# The Randomness Complexity of Parallel Repetition

Kai-Min Chung
*Dept. of Computer Science*
*Cornell University*
*chung@cs.cornell.edu*

Rafael Pass
*Dept. of Computer Science*
*Cornell University*
*rafael@cs.cornell.edu*

***Abstract*—Consider a $m$-round interactive protocol with soundness error $1/2$. How much extra randomness is required to decrease the soundness error to $\delta$ through parallel repetition? Previous work, initiated by Bellare, Goldreich and Goldwasser, shows that for *public-coin* interactive protocols with *statistical soundness*, $m \cdot O(\log(1/\delta))$ bits of extra randomness suffices. In this work, we initiate a more general study of the above question.**

- **We establish the first derandomized parallel repetition theorem for public-coin interactive protocols with *computational soundness* (a.k.a. arguments). The parameters of our result essentially matches the earlier works in the information-theoretic setting.**
- **We show that obtaining even a sub-linear dependency on the number of rounds $m$ (i.e., $o(m) \cdot \log(1/\delta)$) is impossible in the information-theoretic setting, and requires the existence of one-way functions in the computational setting.**
- **We show that non-trivial derandomized parallel repetition for private-coin protocols is impossible in the information-theoretic setting and requires the existence of one-way functions in the computational setting.**

**These results are tight in the sense that parallel repetition theorems in the computational setting can trivially be derandomized using pseudorandom generators, which are implied by the existence of one-way functions.**

***Keywords*-interactive protocols, derandomization, parallel repetition, soundness amplification, randomness extractors**

## I. Introduction

In an *interactive protocol*, two parties, called the prover $P$ and the verifier $V$, receive some common inputs and perhaps some private inputs, toss some random coins, and interact with each other following some prescribed protocol. The prover attempts to convince the verifier $V$ than a certain input $x$ is in a language $L$. The soundness property of an interactive proof requires that when $x \notin L$, $V$ will only accept with bounded (error) probability, even when he interacts with a certain class of adversarial cheating provers $P^*$. Such an upper bound on the error probability of $V$ is called the *soundness error* of the protocol.

Two versions of the soundness property, *statistical soundness* and *computational soundness*, are commonly studied. *Statistical soundness* requires the upper bound on $V$'s error probability (to accept incorrectly) to hold against computationally unbounded adversarial provers, whereas *computational soundness* only requires the soundness to hold against probabilistic polynomial time (PPT). Computational soundness is a weaker requirement than statistical soundness. However, in many settings, requiring only computational soundness allows us to improve the efficiency (e.g., in round complexity or communication complexity). When statistical soundness holds, the protocol is referred to as an interactive *proof*, whereas if only computational soundness holds, the protocol is referred to as an interactive *arguments*.

Ideally, we would like the soundness error to be negligible. But, in many settings, our starting point is a protocol with somewhat large soundness error. For example, to design an interactive proof for a language $L$, it may be easier to first design a protocol with soundness error $1/2$. This leads to the question of *soundness amplification*: Is there a way to decrease the soundness error of a given protocol?

A natural approach to soundness amplification is by *parallel repetition*, i.e., many instances of the protocols are executed in parallel, and the verifier accepts iff all instances accept. We denote by $\Pi^k = (P^k, V^k)$ the $k$-fold parallel repetition of a protocol $\Pi = (P, V)$. It is known that parallel repetition decrease soundness error at an optimal rate (i.e., from $1/2$ to $1/2^k$ under $k$-fold repetition) for interactive proofs [BM88], [Gol01]. For the arguments case, optimal parallel repetition theorems are known for three-message private-coin protocols [BIN97], [CHS05] and for public-coin protocols (i.e., protocols where the verifier does not keep secret) [PV07], [HPWP10], [CL10].

Note, however, that the parallelized verifier $V^k$ uses $k$ times more randomness than the original verifier $V$. We may also consider a *derandomized* parallel verifier $V_G^k$ who generates $V^k$'s coins by applying the function $G$ to a "short" random seed; we usually refer to the function $G$ as a *derandomizer*. A natural question that arises is thus:

*Given an $m$-round protocol with soundness error, say $1/2$, how much extra randomness is required*

Chung is supperted by a Simons Foundation Fellowship.

Pass is supported in part by a Alfred P. Sloan Fellowship, Microsoft New Faculty Fellowship, NSF CAREER Award CCF-0746990, AFOSR YIP Award FA9550-10-1-0093, and DARPA and AFRL under contract FA8750-11-2-0211. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the US government.

*to decrease the soundness error to $\delta$ through parallel repetition?*

This question was first addressed by Bellare, Goldreich, and Goldwasser [BGG93] in 1990. They established that for public-coin interactive proofs with verifier sending $t$-bits messages in each round, $m \cdot (t + O(\log(1/\delta))$ bits of randomness suffice. Their construction is based on the notion of an *averaging sampler* introduced by Bellare and Rompel [BR94]. Relying on more recent constructions of averaging samplers [Zuc97], [RVW01], the extra randomness required can be reduced to $m \cdot O(\log(m/\delta))$.

In this work, we initiate a more general study of the above question. More precisely, we focus on 1) extending the above treatment to both computationally sound protocols and private coin protocols, and 2) investigating the randomness complexity required to perform soundness amplification through parallel repetition in all of these settings.

### A. Our Results

We establish the first derandomized parallel repetition theorem for public-coin interactive protocols with *computational soundness*. The parameters of our result essentially matches the earlier works in the information-theoretic setting.

*Theorem 1:* **(informal)** For every $m, \delta$, there exists a polynomial $k$ and a polynomial-time computable derandomizer $G$, such that for any $m$-round public-coin argument $(P, V)$ with soundness error $1/2$, the protocol $(P^k, V_G^k)$ has soundness error $\delta + \mathsf{ngl}$ and uses only $m \cdot O(\log(m/\delta))$-bits of extra randomness (compared to the original verifier $V$).

We mention that the framework of Bellare *et. al.* [BGG93] does not apply to the computational setting. Rather, we develop a new framework for establishing the above theorem. Our framework applies to both the computational and the information-theoretic settings (i.e., for both proofs and arguments); incidentally, our analysis actually slightly improves the concrete parameters also of earlier works in the information-theoretic setting. Our main technique for proving Theorem 1 is establishing a connection between certain types of, so-called *invertible*, randomness extractors for high entropy sources, and derandomized parallel repetition; roughly speaking, we say that a strong randomness extractor is invertible if we can, given any $r, y$, efficiently sample a uniform $x$ such that $\mathrm{Ext}(x, r) = y$. The extra randomness required in the statement of Theorem 1 corresponds to $m$ times the "entropy-loss" of the invertible extractor used; the parameters in Theorem 1 are then obtained by relying on extractors due to Reingold *et. al.* [RVW01] and Guruswami *et. al.* [GUV09].

We mention that in the computational setting, the existence of a cryptographic pseudorandom generator (PRG) [BM84], [Yao82], which in turn is implied by one-way functions [HILL99] can trivially be used to derandomize interactive protocols. Thus assuming the existence of one-way functions, derandomized parallel repetition doesn't require any additional randomness. The obvious advantage of Theorem 1 is that the derandomization is *unconditional*. A more subtle advantage of our approach is that the resulting protocol remains sound, even if the verifier, at each round, reveals all its random coin tosses. This has two benefits: By slightly changing the prover strategy (to expand the verifier messages using our derandomizer $G$) we can get a derandomized protocol that still is public coin and only increases the verifier communication complexity by an *additive* term of $m \cdot O(\log(m/\delta))$, whereas the PRG solution is private-coin, and increases the verifier communication complexity by a *multiplicative* factor of $k = \log(1/\delta)$.

