRC}(\mathrm{view}, i^*, r)$ *returns* rc*, and let* $\delta_{\mathrm{view},i^*,r} := \mathsf{E}_{Rc}[\delta_{\mathrm{view},i^*,r}(\mathsf{R})]$*, where* Rc *is distributed according to the value a single loop of* $\mathrm{GetNextRC}(\mathrm{view}, i^*, r)$ *induces on* rc*. Given that* $round(\mathrm{view}) < m$*, let* $\gamma_{\mathrm{view},i^*,r}(rc)$ *be the probability that* $\widetilde{\mathrm{V}}^{(k)}$ *accepts (i.e., number of accepting verifier is at least* t*) in a single* view' *sampled by* $\mathrm{GetNextRC}(\mathrm{view}, i^*, r)$ *conditioned on* rc*, and let* $\gamma_{\mathrm{view},i^*,r} := \mathsf{E}_{Rc}[\gamma_{\mathrm{view},i^*,r}(Rc)]$*. Finally, for* $j \in [m]$ *let* $\beta_j := \varepsilon_k(1 - \frac{(3 \cdot j + 2)}{8m})$*.*

Note that $\gamma_{(\mathrm{view},rc),\perp,\perp} = \gamma_{\mathrm{view},\perp,\perp}(rc)$ for $round(\mathrm{view}) < m-1$ and that $\delta_{(\mathrm{view},rc),\perp,\perp} = \gamma_{\mathrm{view},\perp,\perp}(rc)$ for $round(\mathrm{view}) = m-1$. Also note that $\delta_{\perp,\perp,\perp}$ is exactly the cheating probability of ${\mathrm{P}^{(k)}}^*$ and thus equals ε_k. We will use the following claim.

Claim 3.9. *It holds that*

1. $\delta_{\mathrm{view},i^*,r}(rc) \leq (1 + \frac{\mu}{8mn}) \cdot \gamma_{\mathrm{view},i^*,r}(rc)$ *for every value of* $(\mathrm{view}, i^*, r, rc)$
2. $\delta_{\mathrm{view},i^*,r}(rc)(1 \pm (\frac{\mu}{8mn})) \cdot \gamma_{\mathrm{view},i^*,r}(rc)$ *for* $\gamma_{\mathrm{view},i^*,r}(rc) \geq \beta_j$
3. $\delta_{\mathrm{view},i^*,r}(rc) \in O(2^{-n})$ *for* $\gamma_{\mathrm{view},i^*,r}(rc) < \beta_j - \frac{1}{m}$
4. $\delta_{\mathrm{view},i^*,r}(rc)(1 \pm O(\mu)) \cdot \delta_{\mathrm{view}',i^{*\prime},r'}(rc')$ *for* $\gamma_{\mathrm{view},i^*,r}(rc) \cdot (1 \pm \frac{\mu}{8mn}) \cdot \gamma_{\mathrm{view}',i^{*\prime},r'}(rc')$*,*

Proof. An Hoeffding bound yields that for $\gamma_{\mathrm{view},i^*,r}(rc) \geq \varepsilon_k/2$ it holds that $\Pr[\alpha \notin (1 \pm \frac{\mu}{8mn}) \cdot \gamma_{\mathrm{view},i^*,r}(rc)] \in O(2^{-n})$, where α is the value calculated in Line $2.(c).ii$ of GetNextRC. Thus, the first three line of the claim are immediate by the definition of GetNextRC. The last line is immediate for $\gamma_{\mathrm{view},i^*,r}(rc), \gamma_{\mathrm{view}',i^{*\prime},r'}(rc') \notin [\beta_{round(\mathrm{view})} - \frac{1}{m}, \beta_{round(\mathrm{view})}]$. It holds for $\gamma_{\mathrm{view},i^*,r}(rc), \gamma_{\mathrm{view}',i^{*\prime},r'}(rc') \in [\beta_{round(\mathrm{view})} - \frac{1}{m}, \beta_{round(\mathrm{view})}]$, since in this case $\frac{\delta_{\mathrm{view},i^*,r}(rc)}{\delta_{\mathrm{view}',i^{*\prime},r'}(rc')} = \left(\frac{\gamma_{\mathrm{view},i^*,r}(rc)}{\gamma_{\mathrm{view}',i^{*\prime},r'}(rc')}\right)^{8nm}$. Finally, the triangle inequality yields that it also holds for $\gamma_{\mathrm{view}',i^{*\prime},r'}(rc') \in [\beta_{round(\mathrm{view})} - \frac{1}{m}, \beta_{round(\mathrm{view})}]$ and $\gamma_{\mathrm{view},i^*,r}(rc) \notin [\beta_{round(\mathrm{view})} - \frac{1}{m}, \beta_{round(\mathrm{view})}]$. □

Fix view, i^* and r such that $round(\mathrm{view}) < m$, and let Rc be distributed according to the value a single loop of $\mathrm{GetNextRC}(\mathrm{view}, i^*, r)$ induces on rc. An averaging argument yields that $\mathsf{E}_{\gamma_{\mathrm{view},i^*,r}(Rc)}[1_{[\gamma_{\mathrm{view},i^*,r} - \frac{1}{2m}, 1]}] \geq \gamma_{\mathrm{view},i^*,r}/2m$, and thus

$$\delta_{\mathrm{view},i^*,r} \geq \gamma_{\mathrm{view},i^*,r}/2m \ , \tag{3}$$

for $\gamma_{\mathrm{view},i^*,r} > \beta_{round(\mathrm{view})} + \frac{1}{2m}$. It follows that

$$\mathsf{Pr}\left[\mathrm{GetNextRC}(\mathrm{view}, i^*, r) \text{ aborts} \mid \gamma_{\mathrm{view},i^*,r} > \beta_{round(\mathrm{view})} + \frac{1}{2m})\right] \in O(2^{-n}) \tag{4}$$

