How to Delegate Computations: The Power of No-Signaling Proofs

Yael Tauman Kalai * Ran Raz † Ron D. Rothblum ‡

Abstract

We construct a 1-round delegation scheme (i.e., argument-system) for every language computable in time $t = t(n)$, where the running time of the prover is $\mathsf{poly}(t)$ and the running time of the verifier is $n \cdot \mathsf{polylog}(t)$. In particular, for every language in P we obtain a delegation scheme with almost linear time verification. Our construction relies on the existence of a computational sub-exponentially secure private information retrieval (PIR) scheme.

The proof exploits a curious connection between the problem of *computation delegation* and the model of *multi-prover interactive proofs that are sound against no-signaling (cheating) strategies*, a model that was studied in the context of multi-prover interactive proofs with provers that share quantum entanglement, and is motivated by the physical principle that information cannot travel faster than light.

For any language computable in time $t = t(n)$, we construct a multi-prover interactive proof (MIP) that is sound against no-signaling strategies, where the running time of the provers is $\mathsf{poly}(t)$, the number of provers is $\mathsf{polylog}(t)$, and the running time of the verifier is $n \cdot \mathsf{polylog}(t)$.

In particular, this shows that the class of languages that have polynomial-time MIPs that are sound against no-signaling strategies, is exactly EXP. Previously, this class was only known to contain PSPACE.

To convert our MIP into a 1-round delegation scheme, we use the method suggested by Aiello *et al.* (ICALP, 2000), which makes use of a PIR scheme. This method lacked a proof of security. We prove that this method is secure assuming the underlying MIP is secure against no-signaling provers.

*Microsoft Research. Email: yael@microsoft.com

†Weizmann Institute of Science, Israel and the Institute for Advanced Study, Princeton, NJ. Research supported by an Israel Science Foundation grant, by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation, and by NSF grant numbers CCF-0832797, DMS-0835373. Email: ran.raz@weizmann.ac.il

‡Weizmann Institute of Science, Israel. Parts of this research were conducted while visiting Microsoft Research. This research was partially supported by the Minerva Foundation with funds from the Federal German Ministry for Education and Research. Email: ron.rothblum@weizmann.ac.il

Contents

1 Introduction

The problem of delegating computation considers a setting where one party, the *delegator* (or *verifier*), wishes to delegate the computation of a function f to another party, the *worker* (or *prover*). The challenge is that the delegator may not trust the worker, and thus it is desirable to have the worker "prove" that the computation was done correctly. We require that verifying this proof is significantly easier than doing the computation itself; that is, the delegator's running time is significantly smaller than the time complexity of f. Moreover, we require that the running time of the worker is not much larger than the time complexity of f.

The problem of delegating computation became a central problem in cryptography, especially with the increasing popularity of cloud computing, where weak devices use cloud platforms to run their computations.

We focus on the problem of constructing *one-round* delegation protocols, where the delegator wants to verify a statement of the form $x \in \mathcal{L}$. The delegator sends x to the worker together with some query q; then the worker computes $b = \mathcal{L}(x)$, and based on the query q provides a *non-interactive* proof π for the fact that $b = \mathcal{L}(x)$. The delegator should be able to verify the correctness of the proof π very efficiently, and the worker should run in time polynomial in the time it takes to compute f. Throughout this work (similarly to all previous works that consider the problem of one-round delegation), the security requirement is against *computationally bounded* cheating workers. Namely, we consider the computational setting, where the security (i.e., soundness) of our scheme relies on a cryptographic assumption, and the guarantee is that any cheating worker, who cannot break the underlying assumption, cannot prove the correctness of an incorrect statement.

Previously, [GKR08, KR09] proved that (assuming the existence of a sub-exponentially secure computational PIR scheme) any function f that can be computed by a LOGSPACE-uniform circuit C of size $t = t(n)$ and depth $d = d(n)$, has a one-round delegation scheme where the running time of the verifier is $\tilde{O}(n + d)$, and the running time of the prover is $\mathsf{poly}(t)$.[1] Note however that for circuits with large depth d this delegation scheme does not satisfy the efficiency criterion.

A fundamental question is: Do there exist efficient 1-round delegation schemes for *all* deterministic computations? There are several works that (partially) answer this question in the preprocessing model, or under *non-falsifiable* assumptions.[2] We elaborate on these works in Section 1.4.

In this work, we answer the above question positively, by constructing a 1-round delegation scheme for *every* deterministic computation, assuming a sub-exponentially secure computational PIR scheme. More specifically, we show a delegation scheme for every lan-

[1] As is the case with all computationally sound delegation schemes, the runtime of both the prover and the verifier also grows polynomially with the security parameter. To avoid cluttering of notation, throughout this introduction, we omit this dependence on the security parameter.

[2] We note that under non-falsifiable assumptions, there are known positive results even for non-deterministic computations. The focus of this work is on deterministic computations.

guage computable in time $t = t(n)$, where the running time of the verifier is $n \cdot \mathsf{polylog}(t)$, and the running time of the prover is $\mathsf{poly}(t)$. The underlying assumption is that there exists a computational PIR scheme (or an FHE scheme) that cannot be broken in time $t^{\mathsf{polylog}(t)}$ for security parameter $k \leq \mathsf{poly}(n)$.[3]

Our delegation scheme exploits a connection to the seemingly unrelated model of multi-prover interactive proof systems (MIP) in which soundness holds even against *no-signaling* cheating provers. Loosely speaking, no-signaling provers are allowed to use arbitrary strategies (as opposed to local ones, where the reply of each prover is a function only of her own input), as long as their strategies cannot be used for communication between any two disjoint sets of provers.

We show that any MIP that is sound against no-signaling cheating provers can be converted into a 1-round delegation scheme, using a fully-homomorphic encryption scheme (FHE), or alternatively, using a computational private information retrieval (PIR) scheme. We elaborate on this connection in Section 1.2.

We then construct a new MIP, for every deterministic language, with soundness against no-signaling cheating provers. This, together with the transformation above, gives us our 1-round delegation scheme.

1.1 Multi-Prover Interactive Proofs with No-Signaling Provers

The study of MIPs that are secure against no-signaling provers was motivated by the study of MIPs with provers that share entangled quantum states. Recall that no-signaling provers are allowed to use arbitrary strategies, as long as their strategies cannot be used for communication between any two disjoint sets of provers. By the physical principle that information cannot travel faster than light, a consequence of Einstein's special relativity theory, it follows that all the strategies that can be realized by provers that share entangled quantum states are no-signaling strategies.

Moreover, the principle that information cannot travel faster than light is a central principle in physics, and is likely to remain valid in any future ultimate theory of nature, since its violation means that information could be sent from future to past. Therefore, soundness against no-signaling strategies is likely to ensure soundness against provers that obey a future ultimate theory of physics, and not only the current physical theories that we have, that are known to be incomplete.

The study of MIPs that are secure against no-signaling provers is very appealing also because no-signaling strategies have a simple mathematical characterization.

Loosely speaking, in a no-signaling strategy the answer given by each prover is allowed to depend on the queries to all other provers, as long as for any subset of provers S, and any queries given to the provers in S, the distribution of the answers given by the provers in S is independent of all the other queries. Thus, the answer of each prover can depend on the queries to all other provers as a function, but not as a random variable.

[3] In particular, for languages in P we only require a PIR scheme with quasi-polynomial security.

More formally, fix any MIP consisting of ℓ provers, and fix any set of cheating provers $\{P_1^*, \ldots, P_\ell^*\}$ who may see each other's queries (and thus each answer may depend on the queries sent to all the provers). The provers are said to be *no-signaling* if for every subset of provers $\{P_i^*\}_{i \in S}$, and for every two possible query sets $\{q_i\}_{i \in [\ell]}$ and $\{q'_i\}_{i \in [\ell]}$ such that $q_i = q'_i$ for every $i \in S$, it holds that the distributions of answers $\{a_i\}_{i \in S}$ and $\{a'_i\}_{i \in S}$ are *identical*, where $\{a_i\}_{i \in S}$ is the the answers of the provers in S corresponding to the queries $\{q_i\}_{i \in [\ell]}$, and $\{a'_i\}_{i \in S}$ is the answers of the provers in S corresponding to the queries $\{q'_i\}_{i \in [\ell]}$. If we have the slightly weaker guarantee that these two distributions are statistically close, then we say that the provers are *statistically no-signaling*. More specifically, if these two distributions are δ-close, then we say that the provers are δ-no-signaling. We refer the reader to Section 4.3 for details.

No-signaling strategies were first studied in physics in the context of Bell inequalities by Khalfin and Tsirelson [KT85] and Rastall [Ras85], and they gained much attention after they were reintroduced by Popescu and Rohrlich [PR94]. MIPs that are secure against no-signaling provers were extensively studied in the literature (see for example [Ton09, BLM+05, AII06, KKM+08, IKM09, Hol09, Ito10]). However, their precise power was unknown. It was known that they contain PSPACE [IKM09] and are contained in EXP.[4] For the case of *two* provers, Ito [Ito10] showed that the corresponding complexity class is contained in (and therefore equal to) PSPACE. Characterizing the exact power of MIPs (with more than two provers) that are secure against no-signaling provers remained an open problem.

In this work, we solve this open problem by constructing MIPs that are secure against no-signaling strategies (and more generally, statistically no-signaling strategies), for every language in EXP. Moreover, in our construction the provers are *efficient*; i.e., they run in time that is polynomial in the computation time. Specifically, for any language computable in time $t = t(n)$, we construct an MIP that is sound against no-signaling strategies, where the running time of the provers is $\mathsf{poly}(t)$, the number of provers is $\mathsf{polylog}(t)$, and the running time of the verifier is $n \cdot \mathsf{polylog}(t)$. The fact that our MIP is efficient implies that the resulting 1-round delegation scheme is efficient. We note that the previous construction of MIP that is sound against no-signaling strategies for PSPACE [IKM09] is inefficient (the provers run in time exponential in the space of the computation).

1.1.1 The Challenges in Proving Soundness Against No-Signaling Strategies

It is tempting to consider known constructions of MIPs and to try to prove their soundness against no-signaling strategies. However, known constructions of MIPs are usually for NEXP (or the scaled down version for NP). Since MIPs that are secure against no-signaling strategies are contained in EXP, there is no hope to construct such MIPs for NEXP. In particular, all known MIPs for NEXP (or the scaled down version for NP) are not sound against no-signaling strategies.

Indeed, often the trivial strategy, where the provers simply choose random answers that make the verifier accept, is no-signaling. For example, consider the trivial 2-prover interactive

[4] In a nutshell, one can find the best strategy for the provers by solving an exponential size linear program.

proof for graph 3-coloring , where the verifier sends each prover a vertex in the graph, where with probability 1/2 the vertices are the same and with probability 1/2 there is an edge between these vertices, and the provers reply with the color of these vertices. Suppose the graph is not 3-colorable. We argue that the "random accepting strategy" is a no-signaling strategy that is accepted with probability 1. More specifically, the cheating strategy is the following: If both vertices are the same, choose a random color from the set of three legal colors, and both provers send this color to the verifier. Otherwise, choose two different random colors from the set of three legal colors, and each prover sends one of these colors to the verifier. This strategy is clearly accepted with probability 1. Moreover, it is a no-signaling strategy, since the distribution of answers of each prover is uniform, independent of the query to the other prover.

This intuitive argument extends to more sophisticated MIPs and demonstrates the difficulty in proving soundness against no-signaling strategies.

1.2 From Multi-Prover Interactive Proofs to One-Round Delegation

Aiello *et al.* [ABOR00] suggested a method for converting a 1-round MIP into a 1-round delegation scheme, by using a PIR scheme (or an FHE scheme).[5] In this work, we choose to use the terminology of FHE schemes (as opposed to PIR schemes), because we find this terminology to be simpler. However, all our results hold with PIR schemes as well.

In the resulting delegation scheme, the verifier computes all the queries of the MIP verifier, and sends all these queries to the prover, each encrypted under a fresh key, using an FHE scheme. The prover then computes the MIP provers' responses homomorphically over the encrypted queries, that is, underneath the layer of the FHE scheme.

Unfortunately, shortly after this method was introduced, Dwork *et al.* [DLN+04] showed that it may, in general, be insecure. We elaborate further on the work of Dwork *et al.* and their connection to no-signaling soundness in Section 1.4.

Motivated by the work of Aiello *et al.*, Kalai and Raz [KR09] showed that a variant of this method can be used to securely convert any interactive proof into a one-round argument system. The idea is simply to have the verifier send all its (say t) messages in the first round, in the following redundant form: For every $i \in [t]$, all the first i messages are encrypted using a fresh FHE key.[6] The work of [KR09], together with the interactive delegation scheme of Goldwasser *et al.* [GKR08], gives rise to the 1-round delegation protocol for LOGSPACE-uniform low-depth circuits, mentioned above.

We show that the method of Aiello *et al.* [ABOR00] is secure if the underlying MIP is sound against statistically no-signaling strategies. Thus, we reduce the cryptographic

[5] Actually, [ABOR00] suggested to use a PCP. However, as pointed out by [DLN+04] an MIP is more suitable.

[6] The reason the i'th message is encrypted together with the preceding messages, is since the prover's reply may depend on all these messages.

problem of constructing secure one-round delegation schemes, to the information theoretical problem of constructing MIP schemes that are secure against statistically no-signaling provers. Such a reduction allows us to "strip off" the cryptography, and to focus on an information theoretic question of constructing an MIP that is secure against statistically no-signaling provers.

This result generalizes the work of [KR09], since any interactive proof can be seen as an MIP where the verifier sends his first i messages to prover i (it is quite easy to verify that the resulting MIP is secure against statistically no-signaling cheating provers). Moreover, our result significantly simplifies the one of [KR09], which implicitly converts the interactive proof into an MIP scheme and then applies the PIR to the resulting MIP scheme. We believe that due to the lack of the "correct" terminology, the result of [KR09] was relatively complicated, whereas this current result is significantly simpler and more general. We refer the reader to Section 16 for details.

1.3 Summary of Our Results

We show that when applying the method of Aiello *et al.* [ABOR00] to an MIP that is sound against statistically no-signaling cheating provers, then the resulting 1-round delegation protocol is secure (assuming that the underlying FHE is secure against attackers of sub-exponential size).

Informal Theorem 1 (See Theorem 12)**.** *Assuming the existence of an* FHE *scheme with sub-exponential security, there exists an efficient way to convert any* 1*-round* MIP *that is sound against statistically no-signaling cheating provers into a secure 1-round delegation scheme, where the running time of the prover and verifier in the delegation scheme are proportional to the running time of the provers and verifier in the* MIP*.*

Remark. More specifically, the precise assumption needed in Informal Theorem 1 is that there exists an FHE scheme that, for security parameter $k \leq \mathsf{poly}(n)$, is secure against adversaries running in time $2^{O(|a_1|+\ldots+|a_\ell|)}$, where $|a_i|$ is the answer size of the i'th prover in the underlying MIP scheme.

Thus, we reduced the cryptographic problem of constructing secure one-round delegation schemes, to the information theoretical problem of constructing MIP schemes that are secure against statistically no-signaling provers.

We then construct an efficient MIP, that is sound against statistically no-signaling strategies, for every language in EXP.

Informal Theorem 2 (See Theorem 4)**.** *For any language L computable in time $t = t(n)$, there exists an* MIP *that is secure against statistically no-signaling adversaries. The (honest) provers in this* MIP *run in time* $\mathsf{poly}(t)$*, the number of provers and the communication complexity is* $\mathsf{polylog}(t)$*, and the verifier runs in time* $n \cdot \mathsf{polylog}(t)$*.*

We note that our MIP has the additional property that the verifier does not need to know the entire input, but rather only needs to access a few points in the low-degree extension of the input (we refer the reader to Section 4.6 for the definition of low-degree extension). This property, which was also a property of the [GKR08] protocol, is important for applications such as memory delegation [CKLR11].

The above theorem, together with Informal Theorem 1, immediately yields the following corollary:

Informal Theorem 3 (See Theorems 9-11)**.** *Assume the existence of an* FHE *scheme with sub-exponential security. Then, there exists a* 1*-round delegation scheme for any function computable in time* $t = t(n)$*. The prover in this delegation scheme runs in time* $\mathsf{poly}(t)$*, the verifier runs in time* $n \cdot \mathsf{polylog}(t)$*, and the communication complexity is* $\mathsf{polylog}(t)$*.*

Remark. As in Informal Theorem 1, the precise assumption needed for the above theorem is the existence of an FHE scheme that, for security parameter $k \leq \mathsf{poly}(n)$, is secure against adversaries running in time $2^{\mathsf{polylog}(t)}$.

As a special case, Informal Theorem 2 gives soundness against provers that share an entangled quantum state, since such provers are no-signaling. This gives a scheme for delegating computation to a group of workers that cannot communicate with each other (where the parameters are as in Theorem 2). The scheme is information theoretically secure even if the workers share an entangled quantum state. Moreover, the scheme remains secure in any future ultimate theory (that may extend quantum theory) as long as the no-signaling principle remains valid. We note, however, that recent breakthroughs by Ito and Vidick construct MIPs that are secure against provers that share entangled quantum states, for any language in NEXP [IV12, Vid13].

The bulk of technical contribution of this work is in proving Informal Theorem 2. As noted above, proving this theorem requires overcoming several technical hurdles that do not appear in the classical MIP (or PCP) setting. We refer the reader to Section 3 for an overview of our techniques for proving this theorem.

Informal Theorem 1 is mainly a conceptual contribution. Its proof is relatively straightforward, but we find the connection between the seemingly unrelated concepts of delegation and no-signaling soundness to be intriguing.

1.4 Related Work

Our work is greatly inspired by the work of Aiello *et al.* [ABOR00], who propose a general methodology of constructing 1-round delegation schemes, by combining an MIP (or a PCP) with a (computational) PIR scheme. Also very relevant to our work is the work of Dwork *et al.* [DLN+04], who proved that this method is not sound, by giving an example of a PCP for which the resulting one-round delegation scheme is not sound, no matter which PIR scheme (or FHE scheme) is used.

Moreover, [DLN+04] define the notion of a "spooky interaction" which is a behavior of the cheating prover, that on the one hand does not directly contradict the security of the PIR, yet on the other hand is not consistent with answers based on PIR databases. Using our terminology, a spooky behavior is exactly a no-signaling distribution on prover answers that are computed "homomorphically" under the "encrypted" PIR queries.

More importantly, Dwork *et al.* also argue that the soundness of the [ABOR00] technique cannot essentially be based on any MIP (or PCP). However, Dwork *et al.* (and [ABOR00]) were focused on constructing 1-round delegation schemes for non-deterministic languages (such as languages in NEXP or the scaled down version of NP). Indeed, it is implicitly shown in [DLN+04] that languages that can be proved by an MIP with soundness against no-signaling provers are in EXP (and the scaled down version of it is contained in P). Additionally, Gentry and Wichs [GW11] recently showed a negative result, proving that there does not exist a non-interactive delegation scheme for NP with a black-box proof of security under any falsifiable assumption.[7] However, these negative results do not apply to our setting as our delegation scheme is not for all of NP, but rather for languages in P (or, in the scaled up version, in EXP).

Thus, by focusing on deterministic classes (as opposed to non-deterministic ones), we manage to show that the [ABOR00] method is indeed sound in some cases.

Related work on computation delegation. Beyond the works of [GKR08, KR09] that were mentioned earlier, there are many other works on delegating computation that are less relevant to this work. Let us mention a few. In the *interactive* setting, Kilian [Kil92] constructed a 4-message delegation scheme for every function in NEXP. Micali [Mic94] showed that in the so called *random oracle model* this result can be made non-interactive, by relying on the Fiat-Shamir paradigm [FS86]. There are also several results that construct non-interactive delegation schemes under *non-falsifiable* assumptions (as defined by Naor [Nao03]). These works include [Gro10, Lip12, BCCT12a, DFH12, GLR11, BCCT12b, GGPR12] and more. Finally, we mention a series of results that construct non-interactive delegation scheme in the *preprocessing model*, where the verifier is efficient only in the amortized setting. These results include [GGP10, CKV10, AIK10, PRV12]. There are many other results that we do not mention, which consider various different models, or are concerned with practical efficiency.

1.5 Organization

In Section 2, we formally state our results. In Section 3, we provide a high-level overview of our techniques. In Section 4, we formally define the notions that we use throughout this work. In Sections 5 to 8, we construct a *base* PCP with soundness against no-signaling strategies for PSPACE. In Sections 9 to 11, we construct the *augmented* PCP for EXP. In Sections 12 to 14, we show how to transform this PCP into an MIP. In Section 15, we use

[7] The model of [GW11] differs from our model in that they allow the prover the additional power of choosing the instance x after seeing the first message sent by the verifier.

the tools from all previous sections to prove the main information theoretic result. Finally, in Section 16 and Section 17, we show how to transform our MIP into an 1-round delegation scheme.

2 Our Results

We show a general result on MIP proof systems that are secure against no-signaling strategies and show how to use the latter to construct a new 1-round delegation scheme (a.k.a. 1-round argument-system).

Theorem 4. *Suppose that* $\mathcal{L} \in \mathsf{DTIME}(t(n))$*, where* $t = t(n)$ *satisfies* $\mathsf{poly}(n) \leq t \leq \mathsf{exp}(n)$*. Then, for any integer* $(\log t)^c \leq k \leq \mathsf{poly}(n)$*, where* c *is some (sufficiently large) universal constant, there exists an* MIP *for* $\mathcal{L}$ *with* $k \cdot \mathsf{polylog}(t)$ *provers and with soundness error* 2^{-k} *against* $2^{-k \cdot \mathsf{polylog}(t)}$*-no-signaling strategies.*

The verifier runs in time $(n + k^2) \cdot \mathsf{polylog}(t)$ *and the provers run in time* $\mathsf{poly}(t, k)$*. Each query and answer is of length* $k \cdot \mathsf{polylog}(t)$*.*

By setting the parameters $t = \mathsf{poly}(n)$ and $k = \mathsf{polylog}(n)$ we obtain the following corollary:

Corollary 5. *If* $\mathcal{L} \in \mathsf{P}$*, then there exists an* MIP *for* $\mathcal{L}$ *with* $\mathsf{polylog}(n)$ *provers, and with soundness error* $2^{-\mathsf{polylog}(n)}$ *against* $2^{-\mathsf{polylog}(n)}$*-no-signaling strategies. The verifier runs in time* $\tilde{O}(n)$ *and the provers run in time* $\mathsf{poly}(n)$*. Each query and answer is of length* $\mathsf{polylog}(n)$*.*

By setting $t = \mathsf{poly}(n)$ and $k = \sqrt{n}$ we obtain the following corollary:

Corollary 6. *If* $\mathcal{L} \in \mathsf{P}$*, then there exists an* MIP *for* $\mathcal{L}$ *with* $\tilde{O}(\sqrt{n})$ *provers, and with soundness error* $2^{-\sqrt{n}}$ *against* $2^{-\tilde{\Omega}(\sqrt{n})}$*-no-signaling strategies. The verifier runs in time* $\tilde{O}(n)$ *and the provers run in time* $\mathsf{poly}(n)$*. Each query and answer is of length* $\tilde{O}(\sqrt{n})$*.*

A scaled up result is obtained by setting $t = \mathsf{exp}(n)$ and $k = \mathsf{poly}(n)$:

Corollary 7. *If* $\mathcal{L} \in \mathsf{EXP}$*, then there exists an* MIP *for* $\mathcal{L}$ *with* $\mathsf{poly}(n)$ *provers and with soundness error* $2^{-\mathsf{poly}(n)}$ *against* $2^{-\mathsf{poly}(n)}$*-no-signaling strategies. The verifier runs in time* $\mathsf{poly}(n)$ *and the provers run in time* $\mathsf{exp}(n)$*. Each query and answer is of length* $\mathsf{poly}(n)$*.*

Having stated our main information-theoretic results, we proceed to state our main cryptographic results. The following theorems rely on the existence of an (S, δ)-secure FHE scheme, which is an FHE scheme where any $\mathsf{poly}(S)$-size adversary cannot distinguish between an encryption of any two messages with probability greater than δ (see Section 4.7 for a formal definition).[8]

[8] Alternatively, we can rely on the existence of a sufficiently strong cryptographic private information retrieval scheme (PIR), see remark at the end of Section 17.

We first state our general transformation from any MIP that has soundness against no-signaling strategies into a 1-round argument-system.

Theorem 8 (Simplified; for the full statement see Theorem 12)**.** *Suppose that the language $\mathcal{L}$ has an MIP with ϵ soundness against δ-no-signaling strategies and a total of λ communication (to all provers). Let $\tau = \tau(n) \geq \lambda$ be a security parameter, where n denotes the input length of the MIP. For every $S = S(\tau) \geq \tau$ such that $S \geq \max(n, 2^{\lambda})$ and $\delta' = \delta'(\tau)$ such that $\delta' \leq \delta/\lambda$, if there exists an (S, δ') secure FHE, then the language $\mathcal{L}$ has a 1-round argument system with soundness (S, ϵ).*

If the MIP verifier runs in time T_V, then the running time of the resulting verifier is $T_V + \mathsf{poly}(\tau)$. If the running time of each MIP prover is T_P, then the running time of the resulting prover is $\mathsf{poly}(T_P, \tau, n)$. The total communication in the resulting argument-system is of length $\mathsf{poly}(\tau)$.

By combining Theorem 4 with Theorem 8 we obtain the following argument-system:

Theorem 9. *Suppose that $\mathcal{L} \in \mathsf{DTIME}(t(n))$, where $t = t(n)$ satisfies $\mathsf{poly}(n) \leq t \leq \mathsf{exp}(n)$. Let $\tau = \tau(n)$ be a security parameter such that $\log(t) \leq \tau \leq \mathsf{poly}(t)$. Let $S = S(\tau) \geq \tau$ such that $2^{(\log(t))^c} \leq S \leq 2^{\mathsf{poly}(n)}$ and $S \leq 2^{\max(n,\tau)}$, where c is some sufficiently large universal constant. If there exists an $\left(S, 2^{-\sqrt{\log S}}\right)$-secure FHE, then $\mathcal{L}$ has a 1-round argument system with soundness $\left(S, 2^{-\frac{\sqrt{\log S}}{\mathsf{polylog}(t)}}\right)$. The verifier runs in time $n \cdot \mathsf{polylog}(t) + \mathsf{poly}(\tau)$ and the prover runs in time $\mathsf{poly}(t)$. The total communication is of length $\mathsf{poly}(\tau)$.*

We stress that the running time of the verifier in Theorem 9 only depends *poly-logarithmically* on the time that it takes to compute $\mathcal{L}$. We proceed to describe two useful corollaries of Theorem 9.

By setting $t = \mathsf{poly}(n)$, $\tau = n^{\epsilon}$ and $S(\tau) = 2^{(\log(\tau))^c}$ where $\epsilon > 0$ (resp., $c > 0$) is a sufficiently small (resp., large) universal constant and assuming the existence of a quasi-polynomially secure FHE, we obtain a (cryptographic) delegation scheme for P with quasi-linear verification and sublinear communication.

Theorem 10. *Suppose that $\mathcal{L} \in \mathsf{P}$. If there exists a $\left(2^{\mathsf{polylog}(\tau)}, 2^{-\mathsf{polylog}(\tau)}\right)$-secure FHE, then, for every constant $\alpha > 0$, the language $\mathcal{L}$ has a 1-round argument system with soundness $\left(2^{\mathsf{polylog}(n)}, 2^{-\mathsf{polylog}(n)}\right)$. The verifier runs in time $\tilde{O}(n)$ and the prover runs in time $\mathsf{poly}(n)$. The total communication is of length $O(n^{\alpha})$.*

By setting $t = \mathsf{poly}(n)$, $\tau = (\log(n))^c$ and $S(\tau) = 2^{\tau^{\epsilon}}$ where $\epsilon > 0$ is a sufficiently small universal constant, $c > 0$ is a sufficiently large universal constant (that is chosen after ϵ) and assuming the existence of a *sub-exponentially* secure FHE (a stronger assumption than that in Theorem 10) we obtain a (cryptographic) delegation scheme for P with quasi-linear verification but only *poly-logarithmic* communication.

Theorem 11. *Suppose that $\mathcal{L} \in \mathsf{P}$. If there exists a $\left(2^{\tau^{\epsilon}}, 2^{-\tau^{\epsilon/2}}\right)$-secure FHE, where $\epsilon > 0$ is a sufficiently small universal constant, then $\mathcal{L}$ has a 1-round argument system with soundness*

$\left(2^{\mathsf{polylog}(n)}, 2^{-\mathsf{polylog}(n)}\right)$. *The verifier runs in time* $\tilde{O}(n)$ *and the prover runs in time* $\mathsf{poly}(n)$. *The total communication is of length* $\mathsf{polylog}(n)$.

3 Our Techniques

Our techniques can be separated into two parts. The main technical contribution of this work is the construction of an MIP that is sound against statistically no-signaling cheating provers, for any function computable in time t. The number of provers is $\mathsf{polylog}(t)$, each prover runs in time at most $\mathsf{poly}(t)$, and the verifier runs in time $n \cdot \mathsf{polylog}(t)$. This construction, described in Section 3.1, is information theoretic, and does not rely on any cryptographic assumptions.

Then, in Section 3.2 we show how to convert a statistically no-signaling MIP into a one-round argument (while preserving the parameters, up to polynomial factors). The soundness of the resulting one-round argument assumes the existence of a fully homomorphic encryption (FHE) scheme with sub-exponential security.

3.1 Our Statistically No-Signaling MIP

We start by giving an overview of our MIP, and then give the high-level idea for why soundness holds against statistically no-signaling cheating provers. The proof of soundness requires a different approach than the ones taken to prove classical soundness. Indeed, all known MIP's for NEXP (or the scaled down version of NP) are not sound against no-signaling adversaries (see discussion in Section 1.1.1).

The main difference between a classical MIP and a no-signaling MIP is that in a classical MIP once a prover fixes his random tape (if at all he uses randomness), then his answer is a deterministic function of his query. This is not the case in the no-signaling setting, since a prover's answer can depend on the other queries. It is required that the answer of the prover is independent of the other queries *as a random variable* , but it may certainly depend on the other queries as a function. This makes the soundness proof significantly more challenging.

Before presenting the high level ideas of this proof, we first give a high level overview of our MIP.

As a first step in the construction of our MIP, we would like to assume for simplicity that any set of (possibly malicious) provers behave *symmetrically*; namely, any two subsets of provers, who are asked the same questions, answer similarly. Of course, we cannot ensure such a thing, since cheating provers may behave arbitrarily. Instead, this is ensured by defining a new model of no-signaling PCP, as oppose to no-signaling MIP.

Intuitively, a no-signaling PCP is defined like a classical PCP, but where soundness is required to hold also against a *no-signaling* prover, who can see all queries. Loosely speaking, a no-signaling prover, upon receiving any set of queries Q, may reply with answers, where each answer may depend on all the queries in Q *as a function*, but not as a random variable.

Namely, for any set of queries Q and for any subset $Q' \subseteq Q$, the *distribution* of the answers corresponding to the queries Q', should be independent of queries in $Q \setminus Q'$.

Formally, a no-signaling prover consists of a family of distributions $\{\mathcal{A}_Q\}$, where there is one distribution for every "sufficiently small" set of queries Q, and the requirement is that for every subset of queries $Q' \subseteq Q$, the distribution $(\mathcal{A}_Q)|_{Q'}$ (which is the distribution of answers $\mathcal{A}_Q$ restricted to queries in Q') is independent of queries in $Q \setminus Q'$. More generally, a δ-no-signaling family of distributions has the property that for every three sets of queries Q_1, Q_2, Q', such that $Q' \subseteq Q_1$ and $Q' \subseteq Q_2$, the distributions $(\mathcal{A}_{Q_1})|_{Q'}$ and $(\mathcal{A}_{Q_2})|_{Q'}$ are δ-close. We emphasize that in a δ-no-signaling PCP we think of a set of queries Q as an *unordered* set, thus achieving the desired *symmetry*; i.e., the answers do not depend on the order of the queries.

We note, however, that the definition of δ-no-signaling PCP given above, is not complete. One needs to define what is a "sufficiently small set of queries". We define it to be all the query sets with at most $k_{\max}$ queries. k_{max} is an important parameter. The larger k_{max} is, the more limited the cheating provers are.[9] We denote such a PCP by (k_{max}, δ) no-signaling PCP, and define it formally in Section 4.5. We devote most of the technical sections to constructing a (k_{max}, δ)-no-signaling PCP and proving its soundness.

Converting this PCP into a δ-no-signaling MIP is relatively straightforward. The basic idea is that the MIP verifier emulates the PCP verifier, and sends each query to a random prover (that was not yet asked any query). Each prover answers by simulating the (honest) PCP. The parameter k_{max} corresponds to the number of provers in the resulting MIP. In this work, $k_{max} = \mathsf{polylog}(t)$, and thus the number of provers in our MIP is $\mathsf{polylog}(t)$, and the verifier in our one-round argument runs in time $n \cdot \mathsf{polylog}(t)$.

Overview of our underlying PCP. We present our PCP in two steps. First, we construct a "base" PCP for languages in PSPACE. Then we show how to augment this PCP, and construct a PCP for EXP. We prove that both PCPs are sound against statistically no-signalling strategies.

3.1.1 Our Base PCP.

Let $\mathcal{L}$ be a language computable by a (deterministic) Turing machine running in time $t(n)$ and space $s(n)$ on instances of length n. Our base PCP for $\mathcal{L}$ has $k_{max} = \tilde{O}(s(n))$. This PCP is similar to known PCPs (in particular, to the PCP of [Sud00]). The main points of distinction are that in our base PCP each test is repeated k times, where k is a security parameter, and that our PCP is applied for deterministic computations, rather than for non-deterministic computations.

Suppose that the prover needs to prove that $x \in \mathcal{L}$, where x is an instance of length n. The underlying PCP consists of several low degree multi-variate polynomials. The first polynomial is the low-degree extension (defined in Section 4.6) of the entire computation.

[9] Jumping ahead, we note that in this work $k_{max} = \mathsf{polylog}(t)$.

More specifically, let $\mathcal{C}_n$ be a circuit of size $N = O(t(n)s(n))$ that computes $\mathcal{L}$ on inputs of length n. It is known that the circuit $\mathcal{C}_n$ can be made layered, with $O(s(n))$ gates in each layer, and $O(t(n))$ layers.

Assume that the wires of the circuit are indexed by the numbers $1, \ldots, N$, in an order that agrees with the layers of the circuit. In particular, the indexes of wires at layer i are larger than the indexes of wires at layer $i-1$. We assume that $1, \ldots, n$ are the indexes of the n input variables and that N is the index of the output wire. Let $x_1, \ldots, x_N$ be the values of the N wires of the circuit $\mathcal{C}_n$ when computed with input $x = (x_1, \ldots, x_n)$.

The entire computation $x_1, \ldots, x_N$ appears in the PCP encoded using an error correcting code (specifically, using the low-degree extension encoding), so that if a single bit in the computation is incorrect it causes a global affect on the encoding.

In addition, the PCP contains several other low-degree multi-variate polynomials, denoted by $P_0, P_1, \ldots, P_\ell$, which are defined in Section 5. In this overview we ignore these polynomials.

The analysis of our base PCP. The analysis of our base PCP begins with an *error amplification* step, where (loosely speaking) we prove that if there exists a (statistically) no-signaling prover (which is a family of distributions, one for every possible set of queries), that convinces the PCP verifier to accept a statement of the form $\mathcal{C}_n(x) = b$ with some non-negligible probability, then there exists a (statistically) no-signaling prover that convinces a different verifier, called, the *relaxed verifier*, to accept the same statement with probability close to 1 (i.e., with probability $1 - \frac{1}{\mathsf{poly}(t)}$ for any polynomial poly).

This error amplification step, which is a crucial step in our proof, is achieved as follows: Recall that the verifier V repeats each test k times, and accepts if and only if all tests accept. We define a "relaxed" verifier V' that makes the exact same queries as V, but accepts if and only if for each (repeated) test, at least r of the k repetitions are accepting, where r is a parameter. Loosely speaking, we prove that if the verifier V accepts with probability ϵ then the relaxed verifier accepts with probability $1 - \frac{\tilde{O}(2^{-r})}{\epsilon}$, where $\tilde{O}$ hides $\mathsf{polylog}(N)$ factors.

To prove this we argue that if V and V' choose their queries independently then the probability that V accepts and V' rejects is very small. This is true because for each group of k tests we can first choose the $2k$ tests for both V and V', and only then decide which tests go to V and which ones go to V'. Consider the answers for these $2k$ tests. (It is important here that k_{max} is greater than the total number of queries in these $2k$ tests, so that all these queries can be asked simultaneously.) If among the $2k$ tests many are rejected then V rejects with high probability. On the other hand, if among the $2k$ tests only few are rejected (say, less than r) then V' always accepts. We refer the reader to Section 6 for details.

In this overview, we ignore the fact that the relaxed verifier is different than the actual verifier, and assume for simplicity that there is a (statistically) no-signaling prover that convinces the actual verifier to accept with probability $1 - \frac{1}{\mathsf{poly}(t)}$. We will prove that in that case the statement $\mathcal{C}_n(x) = b$ must be correct.

To this end, we first prove that for every (statistically) no-signaling prover, if the PCP

verifier accepts with probability $1 - \frac{1}{\mathsf{poly}(t)}$, then it must be the case that the distributions corresponding to queries in $\{x_1, \ldots, x_N\}$ are *locally consistent*. More specifically, we prove that for every gate in the circuit $\mathcal{C}_n$, and for every set of queries that include the two input wires and the output wire of the gate, the answers of values of the inputs and output wires are consistent with the gate, with very high probability (say, higher than $1 - \frac{1}{t^3}$). We note that this guarantee only requires $k_{max} = \mathsf{polylog}(t)$, and in particular the dependence on the space s is not needed to obtain this local consistency guarantee.[10]

We note that the local consistency guarantee is only true if variables in $\{x_1, \ldots, x_N\}$ are read in a certain way, which uses interpolation and the local decoding properties of the low-degree extension encoding. In this overview, for simplicity, we completely ignore this extra complication and assume that the local consistency guarantee holds as stated above.

From a classical perspective, it seems that the local consistency guarantee should immediately imply global consistency, and thus correctness, by applying a straightforward union bound. However, in the no-signaling setting, this intuition is misleading. The reason is that in order to apply the union bound we need to consider the probability that *all* the local consistency conditions are met *simultaneously*, and make the following argument:

$$
\begin{aligned}
&\Pr[\text{correctness}] \geq \\
&\Pr[\text{all local consistency conditions hold}] = \\
&1 - \Pr[\exists \text{ local consistency condition that does not hold}] \geq \\
&1 - O(t \cdot s) \cdot \Pr[\text{a single local consistency does not hold}] \geq \\
&1 - O(t \cdot s) \cdot \frac{1}{t^3} \geq \\
&1 - O\left(\frac{1}{t}\right).
\end{aligned}
$$

Unfortunately, in the no-signaling setting, this type of calculation is not correct, since it is not clear what it means for *all* the local consistency conditions to hold simultaneously. Recall that there is no PCP in the sky, but rather a set of distributions for each set of queries of size at most k_{max}. Thus, we can check whether at most k_{max} local consistency conditions hold simultaneously, but not more than that, as the relevant random variables are not even defined simultaneously.

We can still use the local consistency guarantee to argue that up to a small probability of error, the probability that the output is correct is at least the probability that *both* children of the output gate are correct (using the local consistency condition). We can then proceed by induction, towards the base of the tableau. However, this will incur an exponential (in t) blowup in the error.

Generally, we cannot afford this exponential blowup in error. Jumping ahead, we note that curiously, in one of the lemmas for our augmented PCP, we do use this analysis (and

[10] Jumping ahead, we note that in the base PCP, $k_{max} = \tilde{\Theta}(s)$ is required to go from local consistency to global consistency.

guarantee) for a specific computation for which the depth of the tableau is relatively small ($O(\log s)$). We elaborate on this point below.

In the analysis of our base PCP, we solve this problem by taking $k_{max} = \tilde{\Theta}(s)$, which enables us to check the correctness of an entire layer of the tableau simultaneously. More specifically, we first check the correctness of the input. The local consistency condition implies that this check passes with probability $1 - \frac{1}{\mathsf{poly}(t)}$. Then we check the first two level simultaneously. This could be done since $k_{max} = \tilde{\Theta}(s)$. The local consistency, together with the correctness of the first level, implies that the second level is correct with probability at least $\left(1 - \frac{1}{\mathsf{poly}(t)}\right)^2$. Then, we check consistency of the second and third levels, and deduce that the third level is correct with probability $\left(1 - \frac{1}{\mathsf{poly}(t)}\right)^3$. This argument continues by induction until the top layer is reached, and the conclusion is that the computation is correct with probability $\left(1 - \frac{1}{\mathsf{poly}(t)}\right)^t$, which is very close to 1, as desired.

This idea indeed works, however, it results with an MIP with $\tilde{\Theta}(s)$ provers, and thus with a one-round argument where the running time of the verifier grows linearly with s. We refer the reader to Section 7 for the formal analysis.

Our goal is to make k_{max} independent of s, and thus eliminate the dependency on the space. Indeed, we manage to "augment" this base PCP, and to prove that our augmented PCP is secure against statistically no-signaling distributions with $k_{max} = \mathsf{polylog}(t)$. This gives rise to an MIP where the verifier runs in time $n \cdot \mathsf{polylog}(t)$, and where the provers run in time $\mathsf{poly}(t)$. This, in turn, gives rise to our one-round delegation scheme, that achieves similar parameters.

3.1.2 Our Augmented PCP.

Recall that our base PCP is a proof that the computation of $\mathcal{C}_n$ was performed correctly, where $\mathcal{C}_n$ is a layered circuit of size $N = O(t(n)s(n))$ (consisting of $O(t(n))$ layers each of size $O(s(n))$), that computes $\mathcal{L}$ on inputs of length n.

The basic idea behind our *augmented* PCP is to run the same base PCP on an *augmented* circuit, denoted by $\mathcal{C}'_n$. Loosely speaking, the circuit $\mathcal{C}'_n$ computes the same function as $\mathcal{C}_n$, but in $\mathcal{C}'_n$ each layer of the circuit is augmented with the low-degree extension of the layer, and with all the low-degree tests corresponding to lines of the low-degree extension. Namely, the circuit $\mathcal{C}'_n$ is the same as $\mathcal{C}_n$, but where after each layer we insert another circuit, denoted by $\mathcal{C}_{\mathsf{LDE}}$, which takes as input the entire layer, and computes the low-degree extension of the layer, and performs *all* the low-degree tests; i.e., for every line in the low-degree extension it checks that the values on that line correspond to a low-degree univariate polynomial.[11] It is known that $\mathcal{C}_{\mathsf{LDE}}$ can be made a circuit of size $\mathsf{poly}(s)$ and depth $O(\log s)$. We refer the reader to Section 9 for a formal description of our augmented PCP.

[11] The low-degree tests are seemingly redundant as the values of the low-degree extension were computed by the circuit. However, since we don't know that the values are computed correctly, the low-degree tests will be very important in our analysis.

The basic idea behind adding the computation of $\mathcal{C}_{\mathsf{LDE}}$ after each layer, is that now the PCP verifier can read a single point in the low-degree extension of a layer, and in some sense, this point contains information about the entire layer. As we argue below, if a random point in the low-degree extension is correct with high probability, then *each* value in the layer is correct with almost the same probability. We elaborate below.

Our analysis. Since our augmented PCP is identical to our base PCP, applied to the augmented circuit $\mathcal{C}'_n$ (as opposed to $\mathcal{C}_n$), the analysis of our base PCP implies that if there exists a (statistically) no-signaling prover that convinces the verifier to accept with probability close to 1, then local consistency holds with probability close to 1. Namely, for any set of queries Q of size at most $k_{max} = \mathsf{polylog}(t)$ and for any subset $Q' \subseteq Q$ of queries corresponding to variables in the tableau of $\mathcal{C}'_n$, the answers to these queries are locally consistent (i.e., they satisfy the constraints imposed by the gates and they satisfy the low-degree tests) with probability close to 1.

We next show how we go from local consistency to global consistency, without increasing the size of k_{max}. To this end, we use the special structure of $\mathcal{C}'_n$; i.e., the fact that it includes all the low-degree extensions and low degree tests.

Fix any layer in $\mathcal{C}_n$. We first argue that if a random point in the low-degree extension of this layer has the correct value with high probability, then the value of every point in the layer is correct with high probability. The idea is the following: Fix any point z in the layer. Consider the line connecting this point to the random point in the low-degree extension. The local consistency condition implies that with high probability the values on this line correspond to a low degree polynomial. Thus, if the value of the point z is incorrect (with significant probability) then most of the points on the line are incorrect (with significant probability), and in particular a random point on the line is incorrect (with significant probability), contradicting our assumption that a random point in the low-degree extension is correct with high probability.