We next show that obtaining even a sub-linear dependency on the number of rounds $m$ is impossible in the information-theoretic, and requires the existence of one-way functions in the computational setting.

*Theorem 2:* **(informal)** Consider some $m > 0, \delta > 0$. There does not exist a derandomizer $G$ and a polynomial $k$ such that for every $m$-round public-coin proof $(P, V)$ with soundness error $1/2$, the protocol $(P^k, V_G^k)$ has soundness error $\delta$ and uses only $m \cdot \log(1/\delta) - O(1)$-bits of extra randomness. For the case of arguments, the existence of such a derandomizer instead implies the existence of one-way functions.

Our lower bound is actually a bit stronger than stated: we present a specific protocol for which every derandomizer must use at least $m \cdot \log(1/\delta) - O(1)$-bits of extra randomness. Additionally, note that we cannot hope to get an unconditional lower-bound in the computational setting, since, as mentioned, assuming the existence of a PRG, parallel repetition theorems can be derandomized without any additional randomness. However, if we require that the derandomizer protocol remains secure even if the verifier reveals its random coins at each round (which is the case for the derandomized protocols from our upper-bound), then the above lower bound holds unconditionally also in the computational setting.

We finally show that non-trivial derandomized parallel repetition for private-coin protocols is impossible in the information-theoretic setting, and requires proving the existence of one-way functions in the computational setting.

*Theorem 3:* **(informal)** Consider some $m \geq 3, \delta > 0$. There does not exist a derandomizer $G$ and a polynomial $k$ such that for every $m$-round proof $(P, V)$ with soundness error $1/2$, the protocol $(P^k, V_G^k)$ has soundness error $\delta$ and uses $o(t \cdot \log(1/\delta))$ bits of randomness, where $t$ is the random complexity of $V$. For the case of arguments, the existence of such a derandomizer instead implies the existence of one-way functions.

As before, we here present a specific protocol for which every derandomizer must use at least $O(t \cdot \log(1/\delta))$ bits of randomness. Again, this results is "tight" as in the computational setting, pseudorandom generators can be used to trivially derandomize parallel repetition theorems also for private-coin interactive protocols.

*Future Work:* Our results establish essentially tight upper and lower bounds on the randomness/soundness trade-off for soundness amplification through parallel repetition. We have not focused on minimizing the number of parallel repetitions needed for such randomness efficient soundness amplification; in other words, we have not focused on minimizing prover communication complexity. In our framework the number of parallel repetitions is 2 to the power of the seed length of the extractor we use. So one method for decreasing the number of parallel repetitions (and thus improving prover communication comlexity) would be to improve the seed length of known (invertible) extractors for high entropy sources, without hurting the entropy loss. We leave open the question of determining the trade-off between randomness, soundness and the number of parallel repetitions.

We have only focused on methods for derandomizing the *original* verifier strategy; that is, we consider a derandomizer that generates random coins, and then runs parallel instances of the original verifier. A less restrictive approach would be to allow the derandomizer to arbitrarily change the parallel verifier strategy (i.e., both how its messages are generated, and its acceptance rule) subject to preserving the completeness condition with respect to the original prover strategy. Our lower bounds do not extend to this setting. We leave open the question of whether more randomness efficient parallel repetition can be performed in this model.

Finally, in this paper we have only focused on establishing direct product theorems, as opposed to "Chernoff-type" theorems. Nonetheless, although we haven't checked the details, it seems that our techniques directly extend also to the Chernoff bound setting by plugging them into the framework of [HPWP10]; we leave the details for future work.

### B. Organization of the Paper

Section II presents some notations and preliminaries on interactive protocols. We define the notion of derandomizers in Section III. In Section IV, we present our main derandomized parallel repetition theorem (formal version of Theorem 1) and our new framework for proving the theorem; the formal proof is deferred to Section VI. We introduce preliminaries on randomness extractors and state the extractors we use in Section V. Finally, we present our lower bounds in Section VII and VIII.

## II. Preliminaries

We use $\mathbb{N}$ to denote the natural numbers $\{0,1,\dots\}$, $[n]$ to denote the set $\{1,\dots,n\}$, and $|x|$ to denote the length of a string $x \in \{0,1\}^*$. By $\mathsf{ngl}$, we mean a function negligible in $n$ (i.e., $1/n^{\omega(1)}$). All log's are base 2 unless otherwise specified. For random variables $X, Y$, we use $\mathsf{P}_X(x)$ to denote $\Pr[X=x]$ and use $\mathsf{P}_{X|Y}(x|y)$ to denote $\Pr[X=x|Y=y]$.

### A. Interactive Protocols

An **interactive protocol** $\Pi$ is a pair of interactive Turing machines, $(P,V)$, where $V$ is probabilistic polynomial time (PPT). $P$ is called the prover, while $V$ is called the verifier. $\langle P,V\rangle(z)$ denotes the random variable (over the randomness of $P$ and $V$) representing $V$'s output at the end of the interaction on common input $z$. We often omit the input $z$ and simply write $\langle P,V\rangle$. We count one round as two message exchanges from one party to the other and back, so a $m$-round protocol consists of $2m$ messages. $\Pi$ is **public-coin** if the verifier's messages are simply independent uniformly random coins.

We are interested in the trade-off between the randomness complexity and the soundness property of protocols. The **randomness complexity** of a protocol $\Pi$ is simply the number of random coins tossed by the verifier. A protocol $\Pi$ for a language $L$ has **statistical** (resp., **computational**) **soundness error** $\varepsilon$, if for every $z \notin L$, for every unbounded (resp., PPT) adversarial prover $P^*$,

$$\Pr[\langle P^*,V\rangle(z)=1] \leq \varepsilon(|z|).$$

$\Pi$ is referred to as an **interactive proof** (resp., **interactive argument**) for $L$ if $\Pi$ has bounded statistical (reps., computational) soundness error.

Let $\Pi^k=(P^k,V^k)$ denote the $k$**-fold parallel repetition** of $\Pi$, where $k$ independent copies of $\Pi$ are executed in parallel and at the end of interaction, $V^k=(V_1,\dots,V_k)$ accepts iff all sub-verifiers $V_i$'s accept.

## III. Definition of Derandomizers

In this section, we introduce our framework for studying the randomness complexity of parallel repetition by formalizing the notion of a *derandomizer*.

Recall that in a parallel repetition $\Pi^k$ of an interactive protocol $\Pi$, $k$ independent copies of the protocols are executed in parallel, and the parallel verifier $V^k$ accepts iff all sub-verifiers $V_i$'s accept. A *derandomizer* is simply an efficiently computable function $G$ that on input a short seed $U_s$ of $s$ bits randomness, generates a random tape for the parallel verifier $V^k$, i.e., $G:\{0,1\}^s \to \{0,1\}^{k\cdot t}$, where $t$ is the randomness complexity of $\Pi$. In other words, $G$ derandomizes the random tape of $V^k$. This induces a derandomized parallel protocol $\Pi_G$ where the derandomized verifier $V_G$ uses the derandomized random tape generated by $G$ to interact with $P^k$.

*Definition 4 (Derandomizer for Interactive Protocols):* Let $\Pi$ be an interactive proof/argument with soundness error $\varepsilon$ and randomness complexity $t$. A **derandomizer** for $\Pi$ is simply an efficiently computable function $G:\{0,1\}^s \to \{0,1\}^{k\cdot t}$. $G$ induces a derandomized

$k$-fold parallel repetition $\Pi_G = (P^k, V_G)$ of $\Pi$, where the derandomized parallel verifier $V_G$ first generates $G(U_s)$ using a uniform seed $U_s$, and then emulates $V^k$ with coins $G(U_s)$.