Proof. (of Claim 3.7) For $j \in [m-1]$ let Γ_j denote the value of $\gamma_{\mathrm{view},\perp,\perp}$ in a random instance of Exp^m, and let Γ_m be the value of $\delta_{\mathrm{view},\perp,\perp}$ in such execution. Since $\Gamma_1 = \varepsilon_k$, Claim 3.9(3) and Equation 4 yield with save but negligible probability, it holds that $\Gamma_j \geq \beta_{j-1} - \frac{1}{m} = \beta_j + \frac{1}{m}$. In particular, Exp^m aborts only with negligible probability. Assuming that $\Gamma_m > \varepsilon_k/2$, the probability that GetNextRC returns rc for which the number of accepting verifiers in (view, rc) is less than $t - \frac{-\log(\varepsilon_k \cdot \mu)}{\mu}$, is bounded by $O(\mu)$. It follows that $\mathsf{P}^m_{\mathrm{View},I^*}(\widetilde{\mathrm{V}}_{I^*} \text{ is accepting in View}) > \frac{t - \frac{-\log(\varepsilon_k \cdot \mu)}{\mu}}{k} - O(\mu) \geq \frac{t}{k} - O(\mu)$. □

Proof. (of Claim 3.6) We prove the claim by proving that $\left\|\mathsf{P}^{\ell}_{\text{View},I^*} - \mathsf{P}^{\ell+1}_{\text{View},I^*}\right\| \in O(\mu)$ for every $\ell \in \{1,\ldots,m-1\}$. For a given (partial) view of $\widetilde{\mathrm{V}}^{(k)}$ view $= (r^k, \mathcal{S}_1, \ldots, \mathcal{S}_\ell)$, we let $k_j(\text{view}) = |\mathcal{S}_{>j}(\text{view})|$ and identify the indices in $\mathcal{S}_{>j}(\text{view})$ with the set $[k_j(\text{view})]$. We call view ***typical*** if for every $j \in [\ell-1]$ is holds that $k_{j+1}(\text{view}) \in (1 \pm \mu) k_j(\text{view}) \cdot p$. By induction, for every $\ell \in \{0,\ldots,m\}$ it holds that $\mathsf{P}^{\ell}_{\text{View},I^*}(\text{View is not typical}) \in O(mn \cdot 2^{-n}/\varepsilon_k) \in O(2^{-n/2})$.

Since the first ℓ rounds of Exp^{ℓ} and $\text{Exp}^{\ell+1}$ are the same, it suffices to prove that $\mathsf{P}^{\ell}_{\text{View},I^*}$ and $\mathsf{P}^{\ell+1}_{\text{View},I^*}$ are close conditioned on a *fixed* value of view as calculated after the ℓ^{th} call to GetNextRC. The proof of Claim 3.7 yields that $\mathsf{P}^{m}_{\text{View},I^*}[\exists j \in [m] \colon \Gamma_j < (\beta_j + \frac{1}{m}) \vee$ the execution aborts] is negligible. Hence, it suffices to prove that each neighbor distributions are closed given that the fixed value of view is typical, and that in the $\ell+1$ call to GetNextRC(view, $\perp, \perp$) done in $\text{Exp}^{\ell+1}$, it holds that $\gamma \geq \beta_{\ell+1} + \frac{1}{m} - O(\mu) > \beta_{\ell+1} + \frac{1}{2m}$.

We assume for simplicity that $m > 1$ and upper bound the statistical difference between $\mathsf{P}^{\ell}_{\text{View},I^*}$ and $\mathsf{P}^{\ell+1}_{\text{View},I^*}$ separately for $\ell = 1$, $2 \leq \ell < m$ and $\ell = m$.

1. For every $\ell \in [m-2]$, it holds that $\left\|\mathsf{P}^{\ell}_{\text{View},I^*} - \mathsf{P}^{\ell+1}_{\text{View},I^*}\right\| \in O(\mu)$.

   Proof. Since the only difference between Exp^{ℓ} and $\text{Exp}^{\ell+1}$ is in the $\ell+1$ call to GetNextRC and in the way i^* is chosen, it suffices to prove the following distributions are close.

   $\mathsf{D}^{0}_{I^*,S} := (I^* \leftarrow \mathcal{S}_{>\ell}(\text{view}), S = \text{GetNextRC}(\text{view}, I^*, \perp)$, and

   $\mathsf{D}^{1}_{I^*,S} := (S = \text{GetNextRC}(\text{view}, \perp, \perp), I^* \leftarrow \mathcal{S}_{>\ell}(\text{view}, S))$

   Consider the hybrid distribution

   $\mathsf{D}^{\frac{1}{2}}_{I^*,S} := (I^* \leftarrow \mathcal{S}_{>\ell}(\text{view}), S = \text{GetNextRC}'(\text{view}, I^*, \perp)$,

   where GetNextRC$'$ is a variant of GetNextRC that in Line $2.(a).i$ does not condition on $i^* \in \mathcal{S}_{round+1}(\text{view}')$ (even if $i^* \neq \perp$). The proof in concluded through by the following two claims.

   Claim 3.10. $\left\|\mathsf{D}^{\frac{1}{2}}_{I^*,S} - \mathsf{D}^{1}_{I^*,S}\right\| \in O(\mu)$.

