From Obfuscation to the Security of Fiat-Shamir for Proofs

Yael Tauman Kalai[1(✉)], Guy N. Rothblum[2], and Ron D. Rothblum[3]

[1] Microsoft Research, Cambridge, USA
yaelism@gmail.com
[2] Weizmann Institute of Science, Rehovot, Israel
[3] MIT, Cambridge, USA

Abstract. The Fiat-Shamir paradigm [CRYPTO'86] is a heuristic for converting three-round identification schemes into signature schemes, and more generally, for collapsing rounds in constant-round public-coin interactive protocols. This heuristic is very popular both in theory and in practice, and its security has been the focus of extensive study.

In particular, this paradigm was shown to be secure in the Random Oracle Model. However, in the plain model, the results shown were mostly negative. In particular, the heuristic was shown to be *insecure* when applied to *computationally sound* proofs (also known as arguments). Moreover, recently it was shown that even in the restricted setting where the heuristic is applied to interactive *proofs* (as opposed to arguments), its soundness cannot be proven via a black-box reduction to any so-called *falsifiable* assumption.

In this work, we give a *positive result* for the security of this paradigm in the *plain model*. Specifically, we construct a hash function for which the Fiat Shamir paradigm is *secure* when applied to proofs (as opposed to arguments), assuming the existence of a sub-exponentially secure indistinguishability obfuscator, the existence of an exponentially secure input-hiding obfuscator for the class of multi-bit point functions, and the existence of a sub-exponentially secure one-way function.

More generally, we construct a hash family that is *correlation intractable* (under the computational assumptions above), solving an open problem originally posed by Canetti, Goldreich and Halevi (JACM, 2004), under the above assumptions.

In addition, we show that our result resolves a long-lasting open problem in about zero-knowledge proofs: It implies that there does not exist a public-coin constant-round zero-knowledge proof with negligible soundness (under the assumptions stated above).

1 Introduction

In 1986, Fiat and Shamir [FS86] proposed a general method for converting any three-round identification (ID) scheme into a signature scheme. This method quickly gained popularity both in theory and in practice, since known ID schemes (in which a sender *interactively* identifies himself to a receiver) are significantly

J. Katz and H. Shacham (Eds.): CRYPTO 2017, Part II, LNCS 10402, pp. 224–251, 2017.
DOI: 10.1007/978-3-319-63715-0_8

simpler and more efficient than known signature schemes, and thus this heuristic gives an efficient and easy way to implement digital signature schemes.

The Fiat-Shamir method is both simple and intuitive: The public key of the signature scheme consists of a pair (pk, H), where pk is a public key corresponding to the underlying ID scheme, and H is a hash function chosen at random from a hash family. To sign a message m, compute a triplet (α, β, γ), such that $\beta = H(\alpha, m)$ and (α, β, γ) is an accepting transcript of the ID scheme with respect to pk.

The main question is:

Is the Fiat-Shamir heuristic sound?

Namely, for what hash function families is the signature scheme, obtained by applying the Fiat-Shamir heuristic to a secure ID scheme, secure against adaptive chosen message attacks?

The intuition for why the heuristic may be sound, is that if H looks like a truly random function, and if all the adversary (i.e., impersonator) can do is use H in a black-box manner, then interacting with H is similar to interacting with the real verifier. This intuition was formalized by Pointcheval and Stern [PS96], and by followup works [OO98, AABN02], who proved that the Fiat-Shamir heuristic is sound in the so-called *Random Oracle Model* (ROM) – when the hash function is modeled by a random oracle [BR93], assuming the underlying ID scheme is sound against passive impersonation attacks.

This led to the belief that if a 2-round protocol, obtained by applying the Fiat-Shamir paradigm, is insecure, then it must be the case that the hash family used is not "secure enough", and the hope was that there exists another hash family that is sufficiently secure. These positive results (in the ROM), together with the popularity and importance of the Fiat-Shamir heuristic, led many researchers to try to prove the security of this paradigm in the plain model (without resorting to random oracles). Unfortunately, these attempts led mainly to negative results.

Goldwasser and Kalai [GK03] proved a negative result, by constructing a (contrived) 3-round public-coin ID scheme, for which the resulting signature scheme obtained by applying the Fiat-Shamir heuristic, is insecure, no matter which hash family is used.

Extending the Fiat-Shamir Heuristic. The Fiat-Shamir heuristic can be used outside the regime of ID and signature schemes. It can be used to convert any constant-round public-coin proof system into a two-round proof system, as follows: In the first round, the verifier sends a hash function H, where H is chosen at random from a hash family; in the second round, the prover sends the entire transcript of the interactive protocol, where the verifier's messages are computed by applying H to the communication so far.

The first work to extend the Fiat-Shamir paradigm to this regime, was the work of Micali [Mic94] on CS-proofs. We note that in this regime, the importance of the Fiat-Shamir heuristic stems from the fact that latency, caused by sending messages back and forth, is often a bottleneck in running cryptographic protocols [MNPS04, BDNP08].

The main question about this (extended) heuristic is therefore:

Is the two-message proof system obtained by applying the Fiat-Shamir heuristic, to a constant-round proof system, sound?

Namely, does there exist an explicit hash family, for which is it infeasible for a (computationally bounded) cheating prover, given an input outside the language and a random function H from the family, to generate an accepting transcript for the original interactive protocol (where each verifier-message is computed by applying H to the communication so far).

Barak [Bar01] gave the first negative result in the "plain model", by constructing a constant-round public-coin protocol, such that for any hash family $\mathcal{H}$, the resulting 2-round protocol, obtained by applying the Fiat-Shamir heuristic to this interactive protocol with respect to $\mathcal{H}$, is not sound.[1] However, the interactive protocol constructed in [Bar01] has only computational soundness, and thus is an *argument system* (as opposed to a proof). This gave rise to the following question:

Is the Fiat-Shamir method secure when applied to interactive ***proofs*** *(as opposed to arguments)?*

Namely, does there exist an explicit hash family for which the transformation, when applied to an *information-theoretically sound* interactive proof, produces a (computationally) sound two-message argument system?

In this work, we give a positive answer to this final question (under strong computational assumptions). Before we present our results in detail, we describe previous works which attempted to answer this question.

Barak, Lindell and Vadhan [BLV06] presented a security property for the Fiat-Shamir hash function which, if realized, would imply the soundness of the Fiat-Shamir paradigm applied to any constant-round public-coin interactive proof system.[2] However, they left open the problem of realizing this security definition under standard hardness assumptions (or under any assumption beyond simply assuming that the definition holds for a given hash function).

Dodis, Ristenpart and Vadhan [DRV12] showed that under specific assumptions regarding the existence of robust randomness condensers for seed-dependent sources, the definitions of [BLV06] can be realized. However, the question of constructing such suitable robust randomness condensers was left open by [DRV12].

On the other hand, Bitansky *et al.* [BDG+13] gave a negative result. They showed that that soundness of the Fiat-Shamir paradigm, even when applied to

[1] We note that the work of [GK03] is a followup work to [Bar01], and builds upon its techniques.

[2] Loosely speaking, a hash family $\{h_s\}$ is said to have this security property if for every probabilistic polynomial time adversary $\mathcal{A}$, that is given a random seed s and outputs an element in the domain of h_s, the random variable $h_s(\mathcal{A}(s))$ conditioned on $\mathcal{A}(s)$ has almost full min entropy.

interactive proofs, cannot be proved via a black-box reduction to any so-called *falsifiable* assumption, a notion defined by Naor [Nao03]).[3,4]

Correlation Intractable Hash Functions. Our results can be cast more generally in the language of *correlation intractability*, a notion defined in the seminal work of Canetti, Goldreich and Halevi [CGH04].

Roughly speaking, a correlation intractable function family is one for which it is infeasible to find input-output pairs that satisfy some "rare" relation. More precisely, a binary relation R is said to be *evasive* if for every value x only negligible fraction of the y values satisfy $(x, y) \in R$. A function family $F = \{f_s\}$ is correlation intractable if for every evasive relation R it is computationally hard, given a description of a random function $f_s \in F$, to find a value x such that $(x, f_s(x)) \in R$.

It was shown in [CGH04] that there does not exist a correlation intractable hash family whose seeds are shorter than the input length. The question of whether there exists a correlation intractable function family whose seeds are larger than the input, remained open. Very recently, [CCR15] construct a function family that is correlation intractable with respect to all relations that are computable in a-priori bounded polynomial complexity (under computational assumptions).

In this work, we construct a correlation intractable hash family with respect to *all* relations (under computational assumptions). We provide a more detailed comparison between our work and that of [CCR15] after we present our result more formally, below.

1.1 Our Results

In this work, we construct a hash family, and prove that the Fiat-Shamir paradigm is *sound* w.r.t. this hash family, when applied to interactive proofs (as opposed to arguments). We also show that the family is correlation intractable. Both results are shown under the following three cryptographic assumptions:

1. The existence of 2^n-secure indistinguishability obfuscation iO, where 2^n is the domain size of the functions being obfuscated.[5]

[3] The formalization of a falsifiable assumption, given in [BDG+13], is similar to the formalization given in [GW11], and differs slightly from the formalization given in [Nao03].

[4] Our assumptions (see Sect. 1.1), which deal with exponential-time (rather than polynomial-time) adversaries, are inherently not falsifiable. Note that [BDG+13] allow an unbounded challenger, but restrict to polynomial-time attackers. In the context of obfuscation, the attacker is the algorithm trying to *break* the security of the obfuscation. We assume hardness against super polynomial-time attackers, and thus our assumptions do not fall into the category ruled out by Bitansky *et al.*

[5] This assumption has been made in many previous works on iO and is referred to as sub-exponential iO, since the security parameter can be polynomially larger than n (which makes 2^n sub-exponential in the security parameter).