We say that $G$ is a $(\varepsilon \mapsto \delta)$**-statistically-sound derandomizer** (resp., **computationally-sound derandomizer**) **for** $\Pi$ if $\Pi_G$ has statistical (resp., computational) soundness error at most $\delta$.

Note that our definition of a derandomizer is general in the sense that we allow the derandomizer to depend on the protocol arbitrarily. Allowing a general definition of a derandomizer makes our lower bound results stronger.

On the other hand, the derandomizer $G$ for public-coin protocols we construct in Theorem 6 works for *all* public-coin protocols of some fixed round complexity and message length. Additionally, as we shall see, there exists an efficient algorithm that given just the rounds complexity $m$ and message length $t$, outputs such a protocol "oblivious" derandomizer $G$. Let us proceed to formalizing the notion of an oblivious derandomizer.

*Definition 5:* Let $\mathcal{C}$ be a class of interactive protocols. $G$ is an $(\varepsilon \mapsto \delta)$**-oblivous statistically-sound** (resp., **computationally-sound**) **derandomizer for** $\mathcal{C}$ if for every protocol $\Pi \in \mathcal{C}$ (for some language $L$) with statistical (resp., computational) soundness error at most $\varepsilon$, the derandomized parallel protocol $\Pi_G$ has statistical (resp., computational) soundness error at most $\delta$.

## IV. Derandomized Parallel Repetition

In this section, we state a formal version of Theorem 1, and give a detailed proof overview. We shall prove Theorem 6 formally in Section VI, after presenting some necessary preliminary on randomness extractors in Section V.

For simplicity of exposition, we consider $m$-round public-coin protocols where the verifier sends a $t$-bit random message at each round. We refer to such protocols as $(m, t)$**-public-coin protocols**. We focus on constructing efficient and oblivious derandomizers for the class of $(m, t)$-public-coin protocols.

*Theorem 6:* For every polynomially bounded $m, t : \mathbb{N} \to \mathbb{N}$, every $\varepsilon, \delta : \mathbb{N} \to (0, 1)$, there exists a $((1-\varepsilon) \mapsto \delta)$-oblivious statistically-sound derandomizer (resp., $((1-\varepsilon) \mapsto \delta + \mathsf{ngl})$-oblivious computationally-sound derandomizer) $G : \{0,1\}^s \to \{0,1\}^{k \cdot m \cdot t}$ for $(m, t)$-public-coin interactive proofs (resp., arguments) with randomness complexity

$$s = m \cdot t + m \cdot O(\log(1/\delta)) + O(\log(m/\varepsilon)),$$

and number of repetition $k = \mathrm{poly}(m, 1/\varepsilon, \log(1/\delta))$. Furthermore, for the case of interactive proofs, the constant in the $O(\log(1/\delta))$ term for the randomness complexity can be set to $(1 + \gamma)$ for an arbitrarily small constant $\gamma$.

Our approach to derandomize parallel repetition of interactive protocols can be viewed as derandomizing the *analysis* of a parallel repetition theorem of Håstad, Pass, Pietrzak, and Wikström [HPWP10]. Therefore, we start by a high level overview of their result and their proof.

Håstad *et. al.* proved an efficient parallel repetition theorem for public-coin interactive arguments, stating that parallel repetition decreases the soundness error at an exponential rate. They proved the theorem by an efficient black-box reduction. Namely, suppose there exists an adversary $P^{k*}$ for the parallel protocol $\Pi^k$ that breaks the soundness with probability $\delta$, then there exists an adversary $P^*$ for the original protocol $\Pi$ such that, given oracle access to $P^{k*}$, $P^*$ can break the soundness with a much higher probability $\varepsilon \gg \delta$, which, in the contrapositive form, shows that the soundness error goes down from $\varepsilon$ to $\delta$ under parallel repetition.

A general framework for such reductions is for the single-instance adversary $P^*$ to interact with the verifier $V$ by simulating the interaction between $P^{k*}$ and $V^k$, where the external verifier $V$ is embedded in some coordinate $V_i$ of $V^k$, and $P^*$ simulates $P^{k*}$ and the remaining $k-1$ sub-verifiers (denoted by $V_{-i}$) of $V^k$ internally and forwards $P^{k*}$'s messages at coordinate $i$ to $V$. The task of $P^*$ is to decide which coordinate to embed $V$, and to choose $k-1$ messages of $V_{-i}$ at each round.

Håstad *et. al.* showed that the following **rejection sampling** strategy $P^*_{rej}$ works. $P^*_{rej}$ simply selects a uniformly random coordinate $i \in [k]$ to embed $V$. At each round $j$, upon receiving the external verifier $V$'s message $x_{j,i}$, $P^*$ repeatedly samples a random continuation of $(P^{k*}, V^k)$ until he finds an **accepting continuation**, i.e., $V^k$ accepts at the end of interaction. Then $P^*$ selects the corresponding messages in the accepting continuation as the messages of $V_{-i}$ at round $j$, and forward the corresponding response of $P^{k*}$ to $V$. If $P^*$ fails to find an accepting continuation, then $P^*$ simply aborts.

To show that the rejection sampling strategy works, we consider a mental experiment, where the external verifier is also aware of the simulated interaction $(P^{k*}, V^k)$, and also uses the rejection sampling strategy to selects his message $x_{j,i}$ at each round $j$. Namely, the verifier also repeatedly samples a random continuation of $(P^{k*}, V^k)$ until a accepting continuation is found, and forwards the corresponding message in the accepting continuation to $P^*_{rej}$. We refer to this mental experiment as the **ideal experiment** $(P^*_{rej}, V^*_{rej})$, in contrast to the **real experiment** $(P^*_{rej}, V)$.

Now, a key observation is that, both parties performing rejection sampling strategy is equivalent to them jointly sampling an accepting interaction. Therefore, in the ideal experiment, the verifier accepts with probability 1 at the end of interaction. The crux of the analysis is to show that the real experiment and the ideal experiment are *statistically close*, using a sampling lemma by Raz [Raz98].

Specifically, let $E$ denote the event of accepting interaction. Recall that $\Pr[E] \geq \delta$. Consider the distribution of $V^k$'s first message $\vec{X}_1 = (X_{1,1}, \ldots, X_{1,k})$. Since $\Pi$ is public-coin, $\vec{X}_1$ is simply a uniform distribution. In the ideal experiment, $P^*_{rej}$ and $V^*_{rej}$ jointly select the first message from the conditional distribution $\vec{X}_1|_E$. In the real experiment, $V$ selects $X_{1,i}$ uniformly without conditioning, and then $P^*_{rej}$ selects the messages of $V_{-i}$ according to distribution $X_{1,-i}|_{E,X_{1,i}}$. Recall that the coordinate $i$ is uniformly random, the statistical distance of the first message between the two experiments is

$$\frac{1}{k}\sum_{i=1}^{k} \mathbf{SD}(X_{1,i}|_E, X_{1,i}),$$

which can be upper bounded by the following Raz's Lemma.

