Arguments of Proximity
[Extended Abstract]

Yael Tauman Kalai[1(✉)] and Ron D. Rothblum[2]

[1] Microsoft Research, Cambridge, USA
yael@microsoft.com
[2] Weizmann Institute of Science, Rehovot, Israel
ron.rothblum@weizmann.ac.il

Abstract. An interactive proof of proximity (IPP) is an interactive protocol in which a prover tries to convince a *sublinear-time* verifier that $x \in \mathcal{L}$. Since the verifier runs in sublinear-time, following the property testing literature, the verifier is only required to reject inputs that are *far* from $\mathcal{L}$. In a recent work, Rothblum *et. al* (STOC, 2013) constructed an IPP for every language computable by a low depth circuit.

In this work, we study the computational analogue, where soundness is required to hold only against a *computationally bounded* cheating prover. We call such protocols *interactive arguments of proximity*.

Assuming the existence of a sub-exponentially secure FHE scheme, we construct a *one-round* argument of proximity for *every language* computable in time t, where the running time of the verifier is $o(n)+\mathsf{polylog}(t)$ and the running time of the prover is $\mathsf{poly}(t)$.

As our second result, assuming sufficiently hard cryptographic PRGs, we give a lower bound, showing that the parameters obtained both in the IPPs of Rothblum *et al.*, and in our arguments of proximity, are close to optimal.

Finally, we observe that any one-round argument of proximity immediately yields a one-round delegation scheme (without proximity) where the verifier runs in *linear* time.

1 Introduction

With the prominent use of computers, tremendous amounts of data are available. For example, hospitals have massive amounts of medical data. This data is very precious as it can be used, for example, to learn important statistics about various diseases. This data is often too large to store locally, and thus is often stored on cloud platforms (or external servers). As a result, if a hospital (which has bounded storage and bounded computational power), wishes to perform some computation on its medical data, it would need to delegate this computation to the cloud. Since the cloud's computation may be faulty, the party delegating the computation (say, the hospital), may want a proof that the computation was done correctly. It is important that this proof can be verified very efficiently,

R. Gennaro and M. Robshaw (Eds.): CRYPTO 2015, Part II, LNCS 9216, pp. 422–442, 2015.
DOI: 10.1007/978-3-662-48000-7_21

and that the prover's running time is not much larger than the time it takes to perform the computation, since otherwise, the solution will not be practical.

This problem is closely related to the problem of computation delegation, where a weak client delegates a computation to a powerful server, and the server needs to provide the client with a proof that the computation was done correctly. In contrast to the current setting, in the setting of computation delegation, the input is thought of as being small and the computation is thought of as being large. The client (verifier) is required to run in time that is proportional to the input size (but much smaller than the time it takes to do the computation), and the powerful server (prover) runs in time polynomially related to the time it takes to do the computation. Indeed the problem of computation delegation is extremely important, and received a lot of attention (e.g., [GKR08, Mic94, Gro10, GGP10, CKV10, AIK10, GLR11, Lip12, BCCT12a, DFH12, BCCT12b, GGPR12, PRV12, KRR13a, KRR13b]).

In reality, however, the input (data) is often very large, and the client cannot even store the data. Hence, we seek a solution in which the client runs in time that is *sub-linear* in the input size. The question is:

If the client cannot read the data, how can he verify the correctness of a computation on the data?

The work of [CKLR11], on memory delegation, considers this setting where the input (thought of as the client's memory) is large, and the client cannot store it locally. However, in memory delegation, it is assumed that the client (verifier) stores a short "commitment" of the input, and then can verify computations in sub-linear time. However, computing such a commitment takes time at least linear in the input length, which is infeasible in many settings.

Recently, Rothblum, Vadhan and Wigderson [RVW13], in their work on interactive proofs of proximity (IPP, a notion first studied by Ergün, Kumar and Rubinfeld [EKR04]), provide a solution where the verifier does not need to know such a commitment. Without such a commitment, the verifier cannot be sure that the computation is correct (since he cannot read the entire input), however they guarantee that the input is "close" to being correct. More specifically, they construct an interactive proof system for every language computable by a (log-space uniform) low depth circuit, where the verifier is given *oracle access* to the input (the data), and the verifier can check whether the input is *close* to being in the language in *sub-linear* time in the input (and linear time in the depth of the computation). We note that in many settings where the data is large (such as medical data) and the goal is to compute some statistics on this data, an approximate solution is acceptable. The work of [RVW13] is the starting point of our work.

1.1 Our Results in a Nutshell

We depart from the interactive proof of proximity setting, and consider *arguments of proximity*. In contrast to proofs of proximity, in an argument of proximity, soundness is required to hold only against *computationally bounded* cheating provers. Namely, the soundness guarantee is that any bounded cheating prover

can convince the verifier to accept an input that is far from the language (in Hamming distance) only with small probability. By relaxing the power of the prover we obtain stronger results.

We construct *one-round* arguments of proximity for every deterministic language (without a dependency on the depth). Namely, fix any $t = t(n)$ and any language $\mathcal{L} \in \mathsf{DTIME}(t(n))$, we construct a one-round argument of proximity for $\mathcal{L}$ where the verifier runs in time $o(n) + \mathsf{polylog}(t)$, assuming the existence of a sub-exponentially secure fully homomorphic encryption (FHE) scheme.

Our one-round argument of proximity is constructed in two steps, and follows the outline of the recent works of Kalai *et al.* [KRR13a,KRR13b]. These works first show how to construct an MIP for all deterministic languages, that is sound against *no-signaling strategies*. Such no-signaling soundness is stronger than the typical notion of soundness, and is inspired by quantum physics and by the principal that information cannot travel faster than light (see Sect. 3.2 for the definition, and [KRR13a,KRR13b] for more background on this notion). They then show how to convert these no-signaling MIPs into one-round arguments.

As our first step, we combine the IPPs of [RVW13], and the no-signaling MIP construction of [KRR13b], to obtain a no-signaling *multi-prover interactive proof of proximity* (MIPP). This construction combines techniques and results of [RVW13] and [KRR13b], and may be of independent interest.

Then, similarly to [KRR13a], we show how to convert any no-signaling MIPP to a one-round argument of proximity. This transformation relies on a heuristic developed by Aiello *et al.* [ABOR00], which uses a (computational) PIR scheme (or a fully homomorphic encryption scheme) to convert any MIP into a one-round argument. This heuristic was proven to be secure in [KRR13a] if the underlying MIP is secure against no-signaling strategies. We extend the result of [KRR13a] to the proximity setting.

Finally, we provide a negative result, which shows that the parameters we obtain for MIPP and the parameters obtained in [RVW13], are somewhat tight. Proving such a lower bound was left as an open problem in [RVW13]. This part contains several new ideas, and is the main technical contribution of this work.

We also show that the parameters in our one-round argument of proximity are somewhat optimal, for arguments with adaptive soundness and are proven to be (adaptively) sound via a black-box reduction to a falsifiable assumption. See the full version for further details.

Linear-Time Delegation. We observe that both proofs and arguments of proximity, aside from being natural notions, can also be used as tools to obtain new results for delegating computation in the standard setting (i.e., where soundness is guaranteed for *every* $x \notin \mathcal{L}$). More specifically, using our results on arguments of proximity and the [RVW13] results on interactive proofs of proximity for low-depth circuits, we can construct (standard) one-round argument-systems for any deterministic computation, and interactive proof systems for low-depth circuits,

where the verifier truly runs in *linear-time*. In contrast, the results of [GKR08] and [KRR13b] only give a *quasi-linear* time verifier.[1]

1.2 Our Results in More Detail

Our main result is a construction of a one-round argument of proximity for any deterministic language. Here, and throughout this work, we use n to denote the input length. Let $t = t(n)$, let $\mathcal{L} \in \mathsf{DTIME}(t)$ be a language. For a proximity parameter $\varepsilon = \varepsilon(n) \in (0,1)$, we denote by ε-IPP an interactive proof for testing ε-proximity to $\mathcal{L}$.[2] Similarly we denote by ε-MIPP a multi-prover interactive proof for testing ε-proximity to $\mathcal{L}$.