   Proof. Let $k_\ell := k_\ell(\text{view})$ and consider the real function $f \colon \{0,1\}^{k_\ell} \mapsto [0,1]$ defined as $f(x) = \delta_{\text{view},\perp,\perp}(x)$. Let $\delta := \mathsf{E}_{U^p_{k_\ell}}[f]$, let $\delta_i := \mathsf{E}_{U^p_{k_\ell,i=0}}[f]$ and LargeEfct $:= \{i \in [k_\ell] \colon \delta_i \notin (1 \pm \mu) \cdot \delta\}$. Our assumption about view yields that $\delta \geq \varepsilon_k/2$, and thus by Lemma 2.3 it holds that $|\text{LargeEfct}| = q \in O(n^2m^2/\mu^2)$. Let $(i,x) \in \text{Supp}(\mathsf{D}^{1}_{I^*,S})$, and assume that $i \notin$ LargeEfct. It follows that,

   $$\mathsf{D}^{1}_{I^*,S}(i,x) := \mathsf{D}^{1}_{S}(x) \cdot \mathsf{D}^{1}_{I^*|S=x}(i) = \left(\frac{C^1}{\delta} \cdot U^p_{k_\ell}(x) \cdot f(x)\right) \cdot \frac{1}{w(x)} \enspace ,$$

   where C^1 is the probability that $\mathsf{D}^{1}_{I^*,s}$ does not abort and $w(x) := |\{i \in k_\ell \colon x[i] = 1\}|$. Similarly, for $\mathsf{D}^{\frac{1}{2}}_{I^*,S}$ it holds that

   $$\mathsf{D}^{\frac{1}{2}}_{I^*,S}(i,x) = \mathsf{D}^{\frac{1}{2}}_{I^*}(i) \cdot \mathsf{D}^{\frac{1}{2}}_{S|I^*=1}(x) = \frac{1}{k_\ell} \cdot \left(\frac{C^{\frac{1}{2}}_i}{\delta_i} \cdot U^p_{k_\ell,i=1}(x) \cdot f(x)\right) \enspace ,$$

   where $C^{\frac{1}{2}}_i$ is the probability that $\mathsf{D}^{\frac{1}{2}}_{I^*=i,S}$ does not abort. Equation 4 yields that $C^1 \geq 1 - O(2^{-n})$, and since $i \notin$ LargeEfct, the former also holds for $C^{\frac{1}{2}}_i$. It follows that $\frac{\mathsf{D}^{1}_{I^*,S}(i,x)}{\mathsf{D}^{\frac{1}{2}}_{I^*,S}(i,x)} \in (1 \pm 2\mu) \cdot (1 \pm \mu) \in (1 \pm O(\mu))$.

We conclude that

$$
\begin{aligned}
&\left\|\mathsf{D}^{\frac{1}{2}}_{I^*,S} - \mathsf{D}^{1}_{I^*,S}\right\| \\
&\leq \quad \mathsf{D}^{\frac{1}{2}}_{I^*}(\mathrm{LargeEfct}) + \sum_{(i,x)\colon\, i \notin \mathrm{LargeEfct}} \left|\mathsf{D}^{1}_{I^*,S}(i,x) - \mathsf{D}^{\frac{1}{2}}_{I^*,S}(i,x)\right| \\
&\leq \quad q/k + O(\mu) \in O(\mu) \ ,
\end{aligned}
$$

where the first inequality is due to Proposition 2.1. □

Claim 3.11. $\left\|\mathsf{D}^{0}_{I^*,S} - \mathsf{D}^{\frac{1}{2}}_{I^*,S}\right\| \in O(\mu)$

Proof. Let f, δ, δ_i and LargeEfct be as in the proof of Claim 3.10. Define $f^0_i : \{0,1\}^{k_\ell} \mapsto [0,1]$ as $f^0_i(x) = \delta_{\mathrm{view},i,\perp}(x)$, and let $\delta^0_i := \mathsf{E}_{U^p_{k_\ell,i=0}}[f^0_i]$. Note that $\mathsf{D}^{\frac{1}{2}}_{I^*,S}(i,x) = \frac{1}{k_\ell} \cdot \frac{f(x)}{\delta_i}$ and that $\mathsf{D}^{0}_{I^*,S}(i,x) = \frac{1}{k_\ell} \cdot \frac{f^0_i(x)}{\delta^0_i}$.

For $\mathcal{M} \subseteq [k_\ell]$, let $k_{\mathcal{M}} := k_\ell - |\mathcal{M}|$, where we identify the indices in $[k_\ell] \setminus \mathcal{M}$ with the set $[k_{\mathcal{M}}]$. Define $g_{\mathcal{M}} : \{0,1\}^{k_{\mathcal{M}}} \mapsto [0,1]$ as $g_{\mathcal{M}}(x) = \gamma_{(\mathrm{view},\mathcal{M}),\perp,\perp}(x)$. Let $\gamma(\mathcal{M}) = \mathsf{E}_{U^p_{k_{\mathcal{M}}}}[g_{\mathcal{M}}]$, and for $i \in [k_{\mathcal{M}}]$ let $\gamma_i(\mathcal{M}) := \mathsf{E}_{U^p_{k_{\mathcal{M}},i=1}}[g_{\mathcal{M}}]$.