*Lemma 7 (Raz's Lemma [Raz98]):* Let $X_1, \ldots, X_k$ be *independent* random variables on a finite domain $U$. Let $E$ be an event over $\vec{X} = (X_1, \ldots, X_k)$. We have

$$\frac{1}{k} \cdot \sum_{i=1}^{k} \mathbf{SD}(X_i|_E, X_i) \le \sqrt{\frac{1}{k} \cdot \log \frac{1}{\Pr[E]}}.$$

Applying the Raz's Lemma to every round together with a hybrid argument, one can show that the statistical distance between the ideal and the real experiments is at most $m \cdot \sqrt{(\log(1/\delta))/k}$, and hence[1]

$$\Pr[(P^*, V) = 1] \ge 1 - m \cdot \sqrt{\frac{\log(1/\delta)}{k}}.$$

It turns out that to derandomize the parallel repetition theorem, it suffices to derandomize the Raz's Lemma in the sense of identifying derandomized distribution $\vec{X} = (X_1, \ldots, X_k)$ such that the conclusion of the lemma remains true. Note that the lemma is applied to the special case where $X_i$'s are uniform. As observed by Shaltiel [Sha10], in this special case, the Raz's Lemma can be derandomized using strong randomness extractors. Recall that $\mathrm{Ext} : \{0,1\}^n \times [k] \to \{0,1\}^t$ is a strong $(n-\Delta, \varepsilon)$-randomness extractor if for every sources $X$ with min-entropy $\mathrm{H}_\infty(X) \ge n - \Delta$, the distribution $(I, \mathrm{Ext}(X, I))$ is $\varepsilon$-close to $(I, U_t)$ in statistical distance, where $I$ is uniformly random seed over $[k]$. Note that

$$\mathbf{SD}((I, \mathrm{Ext}(X, I)), (I, U_t)) = \frac{1}{k} \cdot \sum_{i=1}^{k} \mathbf{SD}(\mathrm{Ext}(X, i), U_t).$$

Therefore, consider the distribution $(X_1, \ldots, X_k) \triangleq (\mathrm{Ext}(U_n, 1), \ldots, \mathrm{Ext}(U_n, k))$ and an event $E$ over $(X_1, \ldots, X_k)$ with $\Pr[E] \ge 2^{-\Delta}$. Let $X = U_n|_E$, and note that $\mathrm{H}_\infty(X) \ge n - \log(1/\Pr[E]) \ge n - \Delta$. By the property of the extractor,

$$\mathbf{SD}((I, \mathrm{Ext}(X, I)), (I, U_t)) = \frac{1}{k} \cdot \sum_{i=1}^{k} \mathbf{SD}(X_i|_E, U_t) \le \varepsilon,$$

which is the desired conclusion we want from the Raz's Lemma. Therefore, the parallel verifier $V^k$ can be derandomized by replacing the independent messages with the outputs of a strong randomness extractor. Namely, at each round, the derandomized verifier $V_G$ samples $X_j \equiv U_n$ and sends $(\mathrm{Ext}(X_j, 1), \ldots, \mathrm{Ext}(X_j, k))$ to $P^k$.

Note that the randomness extractor we need is only required to extract randomness from sources with high min-entropy. On the other hand, we want to minimize the entropy loss $n-\Delta-t$ (corresponds to the extra randomness used) and the seed length (corresponds to the number of repetition). Randomness extractors for high min-entropy sources with very good parameters has been constructed by Reingold, Vadhan, and Wigderson [RVW01].

However, there are two additional issues that we need to address. First, recall that to finish proof of parallel repetition theorem, we need to apply the Raz's Lemma to each round together with a hybrid argument. Except for the first round, there is already some partial interaction $h$ that is determined before the $j$-th message $\vec{X}_j$ is chosen. To handle this issue, Håstad *et. al.* instead used the following generalized Raz's Lemma (formalized by Holenstein [Hol09]).

*Lemma 8 (Generalized Raz's Lemma [Raz98]):* Let $H, X_1, \ldots, X_k$ be independent random variables such that $X_i \equiv U_t$ are uniform[2] for every $i \in [k]$. Let $E$ be an event over $(H, X_1, \ldots, X_k)$ with $\Pr[E] \ge \delta$. Then

$$\frac{1}{k}\sum_{i=1}^{k} \mathbf{SD}((H, X_i)|_E, (H|_E, U_t)) \le \sqrt{\frac{\log(1/\delta)}{k}}.$$

We observe that, the generalized Raz's Lemma can be derandomized using an average-case version of randomness extractor, introduced by Dodis, Ostrovsky, Reyzin, and Smith [DORS08].[3] Informally, an extractor $\mathrm{Ext} : \{0,1\}^n \times [k] \to \{0,1\}^t$ is a strong average-case $(n-\Delta, \varepsilon)$-randomness extractor if for every sources $X$ with "average conditional min-entropy" $\mathrm{H}_\infty(X|H) \ge n - \Delta$ conditioned on some distribution $H$, the distribution $(I, H, \mathrm{Ext}(X, I))$ is $\varepsilon$-close to $(I, H, U_t)$. Namely, Ext can extract $t$ bits of randomness from $X$ even when $X$ only has sufficient average conditional min-entropy.

Now, let $H$ and $U_n$ be independent random variables and let $(X_1, \ldots, X_k) = (\mathrm{Ext}(U_n, 1), \ldots, \mathrm{Ext}(U_n, k))$. Let $E$ be an event over $(H, U_n)$ with $\Pr[E] \ge 2^{-\Delta}$, and let $(\tilde{H}, X) = (H, U_n)|_E$. It can be shown that $\mathrm{H}_\infty(X|\tilde{H}) \ge n - \Delta$, and

[1] The analysis presented here is slightly oversimplified and omits some technical details. Nevertheless, those technical details are irrelevant for the purpose of derandomization and are ignored from the informal discussion here.

[2] As in the basic Raz's Lemma, the generalized Raz's Lemma holds without requiring that $X_i$ being uniform. We state the lemma for uniform $X_i$'s since we only apply the lemma for this case, and it makes the connection to extractors more explicit.

[3] Dodis *et. al.* defined the notion for standard (non-strong) randomness extractor. We require the strong version definition which can be defined readily.

hence, the property of average-case extractor implies

$$\begin{aligned}
&\mathbf{SD}((I,\tilde{H},\mathrm{Ext}(X,I)),(I,\tilde{H},U_t)) \\
&= \frac{1}{k}\cdot\sum_{i=1}^{k}\mathbf{SD}((\tilde{H},\mathrm{Ext}(X,i),(\tilde{H},U_t)) \\
&= \frac{1}{k}\cdot\sum_{i=1}^{k}\mathbf{SD}((H,X_i)|_E,(H|_E,U_t)) \\
&\leq \varepsilon,
\end{aligned}$$

which is the desired conclusion we want from the generalized Raz's Lemma. Therefore, the analysis of Håstad *et. al.* can go through if we derandomize the parallel verifier using average-case extractor.

In fact, as proved by [DORS08], every ordinary randomness extractor is also an average-case extractor with a small loss in parameter. Furthermore, Vadhan [Vad11a] observed that such a parameter loss is actually not necessary. Therefore, the requirement of average-case extractor is not an extra requirement.

The second issue is about the efficiency of the rejection sampling strategy. Note that proving parallel repetition theorem for interactive arguments requires efficient reductions. Recall that upon receiving the external verifier $V$'s message $x_{j,i}$, $P^*_{rej}$ needs to sample a random continuation of $(P^{k*},V_G)$ in order to find an accepting continuation. This requires $P^*_{rej}$ to generates the remaining $k-1$ subverifiers' message, conditioned on the $i$-th verifier's message is $x_{j,i}$.

Recall that $V_G$ generates $x_{j,i}$ according to the distribution $\mathrm{Ext}(U_n,i)$. To ensure that the rejection sampling strategy can be implemented efficiently, we require the extractor to satisfy the following *invertible* property: for every output $y$ and seed $i$, one can efficiently sample a random input $x$ such that $\mathrm{Ext}(x,i)=y$. Fortunately, while not every randomness extractor satisfies the reconstructiblity property, we observe that the high min-entropy extractor constructed in Proposition 5 of Reingold *et. al.* [RVW01] is invertible and achieves very good parameters.