We use the argument above to prove the correctness of the entire computation. The proof is by induction on the depth of the tableau of $\mathcal{C}_n$. In what follows, we denote the probability that local consistency holds by $(1-\epsilon)$, and the reader should think of $\epsilon = \frac{1}{\mathsf{poly}(t)}$. We start with the base case, and claim that each element in the base of the tableau (where the input lies) is correct with probability $1-\epsilon$. This follows from the local consistency condition. Next we claim that if each element in a layer is correct with high probability $(1-\epsilon)$, then any point in the low-degree extension of the layer is correct with probability $(1-\epsilon)^{\mathsf{poly}(s)}$. To this end, we use the analysis where the error increases exponentially with the depth, and we use the fact that the depth of $\mathcal{C}_{\mathsf{LDE}}$ is $O(\log s)$.

For the induction step, we claim that if at layer i a random point in the low-degree extension is correct with some probability p, then a random point in the low-degree extension of layer $i+1$ is correct with probability $\approx p(1-\epsilon)^{\mathsf{poly}(s)}$. Note that this guarantee is strong enough for correctness, since by induction we get that a random point in the low-degree extension of the top layer is correct with probability $\approx (1-\epsilon)^{t\cdot\mathsf{poly}(s)}$ which is close to 1 for $\epsilon = \frac{1}{t^2\cdot\mathsf{poly}(s)}$.

To prove the induction step we use conditional probabilities, and condition on the event that the value of a random point in the low-degree extension of layer i is indeed correct. Conditioned on this event, each element in the i-th layer of $\mathcal{C}_n$ is correct with probability $(1-\epsilon)$ (which is the probability in which local consistency holds). Therefore, conditioned on this event, each element in the $i+1$-th layer of $\mathcal{C}_n$ is correct with probability $(1-\epsilon)^3$. This implies that, conditioned on this event, each element in the low-degree extension of the $i+1$-th layer is correct with probability $(1-\epsilon)^{\mathsf{poly}(s)}$. Therefore, without conditioning, the probability that an element in the low-degree extension of the $i+1$-th layer is correct is $p(1-\epsilon)^{\mathsf{poly}(s)}$.

This analysis does not quite work as is, since there is too big of a loss in the correctness probability when going from a random point in the low-degree extension to a point in the layer of $\mathcal{C}_n$. We fix this by reading several random points in the low-degree extension (rather than just one), and we claim that if all of them are correct with some probability then each point in the layer is correct with essentially the same probability (where the loss here is exponentially small). We refer the reader to Section 10 for details.

3.2 Converting a Statistically No-Signaling MIP into a One-Round Delegation Scheme

In this section we show that the method of Aiello *et al.* [ABOR00], of using a fully homomorphic (FHE) scheme to convert a 1-round MIP into a 1-round delegation scheme, is *sound* if the underlying MIP is secure against δ-no-signaling provers, where the value of δ affects the security requirement of the FHE scheme.[12]

Let us start by recalling their method. Aiello *et al.* proposed to take any MIP and convert it into the following 1-round delegation scheme: The verifier computes all the queries that the MIP verifier would send to the MIP provers, and sends all of these queries to the prover, each encrypted under a fresh and independent key, using an FHE scheme. The prover then answers on behalf of each MIP prover, where each answer is computed *homomorphically* on the corresponding encrypted query.

As mentioned in the introduction, shortly after this method was introduced, Dwork *et al.* [DLN+04] showed that it may, in general, be insecure. In this work, we show that this method in fact is *secure* if the underlying MIP is sound against δ-no-signaling provers.

In a nutshell, our result is obtained by proving that if there exists a cheating prover P^* that breaks the soundness of the 1-round argument, then this prover can be used to construct a δ-no-signaling prover P^{NS} that breaks the soundness of the MIP scheme.

The prover P^{NS} uses P^* in the obvious way: Given a set of queries $(q_1, \ldots, q_\ell)$ it encrypts these queries using fresh and independent keys, and sends the encrypted queries to P^*; upon receiving encrypted answers, it decrypts these answers and sends the decrypted answers $(a_1, \ldots, a_\ell)$ to the MIP verifier.

[12] Aiello *et al.* originally suggested to use a PCP together with a private information retrieval (PIR) scheme to construct a 1-round delegation scheme.

Clearly this strategy breaks the soundness of the MIP verifier, but we need to argue that it is δ-no-signaling. Indeed, we argue that if P^{NS} is *not* δ-no-signaling then the prover P^* can be used to break the underlying FHE scheme. Loosely speaking, by the definition of δ-no-signaling (see Section 4.3), if P^{NS} is *not* δ-no-signaling then there is a subset $S \subset [\ell]$ such that the distribution of the answers $(a_i)_{i \in S}$, conditioned on the corresponding queries $(q_i)_{i \in S}$, depends on the other queries $(q_i)_{i \notin S}$. In other words, these answers give information on the other queries. If this is the case, then indeed one can use P^* to break the FHE scheme.

We note that the above break may take time exponential in the communication complexity of the underlying MIP scheme, since the information obtained from the answers $(a_i)_{i \in S}$, is not necessarily efficiently computable. Therefore, we need to assume that the underlying FHE scheme is secure against adversaries of size $2^{|a_1|+\ldots+|a_\ell|}$. Thus, if we choose the security parameter of the FHE scheme to be polynomially related to the communication complexity, then we need to assume sub-exponential security of the underlying FHE scheme. But one can choose a larger security parameter (resulting in larger communication complexity in the 1-round delegation scheme), and thus relax the security requirement of the FHE scheme. We refer the reader to Section 16 for details.

4 Preliminaries

4.1 Notation

For a vector $a = (a_1, \ldots, a_k)$ and a subset $S \subseteq [k]$, we denote by a_S the sequence of elements of a that are indexed by indices in S, that is, $a_S = (a_i)_{i \in S}$. In general, we denote by a_S a sequence of elements indexed by S, and we denote by a_i the i^{th} coordinate of a vector a.

For a distribution $\mathcal{A}$, we denote by $a \in_R \mathcal{A}$ a random variable distributed according to $\mathcal{A}$ (independently of all other random variables).

We will measure the distance between two distributions by their *statistical distance*, defined as half the l_1-distance between the distributions. We will say that two distributions are δ-*close* if their statistical distance is at most δ.

For a field $\mathbb{F}$ and an integer ℓ, a line L in $\mathbb{F}^\ell$ is an affine function $L : \mathbb{F} \to \mathbb{F}^\ell$. A plain M in $\mathbb{F}^\ell$ is an affine function $M : \mathbb{F}^2 \to \mathbb{F}^\ell$. We say that the line L is orthogonal to the i^{th} coordinate if for every $t_1, t_2 \in \mathbb{F}$, we have $L(t_1)_i = L(t_2)_i$, where $L(t_1)_i, L(t_2)_i$ denote the i^{th} coordinate of the points $L(t_1), L(t_2)$ respectively.

We will sometimes confuse between a set and a multiset. In particular, many times we will refer to a multiset as a set. For example, when we choose a (multi)set of k elements in a certain domain.

We will sometimes write $\Pr_x \Pr_y$ instead of $\Pr_{x,y}$.

4.2 Multi-Prover Interactive Proofs

Let $\mathcal{L}$ be a language and let x be an input of length n. In a one-round k-prover interactive proof, k computationally unbounded provers, $P_1, \ldots, P_k$, try to convince a (probabilistic) $\mathsf{poly}(n)$-time verifier, V, that $x \in \mathcal{L}$. The input x is known to all parties.

The proof consists of only one round. Given x and her random string, the verifier generates k queries, $q_1, \ldots, q_k$, one for each prover, and sends them to the k provers. Each prover responds with an answer that depends only on her own individual query. That is, the provers respond with answers $a_1, \ldots, a_k$, where for every i we have $a_i = P_i(q_i)$. Finally, the verifier decides wether to accept or reject based on the answers that she receives (as well as the input x and her random string).

We say that $(V, P_1, \ldots, P_k)$ is a one-round multi-prover interactive proof system (MIP) for $\mathcal{L}$ if the following two properties are satisfied:

1. **Completeness:** For every $x \in \mathcal{L}$, the verifier V accepts with probability 1, after interacting with $P_1, \ldots, P_k$.
2. **Soundness:** For every $x \notin \mathcal{L}$, and any (computationally unbounded, possibly cheating) provers $P_1^*, \ldots, P_k^*$, the verifier V rejects with probability $\geq 1 - \epsilon$, after interacting with $P_1^*, \ldots, P_k^*$, where ϵ is a parameter referred to as the *error* or *soundness* of the proof system.

Important parameters of an MIP are the number of provers, the length of queries, the length of answers, and the error.

4.2.1 MIPs with Oracle

We will also consider the model of *one-round k-prover interactive proofs with oracle*, where the verifier V is given access to an oracle that computes some fixed function (that may depend on the language $\mathcal{L}$). We require that all queries, to the oracle and the provers, are done simultaneously.

For every n, let $\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}$ be a function (where n', n'' depend on n). We allow the functions ϕ_n to depend on the language $\mathcal{L}$ (but not on the input x).

We define a *one-round multi-prover interactive proof system for $\mathcal{L}$, relative to the oracle $\{\phi_n\}$*, exactly as before, except that now the verifier V is a (probabilistic, $\mathsf{poly}(n)$-time) oracle machine that on input x of length n has free oracle access to the function ϕ_n. The verifier may base her accept/reject decision on queries to the oracle, but the oracle queries are not adaptive, and we do not allow the queries to the provers to depend on the answers of the oracle or the queries to the oracle to depend on the answers of the provers. In other words, we require that all queries, to the oracle and to the provers, are done simultaneously.

We require the same completeness and soundness properties as before.

4.3 No-Signaling MIPs

We will consider a variant of the MIP model, where the cheating provers are more powerful. In the MIP model, each prover answers her own query locally, without knowing the queries that were sent to the other provers. The no-signaling model allows each answer to depend on all the queries, as long as for any subset $S \subset [k]$, and any queries q_S for the provers in S, the distribution of the answers a_S, conditioned on the queries q_S, is independent of all the other queries.

Intuitively, this means that the answers a_S do not give the provers in S information about the queries of the provers outside S, except for information that they already have by seeing the queries q_S.

Formally, denote by D the alphabet of the queries and denote by Σ the alphabet of the answers. For every $q = (q_1, \ldots, q_k) \in D^k$, let $\mathcal{A}_q$ be a distribution over Σ^k. We think of $\mathcal{A}_q$ as the distribution of the answers for queries q.

We say that the family of distributions $\{\mathcal{A}_q\}_{q \in D^k}$ is *no-signaling* if for every subset $S \subset [k]$ and every two sequences of queries $q, q' \in D^k$, such that $q_S = q'_S$, the following two random variables are identically distributed:

- a_S, where $a \in_R \mathcal{A}_q$
- a'_S where $a' \in_R \mathcal{A}_{q'}$

If the two distributions are δ-close, rather than identical, we say that the family of distributions $\{\mathcal{A}_q\}_{q \in D^k}$ is *δ-no-signaling*.

An MIP, $(V, P_1, \ldots, P_k)$ for a language $\mathcal{L}$ (possibly, relative to an oracle $\{\phi_n\}$) is said to have soundness ϵ against no-signaling strategies (or provers) if the following (more general) soundness property is satisfied:

2. **Soundness:** For every $x \notin \mathcal{L}$, and any no-signaling family of distributions $\{\mathcal{A}_q\}_{q \in D^k}$, the verifier V rejects with probability $\geq 1 - \epsilon$, where on queries $q = (q_1, \ldots, q_k)$ the answers are given by $(a_1, \ldots, a_k) \in_R \mathcal{A}_q$, and ϵ is the error parameter.

If the property is satisfied for any δ-no-signaling family of distributions $\{\mathcal{A}_q\}_{q \in D^k}$, we say that the MIP has soundness ϵ against δ-no-signaling strategies (or provers).

4.4 Probabilistically Checkable Proofs

Let $\mathcal{L}$ be a language and let x be an input of length n. Intuitively, a probabilistically checkable proof (PCP) is a proof for $x \in \mathcal{L}$ that can be verified by reading only a small number of its symbols.

Formally, a proof is a vector of symbols $P \in \Sigma^D$, where Σ denotes the alphabet of symbols and D denotes the set of indices. We think of P also as a function $P : D \to \Sigma$ and hence we think of D as a set of possible queries and we think of Σ as a set of possible answers.

A PCP verifier V is a probabilistic $\mathsf{poly}(n)$-time Turing machine that is given access to the input x, as well as an oracle access to the proof $P : D \to \Sigma$.

Given x and her random string, the verifier generates k queries, $q_1, \ldots, q_k \in D$, and queries the proof P in all these places to get answers $a_1 = P(q_1), \ldots, a_k = P(q_k)$. Finally, the verifier decides wether to accept or reject based on the answers that she receives (as well as the input x and her random string).

We say that V is a PCP verifier for $\mathcal{L}$ if the following two properties are satisfied:

1. **Completeness:** For every $x \in \mathcal{L}$, there exists a proof P, such that, the verifier V accepts with probability 1, after querying P.
2. **Soundness:** For every $x \notin \mathcal{L}$, and any proof $P^* : D \to \Sigma$, the verifier V rejects with probability $\geq 1 - \epsilon$, after querying P^*, where ϵ is a parameter referred to as the *error* or *soundness* of the proof system.

Important parameters of a PCP are the length of proof, the number of queries, the length of answers, and the error.

4.4.1 PCPs with Oracle

We will also consider the model of *probabilistically checkable proofs with oracle*, where the verifier V is given access to an additional oracle that computes some fixed function (that may depend on the language $\mathcal{L}$). We require that all queries, to the oracle and to the PCP proof, are done simultaneously.

For every n, let $\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}$ be a function (where n', n'' depend on n). We allow the functions ϕ_n to depend on the language $\mathcal{L}$ (but not on the input x).

We define a *probabilistically checkable proof for $\mathcal{L}$, relative to the oracle $\{\phi_n\}$*, exactly as before, except that now the verifier V is a (probabilistic, $\mathsf{poly}(n)$-time) machine with access to two oracles. The first oracle is the PCP proof P, and in addition, on input x of length n, the verifier has free oracle access to the function ϕ_n. The verifier may base her accept/reject decision on queries to the oracle ϕ_n, but the oracle queries are not adaptive, and we do not allow the queries to P to depend on the answers of the oracle ϕ_n or the queries to the oracle ϕ_n to depend on the answers of P. In other words, we require that all queries, to the oracle ϕ_n and to the PCP proof P, are done simultaneously.

We require the same completeness and soundness properties as before.

4.5 No-Signaling PCPs

We will now define the new notion of PCP with no-signaling soundness, in analogy to MIP with no-signaling soundness.

We will consider a variant of the PCP model, where the cheating proof is more powerful. In the PCP model, each query is answered locally, without knowing the other queries. In

the no-signaling model, we allow each answer to depend on all the queries, as long as for any subset $\{q_1, \ldots, q_d\}$ of queries, the distribution of the answers $(a_1, \ldots, a_d)$, conditioned on the queries $\{q_1, \ldots, q_d\}$, is independent of all the other queries.

Formally, denote by D the alphabet of the queries and denote by Σ the alphabet of the answers. Let k_{max} be some parameter, which is at least the maximal number of queries made by the verifier to the proof. For every subset $Q = \{q_1, \ldots, q_d\} \subset D$, of size $|Q| = d \leq k_{max}$, let $\mathcal{A}_Q$ be a distribution over Σ^Q. We think of $\mathcal{A}_Q$ as the distribution of the answers for the queries $\{q_1, \ldots, q_d\}$.

We say that the family of distributions $\{\mathcal{A}_Q\}_{Q \subset D, |Q| \leq k_{max}}$ is *no-signaling* if for every $Q \subset D$ of size at most k_{max}, and every subset $S \subset Q$, the following two random variables are identically distributed:

- $a \in_R \mathcal{A}_S$
- a'_S, where $a' \in_R \mathcal{A}_Q$

If the two distributions are δ-close, rather than identical, we say that the family of distributions $\{\mathcal{A}_Q\}_{Q \subset D, |Q| \leq k_{max}}$ is *δ-no-signaling*.

A PCP verifier V for a language $\mathcal{L}$ (possibly, relative to an oracle $\{\phi_n\}$) is said to have soundness ϵ against k_{max}-no-signaling strategies (or proofs) if the following (more general) soundness property is satisfied:

2. **Soundness:** For every $x \notin \mathcal{L}$, and any no-signaling family of distributions $\{\mathcal{A}_Q\}_{Q \subset D, |Q| \leq k_{max}}$, the verifier V rejects with probability $\geq 1 - \epsilon$, where on queries $Q = \{q_1, \ldots, q_k\}$, the answers are given by $a_Q \in_R \mathcal{A}_Q$, and ϵ is the error parameter.

If the property is satisfied for any δ-no-signaling family of distributions $\{\mathcal{A}_Q\}_{Q \subset D, |Q| \leq k_{max}}$, we say that the PCP has soundness ϵ against (k_{max}, δ)-no-signaling strategies (or proofs).

Note that k_{max} is an important parameter. The larger k_{max} is, the more limited the cheating proofs are. We will typically take k_{max} to be significantly larger than the maximal number of queries made by the verifier.

4.5.1 A Note on Ordered versus Unordered Sets

The families of distributions $\{\mathcal{A}_Q\}_{Q \subset D, |Q| \leq k_{max}}$ that we consider are defined with *unordered* sets $Q \subset D$. However, sometimes we will have *ordered* sets Q (that is, Q will be a vector of elements); for example, when we need to know which test to apply on which subset of queries, it is important that the set of queries is ordered by the order that the queries were chosen. In these cases, we will abuse notation and denote by Q both the ordered set and the unordered set that corresponds to it. Thus, we will use the notation $\mathcal{A}_Q$ to denote the distribution that corresponds to the unordered set that corresponds to Q. In general, we will sometimes abuse notation between an ordered set and the unordered set that corresponds to it.

4.6 Low Degree Extension

Let $\mathbb{F}$ be a field and $H \subset \mathbb{F}$ a subset of the field. Fix an integer $m \in \mathbb{N}$. A basic fact is that for every function $\phi : H^m \to \mathbb{F}$, there exists a unique extension of ϕ into a function $\hat{\phi} : \mathbb{F}^m \to \mathbb{F}$ (which agrees with ϕ on H^m; i.e., $\hat{\phi}|_{H^m} \equiv \phi$), such that $\hat{\phi}$ is an m-variate polynomial of degree at most $|H| - 1$ in each variable. Moreover, for every $x \in H^m$, there exists a unique m-variate polynomial $\hat{\beta}_x : \mathbb{F}^m \to \mathbb{F}$ of degree $|H| - 1$ in each variable, such that for every function $\phi : H^m \to \mathbb{F}$ it holds that

$$\hat{\phi}(z_1, \ldots, z_m) = \sum_{x \in H^m} \hat{\beta}_x(z_1, \ldots, z_m) \cdot \phi(x).$$

The function $\hat{\phi}$ is called the *low degree extension* of ϕ (with respect to $\mathbb{F}, H, m$).

In the following we assume that all algorithms have access to m, the set H and the field $\mathbb{F}$. We assume that field operations over $\mathbb{F}$ can be computed in time poly-logarithmic in the field size and space that is logarithmic in the field size.

Proposition 4.1 (Cf., e.g., [Rot09, Proposition 3.2.1])**.** *There exists a Turing machine that on input $x \in H^m$, runs in time $\mathsf{poly}(|H|, m, \log|\mathbb{F}|)$ and space $O(\log(|\mathbb{F}|) + \log(m))$, and outputs the polynomial $\hat{\beta}_x : \mathbb{F}^m \to \mathbb{F}$ defined above, represented as an arithmetic circuit over $\mathbb{F}$.*

Moreover, the arithmetic circuit $\hat{\beta}_x$ can be evaluated in time $\mathsf{poly}(|H|, m, \log(|\mathbb{F}|))$ and space $O(\log(|\mathbb{F}|) + \log(m))$. Namely, there exists a Turing machine with the above time and space bounds that given an input pair $(x, z) \in H^m \times \mathbb{F}^m$ outputs $\hat{\beta}_x(z)$.

Proof. Consider the function $\hat{\beta}_x : \mathbb{F}^m \to \mathbb{F}$ defined as:

$$\hat{\beta}_x(z) \stackrel{\text{def}}{=} \prod_{i \in [m]} \prod_{h \in H \setminus \{x_i\}} \frac{z_i - h}{x_i - h}.$$

For every $z \in H^m$ it holds that $\hat{\beta}_x(z) = 1$ if $z = x$ and $\hat{\beta}_x(z) = 0$ otherwise. Thus, for every function $\phi : H^m \to \mathbb{F}$ it holds that $\sum_{x \in H^m} \hat{\beta}_x \cdot \phi(x)$ agrees with ϕ on H^m. Hence, since $\hat{\beta}_x$ has degree $|H| - 1$ in each variable, $\sum_{x \in H^m} \hat{\beta}_x \cdot \phi(x)$ is the (unique) low degree extension of ϕ. □

Proposition 4.2. *Let $\phi : H^m \to \mathbb{F}$ and suppose that ϕ can be evaluated by a Turing Machine in time t and space s. Then, there exists a Turing machine that, given as an input a point $z \in \mathbb{F}^m$, runs in time $|H|^m \left(\mathsf{poly}(|H|, m, \log(|\mathbb{F}|)) + O(t)\right)$ and space $O(m \log(|H|) + s + \log(|\mathbb{F}|))$ and outputs the value $\hat{\phi}(z)$ where $\hat{\phi}$ is the unique low degree extension of ϕ (with respect to $H, \mathbb{F}, m$).*

Proof. The Turing machine computes

$$\hat{\phi}(z) = \sum_{x \in H^m} \hat{\beta}_x(z) \cdot \phi(x)$$

by generating and evaluating $\hat{\beta}_x(z)$ as in Proposition 4.1. □

4.7 Public-Key Encryption and Fully Homomorphic Encryption (FHE)

A *public-key encryption* scheme consists of three probabilistic polynomial-time algorithms $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$. The key generation algorithm Gen, when given as input a security parameter 1^τ, outputs a pair $(\mathsf{pk}, \mathsf{sk})$ of public and secret keys. The encryption algorithm, Enc, on input a public key pk and a message $m \in \{0,1\}^{\mathsf{poly}(\tau)}$, outputs a ciphertext $\hat{m}$, and the decryption algorithm, Dec, when given the ciphertext $\hat{m}$ and the secret key sk, outputs the original message m (with overwhelming probability). We allow the decryption process to fail with negligible probability (over the randomness of all algorithms).

Let $S : \mathbb{N} \to \mathbb{N}$ and $\delta : \mathbb{N} \to [0,1]$ be parameters. A public-key encryption scheme has security (S, δ) if for every family of circuits $\{C_\tau\}_{\tau \in \mathbb{N}}$ of size $\mathsf{poly}(S(\tau))$, for all sufficiently large τ and for any two messages $m, m' \in \{0,1\}^{\mathsf{poly}(\tau)}$ such that $|m| = |m'|$,

$$\left| \Pr_{(\mathsf{pk},\mathsf{sk}) \in_R \mathsf{Gen}(1^\tau)} [C_\tau(\mathsf{pk}, \mathsf{Enc}_{\mathsf{pk}}(m)) = 1] - \Pr_{(\mathsf{pk},\mathsf{sk}) \in_R \mathsf{Gen}(1^\tau)} [C_\tau(\mathsf{pk}, \mathsf{Enc}_{\mathsf{pk}}(m')) = 1] \right| < \delta(\tau)$$

where the probability is also over the random coin tosses of Enc.

Fully homomorphic encryption. The tuple $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec}, \mathsf{Eval})$ is a *fully-homomorphic encryption scheme* if (1) $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is a public-key encryption scheme, and (2) for every key-pair (pk, sk), the probabilistic polynomial-time algorithm Eval, on input the public-key pk, a circuit $C : \{0,1\}^k \to \{0,1\}^\ell$, where $k, \ell \leq \mathsf{poly}(\tau)$ (and τ is the security parameter), and a ciphertext $\hat{m}$ that is an encryption of a message $m \in \{0,1\}^k$ with respect to pk, outputs a string ψ such that the following two conditions hold:

- **Homomorphic Evaluation:** $\mathsf{Dec}_{\mathsf{sk}}(\psi) = C(m)$, except with negligible probability (over the coins of all algorithms).
- **Compactness:** The length of ψ is polynomial in τ, k and ℓ (and is independent of the size of C).

4.8 Interactive Argument Systems

An interactive argument for a language $\mathcal{L}$ consists of a polynomial-time verifier that wishes to verify a statement of the form $x \in \mathcal{L}$, and a prover that helps the verifier to decide. The two parties, given as input $x \in \{0,1\}^n$, interact and at the end of the interaction the verifier either accepts or rejects. We require that if $x \in \mathcal{L}$ then the verifier accepts with high probability but if $x \notin \mathcal{L}$, then no *computationally bounded* prover can convince the verifier to accept with non-negligible (in n) probability.

We focus on 1-round argument systems. Such an argument-system consists of a single message sent from the verifier V to the prover P, followed by a single message sent from the prover to the verifier.

Let $S : \mathbb{N} \to \mathbb{N}$ and $\epsilon : \mathbb{N} \to [0,1]$ be parameters. We say that (V, P) is a one-round argument-system with soundness (S, ϵ) for $\mathcal{L}$ if the following two properties are satisfied:

1. **Completeness:** For every $x \in \mathcal{L}$, the verifier $V(x)$ accepts with overwhelming probability, after interacting with $P(x)$.
2. **Soundness:** For every family of circuits $\{P_n^*\}_{n\in\mathbb{N}}$ of size $\mathsf{poly}(S(n))$, for all sufficiently large $x \notin \mathcal{L}$, the verifier V rejects with probability $\geq 1 - \epsilon(|x|)$, after interacting with $P_{|x|}^*$ on common input x.

5 The Base PCP

5.1 The PCP Proof

Let $\mathcal{L}$ be a language in $\mathsf{DTISP}(t(n), s(n))$, where $\mathsf{poly}(n) \leq t(n) \leq \mathsf{exp}(n)$ and $\log(n) \leq s(n) \leq \mathsf{poly}(n)$. Let x be an input of length n. Since $\mathcal{L} \in \mathsf{DTISP}(t(n), s(n))$, for any n there is a (fanin 2) Boolean circuit $\mathcal{C}_n$ of size $N = O(t(n)s(n))$ that computes $\mathcal{L}$ on inputs of length n. Moreover, the circuit $\mathcal{C}_n$ is layered, with $O(s(n))$ gates in each layer, such that a child of a gate in layer $i+1$ is either an input variable (or a negation of an input variable) or a gate in layer i. Moreover, there is a deterministic Turing machine of space $O(\log N)$ that on input n outputs the (description of the) circuit $\mathcal{C}_n$.

Without loss of generality, we assume that in the circuit $\mathcal{C}_n$ all negations are on input variables, and that the two children of any gate in the circuit are different (this property can be achieved by duplicating each gate in the circuit twice, increasing the number of gates in each layer by a factor of 2).

Also, we assume that the gates of the circuit are indexed by the numbers $1, \ldots, N$, in an order that agrees with the layers of the circuit. In particular, for every gate, the index of the gate is larger than the indexes of its children. We assume that $1, \ldots, n$ are the indexes of the n input variables and $n+1, \ldots, 2n$ are the indexes of their negations. We assume that the circuit has a special output gate indexed by N whose value represents the decision of whether $x \in \mathcal{L}$ (we do not assume that there are no other output gates). We assume that the Turing machine that outputs the (description of the) circuit $\mathcal{C}_n$ outputs the vertices in the order of their index.

Let $w_1, \ldots, w_N$ be variables in $\{0,1\}$ that represent the N wires of the circuit $\mathcal{C}_n$, in the order of their index. In particular, for every gate, the variable that represents the output of the gate appears after the variables that represent the inputs for the gate. Also, $w_1, \ldots, w_n$ represent the n input bits, $w_{n+1}, \ldots, w_{2n}$ represent the negations of the n input bits, and w_N represents the output of the circuit.

Let $\varphi_{\mathcal{C}}(w_1, \ldots, w_N)$ be a 3-CNF Boolean formula that checks that $w_1, \ldots, w_N$ is a correct computation of the circuit $\mathcal{C}_n$ (given the input variables and their negations), by checking that the computation of every gate in the circuit is performed correctly (except for negation

gates on input variables - and recall that we assume that these are the only negation gates in the circuit). More precisely, for every (non-negation) gate in the circuit, the formula $\varphi_{\mathcal{C}}$ contains four clauses that check that the computation of that gate is performed correctly for every possibility for the inputs for the gate. For example, if w_i represents the output of a conjunction gate with inputs that are represented by w_{i_1} and w_{i_2}, we will have the following four clauses in $\varphi_{\mathcal{C}}$ (and note that indeed each of them can be written as a clause):

$$\begin{aligned}
&(w_{i_1} = 0) \wedge (w_{i_2} = 0) \rightarrow (w_i = 0),\\
&(w_{i_1} = 0) \wedge (w_{i_2} = 1) \rightarrow (w_i = 0),\\
&(w_{i_1} = 1) \wedge (w_{i_2} = 0) \rightarrow (w_i = 0),\\
&(w_{i_1} = 1) \wedge (w_{i_2} = 1) \rightarrow (w_i = 1).
\end{aligned}$$

We have $\varphi_{\mathcal{C}}(w_1, \ldots, w_N) = 1$ if and only if $w_1, \ldots, w_N$ is a correct computation of the circuit $\mathcal{C}_n$ (assuming that the input variables are given in $w_1, \ldots, w_n$, and their negations are given in $w_{n+1}, \ldots, w_{2n}$).

For a fixed input $x = (x_1, \ldots, x_n)$, let $\varphi_x(w_1, \ldots, w_{2n}, w_N)$ be a 3-CNF Boolean formula that checks that $(w_1, \ldots, w_n) = (x_1, \ldots, x_n)$, $(w_{n+1}, \ldots, w_{2n}) = (\neg x_1, \ldots, \neg x_n)$ and that $w_N = 1$. More precisely, for every $i \in [n]$, the formula φ_x contains a clause that checks that $w_i = x_i$. For example, if $x_i = 0$, we will have the clause $(w_i = 0) \vee (w_i = 0) \vee (w_i = 0)$ that ensures that $w_i = 0$. In the same way, the formula φ_x contains clauses that check that $w_{n+i} = \neg x_i$, and a clause that checks that $w_N = 1$. We have $\varphi_x(w_1, \ldots, w_{2n}, w_N) = 1$ if and only if $(w_1, \ldots, w_n) = (x_1, \ldots, x_n)$, $(w_{n+1}, \ldots, w_{2n}) = (\neg x_1, \ldots, \neg x_n)$, and $w_N = 1$.

Let $\varphi(w_1, \ldots, w_N)$ be the 3-CNF Boolean formula $\varphi_{\mathcal{C}}(w_1, \ldots, w_N) \wedge \varphi_x(w_1, \ldots, w_{2n}, w_N)$. Thus, $\varphi(w_1, \ldots, w_N) = 1$ if and only if $w_1, \ldots, w_N$ is the computation of the circuit $\mathcal{C}_n$ on the input $x = (x_1, \ldots, x_n)$, and $w_N = 1$. Denote by $x_1, \ldots, x_N$ the computation of the circuit $\mathcal{C}_n$ on the input $x = (x_1, \ldots, x_n)$. Thus, $\varphi(w_1, \ldots, w_N) = 1$ if and only if $(w_1, \ldots, w_N) = (x_1, \ldots, x_N)$, and $x_N = 1$.

Note also that since there is a deterministic Turing machine of space $O(\log N)$ that on input n outputs the description of the circuit $\mathcal{C}_n$, there is a deterministic Turing machine of space $O(\log N)$ that on input n outputs the formula $\varphi_{\mathcal{C}}$.

Let $H = \{0, 1, \ldots, \log N - 1\}$ and let $m = \frac{\log N}{\log \log N}$, so that $N = |H|^m$. (For simplicity and without loss of generality we assume that $\log N$ and $\frac{\log N}{\log \log N}$ are integers, larger than 100). Let $\ell = 3m + 3$. Let $\mathbb{F}$ be a field, such that $4|H|^{10} \leq |\mathbb{F}| \leq 8(\log N)^{10}$.

Since $N = |H|^m$, we can identify $[N]$ and H^m (say, by the lexicographic order on H^m). In what follows we will abuse notation and view $w_1, \ldots, w_N$ and $x_1, \ldots, x_N$ as indexed by $i \in H^m$ (rather than $i \in [N]$). We can hence view $x = (x_1, \ldots, x_N)$ as a function $x : H^m \rightarrow \{0, 1\}$ (given by $x(i) = x_i$, where we identify $[N]$ and H^m).

Define the multi-variate polynomial $X : \mathbb{F}^m \rightarrow \mathbb{F}$ to be the low-degree extension of $x : H^m \rightarrow \{0, 1\}$.

Let $\phi : (H^m)^3 \times \{0, 1\}^3 \rightarrow \{0, 1\}$ be the function where $\phi(i_1, i_2, i_3, b_1, b_2, b_3) = 1$ if and

only if the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in φ. Extend ϕ to be a function $\phi : H^{3m+3} \to \{0,1\}$ by setting it to be 0 for inputs outside of $H^{3m} \times \{0,1\}^3$. Let $\hat{\phi} : \mathbb{F}^\ell \to \mathbb{F}$ be the low-degree extension of ϕ.

Let $\phi_{\mathcal{C}} : (H^m)^3 \times \{0,1\}^3 \to \{0,1\}$ be the function where $\phi_{\mathcal{C}}(i_1, i_2, i_3, b_1, b_2, b_3) = 1$ if and only if the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in $\varphi_{\mathcal{C}}$. Extend $\phi_{\mathcal{C}}$ to be a function $\phi_{\mathcal{C}} : H^{3m+3} \to \{0,1\}$ by setting it to be 0 for inputs outside of $H^{3m} \times \{0,1\}^3$. Let $\hat{\phi}_{\mathcal{C}} : \mathbb{F}^\ell \to \mathbb{F}$ be the low-degree extension of $\phi_{\mathcal{C}}$.

Let $\phi_x : (H^m)^3 \times \{0,1\}^3 \to \{0,1\}$ be the function where $\phi_x(i_1, i_2, i_3, b_1, b_2, b_3) = 1$ if and only if the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in φ_x. Extend ϕ_x to be a function $\phi_x : H^{3m+3} \to \{0,1\}$ by setting it to be 0 for inputs outside of $H^{3m} \times \{0,1\}^3$. Let $\hat{\phi}_x : \mathbb{F}^\ell \to \mathbb{F}$ be the low-degree extension of ϕ_x.

Since the sets of clauses of $\varphi_{\mathcal{C}}$ and φ_x are disjoint, we have $\hat{\phi} = \hat{\phi}_x + \hat{\phi}_{\mathcal{C}}$.

Recall that there is a deterministic Turing machine of space $O(\log N)$ that on input n outputs the formula $\varphi_{\mathcal{C}}$. Hence, by Proposition 4.2, there is a deterministic Turing machine of space $O(\log N)$ that on input $z \in \mathbb{F}^\ell$ outputs $\hat{\phi}_{\mathcal{C}}(z)$. Since ϕ_x is Boolean valued and is zero on all but a fixed set of $2n+1$ points (specifically, the clauses that verify the correctness of the inputs and output), its low degree extension $\hat{\phi}_x$ can be evaluated on a point $z \in \mathbb{F}^\ell$ in time $n \cdot \mathsf{polylog} N$ (by using Proposition 4.1 and iterating only over the set of $O(n)$ potentially non-zero points).

Since for $x \in \mathcal{L}$ we have $\varphi(x_1, \ldots, x_N) = 1$, every clause that appears in φ is satisfied by $(x_1, \ldots, x_N)$. Therefore, if $x \in \mathcal{L}$, for every $z = (i_1, i_2, i_3, b_1, b_2, b_3) \in (H^m)^3 \times H^3 = H^\ell$, we have

$$\hat{\phi}(z) \cdot (X(i_1) - b_1) \cdot (X(i_2) - b_2) \cdot (X(i_3) - b_3) = 0 \tag{1}$$

Let $P_0 : \mathbb{F}^\ell \to \mathbb{F}$ be the multivariate polynomial defined as follows:
For $z = (i_1, i_2, i_3, b_1, b_2, b_3) \in (\mathbb{F}^m)^3 \times \mathbb{F}^3 = \mathbb{F}^\ell$,

$$P_0(z) \triangleq \hat{\phi}(z) \cdot (X(i_1) - b_1) \cdot (X(i_2) - b_2) \cdot (X(i_3) - b_3)$$

Equation (1) implies that if $x \in \mathcal{L}$ then $P_0|_{H^\ell} \equiv 0$. Moreover, the fact that X and $\hat{\phi}$ have degree $< |H|$ in each variable, implies that P_0 has degree $< 2|H|$ in each variable, and hence total degree $< 2|H|\ell$.

Next we define $P_1 : \mathbb{F}^\ell \to \mathbb{F}$. For every $z = (z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$, let

$$P_1(z) = \sum_{h \in H} P_0(h, z_2, \ldots, z_\ell) {z_1}^h$$

Note that if $x \in \mathcal{L}$ then $P_1|_{\mathbb{F} \times H^{\ell-1}} \equiv 0$. More generally, we define by induction $P_1, \ldots, P_\ell : \mathbb{F}^\ell \to \mathbb{F}$ where for every $z = (z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$,

$$P_i(z) = \sum_{h \in H} P_{i-1}(z_1, \ldots, z_{i-1}, h, z_{i+1}, \ldots, z_\ell) {z_i}^h$$

Note that $P_1, \ldots, P_{\ell-1}$ have degree $< 2|H|$ in each variable, and hence total degree $< 2|H|\ell$. Note also that if $x \in \mathcal{L}$ then $P_i|_{\mathbb{F}^i \times H^{\ell-i}} \equiv 0$, and in particular $P_\ell \equiv 0$.

The PCP proof for $x \in \mathcal{L}$ consists of $\ell + 1$ multivariate polynomials: The polynomial $X : \mathbb{F}^m \to \mathbb{F}$ and the ℓ polynomials $P_i : \mathbb{F}^\ell \to \mathbb{F}$, for $i = 0, \ldots, \ell - 1$. To these polynomials we add the polynomial $P_\ell \equiv 0$. The polynomial P_ℓ is not part of the PCP proof (as it is the 0 polynomial) and is added just for simplicity of the notation. When the verifier queries $P_\ell(z)$ she gets 0 automatically.

Let $D_X = \mathbb{F}^m$ be the domain of X, and let $D_0, \ldots, D_\ell$ be $\ell + 1$ copies of $\mathbb{F}^\ell$, the domain of $P_0, \ldots, P_\ell$. We view $D_X, D_0, \ldots, D_\ell$ as the domains of $X, P_0, \ldots, P_\ell$, respectively. Denote,

$$D = D_X \cup D_0 \cup \ldots \cup D_\ell$$

The set D is the alphabet of queries in the PCP. We will refer to D as the domain of the PCP.

5.1.1 Complexity of the Prover

Note that the entire PCP proof can be generated in time $\mathsf{poly}(N) = \mathsf{poly}(t(n))$.

5.2 The PCP Verifier, V

The verifier knows the language $\mathcal{L}$, or more precisely, she knows the Turing machine of space $O(\log N)$ that on input n outputs the description of the circuit $\mathcal{C}_n$. The verifier gets an input x of length n and she wants to verify that $x \in \mathcal{L}$ by querying the PCP proof $X, P_0, P_1, \ldots, P_\ell$.

We will first assume that the verifier has access to the correct values of the function $\hat{\phi} : \mathbb{F}^\ell \to \mathbb{F}$. That is, the verifier can get the correct value of $\hat{\phi}(z)$ for free, for as many points $z \in \mathbb{F}^\ell$ as she wants.

Let $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N$, be a security parameter. (The restriction $k \leq \mathsf{poly}(n)$ is because we would like the running time of the verifier to be at most $\mathsf{poly}(n)$).

Recall that we denote by a_i the i^{th} coordinate of a vector a. In particular, for a line $L : \mathbb{F} \to \mathbb{F}^\ell$, a field element $t \in \mathbb{F}$ and a coordinate $i \in \{1, \ldots, \ell\}$, we denote by $L(t)_i$ the i^{th} coordinate of the point $L(t) \in \mathbb{F}^\ell$. Recall that we say that a line $L : \mathbb{F} \to \mathbb{F}^\ell$ is orthogonal to the i^{th} coordinate if for every $t_1, t_2 \in \mathbb{F}$, we have $L(t_1)_i = L(t_2)_i$.

The verifier V makes the following tests on the PCP proof: a **Low Degree Test for** X; four types of **Low Degree Tests for** P_i; a **Sum Check for** P_i; and a test of **Consistency of** X **and** P_0 (the exact tests are described below). We note that we have four types of low degree tests for P_i, rather than one, just for the simplicity of the analysis. It would be sufficient to do only one test, similar to the low degree test for X (but repeated on $O(k \cdot |\mathbb{F}|^2)$ random lines, rather than k random lines), since all four types of tests that we actually do (and are formally described below) can be embedded in such a test.

Formally, the verifier V makes the following tests, and accepts if the PCP proof passes all of them:

1. **Low Degree Test for X:** Choose k random lines $L_1, \ldots, L_k : \mathbb{F} \to \mathbb{F}^m$. For every $L \in \{L_1, \ldots, L_k\}$, query X on all the points $\{L(t)\}_{t \in \mathbb{F}}$, and check that the univariate polynomial $X \circ L : \mathbb{F} \to \mathbb{F}$ is of degree $< m|H|$.

2. **Low Degree Test for P_i: Type 1 (fixed $L(0)_{i+1}$):** For every $i \in \{0, \ldots, \ell - 1\}$ and every $u \in \mathbb{F}$, choose k random lines $L_1, \ldots, L_k : \mathbb{F} \to \mathbb{F}^\ell$, such that, every line $L \in \{L_1, \ldots, L_k\}$ satisfies $L(0)_{i+1} = u$. For every $L \in \{L_1, \ldots, L_k\}$, query P_i on all the points $\{L(t)\}_{t \in \mathbb{F}}$, and check that the univariate polynomial $P_i \circ L : \mathbb{F} \to \mathbb{F}$ is of degree $< 2\ell|H|$.

3. **Low Degree Test for P_i: Type 2 (orthogonal to the $(i+1)^{th}$ coordinate):** For every $i \in \{0, \ldots, \ell-1\}$, choose k random lines $L_1, \ldots, L_k : \mathbb{F} \to \mathbb{F}^\ell$, such that, every line $L \in \{L_1, \ldots, L_k\}$ is orthogonal to the $(i+1)^{th}$ coordinate. For every $L \in \{L_1, \ldots, L_k\}$, query P_i on all the points $\{L(t)\}_{t \in \mathbb{F}}$, and check that the univariate polynomial $P_i \circ L : \mathbb{F} \to \mathbb{F}$ is of degree $< 2\ell|H|$.

4. **Low Degree Test for P_i: Type 3 (fixed $L(0)_{i+1}$; orthogonal to the i^{th} coordinate):** For every $i \in \{1, \ldots, \ell - 1\}$, and every $u \in \mathbb{F}$, choose k random lines $L_1, \ldots, L_k : \mathbb{F} \to \mathbb{F}^\ell$, such that, every line $L \in \{L_1, \ldots, L_k\}$ is orthogonal to the i^{th} coordinate, and satisfies $L(0)_{i+1} = u$. For every $L \in \{L_1, \ldots, L_k\}$, query P_i on all the points $\{L(t)\}_{t \in \mathbb{F}}$, and check that the univariate polynomial $P_i \circ L : \mathbb{F} \to \mathbb{F}$ is of degree $< 2\ell|H|$.

5. **Low Degree Test for P_i: Type 4 (fixed $L(0)_i$; orthogonal to the $(i+1)^{th}$ coordinate):** For every $i \in \{1, \ldots, \ell - 1\}$, and every $u \in \mathbb{F}$, choose k random lines $L_1, \ldots, L_k : \mathbb{F} \to \mathbb{F}^\ell$, such that, every line $L \in \{L_1, \ldots, L_k\}$ is orthogonal to the $(i+1)^{th}$ coordinate, and satisfies $L(0)_i = u$. For every $L \in \{L_1, \ldots, L_k\}$, query P_i on all the points $\{L(t)\}_{t \in \mathbb{F}}$, and check that the univariate polynomial $P_i \circ L : \mathbb{F} \to \mathbb{F}$ is of degree $< 2\ell|H|$.

6. **Sum Check for P_i:** For every $i \in \{1, \ldots, \ell\}$, choose k random points in $\mathbb{F}^\ell$. For each of these points, $z = (z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$, query P_i, P_{i-1} on all the points $\{(z_1, \ldots, z_{i-1}, t, z_{i+1}, \ldots, z_\ell)\}_{t \in \mathbb{F}}$, and check that for every $t \in \mathbb{F}$,

$$P_i(z_1, \ldots, z_{i-1}, t, z_{i+1}, \ldots, z_\ell) = \sum_{h \in H} P_{i-1}(z_1, \ldots, z_{i-1}, h, z_{i+1}, \ldots, z_\ell) t^h$$

7. **Consistency of X and P_0:** Choose k random points in $\mathbb{F}^\ell$. For each of these points, $z = (i_1, i_2, i_3, b_1, b_2, b_3) \in (\mathbb{F}^m)^3 \times \mathbb{F}^3 = \mathbb{F}^\ell$, query P_0 on the point z and X on the points i_1, i_2, i_3, and check that

$$P_0(z) = \hat{\phi}(z) \cdot (X(i_1) - b_1) \cdot (X(i_2) - b_2) \cdot (X(i_3) - b_3)$$

5.2.1 Complexity of the Verifier

Note that the total number of queries made by V to the PCP proof, as well as the total number of queries made by V to the function $\hat{\phi}$, are both at most $6k\ell|\mathbb{F}|^2$. The time complexity of V is $k \cdot \mathsf{polylog}(N) = k \cdot \mathsf{polylog}(t(n))$.