Recently, several constructions of iO obfuscation were proposed, starting with the work of Garg *et al.* [GGH+13]. However, to date, none of these constructions are known to be provably secure under what is known as a complexity assumption [GK16] or more generally a falsifiable assumption [Nao03]. We mention that [GLSW14] provided a construction and proved its security under the subgroup elimination assumption, which is a complexity assumption (and in particular is a falsifiable assumption). However, this assumption has been refuted in all candidate multi-linear groups.

2. The existence of 2^n-secure puncturable pseudo-random function (PRF) family $\mathcal{F}$, where 2^n is the domain size.
   Puncturable PRFs were defined in [BW13,BGI14,KPTZ13]. The PRF family of [GGM86] is a puncturable PRF family, and thus 2^n-secure puncturable PRFs can be constructed from any sub-exponentially secure one-way function.
3. The existence of an exponentially secure input-hiding obfuscation hideO for the class of multi-bit point functions $\{\mathcal{I}_{n,k}\}$.
   The class $\{\mathcal{I}_{n,k}\}$ consists of functions of the form $I_{\alpha,\beta}$ where $|\alpha| = n$ and $|\beta| = k$, and where $I_{\alpha,\beta}(x) = \beta$ for $x = \alpha$ and $I_{\alpha,\beta}(x) = 0$ otherwise. An obfuscation for this class is said to be input-hiding with T-security if any *poly-size* adversary that is given an obfuscation of a *random* function $I_{\alpha,\beta}$ in this family, guesses α with probability at most T^{-1}. Note that we assume hardness for a distribution where the value β may be correlated with α and furthermore, it may be computationally difficult to find β from α.
   For our construction we require T that is roughly equal to $2^n \cdot \mu$, where μ is the soundness error of the underlying proof system. For example, if we start off with an interactive proof with soundness error 2^{-n^ϵ} (where n is an upper bound the length of prover messages), then we require roughly $T = 2^{n-n^\epsilon}$. For constructing correlation intractable functions, μ is the "evasiveness" of the relation R. That is, for every value x, the fraction of y's satisfying $(x, y) \in R$ is at most μ.
   This assumption was considered in [CD08,BC14], who also provided a candidate construction based on a strong variant of the DDH assumption (we elaborate on this in Sect. 2.4).[6] See further discussion on various notions of point function obfuscation in [BS16].
   We emphasize that we *do not* assume security of the multi-bit point function obfuscation with *auxiliary input*. Indeed, security with auxiliary input is known to be problematic, and, as was shown by Brzuska and Mittelbach [BM14], if iO obfuscation exists then multi-bit point function obfuscation with auxiliary inputs does not exist. We do not allow auxiliary information, and we only assume *input-hiding* (against exponential-time adversaries) for a *random* function from the family (rather than black-box worst-case).

[6] While DDH (and even discrete log) can be broken in time less than 2^n (even in the generic group model - e.g., by the baby-step giant-step algorithm), this does not imply a non-trivial *polynomial-time* attack (i.e., one with success probability greater than $\text{poly}(n)/2^n$).

Theorem 1 *[(Informally Stated, see Theorem 4)]. Under the assumptions above, for any constant-round public-coin interactive proof Π, the resulting 2-message argument Π^{FS}, obtained by applying the Fiat-Shamir paradigm to Π with the function family $\mathsf{iO}(\mathcal{F})$, is sound.*

This theorem provides a general-purpose transformation for reducing interaction in interactive proof systems. Beyond our primary motivation of studying the security of the Fiat-Shamir transformation (and its implications to zero-knowledge proofs), the secure transformation can also serve as an avenue for obtaining new public-coin 2-message argument systems (often referred to as publicly-verifiable non-interactive arguments). For example, it can be applied to the interactive proofs of [RRR16] to obtain arguments for bounded-space polynomial-time computations, with small communication and almost-linear-time verification. We note, however, that prior works [BGL+15] have shown how to construct such arguments for general polynomial-time computations using subexponential iO and one-way functions (without the need for multi-bit point function obfuscation). Nonetheless, one advantage of Theorem 1 is that it can be applied to *any* interactive proof, which may give more efficient arguments for specific languages in P and for languages outside of P.

Cast in the language of correlation intractability, we prove:

Theorem 2 *[(Informally Stated)]. Under the assumptions above, the function family $\mathsf{iO}(\mathcal{F})$ is correlation intractable.*

Here and throughout this work $\mathsf{iO}(\mathcal{F})$ refers to an iO obfuscation of a program that computes the PRF, using a hardwired random seed.

Remark 1. Although outside the scope of this paper, we note that this transformation from interactive proofs to 2-message arguments preserves some secrecy guarantees.

In particular, it is easy to see that the Fiat-Shamir paradigm always preserves witness indistinguishability. Namely, if the underlying interactive proof is witness indistinguishable then the resulting 2-message argument, obtained by applying the Fiat-Shamir method with respect to *any* function family, is also witness indistinguishability. Loosely speaking, this follows from the fact that witness indistinguishability is defined to hold with respect to any cheating (poly-size) verifier.

Moreover, we claim that the Fiat-Shamir paradigm, applied with our function family $\mathsf{iO}(\mathcal{F})$, preserves honest-verifier zero-knowledge. Loosely speaking, this (non-trivial) claim follows from the following argument: To simulate the 2-message argument with respect to some input x, first use the simulator for the interactive proof to obtain a simulated transcript $(m_1, r_1, \ldots, m_c, r_c, m_{c+1})$. Note that this transcript may not be consistent with any hash function from the family. To obtain a simulated transcript for the 2-message argument, we simulate the verifier as sending the iO of a randomly chosen PRF function $f_s \leftarrow \mathcal{F}$, punctured at the points $m_1, (m_1, r_1, m_2), \ldots, (m_1, r_1, \ldots, m_{c_1}, r_{c-1}, m_c)\}$, and hardwire the

values $r_1, r_2 \ldots, r_c$ for these points (respectively). Standard iO techniques can be used to argue that this obfuscated circuit is indistinguishable from $\mathsf{iO}(f_s)$.

As we discuss next, Theorem 1 settles a long lasting open problem about zero-knowledge proofs.

Impossibility of Constant-Round Public-Coin Zero-Knowledge. Hada and Tanaka [HT98] and Dwork *et al.* [DNRS99] observed an intriguing connection between the security of the Fiat-Shamir paradigm and the existence of certain zero-knowledge protocols. In particular, if there exists a constant-round public-coin zero-knowledge proof for a language outside BPP, then the Fiat-Shamir paradigm is not secure when applied to this zero-knowledge proof.[7] Intuitively, this follows from the following observation: Consider the cheating verifier that behaves exactly like the Fiat-Shamir hash function. The fact that the protocol is zero-knowledge implies that there exists a simulator who can simulate the view in an indistinguishable manner. Thus, for elements in the language the simulator generates accepting transcripts. The simulator cannot distinguish between elements in the language and elements outside the language (since the simulator runs in poly-time and the language is outside of BPP). In addition, the protocol is public-coin, which implies that the simulator knows whether the transcript is accepted or not. Hence, it must be the case that the simulator also generates accepting transcripts for elements that are not in the language, which implies that the Fiat-Shamir paradigm is not secure.

Thus, Theorem 1, combined with [DNRS99, Theorem 5.4] implies the following corollary.

Corollary 1. *Under the assumptions above, there does not exist a constant-round public-coin zero-knowledge proof with negligible soundness for languages outside* BPP.

We emphasize that the above negative result not only rule out black-box simulation, but also rules out *non-black-box* simulation. Moreover, as pointed out by [DNRS99], this negative result actually rules out even extremely weak notions of zero-knowledge which they call *ultra weak zero knowledge* (see [DNRS99, Sect. 5]).

In particular, this corollary implies that (under the assumptions above) *parallel repetition of Blum's Hamiltonicity protocol for* NP [Blu87] *is not zero-knowledge*. Previously it was not known whether (in general) parallel repetition preserves zero-knowledge. Our result shows that it does not (under the assumptions above).

The existence of constant-round public-coin zero-knowledge proofs has been a long-standing open question (see, e.g., [GO94, GK96, KPR98, Ros00, CKPR02, BLV06, BGGL01, BL04, Rey01]). For *black-box* zero-knowledge proofs (which means that the simulator only uses the verifier as a black-box), the work of Goldreich and Krawczyk [GK96] ruled out constant-round public-coin protocols

[7] We note that this is how Barak [Bar01] obtained his negative result. He constructed a constant-round public-coin zero-knowledge argument.

(for languages outside of BPP). It is known, however, that non black-box techniques can be quite powerful in the context of zero-knowledge [Bar01]. Under the assumptions stated above, our work rules out *any* constant-round public-coin zero knowledge proof (even non black-box ones).

We note that even for those who are skeptical about the obfuscation assumptions we make, this corollary implies that finding a constant-round public-coin zero-knowledge proof requires overcoming technical barriers, and in particular requires disproving the existence of sub-exponentially secure iO obfuscation, or the existence of exponentially secure input-hiding obfuscation for the class of multi-bit point functions (or, less likely, disproving the existence of sub-exponential OWF).

Comparison to Concurrent Works

Comparison to [CCR15]. As mentioned above, in a concurrent and independent work, Canetti *et al.* [CCR15] construct a correlation intractable function family that withstands all relations computable in a-priori bounded polynomial complexity. More specifically, they construct a function family that is correlation intractable with respect to all evasive relations that can be computed in time p, for any a priori polynomial p, where the size of the functions in the family grows with p.