Let Light $:= \{\mathcal{M} \subset [k_\ell] \colon \gamma(\mathcal{M}) < \varepsilon_k/2m\}$. Let $\mathrm{LargeEfct}_{\mathcal{M}} := \left\{i \in [k_{\mathcal{M}}] \colon \gamma_i(\mathcal{M}) \notin (1 \pm \frac{\mu}{8nm}) \cdot \gamma(\mathcal{M})\right\}$, by Lemma 2.3 we have that $|\mathrm{LargeEfct}_{\mathcal{M}}| = q \in O(n^4m^4/\mu^2)$ for every $\mathcal{M} \notin$ Light. Let $\mathrm{LargeEfct}(i) := \left\{y \in k_\ell \colon i \in \mathrm{LargeEfct}_y\right\}$ and let PotentialLargeEfct be the $\mu \cdot k_\ell$ heaviest indices in k_ℓ according to $\mathsf{E}_{U^p_{k_\ell}}[\gamma \cdot 1_{\mathrm{LargeEfct}(i)}]$, breaking ties arbitrarily. Since $\sum_{i\in[k_\ell]} \mathsf{E}_{U^p_{k_\ell}}[\gamma \cdot 1_{\mathrm{LargeEfct}(i)}] \leq q \cdot \mathsf{E}_{U^p_{k_\ell}}[\gamma]$, for every $i \notin$ PotentialLargeEfct it holds that $\mathsf{E}_{U^p_{k_\ell}}[\gamma \cdot 1_{\mathrm{LargeEfct}(i)}] \leq \frac{q}{|\mathrm{PotentialLargeEfct}|} \cdot \mathsf{E}_{U^p_{k_\ell}}[\gamma]$. An average argument yields that for every $i \notin$ PotentialLargeEfct, it holds that $\mathsf{E}_{U^p_{k_\ell,i=0}}[\gamma \cdot 1_{\mathrm{LargeEfct}(i)}] \leq \frac{1}{p} \cdot \frac{q}{|\mathrm{PotentialLargeEfct}|} \cdot \mathsf{E}_{U^p_{k_\ell}}[\gamma]$. Lemma 2.3 yields that $\gamma_i(\mathcal{M}) \leq \frac{1}{p} \cdot \gamma(\mathcal{M})$ for every $\mathcal{M}$, and therefore for every $i \notin$ PotentialLargeEfct

$$
\begin{aligned}
\mathsf{E}_{U^p_{k_\ell},i=1}[\gamma_i \cdot 1_{\mathrm{LargeEfct}(i)}] &\leq \frac{1}{p^2} \cdot \frac{q}{|\mathrm{PotentialLargeEfct}|} \cdot \mathsf{E}_{U^p_{k_\ell}}[\gamma] \\
&\leq \frac{1}{p^2} \cdot \frac{q}{|\mathrm{PotentialLargeEfct}|} \cdot 2m \cdot \delta \in O(\mu^2 \cdot \delta) \ ,
\end{aligned}
\tag{5}
$$

where the second inequality is due to Equation 3 (recall that our assumption about view yields that $\mathsf{E}_{U^p_{k_\ell}}[\gamma] \geq \beta_{\ell+1} + \frac{1}{2m}$). Hence,

$$
\begin{aligned}
\mathsf{D}^{0}_{S|I^*=i\notin\mathrm{PotentialLargeEfct}}(I^* \in \mathrm{LargeEfct}_S) &\leq \frac{\mathsf{E}_{U^p_{k_\ell,i=0}}[f^0_i \cdot 1_{\mathrm{LargeEfct}(i)}]}{\delta^0_i} \\
&\leq \frac{(1 + \frac{\mu}{8nm}) \cdot \mathsf{E}_{U^p_{k_\ell},i=1}[\gamma_i \cdot 1_{\mathrm{LargeEfct}(i)}]}{\delta^0_i} \in O(\mu^2 \cdot \frac{\delta}{\delta^0_i}) \ ,
\end{aligned}
\tag{6}
$$

where the second inequality is due to Claim 3.9(1). Since for every $(i,\mathcal{M}) \in [k_{\mathcal{M}}] \times \{0,1\}^{k_\ell}$ such that $i \notin \mathrm{LargeEfct}_{\mathcal{M}}$, it holds that $\gamma_i(\mathcal{M}) \in (1 \pm \frac{\mu}{8nm}) \cdot \gamma(\mathcal{M})$, Claim 3.9(4) yields that $f^0_i(\mathcal{M}) \in (1 \pm O(\mu)) \cdot f(\mathcal{M})$. It follows that

$$
\frac{\mathsf{D}^{\frac{1}{2}}_{I^*,S}(i,\mathcal{M})}{\mathsf{D}^{0}_{I^*,S}(i,\mathcal{M})} \in (1 \pm O(\mu)) \cdot \frac{\delta^0_i}{\delta_i} \ , \tag{7}
$$

for every such pair. For $\delta_i \in \Omega(2^{-n}/\mu)$ it holds that

$$\begin{aligned}
\delta_i^0 &= \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0] \\
&= \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \cdot 1_{\text{LargeEfct}(i)}] \cdot \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \mid 1_{\text{LargeEfct}(i)}] \\
&+ (1 - \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \cdot 1_{\text{LargeEfct}(i)}]) \cdot \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \mid \overline{1_{\text{LargeEfct}(i)}}\ ] \\
&\in \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \cdot 1_{\text{LargeEfct}(i)}] \cdot \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \mid 1_{\text{LargeEfct}(i)}] \\
&+ (1 - \mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \cdot 1_{\text{LargeEfct}(i)}]) \cdot (1 \pm O(\mu)) \cdot \mathsf{E}_{U^p_{k_\ell,i=0}}[f \mid \overline{1_{\text{LargeEfct}(i)}}\ ] \\
&\in (1 \pm O(\mu)) \cdot \Big(\delta_i \pm O(\mathsf{E}_{U^p_{k_\ell,i=0}}[f_i^0 \cdot 1_{\text{LargeEfct}(i)}] + \mathsf{E}_{U^p_{k_\ell,i=0}}[f \cdot 1_{\text{LargeEfct}(i)}])\Big) \\
&\in (1 \pm O(\mu)) \cdot (\delta_i \pm O(\mathsf{E}_{U^p_{k_\ell},i=1}[\gamma_i \cdot 1_{\text{LargeEfct}(i)}] + \mathsf{E}_{U^p_{k_\ell},i=1}[\gamma \cdot 1_{\text{LargeEfct}(i)}])
\end{aligned}$$