Theorem 1 (Informal). *Suppose that there exists a sub-exponentially secure* FHE. *Fix a proximity parameter* $\varepsilon \stackrel{\text{def}}{=} n^{-(1-\beta)}$, *for some sufficiently small* $\beta > 0$, *and a security parameter* τ *(polynomially related to* n*).*

There exists a 1-round argument of ε*-proximity for* $\mathcal{L}$, *where the verifier runs in time* $n^{1-\gamma} + \mathsf{polylog}(t) + \mathsf{poly}_{\mathsf{FHE}}(\tau)$, *where* $\gamma > 0$ *is a constant and* $\mathsf{poly}_{\mathsf{FHE}}$ *is a polynomial that depends only on the* FHE *scheme, and makes* $n^{1-\gamma} + \mathsf{polylog}(t)$ *oracle queries to the main input. The prover runs in time* $\mathsf{poly}(t)$. *The total communication is of length* $\mathsf{poly}_{\mathsf{FHE}}(\tau)$.

Note that for languages in $\mathsf{DTIME}\left(2^{n^{\alpha}}\right)$ for sufficiently small $\alpha > 0$ (and in particular for languages in P), the verifier in Theorem 1 runs in *sub-linear* time.

As mentioned previously, this result is obtained in two steps. We first construct an MIPP that is sound against no-signaling strategies, and then show how to convert any such MIPP into a one-round argument of proximity.

Theorem 2 (Informal). *Fix a proximity parameter* $\varepsilon = \varepsilon(n) \in (0,1)$. *There exists an* ε-MIPP *that is secure against no-signaling strategies, where the verifier makes* $q = (1/\varepsilon)^{1+o(1)}$ *oracle queries to the input, the communication complexity* $c = (\varepsilon n)^2 \cdot n^{o(1)} \cdot \mathsf{polylog}(t)$ *and the running time of the verifier is* $(\varepsilon n)^2 \cdot \mathsf{polylog}(t) + \left(\frac{1}{\varepsilon} + \varepsilon n\right)^{1+o(1)}$.

We then show how to convert any no-signaling ε-MIPP to a one-round argument of ε-proximity. In the following we say that a fully homomorphic encryption scheme (FHE) is (T, δ) secure if every family of circuits of size T can break the semantic security of the FHE with probability at most δ.

Theorem 3 (Informal). *Fix a proximity parameter* $\varepsilon = \varepsilon(n) \in (0,1)$. *Suppose that the language* $\mathcal{L}$ *has an* ℓ*-prover* ε-MIPP *that is sound against* δ*-no-signaling strategies, with communication complexity* c. *Suppose that there exists a* $(T, \delta/\ell)$*-secure* FHE, *where* $T \geq 2^c$. *Then* $\mathcal{L}$ *has a 1-round argument of* ε*-proximity where*

[1] Actually, by an observation of Vu *et al.* [VSBW13] (see also [Tha13, Lemma 3]), the verifier in the [GKR08] protocol can be directly implemented in linear-time. However the latter implementation would only guarantee *constant* soundness error.

[2] A string $x \in \{0,1\}^n$ is ε-close to $\mathcal{L}$ if there exists $x' \in \{0,1\}^n \cap \mathcal{L}$ such that $\triangle(x, x') \leq \varepsilon n$, where $\triangle$ denotes the Hamming distance between the two strings.

the running time of the prover and verifier and the communication complexity of the argument system, are proportional to those of the underlying MIPP *scheme.*

We note that the parameters in Theorem 2 are somewhat similar to the parameters of the interactive proof of proximity (IPP) in [RVW13]. In particular, in both constructions it holds that $c \cdot q = \Omega(n)$. The work of [RVW13] shows that this lower bound of $c \cdot q = \Omega(n)$ is inherent for IPPs with 2-messages (and that a weaker bound holds for IPPs with a constant number of rounds), and left open the question of whether this lower bound is inherent for general (multi-round) IPPs.

We resolve this question by showing that for every ε-IPP, and every ε-MIPP that is sound against no-signaling strategies, it must be the case that $c \cdot q = \Omega(n)$. For this result we assume the existence of exponentially hard pseudorandom generators.

Theorem 4 (Informal). *Assume the existence of exponentially hard pseudorandom generators. There exists a constant $\varepsilon > 0$ such that for every $q = q(n) \leq n$, there exists a language $\mathcal{L} \in \mathsf{P}$ such that for every ε-IPP for $\mathcal{L}$, and for every ε-MIPP for $\mathcal{L}$ that sound against no-signaling adversaries, it holds that $q \cdot c = \Omega(n)$, where q is the query complexity and c is the communication complexity.*

In fact, assuming a slightly stronger cryptographic assumption, we can replace $\mathcal{L} \in \mathsf{P}$ with $\mathcal{L} \in \mathsf{NC}_1$ (which shows that the [RVW13] upper bound for log-space uniform NC is essentially tight). See Sect. 4 for details.

We note that the [RVW13] lower bound for 2-message IPPs is unconditional (and in particular they do not assume that the verifier is *computationally* bounded). It remains an interesting open problem to obtain an *unconditional* lower bound for multi-message IPPs.

The parameters we obtain for the one-round argument also satisfy $q \cdot c = \Omega(n)$. We show that these parameters are close to optimal for arguments with adaptive soundness, that are proven sound via a black-box reduction to falsifiable assumptions. We refer the reader to the full version for details.

Finally, using the [RVW13] protocol or the protocol of Theorem 1 we construct delegation schemes in which the verifier runs in *linear-time*.

Theorem 5 (Informal). *For every language in (logspace-uniform)* NC *there exists an interactive proof system in which the verifier runs in time $O(n)$ and the prover runs in time* poly(n).

Theorem 6 (Informal). *Assume that there exists a sub-exponentially secure* FHE. *Then, for every language in* P *there exists a 1-round argument-system in which the verifier runs in time $O(n)$ and the prover runs in time* poly(n).

1.3 Related Work

As mentioned above, the work of [RVW13] and [KRR13a,KRR13b] are most related to ours. Both our work, and the work of [RVW13], lie in the intersection of property-testing and computation delegation. As opposed to property

testing, where an algorithm is required to decide whether an input is close to the language *on its own* in sub-linear time, in our work the algorithm receives a proof, and only needs to verify correctness of the proof in sub-linear time. Thus, our task is significantly easier than the task in property testing. Indeed we get much stronger results. In particular, the works on property testing typically get sub-linear algorithms for specific languages, whereas our result holds *for all deterministic languages*.[3]

Another very related problem is that of constructing a *probabilistically checkable proof of proximity* (PCPP) [BSGH+06] (also known as *assignment testers* [DR06]). A PCPP consists of a prover who publishes a long proof, and a verifier, who gets oracle access to this proof and to the instance x, and needs to decide whether x is close to the language in sub-linear time. The significant difference between PCPP and proofs/argument of proximity is that in the PCPP setting the proof is a fixed string (and cannot be modified adaptively based on the verifier's messages).

The fundamental works of Kilian and Micali [Kil92, Mic94] show how to convert any probabilistically checkable proof (PCP) into a 2-round (4-message) argument. As pointed out by [RVW13], their transformation can be also used to convert any PCPP into a 2-round argument of proximity. Thus, obtaining a 2-round argument of proximity follows immediately by applying the transformation of [Kil92, Mic94] to any PCPP construction. Moreover, the parameters of the resulting 2-round argument are optimal (up to logarithmic factors); i.e., the query complexity, the communication complexity and the runtime of the verifier is $\mathsf{poly}(\log(t), \tau)$ where t is the time it takes to compute if x is in the language, and where τ is the security parameter.