5.3 The Relaxed Verifier, V'

We will now define another verifier for the PCP proof $X, P_0, P_1, \ldots, P_\ell$. We will call the new verifier, the *relaxed verifier* with parameter r, such that $1 \leq r < k$, and denote it by V'. As before, V' knows the language $\mathcal{L}$ and the input x, and we assume that she has access to the correct values of the function $\hat{\phi} : \mathbb{F}^\ell \to \mathbb{F}$.

The verifier V' makes the exact same queries as V, but she accepts in some cases where V rejects.

Recall that V repeated every test k times: The Low Degree Test for X was repeated on k different lines in $\mathbb{F}^m$. The four types of Low Degree Tests for each P_i (and for three of these types, for each $u \in \mathbb{F}$), were each repeated on k different lines in $\mathbb{F}^\ell$. The Sum Check for each P_i (for $i \in \{1, \ldots, \ell\}$) was repeated on k different points in $\mathbb{F}^\ell$. The Consistency of X and P_0 was repeated on k different points in $\mathbb{F}^\ell$.

This gives a partition of all the tests made by V into groups, with exactly k tests in each group. The verifier V accepted if all the tests in all the groups passed. The relaxed verifier, V', accepts if in each group of k tests at least $k - r$ of the tests pass, that is, at most r tests fail.

6 Soundness of V' versus Soundness of V

In this section we will show that if the verifier V can be fooled to accept $x \notin \mathcal{L}$, with very small probability, then the verifier V' can be fooled to accept $x \notin \mathcal{L}$ with probability very close to 1. Intuitively, this makes sense because the relaxed verifier V' accepts even if she rejects some of the tests, as long as the number of tests rejected in each group of k tests is at most r.

Recall that $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N$, is the security parameter of the PCP, and that $1 \leq r < k$ is the parameter of the relaxed verifier V'. Recall that ℓ and $|\mathbb{F}|$ are bounded by $\mathsf{polylog}(N)$.

We will prove the following lemma.

Lemma 6.1. *Assume that V doesn't have soundness ϵ against (k_{max}, δ)-no-signaling strategies, where $\delta < \frac{\epsilon}{8 \cdot |\mathbb{F}|^{6k\ell|\mathbb{F}|^2}}$. Then, V' doesn't have soundness $1 - (10\ell|\mathbb{F}|2^{-r} + 2\delta)/\epsilon$ against (k'_{max}, δ')-no-signaling strategies, where $k'_{max} = k_{max} - 6k\ell|\mathbb{F}|^2$, and $\delta' = 8\delta|\mathbb{F}|^{6k\ell|\mathbb{F}|^2}/\epsilon$.*

Let us first sketch the main techniques that we will use in the proof of the lemma:

The main claim that we will need in order to prove the lemma (Claim 6.2), shows that if V, V' choose their queries independently then the probability that V accepts and V' rejects, when all answers are given by a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S\subset D,|S|\leq k_{max}}$, is very small. This will be true because for each group of k tests we can first choose the $2k$ tests for both V, V' and only then decide which tests go to V and which ones go to V'. If among the $2k$ tests many are rejected then V rejects with high probability. On the other hand, if among the $2k$ tests only few are rejected then V' always accepts on that group.

We will assume that there exists a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S\subset D,|S|\leq k_{max}}$ that fools V with probability larger than ϵ. That is, the verifier V accepts with probability $> \epsilon$, where on queries Q, the answers are given (probabilistically) by $A_Q \in_R \mathcal{A}_Q$. We will construct a δ'-no-signaling family of distributions $\{\mathcal{A}'_S\}_{S\subset D,|S|\leq k'_{max}}$ that fools V' with probability close to 1.

This will be done by fixing a set of queries q for V and answers a_q for the queries in q, such that V accepts on queries q and answers a_q. The queries q will be chosen randomly by the distribution of V, and the answers a_q will be chosen randomly by the distribution $\mathcal{A}_q$, conditioned on the event that V accepts on queries q and answers a_q. The family $\{\mathcal{A}'_S\}$ will be the family $\{\mathcal{A}_S\}$ conditioned on the event that on queries q the answers are a_q.

Formally, for a set S, we denote by $\mathcal{A}_{q\cup S}|_{a_q}$ the distribution of the random element $A \in_R \mathcal{A}_{q\cup S}$, conditioned on the event $A_q = a_q$ (where we assume that the event $A_q = a_q$ occurs with non-zero probability). Since in $\mathcal{A}_{q\cup S}|_{a_q}$ we have that the coordinates indexed by q are fixed to a_q, we think of $\mathcal{A}_{q\cup S}|_{a_q}$, for simplicity of the notations, as a distribution over Σ^S, rather than over $\Sigma^{q\cup S}$, where $\Sigma = \mathbb{F}$ is the alphabet of the answers, (and note that in this distribution the coordinates indexed by $q \cap S$ are fixed to $a_{q\cap S}$). We will define the family of distributions $\{\mathcal{A}'_S\}$ by $\mathcal{A}'_S = \mathcal{A}_{q\cup S}|_{a_q}$.

We assume that for every distribution $\mathcal{A}_S$ (or $\mathcal{A}'_S$) in the family $\{\mathcal{A}_S\}$ (or $\{\mathcal{A}'_S\}$), every query in $S \cap D_\ell$ is answered by 0 with probability 1 (since the polynomial P_ℓ was just the 0 polynomial and was added to the PCP proof for simplicity of notations).

6.1 Proof of Lemma 6.1

Proof. Assume that V doesn't have soundness ϵ against (k_{max}, δ)-no-signaling strategies.

Thus, for some $x \notin \mathcal{L}$, there exists a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S\subset D,|S|\leq k_{max}}$ that fools V with probability larger than ϵ. That is, the verifier V accepts with probability $> \epsilon$, where on queries Q, the answers are given (probabilistically) by $A_Q \in_R \mathcal{A}_Q$.

Let Q be the set of queries chosen randomly by the verifier V and let Q' be the set of queries chosen independently by the verifier V'. Thus, Q, Q' are independent random variables. Let $A \in_R \mathcal{A}_{Q\cup Q'}$ be the (probabilistic) answers for the queries $Q \cup Q'$.

Let $V(Q, A_Q)$ be 1 if V accepts on queries Q and answers A_Q, and 0 otherwise. Let $V'(Q', A_{Q'})$ be 1 if V' accepts on queries Q' and answers $A_{Q'}$, and 0 otherwise. (We assume here that the sets of queries Q, Q' are ordered by the order that the queries were chosen

by the verifiers, so that the sets of queries also define which tests are performed on which queries). We denote by $V(Q, A_Q)$ also the event $V(Q, A_Q) = 1$, and in the same way we denote by $V'(Q', A_{Q'})$ also the event $V'(Q', A_{Q'}) = 1$.

Claim 6.2.

$$\Pr_{Q,Q'} \Pr_{A \in_R \mathcal{A}_{Q \cup Q'}} [V(Q, A_Q) \wedge \neg V'(Q', A_{Q'})] \leq 5\ell|\mathbb{F}| \cdot 2^{-r}$$

(where r is the parameter of the relaxed verifier V').

Proof. Recall that V and V' repeated every test k times, and that this gives a partition of all the tests performed by V and V' into groups, with exactly k tests in each group (see Section 5.3), and the number of groups for each verifier is smaller than $5\ell|\mathbb{F}|$.

Let $Q_{i,j}$ be the set of queries chosen randomly by V in order to perform the j^{th} test in the i^{th} group. Let $Q'_{i,j}$ be the set of queries chosen independently by V' in order to perform the j^{th} test in the i^{th} group. We think of $Q_{i,j}, Q'_{i,j}$ also as tests, rather than just sets of queries. All these tests are performed independently. That is, all the sets in $\{Q_{i,j}\}_{i,j} \cup \{Q'_{i,j}\}_{i,j}$ are independent, as random variables.

Let Q_i be the multiset of tests $\{Q_{i,j}\}_{j \in [k]}$ and let Q'_i be the multiset of tests $\{Q'_{i,j}\}_{j \in [k]}$.

Let $V(Q_i, A_{Q_i})$ be 1 if V accepts all the tests in Q_i (with answers A_{Q_i}), and 0 otherwise. Let $V'(Q'_i, A_{Q'_i})$ be 0 if V' rejects more than r tests in Q'_i (with answers $A_{Q'_i}$), and 1 otherwise. As before, we denote by $V(Q_i, A_{Q_i})$ also the event $V(Q_i, A_{Q_i}) = 1$, and in the same way we denote by $V'(Q'_i, A_{Q'_i})$ also as the event $V'(Q'_i, A_{Q'_i}) = 1$.

Note that if both $V(Q, A_Q)$ and $\neg V'(Q', A_{Q'})$ occur, then there exists i such that V accepts all the tests in Q_i while V' rejects more than r tests in Q'_i. Hence,

$$\Pr_{Q,Q',A}[V(Q, A_Q) \wedge \neg V'(Q', A_{Q'})] \leq \sum_i \Pr_{Q,Q',A}[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i})]$$

Thus, it remains to bound $\Pr[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i})]$ by 2^{-r}, for every i.

Fix i. Let $W_i = \{\{Q_{i,j}, Q'_{i,j}\}\}_{j \in [k]}$. That is, W_i is the partition of the multiset $Q_i \cup Q'_i$ into pairs $\{Q_{i,j}, Q'_{i,j}\}$, without specifying for each pair which test is $Q_{i,j}$ and which one is $Q'_{i,j}$. Note that we could have chosen Q_i, Q'_i by first choosing W_i and only then specifying which test in each pair is $Q_{i,j}$ and which one is $Q'_{i,j}$.

Let $r(W_i)$ be the number of pairs in W_i with at least one test that is rejected by the verifiers. Note that $r(W_i)$ is a random variable that depends on Q, Q', A, but conditioned on W_i it is independent of Q_i, Q'_i (that is, $r(W_i)$ is independent of the specification which test in each pair is $Q_{i,j}$ and which one is $Q'_{i,j}$).

We can now bound

$$\Pr_{Q,Q',A}[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i})] = \mathop{\mathbf{E}}_{W_i, r(W_i)} \left[\Pr_{Q,Q',A}[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i}) \mid W_i, r(W_i)] \right]$$

$\Pr[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i}) \mid W_i, r(W_i)]$ is bounded as follows:

If $r(W_i) < r$ then $[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i})]$ doesn't occur, because $\neg V'(Q'_i, A_{Q'_i})$ doesn't occur (because in order for $\neg V'(Q'_i, A_{Q'_i})$ to occur V' needs to reject at least r tests in Q'_i, which is impossible when $r(W_i) < r$). Hence,

$$\Pr[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i}) \mid W_i, (r(W_i) < r)] = 0$$

For $r(W_i) \geq r$,

$$\Pr[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i}) \mid W_i, (r(W_i) \geq r)] \leq \Pr[V(Q_i, A_{Q_i}) \mid W_i, (r(W_i) \geq r)] \leq 2^{-r},$$

where the second inequality follows because for each pair $\{Q_{i,j}, Q'_{i,j}\} \in W_i$, each test goes to Q_i with probability $1/2$ (independently at random), so the probability that Q_i gets none of the rejected tests is $\leq 2^{-r}$ (because when $r(W_i) \geq r$, there are at least r pairs with at least one rejected test in each pair).

We hence have

$$\Pr_{Q,Q',A}[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i})] = \mathop{\mathbf{E}}_{W_i, r(W_i)} \left[\Pr_{Q,Q',A}[V(Q_i, A_{Q_i}) \wedge \neg V'(Q'_i, A_{Q'_i}) \mid W_i, r(W_i)] \right]$$

$$\leq 2^{-r}$$

□

We will now proceed with the proof of Lemma 6.1. Recall that we assume that the δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S \subset D, |S| \leq k_{max}}$ fools V with probability larger than ϵ.

Recall that $\Sigma = \mathbb{F}$ is the alphabet of the answers. Recall that for a vector $a_Q \in \Sigma^Q$, we denote by $\mathcal{A}_{Q \cup Q'}|_{a_Q}$ the distribution of the random element $A \in_R \mathcal{A}_{Q \cup Q'}$, conditioned on the event $A_Q = a_Q$, (where we assume that the event $A_Q = a_Q$ is obtained with non-zero probability (otherwise we define $\mathcal{A}_{Q \cup Q'}|_{a_Q}$ to be an arbitrary fixed distribution)). Since in $\mathcal{A}_{Q \cup Q'}|_{a_Q}$ we have that the coordinates indexed by Q are fixed to a_Q, we think of $\mathcal{A}_{Q \cup Q'}|_{a_Q}$, for simplicity of the notations, as a distribution over $\Sigma^{Q'}$, rather than over $\Sigma^{Q \cup Q'}$ (and note that in this distribution the coordinates indexed by $Q \cap Q'$ are fixed to $a_{Q \cap Q'}$).

Since $\{\mathcal{A}_S\}$ is a δ-no-signaling family of distributions, the distributions of the following two random variables are δ-close:

- $A \in_R \mathcal{A}_{Q \cup Q'}$
- $\tilde{A}$, where the coordinates indexed by Q of $\tilde{A}$ are chosen by $\tilde{A}_Q \in_R \mathcal{A}_Q$, and the coordinates indexed by Q' of $\tilde{A}$ are chosen by $\tilde{A}_{Q'} \in_R \mathcal{A}_{Q \cup Q'}|_{\tilde{A}_Q}$ (and note that on $Q \cap Q'$ the vectors $\tilde{A}_Q, \tilde{A}_{Q'}$ always agree).

Therefore, by Claim 6.2,

$$5\ell|\mathbb{F}| \cdot 2^{-r} \geq$$

$$\Pr_{Q,Q'} \Pr_{A \in_R \mathcal{A}_{Q \cup Q'}} [V(Q, A_Q) \wedge \neg V'(Q', A_{Q'})] =$$

$$\mathop{\mathbf{E}}_{Q,Q'} \mathop{\mathbf{E}}_{A \in_R \mathcal{A}_{Q \cup Q'}} [V(Q, A_Q) \cdot (1 - V'(Q', A_{Q'}))] \geq$$

$$\mathop{\mathbf{E}}_{Q} \mathop{\mathbf{E}}_{Q'} \mathop{\mathbf{E}}_{\tilde{A}_Q \in_R \mathcal{A}_Q} \mathop{\mathbf{E}}_{\tilde{A}_{Q'} \in_R \mathcal{A}_{Q \cup Q'}|_{\tilde{A}_Q}} [V(Q, \tilde{A}_Q) \cdot (1 - V'(Q', \tilde{A}_{Q'}))] - \delta =$$

$$\mathop{\mathbf{E}}_{Q} \mathop{\mathbf{E}}_{\tilde{A}_Q \in_R \mathcal{A}_Q} \left[V(Q, \tilde{A}_Q) \cdot \mathop{\mathbf{E}}_{Q'} \mathop{\mathbf{E}}_{\tilde{A}_{Q'} \in_R \mathcal{A}_{Q \cup Q'}|_{\tilde{A}_Q}} [1 - V'(Q', \tilde{A}_{Q'})] \right] - \delta$$

That is,

$$\mathop{\mathbf{E}}_{Q} \mathop{\mathbf{E}}_{\tilde{A}_Q \in_R \mathcal{A}_Q} \left[V(Q, \tilde{A}_Q) \cdot \mathop{\mathbf{E}}_{Q'} \mathop{\mathbf{E}}_{\tilde{A}_{Q'} \in_R \mathcal{A}_{Q \cup Q'}|_{\tilde{A}_Q}} [1 - V'(Q', \tilde{A}_{Q'})] \right] \leq 5\ell|\mathbb{F}| \cdot 2^{-r} + \delta \tag{2}$$

The following claim shows that we can fix the values of Q and $\tilde{A}_Q$ to specific values q and $\tilde{a}_q$ that satisfy two desired properties. The first property will be used to show that V' is fooled with high probability. The second one will be used to show that the new family of distributions that we will construct is δ'-no-signaling.

Claim 6.3. *We can fix a set of queries q, and answers $\tilde{a}_q \in \Sigma^q$, such that:*

1. $$\mathop{\mathbf{E}}_{Q'} \mathop{\mathbf{E}}_{\tilde{A}_{Q'} \in_R \mathcal{A}_{q \cup Q'}|_{\tilde{a}_q}} [1 - V'(Q', \tilde{A}_{Q'})] \leq (5\ell|\mathbb{F}| \cdot 2^{-r} + \delta) \cdot \frac{2}{\epsilon}$$

2. $$\Pr_{\tilde{A}_q \in_R \mathcal{A}_q} (\tilde{A}_q = \tilde{a}_q) \geq \frac{\epsilon}{2 \cdot |\Sigma|^{|q|}}$$

Proof. Consider the conditional distribution of $(Q, \tilde{A}_Q) \mid V(Q, \tilde{A}_Q)$, that is, the distribution of $(Q, \tilde{A}_Q)$, where $\tilde{A}_Q \in_R \mathcal{A}_Q$, conditioned on the event $V(Q, \tilde{A}_Q)$. Fix $(q, \tilde{a}_q)$ randomly according to this distribution.

By Equation (2) and Markov inequality, and since

$$\Pr_{Q} \Pr_{\tilde{A}_Q \in_R \mathcal{A}_Q} V(Q, \tilde{A}_Q) > \epsilon,$$

the first part of the claim occurs with probability larger than $1/2$.

Since $\Pr_Q \Pr_{\tilde{A}_Q \in_R \mathcal{A}_Q} V(Q, \tilde{A}_Q) > \epsilon$ and since the number of possibilities for each $\tilde{a}_q$ is $|\Sigma|^{|q|}$, the second part of the claim occurs with probability larger than $1/2$. □

Fix $q, \tilde{a}_q$ from Claim 6.3. Define the family of distributions $\{\mathcal{A}'_S\}_{S \subset D, |S| \leq k'_{max}}$ by

$$\mathcal{A}'_S = \mathcal{A}_{q \cup S}|_{\tilde{a}_q}$$

(where, as before, $\mathcal{A}_{q\cup S}|_{\tilde{a}_q}$ is viewed as a distribution over Σ^S). Note also that $|q| \leq 6k\ell|\mathbb{F}|^2 = k_{max} - k'_{max}$.

By the first part of Claim 6.3,

$$\mathop{\mathbf{E}}_{Q'} \mathop{\mathbf{E}}_{A'_{Q'} \in_R \mathcal{A}'_{Q'}} V'(Q', A'_{Q'}) \geq 1 - (10\ell|\mathbb{F}| \cdot 2^{-r} + 2\delta)/\epsilon$$

That is, V' is fooled with probability of at least $1 - (10\ell|\mathbb{F}| \cdot 2^{-r} + 2\delta)/\epsilon$.

It remains to prove that $\{\mathcal{A}'_S\}$ is a δ'-no-signaling family of distributions.

Claim 6.4. *$\{\mathcal{A}'_S\}$ is a δ'-no-signaling family of distributions, where $\delta' = 8\delta|\Sigma|^{6k\ell|\mathbb{F}|^2}/\epsilon$.*

Proof. Let $S_1 \subset S_2 \subset D$, be such that $|S_2| \leq k'_{max}$. Denote by $(\mathcal{A}'_{S_2})_{S_1}$ and $(\mathcal{A}_{S_2})_{S_1}$ the projections of the distributions $\mathcal{A}'_{S_2}, \mathcal{A}_{S_2}$, respectively, on the coordinates in S_1.

We need to prove that the distributions $\mathcal{A}'_{S_1}$ and $(\mathcal{A}'_{S_2})_{S_1}$ are δ'-close. Without loss of generality, assume that $q \subseteq S_1$. Otherwise, just add q to both S_1, S_2 (this doesn't change the distance between the two distributions because, by the definition of $\mathcal{A}'_{S_1}, \mathcal{A}'_{S_2}$, we just added fixed coordinates to each of the two distribution).

Denote by $\mathcal{A}_{S_1}|_{\tilde{a}_q}$ the distribution of $A \in_R \mathcal{A}_{S_1}$ conditioned on the event $A_q = \tilde{a}_q$, and, in the same way, denote by $\mathcal{A}_{S_2}|_{\tilde{a}_q}$ the distribution of $A \in_R \mathcal{A}_{S_2}$ conditioned on the event $A_q = \tilde{a}_q$, and by $(\mathcal{A}_{S_2})_{S_1}|_{\tilde{a}_q}$ the distribution of $A \in_R (\mathcal{A}_{S_2})_{S_1}$ conditioned on the event $A_q = \tilde{a}_q$.

By the definitions, $\mathcal{A}'_{S_1} = \mathcal{A}_{S_1}|_{\tilde{a}_q}$, and $\mathcal{A}'_{S_2} = \mathcal{A}_{S_2}|_{\tilde{a}_q}$. Thus, we need to prove that $\mathcal{A}_{S_1}|_{\tilde{a}_q}$ and $(\mathcal{A}_{S_2}|_{\tilde{a}_q})_{S_1}$ are δ'-close. Since $(\mathcal{A}_{S_2}|_{\tilde{a}_q})_{S_1} = (\mathcal{A}_{S_2})_{S_1}|_{\tilde{a}_q}$, we need to prove that $\mathcal{A}_{S_1}|_{\tilde{a}_q}$ and $(\mathcal{A}_{S_2})_{S_1}|_{\tilde{a}_q}$ are δ'-close.

Since $\mathcal{A}$ is a δ-no-signaling family, $\mathcal{A}_{S_1}$ and $(\mathcal{A}_{S_2})_{S_1}$ are δ-close.

By the second part of Claim 6.3, and since $\mathcal{A}$ is a δ-no-signaling family, we have that

$$\Pr_{A \in_R \mathcal{A}_{S_1}}(A_q = \tilde{a}_q) \geq \frac{\epsilon}{2 \cdot |\Sigma|^{|q|}} - \delta$$

and

$$\Pr_{A \in_R (\mathcal{A}_{S_2})_{S_1}}(A_q = \tilde{a}_q) \geq \frac{\epsilon}{2 \cdot |\Sigma|^{|q|}} - \delta$$

The proof of the claim thus follows by Proposition 6.5, with $\mu = \mathcal{A}_{S_1}$, $\psi = (\mathcal{A}_{S_2})_{S_1}$, and

$$\alpha = \frac{\epsilon}{2 \cdot |\Sigma|^{|q|}} - \delta \geq \frac{\epsilon}{4 \cdot |\Sigma|^{|q|}} \geq \frac{\epsilon}{4 \cdot |\Sigma|^{6k\ell|\mathbb{F}|^2}}$$

□

Proposition 6.5. *Let δ, α be such that $0 < 2\delta < \alpha \leq 1$. Let $\mu, \psi : \Omega \to \mathbb{R}$ be two probability distributions over a finite set Ω, and assume that μ, ψ are δ-close. Let $E \subset \Omega$ be an event, such that, $\mu(E), \psi(E) \geq \alpha$. Denote by μ_E, ψ_E the conditional distributions μ, ψ, conditioned on the event E. Thus, $\mu_E, \psi_E : E \to \mathbb{R}$ are probability distributions over E.*

Then, μ_E, ψ_E are δ'-close, where $\delta' = 2\delta/\alpha$.

Proof. Denote by $\mu' : E \to \mathbb{R}$ and $\psi' : E \to \mathbb{R}$ the restrictions of μ, ψ to E. That is, for every $e \in E$, we have $\mu'(e) = \mu(e)$ and $\psi'(e) = \psi(e)$. Since μ, ψ are δ-close, $\|\mu' - \psi'\|_1 \leq \|\mu - \psi\|_1 \leq 2\delta$, where $\|\cdot\|_1$ denotes the l_1-norm.

Assume without loss of generality $\mu(E) \geq \psi(E)$. That is, $\frac{\psi(E)}{\mu(E)} \leq 1$.

Assume for a contradiction $\|\mu_E - \psi_E\|_1 > 2\delta'$. Then

$$\delta' < \frac{1}{2}\sum_{e\in E}|\mu_E(e) - \psi_E(e)| = \sum_{\{e|\mu_E(e)\geq\psi_E(e)\}}|\mu_E(e) - \psi_E(e)| \leq$$

$$\sum_{\{e|\mu_E(e)\geq\psi_E(e)\}}\left|\mu_E(e) - \frac{\psi(E)}{\mu(E)}\psi_E(e)\right| \leq \sum_{e\in E}\frac{1}{\mu(E)}\left|\mu(E)\cdot\mu_E(e) - \psi(E)\cdot\psi_E(e)\right|$$

$$\leq \frac{1}{\alpha}\sum_{e\in E}|\mu(E)\cdot\mu_E(e) - \psi(E)\cdot\psi_E(e)|$$

Since $\mu' = \mu(E)\cdot\mu_E$, and $\psi' = \psi(E)\cdot\psi_E$, we get $\delta' < 2\delta/\alpha$. □

This concludes the proof of Lemma 6.1. □

7 Soundness of V' in the Base PCP

In this section we will show that the verifier V' cannot be fooled to accept $x \notin \mathcal{L}$, with probability close to 1.

Recall that $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N$, is the security parameter of the PCP, and that $1 \leq r < k$ is the parameter of the relaxed verifier V'. Recall that ℓ and $|\mathbb{F}|$ are bounded by $\mathsf{polylog}(N)$.

We will prove the following lemma.

Lemma 7.1. *Assume that $k_{max} \geq 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$, where $s = O(s(n))$ is the maximal number of gates in a layer of the circuit $\mathcal{C}_n$. Assume that $\delta < \frac{1}{1000N\ell|\mathbb{F}|}$. Fix $\epsilon = \frac{1}{100N\ell|\mathbb{F}|}$, and note that $\epsilon > 10\max\left(\delta, \frac{2k}{|\mathbb{F}|^{m-2}}\right)$. Assume $r < \frac{k}{20\ell|\mathbb{F}|}$. Then, V' has soundness $1-\epsilon$ against (k_{max}, δ)-no-signaling strategies.*

The rest of the section is devoted for the proof of Lemma 7.1. From now on, through Section 7, fix s, δ, ϵ, r to be as in the statement of Lemma 7.1.

As for the parameter k_{max}, for the proof of Lemma 7.1, we will assume that $k_{max} \geq 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$. We will assume for a contradiction that for some $x \notin \mathcal{L}$, there exists a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S\subset D, |S|\leq k_{max}}$ that fools V' with probability larger than $1-\epsilon$. That is, the verifier V' accepts with probability $> 1-\epsilon$, where on queries Q, the answers are given (probabilistically) by $A \in_R \mathcal{A}_Q$ (see Section 7.7).

However, in most parts of Section 7, a much weaker requirement $k_{max} \geq 6k\ell|\mathbb{F}|^2$ will suffice. Hence, for the rest of the section we fix $k_{max} \geq 6k\ell|\mathbb{F}|^2$ and denote by $\{\mathcal{A}_S\}_{S\subset D,|S|\leq k_{max}}$ a δ-no-signaling family of distributions that makes V' accept x with probability $> 1-\epsilon$. The requirement that $k_{max} \geq 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$ will only be used in Section 7.7, by Lemma 7.36, Lemma 7.37 and by the proof of Lemma 7.1 (and the requirement will be noted therein).

Recall that we denote by D the domain of the PCP (that is, the alphabet of queries in the PCP). Recall that

$$D = D_X \cup D_0 \cup \ldots \cup D_\ell,$$

where $D_X = \mathbb{F}^m$ is viewed as the domain of X, and $D_0, \ldots, D_\ell$ are $\ell+1$ copies of $\mathbb{F}^\ell$, viewed as the domains of $P_0, \ldots, P_\ell$, respectively.

For a set $S \subset D, |S| \leq k_{max}$, we will view the answers $A \in_R \mathcal{A}_S$ as a function $A : S \to \mathbb{F}$. We can view A also as a partial function $A : D \to \mathbb{F}$, and we denote by $A_X, A_0, \ldots, A_\ell$ the restriction of that partial function to $D_X, D_0, \ldots, D_\ell$, respectively.

Recall that we assume that for every distribution $\mathcal{A}_S$ in the family $\{\mathcal{A}_S\}$, every query in $S \cap D_\ell$ is answered by 0 with probability 1 (since the polynomial P_ℓ was just the 0 polynomial and was added to the PCP proof for simplicity of notations).

7.1 Some Immediate Claims

Fix $k_{max}, s, \delta, \epsilon, r$ to be as in the statement of Lemma 7.1. Assume for a contradiction that for some $x \notin \mathcal{L}$, there exists a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S\subset D,|S|\leq k_{max}}$ that fools V' with probability larger than $1-\epsilon$.

We will start by stating an immediate corollary of the fact that $\{\mathcal{A}_S\}$ is a δ-no-signaling family.

Claim 7.2. *Let $S \subset D, |S| \leq k_{max}$ be a set generated by some random process. Let $A \in_R \mathcal{A}_S$. Let $f(S, A)$ be a predicate that is satisfied with probability p (where the probability is over S, A). Let S', Q, such that $S' \subseteq Q \subset D, |Q| \leq k_{max}$, be two sets generated by some random process, such that the distribution of S' is identical to the distribution of S. Let $A' \in_R \mathcal{A}_Q$. Then the probability that $f(S', A'_{S'})$ is satisfied is between $p-\delta$ and $p+\delta$, (where the probability is over S', Q, A').*

Proof. Since $\{\mathcal{A}_S\}$ is a δ-no-signaling family, for every fixed sets $s = s' \subseteq q$,

$$\Pr_{A\in_R\mathcal{A}_s}(f(s, A)) = \Pr_{A'\in_R\mathcal{A}_q}(f(s', A'_{s'})) \mp \delta$$

The claim follows by taking expectation over S on the left hand side and expectation over S', Q on the right hand side. □

Next we will state seven immediate corollaries of the fact that V' accepts with probability $> 1-\epsilon$. The following seven claims correspond to the seven different tests performed by the verifier V'. Each claim states that the corresponding test is satisfied with high probability.

Claim 7.3. Low Degree Test for X:
Let $L_1, \ldots, L_k : \mathbb{F} \to D_X$ be k random lines. Let $S = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1-\epsilon-\delta$, for at least $k-r$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Proof. Note that the set S can be extended to a set Q of queries of the verifier V', where Q is generated by the correct distribution of V', and $S \subset Q$ is the set of queries for the first test performed by V' (that is, the low degree test for X). Let $A' \in_R \mathcal{A}_Q$. Since V' accepts with probability $> 1-\epsilon$ on queries Q and answers A', and in particular this means that the first test of V' passes with probability $> 1-\epsilon$, we have that A' satisfies the claim with probability $> 1-\epsilon$. Formally:

With probability $> 1-\epsilon$, for at least $k-r$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A' \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_k, Q, A'$).

Since $\{\mathcal{A}_S\}$ is a δ-no-signaling family, by Claim 7.2, the same is satisfied for A, rather than A', with probability $> 1-\epsilon-\delta$, rather than $> 1-\epsilon$. □

Claim 7.4. Low Degree Test for P_i (fixed $L(0)_{i+1}$):
Let $i \in \{0, \ldots, \ell-1\}$. Let $u \in \mathbb{F}$. Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that, every line $L \in \{L_1, \ldots, L_k\}$ satisfies $L(0)_{i+1} = u$. Let $S = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1-\epsilon-\delta$, for at least $k-r$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Proof. Similar to the proof of Claim 7.3, using the second test performed by V' (the low degree test for P_i, type 1), rather than the first one. □

Claim 7.5. Low Degree Test for P_i (orthogonal to the $(i+1)^{th}$ coordinate):
Let $i \in \{0, \ldots, \ell-1\}$. Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that, every line $L \in \{L_1, \ldots, L_k\}$ is orthogonal to the $(i+1)^{th}$ coordinate. Let $S = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1-\epsilon-\delta$, for at least $k-r$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Proof. Similar to the proof of Claim 7.3, using the third test performed by V' (the low degree test for P_i, type 2), rather than the first one. □

Claim 7.6. Low Degree Test for P_i (fixed $L(0)_{i+1}$; orthogonal to the i^{th} coordinate):
Let $i \in \{1, \ldots, \ell-1\}$. Let $u \in \mathbb{F}$. Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that, every line $L \in \{L_1, \ldots, L_k\}$ is orthogonal to the i^{th} coordinate, and satisfies $L(0)_{i+1} = u$. Let $S = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1-\epsilon-\delta$, for at least $k-r$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Proof. Similar to the proof of Claim 7.3, using the fourth test performed by V' (the low degree test for P_i, type 3), rather than the first one. □

Claim 7.7. Low Degree Test for P_i (fixed $L(0)_i$; orthogonal to the $(i+1)^{th}$ coordinate):
Let $i \in \{1, \ldots, \ell - 1\}$. Let $u \in \mathbb{F}$. Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that, every line $L \in \{L_1, \ldots, L_k\}$ is orthogonal to the $(i+1)^{th}$ coordinate, and satisfies $L(0)_i = u$. Let $S = \{L_j(t)\}_{j \in [k], t \in \mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1 - \epsilon - \delta$, for at least $k - r$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Proof. Similar to the proof of Claim 7.3, using the fifth test performed by V' (the low degree test for P_i, type 4), rather than the first one. □

Claim 7.8. Sum Check for P_i:
Let $i \in \{1, \ldots, \ell\}$. Let $z_1, \ldots, z_k \in \mathbb{F}^\ell$ be k random points, where $z_j = (z_{j,1}, \ldots, z_{j,\ell}) \in \mathbb{F}^\ell$. Let S^i and S^{i-1} be two copies of the set of points $\{(z_{j,1}, \ldots, z_{j,i-1}, t, z_{j,i+1}, \ldots, z_{j,\ell})\}_{j \in [k], t \in \mathbb{F}} \subset \mathbb{F}^\ell$, and view S^i as a subset of D_i and S^{i-1} as a subset of D_{i-1}. Let $S = S^i \cup S^{i-1} \subset D$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1 - \epsilon - \delta$, for at least $k - r$ of the indices $j \in [k]$, the following is satisfied for every $t \in \mathbb{F}$:

$$A_i(z_{j,1}, \ldots, z_{j,i-1}, t, z_{j,i+1}, \ldots, z_{j,\ell}) = \sum_{h \in H} A_{i-1}(z_{j,1}, \ldots, z_{j,i-1}, h, z_{j,i+1}, \ldots, z_{j,\ell}) t^h$$

(where the probability is over $z_1, \ldots, z_k, A$).

Proof. Similar to the proof of Claim 7.3, using the sixth test performed by V' (the sum check for P_i), rather than the first one. □

Claim 7.9. Consistency of X and P_0:
Let $z_1, \ldots, z_k \in \mathbb{F}^\ell$ be k random points, where $z_j = (i_{j,1}, i_{j,2}, i_{j,3}, b_{j,1}, b_{j,2}, b_{j,3}) \in (\mathbb{F}^m)^3 \times \mathbb{F}^3 = \mathbb{F}^\ell$. Let $S^0 = \{z_j\}_{j \in [k]}$, viewed as a subset of D_0, and let $S^X = \{i_{j,1}, i_{j,2}, i_{j,3}\}_{j \in [k]}$, viewed as a subset of D_X. Let $S = S^0 \cup S^X \subset D$. Let $A \in_R \mathcal{A}_S$. Then, with probability $> 1 - \epsilon - \delta$, for at least $k - r$ of the points $z_j \in \{z_1, \ldots, z_k\}$, the following is satisfied:

$$A_0(z_j) = \hat{\phi}(z_j) \cdot (A_X(i_{j,1}) - b_{j,1}) \cdot (A_X(i_{j,2}) - b_{j,2}) \cdot (A_X(i_{j,3}) - b_{j,3})$$

(where the probability is over $z_1, \ldots, z_k, A$).

Proof. Similar to the proof of Claim 7.3, using the seventh test performed by V' (the consistency of X and P_0), rather than the first one. □

7.2 Additional Notation

Let $\ell' \geq 0$ be an integer. Let $M : \mathbb{F}^2 \to \mathbb{F}^{\ell'}$ be a plain. For every $t_1 \in \mathbb{F}$, denote by $M(t_1, *) : \mathbb{F} \to \mathbb{F}^{\ell'}$ the line $L : \mathbb{F} \to \mathbb{F}^{\ell'}$ defined by $L(t) = M(t_1, t)$. For every $t_2 \in \mathbb{F}$, denote by $M(*, t_2) : \mathbb{F} \to \mathbb{F}^{\ell'}$ the line $L : \mathbb{F} \to \mathbb{F}^{\ell'}$ defined by $L(t) = M(t, t_2)$.

Let $f : \mathbb{F}^2 \to \mathbb{F}$ be a function. For every $t_1 \in \mathbb{F}$, define $f_{(t_1,*)} : \mathbb{F} \to \mathbb{F}$ by $f_{(t_1,*)}(t) = f(t_1, t)$. For every $t_2 \in \mathbb{F}$, define $f_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ by $f_{(*,t_2)}(t) = f(t, t_2)$.

7.3 Consistency of P_0

We will now give a definition that will be central in the rest of the section. Intuitively, a point z satisfies property $\mathcal{Z}(\epsilon', r')$ if when taking k lines through it, with high probability, for most of these lines, the answers correspond to low degree polynomials that "evaluate" the point z to 0.

Definition 7.10. Property $\mathcal{Z}(\epsilon', r')$:
Let $\epsilon' \geq 0$ and $r' \geq 0$. Let $i \in \{0, \ldots, \ell\}$. Let $z \in D_i$.

Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that for every $L \in \{L_1, \ldots, L_k\}$, we have $L(0) = z$. Let $S = \{L_j(t)\}_{j \in [k], t \in \mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$.

Define $A^0 : S \to \mathbb{F}$ by $A^0(z') = A(z')$ for $z' \neq z$ and $A^0(z) = 0$.

We say that the point z satisfies property $\mathcal{Z}(\epsilon', r')$ (also denoted $z \in \mathcal{Z}(\epsilon', r')$) if with probability $\geq 1 - \epsilon'$, for at least $k - r'$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A^0 \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Our main lemma about property $\mathcal{Z}(\epsilon', r')$ is that the property is satisfied, with small ϵ' and r', for any point $z = (z_1, \ldots, z_\ell) \in D_0$, such that, $z_1, \ldots, z_\ell \in H$. (Intuitively, this is analogous to the formula $P_0|_{H^\ell} \equiv 0$, that is satisfied for $x \in \mathcal{L}$).

Lemma 7.11. *For any $z = (z_1, \ldots, z_\ell) \in D_0$, such that, $z_1, \ldots, z_\ell \in H$, we have $z \in \mathcal{Z}(\epsilon', r')$, where $\epsilon' = 8\ell|\mathbb{F}|\epsilon$, and $r' = 8\ell|\mathbb{F}|r$.*

The rest of Subsection 7.3 is devoted for the proof of Lemma 7.11.

7.3.1 Proof of Lemma 7.11

First, we define a variant of property $\mathcal{Z}(\epsilon', r')$, where the random lines are restricted to be orthogonal to the $(i')^{th}$ coordinate. (We will use this property only for $i' \in \{i, i+1\}$).

Definition 7.12. Property $\mathcal{Z}^{i'}(\epsilon', r')$:
Let $\epsilon' \geq 0$ and $r' \geq 0$. Let $i' \in \{1, \ldots, \ell\}$. Let $i \in \{0, \ldots, \ell\}$. Let $z \in D_i$.

Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that for every $L \in \{L_1, \ldots, L_k\}$, we have $L(0) = z$, and L is orthogonal to the $(i')^{th}$ coordinate. Let $S = \{L_j(t)\}_{j \in [k], t \in \mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$.

Define $A^0 : S \to \mathbb{F}$ by $A^0(z') = A(z')$ for $z' \neq z$ and $A^0(z) = 0$.

We say that the point z satisfies property $\mathcal{Z}^{i'}(\epsilon', r')$ (also denoted $z \in \mathcal{Z}^{i'}(\epsilon', r')$) if with probability $\geq 1-\epsilon'$, for at least $k-r'$ of the lines $L \in \{L_1, \ldots, L_k\}$, we have that $A^0 \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Lemma 7.11 will follow easily by Lemma 7.13, Lemma 7.16 and Lemma 7.17.

Lemma 7.13. *For every $\epsilon_1 \geq 0$, every $r_1 \geq 0$, every $i \in \{1, \ldots, \ell-1\}$, and every $z \in D_i$, if $z \in \mathcal{Z}^{i+1}(\epsilon_1, r_1)$ then $z \in \mathcal{Z}^i(\epsilon_2, r_2)$, where $\epsilon_2 = \epsilon_1 + 3|\mathbb{F}|\epsilon$, and $r_2 = r_1 + 2|\mathbb{F}|r$.*

Proof. Assume that $z \in \mathcal{Z}^{i+1}(\epsilon_1, r_1)$.

Let $L_1, \ldots, L_k : \mathbb{F} \to D_i$ be k random lines, such that for every $L \in \{L_1, \ldots, L_k\}$, we have $L(0) = z$, and L is orthogonal to the $(i+1)^{th}$ coordinate. Let $L'_1, \ldots, L'_k : \mathbb{F} \to D_i$ be k random lines, such that for every $L' \in \{L'_1, \ldots, L'_k\}$, we have $L'(0) = z$, and L' is orthogonal to the i^{th} coordinate.

Denote by E the event that for every $j \in [k]$, the lines L_j, L'_j are in general position, that is, the vectors $L_j(1) - L_j(0), L'_j(1) - L'_j(0)$ span a linear subspace of dimension 2 (as vectors in $D_i = \mathbb{F}^\ell$). Note that the event E occurs with probability of at least $1 - \frac{k \cdot 2}{|\mathbb{F}|^{\ell-2}}$.

Let $M_1, \ldots, M_k : \mathbb{F}^2 \to D_i$ be k plains, where $M_j(t_1, t_2) = L_j(t_1) + L'_j(t_2) - z$, (where the addition/substraction are over the vector space $D_i = \mathbb{F}^\ell$).

Let $S = \{M_j(t_1, t_2)\}_{j \in [k], t_1, t_2 \in \mathbb{F}} \subset D_i$. Let $A \in_R \mathcal{A}_S$. Define $A^0 : S \to \mathbb{F}$ by $A^0(z') = A(z')$ for $z' \neq z$ and $A^0(z) = 0$.

We say that M_j is *good* if the following is satisfied:

1. For every $t_1 \in \mathbb{F} \setminus \{0\}$, the function $A^0 \circ M_j(t_1, *) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.
2. For every $t_2 \in \mathbb{F}$, the function $A^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

By Proposition 7.14, (applied with $f = A^0 \circ M_j$ and $d = 2\ell|H|$), if M_j is good then $A^0 \circ L'_j = A^0 \circ M_j(0, *) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

Proposition 7.14. *Let $f : \mathbb{F}^2 \to \mathbb{F}$ be a function. Assume that for every $t_1 \in \mathbb{F} \setminus \{0\}$, the function $f_{(t_1,*)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$, and for every $t_2 \in \mathbb{F}$, the function $f_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$, where $d < |\mathbb{F}|$. Then, $f_{(0,*)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$.*

Proof. For every $t_2 \in \mathbb{F}$, the function $f_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$. Therefore, there exist $a_1, \ldots, a_d \in \mathbb{F}$, (where $a_1, \ldots, a_d$ are the Lagrange interpolation coefficients), such that for every $t_2 \in \mathbb{F}$, we have $f(0, t_2) = \sum_{t=1}^{d} a_t \cdot f(t, t_2)$. That is, $f_{(0,*)} = \sum_{t=1}^{d} a_t \cdot f_{(t,*)}$. Since $f_{(1,*)}, \ldots, f_{(d,*)}$ are univariate polynomials of degree $< d$, their linear combination $f_{(0,*)}$ is also a univariate polynomial of degree $< d$. □

We will show that with high probability, at least $k - r_2$ of the plains $M \in \{M_1, \ldots, M_k\}$ are good (where the probability is over $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$). By Proposition 7.14, this implies that with high probability, at least $k - r_2$ of the lines $L' \in \{L'_1, \ldots, L'_k\}$ satisfy that $A^0 \circ L' : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$).