where in the first '$\in$' follows from Claim 3.9(2) and the fact that $f_i^0(\mathcal{M}), f(\mathcal{M}) \in O(2^{-n})$ for $\mathcal{M} \in$ Light (see Claim 3.9(3)), and the last '$\in$' follows from Claim 3.9(1). The above and Equation 5 yield the following holds for every $i \notin$ PotentialLargeEfct such that $\delta_i \in \Omega(2^{-n}/\mu)$

$$\delta_i^0 \in (1 \pm O(\mu)) \cdot (\delta_i \pm O(\mu^2 \cdot \delta)) \tag{8}$$

We conclude that

$$\begin{aligned}
&\left\| \mathsf{D}^0_{I^*,S} - \mathsf{D}^{\frac{1}{2}}_{I^*,S} \right\| \\
&\quad\le \mathsf{D}^0_{I^*}(\text{Bad}) + \mathsf{D}^0_{I^*,S|I^*\notin\text{Bad}}(I^* \in \text{LargeEfct}_S) + \sum_{(i,x):\ i \notin (\text{LargeEfct}_x\ \cup\ \text{Bad})} \left| \mathsf{D}^{\frac{1}{2}}_{I^*,S}(i,x) - \mathsf{D}^0_{I^*,S}(i,x) \right| \\
&\quad\le +O(\mu) + O(\mu) + O(\mu) \in O(\mu)\ ,
\end{aligned}$$

where Bad := LargeEfct ∪ PotentialLargeEfct □

□

2. $\left\| \mathsf{P}^{m-1}_{\text{View},I^*} - \mathsf{P}^{m}_{\text{View},I^*} \right\| \in O(\mu)$.

   Proof. Note that for every fixing of the parameter view, the statistical difference between $\mathsf{P}^0_{\text{View},I^*}$ and $\mathsf{P}^1_{\text{View},I^*}$ equals the statistical difference between $\mathsf{D}^{\frac{1}{2}}_{I^*,S}$ and $\mathsf{D}^1_{I^*,S}$ (as defined in the proof of 1). Hence, the proof follows the lines of the proof for Claim 3.10. □

3. $\left\| \mathsf{P}^0_{\text{View},I^*} - \mathsf{P}^1_{\text{View},I^*} \right\| \in O(\mu)$.

   Proof. As in the proof of 1, it suffices to show that the following distributions are close:

   $\mathsf{D}^0_{I^*,R^k} := (I^* \leftarrow [k], R^k = \text{GetNextRC}(\perp, I^*, U_{\text{len}}))$, and

   $\mathsf{D}^1_{I^*,S} := (I^* \leftarrow [k], R^k = \text{GetNextRC}(\perp, \perp, \perp))$

   Consider the hybrid distribution

   $\mathsf{D}^{\frac{1}{2}}_{I^*,R^k} := (I^* \leftarrow [k], R^k = \text{GetNextRC}'(\perp, I^*, U_{\text{len}}))$,

   where (as in the proof of 1.) GetNextRC$'$ is a variant of GetNextRC that does not condition in Line $2.(a).i$ on $i^* \in \mathcal{S}_{round+1}(\text{view}')$. We conclude the proof by the following two claims.

   Claim 3.12. $\left\| \mathsf{D}^{\frac{1}{2}}_{I^*,R^k} - \mathsf{D}^1_{I^*,R^k} \right\| \in O(\mu)$.

Proof. Let P_I and $\mathsf{P}_{X^k} := \mathsf{P}_{X_1} \cdots \mathsf{P}_{X_k}$ be the uniform distribution over $[k]$ and $\{0,1\}^{k \cdot \mathrm{len}}$ respectively, and let W be the event that the loop of GetNextRC$(\perp, \perp, \perp)$ returns. Equation 3 yields that $\mathsf{P}_{X^k}(W \mid r^k = X^k) \geq \varepsilon/m$, which yields that

$$\left\| \mathsf{P}_I \cdot \mathsf{P}_{X^k|W} - \mathsf{D}^1_{I^*,R^k} \right\| \in O(2^{-n}) \tag{9}$$

Applying Corollary 2.6 with $f(x^k) = \mathsf{Pr}[W \mid r^k = x^k]$, we have that $\left|\left\{ i \in [k] \colon \mathsf{P}_{X_i}\left(\mathsf{P}_{X^k_{-i}}(W \mid r^k = X^k)) < \varepsilon/2m \right) > \mu \right\}\right| \in O(n^2/\mu^2)$, where $\mathsf{P}_{X^k_{-i}} = \mathsf{P}_{X_1} \cdots \mathsf{P}_{X_{i-1}} \cdot \mathsf{P}_{X_{i+1}} \cdots \mathsf{P}_{X_k}$. It follows that

$$\left\| \mathsf{P}_I \cdot \mathsf{P}_{X_i} \cdot \mathsf{P}_{X^k_{-i}|W,X_i} - \mathsf{D}^{\frac{1}{2}}_{I^*,R^k} \right\| \in O(n^2/\mu^2 \cdot k) \in O(\mu) \ , \tag{10}$$

where $\mathsf{P}_{X^k_{-i}|W,X_i}$ is set arbitrarily in case $\mathsf{P}_{X^k|X^k_i = X_i}(W \mid r^k = X^k) = 0$. We complete the proof using the following result due to Holenstein, restating a lemma of Raz [Raz98, Claim 5.1].