To summarize, we show that parallel repetition of public-coin protocols can be derandomized using randomness extractors that are strong, average-case, and invertible.

## V. Invertible Randomness Extractors

We start with the definition of standard seeded randomness extractor.

*Definition 9 (Min-entropy):* Let $X$ be a finite distribution. The **min-entropy** of $X$ is

$$\mathrm{H}_\infty(X) = -\log \max_{x\in\mathrm{supp}(X)} \mathsf{P}_X(x).$$

*Definition 10 (Strong Randomness Extractors):* A function $\mathrm{Ext}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$ is a **strong** $(k,\varepsilon)$**-extractor** if for every source $X$ over $\{0,1\}^n$ with $\mathrm{H}_\infty(X)\geq k$, the distribution $(U_d,\mathrm{Ext}(X,U_d))$ is $\varepsilon$-close to the uniform distribution $(U_d,U_m)$.

As in the above definition, when we discuss the extractors, we follow the convention in the literature, i.e., $n$ is the length of the source $X$, $k$ is the min-entropy of $X$, $d$ is the seed length, $m$ is the output length, and $\varepsilon$ is the error parameter. Furthermore, $\Delta \triangleq n-k$ is the **entropy deficiency** of $X$, and $\Lambda\triangleq k-m$ (resp., $\Lambda\triangleq k+d-m$) is the **entropy loss** of a strong (resp., non-strong) extractor. An extractor Ext is **explicit** if Ext can be computed in polynomial time.

We need explicit strong randomness extractors for *high* min-entropy source with *short* seed length and *small* entropy loss (and some additional properties we discuss later). Specifically, we think of the entropy deficiency $\Delta$ as independent of $n$, and we require the seed length to be linear in $\log\Delta$ and $\log(1/\varepsilon)$ (so that in our application, the number of repetition is $\mathrm{poly}(\Delta/\varepsilon)$).

As mentioned, we need more general "average-case" extractors, which are able to extract randomness from sources with only sufficient "(average) conditional min-entropy". The following notions are introduced in [DORS08].

*Definition 11 (Conditional Min-entropy):* Let $(H,X)$ be a finite distribution. The **(average) conditional min-entropy** of $(X|H)$ is

$$\begin{aligned}
\mathrm{H}_\infty(X|H) &= -\log\left(\mathop{\mathrm{E}}_{h\leftarrow H}\left[2^{-\mathrm{H}_\infty(X|_{H=h})}\right]\right) \\
&= -\log\left(\sum_h \mathsf{P}_H(h)\cdot\max_x \mathsf{P}_{X|H}(x|h)\right).
\end{aligned}$$

*Definition 12 (Average-case Extractor):* A function $\mathrm{Ext}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$ is a **strong average-case** $(k,\varepsilon)$**-extractor** if for every joint distribution $(H,X)$ over $\{0,1\}^*\times\{0,1\}^n$ with $\mathrm{H}_\infty(X|H)\geq k$, the distribution $(U_d,H,\mathrm{Ext}(X,U_d))$ is $\varepsilon$-close to the distribution $(U_d,H,U_m)$.

Although average-case extractors seem more general, Dodis *et. al.* [DORS08] showed that, if $\mathrm{H}_\infty(X|W)\geq k+\log(1/\varepsilon')$, then a $(k,\varepsilon)$-extractor is still able to extract the randomness from $X$, at the price of increasing the error by $\varepsilon'$. Furthermore, Vadhan [Vad11b] observed that such a $\log(1/\varepsilon')$ loss in parameter is not necessary, as stated in the following lemma.

*Lemma 13 ([Vad11b]):* If $\mathrm{Ext}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$ is a strong $(k,\varepsilon)$-randomness extractor, then Ext is a strong average-case $(k,3\varepsilon)$ extractor.

In addition, we need the extractor to have the following invertible property.

*Definition 14 (Invertibility):* An extractor $\mathrm{Ext}:\{0,1\}^n\times\{0,1\}^d\to\{0,1\}^m$ is **invertible** if there exists an efficient algorithm such that on input $y\in\{0,1\}^m$ and $r\in\{0,1\}^d$, outputs a uniformly random $x\in\{0,1\}^n$ such that $\mathrm{Ext}(x,r)=y$.

We use high-min-entropy extractors from Reingold, Vadhan, Wigderson [RVW01]. Their main extractor construction, when plugged-in the best known explicit strong randomness extractor of Guruswami, Umans, and Vadhan [GUV09] (for general source instead of high-min-entropy ones), yields the following randomness extractor.

*Lemma 15 ([RVW01], [GUV09]):* Let $\gamma > 0$ be an arbitrarily small constant. For every $\Delta, \varepsilon$ and for every sufficiently large $n \geq (1+\gamma) \cdot \Delta + \Omega(\log(1/\varepsilon))$, there exists an explicit strong[4] $(n-\Delta, \varepsilon)$-randomness extractor $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with seed length $d = O(\log \Delta + \log(1/\varepsilon))$ and entropy loss $\gamma \cdot \Delta + O(\log(1/\varepsilon))$.

Note that the above extractor achieves very good seed length and entropy loss, in compared to the information theoretic limit of $\log \Delta + 2\log(1/\varepsilon) - O(1)$ on the seed length and $2\log(1/\varepsilon) - O(1)$ on the entropy loss.

However, one issue with the this extractor is that it seems not invertible. Fortunately, we observed that a more basic version of their construction (specifically, construction stated in Proposition 6.5 of [RVW01]) does yield invertible randomness extractors with slightly worse entropy loss stated as follows.

*Lemma 16 ([RVW01], [GUV09]):* For every $\Delta, \varepsilon$ and for every sufficiently large $n \geq \Omega(\Delta + \log(1/\varepsilon))$, there exists an explicit strong $(n-\Delta, \varepsilon)$-randomness extractor $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with seed length $d = O(\log \Delta + \log(1/\varepsilon))$ and entropy loss $O(\Delta + \log(1/\varepsilon))$.

We refer the refer to the full version of this paper for a proof of invertible property of the above extractor.

### *A. Derandomized Generalized Raz's Lemma*

As mentioned, one of our main observation is that the generalized Raz's Lemma can be derandomized using an average-case randomness extractor. In this section, we formalize and prove our derandomized generalized Raz's Lemma as follows. We will use this lemma to prove our derandomized parallel repetition theorem in the next section.

*Lemma 17 (Derandomized Generalized Raz's Lemma):* Let $\mathrm{Ext} : \{0,1\}^\ell \times \{0,1\}^d \to \{0,1\}^t$ be a strong average-case $(\ell - \Delta, \varepsilon)$-randomness extractor. Let $\delta = 2^{-\Delta}$. Let $(H, X)$ be a joint distribution over $\{0,1\}^* \times \{0,1\}^\ell$ such that $H$ and $X$ are independent, and $X \equiv U_\ell$ is uniform. Let $E$ be an event over $(H, X)$ with $\Pr[E] \geq \delta$. Then,

$$\frac{1}{D} \sum_{i \in \{0,1\}^d} \mathbf{SD}((H, \mathrm{Ext}(X, i))|_E, (H|_E, U_t)) \leq \varepsilon.$$