Claim 7.15. *With probability $\geq 1 - \epsilon_1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta - \frac{2k}{|\mathbb{F}|^{\ell-2}}$, for at least $k - r_1 - 2|\mathbb{F}|r$ of the indices $j \in [k]$, we have that M_j is good.*

Proof. For every $t_1 \in \mathbb{F} \setminus \{0\}$, consider the set of lines $\{M_j(t_1, *)\}_{j\in[k]}$ and note that this is a set of k random lines in D_i, such that, every line $L \in \{M_j(t_1, *)\}_{j\in[k]}$ is orthogonal to the i^{th} coordinate, and satisfies $L(0)_{i+1} = z_{i+1}$. Hence, by Claim 7.6, using also Claim 7.2, with probability $> 1 - \epsilon - 2\delta$, for at least $k - r$ of the indices $j \in [k]$, we have that $A \circ M_j(t_1, *) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$. If, in addition, the event E occurs, then $A^0 \circ M_j(t_1, *) = A \circ M_j(t_1, *)$ and hence $A^0 \circ M_j(t_1, *) : \mathbb{F} \to \mathbb{F}$ is also a univariate polynomial of degree $< 2\ell|H|$.

For every $t_2 \in \mathbb{F} \setminus \{0\}$, consider the set of lines $\{M_j(*, t_2)\}_{j\in[k]}$ and note that this is a set of k random lines in D_i, such that, every line $L \in \{M_j(*, t_2)\}_{j\in[k]}$ is orthogonal to the $(i+1)^{th}$ coordinate, and satisfies $L(0)_i = z_i$. Hence, by Claim 7.7, using also Claim 7.2, with probability $> 1 - \epsilon - 2\delta$, for at least $k - r$ of the indices $j \in [k]$, we have that $A \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$. If, in addition, the event E occurs, then $A^0 \circ M_j(*, t_2) = A \circ M_j(*, t_2)$ and hence $A^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is also a univariate polynomial of degree $< 2\ell|H|$.

Consider the set of lines $\{M_j(*, 0)\}_{j\in[k]}$ and note that $M_j(*, 0) = L_j$. Since $z \in \mathcal{Z}^{i+1}(\epsilon_1, r_1)$, and using also Claim 7.2, with probability $\geq 1 - \epsilon_1 - \delta$, for at least $k - r_1$ of the indices $j \in [k]$, we have that $A^0 \circ M_j(*, 0) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

Recall also that the event E occurs with probability of at least $1 - \frac{2k}{|\mathbb{F}|^{\ell-2}}$.

Adding up all this, by the union bound, we obtain that with probability $\geq 1 - \epsilon_1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta - \frac{2k}{|\mathbb{F}|^{\ell-2}}$, for at least $k - r_1 - 2|\mathbb{F}|r$ of the indices $j \in [k]$, we have that:

1. For every $t_1 \in \mathbb{F} \setminus \{0\}$, $A^0 \circ M_j(t_1, *)$ is a univariate polynomial of degree $< 2\ell|H|$.
2. For every $t_2 \in \mathbb{F} \setminus \{0\}$, $A^0 \circ M_j(*, t_2)$ is a univariate polynomial of degree $< 2\ell|H|$.
3. $A^0 \circ M_j(*, 0)$ is a univariate polynomial of degree $< 2\ell|H|$.

That is, with probability $\geq 1 - \epsilon_1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta - \frac{2k}{|\mathbb{F}|^{\ell-2}}$, for at least $k - r_1 - 2|\mathbb{F}|r$ of the indices $j \in [k]$, we have that M_j is good. □

By Proposition 7.14, Claim 7.15 implies that with probability $\geq 1 - \epsilon_1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta - \frac{2k}{|\mathbb{F}|^{\ell-2}} > 1 - \epsilon_2 + \delta$, at least $k - r_2$ of the lines $L' \in \{L'_1, \ldots, L'_k\}$ satisfy that $A^0 \circ L' : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$ (where the probability is over $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$). Thus, using Claim 7.2, $z \in \mathcal{Z}^i(\epsilon_2, r_2)$.

This concludes the proof of Lemma 7.13. □

Lemma 7.16. *For every $\epsilon_1 \geq 0$, every $r_1 \geq 0$, every $i \in \{0, \ldots, \ell - 1\}$, and every $z \in D_i$, if $z \in \mathcal{Z}^{i+1}(\epsilon_1, r_1)$ then $z \in \mathcal{Z}(\epsilon_2, r_2)$, where $\epsilon_2 = \epsilon_1 + 3|\mathbb{F}|\epsilon$, and $r_2 = r_1 + 2|\mathbb{F}|r$.*

Proof. Similar to the proof of Lemma 7.13, except that we let $L'_1, \ldots, L'_k : \mathbb{F} \to D_i$ be k random lines, such that for every $L' \in \{L'_1, \ldots, L'_k\}$, we have $L'(0) = z$, (without the requirement that L' is orthogonal to the i^{th} coordinate), and we use Claim 7.4 and Claim 7.5, rather than Claim 7.6 and Claim 7.7, in the proof for the equivalent of Claim 7.15. □

Lemma 7.17. *Let $\epsilon_1 \geq 0$. Let $r_1 \geq 0$. Let $i \in \{1, \ldots, \ell\}$. Let $z = (z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$ be a point, such that, $z_i \in H$. For every $t \in \mathbb{F}$, let $z(t) = (z_1, \ldots, z_{i-1}, t, z_{i+1}, \ldots, z_\ell) \in \mathbb{F}^\ell$. Assume that for every $t \in \mathbb{F}$, the point $z(t)$, viewed as a point in D_i, satisfies property $\mathcal{Z}^i(\epsilon_1, r_1)$. Then the point z, viewed as a point in D_{i-1}, satisfies property $\mathcal{Z}^i(\epsilon_2, r_2)$, where $\epsilon_2 = \frac{\epsilon_1}{1-\gamma} + 2|\mathbb{F}|\epsilon$, and $r_2 = \frac{r_1}{1-\gamma} + |\mathbb{F}|r$, and $\gamma = \sqrt{\frac{|H|}{|\mathbb{F}|}}$.*

Proof. Assume that for every $t \in \mathbb{F}$, the point $z(t)$, viewed as a point in D_i, satisfies property $\mathcal{Z}^i(\epsilon_1, r_1)$.

Let $L_1, \ldots, L_k : \mathbb{F} \to \mathbb{F}^\ell$ be k random lines, such that for every $L \in \{L_1, \ldots, L_k\}$, we have $L(0) = 0$, and L is orthogonal to the i^{th} coordinate.

Denote by E the event that for every $j \in [k]$, the line L_j is in a general position (as a line in $\mathbb{F}^\ell$), that is, it's image is not a single point. Note that the event E occurs with probability of at least $1 - \frac{k}{|\mathbb{F}|^{\ell-1}}$.

Let $M_1, \ldots, M_k : \mathbb{F}^2 \to \mathbb{F}^\ell$ be k plains, where $M_j(t_1, t_2) = L_j(t_1) + z(t_2)$, (where the addition is over the vector space $\mathbb{F}^\ell$).

Let S^i and S^{i-1} be two copies of the set of points $\{M_j(t_1, t_2)\}_{j \in [k], t_1, t_2 \in \mathbb{F}} \subset \mathbb{F}^\ell$, and view S^i as a subset of D_i and S^{i-1} as a subset of D_{i-1}. Let $S = S^i \cup S^{i-1} \subset D$. Let $A \in_R \mathcal{A}_S$. Recall that we view A as a function $A : S \to \mathbb{F}$, and we denote by A_i, A_{i-1} the restriction of that function to S^i, S^{i-1}, respectively.

Define $A^0_i : S^i \to \mathbb{F}$ by $A^0_i(z') = A_i(z')$ for $z' \notin \{z(t)\}_{t \in \mathbb{F}}$, and $A^0_i(z') = 0$ for $z' \in \{z(t)\}_{t \in \mathbb{F}}$. Define $A^0_{i-1} : S^{i-1} \to \mathbb{F}$ by $A^0_{i-1}(z') = A_{i-1}(z')$ for $z' \notin \{z(t)\}_{t \in \mathbb{F}}$ and $A^0_{i-1}(z') = 0$ for $z' \in \{z(t)\}_{t \in \mathbb{F}}$.

We say that M_j is *good* if the following is satisfied:

1. For every $t_1 \in \mathbb{F}$, and every $t \in \mathbb{F}$,
$$A^0_i(M_j(t_1, t)) = \sum_{h \in H} A^0_{i-1}(M_j(t_1, h)) t^h$$
2. For at least $|H|$ values $t_2 \in \mathbb{F}$, the function $A^0_i \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

By Proposition 7.18, (applied with $f = A^0_i \circ M_j$, $f' = A^0_{i-1} \circ M_j$ and $d = 2\ell|H|$), if M_j is good then for every $t_2 \in H$, the function $A^0_{i-1} \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

Proposition 7.18. *Let $f : \mathbb{F}^2 \to \mathbb{F}$ and $f' : \mathbb{F}^2 \to \mathbb{F}$ be two functions. Assume that:*

1. *For every $t_1 \in \mathbb{F}$, and every $t \in \mathbb{F}$,*

$$f(t_1, t) = \sum_{h \in H} f'(t_1, h) t^h$$

2. *For at least $|H|$ values $t_2 \in \mathbb{F}$, the function $f_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$.*

Then, for every $t_2 \in H$, the function $f'_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$.

Proof. For every $h \in H$, present the function $f'_{(*,h)} : \mathbb{F} \to \mathbb{F}$ as a univariate polynomial (in the free variable y),

$$f'(y, h) = f'_{(*,h)}(y) = \sum_{s=0}^{|\mathbb{F}|-1} a_{h,s} \cdot y^s$$

where $a_{h,0}, \ldots, a_{h,|\mathbb{F}|-1} \in \mathbb{F}$. Thus, for every $y \in \mathbb{F}$, and every $t \in \mathbb{F}$,

$$f_{(*,t)}(y) = f(y, t) = \sum_{h \in H} f'(y, h) t^h = \sum_{h \in H} \sum_{s=0}^{|\mathbb{F}|-1} a_{h,s} \cdot y^s \cdot t^h = \sum_{s=0}^{|\mathbb{F}|-1} \left(\sum_{h \in H} a_{h,s} \cdot t^h \right) \cdot y^s$$

Assume for a contradiction that for some $s \geq d$, the polynomial $\sum_{h \in H} a_{h,s} \cdot t^h$ is not the identically 0 polynomial, and let s be the largest such index. Since $\sum_{h \in H} a_{h,s} \cdot t^h$ is not identically 0, and its degree is $\leq |H| - 1$, it gives 0 on at most $|H| - 1$ values of $t \in \mathbb{F}$. Hence, the polynomial $f_{(*,t)}(y)$ is of degree $< s$ for at most $|H| - 1$ values of $t \in \mathbb{F}$, which is a contradiction to the assumption that for at least $|H|$ values $t \in \mathbb{F}$, the function $f_{(*,t)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$.

Thus, for every $s \geq d$, the polynomial $\sum_{h \in H} a_{h,s} \cdot t^h$ is the identically 0 polynomial. That is, for every $s \geq d$ and every $h \in H$ we have $a_{h,s} = 0$. Hence, for every $h \in H$, the function $f'_{(*,h)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$. □

We will show that with high probability, at least $k - r_2$ of the plains $M \in \{M_1, \ldots, M_k\}$ are good (where the probability is over $L_1, \ldots, L_k, A$).

Claim 7.19. *With probability $\geq 1 - |\mathbb{F}|\epsilon - 2|\mathbb{F}|\delta - \frac{k}{|\mathbb{F}|^{\ell-1}} - \frac{\epsilon_1+\delta}{1-\gamma}$, for at least $k - |\mathbb{F}|r - \frac{r_1}{1-\gamma}$ of the indices $j \in [k]$, we have that M_j is good, where $\gamma = \sqrt{\frac{|H|}{|\mathbb{F}|}}$.*

Proof. First note that for $t_1 = 0$,

$$A_i^0(M_j(t_1, t)) = \sum_{h \in H} A_{i-1}^0(M_j(t_1, h)) t^h$$

is satisfied trivially (for every $j \in [k]$, and every $t \in \mathbb{F}$), since $A_i^0 \circ M_j(0,*)$ and $A_{i-1}^0 \circ M_j(0,*)$ are the identically 0 function (by the definitions).

For every $t_1 \in \mathbb{F} \setminus \{0\}$, consider the set of points $\{M_j(t_1,0)\}_{j\in[k]}$ and note that this is a set of k random points in $\mathbb{F}^\ell$, such that the i^{th} coordinate of each of these points is 0, (that is, all other coordinates of all these points are uniformly distributed and independent random variables). Note that in Claim 7.8, the i^{th} coordinate of each random point is ignored. Hence, by Claim 7.8, using also Claim 7.2, with probability $> 1-\epsilon-2\delta$, for at least $k-r$ of the indices $j \in [k]$, the following is satisfied for every $t \in \mathbb{F}$:

$$A_i(M_j(t_1,t)) = \sum_{h\in H} A_{i-1}(M_j(t_1,h))t^h$$

If, in addition, the event E occurs, then for every $t \in \mathbb{F}$, we have that, $A_i^0(M_j(t_1,t)) = A_i(M_j(t_1,t))$ and $A_{i-1}^0(M_j(t_1,t)) = A_{i-1}(M_j(t_1,t))$ and hence

$$A_i^0(M_j(t_1,t)) = \sum_{h\in H} A_{i-1}^0(M_j(t_1,h))t^h$$

(and recall that for $t_1 = 0$ this is satisfied trivially).

Recall that the event E occurs with probability of at least $1 - \frac{k}{|\mathbb{F}|^{\ell-1}}$.

Thus, by the union bound, with probability $> 1-|\mathbb{F}|\epsilon-2|\mathbb{F}|\delta-\frac{k}{|\mathbb{F}|^{\ell-1}}$, for at least $k-|\mathbb{F}|r$ of the indices $j \in [k]$, the following is satisfied for every $t_1 \in \mathbb{F}$ and every $t \in \mathbb{F}$:

$$A_i^0(M_j(t_1,t)) = \sum_{h\in H} A_{i-1}^0(M_j(t_1,h))t^h \tag{3}$$

For every $t_2 \in \mathbb{F}$, consider the set of lines $\{M_j(*,t_2)\}_{j\in[k]}$ and note that this is a set of k random lines, such that for every $L \in \{M_j(*,t_2)\}_{j\in[k]}$, we have $L(0) = z(t_2)$, and L is orthogonal to the i^{th} coordinate. Since $z(t_2)$, viewed as a point in D_i, satisfies property $\mathcal{Z}^i(\epsilon_1, r_1)$, and using also Claim 7.2, with probability $\geq 1-\epsilon_1-\delta$, for at least $k-r_1$ of the indices $j \in [k]$, we have that $A_i^0 \circ M_j(*,t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

Since this is true for every $t_2 \in \mathbb{F}$, by Proposition 7.20, applied with $\alpha = \epsilon_1 + \delta$, we obtain the following for any $\gamma < 1$:

with probability $\geq 1-\frac{\epsilon_1+\delta}{1-\gamma}$, for at least $\gamma|\mathbb{F}|$ values $t_2 \in \mathbb{F}$ we have that for at least $k-r_1$ of the indices $j \in [k]$, the function $A_i^0 \circ M_j(*,t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

Proposition 7.20. *Let $\{E_t\}_{t\in\mathbb{F}}$ be a set of events, such that, for every $t \in \mathbb{F}$, $\Pr(E_t) \geq 1-\alpha$. Then, for any $\gamma < 1$, with probability of at least $1-\frac{\alpha}{1-\gamma}$, at least $\gamma|\mathbb{F}|$ events in $\{E_t\}_{t\in\mathbb{F}}$ occur.*

Proof. Let I_t be the characteristic function of the event $\neg E_t$. Let $I = \sum_{t\in\mathbb{F}} I_t$. Thus, $\mathbf{E}[I] \leq \alpha|\mathbb{F}|$. By Markov's inequality, $\Pr[I > (1-\gamma)|\mathbb{F}|] < \alpha/(1-\gamma)$. Thus, with probability of at least $1-\alpha/(1-\gamma)$, at least $\gamma|\mathbb{F}|$ events in $\{E_t\}_{t\in\mathbb{F}}$ occur. □

Thus, with probability $\geq 1-\frac{\epsilon_1+\delta}{1-\gamma}$, for at least $\gamma|\mathbb{F}|$ values $t_2 \in \mathbb{F}$ we have that for at most r_1 of the indices $j \in [k]$, the function $A_i^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is not a univariate polynomial of degree $< 2\ell|H|$.

Since in a $\{0,1\}$-matrix with $\gamma|\mathbb{F}|$ rows and $[k]$ columns, with at most r_1 ones in each row, there are at most $\frac{\gamma|\mathbb{F}|r_1}{\gamma|\mathbb{F}|-|H|}$ columns with more than $\gamma|\mathbb{F}|-|H|$ ones (otherwise, the total number of ones is $> \gamma|\mathbb{F}|r_1$), this implies that with probability $\geq 1-\frac{\epsilon_1+\delta}{1-\gamma}$, for at most $\frac{\gamma|\mathbb{F}|r_1}{\gamma|\mathbb{F}|-|H|}$ indices $j \in [k]$ we have that for less than $|H|$ of the values $t_2 \in \mathbb{F}$ the function $A_i^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

Combined with Equation (3), by the union bound, with probability $> 1-|\mathbb{F}|\epsilon-2|\mathbb{F}|\delta-\frac{k}{|\mathbb{F}|^{\ell-1}}-\frac{\epsilon_1+\delta}{1-\gamma}$, for at least $k-|\mathbb{F}|r-\frac{\gamma|\mathbb{F}|r_1}{\gamma|\mathbb{F}|-|H|}$ of the indices $j \in [k]$, we have that:

1. For every $t_1 \in \mathbb{F}$ and every $t \in \mathbb{F}$:
$$A_i^0(M_j(t_1, t)) = \sum_{h \in H} A_{i-1}^0(M_j(t_1, h))t^h$$
2. For at least $|H|$ of the values $t_2 \in \mathbb{F}$ the function $A_i^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

That is, with probability $\geq 1-|\mathbb{F}|\epsilon-2|\mathbb{F}|\delta-\frac{k}{|\mathbb{F}|^{\ell-1}}-\frac{\epsilon_1+\delta}{1-\gamma}$, for at least $k-|\mathbb{F}|r-\frac{\gamma|\mathbb{F}|r_1}{\gamma|\mathbb{F}|-|H|}$ of the indices $j \in [k]$, we have that M_j is good. In particular, for $\gamma = \sqrt{\frac{|H|}{|\mathbb{F}|}}$, we have that with probability $\geq 1-|\mathbb{F}|\epsilon-2|\mathbb{F}|\delta-\frac{k}{|\mathbb{F}|^{\ell-1}}-\frac{\epsilon_1+\delta}{1-\gamma}$, for at least $k-|\mathbb{F}|r-\frac{r_1}{1-\gamma}$ of the indices $j \in [k]$, we have that M_j is good. □

By Proposition 7.18, Claim 7.19 implies that with probability $\geq 1-|\mathbb{F}|\epsilon-2|\mathbb{F}|\delta-\frac{k}{|\mathbb{F}|^{\ell-1}}-\frac{\epsilon_1+\delta}{1-\gamma} > 1-\epsilon_2+\delta$, at least $k-r_2$ of the indices $j \in [k]$ satisfy that for every $t_2 \in H$, the function $A_{i-1}^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$, (where the probability is over $L_1, \ldots, L_k, A$).

Fix $t_2 = z_i$. Consider the set of lines $\{M_j(*, t_2)\}_{j \in [k]}$ and note that this is a set of k random lines, such that for every $L \in \{M_j(*, t_2)\}_{j \in [k]}$, we have $L(0) = z(t_2) = z$, and L is orthogonal to the i^{th} coordinate.

With probability $> 1-\epsilon_2+\delta$, at least $k-r_2$ of the indices $j \in [k]$ satisfy that the function $A_{i-1}^0 \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$. Thus, using Claim 7.2, the point z, viewed as a point in D_{i-1}, satisfies property $\mathcal{Z}^i(\epsilon_2, r_2)$.

This concludes the proof of Lemma 7.17. □

Combining Lemma 7.17 and Lemma 7.13, we obtain the following lemma.

Lemma 7.21. *Let $\epsilon_1 \geq 0$. Let $r_1 \geq 0$. Let $i \in \{2, \ldots, \ell\}$. Let $z = (z_1, \ldots, z_\ell) \in \mathbb{F}^\ell$ be a point, such that, $z_i \in H$. For every $t \in \mathbb{F}$, let $z(t) = (z_1, \ldots, z_{i-1}, t, z_{i+1}, \ldots, z_\ell) \in \mathbb{F}^\ell$. Assume that for every $t \in \mathbb{F}$, the point $z(t)$, viewed as a point in D_i, satisfies property*

$\mathcal{Z}^i(\epsilon_1, r_1)$. *Then the point* z, *viewed as a point in* D_{i-1}, *satisfies property* $\mathcal{Z}^{i-1}(\epsilon_2, r_2)$, *where* $\epsilon_2 = \frac{\epsilon_1}{1-\gamma} + 5|\mathbb{F}|\epsilon$, *and* $r_2 = \frac{r_1}{1-\gamma} + 3|\mathbb{F}|r$, *and* $\gamma = \sqrt{\frac{|H|}{|\mathbb{F}|}}$.

Proof. Follows by applying Lemma 7.17 and then Lemma 7.13. □

We can now prove Lemma 7.11.

Proof. Recall that we assume that for every distribution $\mathcal{A}_S$ in the family $\{\mathcal{A}_S\}$, every query in $S \cap D_\ell$ is answered by 0 with probability 1 (since the polynomial P_ℓ was just the 0 polynomial and was added to the PCP proof for simplicity of notations). Therefore, any point $z \in D_\ell$, satisfies property $\mathcal{Z}^\ell(\epsilon_\ell, r_\ell)$, where $\epsilon_\ell = 0$ and $r_\ell = 0$.

By inductive application of Lemma 7.21, for any $i \in \{1, \ldots, \ell - 1\}$, any point $z = (z_1, \ldots, z_\ell) \in D_i$, such that, $z_{i+1}, \ldots, z_\ell \in H$, satisfies property $\mathcal{Z}^i(\epsilon_i, r_i)$, where $\epsilon_i = \frac{\epsilon_{i+1}}{1-\gamma} + 5|\mathbb{F}|\epsilon$ and $r_i = \frac{r_{i+1}}{1-\gamma} + 3|\mathbb{F}|r$, and $\gamma = \sqrt{\frac{|H|}{|\mathbb{F}|}}$.

In particular, any point $z = (z_1, \ldots, z_\ell) \in D_1$, such that, $z_2, \ldots, z_\ell \in H$, satisfies property $\mathcal{Z}^1(\epsilon_1, r_1)$, where $\epsilon_1 \leq \frac{5\ell|\mathbb{F}|\epsilon}{(1-\gamma)^\ell} < 6\ell|\mathbb{F}|\epsilon$ and $r_1 \leq \frac{3\ell|\mathbb{F}|r}{(1-\gamma)^\ell} < 6\ell|\mathbb{F}|r$.

Hence, by Lemma 7.17, any point $z = (z_1, \ldots, z_\ell) \in D_0$, such that, $z_1, \ldots, z_\ell \in H$, satisfies property $\mathcal{Z}^1(\epsilon_0, r_0)$, where $\epsilon_0 = \frac{\epsilon_1}{1-\gamma} + 2|\mathbb{F}|\epsilon < 7\ell|\mathbb{F}|\epsilon$, and $r_0 = \frac{r_1}{1-\gamma} + |\mathbb{F}|r < 7\ell|\mathbb{F}|r$.

Finally, by Lemma 7.16, any point $z = (z_1, \ldots, z_\ell) \in D_0$, such that, $z_1, \ldots, z_\ell \in H$, satisfies property $\mathcal{Z}(\epsilon', r')$, where $\epsilon' < 8\ell|\mathbb{F}|\epsilon$, and $r' < 8\ell|\mathbb{F}|r$. □

7.4 Consistency of X

In this subsection we will show that, intuitively, when taking a large number of lines through a point $z \in D_X$, with high probability, there exists a value $v \in \mathbb{F}$, such that for most of these lines, the answers correspond to low degree polynomials that "evaluate" the point z to v.

This is stated formally in Lemma 7.25, Lemma 7.27 and Lemma 7.28. The main goal of the subsection is to prove Lemma 7.27 and Lemma 7.28 (their statements could be read before reading the rest of the subsection). To prove these lemmas, we will need to first prove Lemma 7.22 and Lemma 7.25.

Lemma 7.22. *Let* $\epsilon' = 3|\mathbb{F}|\epsilon$. *Let* $r' = 2|\mathbb{F}|r$. *Let* $z \in D_X$. *Let* $L_1, \ldots, L_k, L'_1, \ldots, L'_k : \mathbb{F} \to D_X$ *be* $2k$ *random lines, such that for every* $L \in \{L_1, \ldots, L_k, L'_1, \ldots, L'_k\}$, *we have* $L(0) = z$. *Let* $S' = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \cup \{L'_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. *Let* $A \in_R \mathcal{A}_{S'}$.

For any $v \in \mathbb{F}$, *define* $A^v : S' \to \mathbb{F}$ *by* $A^v(z') = A(z')$ *for* $z' \neq z$ *and* $A^v(z) = v$.

Then, with probability $\geq 1 - \epsilon'$, *for at least* $k - r'$ *of the indices* $j \in [k]$, *there exists* $v \in \mathbb{F}$, *such that, both* $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ *and* $A^v \circ L'_j : \mathbb{F} \to \mathbb{F}$ *are univariate polynomials of degree* $< m|H|$ *(where the probability is over* $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$*).*

Proof. Let $L_1, \ldots, L_k, L'_1, \ldots, L'_k : \mathbb{F} \to D_X$ be $2k$ random lines, such that for every $L \in \{L_1, \ldots, L_k, L'_1, \ldots, L'_k\}$, we have $L(0) = z$.

Denote by E the event that for every $j \in [k]$, the lines L_j, L'_j are in general position, that is, the vectors $L_j(1) - L_j(0), L'_j(1) - L'_j(0)$ span a linear subspace of dimension 2 (as vectors in $D_X = \mathbb{F}^m$). Note that the event E occurs with probability of at least $1 - \frac{k \cdot 2}{|\mathbb{F}|^{m-2}}$.

Let $M_1, \ldots, M_k : \mathbb{F}^2 \to D_X$ be k plains, where $M_j(t_1, t_2) = L_j(t_1) + L'_j(t_2) - z$, (where the addition/substraction are over the vector space $D_X = \mathbb{F}^m$).

Let $S = \{M_j(t_1, t_2)\}_{j \in [k], t_1, t_2 \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$. For any $v \in \mathbb{F}$, define $A^v : S \to \mathbb{F}$ by $A^v(z') = A(z')$ for $z' \neq z$ and $A^v(z) = v$.

We say that M_j is *good* if the following is satisfied:

1. For every $t_1 \in \mathbb{F} \setminus \{0\}$, the function $A \circ M_j(t_1, *) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For every $t_2 \in \mathbb{F} \setminus \{0\}$, the function $A \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

By Proposition 7.23, (applied with $f = A \circ M_j$ and $d = m|H|$), if the event E occurs and M_j is good then there exists $v \in \mathbb{F}$, such that, $A^v \circ L_j = A^v \circ M_j(*, 0) : \mathbb{F} \to \mathbb{F}$ and $A^v \circ L'_j = A^v \circ M_j(0, *) : \mathbb{F} \to \mathbb{F}$ are both univariate polynomials of degree $< m|H|$.

Proposition 7.23. *Let $f : \mathbb{F}^2 \to \mathbb{F}$ be a function. Assume that for every $t_1 \in \mathbb{F} \setminus \{0\}$, the function $f_{(t_1,*)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$, and for every $t_2 \in \mathbb{F} \setminus \{0\}$, the function $f_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$, where $d < |\mathbb{F}|$. For any $v \in \mathbb{F}$, define $f^v : \mathbb{F}^2 \to \mathbb{F}$ by $f^v(t_1, t_2) = f(t_1, t_2)$ for $(t_1, t_2) \neq (0, 0)$ and $f^v(0, 0) = v$. Then, there exists $v \in \mathbb{F}$, such that, $f^v_{(0,*)} : \mathbb{F} \to \mathbb{F}$ and $f^v_{(*,0)} : \mathbb{F} \to \mathbb{F}$ are both univariate polynomials of degree $< d$.*

Proof. For every $t_2 \in \mathbb{F} \setminus \{0\}$, the function $f_{(*,t_2)} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< d$. Therefore, there exist $a_1, \ldots, a_d \in \mathbb{F}$, (where $a_1, \ldots, a_d$ are the Lagrange interpolation coefficients), such that for every $t_2 \in \mathbb{F} \setminus \{0\}$, we have $f(0, t_2) = \sum_{t=1}^{d} a_t \cdot f(t, t_2)$. Since $f^v(t_1, t_2) = f(t_1, t_2)$ for $(t_1, t_2) \neq (0, 0)$, this implies that for every $t_2 \in \mathbb{F} \setminus \{0\}$ and every $v \in \mathbb{F}$, we have $f^v(0, t_2) = \sum_{t=1}^{d} a_t \cdot f^v(t, t_2)$.

Let $v = \sum_{t=1}^{d} a_t \cdot f(t, 0)$. Since $f^v(0, 0) = v$, we now have for every $t_2 \in \mathbb{F}$ (including $t_2 = 0$), $f^v(0, t_2) = \sum_{t=1}^{d} a_t \cdot f^v(t, t_2)$. That is, $f^v_{(0,*)} = \sum_{t=1}^{d} a_t \cdot f^v_{(t,*)}$. Since $f^v_{(1,*)}, \ldots, f^v_{(d,*)}$ are identical to $f_{(1,*)}, \ldots, f_{(d,*)}$ and are hence univariate polynomials of degree $< d$, their linear combination $f^v_{(0,*)}$ is also a univariate polynomial of degree $< d$.

The proof now follows from Proposition 7.14, applied on the function f^v (with variables t_1, t_2 switched). □

We will show that with high probability, at least $k - r'$ of the plains $M \in \{M_1, \ldots, M_k\}$ are good (where the probability is over $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$). By Proposition 7.23,

this implies that with high probability, for at least $k - r'$ of the indices $j \in [k]$, there exists $v \in \mathbb{F}$, such that, $A^v \circ L_j = A^v \circ M_j(*, 0) : \mathbb{F} \to \mathbb{F}$ and $A^v \circ L'_j = A^v \circ M_j(0, *) : \mathbb{F} \to \mathbb{F}$ are both univariate polynomials of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$).

Claim 7.24. *With probability $\geq 1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta$, for at least $k - 2|\mathbb{F}|r$ of the indices $j \in [k]$, we have that M_j is good.*

Proof. For every $t_1 \in \mathbb{F} \setminus \{0\}$, consider the set of lines $\{M_j(t_1, *)\}_{j \in [k]}$ and note that this is a set of k random lines in D_X. Hence, by Claim 7.3, using also Claim 7.2, with probability $> 1 - \epsilon - 2\delta$, for at least $k - r$ of the indices $j \in [k]$, we have that $A \circ M_j(t_1, *) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

For every $t_2 \in \mathbb{F} \setminus \{0\}$, consider the set of lines $\{M_j(*, t_2)\}_{j \in [k]}$ and note that this is a set of k random lines in D_X. Hence, by Claim 7.3, using also Claim 7.2, with probability $> 1 - \epsilon - 2\delta$, for at least $k - r$ of the indices $j \in [k]$, we have that $A \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Adding up these facts, by the union bound, we obtain that with probability $\geq 1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta$, for at least $k - 2|\mathbb{F}|r$ of the indices $j \in [k]$, we have that:

1. For every $t_1 \in \mathbb{F} \setminus \{0\}$, $A \circ M_j(t_1, *) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For every $t_2 \in \mathbb{F} \setminus \{0\}$, $A \circ M_j(*, t_2) : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

That is, with probability $\geq 1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta$, for at least $k - 2|\mathbb{F}|r$ of the indices $j \in [k]$, we have that M_j is good. □

By Proposition 7.23, and since the event E occurs with probability of at least $1 - \frac{k \cdot 2}{|\mathbb{F}|^{m-2}}$ Claim 7.24 implies that with probability $\geq 1 - 2|\mathbb{F}|\epsilon - 4|\mathbb{F}|\delta - \frac{2k}{|\mathbb{F}|^{m-2}} > 1 - \epsilon' + \delta$, for at least $k - r'$ of the indices $j \in [k]$, there exists $v \in \mathbb{F}$, such that, $A^v \circ L_j = A^v \circ M_j(*, 0) : \mathbb{F} \to \mathbb{F}$ and $A^v \circ L'_j = A^v \circ M_j(0, *) : \mathbb{F} \to \mathbb{F}$ are both univariate polynomials of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_k, L'_1, \ldots, L'_k, A$). Thus, using Claim 7.2, Lemma 7.22 follows. □

Lemma 7.25. *Let $r' = 20|\mathbb{F}|r$. Let $\epsilon' = 4|\mathbb{F}|\epsilon$. Let $z \in D_X$. Let $L_1, \ldots, L_{2k} : \mathbb{F} \to D_X$ be $2k$ random lines, such that for every $L \in \{L_1, \ldots, L_{2k}\}$, we have $L(0) = z$. Let $S = \{L_j(t)\}_{j \in [2k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.*

For any $v \in \mathbb{F}$, define $A^v : S \to \mathbb{F}$ by $A^v(z') = A(z')$ for $z' \neq z$ and $A^v(z) = v$.

Then, with probability $\geq 1 - \epsilon'$, there exists $v \in \mathbb{F}$, such that, for at least $2k - r'$ of the indices $j \in [2k]$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_{2k}, A$).

Proof. Let $L_1, \ldots, L_{2k} : \mathbb{F} \to D_X$ be $2k$ random lines, such that for every $L \in \{L_1, \ldots, L_{2k}\}$, we have $L(0) = z$. Let $S = \{L_j(t)\}_{j\in[2k],t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$. For any $v \in \mathbb{F}$, define $A^v : S \to \mathbb{F}$ by $A^v(z') = A(z')$ for $z' \neq z$ and $A^v(z) = v$.

Denote by E the event that there exists $v \in \mathbb{F}$, such that, for at least $2k - r'$ of the indices $j \in [2k]$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$. We will show that $\Pr(E) \geq 1 - \epsilon'$, as needed (where the probability is over $L_1, \ldots, L_{2k}, A$).

Partition $L_1, \ldots, L_{2k}$ randomly into $L'_1, \ldots, L'_k$ and $L''_1, \ldots, L''_k$. Denote by E' the event that for at least $k - 2|\mathbb{F}|r$ of the indices $j \in [k]$, there exists $v \in \mathbb{F}$, such that, both $A^v \circ L'_j : \mathbb{F} \to \mathbb{F}$ and $A^v \circ L''_j : \mathbb{F} \to \mathbb{F}$ are univariate polynomials of degree $< m|H|$. By Lemma 7.22, $\Pr(E') \geq 1 - 3|\mathbb{F}|\epsilon$.

Claim 7.26. $\Pr(E' \mid \neg E) \leq 2^{-|\mathbb{F}|r/4}$

Proof. For every $v \in \mathbb{F}$, let J_v be the set of indices $j \in [2k]$, such that, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$. Note that for every $v \neq v' \in \mathbb{F}$, $J_v \cap J_{v'} = \emptyset$. If the event $\neg E$ occurs then for every $v \in \mathbb{F}$, $|J_v| < 2k - r'$. Denote by J the largest set J_v and by $\bar{J}$ the complement of J in $[2k]$. Thus, if the event $\neg E$ occurs then $|\bar{J}| > r'$.

Given the sets $\{J_v\}_{v\in\mathbb{F}}$, the probability that E' occurs is the probability that when partitioning $[2k]$ randomly into k pairs, for at least $k - 2|\mathbb{F}|r$ pairs the two indices in the pair are in the same set J_v. Assuming that $|\bar{J}| > r'$, this probability can be bounded by $2^{-|\mathbb{F}|r}$ by the following argument:

Choose the partition as follows: First choose randomly $k' = r'/2$ different indices $j_1, \ldots, j_{k'}$ in $\bar{J}$. Match the indices $j_1, \ldots, j_{k'}$ one by one, each to a random index in $[2k]$ that was still not chosen. Say that $j_t \in \{j_1, \ldots, j_{k'}\}$ is good if it was matched to an index in a set J_v such that $j_t \in J_v$. Finally, extend the partial partition randomly into a partition of $[2k]$ into k pairs. Note that the probability for an index j_t to be good is at most $\frac{k}{2k-r'} < 0.51$, independently of all previous choices of indices. Thus, the probability that at least $k' - 2|\mathbb{F}|r$ indices $j_t \in \{j_1, \ldots, j_{k'}\}$ are good is at most $k' \cdot \binom{k'}{2|\mathbb{F}|r} \cdot 0.51^{k'-2|\mathbb{F}|r} < 2^{-|\mathbb{F}|r/4}$.

Therefore, $\Pr(E' \mid \neg E) < 2^{-|\mathbb{F}|r/4}$. □

We can now bound,

$$1 - 3|\mathbb{F}|\epsilon \leq \Pr(E') \leq \Pr(E' \mid \neg E) + \Pr(E) < \Pr(E) + 2^{-|\mathbb{F}|r/4}$$

Thus,

$$\Pr(E) > 1 - 3|\mathbb{F}|\epsilon - 2^{-|\mathbb{F}|r/4} > 1 - 4|\mathbb{F}|\epsilon$$

□

Lemma 7.27. *Let $r' = 20|\mathbb{F}|r$. Let $\epsilon' = 5|\mathbb{F}|\epsilon$. Let $z \in D_X$. Let $L_1, \ldots, L_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L_1, \ldots, L_k\}$, we have $L(0) = z$. Let $S = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.*

For any $v \in \mathbb{F}$, define $A^v : S \to \mathbb{F}$ by $A^v(z') = A(z')$ for $z' \neq z$ and $A^v(z) = v$.

Then, with probability $\geq 1-\epsilon'$, there exists $v \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_k, A$).

Proof. Follows immediately by Lemma 7.25 and Claim 7.2. □

Lemma 7.28. *Let $r' = 40|\mathbb{F}|r$. Let $\epsilon' = 10|\mathbb{F}|\epsilon$. Let $z \in D_X$. Let $L_1, \ldots, L_{3k} : \mathbb{F} \to D_X$ be $3k$ random lines, such that for every $L \in \{L_1, \ldots, L_{3k}\}$, we have $L(0) = z$. Let $S = \{L_j(t)\}_{j \in [3k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.*

For any $v \in \mathbb{F}$, define $A^v : S \to \mathbb{F}$ by $A^v(z') = A(z')$ for $z' \neq z$ and $A^v(z) = v$.

Then, with probability $\geq 1-\epsilon'$, there exists $v \in \mathbb{F}$, such that, for at least $3k - r'$ of the indices $j \in [3k]$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $L_1, \ldots, L_{3k}, A$).

Proof. Let $L_1, \ldots, L_{3k} : \mathbb{F} \to D_X$ be $3k$ random lines, such that for every $L \in \{L_1, \ldots, L_{3k}\}$, we have $L(0) = z$. Let $S = \{L_j(t)\}_{j \in [3k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$. For any $v \in \mathbb{F}$, define $A^v : S \to \mathbb{F}$ by $A^v(z') = A(z')$ for $z' \neq z$ and $A^v(z) = v$.

Apply Lemma 7.25 twice: once on the set of lines $\{L_1, \ldots, L_{2k}\}$, and once on the set of lines $\{L_{k+1}, \ldots, L_{3k}\}$. By applying Lemma 7.25 twice, and using also Claim 7.2, we know that with probability $\geq 1-\epsilon'$, both of the following are satisfied:

1. There exists $v_1 \in \mathbb{F}$, such that, for at least $2k - 20|\mathbb{F}|r$ of the indices $j \in \{1, \ldots, 2k\}$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. There exists $v_2 \in \mathbb{F}$, such that, for at least $2k - 20|\mathbb{F}|r$ of the indices $j \in \{k+1, \ldots, 3k\}$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Note that if both of the above are satisfied then $v_1 = v_2$, since $2k - 20|\mathbb{F}|r > 3k/2$. Therefore, with probability $\geq 1-\epsilon'$, there exists $v \in \mathbb{F}$, such that, for at least $3k - 40|\mathbb{F}|r$ of the indices $j \in \{1, \ldots, 3k\}$, $A^v \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$. □

7.5 Consistency of X and P_0

Let $i_1, i_2, i_3 \in H^m$. Let $b_1, b_2, b_3 \in \{0,1\}$ be such that $\phi(i_1, i_2, i_3, b_1, b_2, b_3) = 1$, that is, the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in the 3-CNF formula φ.

In this subsection we will show that, intuitively, when taking a large number of lines through each of the points $i_1, i_2, i_3 \in D_X$, with high probability, there exist values $v_1, v_2, v_3 \in \mathbb{F}$, that satisfy $(v_1 - b_1) \cdot (v_2 - b_2) \cdot (v_3 - b_3) = 0$, and such that:

1. For most of the lines through i_1, the answers correspond to low degree polynomials that "evaluate" the point i_1 to v_1.
2. For most of the lines through i_2, the answers correspond to low degree polynomials that "evaluate" the point i_2 to v_2.

3. For most of the lines through i_3, the answers correspond to low degree polynomials that "evaluate" the point i_3 to v_3.

This is stated formally in Lemma 7.30. To prove this lemma, we will need to first prove Lemma 7.29.

Lemma 7.29. *Let $r' = 9\ell|\mathbb{F}|r$. Let $\epsilon' = 9\ell|\mathbb{F}|\epsilon$. Let $z = (i_1, i_2, i_3, b_1, b_2, b_3) \in (H^m)^3 \times H^3 = H^\ell \subset \mathbb{F}^\ell$. We view z as a point in D_0. We view $i_1, i_2, i_3 \in H^m \subset \mathbb{F}^m$ as points in D_X.*

Let $L_1, \ldots, L_k : \mathbb{F} \to D_0$ be k random lines, such that for every $L_j \in \{L_1, \ldots, L_k\}$, we have $L_j(0) = z$. For every $L_j \in \{L_1, \ldots, L_k\}$, the line L_j is a function $L_j : \mathbb{F} \to \mathbb{F}^\ell$. Let $L_j^1 : \mathbb{F} \to \mathbb{F}^m$ be L_j, restricted to coordinates $\{1, \ldots, m\}$. Let $L_j^2 : \mathbb{F} \to \mathbb{F}^m$ be L_j, restricted to coordinates $\{m+1, \ldots, 2m\}$. Let $L_j^3 : \mathbb{F} \to \mathbb{F}^m$ be L_j, restricted to coordinates $\{2m+1, \ldots, 3m\}$. We think of L_j^1, L_j^2, L_j^3 as lines $L_j^1, L_j^2, L_j^3 : \mathbb{F} \to D_X$, and note that $L_j^1(0) = i_1, L_j^2(0) = i_2, L_j^3(0) = i_3$.

Let $S^0 = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_0$. Let $S^X = \{L_j^1(t), L_j^2(t), L_j^3(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. Let $S = S^0 \cup S^X \subset D$. Let $A \in_R \mathcal{A}_S$.

Define $A_0^0 : S^0 \to \mathbb{F}$ by $A_0^0(z') = A_0(z')$ for $z' \neq z$ and $A_0^0(z) = 0$. For any $i \in D_X$ and $v \in \mathbb{F}$, define $A_X^{i\to v} : S^X \to \mathbb{F}$ by $A_X^{i\to v}(i') = A_X(i')$ for $i' \neq i$ and $A_X^{i\to v}(i) = v$.

Then, with probability $\geq 1 - \epsilon'$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, the following is satisfied (where the probability is over $L_1, \ldots, L_k, A$):

1. *$A_0^0 \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.*
2. *$A_X^{i_1\to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.*
3. *$A_X^{i_2\to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.*
4. *$A_X^{i_3\to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.*
5. *$\hat{\phi}(z) \cdot (v_1 - b_1) \cdot (v_2 - b_2) \cdot (v_3 - b_3) = 0$*

Proof. Let $L_1, \ldots, L_k : \mathbb{F} \to D_0$ be k random lines, such that for every $L_j \in \{L_1, \ldots, L_k\}$, we have $L_j(0) = z$. For every $j \in [k]$: Let $L_j^1 : \mathbb{F} \to \mathbb{F}^m$ be L_j, restricted to coordinates $\{1, \ldots, m\}$. Let $L_j^2 : \mathbb{F} \to \mathbb{F}^m$ be L_j, restricted to coordinates $\{m+1, \ldots, 2m\}$. Let $L_j^3 : \mathbb{F} \to \mathbb{F}^m$ be L_j, restricted to coordinates $\{2m+1, \ldots, 3m\}$. We view L_j^1, L_j^2, L_j^3 as $L_j^1, L_j^2, L_j^3 : \mathbb{F} \to D_X$.

Let $S^0 = \{L_j(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_0$. Let $S^X = \{L_j^1(t), L_j^2(t), L_j^3(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. Let $S = S^0 \cup S^X \subset D$. Let $A \in_R \mathcal{A}_S$.

Define $A_0^0 : S^0 \to \mathbb{F}$ by $A_0^0(z') = A_0(z')$ for $z' \neq z$ and $A_0^0(z) = 0$. For any $i \in D_X$ and $v \in \mathbb{F}$, define $A_X^{i\to v} : S^X \to \mathbb{F}$ by $A_X^{i\to v}(i') = A_X(i')$ for $i' \neq i$ and $A_X^{i\to v}(i) = v$.