Lemma 3.13. *([Hol07, Equation 8]) Let $\mathsf{P}_{X^k} := \mathsf{P}_{X_1} \cdots \mathsf{P}_{X_k}$ be a probability distribution over $\mathcal{X}^k$ and let W be an event. Then,*

$$\sum_{j=1}^{k} \left\| \mathsf{P}_{X_j|W} - \mathsf{P}_{X_j} \right\| \leq \sqrt{-k \cdot \log \mathsf{Pr}[W]}.$$

Writing $\mathsf{P}_{X^k|W} = \mathsf{P}_{X_I|W} \mathsf{P}_{X^k_{-I}|X_I,W}$, we have that

$$\left\| \mathsf{P}_I \cdot \mathsf{P}_{X_I} \cdot \mathsf{P}_{X^k_{-I}|W,X_I} - \mathsf{P}_I \cdot \mathsf{P}_{X^k|W} \right\| \in \sqrt{-k \cdot \log \mathsf{Pr}[W]}) \in O(\mu) \ ,$$

and therefore

$$\begin{aligned} & \left\| \mathsf{D}^{\frac{1}{2}}_{I^*,R^k} - \mathsf{D}^1_{I^*,R^k} \right\| \\ \leq \quad & \left\| \mathsf{P}_I \cdot \mathsf{P}_{X^k|W} - \mathsf{D}^1_{I^*,R^k} \right\| + \left\| \mathsf{P}_I \cdot \mathsf{P}_{X_i} \cdot \mathsf{P}_{X^k_{-i}|W,X_i} - \mathsf{D}^{\frac{1}{2}}_{I^*,R^k} \right\| \\ + \quad & \left\| \mathsf{P}_I \cdot \mathsf{P}_{X_I} \cdot \mathsf{P}_{X^k_{-I}|W,X_I} - \mathsf{P}_I \cdot \mathsf{P}_{X^k|W} \right\| \in O(\mu) \ . \end{aligned}$$

□

Claim 3.14. $\left\| \mathsf{D}^0_{I^*,R^k} - \mathsf{D}^{\frac{1}{2}}_{I^*,R^k} \right\| \in O(\mu)$.

Proof. The proof follows similar lines to the proof of Claim 3.11, but involves an additional complication since I^* effects not only the way the "quality" of next message is evaluate (as in the proof Claim 3.11), but also determines which of the coordinates of R^k is sampled only once. In particular, we have to make sure that this change does not increase by too much the probability that a random I^* has large effect.

Define $f, f^0_{i,r} \colon \{0,1\}^{(k-1)\cdot \mathrm{len}} \mapsto [0,1]$ as $f_{i,r}(x) := \delta_{\perp,\perp,\perp}(x[1], \ldots, x[i-1], r, \ldots, x[k-1])$ and $f^0_{i,r}(x) = \delta_{\perp,i,r}(x[1], \ldots, x[i-1], r, \ldots, x[k-1])$. Let $\delta_{i,r} = \mathsf{E}_{(U_n)^{k-1}}[f_{i,r}]$ and let $\delta^0_{i,r} = \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}]$. For $r^k \in \{0,1\}^{k \cdot \mathrm{len}}$, define $g_{r^k} : \{0,1\}^k \mapsto [0,1]$ as $g_{r^k}(x) := \gamma_{(\mathrm{view},\mathcal{M}),\perp,\perp}(x)$. Let $\gamma(r^k) = \mathsf{E}_{U^p_k}[g_{r^k}]$, and let $\gamma_i(r^k) = \mathsf{E}_{U^p_{k,i=1}}[g_{r^k}]$.

Let Light $:= \left\{ r^k \in \{0,1\}^{k \cdot \mathrm{len}} \colon \gamma(r^k) < \varepsilon_k/2m \right\}$ and for $r^k \notin$ Light, let LargeEfct$_{r^k}$ $:= \left\{ i \in [k] \colon \gamma_i(r^k) \notin (1 \pm \frac{\mu}{8nm}) \cdot \gamma(r^k) \right\}$. Lemma 2.3 yields that $|\mathrm{LargeEfct}_{r^k}| = q \in O(n^4 m^4/\mu^2)$. Let LargeEfct$(i) := \left\{ y \in \{0,1\}^{k \cdot \mathrm{len}} \colon i \in \mathrm{LargeEfct}_y \right\}$ and let PotentialLargeEfct be the $\mu \cdot k$ heaviest indices in k according to $\mathsf{E}_{U_{k \cdot \mathrm{len}}}[\gamma \cdot 1_{\mathrm{LargeEfct}(i)}]$, breaking ties arbitrarily.