*Proof:* Let $(\tilde{H}, \tilde{X}) \triangleq (H, X)|_E$. We claim that

$$\mathrm{H}_\infty(\tilde{X}|\tilde{H}) \geq n - \Delta,$$

and use the claim to prove the lemma first. Since $\mathrm{H}_\infty(\tilde{X}|\tilde{H}) \geq n - \Delta$, and Ext is a strong average-case $(n-\Delta, \varepsilon)$-randomness extractor, we know that $(U_d, \tilde{H}, \mathrm{Ext}(\tilde{X}, U_d))$ is $\varepsilon$-close to $(U_d, \tilde{H}, U_m)$. The lemma follows by observing that

$$\begin{aligned} &\mathbf{SD}((U_d, \tilde{H}, \mathrm{Ext}(\tilde{X}, U_d)), (U_d, \tilde{H}, U_m)) \\ &\quad = \frac{1}{D} \sum_{r \in \{0,1\}^d} \mathbf{SD}((\tilde{H}, \mathrm{Ext}(\tilde{X}, r)), (\tilde{H}, U_m)) \\ &\quad = \frac{1}{D} \sum_{r \in \{0,1\}^d} \mathbf{SD}((H, \mathrm{Ext}(X, r))|_E, (H|_E, U_m)) \end{aligned}$$

It remains to show the following claim, which can be proved by observing that $\mathsf{P}_{\tilde{X}|\tilde{H}}(x|h) \leq 2^{-n} / \Pr[E|H = h]$ and applying Bayes' rule. We omit the proof here and refer the reader to the full version of this paper for a proof of this claim. ■

## VI. Proof of Our Main Theorem

In this section, we prove Theorem 6, our derandomized parallel repetition theorems for public-coin interactive proofs and arguments. As outlined in Section IV, we derandomize the parallel verifier using randomness extractors, and we prove Theorem 6 via an efficient black-box reduction a la Håstad *et. al.* [HPWP10]. In Section VI-A, we present a formal description of our construction and prove our main lemma (see Lemma 18). Theorem 6 follows by plugging in the randomness extractor from Lemma 16 (which is invertible but with slightly worse entropy loss; for the case of interactive arguments) and 15 (which has smaller entropy looss but seems not invertible; for the case of interactive proofs) to Lemma 18.

### *A. Our Construction and Main Reduction Lemma*

Let $\mathrm{Ext} : \{0,1\}^\ell \times \{0,1\}^d \to \{0,1\}^t$ be an extractor, and $\Pi$ be a $(m, t)$-public-coin protocol. Let $D = 2^d$ and we identify $\{0,1\}^d$ with $[D]$. We define a $D$-fold derandomized parallel repetition $\Pi_{\mathrm{Ext}} = (P^D, V_{\mathrm{Ext}})$ of $\Pi$, where the derandomized verifier $V_{\mathrm{Ext}}$ is defined in Figure 1.

Let $\mathrm{Ext} : \{0,1\}^\ell \times \{0,1\}^d \to \{0,1\}^t$ be an extractor. Define a derandomized verifier $V_{\mathrm{Ext}}$:

- At each round $j \in [m]$, $V_{\mathrm{Ext}}$ samples a uniformly random $x_j \leftarrow U_\ell$, computes
$$(y_{j,1}, \ldots, y_{j,D}) = (\mathrm{Ext}(x_j, 1), \ldots, \mathrm{Ext}(x_j, D)),$$
and sends $\vec{y}_j = (y_{j,1}, \ldots, y_{j,D})$ to $P^D$.
- At the end of interaction, $V_{\mathrm{Ext}}$ accepts iff all $D$ subverifiers accept.

Figure 1. Formal description of the derandomized verifier $V_{\mathrm{Ext}}$ in $\Pi_{\mathrm{Ext}}$.

Note that this implicitly defines a derandomizer $G : \{0,1\}^{m \cdot \ell} \to \{0,1\}^{m \cdot D \cdot t}$. We note that since $V_{\mathrm{Ext}}$'s messages in different rounds are independent, $V_{\mathrm{Ext}}$ can simply

[4] [RVW01] stated their result for non-strong extractors, but as they mentioned, the strong version result follows readily by using strong extractors in their composition.

send $x_j \in \{0,1\}^\ell$ to $P^D$, who can compute $\vec{y}_j$ by himself. This makes $\Pi_{\mathrm{Ext}}$ remain public-coin and reduces the verifier communication complexity.

We shall show that if Ext is a good randomness extractor, then $G$ is a good oblivious derandomizer. We prove this by a black-box reduction, showing that there is a prover strategy $P^*$ for $\Pi$ such that given oracle access to an adversary $P^{D*}$ that convinces $V_{\mathrm{Ext}}$ with a certain probability, $P^*$ can convince $V$ with a much higher probability. Formally, we prove the following lemma.

*Lemma 18:* Let $\mathrm{Ext} : \{0,1\}^\ell \times \{0,1\}^d \rightarrow \{0,1\}^t$ be a strong, average-case $(\ell - \Delta, \varepsilon)$-randomness extractor. Let $\delta = 2^{-\Delta}$. Let $\Pi$ be a $(m,t)$-public-coin protocol and $\Pi_{\mathrm{Ext}}$ the corresponding derandomized parallel protocol. There exists a prover strategy $P^*$ such that for every $n \in \mathbb{N}$ and common input $z \in \{0,1\}^n$, and every parallel prover strategy $P^{D*}$, the following holds.

1) $\Pr[\langle P^{D*}, V_{\mathrm{Ext}} \rangle (z) = 1] \geq \delta$
$$\Rightarrow \Pr\left[\left\langle P^{*(P^{D*})}, V \right\rangle (z) = 1\right] \geq 1 - 2 \cdot m \cdot \varepsilon.$$
2) If in addition, Ext is invertible, then $P^{*(\cdot)}$ runs in time $\mathrm{poly}(n, \varepsilon^{-1}, \delta^{-1})$ given oracle access to $P^{D*}$.

We proceed to prove the lemma. We first note that we can assume without loss of generality that the parallel prover $P^{D*}$ is deterministic, since we can fix $P^{D*}$'s coins without hurting the success probability of $P^{D*}$ too much. We also note that in this case, the interaction $(P^{D*}, V_{\mathrm{Ext}})$ can be described by $V_{\mathrm{Ext}}$'s coins $(x_1, \ldots, x_m)$.

*Definition of $P^*$.:* We consider a reduction prover $P^*_{rej}$ who interacts with $V$ by simulating the interaction between $P^{D*}$ and $V_{\mathrm{Ext}}$ and uses a rejection sampling strategy of [HPWP10]. More precisely, $P^*_{rej}$ selects a uniformly random coordinate $i \in [D]$ to embed $V$ in $V_{\mathrm{Ext}}$. At each round $j \in [D]$, upon receiving the external verifier $V$'s message, $P^*$ interprets it as $V_i$'s message $y_{j,i}$, and repeatedly samples a random continuation of $(P^{D*}, V_{\mathrm{Ext}})$ until he finds an **accepting continuation**, i.e., $V_{\mathrm{Ext}}$ accepts at the end of interaction. Note that sampling a random continuation amounts to sampling a uniformly random $x_j$ conditioned on $\mathrm{Ext}(x_j, i) = y_{j,i}$, and sampling uniformly random $x_{j+1}, \ldots, x_m$. Once an accepting continuation is found, $P^*_{rej}$ chooses the corresponding $x_j$ to simulate $V_{\mathrm{Ext}}$ at this round, and forward the $i$-th coordinate of $P^{D*}$'s message to $V$. If $P^*_{rej}$ fails to find an accepting continuation in $M \triangleq 1/\delta\varepsilon$ trials, then $P^*_{rej}$ simply aborts. ■

Note that when Ext is invertible, $P^*_{rej}$ can sample a random continuation efficiently, and hence $P^*_{rej}$ runs in time $\mathrm{poly}(n, \varepsilon^{-1}, \delta^{-1})$, as asserted in Lemma 18. The proof of the first item is very similar to the proof of parallel repetition theorem for public-coin protocol in [HPWP10], with the difference that application of the generalized Raz's Lemma is replaced by a derandomized version.