We note that this result does not have any implications to the security of the Fiat-Shamir paradigm, since to prove the security of this paradigm we need a correlation intractable ensemble for relations that cannot be computed in polynomial time. Moreover, we note that since the size of the functions grow with p, leveraging techniques do not seem to apply here.

As mentioned above, our result on the security of the Fiat-Shamir paradigm can be cast more generally in the language of *correlation intractability*. In particular, the hash family that we construct, and with which we prove the security of the Fiat-Shamir paradigm, is correlation intractable (with respect to all relations) under our assumption stated above.

In terms of the assumptions used, [CCR15] assume the existence of sub-exponentially secure indistinguishability obfuscation, the existence of a sub-exponentially secure puncturable PRF family, and the existence of input-hiding obfuscation for the class of evasive functions [BBC+14]. Comparing to the assumptions we make in this work, we also make the first two assumptions. However, we assume input-hiding obfuscation only for multi-bit point functions (a significantly smaller family compared to general evasive functions). On the other hand, we require an exponentially secure input-hiding obfuscation, whereas their work only needs polynomial-time hardness of the input-hiding obfuscation.

Comparison with [MV16]. In an additional independent and concurrent work, Mittelbach and Venturi [MV16] showed a hash function for which the Fiat-Shamir is secure for a very *particular* class of protocols. The class of protocols that they consider in itself does not include any previously-studied protocols. However, [MV16] show an additional transformation for 3 message protocols (on top of Fiat-Shamir) that works when the first message in the underlying

3-message protocol is *independent* (as a function) of the input. Mittelbach and Venturi also show that their transformation, which is based on indistinguishability obfuscation, maintains zero-knowledge, and can be used to obtain signature schemes and NIZKs.

In contrast to [MV16], our primary motivation and goal is showing that the Fiat-Shamir transformation can be used to reduce interaction while preserving soundness. Reducing the interaction in cryptographic protocols and particularly showing that the Fiat-Shamir transform can be proved sound has been a central and widely-studied question in the cryptographic literature. We emphasize that the [MV16] result does *not* yield a method for reducing rounds while preserving soundness.[8]

1.2 Overview

Throughout this overview we focus on proving the security of the Fiat-Shamir paradigm, when applied to 3-round public-coin interactive proofs. The more general case, of any constant number[9] of rounds, is then proved by induction on the number of rounds (we refer the reader to Sect. 4 for details). Consider any 3-round proof Π for a language L. Denote the transcript by (α, β, γ) where α is the first message sent by the prover, β is the random message sent by the verifier, and γ is the final message sent by the prover. Fix any $x \notin L$. The fact that Π is a sound proof means that for every α, for most of the verifier's messages β, there does not exist γ that makes the verifier accept.

The basic idea stems from the original intuition for why the Fiat-Shamir is secure, which is that if we use a hash function H that looks like a truly random function, then all the prover can do is use H in a black-box manner, in which case interacting with H is similar to interacting with the real verifier, and hence security follows.

The first idea that comes to mind is to choose the hash function randomly from a pseudo-random function (PRF) family. However, the security guarantee of a PRF is that given only *black-box* access to a random function f in the PRF family, one cannot distinguish it from a truly random function. No guarantees are given if the adversary is given a succinct circuit for computing f.

Obfuscation to the Rescue. A natural next step is to try to obfuscate f, in the hope that whatever can be learned given the obfuscation of f can also be learned from black-box access to f. However, this requires virtual-black-box (VBB) security, and VBB obfuscation is known not to exist [BGI+12]. Moreover, there are

[8] Indeed, for the class of protocols that [MV16] support, reducing to 2 rounds while preserving soundness (but not necessarily zero-knowledge) is straightforward: Since the prover's first message is not a function of the input, the verifier can compute the prover's first message α for it, and sends α (together with the coins used to generate it) to the prover.

[9] The Fiat Shamir paradigm refers to constant round protocols. Indeed, there are interactive proofs with a super-constant number of rounds (and negligible soundness error) for which the Fiat Shamir paradigm is insecure.

specific PRF families for which VBB obfuscation is impossible [BGI+12]. Further obstacles to VBB obfuscation of PRFs and, more generally, functions with high pseudo-entropy (w.r.t. auxiliary input) are given in [GK05,BCC+14]. Given these obstacles to achieving VBB obfuscation, could we hope to prove security using relaxed notions of obfuscation, such as iO obfuscation? The question is:

Is iO *obfuscation strong enough to prove the security of the Fiat-Shamir paradigm?*

It is well known that iO obfuscation is *not* strong enough to prove the security of the Fiat-Shamir paradigm when applied to computationally sound interactive *arguments*. Indeed the Fiat-Shamir paradigm is known be insecure when applied to arguments as opposed to proofs.[10] In contrast, we show that iO obfuscation (together with additional assumptions) is strong enough to prove security when the Fiat-Shamir paradigm is applied to interactive *proofs* (rather than arguments).

For proving security of the Fiat-Shamir paradigm for *proofs*, consider a cheating prover for the transformed protocol Π^{FS}, who receives the obfuscation $\mathsf{iO}(f_s)$ of a pseudo-random function f_s. Since f_s is a PRF, we know that there will only be a small set Bad_s of inputs α (corresponding to the prover's first message in the proof Π), for which the communication prefix $(\alpha, f_s(\alpha))$ can lead the verifier in the interactive proof to accept (i.e. α's for which there exists γ s.t. $(\alpha, f(\alpha), \gamma)$ is an accepting transcript).

To show the security of the resulting protocol, we now want to claim that the obfuscation *hides* this (small) set Bad_s of inputs, and that a cheating prover P^* cannot find any input $\alpha \in \mathsf{Bad}_s$. Note, however, that iO obfuscation only guarantees that one cannot distinguish between the obfuscation of two functionally equivalent circuits of the same size, and it does not give any hiding guarantees.

Puncturable PRFs to the Rescue? As mentioned above, iO obfuscation does not immediately seem to give any hiding guarantees. Nonetheless, starting with the beautiful work of Sahai and Waters [SW14], iO has proved remarkably powerful in the construction of a huge variety of cryptographic primitives. A basic technique used in order to get a hiding guarantee from iO obfuscation, as pioneered in [SW14], is to use it with a puncturable PRF family.

A puncturable PRF family is a PRF family that allows the "puncturing" of the seed at any point α in the domain of f. Namely, for any point α in the domain, and for any seed s of the PRF, one can generate a "punctured" seed, denoted by $s\{\alpha\}$. This seed allows the computation of f_s anywhere in the domain, except at point α, with the security guarantee that for a random seed s chosen independently of α, the element $f_s(\alpha)$ looks (computationally) random given $(s\{\alpha\}, \alpha)$. The security of iO obfuscation guarantees that one cannot distin-

[10] More specifically, the insecurity is in the sense that there exist contrived interactive arguments such that for any hash family $\mathcal{H}$, applying the Fiat-Shamir paradigm with the hash family $\mathcal{H}$, results in an insecure 2-round protocol [Bar01,GK03].

guish between $\mathsf{iO}(s)$ and $\mathsf{iO}(s\{\alpha\}, \alpha, f_s(\alpha))$,[11] which together with the security of the puncturable PRF, implies that one cannot distinguish between $\mathsf{iO}(s)$ and $\mathsf{iO}(s\{\alpha\}, \alpha, u)$ for a truly random output u. Thus, we managed to use iO, together with the puncturing technique, to generate a circuit for computing f_s that hides the value of $f_s(\alpha)$. We emphasize that this technique crucially relies on the fact that the punctured point α is independent of the seed s, and hence as a result $f_s(\alpha)$ is computationally random.

It is natural to try and use obfuscated puncturable PRFs to show security of the Fiat-Shamir paradigm. Consider the following naive (and flawed) analysis, which loosely speaking proceeds in three steps: Suppose that there exists a poly-size cheating prover P^* that convinces the verifier to accept $x \notin L$. Recall that we denote transcripts by (α, β, γ). The (statistical) soundness of Π implies that for every α, for most of the verifier's messages β, there does not exist γ that makes the verifier accept. For any function f consider the (evasive) relation $R = \{(\alpha, \beta) : \exists\gamma \text{ s.t. } V(x, \alpha, \beta, \gamma) = 1\}$. Suppose that the cheating prover P^*, given $\mathsf{iO}(s)$, outputs α such that $(\alpha, f_s(\alpha)) \in R$, with non-negligible probability.

1. Puncture the PRF at a random point α^* s.t. $\alpha^* \in \mathsf{Bad}_s$, and send the obfuscation of $\mathsf{iO}(s\{\alpha^*\}, \alpha^*, f_s(\alpha^*))$ to the cheating prover P^*. Note that this does not change the functionality.
   Therefore, we can use the (sub-exponential) security of iO to argue that the cheating prover P^* cannot tell where we punctured the PRF, and still succeeds with non-negligible probability. In particular, taking M to be the expected number of α's such that $(\alpha, f_s(\alpha)) \in R$, we have that P^* outputs α^* with probability $\approx 1/M$ (up to $\mathrm{poly}(n)$ factors).[12]
2. Next, we want to use the (sub-exponential) security of the puncturable PRF to argue that the cheating prover P^* cannot distinguish between $(s\{\alpha^*\}, \alpha^*, f_s(\alpha^*))$ and $(s\{\alpha^*\}, \alpha^*, \beta^*)$ where (α^*, β^*) is random in R. Thus, given $\mathsf{iO}(s\{\alpha^*\}, \alpha^*, \beta^*)$ the cheating prover P^* still outputs α^* with probability $\approx 1/M$ (up to $\mathrm{poly}(n)$ factors).
3. In the final step, we argue that α^* is close to uniform (for an appropriate modification of the original protocol) and independent of s. Thus, given $\mathsf{iO}(s\{\alpha^*\}, \alpha^*, \beta^*)$, the cheating prover P^* outputs α^* with probability $\approx 1/M$ (up to $\mathrm{poly}(n)$ factors), where α^* is close to truly random. We want to argue that this contradicts the (sub-exponential) security of iO.