We would like to argue that $\mathsf{D}^0_{I^*,R^k|I^*\notin \mathrm{PotentialLargeEfct}}(I^* \in \mathrm{LargeEfct}_{R^k})$ is small. Unlike the proof of Claim 3.11, the above is note true for every $I^* \notin$ PotentialLargeEfct. Nevertheless, we manage to show that it is so for most values of (i,r). Let PotentialLargeEfctPair := $\left\{(r,i)\colon \mathsf{E}_{U_{k\cdot\mathrm{len}}|U_{k\cdot\mathrm{len}}[i]=r}[\gamma\cdot 1_{\mathrm{LargeEfct}(i)}] > \frac{1}{\mu}\cdot \mathsf{E}_{U_{k\cdot\mathrm{len}}}[\gamma\cdot 1_{\mathrm{LargeEfct}(i)}]\right\}$. A Markov bound yields that

$$\Pr[(U_n, I^*) \in \mathrm{PotentialLargeEfctPair}] \leq \mu \tag{11}$$

An averaging argument yields that for every $i \notin$ PotentialLargeEfct it holds that $\mathsf{E}_{U_{k\cdot\mathrm{len}}}[\gamma\cdot 1_{\mathrm{LargeEfct}(i)}] \leq \frac{q\cdot \mathsf{E}_{U_{k\cdot\mathrm{len}}}[\gamma]}{|\mathrm{PotentialLargeEfct}|}$. Lemma 2.3 yields that $\gamma_i(r^k) \leq \frac{1}{p}\cdot\gamma(r^k)$ for every r^k, and therefore for $(i,r) \notin$ PotentialLargeEfctPair such that $i \notin$ PotentialLargeEfct it holds that

$$\mathsf{E}_{U_{k\cdot\mathrm{len}}|U_{k\cdot\mathrm{len}}[i]=r}[\gamma_i\cdot 1_{\mathrm{LargeEfct}(i)}] \leq \frac{\mu q\cdot \mathsf{E}_{U_{k\cdot\mathrm{len}}}[\gamma]}{|\mathrm{PotentialLargeEfct}|} \in O(\mu\cdot\varepsilon_k)\ , \tag{12}$$

where the last conclusion is due to Equation 3. The rest of the proof follows rather closely that of Claim 3.11. Claim 3.9(1) yields that

$$\begin{aligned} &\mathsf{D}^0_{I^*,R^k|I^*=i\notin \mathrm{PotentialLargeEfct}\wedge R^k[i]=r\wedge(i,r)\notin\mathrm{PotentialLargeEfctPair}}(I^*\in\mathrm{LargeEfct}_{R^k})(13)\\ &\leq \quad \frac{\mathsf{E}_{U_{(k-1)\mathrm{len}}}[f^0_{i,r}\cdot 1_{\mathrm{LargeEfct}(i)}]}{\delta^0_{i,r}}\\ &\leq \quad \frac{(1+\frac{\mu}{8nm})\cdot \mathsf{E}_{U_{k\cdot\mathrm{len}}|U_{k\cdot\mathrm{len}}[i]=r}[\gamma_i\cdot 1_{\mathrm{LargeEfct}(i)}]}{\delta^0_{i,r}} \in O(\mu\cdot\frac{\varepsilon_k}{\delta^0_{i,r}}) \end{aligned}$$

Since for every (i,r^k) such that $i\notin \mathrm{LargeEfct}_{r^k}$, it holds that $\gamma_i(r^k)\in(1\pm\frac{\mu}{8nm})\cdot\gamma(r^k)$. Claim 3.9(4) yields that $f^0_{i,r}(r^k)\in(1\pm O(\mu))\cdot f_{i,r}(r^k)$ and therefore

$$\frac{\mathsf{D}^{\frac{1}{2}}_{I^*,R^k}(i,r^k)}{\mathsf{D}^0_{I^*,R^k}(i,r^k)} \in (1\pm O(\mu))\cdot\frac{\delta^0_{i,r}}{\delta_{i,r}}\ , \tag{14}$$

for every such pair. For $\delta_{i,r}\in\Omega(2^{-n}/\mu)$ it holds that

$$\begin{aligned} \delta^0_{i,r} &= \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}]\\ &= \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\cdot 1_{\mathrm{LargeEfct}(i)}]\cdot \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\mid 1_{\mathrm{LargeEfct}(i)}]\\ &+ (1-\mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\cdot 1_{\mathrm{LargeEfct}(i)}])\cdot \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\mid \overline{1_{\mathrm{LargeEfct}(i)}}\ ]\\ &\in \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\cdot 1_{\mathrm{LargeEfct}(i)}]\cdot \mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\mid 1_{\mathrm{LargeEfct}(i)}]\\ &+ (1-\mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\cdot 1_{\mathrm{LargeEfct}(i)}])\cdot(1\pm O(\mu))\cdot \mathsf{E}_{(U_n)^{k-1}}[f\mid \overline{1_{\mathrm{LargeEfct}(i)}}\ ]\\ &\in (1\pm O(\mu))\cdot\big(\delta_{i,r}\pm O(\mathsf{E}_{(U_n)^{k-1}}[f^0_{i,r}\cdot 1_{\mathrm{LargeEfct}(i)}]+\mathsf{E}_{(U_n)^{k-1}}[f\cdot 1_{\mathrm{LargeEfct}(i)}])\big)\\ &\in (1\pm O(\mu))\cdot(\delta_{i,r}\pm O(\mathsf{E}_{U_{k\cdot\mathrm{len}}|U_{k\cdot\mathrm{len}}[i]=r}[\gamma_i\cdot 1_{\mathrm{LargeEfct}(i)}]+\mathsf{E}_{U_{k\cdot\mathrm{len}}|U_{k\cdot\mathrm{len}}[i]=r}[\gamma\cdot 1_{\mathrm{LargeEfct}(i)}]) \end{aligned}$$

where in the first '$\in$' follows from Claim 3.9(2) and the fact that $f^0_{i,r}(r^{k-1}), f(r^{k-1})\in O(2^{-n})$ for in case $(r^{k-1}[1],\ldots,r^{k-1}[i-1],r,\ldots,r^{k-1}[k-1])\in$ Light (see Claim 3.9(3)), and the last '$\in$' follows from Claim 3.9(1). The above and Equation 12 yields the following for every $(i,r)\notin$ PotentialLargeEfctPair such that $i\notin$ PotentialLargeEfct and $\delta_{i,r}\in\Omega(2^{-n}/\mu)$

$$\delta^0_{i,r}\in(1\pm O(\mu))\cdot(\delta_{i,r}\pm O(\mu\cdot\varepsilon_k)) \tag{15}$$