To lower bound the success probability of $P^*_{rej}$, we refer to the interaction $(P^*_{rej}, V)$ as a **real experiment**, and compare it with an **ideal experiment**, defined as follows.

*Ideal Experiment $(\tilde{P}^*_{rej}, \tilde{V}^*_{rej})$.:* The ideal experiment $(\tilde{P}^*_{rej}, \tilde{V}^*_{rej})$ also simulate the interaction of $(P^{D*}, V_{\mathrm{Ext}})$ as $(P^*_{rej}, V)$, but with the following two differences. First, at each round $j$ the verifier $\tilde{V}^*_{rej}$, instead of choosing the message $y_{j,i}$ uniformly at random, chooses $y_{j,i}$ using rejection sampling as well. Namely, $\tilde{V}^*_{rej}$ repeatedly simulates a random continuation of $(P^{D*}, V_{\mathrm{Ext}})$ until an accepting continuation is found, and choose the corresponding $y_{j,i}$. Second, in the rejection sampling, instead of using bounded number of samples, they sample unbounded number of times until an accepting continuation is found. ■

Let us look at the ideal experiment closely. Let $E$ denote the event of accepting interaction, i.e. $\langle P^{D*}, V_{\mathrm{Ext}} \rangle = 1$. Note that performing rejection sampling is equivalent to selecting a random next message conditioned on $E$, and the interaction $(\tilde{P}^*_{rej}, \tilde{V}^*_{rej})$ is equivalent to choosing a uniformly random accepting interaction of $(P^{D*}, V_{\mathrm{Ext}})$. Recall that the interaction $(P^{D*}, V_{\mathrm{Ext}})$ can be described by $V_{\mathrm{Ext}}$'s randomness $(X_1, \ldots, X_m)$. The outcome of $(\tilde{P}^*_{rej}, \tilde{V}^*_{rej})$ is simply $(X_1, \ldots, X_m)|_E$. Note that $\tilde{V}^*_{rej}$ accepts iff $V_i$ of $V_{\mathrm{Ext}}$ accept, we have $\Pr[\langle \tilde{P}^*_{rej}, \tilde{V}^*_{rej} \rangle = 1] = 1$.

We next argue that the ideal and the real experiments are statistically close, which would give the desired lower bound on the success probability of $P^*$. Recall that there are the following two differences of the real experiment from the ideal one.

1) At each round $j$, $V$ chooses a uniformly random $y_{j,i}$ instead of conditioning on $E$.
2) $P^*_{rej}$ may abort when he fails to find an accepting continuation in $M$ samples.

We will bound the statistical distance incurred by these differences round by round, and combine it using the following hybrid argument. Consider the following hybrid experiments $\mathsf{H}_j$, where in the first $j$ rounds of the interaction, both parties choose messages according to the ideal experiment $(\tilde{P}^*_{rej}, \tilde{V}^*_{rej})$, and for the remaining rounds, they choose messages according to the real experiment $(P^*_{rej}, V)$. Clearly, $\mathsf{H}_0$ is the real experiment, and $\mathsf{H}_m$ is the ideal one. We will argue that the statistical distance between hybrids $\mathsf{H}_{j-1}$ and $\mathsf{H}_j$ is at most $2\varepsilon$ for every $j \in [m]$, and hence the statistical distance between the ideal and real experiments is at most $2m\varepsilon$.

Now, the only difference between $\mathsf{H}_{j-1}$ and $\mathsf{H}_j$ is at the $j$-th round, where the differences are precisely the above two items. We handle them separately, by further considering an intermediate hybrid $\mathsf{H}'_j$, which is the same as $\mathsf{H}_j$, except that at the $j$-th round, the prover chooses his message according to $P^*_{rej}$ instead of $\tilde{P}^*_{rej}$, i.e., he aborts when he fails to find an accepting continuation in $M$ samples.

We first bound the statistical distance between $\mathsf{H}_{j-1}$ and

$\mathsf{H}'_j$, where the only difference is Item (1) at the $j$-th round. We upper bound it by $\varepsilon$ using the derandomized generalized Raz's Lemma (see Lemma 17 in Section V-A) as follows.

Let $H = (X_1, \ldots, X_{j-1})$. Note that the first $j-1$ rounds interaction of both $\mathsf{H}_{j-1}$ and $\mathsf{H}'_j$ is simply $H|_E$, independent of which coordinate $i \in [D]$ played by the verifier. Hence, we can think of the coordinate $i \in [D]$ is chosen uniformly random at the beginning of the $j$-th round. At the $j$-th round, the verifier in $\mathsf{H}_{j-1}$ simply chooses a random $y_{j,i} \leftarrow U_t$, and the verifer in $\mathsf{H}'_j$ chooses $y_{j,i}$ according to $\mathrm{Ext}(X_j, i)|_{H,E}$. The statistical difference is exactly

$$\frac{1}{D} \sum_{i \in \{0,1\}^d} \mathbf{SD}((H, \mathrm{Ext}(X_j, i))|_E, (H|_E, U_t)),$$

which is upper bounded by $\varepsilon$ by Lemma 17.

We next bound the statistical distance between $\mathsf{H}'_j$ and $\mathsf{H}_j$, where the only difference is Item (2) at the $j$-th round. Note that this amounts to bound the aborting probability of $P^*_{rej}$ at the $j$-th round with the history chosen according to the ideal experiment $(X_1, \ldots, X_{j-1}, Y_{j,i})|_E$. This step is exactly the same as [HPWP10], which we repeat as follows. By Lemma 2 of [HPWP10], when the prover uses rejection sampling to find an accepting continuation, the expected number of samples needed is $1/\Pr[E]$ (averaging over the history). Hence, by a Markov inequality, the probability that $P^*_{rej}$ aborts is at most $(1/\Pr[E])/M \leq \varepsilon$. Therefore, the statistical distance between $\mathsf{H}'_j$ and $\mathsf{H}'_j$ is at most $\varepsilon$. We refer the reader to [HPWP10] for further details about computing the expectation.

To summarize, the above hybrid argument shows that the statistical distance between the ideal and the real experiments is at most $2m\varepsilon$. Since $\Pr[\langle \tilde{P}^*_{rej}, \tilde{V}^*_{rej} \rangle = 1] = 1$, it follows that

$$\Pr[\langle P^*_{rej}, V \rangle = 1] \geq 1 - 2 \cdot m \cdot \varepsilon,$$

which completes the proof of Lemma 18. ■

## VII. Lower Bounds for Public-coin Protocols

In this section, we state the following formal version of Theorem 2, which asserts the existence of a specific protocol for which every statistcally-sound derandomizer that reduces soundness error from $1/2$ to $\delta$ must use at least $m \cdot \log(1/\delta) - O(1)$-bits of extra randomness, and for which the existence of more randomness efficient computationally-sound derandomizer implies the existence of one-way functions.

*Theorem 19:* For every polynomially bounded $m : \mathbb{N} \to \mathbb{N}$, there exists a $m$-round public-coin interactive proof $\Pi$ (for the empty language $L$) with (statistical) soundness error at most $1/2$ such that the following holds. For every parameter $\delta : \mathbb{N} \to [0, 1]$, there does not exist a derandomizer $G$ for $\Pi$ that decreases the soundness error to $\delta$ and uses only

$$m \cdot (\log(1/\delta) - 3)$$

extra bits of randomness.

Additionally, the existence of a computationally-sound derandomizer that decreases the soundness error to $\delta - \mathsf{ngl}$ and uses only $m \cdot (\log(1/\delta) - 3)$ extra bits of randomness instead implies the existence of one-way functions.