The focus of this work is on constructing *one-round* arguments of proximity. Unfortunately, our parameters do not match those of the two-round arguments of proximity outlined above. However, we show that using our techniques (i.e., of constructing one-round arguments of proximity from no-signaling MIPPs), our parameters are almost optimal.

Other works that are related to ours are the work of Gur and Rothblum [GR13] on non-interactive proofs of proximity, and of Fischer *et al.* [FGL14] on partial testing. The former studies an NP version of property testing (which can be thought of as a 1-message variant of IPP), whereas the latter studies a model of property testing in which the tester needs to only accept a sub-property (we note that the two notions, which were developed independently, are tightly related, see [GR13, FGL14] for details).

Organization. In this extended abstract we give an overview of our techniques and only prove some of our results. In Sect. 2 we give a high level view of our techniques. In Sect. 3 we formally define arguments of proximity and the other central definitions that are used throughout this work. In Sect. 4 we show our

[3] Indeed, as shown by Goldwasser, Goldreich and Ron [GGR98], there are properties in very low complexity classes that require $\Omega(n)$ queries and running-time in order to test (without the help of a prover).

lower bound for no-signaling MIPPs. See the full version for the missing proofs and formal theorem statements.

2 Our Techniques

2.1 Our Positive Results

To construct arguments of proximity for languages in $\mathsf{DTIME}(t)$, we adapt the technique of [KRR13a] to the "proximity" setting. That is, we first construct an MIPP that has soundness against no-signaling strategies and then employ the technique of Aiello *et al.* [ABOR00] to obtain an argument of proximity. We elaborate on these two steps below. In what follows, we focus for simplicity on languages in P, though everything extends to languages in $\mathsf{DTIME}(t)$.

No-Signaling MIPPs for P. Our first step (which is technically more involved) is a construction of MIPPs that are sound against no-signaling strategies for any language $\mathcal{L} \in \mathsf{P}$. This construction is inspired by (and reminiscent of) the IPP construction of [RVW13]. The starting point for the [RVW13] IPP is the "Muggles" protocol of Goldwasser *et al.* [GKR08], whereas our starting point is the no-signaling MIP of [KRR13b].

The main technical difficulty in using both the [GKR08] and [KRR13b] protocols by a sublinear time verifier is that in both protocols, the verifier needs to compute an error corrected encoding of the input x. More specifically, the verifier needs to compute the low degree extension of x, denoted LDE_x. Since error-correcting codes are very sensitive to changes in the input, a sub-linear algorithm has no hope to compute LDE_x.

The key point is that in both the [GKR08] and the [KRR13b] protocols, it suffices for the verifier to check the value of LDE_x at relatively few *randomly* selected points (this property was also used by [CKLR11] in their work on memory delegation). Hence, it will be useful for us to view both the [GKR08] and [KRR13b] protocols as protocols for producing a sequence of points J in the low degree extension of x and a sequence of corresponding values $\boldsymbol{v}$ with the following properties:

- If $x \in \mathcal{L}$ and the prover(s) honestly follow the protocol then $\mathsf{LDE}_x(J) = \boldsymbol{v}$.
- If $x \notin \mathcal{L}$ then no matter what the cheating prover does (resp., no-signaling cheating prover do), with high probability the verifier outputs $J, \boldsymbol{v}$ such that $\mathsf{LDE}_x(J) \neq \boldsymbol{v}$.

Hence, the verifiers in both protocols first run this subroutine to produce J and $\boldsymbol{v}$ and then accept if and only if $\mathsf{LDE}_x(J) = \boldsymbol{v}$. Remarkably, in both cases, in the protocol that produces J and $\boldsymbol{v}$, the verifier does not need to access x.

The next step in [RVW13] is a parallel repetition of the foregoing protocol in order to reduce the soundness error. Once the soundness error is sufficiently small, [RVW13] argue that for every x that is ε-far from $\mathcal{L}$, no matter what the cheating prover does (in the parallel repetition of the base protocol), the verifier will output $J, \boldsymbol{v}$ such that not only $\mathsf{LDE}_x(J) \neq \boldsymbol{v}$, but furthermore, x is far from

any x' such that $\mathsf{LDE}_{x'}(J) = \boldsymbol{v}$. This steps simply follows by taking a union bound over all x' that are close to x.

We borrow this step almost as-is from [RVW13] except for the following technical difficulty - it is not known whether parallel repetition decreases the soundness error of no-signaling MIP protocols.[4] However, we observe that the [KRR13b] protocol already allows for sufficient flexibility in choosing its soundness error so that the parallel repetition step can be avoided.

The last step of [RVW13] is designing an IPP protocol for a language that they call $\mathsf{PVAL}_{J,\boldsymbol{v}}$ (for "polynomial evaluation"). This language, parameterized by J and $\boldsymbol{v}$, consists of all strings x such that $\mathsf{LDE}_x(J) = \boldsymbol{v}$. Using this IPP for PVAL, the IPP verifier for a language $\mathcal{L}$ first runs the (parallel repetition of the) [GKR08] protocol, to produce $J, \boldsymbol{v}$ as above. Then, the IPP verifier runs the $\mathsf{PVAL}_{J,\boldsymbol{v}}$ protocol and accepts if and only if the PVAL-verifier accepts. If $x \in \mathcal{L}$ then we know that $\mathsf{LDE}_x(J) = \boldsymbol{v}$ and therefore the PVAL-verifier will accept, whereas if x is far from $\mathcal{L}$ then x is far from $\mathsf{PVAL}_{J,\boldsymbol{v}}$ and therefore the PVAL-verifier will reject. Hence the (parallel repetition of the) [GKR08] protocol is sequentially composed with the IPP for PVAL.

For the no-signaling case, we also use the [RVW13] IPP protocol for PVAL. A technical difficulty that arises is that in contrast to the IPP setting in which sequential composition (of two interactive proofs) is trivial, here we need to compose a 1-round no-signaling MIP with an IPP protocol, to produce a no-signalling MIPP. We indeed prove that such a composition holds thereby constructing a no-signaling MIPP as we desire.

From No-Signaling MIPP to Arguments of Proximity. The transformation from a no-signaling MIPP to an argument of proximity is based on the assumption that there exists a fully homomorphic encryption scheme (or alternatively, a computational private information retrieval scheme) and is practically identical to that in [KRR13a]. More specifically, the argument's verifier uses the MIPP verifier to generate a sequence of queries $q_1, \ldots, q_\ell$ to the ℓ provers. It encrypts each query using a fresh encryption key as follows: $\hat{q}_i \leftarrow Enc_{k_i}(q_i)$. The argument's verifier sends all the encrypted queries to the prover. Given $\hat{q}_1, \ldots, \hat{q}_\ell$, the prover uses the homomorphic evaluation algorithm to compute the MIPP answers "underneath" the encryption. It sends these answers back to the verifier, which can decrypt the encrypted answers and decide. As in [KRR13a] we show that if the MIPP is sound against no-signaling strategies then, assuming the semantic security of the FHE, the resulting protocol is sound against computationally bounded adversaries.

Linear-Time Delegation. We show that using the foregoing one-round argument of proximity for every language $\mathcal{L} \in \mathsf{P}$ and good error-correcting codes, one can easily construct a one-round delegation protocol where the verifier runs in *linear* time (in contrast, the verifier in [KRR13b] runs in *quasi*-linear time). A similar observation, in the context of PCPs, was previously pointed out by [EKR04].

[4] Holenstein [Hol09] showed a parallel repetition theorem for no-signaling 2-prover MIPs. It is not known whether this result can be extended to 3 or more provers.