Unfortunately, the argument sketched above is doubly-flawed. In particular, the arguments in Step (2) and Step (3) are simply false. In Step (2) we start with a distribution where f_s is punctured at a point α^* for which $(\alpha^*, f_s(\alpha^*))$ is not (computationally) random, and in fact *the choice of* α^* *depends on the seed* s. We want to argue that this is indistinguishable from the case where we pick

[11] We use $(s\{\alpha\}, \alpha, f_s(\alpha))$ to denote the circuit that on input α outputs the hardwired value $f_s(\alpha)$, and on any other input $x \neq \alpha$ computes $f_s(x)$ using the punctured seed $s\{\alpha\}$.

[12] We think of n as polynomially related to the security parameter, where 2^n is the domain size of f_s.

(α^*, β^*) randomly in R, and then puncture at α^*. It is not a-priori clear why the puncturable PRF or iO would guarantee this indistinguishability. Indeed, the functions generated by these two distributions can be distinguished with some advantage by simply counting the number of input-output pairs that are in R.

Nevertheless, in our analysis (see Lemma 1) we manage to argue that the cheating prover P^*, given $\mathsf{iO}(s\{\alpha^*\}, \alpha^*, \beta^*)$ where (α^*, β^*) is random in R, still outputs α^* with probability significantly higher than $1/2^n$ (i.e., significantly higher than guessing). Indeed, P^* still outputs α^* with probability $\approx 1/M$ (up to $\mathrm{poly}(n)$ factors).

We next move to the flaw in Step (3). The problem here is that puncturing at the point α^* *does not at all hide* α^*. It is also not clear whether the iO obfuscation of the punctured seed hides α^*.

Input-Hiding Obfuscation to the Rescue. We overcome this hurdle by using an exponentially secure input-hiding obfuscation to hide the punctured point.

Namely, we replace $\mathsf{iO}(s\{\alpha^*\}, \alpha^*, \beta^*)$ with $\mathsf{iO}(s, \mathsf{hideO}(\alpha^*, \beta^*))$, where hideO is an exponentially secure input hiding obfuscator, and where we did not change the functionality of the circuit; i.e. the circuit on input x first runs $\mathsf{hideO}(\alpha^*, \beta^*)$ to check if $x = \alpha^*$; if so it outputs β^* and otherwise it outputs $f_s(x)$. The security of iO implies that $P^*(\mathsf{iO}(s, \mathsf{hideO}(\alpha^*, \beta^*)))$ outputs α^* with probability $1/M$ (up to $\mathrm{poly}(n)$ factors).

It remains to note that s is independent of (α^*, β^*), and hence we conclude that there exists a poly-size adversary that given $\mathsf{hideO}(\alpha^*, \beta^*)$ outputs α^* with probability $1/M$ (up to $\mathrm{poly}(n)$ factors). In the last step we replace the distribution of (α^*, β^*) with a distribution where α^* is chosen uniformly at random from $\{0,1\}^n$ and β^* is chosen at random such that $(\alpha^*, \beta^*) \in R$ and prove that still there exists a poly-size adversary that given $\mathsf{hideO}(\alpha^*, \beta^*)$ (where (α^*, β^*) is according to the new distribution) outputs α^* with probability $1/M$ (up to $\mathrm{poly}(n)$ factors). This contradicts the exponential security of the input-hiding obfuscator hideO.

Remark 2. We note that the input-hiding obfuscator *was only used in the security analysis*. It plays no role in the construction itself. This is similar to some other recent uses of indistinguishability obfuscation in the literature.

We note that the idea of using input-hiding obfuscation to hide the punctured point, was also used in [BM14]. However, as opposed to this work, they relied on the obfuscation being secure against auxiliary inputs.

2 Preliminaries

2.1 Indistinguishability

Definition 1. *For any function $T : \mathbb{N} \to \mathbb{N}$ and for any function $\mu : \mathbb{N} \to [0,1]$, we say that $\mu = negl(T)$ if for every constant $c > 0$ there exists $K \in \mathbb{N}$ such that for every $k \geq K$,*

$$\mu(k) \leq T(k)^{-c}.$$

Definition 2. *Two distribution families* $\mathcal{X} = \{\mathcal{X}_\kappa\}_{\kappa\in\mathbb{N}}$ *and* $\mathcal{Y} = \{\mathcal{Y}_\kappa\}_{\kappa\in\mathbb{N}}$ *are said to be* T-indistinguishable *(denoted by* $\mathcal{X} \overset{T}{\approx} \mathcal{Y}$*) if for every circuit family* $D = \{D_\kappa\}_{\kappa\in\mathbb{N}}$ *of size poly*$(T(\kappa))$*,*

$$Adv_D^{\mathcal{X},\mathcal{Y}}(T) \overset{\text{def}}{=} |\Pr[D(x) = 1] - \Pr[D(y) = 1]| = negl(T(\kappa)),$$

where the probabilities are over $x \leftarrow \mathcal{X}_\kappa$ *and over* $y \leftarrow \mathcal{Y}_\kappa$*.*

2.2 Puncturable PRFs

Our construction uses a *puncturable* pseudo-random function (PRF) family [BW13,BGI14,KPTZ13,SW14] that is 2^n-secure (where n is the input length); see the definitions below.

Definition 3 (T-Secure PRF [GGM86]). *Let* $m = m(\kappa)$*,* $n = n(\kappa)$ *and* $k = k(\kappa)$ *be functions of the security parameter* κ*. A PRF family is an ensemble* $\mathcal{F} = \{\mathcal{F}_\kappa\}_{\kappa\in\mathbb{N}}$ *of function families, where* $\mathcal{F}_\kappa = \{f_s : \{0,1\}^n \to \{0,1\}^k\}_{s\in\{0,1\}^m}$*. The PRF* $\mathcal{F}$ *is* T*-secure, for* $T = T(\kappa)$*, if for every poly*(T)*-size (non-uniform) adversary Adv:*

$$\left|Adv^{f_s}(1^\kappa) - Adv^{f}(1^\kappa)\right| = negl(T(\kappa)),$$

where f_s *is a random function in* $\mathcal{F}_\kappa$*, generated using a uniformly random seed* $s \in \{0,1\}^{m(\kappa)}$*, and* f *is a truly random function with domain* $\{0,1\}^n$ *and range* $\{0,1\}^k$*.*

We use 2^n-secure PRF families in our construction (for $k = \text{poly}(n)$). We can construct such PRFs assuming subexponentially hard one-way functions by taking the seed length m to be a sufficiently large polynomial in n. Observe that, since the entire truth table of the function can be constructed in time $\text{poly}(n)\cdot 2^n$, we get that 2^n-security implies that the entire truth table of a PRF f_s is indistinguishable from a uniformly random truth table.[13]

Definition 4 (T-Secure Puncturable PRF [SW14]). *A* T*-secure family of PRFs (as in Definition 3) is* puncturable *if there exist* PPT *procedures* puncture *and* eval *such that*

1. *Puncturing a PRF key* $s \in \{0,1\}^m$ *at a point* $r \in \{0,1\}^n$ *gives a punctured key* $s\{r\}$ *that can still be used to evaluate the PRF at any point* $r' \neq r$

$$\forall r \in \{0,1\}^n, r' \neq r : \Pr_{s,s\{r\}\leftarrow\mathsf{puncture}(s,r)}[\mathsf{eval}(s\{r\}, r') = f_s(r')] = 1$$

[13] The fact that subexponential OWF yield PRFs for which distinguishing the entire truth table from a random truth table the truth table of a random function has been previously noted in the literature, most notably by Razborov and Rudich [RR97] in their work on natural proofs.

2. *For any fixed* $r \in \{0,1\}^n$*, given a punctured key* $s\{r\}$*, the value* $f_s(r)$ *is pseudorandom:*

$$(s\{r\}, r, f_s(r)) \overset{T(\kappa)}{\approx} (s\{r\}, r, u),$$

where $s\{r\}$ *is obtained by puncturing a random seed* $s \in \{0,1\}^{m(\kappa)}$ *at the point* r*, and* u *is uniformly random in* $\{0,1\}^k$*.*

We note that the GGM-based construction of PRFs gives a construction of 2^n-secure puncturable PRFs from any subexponentially hard one-way function [GGM86, HILL99].

2.3 Indistinguishability Obfuscation

Our construction uses an indistinguishability obfuscator iO with 2^{-n} security. A candidate construction was first given in the work of Garg *et al.* [GGH+13].