Finally, let LargefctPair $:= \{(r,i) \colon \delta_{i,r} < \varepsilon_k/2\}$, Corollary 2.6 yields $\mathsf{Pr}[(I^*, U_n)) \in \text{LargefctPair}] \in O(\frac{n^2}{k\cdot\mu^2} + \mu) \in O(\mu)$, and we conclude that

$$\begin{aligned}
&\left\|\mathsf{D}^0_{I^*,R^k} - \mathsf{D}^{\frac{1}{2}}_{I^*,R^k}\right\| \\
\leq\ & \mathsf{D}^0_{I^*}(\text{PotentialLargeEfct}) + \mathsf{D}^0_{I^*,R^k}((I^*, R^k_{I^*}) \in \text{BadPairs}) \\
+\ & \mathsf{D}^0_{I^*,R^k | I^* \notin \text{PotentialLargeEfct} \wedge\ (I^*, R^k_{I^*}) \notin \text{BadPairs}}(I^* \in \text{LargeEfct}_{R^k}) \\
+\ & \sum_{(i,x)\colon\ i \notin (\text{PotentialLargeEfct} \cup \text{LargeEfct}_x) \wedge\ (i,x) \notin \text{BadPairs}} \left|\mathsf{D}^{\frac{1}{2}}_{i,x}(i,x) - \mathsf{D}^0_{I^*,x}(i,x)\right| \\
\in\ & O(\mu)\ ,
\end{aligned}$$

where BadPairs := LargefctPair ∪ PotentialLargeEfctPair. □

□

We conclude that $\left\|\mathsf{P}^0_{\text{View},I^*} - \mathsf{P}^m_{\text{View},I^*}\right\| \leq \sum_{\ell=0}^{m-1} \left\|\mathsf{P}^\ell_{\text{View},I^*} - \mathsf{P}^{\ell+1}_{\text{View},I^*}\right\| \in O(m \cdot \mu)$. □

Acknowledgment

I am very thankful to Oded Goldreich, Thomas Holenstein, Omer Reingold, Salil Vadhan and Alex Samorodnitsky for very useful discussion. In particular, I am in great depth to Alex for helping with the proof of Lemma 2.4. I am also very thankful to Tal Moran, Omer Reingold and Gil Segev for helping me to improve the readability of this text.

References

[BIN97] Mihir Bellare, Russell Impagliazzo, and Moni Naor. Does parallel repetition lower the error in computationally sound protocols? In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science (FOCS)*, 1997.

[BM88] László Babai and Shlomo Moran. Arthur-merlin games: A randomized proof system, and a hierarchy of complexity classes. *J. Comput. Syst. Sci.*, 36(2):254–276, 1988.

[CHS05] Canetti, Halevi, and Steiner. Hardness amplification of weakly verifiable puzzles. In *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005*, volume 2, 2005.

[DP98] Ivan B. Damgård and Birgit Pfitzmann. Sequential iteration arguments and an efficient zero-knowledge argument for NP. In *ICALP: Annual International Colloquium on Automata, Languages and Programming*, 1998.

[FK94] Uriel Feige and Joe Kilian. Two prover protocols: low error at affordable rates. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing (STOC)*, pages 172–183, 1994.

[Gol99] Oded Goldreich. *Modern Cryptography, Probalistic Proofs and Pseudorandomness.* Springer-Verlag, Berlin Heidelberg, 1999.

[GRYY08] Ronen Gradwohl, Omer Reingold, Ariel Yadin, and Amir Yehudayoff. The player's effect. Technical Report arXiv:0805.0400v1, 2008.

[HKM06] Hagggstrom, Kalai, and Mossel. A law of large numbers for weighted majority. *ADVAM: Advances in Applied Mathematics*, 37, 2006.

[Hol07] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC)*. ACM Press, 2007.

[HPPW08] Johan Håstad, Rafael Pass, Krzysztof Pietrzak, and Douglas Wikström. An efficient parallel repetition theorem. Unpublished manuscript, 2008.

[HRVW] Iftach Haitner, Omer Reingold, Salil Vadhan, and Hoeteck Wee. Inaccessible entropy. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC).*

[IJK06] Russell Impagliazzo, Ragesh Jaiswal, and Ragesh Kabanets. Approximately list-decoding direct product codes and uniform hardness amplification. In *Proceedings of the 46th Annual Symposium on Foundations of Computer Science (FOCS)*, 2006.

[PV07] Rafael Pass and Muthuramakrishnan Venkitasubramaniam. An efficient parallel repetition theorem for arthur-merlin games. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC)*, 2007.

[PW07] Krzysztof Pietrzak and Douglas Wikström. Parallel repetition of computationally sound protocols revisited. In Salil P. Vadhan, editor, *Theory of Cryptography, Fourth Theory of Cryptography Conference, TCC 2007*, volume 4392 of *Lecture Notes in Computer Science*, pages 86–102. Springer, 2007.

[Raz98] Ran Raz. A parallel repetition theorem. *Journal of the ACM*, 27(3):763–803, 1998. Preliminary version in *STOC'95.*

[Tal96] Michel Talagrand. How much are increasing sets positively correlated? *Combinatorica*, 16(2):243–258, 1996.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/