Denote by E the event that for every $j \in [k]$, and every $w \in \{1, 2, 3\}$ the line L_j^w is in a general position (as a line in $\mathbb{F}^m$), that is, it's image is not a single point. Note that the event E occurs with probability of at least $1 - \frac{3k}{|\mathbb{F}|^{m-1}}$.

By Lemma 7.11, using Claim 7.2, with probability $\geq 1-8\ell|\mathbb{F}|\epsilon-\delta$, for at least $k-8\ell|\mathbb{F}|r$ of the indices $j \in [k]$, we have that $A_0^0 \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.

By Lemma 7.27, using Claim 7.2, with probability $\geq 1-5|\mathbb{F}|\epsilon-\delta$, there exists $v_1 \in \mathbb{F}$, such that, for at least $k-20|\mathbb{F}|r$ of the indices $j \in [k]$, we have that $A_X^{i_1 \to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

By Lemma 7.27, using Claim 7.2, with probability $\geq 1-5|\mathbb{F}|\epsilon-\delta$, there exists $v_2 \in \mathbb{F}$, such that, for at least $k-20|\mathbb{F}|r$ of the indices $j \in [k]$, we have that $A_X^{i_2 \to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

By Lemma 7.27, using Claim 7.2, with probability $\geq 1-5|\mathbb{F}|\epsilon-\delta$, there exists $v_3 \in \mathbb{F}$, such that, for at least $k-20|\mathbb{F}|r$ of the indices $j \in [k]$, we have that $A_X^{i_3 \to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

For every $t \in \mathbb{F} \setminus \{0\}$, consider the set of points $\{L_j(t)\}_{j\in[k]}$ and note that this is a set of k random points in D_0. Each point $L_j(t) \in \mathbb{F}^\ell$ can be written as

$$L_j(t) = (L_j^1(t), L_j^2(t), L_j^3(t), L_j(t)_{\ell-2}, L_j(t)_{\ell-1}, L_j(t)_\ell) \in (\mathbb{F}^m)^3 \times \mathbb{F}^3 = \mathbb{F}^\ell$$

where $L_j(t)_{\ell-2}, L_j(t)_{\ell-1}, L_j(t)_\ell$ are the last 3 coordinates of $L_j(t)$. By Claim 7.9, using also Claim 7.2, for every $t \in \mathbb{F} \setminus \{0\}$, with probability $> 1-\epsilon-2\delta$, for at least $k-r$ of the indices $j \in [k]$, we have

$$A_0(L_j(t)) = \hat{\phi}(L_j(t)) \cdot (A_X(L_j^1(t)) - L_j(t)_{\ell-2}) \cdot (A_X(L_j^2(t)) - L_j(t)_{\ell-1}) \cdot (A_X(L_j^3(t)) - L_j(t)_\ell)$$

If in addition the event E occurs, this implies that for every $v_1, v_2, v_3 \in \mathbb{F}$,

$$A_0^0(L_j(t)) = \hat{\phi}(L_j(t)) \cdot (A_X^{i_1 \to v_1}(L_j^1(t)) - L_j(t)_{\ell-2}) \cdot (A_X^{i_2 \to v_2}(L_j^2(t)) - L_j(t)_{\ell-1}) \cdot (A_X^{i_3 \to v_3}(L_j^3(t)) - L_j(t)_\ell)$$

Adding up all this, by the union bound, we obtain that with probability $\geq 1-\epsilon'$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k-r'$ of the indices $j \in [k]$, the following is satisfied (where the probability is over $L_1, \ldots, L_k, A$):

1. $A_0^0 \circ L_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< 2\ell|H|$.
2. $A_X^{i_1 \to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A_X^{i_2 \to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $A_X^{i_3 \to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
5. For every $t \in \mathbb{F} \setminus \{0\}$,

$$A_0^0(L_j(t)) =$$

$$\hat{\phi}(L_j(t)) \cdot (A_X^{i_1 \to v_1}(L_j^1(t)) - L_j(t)_{\ell-2}) \cdot (A_X^{i_2 \to v_2}(L_j^2(t)) - L_j(t)_{\ell-1}) \cdot (A_X^{i_3 \to v_3}(L_j^3(t)) - L_j(t)_\ell)$$

Note that since both sides of the equation are polynomials of degree $< |\mathbb{F}|$ in the variable t, the equation must be satisfied for $t = 0$ as well. Substituting $t = 0$, since $L_j(0) = z$, we have

$$0 = \hat{\phi}(z) \cdot (v_1 - b_1) \cdot (v_2 - b_2) \cdot (v_3 - b_3)$$

□

Lemma 7.30. *Let $r' = 9\ell|\mathbb{F}|r$. Let $\epsilon' = 9\ell|\mathbb{F}|\epsilon + \delta$. Let $i_1, i_2, i_3 \in H^m$. We view i_1, i_2, i_3 as points in D_X. Let $b_1, b_2, b_3 \in \{0,1\}$ be such that $\phi(i_1, i_2, i_3, b_1, b_2, b_3) = 1$, that is, the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in the 3-CNF formula φ.*

Let $L_1^1, \ldots, L_k^1 : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L_1^1, \ldots, L_k^1\}$, we have $L(0) = i_1$. Let $L_1^2, \ldots, L_k^2 : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L_1^2, \ldots, L_k^2\}$, we have $L(0) = i_2$. Let $L_1^3, \ldots, L_k^3 : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L_1^3, \ldots, L_k^3\}$, we have $L(0) = i_3$.

Let $S = \{L_j^1(t), L_j^2(t), L_j^3(t)\}_{j \in [k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i \to v} : S \to \mathbb{F}$ by $A^{i \to v}(i') = A(i')$ for $i' \neq i$ and $A^{i \to v}(i) = v$.

Then, with probability $\geq 1 - \epsilon'$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, the following is satisfied
(where the probability is over $L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3, A$):

1. *$A^{i_1 \to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.*
2. *$A^{i_2 \to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.*
3. *$A^{i_3 \to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.*
4. *$(v_1 - b_1) \cdot (v_2 - b_2) \cdot (v_3 - b_3) = 0$*

Proof. Follows immediately by Lemma 7.29, applied for the point $z = (i_1, i_2, i_3, b_1, b_2, b_3)$, and Claim 7.2. □

7.6 Property $\mathcal{R}(\epsilon', r')$

Recall that we have a (fanin 2) Boolean circuit $\mathcal{C}_n$ of size $N = O(t(n)s(n))$ that computes $\mathcal{L}$ on inputs of length n. The circuit $\mathcal{C}_n$ is layered, with $O(s(n))$ gates in each layer, such that a child of a gate in layer $i+1$ is either an input variable (or a negation of an input variable) or a gate in layer i. Recall that we assume that in the circuit $\mathcal{C}_n$ all negations are on input variables, and that the two children of any gate in the circuit are different.

Recall that the gates of the circuit are indexed by the numbers $1, \ldots, N$, in an order that agrees with the layers of the circuit. We assume that $1, \ldots, n$ are the indexes of the n input variables and $n+1, \ldots, 2n$ are the indexes of their negations, and that N is the index of the special output gate.

Recall that $\varphi(w_1, \ldots, w_N)$ is a 3-CNF Boolean formula, such that, $\varphi(w_1, \ldots, w_N) = 1$ if and only if $w_1, \ldots, w_N$ is the computation of the circuit $\mathcal{C}_n$ on the input $x = (x_1, \ldots, x_n)$, and $w_N = 1$. Denote by $x_1, \ldots, x_N$ the computation of the circuit $\mathcal{C}_n$ on the input $x = (x_1, \ldots, x_n)$. Thus, $\varphi(w_1, \ldots, w_N) = 1$ if and only if $(w_1, \ldots, w_N) = (x_1, \ldots, x_N)$, and $x_N = 1$.

Recall that since $N = |H|^m$, we identify $[N]$ and H^m by the lexicographic order on H^m, and view $w_1, \ldots, w_N$ and $x_1, \ldots, x_N$ as indexed by $i \in H^m$ (rather than $i \in [N]$). We hence view $x = (x_1, \ldots, x_N)$ as a function $x : H^m \to \{0,1\}$ (given by $x(i) = x_i$, where we identify $[N]$ and H^m).

Recall that $\phi : (H^m)^3 \times \{0,1\}^3 \to \{0,1\}$ is a function where $\phi(i_1, i_2, i_3, b_1, b_2, b_3) = 1$ if and only if the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in φ, and $\hat{\phi} : \mathbb{F}^\ell \to \mathbb{F}$ is the low-degree extension of ϕ.

We will now give a definition that will be central in the rest of the section. Intuitively, a subset $B \subset H^m \subset D_X$ satisfies property $\mathcal{R}(\epsilon', r')$ if when taking k lines through every point in B, with high probability, for every point $i \in B$, for most of the lines through the point i, the answers correspond to low degree polynomials that "evaluate" the point i to x_i.

To make sure that the property is well defined, we will limit ourselves to sets $B \subset H^m$ such that $k|B||\mathbb{F}| \leq k_{max}$. Since we identify H^m and $[N]$, we view each set B also as a subset of $[N]$. We will think of every set B also as a subset of D_X.

Let $\mathcal{B}$ be the set of all subsets $B \subset [N]$, such that, $k|B||\mathbb{F}| \leq k_{max}/2$.

Definition 7.31. Property $\mathcal{R}(\epsilon', r')$:
Let $\epsilon' \geq 0$ and $r' \geq 0$. Let $B \subset H^m$ be such that $k|B||\mathbb{F}| \leq k_{max}$. We view B as a subset of D_X.

For every $i \in B$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i \in B, j \in [k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i \to v} : S \to \mathbb{F}$ by $A^{i \to v}(i') = A(i')$ for $i' \neq i$ and $A^{i \to v}(i) = v$.

We say that the set B satisfies property $\mathcal{R}(\epsilon', r')$ (also denoted $B \in \mathcal{R}(\epsilon', r')$) if with probability $\geq 1 - \epsilon'$, for every $i \in B$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $\{L^i_j\}_{i \in B, j \in [k]}, A$).

We think of the empty set as satisfying $\mathcal{R}(\epsilon', r')$ for any ϵ', r'.

In all that comes below, we fix

$$r' = 9\ell|\mathbb{F}|r$$

Lemma 7.32. *Let $i \in [2n]$. Then, $\{i\} \in \mathcal{R}(\epsilon', r')$, where $\epsilon' = 10\ell|\mathbb{F}|\epsilon$, and $r' = 9\ell|\mathbb{F}|r$.*

Proof. We will give the proof for $i \in [n]$, such that, $x_i = 0$. The proof for $i \in [n]$, such that, $x_i = 1$, and for $i \in \{n+1, \ldots, 2n\}$ is similar.

Recall that for every $i \in [n]$, the formula φ contains a clause that checks that $w_i = x_i$. For example, if $x_i = 0$, we have the clause $(w_i = 0) \vee (w_i = 0) \vee (w_i = 0)$ that ensures that $w_i = 0$.

Let $L^1_1, \ldots, L^1_k, L^2_1, \ldots, L^2_k, L^3_1, \ldots, L^3_k : \mathbb{F} \to D_X$ be $3k$ random lines, such that for every line $L \in \{L^1_1, \ldots, L^1_k, L^2_1, \ldots, L^2_k, L^3_1, \ldots, L^3_k\}$, we have $L(0) = i$.

Let $S = \{L_j^1(t), L_j^2(t), L_j^3(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

By Lemma 7.30, with probability $\geq 1 - 9\ell|\mathbb{F}|\epsilon - \delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, the following is satisfied
(where the probability is over $L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3, A$):

1. $A^{i\to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{i\to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{i\to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $v_1 \cdot v_2 \cdot v_3 = 0$.

On the other hand, by Lemma 7.28, with probability $\geq 1 - 10|\mathbb{F}|\epsilon$, there exists $v \in \mathbb{F}$, such that, for at least $3k - 40|\mathbb{F}|r$ of the lines $L \in \{L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3\}$, $A^{i\to v} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Thus, by the union bound, with probability $\geq 1 - 9\ell|\mathbb{F}|\epsilon - 10|\mathbb{F}|\epsilon - \delta$, there exist $v_1, v_2, v_3, v \in \mathbb{F}$, such that, $v_1 \cdot v_2 \cdot v_3 = 0$, and

1. For at least $k - r'$ of the indices $j \in [k]$, $A^{i\to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For at least $k - r'$ of the indices $j \in [k]$, $A^{i\to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. For at least $k - r'$ of the indices $j \in [k]$, $A^{i\to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. For at least $3k - 40|\mathbb{F}|r$ of the lines $L \in \{L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3\}$, $A^{i\to v} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Since, $40|\mathbb{F}|r + r' < k$, this implies $v = v_1 = v_2 = v_3$, and hence $v = 0$.

Thus, with probability $\geq 1 - 9\ell|\mathbb{F}|\epsilon - 10|\mathbb{F}|\epsilon - \delta > 1 - \epsilon' + \delta$, for at least $k - r'$ of the lines $L \in \{L_1^1, \ldots, L_k^1\}$, $A^{i\to 0} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

The proof of the lemma hence follows by Claim 7.2. □

Lemma 7.33. *Let $i_1, i_2, i_3 \in [N]$ be such that the gate indexed by i_1 in the circuit $\mathcal{C}_n$ has children indexed by i_2, i_3. Let $B \in \mathcal{B}$ be such that $i_2, i_3 \in B$. Assume that $B \in \mathcal{R}(\epsilon', r')$, where $r' = 9\ell|\mathbb{F}|r$. Then $B \cup \{i_1\} \in \mathcal{R}(\epsilon'', r')$, where $\epsilon'' = \epsilon' + 9\ell|\mathbb{F}|\epsilon + 3\delta$.*

Proof. Let $B' = B \cup \{i_1\}$. For every $i \in B'$, let $L_1^i, \ldots, L_k^i : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L_1^i, \ldots, L_k^i\}$, we have $L(0) = i$.

Let $S = \{L_j^i(t)\}_{i\in B', j\in[k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

Since the gate indexed by i_1 in the circuit $\mathcal{C}_n$ has children indexed by i_2, i_3, the formula φ contains the clause $(w_{i_2} = x_{i_2}) \wedge (w_{i_3} = x_{i_3}) \to (w_{i_1} = x_{i_1})$.

By Lemma 7.30 and Claim 7.2, with probability $\geq 1-9\ell|\mathbb{F}|\epsilon-2\delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, the following is satisfied (where the probability is over $\{L^i_j(t)\}_{i\in B', j\in[k], t\in\mathbb{F}}, A$):

1. $A^{i_1\to v_1} \circ L^{i_1}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{i_2\to v_2} \circ L^{i_2}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{i_3\to v_3} \circ L^{i_3}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $(v_2 = x_{i_2}) \wedge (v_3 = x_{i_3}) \to (v_1 = x_{i_1})$

On the other hand, since the set B satisfies property $\mathcal{R}(\epsilon', r')$, using Claim 7.2, with probability $\geq 1 - \epsilon' - \delta$, for every $i \in B$: For at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Thus, by the union bound, with probability $\geq 1-\epsilon'-9\ell|\mathbb{F}|\epsilon-3\delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, $(v_2 = x_{i_2}) \wedge (v_3 = x_{i_3}) \to (v_1 = x_{i_1})$, and

1. For at least $k-r'$ of the indices $j \in [k]$, $A^{i_1\to v_1} \circ L^{i_1}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For at least $k-r'$ of the indices $j \in [k]$, $A^{i_2\to v_2} \circ L^{i_2}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. For at least $k-r'$ of the indices $j \in [k]$, $A^{i_3\to v_3} \circ L^{i_3}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. For every $i \in B$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Since, $r' + r' < k$, this implies $v_2 = x_{i_2}$, $v_3 = x_{i_3}$ and hence also $v_1 = x_{i_1}$.

Thus, with probability $\geq 1 - \epsilon' - 9\ell|\mathbb{F}|\epsilon - 3\delta$,

1. For at least $k-r'$ of the indices $j \in [k]$, $A^{i_1\to x_{i_1}} \circ L^{i_1}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For every $i \in B$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Thus $B' \in \mathcal{R}(\epsilon'', r')$ □

Lemma 7.34. *Let $B_1, B_2 \in \mathcal{B}$. If $B_1 \in \mathcal{R}(\epsilon_1, r')$ and $B_2 \in \mathcal{R}(\epsilon_2, r')$ then $B_1 \cup B_2 \in \mathcal{R}(\epsilon', r')$, where $\epsilon' = \epsilon_1 + \epsilon_2 + 2\delta$.*

Proof. Follows immediately by the union bound and (two applications of) Claim 7.2. □

Lemma 7.35. *Let $B_1, B_2 \in \mathcal{B}$. If $B_1 \subset B_2$ and $B_2 \in \mathcal{R}(\epsilon_2, r')$ then $B_1 \in \mathcal{R}(\epsilon_1, r')$, where $\epsilon_1 = \epsilon_2 + \delta$.*

Proof. Follows immediately by Claim 7.2. □

7.7 Proof of Lemma 7.1

Lemma 7.1 will be superseded by Lemma 10.1. We include its proof since: (1) it is simpler than the proof of Lemma 10.1, (2) it allows for a more modular proof and (3) what remains to be shown is relatively short.

For the rest of Section 7, we assume that $k_{max} \geq 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$. We assume for a contradiction that for some $x \notin \mathcal{L}$, there exists a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S \subset D, |S| \leq k_{max}}$ that fools V' with probability larger than $1 - \epsilon$. That is, the verifier V' accepts with probability $> 1-\epsilon$, where on queries Q, the answers are given (probabilistically) by $A \in_R \mathcal{A}_Q$.

For every $i \in \{2n, \ldots, N\}$, define $B_i \in \mathcal{B}$ as follows: B_i contains all the indexes $2n < i' \leq i$, such that, in the circuit $\mathcal{C}_n$, the gate indexed by i' is either in the same layer as the gate indexed by i, or in the previous layer. Note that $B_{2n} = \emptyset$ (this was added for the simplicity of the notation) and recall that we think of the empty set as satisfying $\mathcal{R}(\epsilon', r')$ for any ϵ'.

Lemma 7.36. *Assume that $k_{max} \geq 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$. Let $i \in \{2n+1, \ldots, N\}$. If $B_{i-1} \in \mathcal{R}(\epsilon', r')$, where $r' = 9\ell|\mathbb{F}|r$, then $B_i \in \mathcal{R}(\epsilon'', r')$, where $\epsilon'' = \epsilon' + 30\ell|\mathbb{F}|\epsilon$.*

Proof. Denote by i_1, i_2 the indexes of the two children of the gate indexed by i in the circuit $\mathcal{C}_n$. Note that $\{i_1, i_2\} \subset [2n] \cup B_{i-1}$. Denote $B' = \{i_1, i_2\} \cap [2n]$. Note also that $B_i \subseteq B_{i-1} \cup \{i\}$.

By Lemma 7.32 and Lemma 7.34, $B' \in \mathcal{R}(20\ell|\mathbb{F}|\epsilon + 2\delta, r')$.

Hence, by Lemma 7.34, $B' \cup B_{i-1} \in \mathcal{R}(\epsilon' + 20\ell|\mathbb{F}|\epsilon + 4\delta, r')$.

Hence, by Lemma 7.33, $B' \cup B_{i-1} \cup \{i\} \in \mathcal{R}(\epsilon' + 29\ell|\mathbb{F}|\epsilon + 7\delta, r')$.

Hence, by Lemma 7.35, $B_i \in \mathcal{R}(\epsilon' + 29\ell|\mathbb{F}|\epsilon + 8\delta, r')$. □

Lemma 7.37. *Assume that $k_{max} \geq 4sk|\mathbb{F}|+6k\ell|\mathbb{F}|^2$. Then, $B_N \in \mathcal{R}(\epsilon', r')$, where $r' = 9\ell|\mathbb{F}|r$ and $\epsilon' = 30N\ell|\mathbb{F}|\epsilon = 0.3$.*

Proof. Follows immediately by an inductive application of Lemma 7.36, and since $\epsilon = \frac{1}{100N\ell|\mathbb{F}|}$. □

Proof of Lemma 7.1

Proof. Let $r' = 9\ell|\mathbb{F}|r$.

Consider the point $N \in [N]$, viewed as a point in $H^m \subset D_X$. Recall that the formula φ contains a clause $(w_N = 1) \vee (w_N = 1) \vee (w_N = 1)$ that checks that $w_N = 1$.

Let $L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3 : \mathbb{F} \to D_X$ be $3k$ random lines, such that for every line $L \in \{L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3\}$, we have $L(0) = N$.

Let $S = \{L_j^1(t), L_j^2(t), L_j^3(t)\}_{j\in[k],t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $v \in \mathbb{F}$, define $A^{N\to v} : S \to \mathbb{F}$ by $A^{N\to v}(i') = A(i')$ for $i' \neq N$ and $A^{N\to v}(N) = v$.

By Lemma 7.30, with probability $\geq 1 - 9\ell|\mathbb{F}|\epsilon - \delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, the following is satisfied
(where the probability is over $L_1^1, \ldots, L_k^1, L_1^2, \ldots, L_k^2, L_1^3, \ldots, L_k^3, A$):

1. $A^{N\to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{N\to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{N\to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $(v_1 - 1) \cdot (v_2 - 1) \cdot (v_3 - 1) = 0$.

On the other hand, by (three applications of) Lemma 7.37 and Claim 7.2:

1. With probability $\geq 1 - 0.3 - 2\delta$, for at least $k - r'$ of the lines $L \in \{L_1^1, \ldots, L_k^1\}$, we have that $A^{N\to x_N} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. With probability $\geq 1 - 0.3 - 2\delta$, for at least $k - r'$ of the lines $L \in \{L_1^2, \ldots, L_k^2\}$, we have that $A^{N\to x_N} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. With probability $\geq 1 - 0.3 - 2\delta$, for at least $k - r'$ of the lines $L \in \{L_1^3, \ldots, L_k^3\}$, we have that $A^{N\to x_N} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Thus, by the union bound, with probability $> 0.1 - 9\ell|\mathbb{F}|\epsilon - 7\delta > 0$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, $(v_1 - 1) \cdot (v_2 - 1) \cdot (v_3 - 1) = 0$, and

1. For at least $k - r'$ of the indices $j \in [k]$, $A^{N\to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For at least $k - r'$ of the indices $j \in [k]$, $A^{N\to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. For at least $k - r'$ of the indices $j \in [k]$, $A^{N\to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

4. For at least $k - r'$ of the lines $L \in \{L_1^1, \ldots, L_k^1\}$, we have that $A^{N \to x_N} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

5. For at least $k - r'$ of the lines $L \in \{L_1^2, \ldots, L_k^2\}$, we have that $A^{N \to x_N} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

6. For at least $k - r'$ of the lines $L \in \{L_1^3, \ldots, L_k^3\}$, we have that $A^{N \to x_N} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Since, $r' + r' < k$, this implies $x_N = v_1 = v_2 = v_3$, and hence $x_N = 1$. Thus, the original input x is in the language $\mathcal{L}$. □

8 Soundness of V in the Base PCP

Lemma 8.1 will be superseded by Lemma 11.1. We include its proof for completeness.

Recall that $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N$, is the security parameter of the PCP, and that $1 \leq r < k$ is the parameter of the relaxed verifier V'. Recall that ℓ and $|\mathbb{F}|$ are bounded by $\mathsf{polylog}(N)$.

Lemma 8.1. *For a security parameter $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N$, fix the following parameters: Let $r = \frac{k}{40\ell|\mathbb{F}|}$. Let $\epsilon = 2^{-r/2}$. Let $k_{max} = 4sk|\mathbb{F}| + 12k\ell|\mathbb{F}|^2$, where $s = O(s(n))$ is the maximal number of gates in a layer of the circuit $\mathcal{C}_n$. Let $\delta = \frac{1}{|\mathbb{F}|^{8k\ell|\mathbb{F}|^2}}$. Then, V has soundness ϵ against (k_{max}, δ)-no-signaling strategies.*

Proof. Assume for a contradiction that V doesn't have soundness ϵ against (k_{max}, δ)-no-signaling strategies. By Lemma 6.1, since $\delta < \frac{\epsilon}{8 \cdot |\mathbb{F}|^{6k\ell|\mathbb{F}|^2}}$, we know that V' (with parameter r) doesn't have soundness $1 - \epsilon'$ against (k'_{max}, δ')-no-signaling strategies, where $k'_{max} = k_{max} - 6k\ell|\mathbb{F}|^2 = 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$, and $\delta' = 8\delta|\mathbb{F}|^{6k\ell|\mathbb{F}|^2}/\epsilon < \frac{1}{|\mathbb{F}|^{k\ell|\mathbb{F}|^2}}$, and $\epsilon' = (10\ell|\mathbb{F}|2^{-r} + 2\delta)/\epsilon < \frac{1}{100N\ell|\mathbb{F}|}$.

Hence V' (with parameter r) doesn't have soundness $1 - \epsilon'$ against (k'_{max}, δ')-no-signaling strategies, where $k'_{max} = 4sk|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$, and $\delta' = \frac{1}{|\mathbb{F}|^{k\ell|\mathbb{F}|^2}}$, and $\epsilon' = \frac{1}{100N\ell|\mathbb{F}|}$.

This contradicts Lemma 7.1. □

9 The Augmented PCP

In this section we describe the construction of the augmented PCP system, based on the base PCP system described in Section 5.

Let $\mathcal{L}$ be a language in $\mathsf{DTISP}(t(n), s(n))$, where $\mathsf{poly}(n) \leq t(n) \leq \mathsf{exp}(n)$ and $\max(n, \log(t(n))) \leq s(n) \leq t(n)$. Let x be an input of length n. Since $\mathcal{L} \in \mathsf{DTISP}(t(n), s(n))$, for any n there is a (fanin 2) Boolean circuit $\mathcal{C}_n$ of size $N = O(t(n)s(n))$ that computes $\mathcal{L}$ on inputs of

length n. Moreover, the circuit $\mathcal{C}_n$ is layered, with at most $t = O(t(n))$ layers that consist of $s = O(s(n))$ gates each. For simplicity, we think of the input variables $x_1, \ldots, x_n$ and their negations as being included in each layer of $\mathcal{C}_n$ (since $s \geq n$ this property can be achieved by increasing s by a constant factor)

We augment each circuit $\mathcal{C}_n$ to produce a circuit $\mathcal{C}'_n$ as follows.

Let $\mathbb{G}$ be a finite field of characteristic 2 and size $|\mathbb{G}| = \Theta(\log^2 s)$. Fix an arbitrary set $H_{\mathbb{G}} \subset \mathbb{G}$ of size $|H_{\mathbb{G}}| = \log s$ and a dimension $m_{\mathbb{G}} = \frac{\log s}{\log\log s}$ (such that $|H_{\mathbb{G}}|^{m_{\mathbb{G}}} = s$ and $m_{\mathbb{G}} \cdot |H_{\mathbb{G}}| < \frac{|\mathbb{G}|-1}{2}$). (For simplicity and without loss of generality we assume that $\log s$ and $\frac{\log s}{\log\log s}$ are integers). We construct a circuit $\mathcal{C}_{\mathsf{LDE}} : \{0,1\}^s \to \{0,1\}^{\mathsf{poly}(s)}$ by the following two-step process:

1. Given an input $\alpha \in \{0,1\}^s$, the circuit $\mathcal{C}_{\mathsf{LDE}}$ first computes the LDE $\hat{\alpha}$ of α w.r.t. $\mathbb{G}, H_{\mathbb{G}}, m_{\mathbb{G}}$. Recall that the polynomial $\hat{\alpha} : \mathbb{G}^{m_{\mathbb{G}}} \to \mathbb{G}$ is the (unique) individual degree $|H_{\mathbb{G}}|-1$ polynomial that agrees with α on $H_{\mathbb{G}}^{m_{\mathbb{G}}}$ (when α is interpreted as the truth table of a function $\alpha : H_{\mathbb{G}}^{m_{\mathbb{G}}} \to \{0,1\}$), see Section 4.6. Denote the output of this step by α'. We note that α' can be computed by a Boolean circuit of size $\mathsf{poly}(|\mathbb{G}|^{m_{\mathbb{G}}}) = \mathsf{poly}(s)$ and depth $O\big(m_{\mathbb{G}} \cdot \log(|\mathbb{G}|) + \log m_{\mathbb{G}} \cdot \mathsf{polylog}(|\mathbb{G}|)\big) = O(\log(s))$ (see Appendix A).

2. As its second (seemingly redundant) step, the circuit $\mathcal{C}_{\mathsf{LDE}}$ verifies that the restriction of α' to *every* line $L : \mathbb{G} \to \mathbb{G}^{m_{\mathbb{G}}}$ is a degree $m_{\mathbb{G}}|H_{\mathbb{G}}|$ univariate polynomial. That is, for every line L, the circuit $\mathcal{C}_{\mathsf{LDE}}$ checks that the function $\alpha' \circ L$ is a degree $m_{\mathbb{G}}|H_{\mathbb{G}}|$ univariate polynomial. For every such line L there is a corresponding output bit of $\mathcal{C}_{\mathsf{LDE}}$ that equals 1 if $\alpha' \circ L$ has low degree and 0 otherwise. (Indeed, if α' is in fact the LDE of α then every output bit of $\mathcal{C}_{\mathsf{LDE}}$ should have value 1.)

   We note that testing the degree of a univariate function $f : \mathbb{G} \to \mathbb{G}$ can be done by a Boolean circuit of size $\mathsf{poly}(|\mathbb{G}|) = \mathsf{polylog}(s)$ and depth $\mathsf{polylog}(|\mathbb{G}|) = \mathsf{polylog}(\log(s))$.

We denote by d the depth of $\mathcal{C}_{\mathsf{LDE}}$ and note that $d = O(\log s)$. We also note that the circuit $\mathcal{C}_{\mathsf{LDE}}$ has size $\mathsf{poly}(s)$ but if that size is smaller than $2^d \cdot s \cdot \log^5(t)$ then we (artificially) increase the size of $\mathcal{C}_{\mathsf{LDE}}$ to be $2^d \cdot s \cdot \log^5(t)$ while maintaining the depth d (by simply adding dummy gates).[13] We assume that $\mathcal{C}_{\mathsf{LDE}}$ contains no negation gates, but may contain *arbitrary* fan-in 2 Boolean gates. We also note that $\mathcal{C}_{\mathsf{LDE}}$ can be generated by a Turing machine in space $O(\log s)$.

The circuit $\mathcal{C}'_n$ is constructed by adding to $\mathcal{C}_n$ the computation of $\mathcal{C}_{\mathsf{LDE}}$ on *every* layer of $\mathcal{C}_n$. Thus, the circuit $\mathcal{C}'_n$ is composed of t layers, where each layer consists of the corresponding layer of $\mathcal{C}_n$ and the computation of $\mathcal{C}_{\mathsf{LDE}}$ on that layer. That is, the first layer consists of the first layer of $\mathcal{C}_n$ and the computation of $\mathcal{C}_{\mathsf{LDE}}$ on the first layer of $\mathcal{C}_n$ and for each $\mu \in \{2, \ldots, t\}$, the μ-th layer of $\mathcal{C}'_n$ consists of the computation of the μ-th layer of $\mathcal{C}_n$ from the $(\mu-1)$-th layer of $\mathcal{C}'_n$ and the computation of $\mathcal{C}_{\mathsf{LDE}}$ of the μ-th layer of $\mathcal{C}_n$. We denote the size of $\mathcal{C}'_n$ by N'. Recall that the input variables $x_1, \ldots, x_n$ and their negations are included

[13] This step can actually be avoided and we do it solely for convenience.

in each layer of $\mathcal{C}_n$. Thus, the layer μ of $\mathcal{C}'_n$ can be computed directly from layer $\mu-1$ of $\mathcal{C}'_n$. Note that $\mathcal{C}'_n$ has depth $t \cdot d$ and that

$$s \cdot t \cdot 2^d \cdot \log^5(t) \leq N' \leq \mathsf{poly}(N).$$

We call the gate indexed by N' the special output gate and note that its value represents the decision of whether $x \in \mathcal{L}$. We also assume without loss of generality that in the circuit $\mathcal{C}'_n$ all negations are on input variables, and that the two children of any gate in the circuit are different (this property can be achieved by duplicating each gate in the circuit twice, increasing the number of gates in each layer by a factor of 2). Note however that $\mathcal{C}'_n$ contains arbitrary fan-in 2 Boolean gates. Lastly, we note that there exists an $O(\log N')$ space Turing machine that on input $n \in \mathbb{N}$ outputs the circuit $\mathcal{C}'_n$.

For every layer $\mu \in [t]$ we denote by $\beta_\mu \subset [N']$ the set of indices of gates in $\mathcal{C}'_n$ that are associated with the LDE of the μ-th layer of $\mathcal{C}_n$. For $z \in \mathbb{G}^{m_\mathbb{G}}$, we denote by $\beta_\mu[z] \subset \beta_\mu$ the set of indices of the $\log_2|\mathbb{G}|$ gates associated with the point z in the computation of the LDE of layer μ in $\mathcal{C}'_n$. For a sequence of indices $Z \subset \mathbb{G}^{m_\mathbb{G}}$ we denote by $\beta_\mu[Z] \stackrel{\text{def}}{=} \cup_{z \in Z} \beta_\mu(z)$.

We construct the formulas $\varphi, \varphi_\mathcal{C}, \varphi_x$, as well as the parameters H, $\mathbb{F}$, m and ℓ, exactly as in Section 5 but with respect to the circuit $\mathcal{C}'_n$ (of size N') rather than $\mathcal{C}_n$ (of size N). We also construct the corresponding functions $\phi, \phi_{\mathcal{C}'}, \phi_x$ and $\hat{\phi}, \hat{\phi}_{\mathcal{C}'}, \hat{\phi}_x$ as in Section 5.

In addition, we construct a formula $\varphi_{extra}(w_1, \ldots, w_{N'})$ as follows. For every $i \in [N']$, the formula φ_{extra} contains a (seemingly redundant) clause that verifies that w_i has a *Boolean* value. Additionally, for every $\mu \in [t]$, we add to φ_{extra} clauses that verify that each one of the output gates of the corresponding $\mathcal{C}_{\mathsf{LDE}}$ circuit of layer μ has value 1. In other words,

$$\varphi_{extra}(w_1, \ldots, w_{N'}) = \bigwedge_{i \in [N']} \big((w_i = 0) \vee (w_i = 0) \vee (w_i = 1)\big) \wedge \bigwedge_{i \text{ is output gate of } \mathcal{C}_{\mathsf{LDE}}} \big((w_i = 1) \vee (w_i = 1) \vee (w_i = 1)\big).$$

Let $\phi_{extra} : (H^m)^3 \times \{0,1\}^3 \to \{0,1\}$ be the function where $\phi_{extra}(i_1, i_2, i_3, b_1, b_2, b_3) = 1$ if and only if the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in φ_{extra}. Extend ϕ_{extra} to be a function $\phi_{extra} : H^{3m+3} \to \{0,1\}$ by setting it to be 0 for inputs outside of $H^{3m} \times \{0,1\}^3$. Let $\hat{\phi}_{extra} : \mathbb{F}^\ell \to \mathbb{F}$ be the low-degree extension of ϕ_{extra}.

Since there is an $O(\log N')$ space deterministic Turing machine that on input n outputs φ_{extra}, by Proposition 4.2, the function $\hat{\phi}_{extra}$ can be computed in $O(\log N')$ space.

Let $\varphi' = \varphi \wedge \varphi_{extra}$ and let $\phi' : (H^m)^3 \times \{0,1\}^3 \to \{0,1\}$ be the function where $\phi'(i_1, i_2, i_3, b_1, b_2, b_3) = 1$ if and only if the clause $(w_{i_1} = b_1) \vee (w_{i_2} = b_2) \vee (w_{i_3} = b_3)$ appears in φ'. Extend ϕ' to be a function $\phi' : H^{3m+3} \to \{0,1\}$ by setting it to be 0 for inputs outside of $H^{3m} \times \{0,1\}^3$. Let $\hat{\phi}' : \mathbb{F}^\ell \to \mathbb{F}$ be the low-degree extension of ϕ'.

Since the sets of clauses of φ and φ_{extra} are disjoint, we have $\hat{\phi}' = \hat{\phi} + \hat{\phi}_{extra} = \hat{\phi}_x + \hat{\phi}_{\mathcal{C}'} + \hat{\phi}_{extra}$.

The PCP proof (i.e., the polynomials $X, P_0, \ldots, P_\ell$) is constructed exactly as in Section 5.1 except that we use the circuit $\mathcal{C}'_n$ and the formula φ' (rather than $\mathcal{C}_n$ and φ). The PCP verifier

V (resp., the relaxed verifier V') is constructed exactly as in Section 5.2 (resp., Section 5.3) with respect to the new PCP proof.

10 Soundness of V' in the Augmented PCP

In this section we will show that the relaxed verifier V' cannot be fooled to accept $x \notin \mathcal{L}$, with probability close to 1.

Recall that $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N'$, is the security parameter of the PCP, and that $1 \leq r < k$ is the parameter of the relaxed verifier V'. Recall that ℓ and $|\mathbb{F}|$ are bounded by $\mathsf{polylog}(N')$.

Recall that t is the depth of the (original) circuit $\mathcal{C}_n$ and that $d = O(\log s)$ is the depth of the circuit $\mathcal{C}_{\mathsf{LDE}}$. We will prove the following lemma.

Lemma 10.1. *Assume that* $k_{max} \geq k\mathsf{polylog}(s)\log(t)|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$. *Assume that* $\delta < \frac{1}{1000N'\ell|\mathbb{F}|}$. *Fix* $\epsilon = \frac{1}{100N'\ell|\mathbb{F}|}$, *and note that* $\epsilon > 10\max\left(\delta, \frac{2k}{|\mathbb{F}|^{m-2}}\right)$. *Assume* $r < \frac{k}{20\ell|\mathbb{F}|}$. *Then,* V' *has soundness* $1-\epsilon$ *against* (k_{max}, δ)*-no-signaling strategies.*

The rest of the section is devoted to the proof of Lemma 10.1. From now on, through Section 10, fix $k_{max}, \delta, \epsilon, r$ to be as in the statement of Lemma 10.1 and fix

$$r' = 9\ell|\mathbb{F}|r$$

(note that $r' < k/2$). We also fix a parameter

$$\nu = 10(\log(t) + d).$$

We will assume for a contradiction that for some $x \notin \mathcal{L}$, there exists a δ-no-signaling family of distributions $\{\mathcal{A}_S\}_{S \subset D, |S| \leq k_{max}}$ that fools V' with probability larger than $1-\epsilon$. That is, the verifier V' accepts with probability $> 1-\epsilon$, where on queries Q, the answers are given (probabilistically) by $A \in_R \mathcal{A}_Q$.

10.1 Reading Multiple Points Together

Let $B \subseteq H^m$ and let $\alpha : B \to \{0,1\}$. We think of B as specifying a subset of the variables of the formula φ' and of α as an assignment to B. We say that α is *consistent* with respect to B if it satisfies all the clauses of φ' in which *only* variables in B appear.

Lemma 10.2. *Let* $B \subseteq H^m$ *such that* $3k|\mathbb{F}||B| < k_{max}$.

For every $i \in B$, *let* $L_1^i, \ldots, L_k^i : \mathbb{F} \to D_X$ *be* k *random lines, such that for every* $L \in \{L_1^i, \ldots, L_k^i\}$, *we have* $L(0) = i$. *Let* $S = \{L_j^i(t)\}_{i \in B, j \in [k], t \in \mathbb{F}} \subset D_X$. *Let* $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ *and* $v \in \mathbb{F}$, *define* $A^{i \to v} : S \to \mathbb{F}$ *by* $A^{i \to v}(i') = A(i')$ *for* $i' \neq i$ *and* $A^{i \to v}(i) = v$.

Then, with probability $\geq 1 - 200|B|^3\ell|\mathbb{F}|\epsilon$*, there exists* $\alpha : B \to \{0,1\}$ *that is consistent with respect to* B*, such that for every* $i \in B$*, for at least* $k - r'$ *of the indices* $j \in [k]$*, it holds that* $A^{i\to\alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ *is a univariate polynomial of degree* $< m|H|$ *(where the probability is over* $\{L^i_j\}_{i\in B, j\in[k]}, A$*).*

Proof. For every $i \in B$, let $L^i_1, \ldots, L^i_k, L^i_{k+1}, \ldots, L^i_{2k}, L^i_{2k+1}, \ldots, L^i_{3k} : \mathbb{F} \to D_X$ be $3k$ random lines, such that for every $L \in \left\{L^i_j\right\}_{j\in[3k]}$, we have $L(0) = i$. Let $S = \{L^i_j(t)\}_{i\in B, j\in[3k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

By Lemma 7.28, using also Claim 7.2, for every $i \in B$, with probability $\geq 1 - 10|\mathbb{F}|\epsilon - \delta$, there exists $v_i \in \mathbb{F}$, such that, for at least $3k - r'$ of the indices $j \in [3k]$, $A^{i\to v_i} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $\{L^i_j\}_{i\in B, j\in[3k]}, A$).[14]

Let $\psi = (w_{i_1} = b_1 \vee w_{i_2} = b_2 \vee w_{i_3} = b_3)$ be a clause in φ' such that $i_1, i_2, i_3 \in B$. By Lemma 7.30 (using also Claim 7.2), with probability $\geq 1 - 9\ell|\mathbb{F}|\epsilon - 2\delta$ there exist $v^{(\psi)}_1, v^{(\psi)}_2, v^{(\psi)}_3 \in \mathbb{F}$ such that for at least $k - r'$ of the indices $j \in [k]$ it holds that:

1. $A^{i_1\to v^{(\psi)}_1} \circ L^{i_1}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{i_2\to v^{(\psi)}_2} \circ L^{i_2}_{k+j} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{i_3\to v^{(\psi)}_3} \circ L^{i_3}_{2k+j} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $(v^{(\psi)}_1 - b_1) \cdot (v^{(\psi)}_2 - b_2) \cdot (v^{(\psi)}_3 - b_3) = 0$.

By the union bound, and since B contains at most $8|B|^3$ clauses, with probability $\geq 1 - |B|(10|\mathbb{F}|\epsilon + \delta) - 8|B|^3(9\ell|\mathbb{F}|\epsilon + 2\delta) > 1 - 100|B|^3\ell|\mathbb{F}|\epsilon$, for every $i \in B$ there exists $v_i \in \mathbb{F}$ and for every clause $\psi = (w_{i_1} = b_1 \vee w_{i_2} = b_2 \vee w_{i_3} = b_3)$ in φ' that contains only variables from B, there exist $v^{(\psi)}_1, v^{(\psi)}_2, v^{(\psi)}_3 \in F$ such that:

1. For at least $3k - r'$ of the indices $j \in [3k]$, $A^{i\to v_i} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For at least $k - r'$ of the indices $j \in [k]$,
   (a) $A^{i_1\to v^{(\psi)}_1} \circ L^{i_1}_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
   (b) $A^{i_2\to v^{(\psi)}_2} \circ L^{i_2}_{k+j} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
   (c) $A^{i_3\to v^{(\psi)}_3} \circ L^{i_3}_{2k+j} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
   (d) $(v^{(\psi)}_1 - b_1) \cdot (v^{(\psi)}_2 - b_2) \cdot (v^{(\psi)}_3 - b_3) = 0$.

[14] We note that the statement of Lemma 7.28 refers to a smaller value of r' but of course, in particular, it also holds for larger values of r'.

where the probability is over $\{L^i_j\}_{i\in B, j\in[3k]}, A$. But since $r' + r' < k$, the latter implies that for every clause $\psi = (w_{i_1} = b_1 \vee w_{i_2} = b_2 \vee w_{i_3} = b_3)$ it holds that $v^{(\psi)}_1 = v_{i_1}$, $v^{(\psi)}_2 = v_{i_2}$ and $v^{(\psi)}_3 = v_{i_3}$ and in particular, $(v_{i_1} - b_1) \cdot (v_{i_2} - b_2) \cdot (v_{i_3} - b_3) = 0$. Furthermore, since for every $i \in B$ there is a clause of the form $i_1 = i_2 = i_3 = i$ and $b_1 = b_2 = 0$ and $b_3 = 1$ (indeed, these clauses were added in φ_{extra} to ensure a Boolean value), for every $i \in B$ it holds that $v_i \cdot v_i \cdot (v_i - 1) = 0$ and so $v_i \in \{0, 1\}$.

Thus, with probability $\geq 1 - 100|B|^3\ell|\mathbb{F}|\epsilon$, there exists an assignment $\alpha : B \to \{0, 1\}$ that is consistent with respect to B such that for every $i \in B$, for at least $3k - r'$ of the indices $j \in [3k]$ it holds that $A^{i\to\alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $\{L^i_j\}_{i\in B, j\in[3k]}, A$).