Let $\mathcal{L} \in \mathsf{P}$ and consider $\mathcal{L}' = \{\mathsf{ECC}(x) : x \in \mathcal{L}\}$ where ECC is an error correcting code with constant rate, constant relative distance, linear-time encoding and polynomial-time decoding[5]. Then, $\mathcal{L}' \in \mathsf{P}$ and so it has an argument of proximity with a sublinear-time verifier. We construct a delegation scheme for $\mathcal{L}$ by having both the verifier and the prover compute $x' = \mathsf{ECC}(x)$ and run the argument of proximity protocol with respect to x'. Since the argument of proximity verifier runs in sublinear time, and $\mathsf{ECC}(x)$ can be computed in linear-time, the resulting delegation verifier runs in linear-time. Soundness follows from the fact that a cheating prover that convinces the argument-system verifier to accept $x \notin L$ can be used to convince the argument-of-proximity verifier to accept $\mathsf{ECC}(x)$ which is indeed far from $\mathcal{L}'$.

A similar result can be obtained for interactive proofs for low-depth computation based on the results of [RVW13] by using an error-correcting code that can be decoded in logarithmic-depth (such a code was constructed by Spielman [Spi96]).

2.2 Our Negative Results

We prove that assuming the existence of exponentially hard pseudorandom generators, there exists a constant $\varepsilon > 0$ for which there does not exist a no-signaling ε-MIPP for all of P with query complexity q and communication complexity c such that $q \cdot c = o(n)$ (where n is the input length). We also show a similar result for ε-IPP.

We start by focusing on our lower bound for MIPP. The high-level idea is the following: Suppose (towards contradiction) that every language in P has a no-signaling MIPP with query complexity q and communication complexity c where $q \cdot c = o(n)$. The fact that $q = o(n)$ implies that (for every language in P), there is some set of coordinates $S \subseteq [n]$ of size $O(n/q)$ that with high (constant) probability the verifier does not query.

As a first step, suppose for the sake of simplicity that there is a fixed (universal) set of coordinates $S \subseteq [n]$ such that with high probability the verifier never queries the coordinates in S, for every language in P (for example, if the verifier's queries are non-adaptive and are generated before it communicates with the prover, then such a set S must exist). We derive a contradiction by showing that one can use the no-signaling MIPP to construct a no-signaling MIP for languages in $\mathsf{NP}\backslash\mathsf{P}$ with communication $c = o(n)$. The latter was shown to be impossible, assuming that $\mathsf{NP} \not\subseteq \mathsf{DTIME}(2^{o(n)})$ [DLN+04] (see also [Ito10]).

The basic idea is the following: Take any language $\mathcal{L} \in \mathsf{NP}\backslash\mathsf{P}$ that is assumed to be hard to compute in time $2^{o(n)}$, and convert it into the language $\mathcal{L}' \in \mathsf{P}$, defined as follows: $x' \in \mathcal{L}'$ if and only if x'_S is a valid witness of $x'_{[n]\setminus S}$ in the underlying NP language $\mathcal{L}$. The no-signaling MIP for $\mathcal{L}$ will simply be the no-signaling ε-MIPP for $\mathcal{L}'$, where the MIP verifier simulates the ε-MIPP verifier with oracle access to x' where $x'_{[n]\setminus S} = x$, and $x'_S = 0^{|S|}$. Note that the MIP verifier, which takes as input x (supposedly in $\mathcal{L}$), cannot (efficiently) generate a

[5] Such codes are known to exist, see, e.g., [Spi96].

corresponding witness w and set $x'_S = w$. But the point is that it does not need to, since S was chosen so that with high probability the MIPP verifier for $\mathcal{L}'$ will not query x' on coordinates in S.

There are several problems with this approach. First, the witness can be very long compared to x, and the set S may be very small compared to n. In this case we will not be able to fit the entire witness in the coordinate set S. Second, after running the MIPP, the verifier is convinced that x' is close to an instance in $\mathcal{L}'$. However, this does not imply that x is in $\mathcal{L}$ (and can only imply that x is close to $\mathcal{L}$).

One can fix these two problems with a single solution: Instead of setting $x'_{[n]\setminus S} = x$ we set $x'_{[n]\setminus S} = \mathsf{ECC}(x)$, where ECC is a error-correcting code with efficient encoding, that is resilient to 2ε-fraction of errors. Now, we can take $\mathsf{ECC}(x)$ so that $|\mathsf{ECC}(x)|$ is very large compared to $|w|$, so that we can fit all of the witness in the coordinate set S. Moreover, if $|\mathsf{ECC}(x)| > |w|$ then if x' is ε-close to $\mathcal{L}'$ then $x'_{[n]\setminus S}$ is 2ε-close to $\mathcal{L}$. This, together with the fact that $\mathsf{ECC}(x)$ is resilient to 2ε-fraction of errors implies that the encoded element is indeed in $\mathcal{L}$.

The foregoing idea indeed seems to work if there was a fixed (universal) set S that the MIPP verifier does not query (with high probability). However, this is not necessarily the case, and this set S may be different for different languages in P. In particular, we cannot claim that for the language $\mathcal{L}'$ the set S is exactly where the witness lies. Namely, it may be that the verifier in the underlying MIPP always queries some coordinates in S.

We solve this problem by using repetitions. Namely, every element $x' \in \mathcal{L}'$ will consist of many instances (encoded using an error-correcting code) along with many witnesses; i.e., $x' = (\mathsf{ECC}(x_1, \ldots, x_m), w_1, \ldots, w_m)$, where each w_j is a witness for the NP statement $x_j \in \mathcal{L}$. Now, suppose that the verifier makes q queries to x' (where $q = o(n)$). Then if we take $m = 4q$ then we know that $3/4$ of the (x_j, w_j)'s are not queried.

As above, we derive a contradiction by showing that one can use the no-signaling MIPP to construct a no-signaling MIP for languages in $\mathsf{NP} \setminus \mathsf{P}$ with $o(n)$ communication, (which is known to be impossible for languages that cannot be computed in time $2^{o(n)}$ [DLN+04,Ito10]). However, now the no-signaling MIP construction will be different: Given an instance x (supposedly in $\mathcal{L}$), the MIP verifier will choose a random $i^* \in_R [m]$, along with m random instance and witness pairs $(x_1, w_1), \ldots, (x_m, w_m)$, where $x_{i^*} = x$ and w_{i^*} can be arbitrary (assumed not to be queried).

We need to argue that with probability at least $3/4$ the verifier will not query the coordinates of w_{i^*}, and thus with probability at least $3/4$ the MIP verifier will successfully simulate the MIPP verifier. If the queries of the MIPP verifier were chosen before interacting with the prover then this would follow immediately from the fact that $i^* \in [m]$ is chosen at random. However, the MIPP verifier may choose its oracle queries after interacting with the MIPP provers, and therefore we need to argue that the MIPP provers also do not know i^*. Note that the MIPP provers see all of $x_1, \ldots, x_m$. Hence, in order to claim that the provers

cannot guess i^* it needs to be the case that x is distributed identically to the other $x_1, \ldots, x_m$.

Hence, we seek a language $\mathcal{L} \in \mathsf{NP}\backslash\mathsf{P}$ for which there exists a distribution $\mathcal{D}$ (distributed over $\mathcal{L}$) such that:

1. It is computationally hard to distinguish between $x \in_R \mathcal{D}$ and $x \notin \mathcal{L}$ (i.e., $\mathcal{L}$ is hard on the average); and
2. $x \in_R \mathcal{D}$ can be sampled together with a corresponding NP witness.

We note that the first requirement is needed to obtain a contradiction (and replaces the weaker assumption that $\mathcal{L} \in \mathsf{NP}\backslash\mathsf{P}$) whereas the second assumption is required so that we can sample $x_1, \ldots, x_m$ (together with the corresponding witnesses) so that MIPP protocol cannot distinguish between x and any of the x_j's (thereby hiding i^*). In can be easily verified that both requirement are met by considering $\mathcal{D}$ which is the output of a cryptographic pseudorandom generator (PRG). Hence the language $\mathcal{L}$ that we use is precisely the output of such a PRG.