Definition 5 (T-secure Indistinguishability Obfuscator [BGI+12]).
Let $T : \mathbb{N} \to \mathbb{N}$ *be a function. Let* $\mathbb{C} = \{\mathbb{C}_n\}_{n\in\mathbb{N}}$ *be a family of polynomial-size circuits, where* $\mathbb{C}_n$ *is a set of boolean circuits operating on inputs of length* n*. Let* iO *be a* PPT *algorithm, which takes as input a circuit* $C \in \mathbb{C}_n$ *and a security parameter* $\kappa \in \mathbb{N}$*, and outputs a boolean circuit* $\mathsf{iO}(C)$ *(not necessarily in* $\mathbb{C}$*).*

iO *is a* T*-secure* indistinguishability obfuscator *for* $\mathbb{C}$ *if it satisfies the following properties:*

1. Preserving Functionality: *For every* $n, \kappa \in \mathbb{N}$*,* $C \in \mathbb{C}_n$*,* $x \in \{0,1\}^n$*:*

$$(\mathsf{iO}(C, 1^\kappa))(x) = C(x).$$

2. Indistinguishable Obfuscation: *For every two sequences of circuits* $\{C^1_n\}_{n\in\mathbb{N}}$ *and* $\{C^2_n\}_{n\in\mathbb{N}}$*, such that for every* $n \in \mathbb{N}$*,* $|C^1_n| = |C^2_n|$*,* $C^1_n \equiv C^2_n$*, and* $C^1_n, C^2_n \in \mathbb{C}_n$*, and for every polynomially-bounded function* $m : \mathbb{N} \to \mathbb{N}$ *it holds that:*

$$\left(1^\kappa, \mathsf{iO}(C^1_{m(\kappa)}, 1^\kappa)\right) \overset{T(\kappa)}{\approx} \left(1^\kappa, \mathsf{iO}(C^2_{m(\kappa)}, 1^\kappa)\right).$$

2.4 Input-Hiding Obfuscation

An input-hiding obfuscator for a class of circuits $\mathbb{C}$, as defined by Barak *et al.* [BBC+14], has the security guarantee that given an obfuscation of a randomly drawn circuit in the family $\mathbb{C}$, it is hard for an adversary to find an accepting input. In our work, we consider input-hiding obfuscation for the class of multi-bit point functions. A multi-bit point function $I_{x,y}$ is defined by an input $x \in \{0,1\}^n$, and an output $y \in \{0,1\}^k$. $I_{x,y}$ outputs y on input x, and 0 on all other inputs. Informally, we assume that given the obfuscation of $I_{x,y}$ for a uniformly random x and an arbitrary y, it is hard for an adversary to recover x.

Definition 6 (T-secure Input-Hiding Obfuscator [BBC+14]). *Let $T : \mathbb{N} \to \mathbb{N}$ be a function, and let $\mathbb{C} = \{\mathbb{C}_n\}_{n\in\mathbb{N}}$ be a family of poly-size circuits, where $\mathbb{C}_n$ is a set of boolean circuits operating on inputs of length n. A* PPT *obfuscator* hideO *is a T-secure input-hiding obfuscator for $\mathbb{C}$, if it satisfies the preserving functionality requirement of Definition 5, as well as the following security requirement. For every poly-size (non-uniform) adversary Adv and all sufficiently large n,*

$$\Pr_{C\leftarrow\mathbb{C}_n,\mathsf{hideO}}[C(Adv(\mathsf{hideO}(C))) \neq 0] \leq T^{-1}(n).$$

We emphasize that (unlike other notions of T-security used in this work), we only allow the adversary for a T-secure input hiding obfuscation to run in polynomial time. Nevertheless, depending on the function T, the definition of T-secure input hiding is quite strong. In particular, for the typical case of proof-systems with soundness 2^{n^ϵ} (where $\epsilon > 0$ is a constant) we will assume input-hiding obfuscation for $T = 2^{n-n^\epsilon}$, which means that a polynomial-time adversary can only do sub-exponentially better than the trivial attack that picks random inputs until it finds an accepting input (this attack succeeds with probability $\text{poly}(n)/2^n$). This is also why we do not separate the security parameter from the input length (the adversary can always succeed with probability 2^{-n}, assuming there exists an accepting input).

We assume input-hiding obfuscation for the class of multi-bit point functions (see above), where the point x is drawn uniformly at random, and the output y is arbitrary. In particular, we do not assume that the collection $\mathbb{C}$ of pairs (x, y) can be sampled efficiently, only that its marginal distribution on x is uniform.

Assumption 3 (T-secure Input-Hiding for Multi-Bit Point Functions). *Let $T, k : \mathbb{N} \to \mathbb{N}$ be functions. An obfuscator* hideO *is a T-secure input-hiding obfuscator for (n, k)-multi-bit point functions if for every collection $\mathbb{C}$ as below,* hideO *is a T-secure input-hiding obfuscator for $\mathbb{C}$. In the collection $\mathbb{C}$, for every $n \in \mathbb{N}$, every function $I_{x,y} \in \mathbb{C}_n$ has $x \in \{0,1\}^n, y \in \{0,1\}^{k(n)}$, and the marginal distribution of a random draw from $\mathbb{C}_n$ on x is uniform.*

The assumption is strong in that we do not assume that a random function in $\mathbb{C}$ can be sampled efficiently, or that the output y is an efficient function of the input x. This assumption was studied in [CD08, BC14]. A candidate construction (in the standard model) was provided in [CD08]. Loosely speaking, their construction is an extension of the point function obfuscation of Canetti [Can97], where the obfuscation of $I_{x,y}$ consists of a pair of the form (r, r^x), together with k pairs of the form $(r_i, r_i^{\alpha_i})$ where $\alpha_i = x$ if $y_i = 1$ and is uniformly random otherwise. It was proved in [BC14] that this construction is secure in the generic group model, where the inversion probability is at most $\text{poly}(n) \cdot 2^{-n}$.

2.5 Interactive Proofs and Arguments

An interactive proof, as introduced by Goldwasser, Micali and Rackoff [GMR89], is a protocol between two parties, a computationally unbounded prover and a

polynomial-time verifier. Both parties have access to an input x and the prover tries to convince the verifier that $x \in L$. Formally an interactive proof is defined as follows:

Definition 7 (Interactive Proof [GMR89]). *An r-message interactive proof for the language L is an r-message protocol between the verifier V, which is polynomial-time, and a prover P, which is computationally unbounded. We require that the following two conditions hold:*

- **Completeness**: *For every $x \in L$, if V interacts with P on common input x, then V accepts with probability at least $2/3$.*
- **Soundness**: *For every $x \notin L$ and every (computationally unbounded) cheating prover strategy $\tilde{P}$, the verifier V accepts when interacting with $\tilde{P}$ with probability at most $1/3$.*

We say that an interactive-proof is public-coin if all messages sent from V to P consist of fresh random coins tosses. Also, recall that the constants $1/3$ and $2/3$ are arbitrary and can be amplified by (e.g., parallel) repetition.

Interactive Arguments. An interactive argument is defined similarly to an interactive proof except that the soundness condition is only required to hold for cheating provers that run in polynomial time. We also require that the honest prover run in polynomial-time, given the witness as an auxiliary input.

Definition 8 (Interactive Argument). *An r-message argument for the language $L \in \mathsf{NP}$ is an r-message protocol between a verifier V and a prover P, both of which are polynomial-time algorithms. We require that the following two conditions hold:*

- **Completeness**: *There exists a negligible function $negl$ such that for every $x \in L$, if V interacts with P on common input x, where P is given in addition an NP witness w for $x \in L$, then V accepts with probability at least $1 - negl(|x|)$.*
- **Soundness**: *For every polynomial-size cheating prover strategy $\tilde{P}$ and for every $x \notin L$, the verifier V accepts when interacting with $\tilde{P}$ on common input x, with probability at most $negl(|x|)$.*

We remark that in contrast to Definition 7, here we require negligible completeness and soundness errors. This is because parallel repetition does not necessarily decrease the soundness error for interactive arguments [BIN97]. We further remark that it is common to add a security parameter to the definition of argument systems so as to allow obtaining strong security guarantees even for short inputs. For simplicity of notations however we refrain from introducing a security parameter and note that better security guarantees for short inputs can be simply obtained by padding the input.

2.6 The Fiat-Shamir Paradigm

In this section, we recall the Fiat-Shamir paradigm. For the sake of simplicity of notation, we describe this paradigm when applied to 3-round (as opposed to arbitrary constant round) public-coin protocols. Let $\Pi = (P, V)$ be a 3-round public-coin proof system for an NP language L. We denote its transcripts by (α, β, γ), where β are the messages sent by the verifier, and α, γ are the messages sent by the prover. We denote by n the length of α (i.e., $\alpha \in \{0,1\}^n$), and we denote by k the length of β (i.e., $\beta \in \{0,1\}^k$). We assume that $k \leq \text{poly}(n)$ (since otherwise we can just pad).

Let $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ be an ensemble of hash functions, such that for every $n \in \mathbb{N}$ and for every $h \in \mathcal{H}_n$,

$$h : \{0,1\}^n \to \{0,1\}^k.$$

We define Π^{FS}, with respect to the hash family $\mathcal{H}$ to be the 2-round protocol obtained by applying the Fiat-Shamir transformation to Π using $\mathcal{H}$. A formal presentation of the "collapsed" protocol $\Pi^{\mathsf{FS}} = (P^{\mathsf{FS}}, V^{\mathsf{FS}})$ is in Fig. 1.

Remark 3. We emphasize that our main result is that the Fiat-Shamir paradigm *in its original formulation* (as presented in Fig. 1) is secure when applied to interactive proofs and when using a *particular hash function* (based on the assumption mentioned above).