The lemma follows from Claim 7.2. □

10.2 The Main Lemma

Fix a layer $\mu \in [t]$ of the circuit $\mathcal{C}'_n$. Recall that:

- $\beta_\mu \subset [N']$ refers to the set of indices of gates in $\mathcal{C}'_n$ that are associated with the LDE of the μ-th layer of $\mathcal{C}_n$.
- For every point $z \in \mathbb{G}^{m_\mathbb{G}}$ we denote by $\beta_\mu[z] \subset \beta_\mu$ the set of indices of the $\log_2 |\mathbb{G}|$ gates associated with the point z in the computation of the LDE of layer μ in $\mathcal{C}'_n$. For a sequence $Z \subset \mathbb{G}^{m_\mathbb{G}}$ we denote by $\beta_\mu[Z] \stackrel{\text{def}}{=} \cup_{z\in Z}\beta_\mu(z)$.
- The values $x_1, \ldots, x_{N'}$ denote the computation of the circuit $\mathcal{C}'_n$ on the input $(x_1, \ldots, x_n)$.

Lemma 10.3. *Let $\lambda \in \beta_\mu$ be a fixed point and let $Z = (z_1, \ldots, z_\nu)$ be a sequence of ν points, where each point z_i is uniformly distributed in $\mathbb{G}^{m_\mathbb{G}}$.*

Let $B = \beta_\mu[Z] \cup \{\lambda\}$. We view the random variable B as being distributed over subsets of $H^m \subset D_X$.

For every $i \in B$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i\in B, j\in[k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

Let $\eta > 0$. Suppose that with probability $\geq 1 - \eta$, for every $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i\in B, j\in[k]}, A$).

Then, with probability $\geq 1 - \eta - \mathsf{polylog}(s) \cdot \nu^3\ell|\mathbb{F}|\epsilon - 2^{-\nu}$, for every $i \in \{\lambda\} \cup \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i\in B, j\in[k]}, A$).

Proof. Let $z_0 \in \mathbb{G}^{m_{\mathbb{G}}}$ be the point such that $\lambda \in \beta_\mu[z_0]$ (i.e., λ belongs to the block associated with the point z_0). Let $Z = (z_1, \ldots, z_\nu)$ be a sequence of ν uniformly distributed points in $\mathbb{G}^{m_{\mathbb{G}}}$. For every $i \in [\nu]$, let $L_{z_0,z_i} : \mathbb{G} \to \mathbb{G}^{m_{\mathbb{G}}}$ be the line $L_{z_0,z_i}(t) = (z_i - z_0) \cdot t + z_0$ (i.e., the line that passes through the points z_0 and z_i).

For every line L_{z_0,z_i} let $B_{z_0,z_i} \subset [N']$ be the indices of all gates that are associated with the verification in $\mathcal{C}'_n$ that the LDE of layer μ restricted to the line L_{z_0,z_i} is a degree $m_{\mathbb{G}}|H_{\mathbb{G}}|$ univariate polynomial (recall that such gates are a part of the $\mathcal{C}_{\mathsf{LDE}}$ circuit of the μ-th layer of $\mathcal{C}'_n$, see Section 9). Let $B' = \cup_{i\in[\nu]} B_{z_o,z_i}$. Note that $|B'| = \nu \cdot \mathsf{polylog}(s)$ (since the verification can be implemented by a Boolean circuit of size $\mathsf{polylog}(s)$, see Section 9).

For every assignment $\alpha : B' \to \{0,1\}$, we denote by $\alpha_{\mathbb{G}} : \mathbb{G}^{m_{\mathbb{G}}} \to \mathbb{G}$ the partial function[15] $\alpha_{\mathbb{G}}(\zeta) \stackrel{\text{def}}{=} \alpha(\beta_\mu[\zeta]) \in \{0,1\}^{\log_2|\mathbb{G}|}$ (where $\alpha_{\mathbb{G}}$ is only defined over $\cup_{i\in[\nu]}\{L_{z_0,z_i}(u) \ : \ u \in \mathbb{G}\}$). We say that α is *consistent* (w.r.t. the sequence Z) if for every $i \in [\nu]$ the function $\alpha_{\mathbb{G}} \circ L_{z_0,z_i}$ is a degree $m_{\mathbb{G}}|H_{\mathbb{G}}|$ (univariate) polynomial. We say that the assignment $\alpha : B' \to \{0,1\}$ is *correct* at the point $\zeta \in \cup_{i\in[\nu]}\{L_{z_0,z_i}(u) \ : \ u \in \mathbb{G}\}$ if for every $i \in \beta_\mu(\zeta) \subseteq B'$ it holds that $\alpha(i) = x_i$. We say that the assignment α is *correct* at a sequence of points if it is correct at every point in the sequence.

For every $i \in B'$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$. Let $S = \{L^i_j(t)\}_{i\in B', j\in[k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

Since the $\mathcal{C}_{\mathsf{LDE}}$ circuit of layer μ verifies that each line L_{z_0,z_i} has low degree, by applying Lemma 10.2 to the set B' (while noting that $3k|\mathbb{F}||B'| < k_{max}$) we have that, with probability $\geq 1 - 200(\nu \cdot \mathsf{polylog}(s))^3 \ell|\mathbb{F}|\epsilon$, (over $\{L^i_j\}_{i\in B', j\in[k]}, A$), there exists a *consistent* assignment $\alpha : B' \to \{0,1\}$ such that for every $i \in B'$, for at least $k - r'$ of the indices $j \in [k]$, it holds that $A^{i\to\alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

On the other hand, by the lemma's hypothesis (using also Claim 7.2), with probability $\geq 1 - \eta - \delta$ (over $Z, \{L^i_j\}_{i\in B', j\in[k]}, A$), for every $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Let E be the event that there exists a *consistent* assignment $\alpha : B' \to \{0,1\}$ *that is correct on* Z such that for every $i \in B'$, for at least $k - r'$ of the indices $j \in [k]$, it holds that $A^{i\to\alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

By the union bound (and using the fact that $r' + r' < k$),

$$\Pr[E] \geq 1 - \eta - \mathsf{polylog}(s) \cdot \nu^3 \ell|\mathbb{F}|\epsilon.$$

where the probability is over $Z, \{L^i_j\}_{i\in B', j\in[k]}, A$.

Let E' be the event there exists a consistent assignment $\alpha : B' \to \{0,1\}$ *that is incorrect*

[15] We use $\alpha(\beta_\mu[\zeta])$ to denote the element in $\mathbb{G}$ that is obtained by considering the assignment α applied to the gates indexed by $\beta_\mu[\zeta]$ in $\mathcal{C}'_n$ and interpreting the resulting $\log_2 \mathbb{G}$ string as the corresponding element in $\mathbb{G}$.

at the point z_0 such that for every $i \in B'$, for at least $k - r'$ of the indices $j \in [k]$, it holds that $A^{i \to \alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Consider the event $E \wedge E'$. If both E and E' occur then, by their definitions:

1. There exists a consistent assignment $\alpha : B' \to \{0,1\}$ that is correct on Z such that for every $i \in B'$, for at least $k - r'$ of the indices $j \in [k]$, it holds that $A^{i \to \alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. There exists a consistent assignment $\alpha' : B' \to \{0,1\}$ that is incorrect at the point z_0 such that for every $i \in B'$, for at least $k - r'$ of the indices $j \in [k]$, it holds that $A^{i \to \alpha'(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

However, since $r' + r' < k$ the assignment α must agree with α' on every $i \in B'$. Thus, if the event $E \wedge E'$ occurs then there exists a *single* consistent assignment $\alpha : B' \to \{0,1\}$ that is correct on Z *and* incorrect at the point z_0 such that for every $i \in B'$, for at least $k - r'$ of the indices $j \in [k]$, it holds that $A^{i \to \alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

We proceed to compute the probability that the event $E \wedge E'$ occurs. First observe that the sequence Z can be generated by using the following random process. First a sequence $Z' = (z'_1, \ldots, z'_\nu)$ of ν uniformly random points in $\mathbb{G}^{m_\mathbb{G}}$ is selected. For every $i \in [\nu]$, let $L_{z_0, z'_i}(t) = (z'_i - z_0)t + z_0$ be the line that passes through the points z_0 and z'_i. For every $i \in [\nu]$, the point z_i is selected by choosing at random $u_i \in \mathbb{G} \backslash \{0\}$ and setting $z_i = L_{z_0, z'_i}(u_i)$. Note that each one of the sequences Z and Z' is a sequence of ν uniformly distributed points in $\mathbb{G}^{m_\mathbb{G}}$.

Let $\chi : B' \to \{0,1\}$ denote the assignment of correct values to B'. That is, for every $i \in B'$, it holds that $\chi(i) = x_i$. Note that Z' already determines the set B' and that $Z', \{L^i_j\}_{i \in B', j \in [k]}, A$ already determine whether the event E' occurs (regardless of the choice of Z). Furthermore, if $Z', \{L^i_j\}_{i \in B', j \in [k]}, A$ are such that the event E' occurs, then the assignment α (guaranteed by E') is consistent and *incorrect* at the point z_0. Thus, for every $i \in [\nu]$ the two polynomials $\alpha_\mathbb{G} \circ L_{z_0, z'_i}$ and $\chi_\mathbb{G} \circ L_{z_0, z'_i}$ differ (at the point 0) and have degree at most $m_\mathbb{G}|H_\mathbb{G}|$. Hence, the two polynomials can agree on at most $m_\mathbb{G}|H_\mathbb{G}| < \frac{|\mathbb{G}|-1}{2}$ points, or in other words, for every $i \in [\nu]$ the assignment α is correct on less than half of the points on the line L_{z_0, z'_i}. Thus, we have:

$$\begin{aligned}
\Pr_{Z, \{L^i_j\}_{i \in B', j \in [k]}, A}[E \wedge E'] &= \mathop{\mathbf{E}}_{Z', \{L^i_j\}_{i \in B', j \in [k]}, A}\left[\Pr_{u_1, \ldots, u_\nu}[E \wedge E']\right] \\
&= \Pr[E'] \cdot \mathop{\mathbf{E}}_{Z', \{L^i_j\}_{i \in B', j \in [k]}, A}\left[\Pr_{u_1, \ldots, u_\nu}[E \wedge E'] \middle| E'\right] + \\
&\quad \Pr[\neg E'] \cdot \mathop{\mathbf{E}}_{Z', \{L^i_j\}_{i \in B', j \in [k]}, A}\left[\Pr_{u_1, \ldots, u_\nu}[E \wedge E'] \middle| \neg E'\right]
\end{aligned}$$

However, if Z', $\{L^i_j\}_{i \in B', j \in [k]}$ and A are such that $\neg E'$ occurs then $\Pr_{u_1, \ldots, u_\nu}[E \wedge E'] = 0$. On the other hand, by the foregoing discussion, if Z', $\{L^i_j\}_{i \in B', j \in [k]}$, and A are such that E'

occurs then $\Pr_{u_1,\ldots,u_\nu}[E \wedge E'] \leq 2^{-\nu}$. Thus:

$$\Pr_{Z,\{L^i_j\}_{i\in B',j\in[k]},A}[E \wedge E'] \leq \Pr[E'] \cdot 2^{-\nu} \leq 2^{-\nu}$$

and so

$$\Pr[E \wedge \neg E'] = \Pr[E] - \Pr[E \wedge E'] \geq 1 - \eta - \mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon - 2^{-\nu}.$$

In other words, with probability $1-\eta-\mathsf{polylog}(s)\cdot\nu^3\ell|\mathbb{F}|\epsilon-2^{-\nu}$, there exists a consistent assignment $\alpha : B' \to \{0,1\}$ *that is correct on* Z *and on* z_0 such that for every $i \in B'$, for at least $k-r'$ of the indices $j \in [k]$, it holds that $A^{i\to\alpha(i)} \circ L^i_j : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$. The lemma follows by Claim 7.2.

□

10.3 Some Useful Claims

Claim 10.4. *Let $S \subset D, |S| \leq k_{max}$ be a set generated by some random process. Let $A \in_R \mathcal{A}_S$. Let $g(S,A)$ be a predicate such that $\Pr_{A,S}[g(S,A)] \geq 1/2$. Let $f(S,A)$ be a predicate such that $\Pr_{A,S}[f(S,A) \mid g(S,A)] = p$. Let S', Q, such that $S' \subseteq Q \subset D, |Q| \leq k_{max}$, be two sets generated by some random process, such that the distribution of S' is identical to the distribution of S. Let $A' \in_R \mathcal{A}_Q$. Then,*

$$p - 4\delta \leq \Pr_{A',S',Q}\left[f(S', A'_{S'}) \mid g(S', A'_{S'})\right] \leq p + 4\delta$$

.

Proof. Denote:

$$\begin{aligned}
a &\stackrel{\text{def}}{=} \Pr_{S,A\in_R\mathcal{A}_s}[f(S,A) \wedge g(S,A)] \\
b &\stackrel{\text{def}}{=} \Pr_{S,A\in_R\mathcal{A}_s}[g(S,A)] \\
c &\stackrel{\text{def}}{=} \Pr_{Q,S',A'\in_R\mathcal{A}_Q}[f(S', A'_{S'}) \wedge g(S', A'_{S'})] \\
d &\stackrel{\text{def}}{=} \Pr_{Q,S',A'\in_R\mathcal{A}_Q}[g(S', A'_{S'})]
\end{aligned}$$

By Claim 7.2, $|a-c| < \delta$ and $|b-d| < \delta$ (and in particular $d \geq 1/2 - \delta > 0.4$ and therefore the conditional probability space in the lemma's conclusion is non-empty). Note that:

$$\begin{aligned}
\frac{a}{b} &= \Pr_{S,A\in_R\mathcal{A}_s}[f(S,A) \mid g(S,A)] \\
\frac{c}{d} &= \Pr_{Q,S',A'\in_R\mathcal{A}_Q}[f(S', A'_{S'}) \mid g(S', A'_{S'})].
\end{aligned}$$

Using elementary manipulations we have that,

$$\left|\frac{a}{b}-\frac{c}{d}\right| = \frac{|ad-bc|}{bd} = \frac{|ad-cd+cd-bc|}{bd} \leq \frac{d|a-c|+c|d-b|}{bd} \leq \frac{|a-c|}{b}+\frac{|d-b|}{b}\cdot c/d \leq 4\delta$$

where the last inequality uses also the hypothesis that $b \geq 1/2$ and the fact that $c \leq d$. □

Claim 10.5. *Let $\gamma \geq 0$ and let A and B be events over the same probability space such that $\Pr[A] \geq 1-\gamma$ and $\Pr[B] \geq \frac{1}{2}$. Then $\Pr[A|B] \geq 1-2\gamma$.*

Proof.

$$\Pr[A|B] = \frac{\Pr[A \wedge B]}{\Pr[B]} \geq \frac{\Pr[A]+\Pr[B]-1}{\Pr[B]} \geq 1-\frac{\gamma}{\Pr[B]} \geq 1-2\gamma.$$

□

Claim 10.6. *Let $\gamma, \eta < 1$ and let A and B be events over the same probability space such that $\Pr[B] \geq 1-\gamma$ and $\Pr[A|B] \geq 1-\eta$. Then $\Pr[A] \geq 1-\gamma-\eta$.*

Proof.

$$\Pr[A] \geq \Pr[A \wedge B] = \Pr[A|B] \cdot \Pr[B] \geq (1-\gamma)(1-\eta) \geq 1-\gamma-\eta.$$

□

Claim 10.7 (Union Bound under Conditioning)**.** *Let A, B and C be events over the same probability space, such that C has non-zero probability. Then:*

$$\Pr[A \vee B|C] \leq \Pr[A|C] + \Pr[B|C].$$

Proof.

$$\Pr[A \vee B|C] = \frac{\Pr[(A \vee B) \wedge C]}{\Pr[C]} \leq \frac{\Pr[(A \wedge C)] + \Pr[(B \wedge C)]}{\Pr[C]} = \Pr[A|C] + \Pr[B|C].$$

□

10.4 The Property $\mathcal{R}_\mu$ and making Progress under Conditioning

We will now give two definitions that will be central in the rest of the section. The first definition is analogous to the definition of the property $\mathcal{R}$ (Definition 7.31) in Section 7.6.

Definition 10.8. Property $\mathcal{R}_\mu(\epsilon', r')$:
Let $\mu \in [t]$, $\epsilon' \geq 0$ and $r' \geq 0$.

Let $B \subseteq [N']$ such that $(|B|+\nu \cdot \log_2|\mathbb{G}|)\cdot k|\mathbb{F}| < k_{max}$. Let Z be a sequence of ν uniformly distributed points in $\mathbb{G}^{m_\mathbb{G}}$. Let $B' = B \cup \beta_\mu[Z]$. We view both B and B' as being distributed over subsets of $H^m \subseteq D_X$.

For every $i \in B'$, let $L_1^i, \ldots, L_k^i : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L_1^i, \ldots, L_k^i\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i\in B', j\in[k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

Denote by E the event that for every point $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

We say that the set B satisfies property $\mathcal{R}_\mu(\epsilon', r')$ (also denoted $B \in \mathcal{R}_\mu(\epsilon', r')$) if, conditioned on the event E, with probability $\geq 1 - \epsilon'$, for every point $i \in B$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i\in B, j\in[k]}, A$).

To ensure that R_μ is well defined, if $\Pr[E] = 0$, then no set B is said to satisfy $\mathcal{R}_\mu(\epsilon', r')$.

Definition 10.9. ***p*-good layers:**
Let $\mu \in [t]$. Let Z be a sequence of ν uniformly distributed points in $\mathbb{G}^{m_\mathbb{G}}$.

For every $i \in \beta_\mu[Z]$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i\in\beta_\mu[Z], j\in[k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

*We say that the layer μ is p-*good *for $p \in [0, 1]$ if, with probability $\geq p$, for every point $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i\in\beta_\mu[Z], j\in[k]}, A$).*

Lemma 10.10. *Let $i_1, i_2, i_3 \in [N']$ be such that the gate indexed by i_1 in the circuit $\mathcal{C}'_n$ has children indexed by i_2, i_3. Let $\mu \in [t]$ be a 0.9-good layer. If $\{i_2\}, \{i_3\} \in \mathcal{R}_\mu(\epsilon', r')$, then $\{i_1\} \in \mathcal{R}_\mu(\epsilon'', r')$ where $\epsilon'' = 2\epsilon' + 34\ell|\mathbb{F}|\epsilon$.*

Proof. Let Z be a sequence of ν uniformly distributed points in $\mathbb{G}^{m_\mathbb{G}}$. Let $B = \{i_1, i_2, i_3\} \cup \beta_\mu[Z]$. We view B as being distributed over subsets of $H^m \subseteq D_X$.

For every $i \in B$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i\in B, j\in[k], t\in\mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i\to v} : S \to \mathbb{F}$ by $A^{i\to v}(i') = A(i')$ for $i' \neq i$ and $A^{i\to v}(i) = v$.

Denote by E the event that for every point $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i\to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$. Since μ is a 0.9-good layer (using also Claim 7.2), the event E occurs with probability $\geq 0.9 - \delta > 1/2$ (where the probability is over $Z, \{L^i_j\}_{i\in B, j\in[k]}, A$).

Since $\{i_2\} \in \mathcal{R}_\mu(\epsilon', r')$, using also Claim 10.4, conditioned on the event E occurring, with probability $\geq 1 - \epsilon' - 4\delta$ for at least $k - r'$ of the lines $L \in \{L^{i_2}_1, \ldots, L^{i_2}_k\}$, we have that

$A^{i_2 \to x_{i_2}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L_j^i\}_{i \in B, j \in [k]}, A$).

Similarly, since $i_3 \in \mathcal{R}_\mu(\epsilon', r')$, conditioned on the event E occurring, with probability $\geq 1 - \epsilon' - 4\delta$ for at least $k - r'$ of the lines $L \in \{L_1^{i_3}, \ldots, L_k^{i_3}\}$, we have that $A^{i_3 \to x_{i_3}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L_j^i\}_{i \in B, j \in [k]}, A$).

Since the gate indexed by i_1 in the circuit $\mathcal{C}'_n$ has children indexed by i_2, i_3, the formula φ contains the clause $(w_{i_2} = x_{i_2}) \wedge (w_{i_3} = x_{i_3}) \to (w_{i_1} = x_{i_1})$. Thus, by Lemma 7.30 (using also Claim 7.2), with probability $\geq 1 - 9\ell|\mathbb{F}|\epsilon - 2\delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$,

1. $A^{i_1 \to v_1} \circ L_j^{i_1} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{i_2 \to v_2} \circ L_j^{i_2} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{i_3 \to v_3} \circ L_j^{i_3} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $(v_2 = x_{i_2}) \wedge (v_3 = x_{i_3}) \to (v_1 = x_{i_1})$

Thus, using Claim 10.5, conditioned on the event E occurring, with probability $\geq 1 - 18\ell|\mathbb{F}|\epsilon - 4\delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$,

1. $A^{i_1 \to v_1} \circ L_j^{i_1} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{i_2 \to v_2} \circ L_j^{i_2} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{i_3 \to v_3} \circ L_j^{i_3} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $(v_2 = x_{i_2}) \wedge (v_3 = x_{i_3}) \to (v_1 = x_{i_1})$.

Thus, by the union bound under conditioning (Claim 10.7), conditioned on the event E occurring, with probability $\geq 1 - 2\epsilon' - 18\ell|\mathbb{F}|\epsilon - 12\delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that $(v_2 = x_{i_2}) \wedge (v_3 = x_{i_3}) \to (v_1 = x_{i_1})$ and:

- For at least $k - r'$ of the lines $L \in \{L_1^{i_2}, \ldots, L_k^{i_2}\}$, we have that $A^{i_2 \to x_{i_2}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
- For at least $k - r'$ of the lines $L \in \{L_1^{i_3}, \ldots, L_k^{i_3}\}$, we have that $A^{i_3 \to x_{i_3}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
- For at least $k - r'$ of the indices $j \in [k]$,
  1. $A^{i_1 \to v_1} \circ L_j^{i_1} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
  2. $A^{i_2 \to v_2} \circ L_j^{i_2} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
  3. $A^{i_3 \to v_3} \circ L_j^{i_3} : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Since $r'+r' < k$, this implies that $v_2 = x_{i_2}$, $v_3 = x_{i_3}$ and hence $v_1 = x_{i_1}$. Thus, conditioned on the event E occurring, with probability $\geq 1 - 2\epsilon' - 18\ell|\mathbb{F}|\epsilon - 12\delta > 1 - 2\epsilon' - 30\ell|\mathbb{F}|\epsilon$, for at least $k - r'$ of the lines $L \in \{L_1^{i_1}, \ldots, L_k^{i_1}\}$, we have that $A^{i_1 \to x_{i_1}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L_j^i\}_{i \in B, j \in [k]}, A$).

The lemma follows by an application of Claim 10.4. □

If C is a circuit, we say that a subset B of the gates of C is a sub-circuit of C if for every gate $g \in B$ either both of its children are in B or both are not in B. Gates in B whose children are not in B are called input gates of B and gates in B who are not children of any gate in B are called output gates of B. We say that the sub-circuit B has depth Δ if the longest path from an output gate of B to an input gate of B is of length Δ.

Using Lemma 10.10, we are ready to prove the following lemma.

Lemma 10.11. *Let $B \subset [N']$ be a sub-circuit of $\mathcal{C}'_n$ of depth Δ with input gates $B_I \subset B$ and output gates $B_O \subset B$, such that $(|B_O| + \nu \cdot \log_2 |\mathbb{G}|) \cdot k|\mathbb{F}| < k_{max}$. Let $\mu \in [t]$ be a 0.9-good layer of the circuit. If for all $i \in B_I$ it holds that $\{i\} \in \mathcal{R}_\mu(\epsilon', r')$ then $B_O \in \mathcal{R}_\mu(\epsilon'', r')$, where $\epsilon'' = |B_O| \cdot 2^\Delta \cdot (2\epsilon' + 38\ell|\mathbb{F}|\epsilon)$.*

Proof. For every $i \in B_O$, by iterated applications of Lemma 10.10, it holds that $\{i\} \in \mathcal{R}_\mu\left(2^\Delta \cdot (2\epsilon' + 34\ell|\mathbb{F}|\epsilon), r'\right)$. The lemma follows from $|B_O|$ applications of Claim 10.4, and the union bound under conditioning (Claim 10.7). □

We also prove (simpler) variants of Lemma 10.10 and Lemma 10.11 with respect to the property $\mathcal{R}$ (rather than $\mathcal{R}_\mu$), see Definition 7.31. Recall that, intuitively, a subset $B \subset H^m \subset D_X$ satisfies property $\mathcal{R}(\epsilon', r')$ if when taking k lines through every point in B, with high probability, for every point $i \in B$, for most of the lines through the point i, the answers correspond to low degree polynomials that "evaluate" the point i to x_i.

Lemma 10.12. *Let $i_1, i_2, i_3 \in [N']$ be such that the gate indexed by i_1 in the circuit $\mathcal{C}'_n$ has children indexed by i_2, i_3. If $\{i_2\}, \{i_3\} \in \mathcal{R}(\epsilon', r')$, then $\{i_1\} \in \mathcal{R}(\epsilon'', r')$ where $\epsilon'' = 2\epsilon' + 15\ell|\mathbb{F}|\epsilon$.*

Proof. If $\{i_2\}, \{i_3\} \in \mathcal{R}(\epsilon', r')$, then by Lemma 7.34, it holds that $\{i_2, i_3\} \in \mathcal{R}(2\epsilon' + 2\delta, r')$. Since the gate indexed by i_1 in the circuit $\mathcal{C}'_n$ has children indexed by i_2, i_3, by Lemma 7.33, it holds that $\{i_1, i_2, i_3\} \in \mathcal{R}(2\epsilon' + 9\ell|\mathbb{F}|\epsilon + 5\delta, r')$. The lemma follows from Lemma 7.35. □

Lemma 10.13. *Let $B \subset [N']$ be a sub-circuit of $\mathcal{C}'_n$ of depth Δ with input gates $B_I \subset B$ and output gates $B_O \subset B$, such that $|B_O|k|\mathbb{F}| < k_{max}$. If for all $i \in B_I$ it holds that $\{i\} \in \mathcal{R}(\epsilon', r')$ then $B_O \in \mathcal{R}(\epsilon'', r')$, where $\epsilon'' = |B_O| \cdot 2^\Delta \cdot (2\epsilon' + 16\ell|\mathbb{F}|\epsilon)$.*

Proof. For every $i \in B_O$, by iterated applications of Lemma 10.12, it holds that $\{i\} \in \mathcal{R}\left(2^\Delta \cdot (2\epsilon' + 15\ell|\mathbb{F}|\epsilon), r'\right)$. The lemma follows from $|B_O|$ applications of Claim 7.2, and the union bound. □

10.5 Proof of Lemma 10.1

In this section we complete the proof of Lemma 10.1. We first show that if layer $\mu-1$ is a good layer then layer μ is also good. Then we derive that the top layer is good and use that to contradict our assumption that $x \notin \mathcal{L}$.

Lemma 10.14. *Let $\mu \in [t]$ be a 0.9-good layer. Then, for every $\lambda \in \beta_\mu$ it holds that $\lambda \in \mathcal{R}_\mu(\epsilon', r')$, where $\epsilon' = \mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon + 2^{-\nu+1}$.*

Proof. Let $\lambda \in \beta_\mu$. Let Z be a sequence of ν uniformly distributed points in $\mathbb{G}^{m_\mathbb{G}}$. Let $B = \{\lambda\} \cup \beta_\mu[Z]$.

For every $i \in B$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i \in B, j \in [k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i \to v} : S \to \mathbb{F}$ by $A^{i \to v}(i') = A(i')$ for $i' \neq i$ and $A^{i \to v}(i) = v$.

Denote by E the event that for every point $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i \in B, j \in [k]}, A$). Suppose that $\Pr[E] = 1 - \eta$ for some $\eta \in [0, 1]$. Note that by the hypothesis that μ is 0.9-good, using also Claim 7.2, $\eta < 0.1 + \delta < 1/2$.

Denote by $E' \subset E$ the event that for every point $i \in \{\lambda\} \cup \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i \in B, j \in [k]}, A$). By Lemma 10.3, $\Pr[E'] \geq 1 - \eta - \mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon - 2^{-\nu}$.

Thus, the probability that the event E' occurs conditioned on the event E is at least:

$$\Pr[E'|E] = \frac{\Pr[E']}{\Pr[E]} \geq \frac{1 - \eta - \mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon - 2^{-\nu}}{1 - \eta} \geq 1 - \mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon - 2^{-\nu+1}$$

(where the last inequality follows from the fact that $\eta \leq 1/2$) and the lemma follows. □

Lemma 10.15. *Let $\mu \in [t-1]$ be a 0.9-good layer. Then, for every set $B \subseteq \beta_{\mu+1}$ of points that belong to layer $\mu+1$, such that $(|B| + \nu \cdot \log_2 |\mathbb{G}|) \cdot k |\mathbb{F}| < k_{max}$, it holds that $B \in \mathcal{R}_\mu(\epsilon', r')$, where $\epsilon' = |B| \cdot 2^d \cdot \left(\mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon + 2^{-\nu+3}\right)$.*

Proof. Consider the sub-circuit that computes layer $\mu + 1$ from layer μ. Recall that this sub-circuit has depth $d + 1$ and first computes the $\mu + 1$-th layer of $\mathcal{C}_n$, in depth 1, and then applies a C_{LDE} circuit, of depth d (see Section 9). Let $B_I \subseteq \beta_\mu$ be the variables associated with the inputs of this sub-circuit. By Lemma 10.14, for every $\lambda \in B_I$ it holds that $\lambda \in \mathcal{R}_\mu(\mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon + 2^{-\nu+1}, r')$.

The lemma follows from Lemma 10.11. □

Lemma 10.16. *If a layer $\mu \in [t-1]$ is $(1-\epsilon')$-good, for some $\epsilon' < 0.1$, then the layer $\mu+1$ is $(1-\epsilon'')$-good, where $\epsilon'' = \epsilon' + 2^d \cdot \mathsf{polylog}(s) \cdot \left(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\right)$.*

Proof. Let Z and Z' be two sequences of ν uniformly distributed points in $\mathbb{G}^{m_{\mathbb{G}}}$. Let $B = \beta_\mu[Z] \cup \beta_{\mu+1}[Z']$. We view B as being distributed over subsets of D_X.

For every $i \in B$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i \in B, j \in [k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i \to v} : S \to \mathbb{F}$ by $A^{i \to v}(i') = A(i')$ for $i' \neq i$ and $A^{i \to v}(i) = v$.

Denote by E the event that for every point $i \in \beta_\mu[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $\{Z, Z', L^i_j\}_{i \in B, j \in [k]}, A$). By the hypothesis that μ is $(1-\epsilon')$-good, and using Claim 7.2, the event E occurs with probability $\geq 1 - \epsilon' - \delta$.

By Lemma 10.15 (using the fact that $\epsilon' < 0.1$), it holds that $\beta_{\mu+1}[Z'] \in \mathcal{R}_\mu\Big((\log_2 \mathbb{G} \cdot \nu) \cdot 2^d \cdot \left(\mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon + 2^{-\nu+3}\right), r'\Big)$. In other words, conditioned on the event E, with probability $\geq 1 - 2^d \mathsf{polylog}(s) \cdot \left(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\right)$, for every point $i \in \beta_{\mu+1}[Z']$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, Z', \{L^i_j\}_{i \in B, j \in [k]}, A$).

However, since E occurs with high probability, we can remove the conditioning as follows. Toward this end, we apply Claim 10.6 and obtain that with probability $\geq 1 - 2^d \cdot \mathsf{polylog}(s) \cdot \left(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\right) - \epsilon' - \delta$, for every point $i \in \beta_{\mu+1}[Z']$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, Z', \{L^i_j\}_{i \in B, j \in [k]}, A$). The lemma follows by Claim 7.2. □

Recall that $1, \ldots, n$ are the indexes of the n input variables and $n+1, \ldots, 2n$ are the indexes of their negations.

Lemma 10.17. *The first layer of $\mathcal{C}'_n$ is $(1-\epsilon')$-good, where $\epsilon' = 2^d \mathsf{polylog}(s) \cdot \nu \ell |\mathbb{F}| \epsilon$.*

Proof. By Lemma 7.32, for every $i \in [2n]$ it holds that $\{i\} \in \mathcal{R}\left(10\ell|\mathbb{F}|\epsilon, r'\right)$. Let $\mathcal{C}^{(1)}_{\mathsf{LDE}}$ be the $\mathcal{C}_{\mathsf{LDE}}$ circuit of the first layer of $\mathcal{C}'_n$. Note that the inputs of $\mathcal{C}^{(1)}_{\mathsf{LDE}}$ are associated with the variables $i \in [2n]$ and that $\mathcal{C}^{(1)}_{\mathsf{LDE}}$ has depth d. Thus, by Lemma 10.13, for every sequence of ν points Z in $\mathbb{G}^{m_{\mathbb{G}}}$ it holds that $\beta_1[Z] \in \mathcal{R}\Big(\left(\nu \cdot \log_2(|\mathbb{G}|)\right) \cdot 2^d \cdot \mathsf{polylog}(s) \cdot \ell |\mathbb{F}| \epsilon, r'\Big)$ and the lemma follows. □

Lemma 10.18. *The top layer of $\mathcal{C}'_n$ (i.e., the t-th layer) is $(1-\epsilon')$-good, where $\epsilon' \leq t \cdot 2^d \cdot \mathsf{polylog}(s)\left(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\right)$.*

Proof. By induction, using Lemma 10.17 and Lemma 10.16. □

Recall that N' is the index of the special output gate.

Lemma 10.19. *$\{N'\} \in \mathcal{R}(\epsilon', r')$, where $\epsilon' = t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big)$.*

Proof. Let Z be a sequence of ν uniformly distributed points in $\mathbb{G}^{m_{\mathbb{G}}}$. Let $B = \{N'\} \cup \beta_t[Z]$. Note that the point N' belongs to layer t. We view B as being distributed over subsets of $H^m \subseteq D_X$.

For every $i \in B$, let $L^i_1, \ldots, L^i_k : \mathbb{F} \to D_X$ be k random lines, such that for every $L \in \{L^i_1, \ldots, L^i_k\}$, we have $L(0) = i$.

Let $S = \{L^i_j(t)\}_{i \in B, j \in [k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $i \in D_X$ and $v \in \mathbb{F}$, define $A^{i \to v} : S \to \mathbb{F}$ by $A^{i \to v}(i') = A(i')$ for $i' \neq i$ and $A^{i \to v}(i) = v$.

Denote by E the event that for every point $i \in \beta_t[Z]$, for at least $k - r'$ of the lines $L \in \{L^i_1, \ldots, L^i_k\}$, we have that $A^{i \to x_i} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$. By Lemma 10.18, and using Claim 7.2, the event E occurs with probability at least $1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big)$.

Since by our setting of parameters $t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big) < 0.1$, the layer t is 0.9 good and so, by Lemma 10.14, it holds that $N' \in \mathcal{R}_t(\mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon + 2^{-\nu+1}, r')$. In other words, conditioned on the event E, with probability $\geq 1 - \mathsf{polylog}(s) \cdot \nu^3 \ell |\mathbb{F}| \epsilon - 2^{-\nu+1}$, for at least $k - r'$ of the lines $L \in \{L^{N'}_1, \ldots, L^{N'}_k\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i \in B, j \in [k]}, A$).

Hence, by Claim 10.6, with probability $\geq 1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big)$, for at least $k - r'$ of the lines $L \in \{L^{N'}_1, \ldots, L^{N'}_k\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$ (where the probability is over $Z, \{L^i_j\}_{i \in B, j \in [k]}, A$). The lemma follows by Claim 7.2. □

Proof of Lemma 10.1

The following proof is similar to the proof of Lemma 7.1 (in Section 7.7) but differs in the actual parameters, and in the use of Lemma 10.19 (rather than Lemma 7.37).

Proof. Consider the point $N' \in [N']$, viewed as a point in $H^m \subset D_X$. Recall that the formula φ' contains a clause $(w_{N'} = 1) \vee (w_{N'} = 1) \vee (w_{N'} = 1)$ that checks that $w_{N'} = 1$.

Let $L^1_1, \ldots, L^1_k, L^2_1, \ldots, L^2_k, L^3_1, \ldots, L^3_k : \mathbb{F} \to D_X$ be $3k$ random lines, such that for every line $L \in \{L^1_1, \ldots, L^1_k, L^2_1, \ldots, L^2_k, L^3_1, \ldots, L^3_k\}$, we have $L(0) = N'$.

Let $S = \{L^1_j(t), L^2_j(t), L^3_j(t)\}_{j \in [k], t \in \mathbb{F}} \subset D_X$. Let $A \in_R \mathcal{A}_S$.

For any $v \in \mathbb{F}$, define $A^{N' \to v} : S \to \mathbb{F}$ by $A^{N' \to v}(i') = A(i')$ for $i' \neq N'$ and $A^{N' \to v}(N') = v$.

By Lemma 7.30, with probability $\geq 1 - 9\ell |\mathbb{F}| \epsilon - \delta$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, for at least $k - r'$ of the indices $j \in [k]$, the following is satisfied
(where the probability is over $L^1_1, \ldots, L^1_k, L^2_1, \ldots, L^2_k, L^3_1, \ldots, L^3_k, A$):

1. $A^{N' \to v_1} \circ L_j^1 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. $A^{N' \to v_2} \circ L_j^2 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. $A^{N' \to v_3} \circ L_j^3 : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. $(v_1 - 1) \cdot (v_2 - 1) \cdot (v_3 - 1) = 0$.

On the other hand, by (three applications of) Lemma 10.19 and Claim 7.2:

1. With probability $\geq 1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big) - \delta$, for at least $k - r'$ of the lines $L \in \{L_1^1, \ldots, L_k^1\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. With probability $\geq 1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big) - \delta$, for at least $k - r'$ of the lines $L \in \{L_1^2, \ldots, L_k^2\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. With probability $\geq 1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big) - \delta$, for at least $k - r'$ of the lines $L \in \{L_1^3, \ldots, L_k^3\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Thus, by the union bound, with probability $\geq 1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big) > 0$, there exist $v_1, v_2, v_3 \in \mathbb{F}$, such that, $(v_1 - 1) \cdot (v_2 - 1) \cdot (v_3 - 1) = 0$, and

1. For at least $k - r'$ of the lines $L \in \{L_1^1, \ldots, L_k^1\}$, we have that $A^{N' \to v_1} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
2. For at least $k - r'$ of the lines $L \in \{L_1^2, \ldots, L_k^2\}$, we have that $A^{N' \to v_2} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
3. For at least $k - r'$ of the lines $L \in \{L_1^3, \ldots, L_k^3\}$, we have that $A^{N' \to v_3} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
4. For at least $k - r'$ of the lines $L \in \{L_1^1, \ldots, L_k^1\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
5. For at least $k - r'$ of the lines $L \in \{L_1^2, \ldots, L_k^2\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.
6. For at least $k - r'$ of the lines $L \in \{L_1^3, \ldots, L_k^3\}$, we have that $A^{N' \to x_{N'}} \circ L : \mathbb{F} \to \mathbb{F}$ is a univariate polynomial of degree $< m|H|$.

Since, $r' + r' < k$, this implies that $x_{N'} = v_1 = v_2 = v_3$, and hence $x_{N'} = 1$. Since by our setting of parameters

$$1 - t \cdot 2^d \cdot \mathsf{polylog}(s)\big(\nu^4 \ell |\mathbb{F}| \epsilon + 2^{-\nu/2}\big) > 0,$$

the original input x is in the language $\mathcal{L}$. □

11 Soundness of V in the Augmented PCP

This section is similar to Section 8, but with respect to the augmented PCP.

Recall that $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N'$, is the security parameter of the PCP, and that $1 \leq r < k$ is the parameter of the relaxed verifier V'. Recall that ℓ and $|\mathbb{F}|$ are bounded by $\mathsf{polylog}(N')$.

Lemma 11.1. *For a security parameter $k \leq \mathsf{poly}(n)$, such that $4|\mathbb{F}|^4 \leq k \leq N'$, fix the following parameters: Let $r = \frac{k}{40\ell|\mathbb{F}|}$. Let $\epsilon = 2^{-r/2}$. Let $k_{max} = k \cdot \mathsf{polylog}(s) \cdot \log(t)|\mathbb{F}| + 12k\ell|\mathbb{F}|^2$. Let $\delta = \frac{1}{|\mathbb{F}|^{8k\ell|\mathbb{F}|^2}}$. Then, V has soundness ϵ against (k_{max}, δ)-no-signaling strategies.*

The proof of Lemma 11.1 is similar to the proof of Lemma 8.1, but based on Lemma 10.1 (rather than Lemma 7.1).

Proof. Assume for a contradiction that V doesn't have soundness ϵ against (k_{max}, δ)-no-signaling strategies. By Lemma 6.1, since $\delta < \frac{\epsilon}{8 \cdot |\mathbb{F}|^{6k\ell|\mathbb{F}|^2}}$, we know that V' (with parameter r) doesn't have soundness $1-\epsilon'$ against (k'_{max}, δ')-no-signaling strategies, where $k'_{max} = k_{max} - 6k\ell|\mathbb{F}|^2 = k \cdot \mathsf{polylog}(s) \cdot \log(t)|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$, and $\delta' = 8\delta|\mathbb{F}|^{6k\ell|\mathbb{F}|^2}/\epsilon < \frac{1}{|\mathbb{F}|^{k\ell|\mathbb{F}|^2}}$, and $\epsilon' = (10\ell|\mathbb{F}|2^{-r} + 2\delta)/\epsilon < \frac{1}{100N'\ell|\mathbb{F}|}$.

Hence V' (with parameter r) doesn't have soundness $1-\epsilon'$ against (k'_{max}, δ')-no-signaling strategies, where $k'_{max} = k \cdot \mathsf{polylog}(s) \cdot \log(t)|\mathbb{F}| + 6k\ell|\mathbb{F}|^2$, and $\delta' = \frac{1}{|\mathbb{F}|^{k\ell|\mathbb{F}|^2}}$, and $\epsilon' = \frac{1}{100N'\ell|\mathbb{F}|}$.

This contradicts Lemma 10.1. □

12 From No-Signaling PCP to No-Signaling MIP

In this section we show how to transform a PCP that has soundness against (k_{max}, δ)-no-signaling strategies into an analogous MIP that uses k_{max} provers and has soundness against δ-no-signaling strategies.

Recall that a PCP (resp., MIP) relative to an oracle $\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}$ is a PCP (resp., MIP) in which the verifier has oracle access to the function ϕ_n (see Section 4).

Lemma 12.1. *Let $\mathcal{L}$ be a language and suppose that $\mathcal{L}$ has a PCP with soundness ϵ against (k_{max}, δ)-no-signaling strategies relative to an oracle $\{\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}\}_n$ (where n is the input length). Let D be the query alphabet, Σ be the answer alphabet, $k \leq k_{max}$ be the number of PCP queries and ℓ be the number of oracle queries. Then, $\mathcal{L}$ has an MIP relative to the same oracle $\{\phi_n\}$ with soundness ϵ against δ-no-signaling strategies. The MIP uses k_{max} provers, query alphabet D, answer alphabet Σ and ℓ oracle queries.*

Furthermore, if the running time of the PCP verifier is T_V then the running time of the MIP verifier is $O\left(T_V + k_{max} \cdot (\log|D| + \log|\Sigma|)\right)$ and if the PCP proof can be generated in time T_P then the running time of each of the MIP provers is $O(T_P)$.

Proof. Let V be the PCP verifier for $\mathcal{L}$ and let G_P be an algorithm that on input $x \in \mathcal{L}$ generates the PCP proof $P = G_P(x)$. We use V and G_P to construct an MIP for $\mathcal{L}$ (relative to the oracle $\{\phi_n\}$) that is sound against δ-no-signaling strategies.

We think of the PCP verifier V as being composed of two algorithms V_1 and V_2. The first algorithm, V_1, on input x of length n and a random string r generates a set $Q \subset D$ of k queries to the PCP and a set $Q_\phi \subset \{0,1\}^{n'}$ of ℓ queries to the oracle. The second algorithm, V_2, given x, the same random string r, the k answers $A \in \Sigma^Q$ (of the PCP) and oracle answers $A_\phi \in (\{0,1\}^{n''})^{Q_\phi}$, decides whether to accept the proof. We also assume (without loss of generality) that the algorithm G_P is deterministic. (Since completeness holds with probability 1, we can de-randomize G_P by fixing its random string arbitrarily.)

We first describe the k_{max} (honest) MIP provers' strategies and then proceed to describe the MIP verifier's strategy. Given an input $x \in \mathcal{L}$, each MIP prover (individually) computes the (deterministic) PCP proof $P = G_P(x)$ and given a query $q \in D$ just answers with $P(q)$.