Indeed, we can only argue that our no-signaling MIP has *average-case* completeness (with respect to the distribution $\mathcal{D}$), since if $x \in \mathcal{L}$ is distributed differently from $(x_1, \ldots, x_m)$ then the verifier of the MIPP may always query the coordinates where the witness of x is embedded, in which case the MIP verifier will fail to simulate. However, for random $x \in_R \mathcal{L}$ the provers (and verifier) in the MIPP cannot guess i^* with any non-negligible advantage, and therefore the verifier will not query the coordinates of w_{i^*} with probability at least $3/4$, in which case the MIP verifier will succeed in simulating the underlying ε-MIPP verifier. We refer the reader to Sect. 4 for further details.

A Lower Bound for IPP. To obtain a multiplicative lower bound for IPP, we follow the same paradigm outlined above for MIPP's with no-signaling soundness. More specifically, we consider a language $\mathcal{L} \in \mathsf{NP}$ and the corresponding language

$$\mathcal{L}' = \left\{ \left(\mathsf{ECC}(x_1, \ldots, x_m), w_1, \ldots, w_m\right) \,:\, w_j \text{ is an } \mathsf{NP}\text{-witness for } x_j \right\}$$

as above. We show that an IPP protocol for $\mathcal{L}'$ implies a (standard) interactive-proof for $\mathcal{L}$ with similar communication complexity. Here we obtain a contradiction by arguing that (assuming exponential hardness) there are languages in $\mathsf{NP}\backslash\mathsf{P}$ for which every interactive proof require $\Omega(n)$ communication. The latter is based on the proof that $\mathsf{IP} \subseteq \mathsf{PSPACE}$ (i.e., the "easy" direction in the $\mathsf{IP} = \mathsf{PSPACE}$ theorem).

Given the [RVW13] positive result of IPP for low depth computations, we would like to show that our lower bound is not just for languages in P but even for languages, say, in NC_1 (thereby showing that the [RVW13] result is tight). To do so we observe that if (1) the error correcting code that we use has an encoding procedure that can be computed by an NC_1 circuit and (2) the cryptographic PRG can be computed in NC_1, then indeed $\mathcal{L}' \in \mathsf{NC}_1$.

A Lower Bound for One-Round Arguments of Proximity. For one-round arguments of proximity, we show a similar lower-bound of $q \cdot c = \Omega(n)$, assuming

the argument has *adaptive* soundness, and the proof of (adaptive) soundness is via a *black-box reduction* to some *falsifiable* cryptographic assumption.

Loosely speaking, a cryptographic assumption is falsifiable (a notion due to Naor [Nao03]) if there is an *efficient* way to "falsify it", i.e., to demonstrate that it is false. We note that most standard cryptographic assumptions (e.g., one-way functions, public-key encryption, LWE etc.) are falsifiable. A black-box reduction of one cryptographic primitive to another, is a reduction that, using black-box access to any (possibly inefficient) adversary for the first primitive, breaks the security of the second primitive.

Similarly to the MIPP and IPP lower bounds, we consider the languages $\mathcal{L}$ and $\mathcal{L}'$, as above, where $\mathcal{L} \in \mathsf{NP}$ is exponentially hard on average and $\mathcal{L} \in \mathsf{P}$. We prove that if there exists an adaptively sound one-round argument of proximity for $\mathcal{L}'$ with $q \cdot c = o(n)$ then there exists an adaptively sound one-round argument for $\mathcal{L}$ with $o(n)$ communication (in the crs model).

We then rely on a result of Gentry and Wichs [GW11], which shows that there does not exist a one-round argument for exponentially hard (on average) NP languages, with adaptive soundness and black-box reduction to a falsifiable assumption.

We conclude that P does not have an adaptively sound one-round argument of proximity with $q \cdot c = o(n)$, and a black-box reduction to a falsifiable assumption. We refer the reader to the full version for details.

3 Definitions

In this section we define arguments of proximity and MIPs of proximity (with soundness against no-signaling strategies). See the full version for additional standard definitions.

Notation. For $x, y \in \{0,1\}^n$, we denote the Hamming distance of x and y by $\Delta(x,y) \stackrel{\text{def}}{=} |\{i \in [n] \colon x_i \neq y_i\}|$. We say that x is ε-close to y if $\Delta(x,y) \leq \delta$. We say that x is ε-close to a set $S \subseteq \{0,1\}^n$ if there exists $y \in S$ such that x is ε-close to y.

If A is an oracle machine, we denote by $A^x(z)$ the output of A when given oracle access to x and explicit access to z.

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

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.

3.1 Arguments of Proximity

An interactive argument of proximity for a language $\mathcal{L}$ consists of a polynomial-time verifier that wishes to verify that x is close (in Hamming distance) to some

x' such that $x' \in \mathcal{L}$, and a prover that helps the verifier to decide. The verifier is given as input $n \in \mathbb{N}$, a proximity parameter $\varepsilon = \varepsilon(n) > 0$ and oracle access to $x \in \{0,1\}^n$ (and its oracle queries are counted). The prover gets as input ε and x. The two parties 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 is ε-far from $\mathcal{L}$, then no *computationally bounded* prover can convince the verifier to accept with non-negligible (in n) probability.

We focus on 1-round arguments of proximity 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 $\varepsilon = \varepsilon(n) \in (0,1)$ be a proximity parameter. Let $T : \mathbb{N} \to \mathbb{N}$ and $s : \mathbb{N} \to [0,1]$ be parameters. We say that (V,P) is a one-round argument of ε-proximity for $\mathcal{L}$, with soundness (T,s), if the following two properties are satisfied:

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

3.2 Multi-prover Interactive Proofs (MIP)

Let $\mathcal{L}$ be a language and let x be an input of length n. In a one-round ℓ-prover interactive proof, ℓ computationally unbounded provers, $P_1,\ldots,P_\ell$, 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 its random string, the verifier generates ℓ queries, $q_1,\ldots,q_\ell$, one for each prover, and sends them to the ℓ provers. Each prover responds with an answer that depends only on its own individual query. That is, the provers respond with answers $a_1,\ldots,a_\ell$, 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 it receives (as well as the input x and its random string).

We say that $(V,P_1,\ldots,P_\ell)$ is a one-round multi-prover interactive proof system (MIP) for $\mathcal{L}$, with completeness $c \in [0,1]$ and soundness $s \in [0,1]$ (think of $s < c$) if the following two properties are satisfied:

1. **Completeness:** For every $x \in \mathcal{L}$, the verifier V accepts with probability c, over the random coins of $V,P_1,\ldots,P_\ell$, after interacting with $P_1,\ldots,P_\ell$, where c is a parameter referred to as the *completeness* of the proof system.
2. **Soundness:** For every $x \notin \mathcal{L}$, and any (computationally unbounded, possibly cheating) provers $P^*_1,\ldots,P^*_\ell$, the verifier V rejects with probability $\geq 1-s$, over the random coins of V, after interacting with $P^*_1,\ldots,P^*_\ell$, where s 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. We say that the proof-system has *perfect completeness* If completeness hold with probability 1 (i.e. $c = 1$).

No-Signaling MIP. We will consider a variant of the MIP model, where the cheating provers are more powerful. In the MIP model, each prover answers its 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 $S \subset [\ell]$, 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 $q = (q_1, \ldots, q_\ell) \in D^\ell$, let $\mathcal{A}_q$ be a distribution over Σ^ℓ. We think of $\mathcal{A}_q$ as the distribution of the answers for queries q.

We say that the family of distributions $\{\mathcal{A}_q\}_{q \in D^\ell}$ is *no-signaling* if for every subset $S \subset [\ell]$ and every two sequences of queries $q, q' \in D^\ell$, 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 $\{\mathcal{A}_q\}_{q \in D^\ell}$ is *δ-no-signaling*.

An MIP $(V, P_1, \ldots, P_\ell)$ for a language $\mathcal{L}$ is said to have soundness s 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 $\{\mathcal{A}_q\}_{q \in D^\ell}$, the verifier V rejects with probability $\geq 1 - s$, where on queries $q = (q_1, \ldots, q_\ell)$ the answers are given by $(a_1, \ldots, a_\ell) \in_R \mathcal{A}_q$, and s is the soundness parameter.