Protocol $\Pi^{\mathsf{FS}}(1^n, x)$ for an NP Language L

Prover's Input: Statement x and a witness w for $x \in L$.
Verifier's Input: Statement x.
$V^{\mathsf{FS}} \to P^{\mathsf{FS}}$: The verifier V^{FS} chooses a random $h \leftarrow \mathcal{H}_n$, and sends h to the prover P^{FS}.
$P^{\mathsf{FS}} \to V^{\mathsf{FS}}$: The prover P^{FS} simulates an execution with the prover P of Π in the following way:
- Choose a random tape for P and continue the emulation of (P, V) by running P. Let $\alpha \in \{0,1\}^n$ be the first message sent by P in Π.
- Compute $h(\alpha) = \beta$.
- Continue the emulation of P assuming P received β as the second message from V^{FS}. Let γ be the third message sent by P.

Send (α, β, γ) to the verifier V^{FS}.
Verification: The verifier V^{FS} accepts if and only if:
- $h(\alpha) = \beta$.
- V accepts the transcript (α, β, γ).

Fig. 1. Collapsing a 3-round Protocol $\Pi = (P, V)$ into a 2-round Protocol $\Pi^{\mathsf{FS}} = (P^{\mathsf{FS}}, V^{\mathsf{FS}})$ using $\mathcal{H}$

3 Security of Fiat-Shamir for 3-Message Proofs

We show an instantiation of the Fiat-Shamir paradigm that is sound when it is applied to interactive proofs (as opposed to arguments). Taking n to be a bound on the message lengths of the prover in Π, our instantiation assumes the existence of a 2^n-secure indistinguishability obfuscation scheme iO, a 2^n-secure

puncturable PRF family $\mathcal{F}$, and a 2^n-secure input-hiding obfuscation for the class of multi-bit point functions $\mathcal{I}_{n,k}$.

For clarity of exposition, we first show that our instantiation is secure for 3-round public-coin interactive proofs. This is the regime for which the Fiat-Shamir paradigm was originally suggested. We then build on the proof for the 3-message case (or rather the 4-message case, see below), and prove security for any constant number of rounds.

Theorem 4 (Fiat-Shamir for 3-message Proofs). *Let Π be a public-coin 3-message interactive proof system with negligible soundness error. Let n be an upper bound on the input length and the length of the prover's messages and let $k \leq poly(n)$ be an upper bound on the length of the verifier's messages.*

Assume the existence of a 2^n-secure puncturable PRF family $\mathcal{F}$, the existence of a 2^n-secure Indistinguishability Obfuscation iO, and the existence of a secure input-hiding obfuscation for the class of multi-bit point functions $\{\mathcal{I}_{n,k}\}$ with security $T = 2^n \cdot negl(n)$.

Then, the resulting 2-round argument Π^{FS}, obtained by applying the Fiat-Shamir paradigm (see Fig. 1) to Π with the function family $\mathsf{iO}(\mathcal{F})$, is secure.

(Recall that we defined $\mathsf{iO}(\mathcal{F})$ as the iO obfuscation of a program that computes the PRF, using a hardwired random seed.)

In Sect. 4 we prove the security of the Fiat-Shamir paradigm when applied to any constant round interactive proof. To prove the general (constant round) case, we need to rely on a more general (and more technical) variation of Theorem 4. First, we rely on the security of the Fiat-Shamir paradigm for any 4-round interactive proof Π where the first message is sent by the verifier. In the transformed protocol Π^{FS}, the first message of the verifier consists of the first message as in Π, along with a Fiat-Shamir hash function, which will be applied to the prover's first message. In addition, in the generalized theorem we allow the verifier in the original protocol Π to run in time $2^{O(n)}$.

We state the generalized theorem below.

Theorem 5 (Theorem 4, more General Statement). *Let Π be a 4-message public-coin interactive proof system, where the first message is sent by the verifier. Let n be an upper bound on the input length*[14] *and the lengths of the prover's messages, let $k \leq poly(n)$ be a bound on the verifier's messages, let $\mu(n) = negl(n)$ be the soundness error*[15] *error, and assume that the verifier runs in time at most $2^{O(n)}$.*

[14] We remark that the reason we bound the input length is solely because we use a simplified definition of argument system that does not have a security parameter, and we are aiming for argument systems with soundness that is negligible in the *input length*.

[15] Since parallel repetition decreases the soundness error of interactive *proofs* at an exponential rate, we may assume without loss of generality that the soundness error is negligible in n.

Assume the existence of a 2^n-secure puncturable PRF family $\mathcal{F}$, the existence of a 2^n-secure Indistinguishability Obfuscation iO, and the existence of a input-hiding obfuscation for the class of multi-bit point functions $\{\mathcal{I}_{n,k}\}$ that is T-secure for every $T = 2^n \cdot \mu/\nu$, where ν is any non-negligible function.

Then the resulting 2-round argument Π^{FS}, obtained by applying the Fiat-Shamir paradigm[16] to Π with the function family $\mathsf{iO}(\mathcal{F})$, is secure.

We remark that $\mu \cdot 2^n \cdot \mathrm{poly}(n)$ is a shorthand for a function T such that for every $c > 0$ and all sufficiently large $n \in \mathbb{N}$ it holds that $T(n) \geq \mu(n) \cdot 2^n \cdot n^c$.

Proof (Proof of Theorem 5). Fix any 4-round interactive proof $\Pi = (P, V)$ as claimed in the theorem statement. Let $\mu = \mathrm{negl}(n)$ be the soundness error of Π.

Suppose for the sake of contradiction that there exists a poly-size cheating prover P^* who breaks the soundness of the protocol Π^{FS} with respect to some $x^* \notin L$ with non-negligible probability ν. We will use P^* to eventually break the security of the input-hiding obfuscation, while using along the way the soundness of Π as well as the security of the PRF $\mathcal{F}$ and Indistinguishability Obfuscator iO.

There must exist a choice for the verifier's first message τ in Π, such that the following two conditions hold: (i) Even conditioned on the first part of the first message in Π^{FS} being τ, the cheating prover P^* still breaks the soundness of the protocol Π^{FS} on x^* with probability at least $(\nu/2)$, and (ii) even conditioned on the first message in Π being τ, the original protocol Π still has soundness error at most $(2\mu/\nu)$. Such a τ must exist because at least a $(\nu/2)$-fraction of the messages must satisfy condition (i) (otherwise P^* cannot break Π^{FS} with total probability ν), and the fraction that do not satisfy condition (ii) must be smaller than $(\nu/2)$ (otherwise the soundness of Π is smaller than μ).

Fix the verifier's first message to always be τ (both in the original and in the transformed protocols). We have that:

$$\Pr_{s,\mathsf{iO}}\left[P^*(\tau, \mathsf{iO}(s)) = (\alpha, \gamma) \text{ s.t. } V(x^*, \tau, \alpha, f_s(\alpha), \gamma) = 1\right] \geq \nu/2, \tag{3.1}$$

where $\mathsf{iO}(s)$ refers to the iO obfuscation of a random function f_s from the family $\mathcal{F}$.

The relaxed verifier and its properties. To obtain a contradiction, we analyze a relaxed verifier V' (which is only used in the security analysis). The relaxed verifier accepts a transcript (α, β, γ) if the original verifier V would accept, or if the first $\lceil \log(\nu/(2\mu)) \rceil$ bits of β are all 0 (where recall that μ is the soundness

[16] For 4-message proofs, the same paradigm as in Fig. 1 is used, except that the verifier also sends its first message from the base proof-system (i.e., a random string) in the first round.

error of Π).[17] In particular, whenever V accepts, the relaxed verifier V' also accepts, and so:

$$\Pr_{s,\mathsf{iO}}\left[P^*(\tau,\mathsf{iO}(s)) = (\alpha,\gamma) \text{ s.t. } V'(x^*,\tau,\alpha,f_s(\alpha),\gamma) = 1\right] \geq \nu/2. \tag{3.2}$$

We take μ' to be the soundness of the interactive proof (P,V') (after τ is fixed), which runs the relaxed verifier. Observe that by a union bound

$$\mu' \leq (2\mu/\nu) + 2^{-\lceil \log(\nu/(2\mu))\rceil} \leq 4\mu/\nu,$$

(in particular if μ is negligible, then so is μ').

We define:

$$\mathsf{ACC} = \left\{(\alpha,\beta) : \exists\gamma \text{ s.t. } V'(x^*,\tau,\alpha,\beta,\gamma) = 1\right\}$$

Observe that membership in ACC can be computed in time $2^n \cdot \mathrm{poly}(n) = 2^{O(n)}$ by enumerating over all γ's and running V'. Equation (3.2) implies that there exists a poly-size adversary $\mathcal{A}$ (that just outputs the first part of P^*'s output) such that:

$$\Pr_{s,\mathsf{iO}}\left[\mathcal{A}(\mathsf{iO}(s)) \text{ outputs some } \alpha \text{ s.t. } \big(\alpha, f_s(\alpha)\big) \in \mathsf{ACC}\right] \geq \nu/2. \tag{3.3}$$

Using Eq. (3.3) we prove our main lemma.

Lemma 1.

$$\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\mathcal{A}\big(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*)\big) = \alpha^* \;\middle|\; \big(\alpha^*,u^*\big) \in \mathsf{ACC}\right] \geq 2^{-n+2}\cdot\nu/\mu'$$

where α^* and u^* are uniformly distributed (in $\{0,1\}^n$ and $\{0,1\}^k$, respectively) and $\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*)$ refers to an iO obfuscation of the program that contains the seed s punctured at the point α^*, and on input α first checks if $\alpha = \alpha^*$ and if so outputs u^* and otherwise outputs $f_s(\alpha)$.