The MIP verifier, on input x and a random string r, first runs $V_1(x, r)$ to obtain a set of k PCP queries $Q = \{q_1, \ldots, q_k\}$ and ℓ oracle queries Q_ϕ. The set Q is then used to construct a sequence $w \in D^{k_{max}}$ of k_{max} queries as follows. Initially, every entry of w is set to an arbitrary fixed value $z \in D$. Then, the verifier embeds the k queries of Q at random in w (which is of length $k_{max} \geq k$). Formally, for every set $Q \subset D$, every 1-to-1 function $\pi : Q \to [k_{max}]$ and every subset $S \subseteq Q$, let $w_{S,\pi} \in D^{k_{max}}$ be defined as follows. For every $i \in [k_{max}]$, if there exists $q \in S$ such that $i = \pi(q)$ then $(w_{S,\pi})_i = q$ and otherwise $(w_{S,\pi})_i = z$. The verifier chooses at random a 1-to-1 function $\pi : Q \to [k_{max}]$ and sets $w = w_{Q,\pi}$. The verifier then sends w to the k_{max} provers where prover i gets w_i. Simultaneously, the verifier queries the oracle ϕ_n at the points Q_ϕ.

Once the k_{max} provers respond with their answers $\alpha \in \Sigma^{k_{max}}$ (where the answer of the i^{th} prover is α_i) and the oracle responds with A_ϕ, the MIP verifier constructs $A \in \Sigma^Q$ by setting $A_q = \alpha_{\pi(q)}$ for every $q \in Q$. Formally, for every $S \subseteq Q$, let $T_{S,\pi} : \Sigma^{k_{max}} \to \Sigma^S$ be defined as $(T_{S,\pi}(\alpha))_q = \alpha_{\pi(q)}$ for every $q \in S$. The verifier sets $A = T_{Q,\pi}(\alpha)$ and outputs the result of V_2 on input (x, r, A, A_ϕ).

To see that (perfect) completeness holds, observe that the honest MIP provers (that get queries in Q) answer according to the PCP. Likewise, the oracle queries and answers are also exactly as in the PCP and therefore if $x \in \mathcal{L}$ then V_2 accepts and we obtain perfect completeness. We proceed to show that soundness holds against δ-no-signaling strategies.

Suppose that for some $x \notin \mathcal{L}$, there exists a δ-no-signaling family of distributions $\mathcal{A} = \{\mathcal{A}_u\}_{u \in D^{k_{max}}}$ that makes the MIP verifier accept with probability at least ϵ. By the construction of the MIP system this implies that:

$$\Pr_{\alpha \in_R \mathcal{A}_w, r, \pi} [V_2(x, r, A, A_\phi) = 1] \geq \epsilon \tag{4}$$

where (Q, Q_ϕ) is the output of $V_1(x, r)$, the function $\pi : Q \to [k_{max}]$ is the random 1-to-1 function, $w = w_{Q,\pi}$, $A = T_{Q,\pi}(\alpha)$, and A_ϕ are the answers of the oracle ϕ_n on the points Q_ϕ.

We use $\mathcal{A}$ to construct a family of distributions $\mathcal{B} = \{\mathcal{B}_Q\}_{Q \subset D, |Q| \leq k_{max}}$ that violates the (k_{max}, δ)-no-signaling soundness of the PCP. For every set Q of size at most k_{max}, the

distribution $\mathcal{B}_Q$ is defined by first sampling a random 1-to-1 function $\pi : Q \to [k_{max}]$, setting $w = w_{Q,\pi}$, then sampling α from $\mathcal{A}_w$ and outputting $T_{Q,\pi}(\alpha)$. Note that for every $b \in \Sigma^Q$ it holds that

$$\Pr_{\beta \in_R \mathcal{B}_Q}[\beta = b] = \mathop{\mathbf{E}}_{\pi}\left[\Pr_{\alpha \in_R \mathcal{A}_w}[A = b]\right],$$

where $\pi : Q \to [k_{max}]$ is a random 1-1 function, $w = w_{Q,\pi}$ and $A = T_{Q,\pi}(\alpha)$.

We first show that the family of distributions $\mathcal{B} = \{\mathcal{B}_Q\}_{Q \subset D, |Q| \leq k_{max}}$ fools the PCP verifier into accepting with probability at least ϵ, and then proceed to show that $\mathcal{B}$ is δ-no-signaling. Indeed, by the definition of $\mathcal{B}$,

$$\Pr_{\beta \in_R \mathcal{B}_Q, r}[V_2(x, r, \beta, A_\phi) = 1] = \Pr_{\alpha \in_R \mathcal{A}_w, r, \pi}[V_2(x, r, A, A_\phi) = 1]$$

where (Q, Q_ϕ) is the output of $V_1(x, r)$, the function $\pi : Q \to [k_{max}]$ is the random 1-to-1 function, $w = w_{Q,\pi}$, $A = T_{Q,\pi}(\alpha)$, and A_ϕ are the answers of the oracle ϕ_n on the points Q_ϕ. Thus, by Eq. (4), the PCP verifier accepts $x \notin \mathcal{L}$ with probability at least ϵ.

The next claim shows that $\mathcal{B}$ is δ-no-signaling and therefore we have a contradiction to our assumption that the PCP has ϵ-soundness against (k_{max}, δ)-no-signaling strategies.

Claim 12.2. *The family of distributions $\mathcal{B} = \{\mathcal{B}_Q\}_{Q \subset D, |Q| \leq k_{max}}$ is δ-no-signaling.*

Proof. To show that $\mathcal{B}$ is δ-no-signaling we need to show that for every $Q \subset D$ of size at most k_{max} and every $S \subset Q$ it holds that

$$\frac{1}{2} \sum_{b \in \Sigma^S} \left| \Pr_{\beta \in_R \mathcal{B}_S}[\beta = b] - \Pr_{\beta \in_R \mathcal{B}_Q}[\beta_S = b] \right| \leq \delta.$$

For every $b \in \Sigma^S$ it holds that

$$\Pr_{\beta \in_R \mathcal{B}_S}[\beta = b] = \mathop{\mathbf{E}}_{\pi'}\left[\Pr_{\alpha' \in_R \mathcal{A}_{w'}}[T_{S,\pi'}(\alpha') = b]\right] = \mathop{\mathbf{E}}_{\pi}\left[\Pr_{\alpha \in_R \mathcal{A}_w}[T_{S,\pi}(\alpha) = b]\right] = \mathop{\mathbf{E}}_{\pi}\left[\Pr_{\alpha \in_R \mathcal{A}_w}[(T_{Q,\pi}(\alpha))_S = b]\right] \tag{5}$$

where $\pi' : S \to [k_{max}]$ and $\pi : Q \to [k_{max}]$ are random 1-to-1 functions, $w' = w_{S,\pi'}$ and $w = w_{S,\pi}$, the second equality follows from the fact that π, restricted to S, is distributed identically to π', and the last equality follows from the fact that $T_{S,\pi}(\alpha) = (T_{Q,\pi}(\alpha))_S$.

On the other hand, using elementary operations and linearity of expectation, for every

$b \in \Sigma^S$ it holds that

$$\begin{aligned}
\Pr_{\beta \in_R \mathcal{B}_Q}[\beta_S = b] &= \sum_{b' \in \Sigma^Q \text{ s.t. } b'_S = b} \Pr_{\beta \in_R \mathcal{B}_Q}[\beta = b'] \\
&= \sum_{b' \in \Sigma^Q \text{ s.t. } b'_S = b} \mathop{\mathbf{E}}_{\pi}\left[\Pr_{\alpha \in_R \mathcal{A}_{w''}}[T_{Q,\pi}(\alpha) = b']\right] \\
&= \mathop{\mathbf{E}}_{\pi}\left[\sum_{b' \in \Sigma^Q \text{ s.t. } b'_S = b} \Pr_{\alpha \in_R \mathcal{A}_{w''}}[T_{Q,\pi}(\alpha) = b']\right] \\
&= \mathop{\mathbf{E}}_{\pi}\left[\Pr_{\alpha \in_R \mathcal{A}_{w''}}[(T_{Q,\pi}(\alpha))_S = b]\right], \qquad (6)
\end{aligned}$$

where $\pi : Q \to [k_{max}]$ is a random 1-1 function and $w'' = w_{Q,\pi}$. Using Eq. (5) and Eq. (6), we obtain that:

$$\begin{aligned}
\frac{1}{2}\sum_{b \in \Sigma^S}\left|\Pr_{\beta \in_R \mathcal{B}_S}[\beta = b] - \Pr_{\beta \in_R \mathcal{B}_Q}[\beta_S = b]\right| &= \frac{1}{2}\sum_{b \in \Sigma^S}\left|\mathop{\mathbf{E}}_{\pi}\left[\Pr_{\alpha \in_R \mathcal{A}_w}[(T_{Q,\pi}(\alpha))_S = b] - \Pr_{\alpha'' \in_R \mathcal{A}_{w''}}[(T_{Q,\pi}(\alpha''))_S = b]\right]\right| \\
&\leq \mathop{\mathbf{E}}_{\pi}\left[\frac{1}{2}\sum_{b \in \Sigma^S}\left|\Pr_{\alpha \in_R A_w}[(T_{Q,\pi}(\alpha))_S = b] - \Pr_{\alpha'' \in_R \mathcal{A}_{w''}}[(T_{Q,\pi}(\alpha''))_S = b]\right|\right] \\
&= \mathop{\mathbf{E}}_{\pi}\left[\frac{1}{2}\sum_{b \in \Sigma^S}\left|\Pr_{\alpha \in_R A_w}[\alpha_{\pi(S)} = b] - \Pr_{\alpha'' \in_R \mathcal{A}_{w''}}[\alpha''_{\pi(S)} = b]\right|\right] \\
&\leq \delta,
\end{aligned}$$

where $\pi : Q \to [k_{max}]$ is a random 1-1 function, $w = w_{S,\pi}$, $w'' = w_{Q,\pi}$ and the last inequality follows from the fact that $w_{\pi(S)} = w''_{\pi(S)}$ and our assumption that $\mathcal{A}$ is δ-no-signaling. Thus, $\mathcal{B}$ is δ-no-signaling. This concludes the proof of Claim 12.2 □

This concludes the proof of Lemma 12.1 □

13 A No-Signaling MIP for PSPACE with an Inefficient Prover

In this section we construct MIP protocols that have no-signaling soundness for languages that can be computed in bounded space. The protocol's (honest) provers are inefficient and run in time exponential in the space bound. This protocol will prove useful in Section 14 where we apply it to logspace computations (so that the provers run in polynomial time). We note that a similar result was obtained both by [KR09] and (independently) by [IKM09].

As a first step we show how to construct MIPs with no-signaling soundness for languages in IP (Lemma 13.1). We later use (a strong version, due to [GKR08], of) the IP = PSPACE [LFKN92, Sha92] theorem to obtain the required result (Lemma 13.3).

Lemma 13.1. *If a language $\mathcal{L}$ has an ℓ-round public-coin interactive proof-system with soundness ϵ (and perfect completeness), then for every $\delta \geq 0$, the language $\mathcal{L}$ has a 1-round ℓ-prover* MIP *with soundness $\epsilon + \delta\ell$ against δ-no-signaling strategies. If Λ is the length of the longest message in the interactive proof then the* MIP *has query and answer alphabet $\{0,1\}^{\ell \cdot \Lambda}$.*

Furthermore, if the running time of the interactive-proof verifier is T_V then the running time of the MIP *verifier is $O(\ell \cdot T_V)$. If the running time of the interactive-proof prover is T_P then the running time of each* MIP *prover is $O(\ell \cdot T_P)$.*

Proof. Let $\delta \geq 0$ and let (P, V) be an ℓ-round public-coin interactive proof for a language $\mathcal{L}$. Let Λ be the length of the longest message in the interactive proof. Let m_i denote the message sent from the verifier to the prover in the i^{th} round and let b_i denote the prover's response to m_i. Since the protocol is public-coin, we assume that the messages $m_1, \ldots, m_\ell$ are generated by the verifier in the beginning of the protocol and in particular, they do not depend on the prover's answers. We also assume without loss of generality that the *honest* prover's response b_i to the i^{th} message m_i depends only on m_i and x (and not on $m_1, \ldots, m_{i-1}$).[16] We construct a 1-round MIP $(V', P'_1, \ldots, P'_\ell)$ with δ no-signaling soundness for $\mathcal{L}$ as follows.

The verifier V' generates the ℓ messages $m_1, \ldots, m_\ell$ and for every $i \in [\ell]$, it sends m_i to the prover P'_i. The prover P'_i answers the query m_i by b_i which is computed by the next message function of P at round i and with respect to m_i and x. To decide whether to accept, the verifier V' simply runs $V(x, m_1, \ldots, m_\ell, b_1, \ldots, b_\ell)$.

To show that (perfect) completeness holds, observe that for $x \in L$ the probability that V' outputs 1 after interacting with $P'_1, \ldots, P'_\ell$ equals the probability that V outputs 1 after interacting with P. We proceed to prove that no-signaling soundness holds.

Suppose toward a contradiction that there exists a δ-no-signaling cheating strategy $\{\mathcal{A}_q\}_{q \in (\{0,1\}^\Lambda)^\ell}$ that breaks the soundness of V' with probability $\epsilon + \delta\ell$. We use the latter to construct a cheating prover P^* for the interactive proof that breaks soundness with probability at least ϵ (contradicting our assumption on the soundness of V).

The cheating prover P^* is defined as follows. Given V's first message m_1, the cheating prover selects at random $(b_1^{(1)}, \ldots, b_\ell^{(1)}) \in_R \mathcal{A}_{m_1,*,\ldots,*}$, where $*$ denotes an arbitrary fixed string (e.g., the string 0^Λ). It saves only $b_1 \stackrel{\text{def}}{=} b_1^{(1)}$ and sends b_1 to the verifier. The verifier answers with m_2. After receiving m_2, the prover selects $(b_1^{(2)}, \ldots, b_\ell^{(2)}) \in_R \mathcal{A}_{m_1,m_2,*,\ldots,*}$ conditioned on $b_1^{(2)} = b_1$. It saves only $b_2 \stackrel{\text{def}}{=} b_2^{(2)}$ and sends b_2 to the verifier. Generally, after getting the i^{th} message m_i, the prover selects $(b_1^{(i)}, \ldots, b_\ell^{(i)}) \in_R \mathcal{A}_{m_1,\ldots,m_i,*,\ldots,*}$ conditioned on $b_1^{(i)}, \ldots, b_{i-1}^{(i)} = b_1, \ldots, b_{i-1}$ and sends $b_i \stackrel{\text{def}}{=} b_i^{(i)}$ to the prover.

Before proceeding we note that in the above process it might happen that the conditional probability space is empty. In such a case the prover P^* just sends a special symbol $\perp$.

[16] This can be easily achieved by having the verifier resend its previous messages at every round. Note that this increases the length of each message by a factor of ℓ.

We show that for every ℓ messages $m_1, \ldots, m_\ell$, the distribution of the answers $b_1, \ldots, b_\ell$ described above is $\delta\ell$-close to the distribution $\mathcal{A}_{m_1,\ldots,m_\ell}$. This follows from the following claim by setting $i = \ell$.

Claim 13.2. *Fix ℓ messages $m_1, \ldots, m_\ell \in (\{0,1\}^\Lambda)^\ell$. Let B_i denote the distribution of the first i elements in $\mathcal{A}_{m_1,\ldots,m_i,*,\ldots,*}$. Then, for every $0 \leq i \leq \ell$, the distribution $(b_1, \ldots, b_i)$ is δi-close to B_i.*

Proof. We prove the claim by induction. The base case $i = 0$ is trivial and so we proceed to the inductive step. Suppose that the claim holds for some i. For every $\beta_1, \ldots, \beta_i \in \{0,1\}^\Lambda$, consider the random variable $X_{i+1}(\beta_1, \ldots, \beta_i)$ defined by the following random process: select $(z_1, \ldots, z_{i+1})$ according to the distribution B_{i+1} conditioned on $z_1, \ldots, z_i = \beta_1, \ldots, \beta_i$ and output z_{i+1}. As before, if the conditional probability space is empty then output $\perp$.

Note that by the definition of P^*, the message b_{i+1} is distributed exactly as $X_{i+1}(b_1, \ldots, b_i)$. Therefore, by the inductive hypothesis, the distributions

- $b_1, \ldots, b_{i+1}$; and
- $B_i, X_{i+1}(B_i)$

are δi-close. Since $\mathcal{A}$ is δ-no-signaling, the distribution B_i is δ-close to the distribution obtained by taking the i first elements of B_{i+1}. Therefore, the distributions

- $B_i, X_{i+1}(B_i)$; and
- B_{i+1}

are δ-close. Thus, $(b_1, \ldots, b_{i+1})$ and B_{i+1} are $\delta(i+1)$-close. This completes the proof of Claim 13.2. □

By our assumption, the soundness of V' is violated with probability $\epsilon + \delta\ell$ when the answers that it receives are distributed according to $\mathcal{A}_{m_1,\ldots,m_\ell}$. Therefore, by Claim 13.2, the soundness of V is violated with probability at least ϵ when it receives the answers $b_1, \ldots, b_\ell$, in contradiction to our assumption on the soundness of V. □

Using Lemma 13.1, we can prove the following useful lemma.

Lemma 13.3. *If $\mathcal{L}$ can be computed by a Turing machine in space $s \stackrel{\text{def}}{=} s(n) \geq n$ (where n is the input length) then, for every $t \geq 1$, the language $\mathcal{L}$ has a $\mathsf{poly}(s)$-prover MIP with soundness $\mathsf{poly}(s) \cdot 2^{-t}$ against 2^{-t}-no-signaling strategies. The query and answer alphabets are $\{0,1\}^{t \cdot \mathsf{poly}(s)}$.*

Furthermore, the verifier runs in time $t \cdot \mathsf{poly}(s)$ and the (honest) provers run in time $t \cdot \mathsf{poly}(2^s)$.

Proof. Goldwasser *et al.* [GKR08] (see also [Rot09, Corollary 3.4.8]) show that if $\mathcal{L}$ can be computed in space s, then $\mathcal{L}$ has a $\mathsf{poly}(s)$-round public-coin interactive proof with soundness error $\frac{1}{2}$. The verifier's running time is $\mathsf{poly}(s)$ and the prover's running time is $\mathsf{poly}(2^s)$. The length of each message is $\mathsf{poly}(s)$.[17]

To this base protocol we apply a $O(t)$-fold parallel repetition (see, e.g., [Gol08, Exercise 9.1] or [Gol99, Appendix C.1]) which produces a $\mathsf{poly}(s)$-round public-coin interactive proof with soundness error 2^{-2t}. The verifier's running time is $t \cdot \mathsf{poly}(s)$ and the prover's running time is $t \cdot \mathsf{poly}(2^s)$. The length of each message is $t \cdot \mathsf{poly}(s)$. The lemma follows by applying Lemma 13.1 with $\delta = 2^{-t}$. □

14 Simulating an MIP Oracle

In this section we show that if a language $\mathcal{L}$ has an MIP with soundness against no-signaling strategies relative to an oracle $\{\phi_n\}$ and the function $\{\phi_n\}$ can be computed by a Turing machine that uses only a small amount of space, then the oracle can essentially be simulated and $\mathcal{L}$ has an MIP with soundness against no-signaling strategies *without* an oracle.

Lemma 14.1. *Let $\mathcal{L}$ be a language and suppose that $\mathcal{L}$ has an* MIP *relative to an oracle $\{\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}\}_n$ (where n is the input length) with soundness error ϵ against δ-no-signaling strategies. Let k be the number of provers and ℓ be the number of oracle queries used by the* MIP*. Suppose further that the function $\{\phi_n\}$ can be computed by a Turing machine in linear space (i.e., in space $O(n')$). Then, for every $t \geq 1$, the language $\mathcal{L}$ has an* MIP *protocol* without an oracle *that has soundness $\epsilon + \ell \cdot \mathsf{poly}(n^*) \cdot 2^{-t}$, where $n^* = n' + \log(n'')$, against $\min(\delta, 2^{-t})$-no-signaling strategies. The resulting* MIP *uses $k + \ell \cdot \mathsf{poly}(n^*)$ provers.*

Furthermore, if the original MIP *verifier runs in time T_V then the resulting verifier runs in time $T_V + O(\ell \cdot t \cdot \mathsf{poly}(n^*))$. If the original* MIP *provers run in time T_P then the resulting provers run in time $T_P + O(\ell \cdot t \cdot \mathsf{poly}(2^{n^*}))$. If the original* MIP *has query alphabet D and answer alphabet Σ then the resulting* MIP *has query alphabet $D \cup \{0,1\}^{t \cdot \mathsf{poly}(n^*)}$ and answer alphabet $\Sigma \cup \{0,1\}^{t \cdot \mathsf{poly}(n^*)}$.*

The high level approach is to use Lemma 13.3 to transform each oracle query into an additional MIP with no-signaling soundness and to show that composing these MIP protocols maintains the no-signaling soundness. The rest of this section is devoted to the (straightforward and somewhat tedious) proof of Lemma 14.1.

Simulating a single query. To prove Lemma 14.1, we first show that if the oracle can be computed by an MIP protocol with soundness against no-signaling strategies, then the oracle queries can be removed one by one (Lemma 14.2). For simplicity and since it suffices for our purposes, in the following we replace the oracle $\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}$ with an equivalent

[17] Indeed, the advantage in using the [GKR08] protocol is that the running time of the prover is $\mathsf{poly}(2^s)$ rather than $2^{\mathsf{poly}(s)}$ as in the classical [LFKN92, Sha92] protocol.

oracle $\phi'_n : \{0,1\}^{n^*} \to \{0,1\}$ that returns Boolean valued answers, where $n^* = n' + \log(n'')$. The oracle ϕ'_n on input $(z,i) \in \{0,1\}^{n'+\log(n'')}$ simply outputs the i-th bit of $\phi_n(z)$.

We note that the requirement in Lemma 14.2 will be that the oracle function $\{\phi'_n\}$ can be *computed*, rather than *decided*, by an MIP protocol (with soundness against no-signaling strategies). This means that both the language $\mathcal{L}_{\phi'} = \{z \in \{0,1\}^{n^*} : \phi'_n(z) = 1\}$ and the complement language $\overline{\mathcal{L}}_{\phi'}$ have MIP protocols with no-signaling soundness. However, it will be convenient for us to assume that there is a *single* protocol for computing $\{\phi'_n\}$ with no-signaling soundness, a notion that will be defined next. Indeed, as will be shown in Claim 14.5, the existence of MIP protocols with no-signaling soundness for both $\mathcal{L}_{\phi'}$ and $\overline{\mathcal{L}}_{\phi'}$ implies a single protocol for computing $\{\phi'_n\}$.

Multi-prover protocols for computing a function. In a one-round k-prover interactive protocol for *computing* a function $f : \{0,1\}^* \to \{0,1\}^*$, there are k computationally unbounded provers, $P_1, \ldots, P_k$, that try to convince a (probabilistic) polynomial-time verifier, V, of the value of $f(x)$ where the input $x \in \{0,1\}^*$ is known to all parties.

The interaction is similar to that in a one-round MIP (see Section 4.2). Given x and her random string, the verifier generates k queries, $q_1, \ldots, q_k$, one for each prover, and sends them to the k provers. The provers respond with answers $a_1, \ldots, a_k$. Finally, the verifier, based on the answers that she receives (as well as the input x and her random string), either outputs a value (which is supposed to equal $f(x)$) or outputs a special abort symbol $\perp$.

Denote by D the query alphabet and by Σ the answer alphabet. We say that $(V, P_1, \ldots, P_k)$ is a one-round k-prover protocol for computing f, with soundness ϵ against δ-no-signaling strategies, if the following two properties are satisfied:

1. **Completeness:** For every $x \in \{0,1\}^*$, the verifier V outputs $f(x)$ with probability 1, after interacting with $P_1, \ldots, P_k$.
2. **Soundness:** For every $x \in \{0,1\}^*$, and any δ-no-signaling family of distributions $\{\mathcal{A}_q\}_{q \in D^k}$ (where $\mathcal{A}_q$ is distributed over Σ^k, for every $q \in D^k$), with probability $\geq 1-\epsilon$, the verifier V outputs either $f(x)$ or $\perp$, where on queries $q = (q_1, \ldots, q_k)$ the answers are given probabilistically by $(a_1, \ldots, a_k) \in_R \mathcal{A}_q$.

We are now ready to state and prove Lemma 14.2.

Lemma 14.2. *Let $\mathcal{L}$ be a language and suppose that $\mathcal{L}$ has an MIP relative to an oracle $\{\phi'_n : \{0,1\}^{n^*} \to \{0,1\}\}_n$ (where n is the input length) with soundness ϵ against δ-no-signaling strategies. Let k be the number of provers and $\ell > 0$ be the number of oracle queries used by the MIP. Suppose further that the function $\{\phi'_n\}$ can be computed by a one-round k'-prover protocol with soundness ϵ' against δ'-no-signaling strategies. Then, $\mathcal{L}$ has an MIP relative to the same oracle $\{\phi'_n\}$ with ϵ'' soundness against δ''-no-signaling strategies where $\epsilon'' = \epsilon + \epsilon' + \delta''$ and $\delta'' = \min(\delta, \delta')$. The resulting MIP uses $k + k'$ provers but only $\ell - 1$ oracle queries.*

Furthermore, if the original MIP *verifier for* $\mathcal{L}$ *runs in time* T_V*, and the verifier of the* MIP *for* $\{\phi'_n\}$ *runs in time* T'_V*, then the resulting verifier runs in time* $T_V + O(T'_V + k'n^*)$*. If the provers of the* MIP *for* $\mathcal{L}$ *run in time* T_P *and the provers of the* MIP *for* $\{\phi'_n\}$ *run in time* T'_P *then the resulting provers run in time* $T_P + O(T'_P + n^*)$*. If the original* MIP *has query alphabet* D *and answer alphabet* Σ *and the oracle* MIP *has query alphabet* D' *and answer alphabet* Σ' *then the resulting* MIP *has query alphabet* $D'' = D \cup (\{0,1\}^{n^*} \times D')$ *and answer alphabet* $\Sigma'' = \Sigma \cup \Sigma'$*.* [18]

The high level idea is that if the oracle can be computed by a no-signaling MIP protocol then an oracle query can just be simulated by adding sufficiently many provers and running the multi-prover protocol for the oracle function with the additional provers. The no-signaling soundness property guarantees that revealing the oracle query to the provers does not harm soundness (too much).

Proof of Lemma 14.2. Let $(V, P_1, \ldots, P_k)$ be the MIP for $\mathcal{L}$ relative to the oracle $\{\phi'_n : \{0,1\}^{n^*} \to \{0,1\}\}$ that has soundness ϵ against δ-no-signaling strategies. Let D be the query alphabet and Σ the answer alphabet. Let $(V', P'_1, \ldots, P'_{k'})$ be the k'-prover interactive-protocol for computing $\{\phi'_n\}$ with soundness ϵ' against δ'-no-signaling strategies. Recall that this means that when interacting with the honest provers, $V'(z)$ outputs $\phi'_n(z)$ (with probability 1) and that no δ'-no-signaling cheating strategy can convince $V'(z)$ to output anything other than $\phi'_n(z)$ or $\perp$, with probability greater than ϵ'. Let D' be the query alphabet and Σ' the answer alphabet of $(V', P'_1, \ldots, P'_{k'})$. Let $\Sigma'' = \Sigma \cup \Sigma'$. It will be convenient for us to extend the answer alphabets of both protocols to Σ''.[19]

Since we deal with 1-round protocols, it will be convenient to think of each one of our verifiers as being composed of two algorithms (that share their randomness) - the query generation step and the verification step. Specifically, we think of V as being composed of two algorithms V_1 and V_2. The first algorithm, V_1, on input x and a random string r, outputs a sequence of k prover-queries $\mathbf{q} \in D^k$ and a sequence of ℓ oracle-queries $\mathbf{q}' \in (\{0,1\}^{n^*})^\ell$. The second algorithm, V_2, on input x, the same random string r, prover answers $\mathbf{a} \in (\Sigma'')^k$ and oracle answers $\mathbf{b} \in \{0,1\}^\ell$ outputs a bit representing whether $x \in \mathcal{L}$. Similarly, we think of V' as being composed of two algorithms V'_1 and V'_2. The first algorithm, V'_1, on input $q^* \in \{0,1\}^{n^*}$ and a random string s, outputs a sequence of k' queries $\mathbf{w} \in (D')^{k'}$. The second algorithm, V'_2, on input q^*, the same random string s, and answers $\mathbf{z} \in (\Sigma'')^{k'}$, outputs the result (which is supposed to be equal to $\phi'_n(q^*)$).

We construct an MIP protocol for $\mathcal{L}$ with only $\ell - 1$ oracle queries (but using $k + k'$ provers) as follows. The verifier V'' is composed of two steps, where V''_1 denotes the query generation step and V''_2 denotes the verification step. The first algorithm, V''_1, on input x and the random string (r, s), first invokes $V_1(x, r)$ to obtain k prover-queries $\mathbf{q} = (q_1, \ldots, q_k) \in D^k$ and ℓ oracle-queries $\mathbf{q}' = (q'_1, \ldots, q'_\ell) \in (\{0,1\}^{n^*})^\ell$. For every $i \in [k]$, the query q_i is sent directly to the i-th prover and the oracle queries $q'_2, \ldots, q'_\ell$ are sent directly to the oracle

[18] Note that we do not assume that $D \cap (\{0,1\}^{n^*} \times D') = \emptyset$ nor that $\Sigma \cap \Sigma' = \emptyset$.

[19] This can be done by having the verifiers reject immediately if they see symbols that are not in their original alphabets.

ϕ'_n. The query $q^* \stackrel{\text{def}}{=} q'_1$ is handled differently (to avoid an additional oracle query). The k' additional prover queries are generated by invoking $V'_1(q^*, s)$, to obtain a sequence of k' queries $\mathbf{w} = (w_1, \ldots, w_{k'}) \in (D')^{k'}$. For every $i \in [k']$ the query (q^*, w_i) is sent to the $(k+i)$-th prover.

We denote the query alphabet by $D'' \stackrel{\text{def}}{=} D \cup (\{0,1\}^{n^*} \times D')$. For every sequence $\omega \in (D')^{k'}$ we denote by $\overline{\omega}(q^*) \stackrel{\text{def}}{=} (q^*, \omega_1), \ldots, (q^*, \omega_{k'}) \in (D'')^{k'}$. Thus, the sequence of queries sent by the verifier is $(\mathbf{q}, \overline{\mathbf{w}}(q^*)) \in (D'')^{k+k'}$.

The honest provers operate as follows. The first k provers operate exactly the same as the provers $P_1, \ldots, P_k$ in the original MIP for $\mathcal{L}$. That is, for every $i \in [k]$, the i-th prover, on input x and query q_i, answers with $a_i = P_i(x, q_i)$. We denote $\mathbf{a} \stackrel{\text{def}}{=} (a_1, \ldots, a_k) \in (\Sigma'')^k$. The last k' provers answer their queries as follows. For every $i \in [k']$, the $(k+i)$-th prover, given the query (q^*, w_i), answers its query with $z_i = P'_i(q^*, w_i)$. We denote $\mathbf{z} \stackrel{\text{def}}{=} (z_1, \ldots, z_{k'}) \in (\Sigma'')^{k'}$.

To decide whether to accept, on input x, the random string (r, s), prover answers $(\mathbf{a}, \mathbf{z}) \in (\Sigma'')^{k+k'}$, and oracle answers $b_2, \ldots, b_\ell \in \{0,1\}$, the algorithm V''_2 first recomputes q^* from x and r and computes $b^* \stackrel{\text{def}}{=} V'_2(q^*, s, \mathbf{z})$. If $b^* = \perp$, then V''_2 rejects. Otherwise, V''_2 outputs the result of $V_2(x, r, \mathbf{a}, \mathbf{b})$, where $\mathbf{b} = (b^*, b_2, \ldots, b_\ell)$. In other words, the verifier computes the result of the original verification process when the answers to $\mathbf{q}$ are $\mathbf{a}$, the answer to the first oracle query is b^* and the answers to the rest of the oracle queries $q'_2, \ldots, q'_\ell$ are (respectively) $b_2, \ldots, b_\ell$.

We first argue that the resulting MIP has perfect completeness and then proceed to prove soundness against δ''-no-signaling strategies. To show that completeness holds, observe that when the verifier V'' interacts with the honest provers on input $x \in \mathcal{L}$, since the protocol $(V', P'_1, \ldots, P'_{k'})$ has perfect completeness, it holds that $b^* = \phi'_n(q^*)$ and therefore V''_2 runs V_2 with the correct oracle answers. The completeness of the protocol follows from the completeness of $(V, P_1, \ldots, P_k)$.

To show that δ''-no-signaling soundness holds, assume for a contradiction that there exists some δ''-no-signaling cheating strategy $\{\mathcal{A}_{(\chi,\omega)}\}_{(\chi,\omega)\in(D'')^{k+k'}}$ that fools V'' into accepting $x \notin \mathcal{L}$ with probability ϵ''. That is,

$$\Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V''_2(x, (r,s), (\mathbf{a}, \mathbf{z}), (b_2, \ldots, b_\ell)) = 1\right] \geq \epsilon''$$

where $\mathbf{q}, q^*, \mathbf{w}, b_2, \ldots, b_\ell$ are constructed as above. Using elementary manipulations we have that

$$
\begin{aligned}
\epsilon'' \leq & \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V_2''(x,(r,s),(\mathbf{a},\mathbf{z}),(b_2,\ldots,b_\ell))=1\right] \\
= & \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[\left(V_2''(x,(r,s),(\mathbf{a},\mathbf{z}),(b_2,\ldots,b_\ell))=1\right) \wedge \left(b^* \neq \phi_n'(q^*)\right)\right] + \\
& \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[\left(V_2''(x,(r,s),(\mathbf{a},\mathbf{z}),(b_2,\ldots,b_\ell))=1\right) \wedge \left(b^* = \phi_n'(q^*)\right)\right] \\
\leq & \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[b^* \notin \{\phi_n'(q^*),\perp\}\right] + \\
& \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[\left(V_2''(x,(r,s),(\mathbf{a},\mathbf{z}),(b_2,\ldots,b_\ell))=1\right) \wedge \left(b^* = \phi_n'(q^*)\right)\right] \qquad (7)
\end{aligned}
$$

where $\mathbf{q}, q^*, \mathbf{w}, b_2, \ldots, b_\ell$ are as above and $b^* = V_2'(q^*, s, \mathbf{z})$. Lemma 14.2 follows from the following two claims (Claim 14.3 and Claim 14.4), that analyze the last two terms separately.

Claim 14.3.

$$
\Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[b^* \notin \{\phi_n'(q^*),\perp\}\right] < \epsilon'
$$

Proof. Suppose otherwise. That is:

$$
\Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V_2'(q^*,s,\mathbf{z}) \notin \{\phi_n'(q^*),\perp\}\right] \geq \epsilon'.
$$

Then, by an averaging argument, there exists a fixed value of r for which:

$$
\Pr_{\substack{s \\ (\mathbf{a},\mathbf{z})\in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V_2'(q^*,s,\mathbf{z}) \notin \{\phi_n'(q^*),\perp\}\right] \geq \epsilon' \qquad (8)
$$

where $\mathbf{q}, q^*$ are fixed (based on the value of r), and $\mathbf{w} = V_1'(q^*, s)$. For the rest of the proof of Claim 14.3, we use $r, \mathbf{q}, q^*$ to refer to the foregoing fixed values.

We construct a δ'-no-signaling strategy $\mathcal{B} = \{\mathcal{B}_\omega\}_{\omega\in(D')^{k'}}$ that on input q^*, fools V' into outputting a value that is neither $\phi_n'(q^*)$ nor $\perp$, with probability $\geq \epsilon'$, contradicting our assumption on the soundness of $(V', P_1', \ldots, P_{k'}')$. For every $\omega \in (D')^{k'}$, the distribution $\mathcal{B}_\omega$ is defined by sampling $(\mathbf{a}, \mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\omega}(q^*))}$ and outputting $\mathbf{z}$.

To show that $\mathcal{B}$ violates the δ'-no-signaling soundness of $(V', P_1', \ldots, P_{k'}')$, note that by Eq. (8), the probability that V_2', on input q^*, the random string s and given answers $\mathbf{z} \in_R \mathcal{B}_{\mathbf{w}}$, where $\mathbf{w} = V_1'(q^*, s)$, outputs a value that is neither $\phi_n'(q^*)$ nor $\perp$ is at least ϵ'.

We proceed to show that $\mathcal{B}$ is δ'-no-signaling. Let $S \subset [k']$ and $\omega, \omega' \in (D')^{k'}$, such that $\omega_S = \omega'_S$. Suppose that the statistical distance between $\mathbf{z}_S$ and $\mathbf{z}'_S$ is more than δ, where

$\mathbf{z} \in_R \mathcal{B}_\omega$ and $\mathbf{z}' \in_R \mathcal{B}_{\omega'}$. Hence,

$$\delta' < \frac{1}{2} \sum_{\beta \in (\Sigma'')^S} \left| \Pr_{\mathbf{z} \in_R \mathcal{B}_\omega} [\mathbf{z}_S = \beta] - \Pr_{\mathbf{z}' \in_R \mathcal{B}_{\omega'}} [\mathbf{z}'_S = \beta] \right|$$

$$= \frac{1}{2} \sum_{\beta \in (\Sigma'')^S} \left| \Pr_{(\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\omega}(q^*))}} [\mathbf{z}_S = \beta] - \Pr_{(\mathbf{a}',\mathbf{z}') \in_R \mathcal{A}_{(\mathbf{q},\overline{\omega'}(q^*))}} [\mathbf{z}'_S = \beta] \right|. \tag{9}$$

Let $S' = \{k + i : i \in S\}$. Then, by Eq. (9) the projections of the distributions $\mathcal{A}_{(\mathbf{q},\overline{\omega}(q^*))}$ and $\mathcal{A}_{(\mathbf{q},\overline{\omega'}(q^*))}$ to coordinates in S' are δ'-far. Since $(\mathbf{q}, \overline{\omega}(q^*))_{S'} = (\mathbf{q}, \overline{\omega'}(q^*))_{S'}$ and $\delta'' \leq \delta'$, this contradicts our assumption that $\mathcal{A}$ is δ''-no-signaling.

This concludes the proof of Claim 14.3. □

Claim 14.4.

$$\Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[\left(V''_2(x, (r,s), (\mathbf{a}, \mathbf{z}), (b_2, \ldots, b_\ell)) = 1 \right) \wedge \left(b^* = \phi'_n(q^*) \right) \right] < \epsilon + \delta''$$

Proof. Suppose otherwise. Thus, by the definition of V''_2,

$$\begin{aligned}
\epsilon + \delta'' \leq & \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[\left(V''_2(x, (r,s), (\mathbf{a}, \mathbf{z}), (b_2, \ldots, b_\ell)) = 1 \right) \wedge \left(b^* = \phi'_n(q^*) \right) \right] \\
= & \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V''_2(x, (r,s), (\mathbf{a}, \mathbf{z}), (b_2, \ldots, b_\ell)) = 1 \;\middle|\; b^* = \phi'_n(q^*) \right] \cdot \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[b^* = \phi'_n(q^*) \right] \\
= & \Pr_{\substack{r,s, \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V_2(x, r, \mathbf{a}, \phi'_n(\mathbf{q}')) = 1 \;\middle|\; b^* = \phi'_n(q^*) \right] \cdot \Pr_{\substack{r,s \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[b^* = \phi'_n(q^*) \right] \\
= & \Pr_{\substack{r,s, \\ (\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}} \left[V_2(x, r, \mathbf{a}, \phi'_n(\mathbf{q}')) = 1 \right] \qquad (10)
\end{aligned}$$

where $\mathbf{q}, \mathbf{q}', q^*, \mathbf{w}, b^*$ are as above, and $\phi'_n(\mathbf{q}') = (\phi'_n(q'_1), \ldots, \phi'_n(q'_\ell))$ (i.e., the correct oracle answers).

We argue that Eq. (10) contradicts the δ-no-signaling soundness of V. Toward this end, we construct a cheating strategy $\mathcal{B} = \{\mathcal{B}_\chi\}_{\chi \in D^k}$ that fools V into accepting $x \notin \mathcal{L}$, with probability $\geq \epsilon$. Let $\sigma \in (D'')^{k'}$ be an arbitrary fixed value. For every $\chi \in D^k$, the distribution $\mathcal{B}_\chi$ is defined by sampling $(\mathbf{a}, \mathbf{z}) \in_R \mathcal{A}_{(\chi,\sigma)}$ and outputting $\mathbf{a}$.

We first show that $\mathcal{B}$ is δ-no-signaling and proceed to show that it violates the soundness of $(V, P_1, \ldots, P_k)$. Let $S \subset [k]$ and $\chi, \chi' \in D^k$ such that $\chi_S = \chi'_S$. Suppose toward a contradiction that the statistical distance between $\mathbf{a}_S$ and $\mathbf{a}'_S$ is more than δ, where $\mathbf{a} \in_R \mathcal{B}_\chi$ and $\mathbf{a}' \in_R \mathcal{B}_{\chi'}$. Hence,

$$\delta < \frac{1}{2} \sum_{\beta \in (\Sigma'')^S} \left| \Pr_{\mathbf{a} \in_R \mathcal{B}_\chi} [\mathbf{a}_S = \beta] - \Pr_{\mathbf{a}' \in_R \mathcal{B}_{\chi'}} [\mathbf{a}'_S = \beta] \right|$$

$$= \frac{1}{2} \sum_{\beta \in (\Sigma'')^S} \left| \Pr_{(\mathbf{a},\mathbf{z}) \in_R \mathcal{A}_{(\chi,\sigma)}} [\mathbf{a}_S = \beta] - \Pr_{(\mathbf{a}',\mathbf{z}') \in_R \mathcal{A}_{(\chi',\sigma)}} [\mathbf{a}'_S = \beta] \right|.$$

In particular, the projections of the distributions $\mathcal{A}_{(\chi,\sigma)}$ and $\mathcal{A}_{(\chi',\sigma)}$ to coordinates in S are δ-far. Since $(\chi,\sigma)_S = (\chi',\sigma)_S$ and $\delta'' \leq \delta$, this contradicts our assumption that $\mathcal{A}$ is δ''-no-signaling.

We proceed to show that $\{\mathcal{B}_\chi\}_{\chi\in D^k}$ fools V into accepting $x \notin \mathcal{L}$ with probability $\geq \epsilon$. Assume for a contradiction that

$$\Pr_{\substack{r,\\ \mathbf{a}'\in_R\mathcal{B}_\mathbf{q}}}\left[V_2(x,r,\mathbf{a}',\phi'_n(\mathbf{q}'))=1\right] < \epsilon. \tag{11}$$

where $\mathbf{q},\mathbf{q}'$ are as above. Combining Eq. (10) and Eq. (11) we have that:

$$\mathop{\mathbf{E}}_{r,s}\left[\Pr_{(\mathbf{a},\mathbf{z})\in_R\mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}\left[V_2(x,r,\mathbf{a},\phi'_n(\mathbf{q}'))=1\right] - \Pr_{(\mathbf{a}',\mathbf{z}')\in_R\mathcal{A}_{(\mathbf{q},\sigma)}}\left[V_2(x,r,\mathbf{a}',\phi'_n(\mathbf{q}'))=1\right]\right] > \delta''$$

where $\mathbf{q},q^*,\mathbf{q}',\phi'_n(\mathbf{q}')$ are as above and $\mathbf{w} = V'_1(q^*,s)$.

By an averaging argument, there exist *fixed* values for r and s such that:

$$\Pr_{(\mathbf{a},\mathbf{z})\in_R\mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}}\left[V_2(x,r,\mathbf{a},\phi'_n(\mathbf{q}'))=1\right] - \Pr_{(\mathbf{a}',\mathbf{z}')\in_R\mathcal{A}_{(\mathbf{q},\sigma)}}\left[V_2(x,r,\mathbf{a}',\phi'_n(\mathbf{q}'))=1\right] > \delta'' \tag{12}$$

where $\mathbf{q},q^*,\mathbf{q}',\phi'_n(\mathbf{q}')$ and $\mathbf{w}$ are *fixed* based on the fixed values of r and s.

Equation (12) gives a statistical test that distinguishes between the projections of the distributions $\mathcal{A}_{(\mathbf{q},\sigma)}$ and $\mathcal{A}_{(\mathbf{q},\overline{\mathbf{w}}(q^*))}$ to coordinates in $[k]$ with gap $> \delta''$, contradicting our assumption that $\mathcal{A}$ is δ''-no-signaling. Hence, $\mathcal{B}$ fools V into accepting $x \notin \mathcal{L}$ with probability $\geq \epsilon$. This concludes the proof of Claim 14.4. □

This concludes the proof of Lemma 14.2. □

To prove Lemma 14.1 we also need the following straightforward claim.