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

MIP of Proximity (MIPP). Let $\mathcal{L}$ be a language, let x be an input of length n (which we refer to as the main input) and let $\varepsilon = \varepsilon(n) \in (0, 1)$ be a proximity parameter. In a one-round ℓ-prover interactive proof of proximity, ℓ computationally unbounded provers, $P_1, \ldots, P_\ell$, try to convince a (probabilistic) polynomial-time verifier, V, that the input x is ε-close (in relative Hamming distance) to some $x' \in \mathcal{L}$. The provers have free access to n, ε and x. The verifier has free access to n and ε and oracle access to x (and the number of oracle queries is counted).

We say that $(V, P_1, \ldots, P_\ell)$ is a one-round multi-prover interactive proof system of ε-proximity (ε-MIPP) for $\mathcal{L}$, with completeness $c \in [0, 1]$ and soundness $s \in [0, 1]$, if the following properties are satisfied:

1. **Running Time:** The verifier runs in polynomial time, i.e., time polynomial in the communication complexity and the number of oracle queries.

2. **Completeness:** For every $x \in \mathcal{L}$ the verifier V accepts with probability c, after interacting with $P_1, \ldots, P_\ell$.
3. **Soundness:** For every x that is ε-far from $\mathcal{L}$, and any (computationally unbounded, possibly cheating) provers $P_1^*, \ldots, P_\ell^*$, the verifier V rejects with probability $\geq 1 - s$, after interacting with $P_1^*, \ldots, P_\ell^*$.

We denote such a proof system by ε-MIPP (and omit the soundness and completeness parameters from the notation). We say that the proof-system has *perfect completeness* if completeness hold with probability 1 (i.e. $c = 1$). The parameters we are mainly interested in are the query complexity and the communication complexity.

No-Signaling MIPP. An ε-MIPP, $(V, P_1, \ldots, P_\ell)$ for a language $\mathcal{L}$ is said to have soundness s against no-signaling strategies (or provers) if the following (more general) soundness property is satisfied:

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

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

4 Lower Bound for No-Signaling MIPP

In this section we prove a lower bound, showing that there does not exist a no-signaling MIPP for all of P with query complexity q and communication complexity c such that $q \cdot c = o(n)$ (where n is the input length). More specifically, for every q we construct a language $\mathcal{L}$ in P and prove that if exponentially hard pseudo-random generators exist then for any no-signaling ε-MIPP for $\mathcal{L}$ with query complexity q and communication complexity c, it must be the case that $q \cdot c = \Omega(n)$. In the full version we show how to extend the result to IPPs and to arguments of proximity.

In what follows we denote by τ the security parameter.

Definition 1. *A pseudo-random generator $G : \{0,1\}^n \to \{0,1\}^{\ell(n)}$ (with stretch $\ell(n) > n$) is said to be* exponentially hard *if for every circuit family $\{\mathcal{A}_\tau\}_\tau$ of size $2^{o(\tau)}$,*

$$\left| \Pr_{s \in_R \{0,1\}^\tau} [\mathcal{A}_\tau(1^\tau, G(s)) = 1] - \Pr_{y \in_R \{0,1\}^{\ell(\tau)}} [\mathcal{A}_\tau(1^\tau, y) = 1] \right| = \mathrm{negl}(\tau).$$

Theorem 7. *Assume the existence of exponentially hard pseudo-random generators. There exists a constant $\varepsilon > 0$ such that for every $q = q(n) \leq n$, there exists a language $\mathcal{L} \in$ P such that every* MIPP *for testing ε-proximity to $\mathcal{L}$ with completeness $2/3$, soundness $1/3$, query complexity q and communication complexity c it holds that $q \cdot c = \Omega(n)$.*

Remark 1. The above theorem holds with respect to any constant completeness parameter $c > 0$ and constant soundness parameter s such that $s < c$, and we chose $c = 2/3$ and $s = 1/3$ only for the sake of concreteness.

Remark 2. The assumption in Theorem 7 can be reduced to sub-exponentially hard pseudo-random generators (i.e., it is infeasible for circuits of size 2^{τ^δ} to distinguish the output of the generator from uniform, for some $\delta > 0$), rather than exponential hardness, at the cost of a weaker implication (i.e., $q \cdot c = \Omega(n^\delta)$).

Proof of Theorem 7. We start by defining the notion of average-case no-signaling MIP (in the crs model), which is used in the proof of Theorem 7. We note that this average-case completeness seems too weak for applications and we define this weak notion only for the sake of the proof of Theorem 7.

Definition 2. *An* average-case *no-signaling* MIP *in the* common random string *(*crs*) model, for a language $\mathcal{L}$, with completeness c and soundness s, consists of $(V, P_1, \ldots, P_\ell, \mathsf{crs})$, where as before V is the verifier, $P_1, \ldots, P_\ell$ are the provers, and* crs *is a* common random string *of length* $\mathsf{poly}(n)$*, chosen uniformly at random and given to all parties. In particular, V's queries and decision may depend on the* crs*, and the answers generated by both honest and cheating provers may depend on the* crs*. The following completeness and soundness conditions are required:*

- **Average-Case Completeness.** *For all sufficiently large $n \in \mathbb{N}$,*

$$\Pr\left[(V, P_1, \ldots, P_\ell)(x, \mathsf{crs}) = 1\right] \geq c,$$

  where the probability is over uniformly distributed $x \in_R \mathcal{L} \cap \{0,1\}^n$*, over uniformly generated* $\mathsf{crs} \in_R \{0,1\}^{\mathsf{poly}(n)}$*, and over the random coin tosses of the verifier V.*
- **Soundness Against No-Signaling Provers.** *For every $x \notin \mathcal{L}$, and every family of distributions $\{\mathcal{A}_{q,\mathsf{crs}}\}_{q \in D^\ell, \mathsf{crs} \in \{0,1\}^{\mathsf{poly}(n)}}$ such that for every* $\mathsf{crs} \in \{0,1\}^{\mathsf{poly}(n)}$ *the family of distributions $\{A_{q,\mathsf{crs}}\}_{q \in D^\ell}$ is no-signaling, the verifier V rejects with probability $\geq 1 - s$, where the answers corresponding to (q, crs) are given by $(a_1, \ldots, a_\ell) \in_R \mathcal{A}_{q,\mathsf{crs}}$.*

The following proposition, which we use in the proof of Theorem 7, follows from [DLN+04] (see also [Ito10]).

Proposition 1. *Suppose that a language $\mathcal{L}$ has an average-case no-signaling* MIP *in the* crs *model, communication complexity $c = c(n)$ (where n is the instance length), and with constant completeness and soundness (where the soundness parameter is smaller than the completeness parameter). Then, there exists a randomized algorithm D that runs in time* $\mathsf{poly}(n, 2^c)$ *such that:*

- *For every $n \in \mathbb{N}$,*

$$\Pr_{x \in_R \mathcal{L} \cap \{0,1\}^n}[D(x) = 1] \geq 2/3$$

  where the probability is also over the coin tosses of D.

– *For every $x \notin \mathcal{L}$ it holds that*

$$\Pr[D(x) = 1] \leq 1/3$$

where the probability is over the coins tosses of D.

We note that [DLN+04,Ito10] did not consider the crs model nor average-case completeness, but the claim extends readily to this setting as well.

We are now ready to prove Theorem 7.

Proof of Theorem 7. Assume that there exists a pseudo-random generator (PRG), denoted by $G : \{0,1\}^\tau \to \{0,1\}^{2\tau}$, that is exponentially secure. Namely, every adversary of size $2^{o(\tau)}$ cannot distinguish between uniformly distributed $r \in_R \{0,1\}^{2\tau}$ and $G(s)$ for uniformly distributed $s \in_R \{0,1\}^\tau$, with non-negligible advantage. For sake of simplicity, we assume that G is injective[6].