Proof. We prove the lemma by analyzing the probability that the event

$$\Big(\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*)) = \alpha^*\Big) \;\wedge\; \Big(\big(\alpha^*,u^*\big) \in \mathsf{ACC}\Big)$$

occurs.

By the exponential hardness of the puncturable PRF, and the fact that membership in ACC is computable in $2^{O(n)}$ time, we have that

[17] In the original protocol Π, it may be the case that different messages α sent by the prover can lead the verifier to accept with different probabilities. E.g., some specific α's may lead the verifier to accept with probability μ and others with probability 0. This presents a technical difficulty later in the proof and so we construct the relaxed verifier V' so that every string α leads it to accept with roughly the same probability (up to a small multiplicative constant) without increasing the soundness error by too much.

$$\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\begin{array}{c}\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*))=\alpha^*\\ \wedge\\ (\alpha^*,u^*)\in\mathsf{ACC}\end{array}\right]\geq\Pr_{s,\alpha^*,\mathsf{iO}}\left[\begin{array}{c}\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,f_s(\alpha^*)))=\alpha^*\\ \wedge\\ (\alpha^*,f_s(\alpha^*))\in\mathsf{ACC}\end{array}\right]-2^{-2n}. \tag{3.4}$$

Further applying the exponential hardness of the iO scheme (and the fact that membership in ACC can be decided in $2^{O(n)}$ time), we get that:

$$\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\begin{array}{c}\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*))=\alpha^*\\ \wedge\\ (\alpha^*,u^*)\in\mathsf{ACC}\end{array}\right]\geq\Pr_{s,\alpha^*,\mathsf{iO}}\left[\begin{array}{c}\mathcal{A}(\mathsf{iO}(s))=\alpha^*\\ \wedge\\ (\alpha^*,f_s(\alpha^*))\in\mathsf{ACC}\end{array}\right]-2\cdot 2^{-2n}. \tag{3.5}$$

Using elementary probability theory, we have that:

$$\begin{aligned}\Pr_{s,\alpha^*,\mathsf{iO}}\left[\begin{array}{c}\mathcal{A}(\mathsf{iO}(s))=\alpha^*\\ \wedge\\ (\alpha^*,f_s(\alpha^*))\in\mathsf{ACC}\end{array}\right]&=\Pr_{s,\alpha^*,\mathsf{iO}}\left[\bigcup_\alpha\left((\mathcal{A}(\mathsf{iO}(s))=\alpha^*)\wedge((\alpha^*,f_s(\alpha^*))\in\mathsf{ACC})\wedge(\alpha^*=\alpha)\right)\right]\\ &=\sum_\alpha\Pr_{s,\alpha^*,\mathsf{iO}}\left[((\mathcal{A}(\mathsf{iO}(s))=\alpha)\wedge((\alpha,f_s(\alpha))\in\mathsf{ACC})\wedge(\alpha^*=\alpha))\right]\\ &=2^{-n}\sum_\alpha\Pr_{s,\mathsf{iO}}\left[(\mathcal{A}(\mathsf{iO}(s))=\alpha)\wedge((\alpha,f_s(\alpha))\in\mathsf{ACC})\right]\\ &=2^{-n}\Pr_{s,\mathsf{iO}}\left[\mathcal{A}(\mathsf{iO}(s))\text{ outputs some }\alpha\text{ s.t. }(\alpha,f_s(\alpha))\in\mathsf{ACC}\right]\\ &\geq 2^{-n}\cdot\nu/2\end{aligned}$$

where the last inequality is by Eq. (3.3). Thus, we have that:

$$\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\begin{array}{cc}\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*))=\alpha^* & \\ \wedge & s\\ (\alpha^*,u^*)\in\mathsf{ACC} & \end{array}\right]\geq\frac{1}{4}\cdot 2^{-n}\cdot\nu.$$

By the soundness of the underlying proof-system, it holds thats $\Pr_{\alpha^*,u^*}[(\alpha^*,u^*)\in\mathsf{ACC}]\leq\mu'$ (since otherwise a cheating prover could violate soundness by just sending a random α^*).[18]

Let $\zeta=\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*))=\alpha^*\;\middle|\;(\alpha^*,u^*)\in\mathsf{ACC}\right]$.
Then, by definition of conditional probability we have that

$$\zeta=\frac{\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\begin{array}{c}\mathcal{A}(\mathsf{iO}(s\{\alpha^*\},\alpha^*,u^*))=\alpha^*\\ \wedge\\ (\alpha^*,u^*)\in\mathsf{ACC}\end{array}\right]}{\Pr_{\alpha^*,u^*}[(\alpha^*,u^*)\in\mathsf{ACC}]}\geq\frac{1}{4}\cdot 2^{-n}\cdot\nu/\mu',$$

and the lemma follows.

[18] It may at first seem odd that we only use the soundness of the underlying proof-system with respect to a cheating prover that just sends a random message α^*. Recall however that here we consider the *relaxed* verifier who, by design, has a (roughly) similar acceptance probability given any string α.

We are now ready to use (and break) our input-hiding obfuscator hideO. Lemma 1, together with the 2^n-security of the iO implies that

$$\Pr_{s,\alpha^*,u^*,\mathsf{iO}}\left[\mathcal{A}(\mathsf{iO}(s,\mathsf{hideO}(\alpha^*,u^*))) = \alpha^* \;\middle|\; (\alpha^*,u^*) \in \mathsf{ACC}\right] \geq \frac{1}{4}\cdot 2^{-n}\cdot \nu/\mu' - 2^{-n} \geq \frac{1}{8}\cdot 2^{-n}\cdot \nu/\mu', \tag{3.6}$$

where α^* and u^* are uniformly distributed and $\mathsf{iO}(s,\mathsf{hideO}(\alpha^*,u^*))$ refers to the iO obfuscation of the program that contains a seed s for a PRF (in its entirety), and the input-hiding obfuscation $\mathsf{hideO}(\alpha^*,u^*)$ of a multi-bit point function that on input α^* outputs u^*. The program uses the input-hiding obfuscation to check if its input equals α^*, and if so outputs the same value as $\mathsf{hideO}(\alpha^*,u^*)$. Otherwise the program behaves like the PRF.

Equation (3.6) is almost what we want. Namely, an adversary that given access to $\mathsf{hideO}(\alpha^*,u^*)$ produces α^* with probability $\omega(\mathrm{poly}(n)/2^n)$ (since ν is inverse polynomial and μ is a negligible function). The only remaining problem is that the distribution of (α^*,u^*) is not quite what we need. More specifically, in Eq. (3.6) (α^*,u^*) are distributed uniformly conditioned on $(\alpha^*,u^*)\in\mathsf{ACC}$, whereas we need for the marginal distribution of α to be uniform in order to break the hideO obfuscation. Using the properties of the *relaxed* verifier, we show that these two distributions are actually closely related.

We define the following two distributions. The distribution $\mathcal{T}_1$ is obtained by jointly picking a pair (α,β) uniformly from ACC (this is the distribution from which (α^*,u^*) are sampled from in Eq. (3.6)). $\mathcal{T}_2$ is the distribution obtained by picking a uniformly random $\alpha\in\{0,1\}^n$ and then a random β conditioned on $(\alpha,\beta)\in\mathsf{ACC}$ (i.e. the marginal distribution on α is uniform). For $\alpha^*\in\{0,1\}^n$, $\beta^*\in\{0,1\}^k$, we use $\mathcal{T}_1[\alpha^*,\beta^*]$ and $\mathcal{T}_2[\alpha^*,\beta^*]$ to denote the probability of the pair (α^*,β^*) by $\mathcal{T}_1$ and by $\mathcal{T}_2$ (respectively).

Proposition 1. *For any $\alpha^*\in\{0,1\}^n$ and $\beta^*\in\{0,1\}^k$:*

$$\mathcal{T}_2[\alpha^*,\beta^*] \geq \frac{1}{4}\mathcal{T}_1[\alpha^*,\beta^*]$$

Proof. For every α^* denote by:

$$S_{\alpha^*} = \left\{\beta^*\in\{0,1\}^k : (\alpha^*,\beta^*)\in\mathsf{ACC}\right\}.$$

By construction of the relaxed verifier V', we know that for every $\alpha\in\{0,1\}^n$ it holds that

$$\frac{\mu}{\nu} \leq \frac{|S_\alpha|}{2^k} \leq \frac{4\mu}{\nu}.$$

In particular, for any $\alpha,\alpha^*\in\{0,1\}^n$:

$$|S_\alpha| \geq \frac{1}{4}|S_{\alpha^*}|.$$

Now we have that:

$$\mathcal{T}_1[\alpha^*, \beta^*] = \frac{1}{\sum_{\alpha \in \{0,1\}^n} |S_\alpha|} \leq \frac{4}{\sum_{\alpha \in \{0,1\}^n} |S_{\alpha^*}|} = \frac{4}{2^n \cdot |S_{\alpha^*}|} = 4 \cdot \mathcal{T}_2[\alpha^*, \beta^*] \tag{3.7}$$

In particular, drawing by $\mathcal{T}_2$ rather than $\mathcal{T}_1$ can only decrease the success probability of $\mathcal{A}$ by a multiplicative factor of 4. Moreover, when drawing by $\mathcal{T}_2$, the marginal distribution on α^* *is* uniform. Thus Proposition 1 and Eq. (3.6) imply that there exists a poly-size adversary $\mathcal{A}$, such that

$$\Pr_{(\alpha^*, u^*) \leftarrow \mathcal{T}_2, \mathsf{hideO}}[\mathcal{A}(\mathsf{hideO}(\alpha^*, u^*)) = \alpha^*] \geq \frac{1}{32} \cdot \frac{\nu}{\mu' \cdot 2^n}$$

where α^* drawn by $\mathcal{T}_2$ *is* uniformly random. Since ν is a non-negligible function and $\mu' = O(\mu/\nu)$, this contradicts the security of the input-hiding obfuscation hideO.