Claim 14.5. *Let $\mathcal{L}$ be a language and suppose that both $\mathcal{L}$ and $\overline{\mathcal{L}}$ (i.e., the complement language of $\mathcal{L}$) have* MIP *protocols with soundness ϵ against δ-no-signaling strategies. Assume that each of the* MIP *protocols uses k provers. Then, there exists a $2k$-prover interactive protocol for* computing *the function $\mathcal{L}(x) : \{0,1\}^* \to \{0,1\}$, where $\mathcal{L}(x) = 1$ if and only if $x \in \mathcal{L}$, with soundness ϵ against δ-no-signaling strategies. If both of the original* MIP *protocols use query alphabet D and answer alphabet Σ then the resulting $2k$-prover protocol has query alphabet D and answer alphabet $\Sigma \cup \{\perp\}$, where $\perp \notin \Sigma$ is a special symbol.*

Furthermore, if each of the original MIP *verifiers runs in time T_V then the resulting verifier runs in time $O(T_V + k \cdot \log(|\Sigma|))$ and if each of the original* MIP *(honest) provers runs in time T_P then the resulting provers run in time $O(T_P + T_\mathcal{L} + \log(|\Sigma|))$, where $T_\mathcal{L}$ is the time that it takes for a Turing machine to compute $\mathcal{L}(x)$.*

Proof. Let $(V,P_1,\ldots,P_k)$ be the MIP for $\mathcal{L}$ and let $(V',P'_1,\ldots,P'_k)$ be the MIP for $\overline{\mathcal{L}}$. We assume that V and V' are composed of two algorithms, a query generation algorithm and a verification algorithm. The query generation algorithm V_1 (resp., V'_1) on input x and a

random string r (resp., r'), outputs k queries $\mathbf{q} = (q_1, \ldots, q_k) \in D^k$ (resp., $\mathbf{q}' = (q'_1, \ldots, q'_k) \in D^k$). The verification algorithm V_2 (resp., V'_2), on input x, the same random string r (resp., r') and k answers $\mathbf{a} = (a_1, \ldots, a_k) \in \Sigma^k$ (resp., $\mathbf{a}' = (a'_1, \ldots, a'_k) \in \Sigma^k$), outputs a bit representing whether to accept or reject the statement $x \in \mathcal{L}$ (resp., $x \notin \mathcal{L}$).

We construct a $2k$-prover protocol for *computing* $\mathcal{L}$ as follows. The first k provers are the same as $P_1, \ldots, P_k$ except that they first verify that $x \in \mathcal{L}$. If $x \notin \mathcal{L}$, then they answer with the special symbol $\perp$. Similarly, the last k provers are the same as $P'_1, \ldots, P'_k$ except that they first verify that $x \notin \mathcal{L}$. If $x \in \mathcal{L}$, then they send the special symbol $\perp$.

On input x and a random string (r, r') the query generation algorithm V''_1 computes $\mathbf{q} = V_1(x, r)$ and $\mathbf{q}' = V'_1(x, r')$, where $\mathbf{q}, \mathbf{q}' \in D^k$. For every $i \in [k]$, the query q_i is sent to the i-th prover and the query q'_i is sent to the $(k+i)$-th prover. Given the provers' answers $(\mathbf{a}, \mathbf{a}') \in (\Sigma \cup \{\perp\})^{k+k}$, the verification algorithm V''_2 works as follows:

1. If $V_2(x, r, \mathbf{a}) = 1$ and all the entries of $\mathbf{a}'$ are equal to $\perp$, then output 1 and halt.[20]
2. If $V'_2(x, r', \mathbf{a}') = 1$ and all the entries of $\mathbf{a}$ are equal to $\perp$, then output 0 and halt.
3. Output $\perp$ and halt.

To see that completeness holds note that if $x \in \mathcal{L}$ then the last k provers will send $\perp$ and, by the completeness of $(V, P_1, \ldots, P_k)$ the verifier will output 1. If $x \notin \mathcal{L}$ then the first k provers will send $\perp$ and, by the completeness of $(V', P'_1, \ldots, P'_k)$, the verifier will output 0. We proceed to show that soundness against δ-no-signaling strategies holds.

Fix $x \in \{0,1\}^*$ and assume toward a contradiction that there exists a δ-no-signaling strategy $\{\mathcal{A}_{(u,u')}\}_{(u,u') \in D^{k+k}}$ such that

$$\Pr_{\substack{r,r' \\ (\mathbf{a},\mathbf{a}') \in_R \mathcal{A}_{(\mathbf{q},\mathbf{q}')}}} [V''_2(x, (r, r'), (\mathbf{a}, \mathbf{a}')) \notin \{\mathcal{L}(x), \perp\}] \geq \epsilon,$$

where $\mathbf{q} = V_1(x, r)$ and $\mathbf{q}' = V'_1(x, r')$. For simplicity let us assume that $x \notin \mathcal{L}$. The case that $x \in \mathcal{L}$ is handled analogously (using the soundness of V', rather than the soundness of V). Thus,

$$\Pr_{\substack{r,r' \\ (\mathbf{a},\mathbf{a}') \in_R \mathcal{A}_{(\mathbf{q},\mathbf{q}')}}} [V''_2(x, (r, r'), (\mathbf{a}, \mathbf{a}')) \notin \{0, \perp\}] \geq \epsilon.$$

In particular, by the definition of V'':

$$\Pr_{\substack{r,r' \\ (\mathbf{a},\mathbf{a}') \in_R \mathcal{A}_{(\mathbf{q},\mathbf{q}')}}} [V_2(x, r, \mathbf{a}) = 1] \geq \epsilon. \tag{13}$$

where $\mathbf{q} = V_1(x, r)$ and $\mathbf{q}' = V'_1(x, r')$.

[20] If one of the entries of $\mathbf{a}$ (resp., $\mathbf{a}'$) is $\perp$, then we define $V_2(x, r, \mathbf{a}) = 0$ (resp., $V'_2(x, r', \mathbf{a}') = 0$).

By an averaging argument, Eq. (13) implies that there exists a fixed value of r' such that

$$\Pr_{(\mathbf{a},\mathbf{a}')\in_R \mathcal{A}_{(\mathbf{q},\mathbf{q}')}}[V_2(x,r,\mathbf{a})=1] \geq \epsilon. \tag{14}$$

where $\mathbf{q}' = V_1'(x,r')$ is a fixed value and $\mathbf{q} = V_1(x,r)$. For the rest of the proof of Claim 14.5 we fix r' and $\mathbf{q}'$ as above.

We use $\mathcal{A}$ to construct a δ-no-signaling strategy $\mathcal{B} = \{\mathcal{B}_u\}_{u\in D^k}$ that fools V into accepting $x \notin \mathcal{L}$ with probability $\geq \epsilon$. For every $u \in D^k$, the distribution $\mathcal{B}_u$ is defined by sampling $(\mathbf{a},\mathbf{a}') \in_R \mathcal{A}_{(u,\mathbf{q}')}$ and outputting $\mathbf{a}$.

We first show that $\mathcal{B}$ violates the soundness of V_2 and then show that it is δ-no-signaling. Indeed, by the definition of $\mathcal{B}$ and using Eq. (14) it holds that

$$\Pr_{\mathbf{a}\in_R\mathcal{B}_\mathbf{q}}[V_2(x,r,\mathbf{a})=1] = \Pr_{(\mathbf{a},\mathbf{a}')\in_R\mathcal{A}_{(\mathbf{q},\mathbf{q}')}}[V_2(x,r,\mathbf{a})=1] \geq \epsilon.$$

We proceed to show that $\mathcal{B}$ is δ-no-signaling. Let $S \subset [k]$ and let $u_1, u_2 \in D^k$ such that $(u_1)_S = (u_2)_S$, and suppose that the statistical distance between $(\mathbf{a}_1)_S$ and $(\mathbf{a}_2)_S$ is more than δ, where $\mathbf{a}_1 \in_R \mathcal{B}_{u_1}$ and $\mathbf{a}_2 \in_R \mathcal{B}_{u_2}$. Then:

$$\begin{aligned}
\delta &< \frac{1}{2}\sum_{\beta\in\Sigma^S}\left|\Pr_{\mathbf{a}_1\in_R\mathcal{B}_{u_1}}[(\mathbf{a}_1)_S=\beta] - \Pr_{\mathbf{a}_2\in_R\mathcal{B}_{u_2}}[(\mathbf{a}_2)_S=\beta]\right| \\
&= \frac{1}{2}\sum_{\beta\in\Sigma^S}\left|\Pr_{(\mathbf{a}_1,\mathbf{a}_1')\in_R\mathcal{A}_{(u_1,\mathbf{q}')}}[(\mathbf{a}_1)_S=\beta] - \Pr_{(\mathbf{a}_2,\mathbf{a}_2')\in_R\mathcal{A}_{(u_2,\mathbf{q}')}}[(\mathbf{a}_2)_S=\beta]\right|.
\end{aligned}$$

Thus, the projections of the distributions $\mathcal{A}_{(u_1,\mathbf{q}')}$ and $\mathcal{A}_{(u_2,\mathbf{q}')}$ to coordinates in S are δ-far. Since $(u_1,\mathbf{q}')_S = (u_2,\mathbf{q}')_S$, this contradicts our assumption that $\mathcal{A}$ is δ-no-signaling.

This concludes the proof of Claim 14.5. □

Using Lemma 13.3, Lemma 14.2 and Claim 14.5 we are ready to prove Lemma 14.1.

Proof of Lemma 14.1. As a first step we replace the oracle $\{\phi_n : \{0,1\}^{n'} \to \{0,1\}^{n''}\}$ with a Boolean valued oracle $\{\phi_n' : \{0,1\}^{n^*} \to \{0,1\}\}$, where $n^* = n' + \log(n'')$, by having the oracle ϕ_n' on input $(z,i) \in \{0,1\}^{n'+\log(n'')}$ simply output the i-th bit of $\phi_n(z)$. Note that this step increases the number of oracle queries from ℓ to $\ell \cdot \log(n'')$.

Fix $t \geq 1$. Since $\{\phi_n'\}$ can be computed in linear (i.e., $O(n^*)$) space, by Lemma 13.3, the language $\mathcal{L}_{\phi'} = \{z \in \{0,1\}^* : \phi'(z) = 1\}$ has a $\mathsf{poly}(n^*)$-prover MIP with soundness error $\mathsf{poly}(n^*)\cdot 2^{-t}$ against 2^{-t}-no-signaling strategies. The verifier runs in time $t\cdot\mathsf{poly}(n^*)$ and the (honest) provers run in time $t\cdot\mathsf{poly}(2^{n^*})$. The query and answer alphabets are $\{0,1\}^{t\cdot\mathsf{poly}(n^*)}$. Similarly, the complement language $\overline{\mathcal{L}}_{\phi'}$ can also be computed in space $O(n^*)$ and so it has an MIP with the same parameters.

Thus, by Claim 14.5, there exists a $\mathsf{poly}(n^*)$-prover interactive protocol for computing $\{\phi_n'\}$ with soundness $\mathsf{poly}(n^*)\cdot 2^{-t}$ against 2^{-t}-no-signaling strategies. The verifier of the

resulting protocol runs in time $t \cdot \mathsf{poly}(n^*)$ and the provers run in time $t \cdot \mathsf{poly}(2^{n^*})$. The query and answer alphabets are $\{0,1\}^{t \cdot \mathsf{poly}(n^*)}$.

The lemma follows by applying Lemma 14.2 iteratively $\ell \cdot \log(n'')$ times to the original MIP for $\mathcal{L}$ to remove all of the oracle queries. The resulting MIP has soundness $\epsilon + \ell \cdot \mathsf{poly}(n^*) \cdot \left(2^{-t} + \min(\delta, 2^{-t})\right)$ against $\min(\delta, 2^{-t})$-no-signaling provers. The MIP uses $k + \ell \cdot \mathsf{poly}(n^*)$ provers. The verifier runs in time $T_V + O(\ell \cdot t \cdot \mathsf{poly}(n^*))$ and each prover runs in time $T_P + O(\ell \cdot t \cdot \mathsf{poly}(2^{n^*}))$. The query alphabet is $D \cup \{0,1\}^{t \cdot \mathsf{poly}(n^*)}$ and the answer alphabet is $\Sigma \cup \{0,1\}^{t \cdot \mathsf{poly}(n^*)}$.[21] □

15 Proof of Theorem 4

Using the tools developed in the previous sections, we are finally ready to prove Theorem 4.

Theorem 4. *Suppose that $\mathcal{L} \in \mathsf{DTIME}(t(n))$, where $t = t(n)$ satisfies $\mathsf{poly}(n) \leq t \leq \mathsf{exp}(n)$. Then, for any integer $(\log t)^c \leq k \leq \mathsf{poly}(n)$, where c is some (sufficiently large) universal constant, there exists an MIP for $\mathcal{L}$ with $k \cdot \mathsf{polylog}(t)$ provers and with soundness error 2^{-k} against $2^{-k \cdot \mathsf{polylog}(t)}$-no-signaling strategies.*

The verifier runs in time $(n + k^2) \cdot \mathsf{polylog}(t)$ and the provers run in time $\mathsf{poly}(t, k)$. Each query and answer is of length $k \cdot \mathsf{polylog}(t)$.

Proof. Let $\mathcal{L} \in \mathsf{DTIME}(t(n))$, where $\mathsf{poly}(n) \leq t(n) \leq \mathsf{exp}(n)$. Then, $\mathcal{L} \in \mathsf{DTISP}(t(n), s(n))$ where $\max(n, \log(t(n))) \leq s(n) \leq t(n)$. Fix $t = t(n)$ and $s = s(n)$.

Let $\mathcal{C}_n$ be a circuit on n inputs of size $N = O(t \cdot s)$ that computes $\mathcal{L}$ and let $\mathcal{C}'_n$ be the augmented circuit of size $N' = \mathsf{poly}(N)$, as described in Section 9. Let the parameters ℓ and $\mathbb{F}$ be as defined in Section 9.

Let $k' \leq \mathsf{poly}(n)$ be an integer such that $4|\mathbb{F}|^4 \leq k' \leq N'$. Consider the augmented PCP of Section 9, with respect to $\mathcal{C}'_n$, and the security parameter k'. Since $4|\mathbb{F}|^4 \leq k' \leq N'$, by Lemma 11.1, the PCP verifier has soundness ϵ' against (k_{max}, δ')-no-signaling strategies where:

[21] Note that the alphabet sizes do not increase on every iteration.

$$
\begin{aligned}
|\mathbb{F}| &\leq 8(\log(N')^{10} \\
\ell &= 3\frac{\log(N')}{\log\log(N')} + 3 \\
r &= \frac{k'}{40\ell|\mathbb{F}|} \\
\epsilon' &= 2^{-r/2} \\
k_{max} &= k' \cdot \mathsf{polylog}(s) \cdot \log(t)|\mathbb{F}| + 12k'\ell|\mathbb{F}|^2 \\
\delta' &= \frac{1}{|\mathbb{F}|^{8k'\ell|\mathbb{F}|^2}}.
\end{aligned}
$$

Recall that this PCP is relative to an oracle $\hat{\phi}' : \mathbb{F}^\ell \to \mathbb{F}$ (see Section 5 and Section 9). As noted in Section 5.2.1, the total number of PCP queries as well as the total number of oracle queries is at most $6k'\ell|\mathbb{F}|^2 \leq k'\mathsf{polylog}N'$ and the running time of the verifier is $k'\mathsf{polylog}N'$. As noted in Section 5.1.1 (see also Section 9), the PCP can be generated in time $\mathsf{poly}(N')$. The query alphabet is of size at most $\mathsf{poly}(N')$ and the answer alphabet is of size $|\mathbb{F}| \leq \mathsf{polylog}N'$.

As a first step, we transform the PCP into an MIP. By applying Lemma 12.1, we obtain an MIP (relative to the same oracle) with soundness ϵ' against δ'-no-signaling strategies. The MIP uses $k_{max} \leq k'\mathsf{polylog}N'$ provers and $k'\mathsf{polylog}N'$ oracle queries. The query and answer alphabets remain unchanged. The running time of the MIP verifier is: $O(k'\mathsf{polylog}N' + k_{max} \log N') \leq k'\mathsf{polylog}N'$. The running time of each MIP prover is $\mathsf{poly}(N')$.

Recall that $\hat{\phi}' = \hat{\phi}_x + \hat{\phi}_{\mathcal{C}'} + \hat{\phi}_{extra}$ where $\hat{\phi}_x$, $\hat{\phi}_{\mathcal{C}'}, \hat{\phi}_{extra}$ are the low degree extensions of ϕ_x, $\phi_{\mathcal{C}'}$ and ϕ_{extra} respectively (see Section 5 and Section 9). As our second step, we replace the use of the oracle $\hat{\phi}'$ in the MIP with the oracle $\hat{\phi}_{\mathcal{C}'} + \hat{\phi}_{extra}$. This is done by replacing each oracle query $\hat{\phi}'(z)$, by first querying the new oracle $(\hat{\phi}_{\mathcal{C}'} + \hat{\phi}_{extra})(z)$ and adding $\hat{\phi}_x(z)$, which is computed directly by the verifier, to the result. As noted in Section 5, the function $\hat{\phi}_x$ can be evaluated in time $n \cdot \mathsf{polylog}N'$. Thus, the resulting verifier runs in time $k'\mathsf{polylog}N' + n\mathsf{polylog}N'$ and all other parameters of the MIP remain unchanged.

At this point we apply Lemma 14.1 to obtain an MIP *without an oracle*. Note that, as pointed out in Section 5 (resp., Section 9), the function $\hat{\phi}_{\mathcal{C}'}$ (resp., $\hat{\phi}_{extra}$) can be computed in space that is linear in its input (i.e., $O(\log(|\mathbb{F}|^\ell)) = O(\log N')$ space). Therefore, we can apply Lemma 14.1, with respect to a parameter $t = \log_2(1/\delta')$, to obtain an MIP *without an oracle* that has soundness ϵ against δ-no-signaling strategies where

$$
\begin{aligned}
\epsilon &= \epsilon' + k'\delta'\mathsf{polylog}N' \\
\delta &= \delta'
\end{aligned}
$$

The resulting MIP uses $k'\mathsf{polylog}N'$ provers. The resulting MIP verifier runs in time

$$(n + k')\mathsf{polylog}N' + O(k'\mathsf{polylog}N' \log(1/\delta)\mathsf{polylog}N') \leq (n + k'^2)\mathsf{polylog}N'$$

and the resulting provers run in time

$$\mathsf{poly}(N') + O(k'\mathsf{polylog}N' \cdot \log(1/\delta) \cdot \mathsf{poly}(2^{O(\log N')})) \leq \mathsf{poly}(N').$$

The query alphabet is of size $\mathsf{poly}(N') + 2^{\log(1/\delta)\cdot\mathsf{polylog}(N')} \leq 2^{k'\cdot\mathsf{polylog}N'}$ and the answer alphabet is of size $\mathsf{polylog}N' + 2^{\log(1/\delta)\cdot\mathsf{polylog}(N')} \leq 2^{k'\cdot\mathsf{polylog}N'}$.

The theorem follows by setting $k' = k \cdot \mathsf{polylog}N'$ and noting that $N' = \mathsf{poly}(t)$.[22] □

16 From No-Signaling MIP's to One Round Arguments

In this section we show how to transform any MIP that has soundness against no-signaling strategies into a 1-round argument system, using a fully-homomorphic encryption scheme (FHE) (or alternatively, a computational private information retrieval (PIR) scheme).

Theorem 12. *Suppose that the language $\mathcal{L}$ has an ℓ prover MIP that has ϵ soundness against δ-no-signaling strategies. Let D be the query alphabet and Σ be the answer alphabet of the MIP. Let $\tau = \tau(n) \geq \max(\ell, \log(|\Sigma|), \log(|D|))$ be a security parameter, where n denotes the input length of the MIP. For every $S = S(\tau) \geq \tau$ such that $S \geq \max(n, 2^{\ell\log(|\Sigma|)})$ and $\delta' = \delta'(\tau)$ such that $\delta' \leq \delta/\ell$, if there exists an (S, δ')-secure FHE, then the language $\mathcal{L}$ has a 1-round argument system with soundness (S, ϵ).*

If the MIP verifier runs in time T_V, then the running time of the resulting verifier is $T_V + T_{\mathsf{FHE}}(\tau)$ where T_{FHE} is a polynomial that depends only on the encryption scheme (and not on the language $\mathcal{L}$). If the running time of each MIP prover is T_P, then the running time of the resulting prover is $\mathsf{poly}(T_P, \tau, n)$. The total communication in the resulting argument-system is of length $\mathsf{poly}(\tau)$.

Proof. Let $(V, P_1, \ldots, P_\ell)$ be an ℓ prover MIP for $\mathcal{L}$ with soundness ϵ against δ-no-signaling strategies. Let D be the query alphabet and Σ be the answer alphabet. Since $(V, P_1, \ldots, P_\ell)$ is a 1-round protocol, it will be convenient for us to think of V as being composed of two algorithms that use the same randomness, V_1 and V_2. The first algorithm, V_1, on input x and the random string r outputs a sequence of ℓ queries $\mathbf{q} \in D^\ell$. The second algorithm, V_2, on input x, the same random string r and answers $\mathbf{a} \in \Sigma^\ell$ outputs a bit that represents whether it believes that $x \in \mathcal{L}$. We assume without loss of generality that the provers algorithms $P_1, \ldots, P_\ell$ are deterministic.

Let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec}, \mathsf{Eval})$ be an FHE and let $\tau = \tau(n)$ be a security parameter. We use the MIP and FHE to construct an argument system (V', P') as follows. The verifier, given as input x, first invokes V_1 on input x and a random string r to obtain a sequence of ℓ queries $\mathbf{q} = (q_1, \ldots, q_\ell) \in D^\ell$. Then, for every $i \in [\ell]$, the verifier invokes $\mathsf{Gen}(1^\tau)$, where $\tau = \tau(|x|)$, to obtain a key-pair $(\mathsf{pk}_i, \mathsf{sk}_i)$. The verifier then runs $\mathsf{Enc}_{\mathsf{pk}_i}(q_i)$ to obtain a ciphertext $\hat{q}_i$, for

[22] Note that we assumed that $k' < N'$. If this is not the case then we can increase N' by adding sufficiently many dummy gates to the circuit $\mathcal{C}'_n$.

every $i \in [\ell]$. We denote $\mathsf{pk} = (\mathsf{pk}_1, \ldots, \mathsf{pk}_\ell)$, and $\hat{\mathbf{q}} \stackrel{\text{def}}{=} (\hat{q}_1, \ldots, \hat{q}_\ell)$. The verifier sends $(\mathsf{pk}, \hat{\mathbf{q}})$ to P'.

The prover P' is given as input x and a message $(\mathsf{pk}, \hat{\mathbf{q}})$ from the verifier. For every $i \in [\ell]$, let $\mathcal{C}_{x,i} : D \to \Sigma$ be a Boolean circuit that on input q computes the function $P_i(x, q)$. For every $i \in [\ell]$, the prover P' computes $\hat{a}_i = \mathsf{Eval}(\mathsf{pk}_i, C_{x,i}, \hat{q}_i)$. The prover sends $\hat{\mathbf{a}} \stackrel{\text{def}}{=} (\hat{a}_1, \ldots, \hat{a}_\ell)$ to the verifier.

The verifier V', given the message $\hat{\mathbf{a}}$ from the prover, computes $a_i = \mathsf{Dec}_{\mathsf{sk}_i}(\hat{a}_i)$, for every $i \in [\ell]$. The verifier outputs the result of $V_2(x, (a_1, \ldots, a_\ell), r)$, where r is the same random string used by V_1 to generate the queries.

We proceed to show that (V', P') is an argument system with soundness (S, ϵ) (see definition in Section 4.8).

Completeness. Let $x \in \mathcal{L}$. By the construction and the correctness of the FHE protocol, for every $i \in [\ell]$ it holds that $a_i = P_i(x, q_i)$, with overwhelming probability. When V_2 is invoked with the answers of the honest provers $P_1, \ldots, P_\ell$, by the (perfect) completeness of the MIP, the verifier V outputs 1. Hence, V' accepts with overwhelming probability.

Soundness. Let $\{P^*_n\}_{n \in \mathbb{N}}$ be a family of circuits of size at most $\mathsf{poly}(S(n))$ such that there exist infinitely many $x \notin \mathcal{L}$ such that

$$\Pr[(P^*_{|x|}, V')(x) = 1] > \epsilon, \tag{15}$$

where $(P^*_{|x|}, V')(x)$ denotes the output of V' after interacting with the prover $P^*_{|x|}$ on common input x (and the probability is over the random coins of V'). We show a contradiction by constructing a δ-no-signaling (cheating) strategy that fools the underlying MIP verifier V into accepting some $x \notin \mathcal{L}$ with probability greater than ϵ.

For every $x \notin \mathcal{L}$, consider an MIP prover strategy $\mathcal{A}^{(x)} \stackrel{\text{def}}{=} \{\mathcal{A}^{(x)}_{\mathbf{q}}\}_{\mathbf{q} \in D^\ell}$, where for every $\mathbf{q} = (q_1, \ldots, q_\ell) \in D^\ell$, the distribution $\mathcal{A}^{(x)}_{\mathbf{q}}$ is sampled as follows. First, for every $i \in [\ell]$ invoke $\mathsf{Gen}(1^\tau)$, where $\tau = \tau(|x|)$, to obtain $(\mathsf{pk}_i, \mathsf{sk}_i)$ and compute $\hat{q}_i \in_R \mathsf{Enc}_{\mathsf{pk}_i}(q_i)$. Then, compute $\hat{\mathbf{a}} = (\hat{a}_1, \ldots, \hat{a}_\ell) \in_R P^*_{|x|}(x, (\mathsf{pk}, \hat{\mathbf{q}}))$, where $\mathsf{pk} = (\mathsf{pk}_1, \ldots, \mathsf{pk}_\ell)$ and $\hat{\mathbf{q}} = (\hat{q}_1, \ldots, \hat{q}_\ell)$, and for every $i \in [\ell]$, set $a_i = \mathsf{Dec}_{\mathsf{sk}_i}(\hat{a}_i)$. Finally, output $\mathbf{a} \stackrel{\text{def}}{=} (a_1, \ldots, a_\ell)$.

By the definition of $\mathcal{A}^{(x)}$ and V' and using Eq. (15), for infinitely many $x \notin \mathcal{L}$, it holds that

$$\Pr_{\substack{r \\ \mathbf{a} \in_R \mathcal{A}^{(x)}_{\mathbf{q}}}} [V_2(x, \mathbf{a}, r) = 1] = \Pr\left[(P^*_{|x|}, V')(x) = 1\right] > \epsilon$$

where $\mathbf{q} = V_1(x, r)$. It remains to be shown that for all sufficiently large $x \notin \mathcal{L}$, the strategy $\mathcal{A}^{(x)}$ is δ-no-signaling.

We need to prove that for all sufficiently large x, every $S \subseteq [\ell]$, and every two sequences of queries $\mathbf{q} = (q_1, \ldots, q_\ell) \in D^\ell$ and $\mathbf{q}' = (q'_1, \ldots, q'_\ell) \in D^\ell$ such that $\mathbf{q}_S = \mathbf{q}'_S$ (i.e., $q_i = q'_i$ for all $i \in S$), the following two distributions are δ-close:

- $\mathbf{a}_S$ where $\mathbf{a} \in_R \mathcal{A}^{(x)}_{\mathbf{q}}$; and
- $\mathbf{a}'_S$ where $\mathbf{a}' \in_R \mathcal{A}^{(x)}_{\mathbf{q}'}$.

Toward this end, assume for a contradiction that for infinitely many x this is not the case. That is, for infinitely many x, there exist corresponding S, $\mathbf{q}$, $\mathbf{q}'$ and a distinguisher $\mathcal{D}_x$ such that

$$\left| \Pr_{\mathbf{a} \in_R \mathcal{A}^{(x)}_{\mathbf{q}}} [\mathcal{D}_x(\mathbf{a}_S) = 1] - \Pr_{\mathbf{a}' \in_R \mathcal{A}^{(x)}_{\mathbf{q}'}} [\mathcal{D}_x(\mathbf{a}'_S) = 1] \right| > \delta. \tag{16}$$

Since $\mathcal{D}_x$ takes as input a string of length at most $\ell \cdot \log(|\Sigma|)$, it can be implemented by a circuit of size at most $\mathsf{poly}(2^{\ell \cdot \log(|\Sigma|)})$. We use $\{\mathcal{D}_x\}_x$ to construct a family of circuits $\{\mathcal{C}_\tau\}_\tau$ that breaks the security of the underlying FHE scheme.

For every x as above and for $\tau = \tau(|x|)$, let $\mathcal{C}_\tau$ be a circuit that takes as input a set of public-keys $\{\mathsf{pk}_i\}_{i \in [\ell] \setminus S}$ (with respect to security parameter τ) and a set of ciphertexts $\{c_i\}_{i \in [\ell] \setminus S}$. We show that the circuit $\mathcal{C}_\tau$ distinguishes between the case that (1) each c_i was sampled from $\mathsf{Enc}_{\mathsf{pk}_i}(q_i)$; and the case that (2) each c_i was sampled from $\mathsf{Enc}_{\mathsf{pk}_i}(q'_i)$. The circuit $\mathcal{C}_\tau$ works as follows:

1. For every $i \in S$, sample $(pk_i, sk_i) \in_R \mathsf{Gen}(1^\tau)$ and $c_i \in_R \mathsf{Enc}_{pk_i}(q_i)$. Set $\mathsf{pk} = (pk_1, \ldots, pk_\ell)$ and $\mathbf{c} = (c_1, \ldots, c_\ell)$ (note that for $i \notin S$, the values pk_i and c_i are given as input to the circuit).
2. Compute $\hat{\mathbf{a}} \stackrel{\text{def}}{=} (\hat{a}_1, \ldots, \hat{a}_\ell) = P^*_{|x|}(x, \mathsf{pk}, \mathbf{c})$, where $P^*_{|x|}(x, \mathsf{pk}, \mathbf{c})$ denotes the output of $P^*_{|x|}$ given as input x and a message $(\mathsf{pk}, \mathbf{c})$. (Note that x is fixed inside the description of $\mathcal{C}_\tau$.)
3. For every $i \in S$, set $a_i = \mathsf{Dec}_{sk_i}(\hat{a}_i)$.
4. Output $\mathcal{D}_x(\mathbf{a}_S)$ where $\mathbf{a}_S = (a_i)_{i \in S}$.

Note that $\mathcal{C}_\tau$ has size $\mathsf{poly}\big(2^{\ell \cdot \log(|\Sigma|)}, \tau, S(\tau), |x|\big) \leq \mathsf{poly}(S(\tau))$.

By Eq. (16), the circuit $\mathcal{C}_\tau$ distinguishes between the two cases with probability δ for infinitely many values of τ. By a standard hybrid argument we obtain a circuit that breaks the semantic security of the encryption scheme with distinguishing gap at least $\delta/\ell \geq \delta'(\tau)$ in contradiction to our assumption. Thus, we obtain that for all sufficiently large $x \notin \mathcal{L}$, the prover strategy $\mathcal{A}^{(x)}$ is δ-no-signaling and the lemma follows. □

17 Delegation for P

Using all the results above, we are ready to prove Theorem 9.

Theorem 9. *Suppose that $\mathcal{L} \in \mathsf{DTIME}(t(n))$, where $t = t(n)$ satisfies $\mathsf{poly}(n) \leq t \leq \mathsf{exp}(n)$. Let $\tau = \tau(n)$ be a security parameter such that $\log(t) \leq \tau \leq \mathsf{poly}(t)$. Let $S = S(\tau) \geq \tau$ such*

that $2^{(\log(t))^c} \leq S \leq 2^{\mathsf{poly}(n)}$ *and* $S \leq 2^{\max(n,\tau)}$*, where* c *is some sufficiently large universal constant. If there exists an* $\left(S, 2^{-\sqrt{\log S}}\right)$*-secure* FHE*, then* $\mathcal{L}$ *has a 1-round argument system with soundness* $\left(S, 2^{-\frac{\sqrt{\log S}}{\mathsf{polylog}(t)}}\right)$*. The verifier runs in time* $n \cdot \mathsf{polylog}(t) + \mathsf{poly}(\tau)$ *and the prover runs in time* $\mathsf{poly}(t)$*. The total communication is of length* $\mathsf{poly}(\tau)$*.*

Proof. Suppose that $\mathcal{L} \in \mathsf{DTIME}(t(n))$, where $t = t(n)$ satisfies $\mathsf{poly}(n) \leq t \leq \mathsf{exp}(n)$. Let $\tau = \tau(n)$ be a security parameter such that $\log(t) \leq \tau \leq \mathsf{poly}(t)$. Let $S = S(\tau)$ such that $2^{(\log(t))^{c''}} \leq S \leq 2^{\mathsf{poly}(n)}$ and $S \leq 2^{\max(n,\tau)}$, where $c'' = 2(c + c')$, the constant c is as in Theorem 4 and c' is some sufficiently large universal constant. Let $\delta \stackrel{\text{def}}{=} 2^{-\sqrt{\log S}}$ and $k \stackrel{\text{def}}{=} \frac{\sqrt{\log S}}{(\log(t))^{c'}}$. Note that by the restriction on S, and our setting of k and δ, it holds that:

1. $(\log(t))^c \leq k \leq \mathsf{poly}(n)$.
2. $S \geq \max\left(n, 2^{k^2 \mathsf{polylog}(t)}\right)$.
3. $\delta \leq 2^{-k\mathsf{polylog}(t)}$.

(where the latter two conditions are obtained by setting c' to be a sufficiently large constant).

By applying Theorem 4 (with respect to the parameter k) to the language $\mathcal{L}$, we obtain an MIP for $\mathcal{L}$ with $k \cdot \mathsf{polylog}(t)$ provers and with soundness error 2^{-k} against $2^{-k \cdot \mathsf{polylog}(t)}$-no-signaling strategies. The verifier of the MIP runs in time $(n + k^2) \cdot \mathsf{polylog}(t)$ and the provers run in time $\mathsf{poly}(t, k)$. Each query and answer is of length $k \cdot \mathsf{polylog}(t)$.

Assume that there exists an (S, δ)-secure FHE. By Theorem 12 (and our setting of k, S and δ), we obtain that $\mathcal{L}$ has a 1-round argument system with soundness $(S, 2^{-k})$. The running time of the verifier is $n \cdot \mathsf{polylog}(t) + \mathsf{poly}(\tau)$ and the running time of the prover is $\mathsf{poly}(t)$. The total communication is of length $\mathsf{poly}(\tau)$.

□

Replacing FHE with PIR. As noted in the introduction, Theorem 12 and Theorem 9 can be based on the assumption that a (sufficiently hard) PIR scheme exists rather than a full-blown FHE. Indeed, instead of encrypting the MIP queries, the verifier can send them encapsulated inside PIR queries. The prover, instead of homomorphically evaluating the MIP prover algorithm on encrypted queries, can pre-compute the answers to every possible query and answer according to a corresponding PIR database. However, one must be careful since in the straightforward implementation, the running time of the prover is exponential in the communication complexity of the underlying MIP. This is a real concern in our protocol since the MIP has poly-logarithmic communication complexity (which translates to a quasi-polynomial running time of the prover). We resolve this issue by noting that the next message function of the prover depends only on a logarithmic number of bits and therefore the PIR database can be constructed in polynomial-time.

References

[ABOR00] William Aiello, Sandeep Bhatt, Rafail Ostrovsky, and S. Raj. Rajagopalan. Fast verification of any remote procedure call: Short witness-indistinguishable one-round proofs for NP. In *ICALP: Annual International Colloquium on Automata, Languages and Programming*, 2000.

[AII06] David Avis, Hiroshi Imai, and Tsuyoshi Ito. On the relationship between convex bodies related to correlation experiments with dichotomic observables. *Journal of Physics A: Mathematical and General, 39(36)*, 39(36):11283, 2006.

[AIK10] Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. From secrecy to soundness: Efficient verification via secure computation. In *ICALP (1)*, pages 152–163, 2010.

[BCCT12a] Nir Bitansky, Ran Canetti, Alessandro Chiesa, and Eran Tromer. From extractable collision resistance to succinct non-interactive arguments of knowledge, and back again. In *ITCS*, pages 326–349, 2012.

[BCCT12b] Nir Bitansky, Ran Canetti, Alessandro Chiesa, and Eran Tromer. Recursive composition and bootstrapping for snarks and proof-carrying data. *IACR Cryptology ePrint Archive*, 2012:95, 2012.

[BLM+05] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. Nonlocal correlations as an information-theoretic resource. *Physical Review A, 71(022101)*, 71(2):022101, 2005.

[CKLR11] Kai-Min Chung, Yael Tauman Kalai, Feng-Hao Liu, and Ran Raz. Memory delegation. In *CRYPTO*, pages 151–168, 2011.

[CKV10] Kai-Min Chung, Yael Tauman Kalai, and Salil P. Vadhan. Improved delegation of computation using fully homomorphic encryption. In *CRYPTO*, pages 483–501, 2010.

[DFH12] Ivan Damgård, Sebastian Faust, and Carmit Hazay. Secure two-party computation with low communication. In *TCC*, pages 54–74, 2012.

[DLN+04] Cynthia Dwork, Michael Langberg, Moni Naor, Kobbi Nissim, and Omer Reingold. Succinct proofs for NP and spooky interactions. Unpublished manuscript, available at http://www.cs.bgu.ac.il/~kobbi/papers/spooky_sub_crypto.pdf, 2004.

[FS86] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In *CRYPTO*, pages 186–194, 1986.

[GGP10] Rosario Gennaro, Craig Gentry, and Bryan Parno. Non-interactive verifiable computing: Outsourcing computation to untrusted workers. In *CRYPTO*, pages 465–482, 2010.

[GGPR12] Rosario Gennaro, Craig Gentry, Bryan Parno, and Mariana Raykova. Quadratic span programs and succinct NIZKs without PCPs. *IACR Cryptology ePrint Archive*, 2012:215, 2012.

[GKR08] Shafi Goldwasser, Yael Tauman Kalai, and Guy N. Rothblum. Delegating computation: interactive proofs for muggles. In *STOC*, pages 113–122, 2008.

[GLR11] Shafi Goldwasser, Huijia Lin, and Aviad Rubinstein. Delegation of computation without rejection problem from designated verifier cs-proofs. *IACR Cryptology ePrint Archive*, 2011:456, 2011.

[Gol99] Oded Goldreich. *Modern cryptography, probabilistic proofs and pseudorandomness*, volume 17 of *Algorithms and Combinatorics*. Springer-Verlag, 1999.

[Gol08] Oded Goldreich. *Computational complexity - a conceptual perspective*. Cambridge University Press, 2008.

[Gro10] Jens Groth. Short pairing-based non-interactive zero-knowledge arguments. In *ASIACRYPT*, pages 321–340, 2010.

[GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In *STOC*, pages 99–108, 2011.

[Hol09] Thomas Holenstein. Parallel repetition: Simplification and the no-signaling case. *Theory of Computing*, 5(1):141–172, 2009.

[IKM09] Tsuyoshi Ito, Hirotada Kobayashi, and Keiji Matsumoto. Oracularization and two-prover one-round interactive proofs against nonlocal strategies. In *IEEE Conference on Computational Complexity*, pages 217–228, 2009.

[Ito10] Tsuyoshi Ito. Polynomial-space approximation of no-signaling provers. In *ICALP (1)*, pages 140–151, 2010.

[IV12] Tsuyoshi Ito and Thomas Vidick. A multi-prover interactive proof for NEXP sound against entangled provers. *CoRR*, abs/1207.0550, 2012.

[Kil92] Joe Kilian. A note on efficient zero-knowledge proofs and arguments (extended abstract). In *STOC*, pages 723–732, 1992.

[KKM+08] Julia Kempe, Hirotada Kobayashi, Keiji Matsumoto, Ben Toner, and Thomas Vidick. Entangled games are hard to approximate. In *FOCS*, pages 447–456, 2008.

[KR09] Yael Tauman Kalai and Ran Raz. Probabilistically checkable arguments. In *CRYPTO*, pages 143–159, 2009.

[KT85] Leonid A. Khalfin and Boris S. Tsirelson. Quantum and quasi-classical analogs of Bell inequalities. In *In Symposium on the Foundations of Modern Physics*, pages 441–460, 1985.

[LFKN92] Carsten Lund, Lance Fortnow, Howard J. Karloff, and Noam Nisan. Algebraic methods for interactive proof systems. *J. ACM*, 39(4):859–868, 1992.

[Lip12] Helger Lipmaa. Progression-free sets and sublinear pairing-based non-interactive zero-knowledge arguments. In *TCC*, pages 169–189, 2012.

[Mic94] Silvio Micali. CS proofs (extended abstracts). In *FOCS*, pages 436–453, 1994.

[Nao03] Moni Naor. On cryptographic assumptions and challenges. In *CRYPTO*, pages 96–109, 2003.

[PR94] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[PRV12] Bryan Parno, Mariana Raykova, and Vinod Vaikuntanathan. How to delegate and verify in public: Verifiable computation from attribute-based encryption. In *TCC*, pages 422–439, 2012.

[Ras85] Peter Rastall. Locality, Bell's theorem, and quantum mechanics. *Foundations of Physics*, 15(9):963–972, 1985.

[Rot09] Guy N. Rothblum. *Delegating computation reliably: paradigms and constructions*. PhD thesis, Massachusetts Institute of Technology, 2009.

[Sha92] Adi Shamir. IP = PSPACE. *J. ACM*, 39(4):869–877, 1992.

[Sud00] Madhu Sudan. Probabilistically checkable proofs - lecture notes, 2000. Available at http://people.csail.mit.edu/madhu/pcp/pcp.ps.

[Ton09] Ben Toner. Monogamy of non-local quantum correlations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 465(2101):59–69, 2009.

[Vid13] Thomas Vidick. Three-player entangled xor games are np-hard to approximate. In *FOCS*, 2013.

A Computing LDE over Characteristic 2 Fields

Recall that $\mathbb{G}$ is a finite field of characteristic 2, $H_{\mathbb{G}} \subseteq \mathbb{G}$ is an arbitrary subset of $\mathbb{G}$ and $m_{\mathbb{G}}$ is the dimension.

Proposition A.1. *There exists a Turing Machine that runs in time* $\mathsf{poly}(|\mathbb{G}|^{m_{\mathbb{G}}})$ *and space* $O\big(m_{\mathbb{G}} \cdot \log(|\mathbb{G}|) + \mathsf{polylog}(|\mathbb{G}|)\big)$ *and outputs a Boolean circuit of depth* $O\big(m_{\mathbb{G}} \cdot \log(|\mathbb{G}|) + \log m_{\mathbb{G}} \cdot \mathsf{polylog}(|\mathbb{G}|)\big)$ *and size* $\mathsf{poly}(|\mathbb{G}|^{m_{\mathbb{G}}})$ *that on input a truth table of a function* $\alpha : H_{\mathbb{G}}^{m_{\mathbb{G}}} \to \{0,1\}$ *outputs the truth table of the* LDE $\hat{\alpha} : \mathbb{G}^{m_{\mathbb{G}}} \to \mathbb{G}$ *of* α.

Proof. By the proof of Proposition 4.1,

$$\hat{\alpha}(z) = \sum_{x \in H_{\mathbb{G}}^{m_{\mathbb{G}}}} \hat{\beta}_x(z) \cdot \alpha(x) \tag{17}$$

where each $\hat{\beta}_x$ can be computed by an arithmetic circuit (over $\mathbb{G}$) of depth $O(\log(m_{\mathbb{G}}) + \log(|H_{\mathbb{G}}|))$ and size $\mathsf{poly}(|H_{\mathbb{G}}|, m_{\mathbb{G}})$ and each arithmetic circuit can be generated (by a Turing Machine) in time $\mathsf{poly}(|H_{\mathbb{G}}|, m_{\mathbb{G}}, \log|\mathbb{G}|)$ and in space $O(\log(|\mathbb{G}|) + \log(m_{\mathbb{G}}))$.

Since the field operations can be implemented by Boolean circuits of depth $\mathsf{polylog}(|\mathbb{G}|)$ and size $\mathsf{polylog}(|\mathbb{G}|)$, we can replace each arithmetic circuit by a Boolean circuit of depth $\mathsf{polylog}(|\mathbb{G}|) \cdot \log(m_{\mathbb{G}})$ and size $\mathsf{poly}(|H_{\mathbb{G}}|, m_{\mathbb{G}}, \log(|\mathbb{G}|))$. Each Boolean circuit can be generated in time $\mathsf{poly}(|H_{\mathbb{G}}|, m_{\mathbb{G}}, \log(|\mathbb{G}|))$ and in space $O(\mathsf{polylog}(|\mathbb{G}|) + \log(m_{\mathbb{G}}))$.

The sum of the terms in Eq. 17 can be computed by an arithmetic circuit of depth $O(\log(|H_{\mathbb{G}}|^{m_{\mathbb{G}}}))$ and size $O(|H_{\mathbb{G}}|^{m_{\mathbb{G}}})$. Moreover, since *addition* over $\mathbb{G}$ can be computed by a *constant* depth (fan-in 2) Boolean circuit (because $\mathbb{G}$ has characteristic 2), the sum can be computed by a Boolean circuit of depth $O(\log(|H_{\mathbb{G}}|^{m_{\mathbb{G}}}))$ and size $\mathsf{polylog}(|\mathbb{G}|) \cdot O(|H_{\mathbb{G}}|^{m_{\mathbb{G}}})$. The latter Boolean circuit can be generated in time $\mathsf{polylog}(|\mathbb{G}|) \cdot O(|H_{\mathbb{G}}|^{m_{\mathbb{G}}})$ and space $O(\log(|H_{\mathbb{G}}|^{m_{\mathbb{G}}}))$. □