Let $\varepsilon > 0$ be a constant for which there exists a (good) error-correcting-code, denoted by ECC, with constant rate and efficient encoding that is resilient to (2ε)-fraction of adversarially chosen errors.

Fix any query complexity $q = o(n)$.[7] We show that there exists a language $\mathcal{L} \in \mathsf{P}$ such that for every no-signaling ε-MIPP for $\mathcal{L}$ with query complexity q and communication complexity c (and completeness $\frac{2}{3}$ and soundness $\frac{1}{3}$) it must be the case that $q \cdot c = \Omega(n)$.

Consider the following language:

$$\mathcal{L} = \left\{(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m) : \ \forall i \in [m], G(s_i) = r_i\right\},$$

where $m = 4q$ and $\tau = |s_i| = \Theta(n/q)$, where $n = |(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)|$. The fact that $|s_i| = \Theta(n/q)$ follows from the fact that ECC has constant rate (i.e., $|\mathsf{ECC}(z)| = O(|z|)$).

The fact that ECC is efficiently decodable and G is efficiently computable implies that $\mathcal{L} \in \mathsf{P}$. Suppose for contradiction that there exists a no-signaling ε-MIPP for $\mathcal{L}$, denoted by $(V, P_1, \ldots, P_\ell)$, with communication complexity c such that $c = o(n/q)$.

Consider the following NP language

$$\mathcal{L}_G = \{r : \exists s \text{ s.t. } G(s) = r\}.$$

Proposition 1, together with the fact that G is exponentially secure, implies that $\mathcal{L}_G$ does not have an average-case MIP in the crs model with soundness against no-signaling strategies, with communication complexity $o(\tau)$ for instances of length τ.

We obtain a contradiction by constructing an average-case MIP in the crs model with soundness against no-signaling strategies, with communication complexity $o(\tau)$. To this end, consider the following MIP in the crs model for $\mathcal{L}_G$, denoted by $(V', P'_1, \ldots, P'_\ell, \mathsf{crs})$.

[6] We note that this assumption can be easily removed by replacing the use of the uniform distribution over the language $\mathcal{L}'$ (defined below) with the distribution $G(s)$ for $s \in_R \{0,1\}^\tau$.

[7] Note that for $q = \Omega(n)$ the theorem is trivially true.

- The crs consists of m uniformly distributed seeds $s_1, \ldots, s_m \in_R \{0,1\}^\tau$, and a random coordinate $i \in_R [m]$.
- The verifier V', on input $r \in \{0,1\}^{2\tau}$, does the following:
  1. Let $r_i = r$, and for every $j \in [m] \setminus \{i\}$, let $r_j = G(s_j)$.
  2. Emulate V with oracle access to $(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)$. (Note that with overwhelming probability $r \neq G(s_i)$, and thus $r_i \neq G(s_i)$. However V will not notice this unless it queries coordinates that belong to s_i.)
- The provers $P'_1, \ldots, P'_\ell$, emulate $P_1, \ldots, P_\ell$ on input $(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)$, while setting $r_i = r$ and setting $s_i = s$ where $r = G(s)$ (assuming that such s exists).[8] If such s does not exist then the provers $P'_1, \ldots, P'_\ell$ send a reject message, and abort.

Note that the communication complexity of $(V', P'_1, \ldots, P'_\ell, \mathsf{crs})$ is equal to the communication complexity of $(V, P_1, \ldots, P_\ell, \mathsf{crs})$, denoted by c. By our assumption, $c = o(n/q) = o(\tau)$, as desired.

Average-Case Completeness. We need to prove that $\Pr[(V', P'_1, \ldots, P'_\ell)(r, \mathsf{crs}) = 1] \geq \frac{1}{2}$, where the probability is over *uniformly distributed* $r \in_R (\mathcal{L}_G)_\tau$, over uniformly generated $\mathsf{crs} = (s_1, \ldots, s_m, i)$ where each $s_j \in_R \{0,1\}^\tau$, $i \in_R [m]$, and over the random coin tosses of the verifier V.

Let GOOD denote the event that V' does not query any of the coordinates that belong to s_i, where $i \in [m]$ is the random coordinate chosen by V'. Notice that for every $r \in \mathcal{L}_G$,

$$\Pr\left[(V', P'_1, \ldots, P'_\ell)(r, \mathsf{crs}) = 1 \mid \mathsf{GOOD}\right] =$$

$$\Pr\left[(V, P_1, \ldots, P_\ell)(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m) = 1 \mid s_i \text{ is not queried}\right] \geq \frac{2}{3}$$

where the probabilities are over a uniformly distributed crs and the random coin tosses of V' and V, and where in the second equation $r_i = r$ and $s_i = s$, where $r = G(s)$. Recall that the fact that $r \in \mathcal{L}_G$ implies that such s exists.

The fact that

$$\Pr[(V', P'_1, \ldots, P'_\ell)(r, \mathsf{crs}) = 1] \geq \Pr[(V', P'_1, \ldots, P'_\ell)(r, \mathsf{crs}) = 1 \mid \mathsf{GOOD}] \cdot \Pr[\mathsf{GOOD}]$$

implies that it suffices to prove that $\Pr[\mathsf{GOOD}] \geq \frac{3}{4}$, where the probability is over uniformly distributed $r \in_R \mathcal{L}_G$, uniformly distributed crs, and over the random coin tosses of V'.

Note that $r_1, \ldots, r_m$ are all distributed identically to r, and thus $V, P_1, \ldots, P_\ell$, which all receive as input $(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)$, where $r_i = r$, do not have any advantage in guessing i (here we crucially use the fact that the MIPP provers are not given access to the crs). Therefore, since V makes at most q queries,

[8] This step can be done by a brute force search (since the honest provers are also computationally unbounded). Nevertheless, we note that typically in proof-systems for languages in NP the prover is given the NP witness and so this step can also be done efficiently.

and since $m = 4q$, it follows from the union bound that V queries any location of s_i with probability at most $\frac{q}{m} = \frac{1}{4}$. Hence, $\Pr[\mathsf{GOOD}] \geq \frac{3}{4}$ and (average-case) completeness follows.

Soundness Against No-Signaling Strategies. We prove that for every $r \notin \mathcal{L}_G$, every $\mathsf{crs} = (s_1, \ldots, s_m, i)$, and every no-signaling cheating strategy $P^{\mathsf{NS}} = (P_1^*, \ldots, P_\ell^*)$, it holds that $\Pr[(V', P^{\mathsf{NS}})(r, \mathsf{crs}) = 1] \leq \frac{1}{3}$, where the probability is over the random coin tosses of V' and P^{NS}.

To this end, fix any $r \notin \mathcal{L}_G$ and any $\mathsf{crs} = (s_1, \ldots, s_m, i)$ where each $s_j \in \{0,1\}^\tau$ and $i \in [m]$. Suppose for the sake of contradiction that there exists a no-signaling cheating strategy $P^{\mathsf{NS}} = (P_1^*, \ldots, P_\ell^*)$ such that $\Pr[(V', P^{\mathsf{NS}})(r, \mathsf{crs}) = 1] > \frac{1}{3}$, where the probability is over the random coin tosses of V' and P^{NS}.

Recall that V' runs V on input $(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)$, where $r_i = r$ and where $r_j = G(s_j)$ for every $j \in [m] \setminus \{i\}$. We prove that there exists a no-signaling cheating strategy, denoted by $\hat{P}^{\mathsf{NS}}$, such that

$$\Pr\left[\left(V, \hat{P}^{\mathsf{NS}}\right)(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m) = 1\right] > \frac{1}{3}, \tag{1}$$

where the probability is over the random coin tosses of V and $\hat{P}^{\mathsf{NS}}$.