Due to the space limit, we present a formal description of the protocol $\Pi$ asserted in Theorem 19 together with a brief outline of the analysis, but defer a formal proof of the theorem to the full version of this paper. A formal description of the desired protocol $\Pi$ can be found in Figure 2 with the parameter $t$ set to be $\lceil \log m \rceil + 1$.

Let $m, t$ be parameters. Define $(m, t)$**-Guess-Next-Message Protocol** $\Pi = (P, V)$:

On a common input $z \in \{0,1\}^n$

- For each round $i = 1, \ldots, m$,
  - $P$ sends a random message $y_i \in \{0,1\}^t$ to $V$.
  - $V$ sends a random message $x_i \in \{0,1\}^t$ to $P$.
- $V$ accepts iff there exists some $x_i = y_i$.

Figure 2. Formal description of an $(m, t)$-Guess-Next-Message protocol $\Pi$.

In the above protocol, the prover is allowed to guess the verifier's message $m$ times, and the prover succeeds if he ever guesses correctly. It is not hard to show that by setting $t = \lceil \log m \rceil + 1$, the soundness error is between $1/4$ and $1/2$. Theorem 19 is proved by showing that for every derandomizer $G : \{0,1\}^s \to \{0,1\}^{k \cdot (m \cdot t)}$ for $\Pi$, there exists an adversary $P^{k*}$ such that

$$\Pr[\langle P^{k*}, V_G \rangle = 1] \geq \min \left\{ m \cdot 2^{-(s/m)-1}, 1/2 \right\},$$

and showing that this adversary $P^{k*}$ can be approximated efficiently with inverse polynomially small error assuming that one-way functions do not exist.

## VIII. Lower Bounds for Private-coin Protocols

In this section, we present our impossibility result for nontrivial derandomization of the parallel repetition of private-coin protocols. We exhibit a sepcific private-coin protocol such that parallel repetition of the protocol cannot be derandomized non-trivially in the following strong sense – decreasing the randomness complexity of the parallel verifier by one would increase the soundness error by a factor of two. Formally, we state the following theorem.

*Theorem 20:* For every polynomially bounded $t : \mathbb{N} \to \mathbb{N}$, there exists a 3-message private-coin interactive proof $\Pi$ (for the empty language $L$) with randomness complexity $t$ and (statistical) soundness error $1/2$ such that the following holds. For every polynomially bounded $s, k : \mathbb{N} \to \mathbb{N}$ and every efficient $G : \{0,1\}^s \to \{0,1\}^{k \cdot t}$, the corresponding derandomized parallel protocol $\Pi_G$ has (statistical) soundness error at least

$$\varepsilon \triangleq \frac{2^{k\cdot(t-1)}}{2^s + 2^{k\cdot(t-1)}} \geq \min\left\{2^{-(s-(t-1)\cdot k+1)}, 1/2\right\}.$$

Furthermore, if $\Pi_G$ has computational soundness error less than $\varepsilon - \mathsf{ngl}$, then one-way functions exist.

Note that by padding dummy messages, Theorem 20 can be extended to $m$-message protocols with $m > 3$ readily. We mention that Theorem 3 can be derived from Theorem 20 readily as a simple corollary.

Due to the space limit, we present a formal description of the protocol $\Pi$ asserted in Theorem 20, but defer the proof of the theorem to the full version of this paper. A formal description of the desired protocol $\Pi$ can be found in Figure 2 with the parameter $\ell$ set to be $t-1$.

Let $t, \ell$ be parameters, and $\mathcal{H} = \{h : \{0,1\}^t \to \{0,1\}^\ell\}$ be a pair-wise independent hash function family. Define $(t,\ell)$**-Guess-with-Hint Protocol** $\Pi = (P, V)$:

On a common input $z \in \{0,1\}^n$

- $P$ picks a random $h \leftarrow \mathcal{H}$ and sends $h$ to $V$.
- $V$ sends $z = h(x) \in \{0,1\}^\ell$ to $P$, where $x \in \{0,1\}^t$ is the random coins of $V$.
- $P$ sends $y \in \{0,1\}^t$ to $V$.

At the end, $V$ accepts iff $x = y$.

Figure 3. Formal description of a $(t,\ell)$-Guess-with-Hint protocol $\Pi$.

## Acknowledgments

We thank Salil Vadhan for very useful discussions, and in particular, for suggesting to use the notion of averaging-case randomness extractors to simplify our original proof. We also thank Feng-Hao Liu for the collaboration in the early stage of this research, and anonymous reviewers for useful suggestions.

## References

[1] L. Babai and S. Moran, "Arthur-merlin games: A randomized proof system, and a hierarchy of complexity classes," *J. Comput. Syst. Sci.*, vol. 36, no. 2, pp. 254–276, 1988.

[2] M. Bellare, O. Goldreich, and S. Goldwasser, "Randomness in interactive proofs," *Computational Complexity*, vol. 3, no. 4, pp. 319–354, 1993.

[3] M. Bellare and J. Rompel, "Randomness-efficient oblivious sampling," in *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science*, 1994, pp. 276–287.

[4] M. Bellare, R. Impagliazzo, and M. Naor, "Does parallel repetition lower the error in computationally sound protocols?" in *FOCS*, 1997, pp. 374–383.

[5] M. Blum and S. Micali, "How to generate cryptographically strong sequences of pseudo-random bits," *SIAM J. Comput.*, vol. 13, no. 4, pp. 850–864, 1984.

[6] R. Canetti, S. Halevi, and M. Steiner, "Hardness amplification of weakly verifiable puzzles," in *TCC*, 2005, pp. 17–33.

[7] K.-M. Chung and F.-H. Liu, "Parallel repetition theorems for interactive arguments," in *TCC*, 2010, pp. 19–36.

[8] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM J. Comput.*, vol. 38, no. 1, pp. 97–139, 2008.

[9] O. Goldreich, *The Foundations of Cryptography - Volume 1*. Cambridge University Press, 2001.

[10] V. Guruswami, C. Umans, and S. P. Vadhan, "Unbalanced expanders and randomness extractors from parvaresh–vardy codes," *J. ACM*, vol. 56, no. 4, 2009.

[11] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby, "A pseudorandom generator from any one-way function," *SIAM J. Comput.*, vol. 28, no. 4, pp. 1364–1396, 1999.

[12] J. Håstad, R. Pass, D. Wikström, and K. Pietrzak, "An efficient parallel repetition theorem," in *TCC*, ser. Lecture Notes in Computer Science, D. Micciancio, Ed., vol. 5978. Springer, 2010, pp. 1–18.

[13] T. Holenstein, "Parallel repetition: Simplification and the no-signaling case," *Theory of Computing*, vol. 5, no. 1, pp. 141–172, 2009.

[14] R. Pass and M. Venkitasubramaniam, "An efficient parallel repetition theorem for arthur-merlin games," in *STOC*, 2007, pp. 420–429.

[15] R. Raz, "A parallel repetition theorem," *SIAM J. Comput.*, vol. 27, no. 3, pp. 763–803, 1998.

[16] O. Reingold, S. P. Vadhan, and A. Wigderson, "Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 8, no. 18, 2001.

[17] R. Shaltiel, "Derandomized parallel repetition theorems for free games," in *IEEE Conference on Computational Complexity*, 2010, pp. 28–37.

[18] S. Vadhan, 2011, personal communication.

[19] S. P. Vadhan, *Pseudorandomness*. Now Publishers, 2011.

[20] A. C. Yao, "Theory and applications of trapdoor functions," in *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, 1982, pp. 80–91.

[21] D. Zuckerman, "Randomness-optimal oblivious sampling," *Random Struct. Algorithms*, vol. 11, no. 4, pp. 345–367, 1997.