The cheating strategy $\hat{P}^{\mathsf{NS}}$ simply emulates P^{NS}. Namely, $\hat{P}^{\mathsf{NS}}$, upon receiving queries $(q_1, \ldots, q_\ell)$, will emulate $P^{\mathsf{NS}}(r, \mathsf{crs})$ upon receiving $(q_1, \ldots, q_\ell)$, where $r = r_i$ and $\mathsf{crs} = (s_1, \ldots, s_m, i)$. Note that $\hat{P}^{\mathsf{NS}}$ simulates P^{NS} perfectly, and therefore indeed Equation (1) holds. Also note that the fact that P^{NS} is a no-signaling strategy immediately implies that $\hat{P}^{\mathsf{NS}}$ is also a no-signaling strategy.

To get a contradiction, it thus remains to show that $(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)$ is ε-far from $\mathcal{L}$. Indeed, the fact that ECC is an error correcting code resilient to 2ε-fraction of adversarial errors, together with the fact that $r \notin \mathcal{L}_G$ implies that $(\mathsf{ECC}(r_1, \ldots, r_m), s_1, \ldots, s_m)$ is ε-far from $\mathcal{L}$, as desired. □

Acknowledgments.. We thank Guy Rothblum for pointing out to us the question about arguments of proximity for P - the question that initiated this work. The second author was supported by the Israel Science Foundation (grant No. 671/13).

References

[ABOR00] Aiello, W., Bhatt, S., Ostrovsky, R., Rajagopalan, S.R.: Fast verification of any remote procedure call: short witness-indistinguishable one-round proofs for NP. In: Welzl, E., Montanari, U., Rolim, J.D.P. (eds.) ICALP 2000. LNCS, vol. 1853, pp. 463–474. Springer, Heidelberg (2000)

[AIK10] Applebaum, B., Ishai, Y., Kushilevitz, E.: From secrecy to soundness: efficient verification via secure computation. In: Abramsky, S., Gavoille, C., Kirchner, C., Meyer auf der Heide, F., Spirakis, P.G. (eds.) ICALP 2010. LNCS, vol. 6198, pp. 152–163. Springer, Heidelberg (2010)

[BCCT12a] Bitansky, N., Canetti, R., Chiesa, A., Tromer, E.: From extractable collision resistance to succinct non-interactive arguments of knowledge, and back again. In: ITCS, pp. 326–349 (2012)

[BCCT12b] Bitansky, N., Canetti, R., Chiesa, A., Tromer, E.: Recursive composition and bootstrapping for snarks and proof-carrying data. IACR Cryptology ePrint Archive 2012:95 (2012)

[BSGH+06] Ben-Sasson, E., Goldreich, O., Harsha, P., Sudan, M., Vadhan, S.P.: Robust PCPs of proximity, shorter PCPs, and applications to coding. SIAM J. Comput. **36**(4), 889–974 (2006)

[CKLR11] Chung, K.-M., Kalai, Y.T., Liu, F.-H., Raz, R.: Memory delegation. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 151–168. Springer, Heidelberg (2011)

[CKV10] Chung, K.-M., Kalai, Y., Vadhan, S.: Improved delegation of computation using fully homomorphic encryption. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 483–501. Springer, Heidelberg (2010)

[DFH12] Damgård, I., Faust, S., Hazay, C.: Secure two-party computation with low communication. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 54–74. Springer, Heidelberg (2012)

[DLN+04] Dwork, C., Langberg, M., Naor, M., Nissim, K., Reingold, O.: Succinct proofs for NP and spooky interactions. Unpublished manuscript (2004). http://www.cs.bgu.ac.il/kobbi/papers/spooky_sub_crypto.pdf

[DR06] Dinur, I., Reingold, O.: Assignment testers: Towards a combinatorial proof of the PCP theorem. SIAM J. Comput. **36**(4), 975–1024 (2006)

[EKR04] Funda Ergün, Ravi Kumar, and Ronitt Rubinfeld: Fast approximate probabilistically checkable proofs. Inf. Comput. **189**(2), 135–159 (2004)

[FGL14] Fischer, E., Goldhirsh, Y., Lachish, O.: Partial tests, universal tests and decomposability. In: ITCS, pp. 483–500 (2014)

[GGP10] Gennaro, R., Gentry, C., Parno, B.: Non-interactive verifiable computing: outsourcing computation to untrusted workers. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 465–482. Springer, Heidelberg (2010)

[GGPR12] Gennaro, R., Gentry, C., Parno, B., Raykova, M.: Quadratic span programs and succinct NIZKs without PCPs. IACR Cryptology ePrint Archive 2012:215 (2012)

[GGR98] Goldreich, O., Goldwasser, S., Ron, D.: Property testing and its connection to learning and approximation. J. ACM (JACM) **45**(4), 653–750 (1998)

[GKR08] Goldwasser, S., Kalai, Y.T., Rothblum, G.N.: Delegating computation: interactive proofs for muggles. In: STOC, pp. 113–122 (2008)

[GLR11] Goldwasser, S., Lin, H., Rubinstein, A.: Delegation of computation without rejection problem from designated verifier cs-proofs. IACR Cryptology ePrint Archive 2011:456 (2011)

[GR13] Gur, T., Rothblum, R.D.: Non-interactive proofs of proximity. Electronic Colloquium on Computational Complexity (ECCC), 20:78 (2013)

[Gro10] Groth, J.: Short pairing-based non-interactive zero-knowledge arguments. In: Abe, M. (ed.) ASIACRYPT 2010. LNCS, vol. 6477, pp. 321–340. Springer, Heidelberg (2010)

[GW11] Gentry, C., Wichs, D.: Separating succinct non-interactive arguments from all falsifiable assumptions. In: STOC, pp. 99–108 (2011)

[Hol09] Holenstein, T.: Parallel repetition: simplification and the no-signaling case. Theory Comput. **5**(1), 141–172 (2009)

[Ito10] Ito, T.: Polynomial-space approximation of no-signaling provers. In: Abramsky, S., Gavoille, C., Kirchner, C., Meyer auf der Heide, F., Spirakis, P.G. (eds.) ICALP 2010. LNCS, vol. 6198, pp. 140–151. Springer, Heidelberg (2010)

[Kil92] Kilian, J.: A note on efficient zero-knowledge proofs and arguments (extended abstract). In: STOC, pp. 723–732 (1992)

[KRR13a] Kalai, Y.T., Raz, R., Rothblum, R.D.: Delegation for bounded space. In: STOC, pp. 565–574 (2013)

[KRR13b] Kalai, Y.T., Raz, R., Rothblum, R.D.: How to delegate computations: The power of no-signaling proofs. Electronic Colloquium on Computational Complexity (ECCC), 20:183 (2013)

[Lip12] Lipmaa, H.: Progression-free sets and sublinear pairing-based non-interactive zero-knowledge arguments. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 169–189. Springer, Heidelberg (2012)

[Mic94] Micali, S.: CS proofs (extended abstracts). In: FOCS, pp. 436–453 (1994)

[Nao03] Naor, M.: On cryptographic assumptions and challenges. In: Boneh, D. (ed.) CRYPTO 2003. LNCS, vol. 2729, pp. 96–109. Springer, Heidelberg (2003)

[PRV12] Parno, B., Raykova, M., Vaikuntanathan, V.: How to delegate and verify in public: verifiable computation from attribute-based encryption. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 422–439. Springer, Heidelberg (2012)

[RVW13] Rothblum, G.N., Vadhan, S.P., Wigderson, A.: Interactive proofs of proximity: delegating computation in sublinear time. In: STOC, pp. 793–802 (2013)

[Spi96] Spielman, D.A.: Linear-time encodable and decodable error-correcting codes. IEEE Trans. Inf. Theory **42**(6), 1723–1731 (1996)

[Tha13] Thaler, J.: Time-optimal interactive proofs for circuit evaluation. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part II. LNCS, vol. 8043, pp. 71–89. Springer, Heidelberg (2013)

[VSBW13] Vu, V., Setty, S.T.V., Blumberg, A.J., Walfish, M.: A hybrid architecture for interactive verifiable computation. In: IEEE Symposium on Security and Privacy, pp. 223–237 (2013)