4 Security of Fiat-Shamir for Multi-round Proofs

In this section we show a secure instantiation of the Fiat-Shamir methodology for transforming any constant-round interactive proof into a 2-round computationally-sound argument. We assume for the sake of simplicity, and without loss of generality, that the verifier always sends the first message, and thus consider interactive protocols with an even number of rounds. Namely, for any constant $c \geq 2$, we consider a $2c$-round interactive proof $\Pi = (P, V)$. We assume without loss of generality that all of the prover's messages are of the same length, and denote this length by n (i.e. $\forall i, \alpha_i \in \{0,1\}^n$). Similarly, we assume without loss of generality that all of the verifier's messages are of the same length, and denote this length by k (i.e. $\forall i, \beta_i \in \{0,1\}^k$). We assume without loss of generality that $k \leq n$. All these assumptions are only for the simplicity of notations, and can be easily achieved by padding.

For every $i \in [c-1]$, let $\{\mathcal{F}_n^{(i)}\}_{n \in \mathbb{N}}$ be an ensemble of hash functions, such that for every $n \in \mathbb{N}$ and for every $f^{(i)} \in \mathcal{F}_n$,

$$f^{(i)} : \{0,1\}^{i \cdot (n+k)} \to \{0,1\}^k.$$

We assume without loss of generality that there exists a polynomial p such that for every $i \in [c-1]$ and for every $n \in \mathbb{N}$,

$$\mathcal{F}_n^{(i)} = \{f_s^{(i)}\}_{s \in \{0,1\}^{p(n)}}.$$

We define Π^{FS} to be the 2-round protocol obtained by applying the multi-round Fiat-Shamir transformation to Π using $(\mathsf{iO}(f_{s_1}^{(1)}), \ldots, \mathsf{iO}(f_{s_{c-1}}^{(c-1)}))$, where $f_{s_i}^{(i)} \leftarrow \mathcal{F}_n^{(i)}$ for every $i \in [c-1]$. The security of Π^{FS} is shown in Theorem 6 below.

Theorem 6 (Fiat-Shamir Transform for Multi-Round Interactive Proofs). *Let $\mu : \mathbb{N} \to [0,1]$ be a function. Assume the existence of a 2^n-secure puncturable PRF family $\mathcal{F}$, assume the existence of a 2^n-secure Indistinguishability Obfuscation, and assume the existence of an input-hiding obfuscation for the class of multi-bit point functions $\{\mathcal{I}_{n,k}\}$ that is T-secure for any $T = 2^n \cdot \mu/\nu$, where ν is any non-negligible function.*

Then for any constant $c \in \mathbb{N}$ such that $c \geq 2$, and any $2c$-round interactive proof Π with soundness μ, the resulting 2-round argument Π^{FS}, obtained by applying the multi-round Fiat-Shamir transformation to Π with the function family $\mathsf{iO}(\mathcal{F})$, is secure.

Proof. The proof is by induction on $c \in \mathbb{N}$, for $c \geq 2$. The base case $c = 2$ follows immediately from Theorem 4. Suppose the theorem statement is true for $< c$ rounds, and we will prove that it is true for c rounds.

To this end, fix any $2c$-round interactive proof Π for proving membership in a language L. Suppose for the sake of contradiction that Π^{FS} is not secure. Namely, there exists a poly-size cheating prover P^* and there exists $x^* \notin L$ such that P^* succeeds in convincing the verifier of Π^{FS} that $x^* \in L$ with non-negligible probability. We assume without loss of generality that P^* is deterministic.

Consider the following protocol Ψ for proving membership in L, which consists of $2c-2$ rounds: In the first round the verifier chooses the first message that it would have sent in Π, which we denote by β_0. In addition, it chooses a random seed $s_1 \leftarrow \{0,1\}^{p(n)}$, and sends to the prover the pair $(\beta_0, \mathsf{iO}(f_{s_1}^{(1)}))$. Then, the prover chooses $(\alpha_1, \beta_1, \alpha_2)$ such that $\beta_1 = f_{s_1}^{(1)}(\alpha_1)$, and such that α_1 and α_2 are chosen as in Π. It sends $(\alpha_1, \beta_1, \alpha_2)$ to the verifier. Then the prover and verifier continue to execute the protocol Π interactively, conditioned on $(\beta_0, \alpha_1, \beta_1, \alpha_2)$. Finally, the verifier accepts if and only if the verifier of Π would have accepted the resulting transcript and $\beta_1 = f_{s_1}^{(1)}(\alpha_1)$.

Consider the protocol Ψ_{P^*}, in which we fix the first message from the prover in Ψ to be the message $(\alpha_1, \beta_1, \alpha_2)$ generated by P^* in Π^{FS}. If Ψ_{P^*} is a sound proof then, by our induction hypothesis $(\Psi_{P^*})^{\mathsf{FS}}$ is sound. However, note that P^* can be trivially converted into a cheating prover that breaks the soundness of $(\Psi_{P^*})^{\mathsf{FS}}$, contradicting our induction hypothesis that the Fiat-Shamir transformation is sound for interactive proofs with $2(c-1)$ rounds (with the function family $\mathsf{iO}(\mathcal{F})$). Thus, it must be the case that Ψ_{P^*} is not a sound proof. Namely, there exists a (possibly inefficient) cheating prover P^{**}, an element $x^* \notin L$, and a polynomial q, such that P^{**} convinces the verifier of Ψ_{P^*} to accept x^* with probability $\geq 1/q(\kappa)$ for infinitely many $\kappa \in \mathbb{N}$.

Consider the 4-round protocol Φ, which consists of the first 4 rounds of Π, denoted by $(\beta_0, \alpha_1, \beta_1, \alpha_2)$. Given a transcript $(\beta_0, \alpha_1, \beta_1, \alpha_2)$ the verifier of Φ accepts if and only if there exists a strategy of the (cheating) prover of Π that causes the verifier of Π to accept with probability $\geq 1/q(\kappa)$ conditioned on the first 4-rounds of Π being $(\beta_0, \alpha_1, \beta_1, \alpha_2)$. Note that the verifier of Φ runs in time $\mathrm{poly}(2^{c(n+k)}) = 2^{O(n)}$. The statistical soundness of Π implies that Φ is also statistically sound. Note however that Φ^{FS} is not computationally sound. To see this, consider a poly-size cheating prover for Φ^{FS} that sends the message

$(\alpha_1, \beta_1, \alpha_2)$ that P^* sends in Π. By the fact that Ψ_{P^*} is not sound (since P^{**} breaks its soundness), the verifier of Φ^{FS} will accept $x^* \notin \text{Ł}$. This is in contradiction to Theorem 5 (where we used the fact that Theorem 5 holds even for verifiers running in time $2^{O(n)}$).

Acknowledgments. We thank an anonymous reviewer for suggesting, and allowing us to use, a significant simplification to our original proof. We also thank the reviewers for their useful comments and especially for pointing out an error in a previous version of the proof of Theorem 6.

This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing, supported by the Simons Foundation and by the DIMACS/Simons Collaboration in Cryptography through NSF grant #CNS-1523467.

The third author was also partially supported by the grants: NSF MACS - CNS-1413920, DARPA IBM - W911NF-15-C-0236, SIMONS Investigator award Agreement Dated 6-5-12 and DARPA NJIT - W911NF-15-C-0226.

References

[AABN02] Abdalla, M., An, J.H., Bellare, M., Namprempre, C.: From identification to signatures via the fiat-shamir transform: minimizing assumptions for security and forward-security. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 418–433. Springer, Heidelberg (2002). doi:10.1007/3-540-46035-7_28

[Bar01] Barak, B.: How to go beyond the black-box simulation barrier. In: FOCS, pp. 106–115 (2001)

[BBC+14] Barak, B., Bitansky, N., Canetti, R., Kalai, Y.T., Paneth, O., Sahai, A.: Obfuscation for evasive functions. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 26–51. Springer, Heidelberg (2014). doi:10.1007/978-3-642-54242-8_2

[BC14] Bitansky, N., Canetti, R.: On strong simulation and composable point obfuscation. J. Cryptol. **27**(2), 317–357 (2014)

[BCC+14] Bitansky, N., Canetti, R., Cohn, H., Goldwasser, S., Kalai, Y.T., Paneth, O., Rosen, A.: The impossibility of obfuscation with auxiliary input or a universal simulator. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014. LNCS, vol. 8617, pp. 71–89. Springer, Heidelberg (2014). doi:10.1007/978-3-662-44381-1_5

[BDG+13] Bitansky, N., Dachman-Soled, D., Garg, S., Jain, A., Kalai, Y.T., López-Alt, A., Wichs, D.: Why "Fiat-Shamir for Proofs" lacks a proof. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785, pp. 182–201. Springer, Heidelberg (2013). doi:10.1007/978-3-642-36594-2_11

[BDNP08] Ben-David, A., Nisan, N., Pinkas, B.: Fairplaymp: a system for secure multi-party computation. In: ACM Conference on Computer and Communications Security, pp. 257–266 (2008)

[BGGL01] Barak, B., Goldreich, O., Goldwasser, S., Lindell, Y.: Resettably-sound zero-knowledge and its applications. In: 42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14–17 October 2001, Las Vegas, Nevada, USA, pp. 116–125 (2001)

[BGI+12] Barak, B., Goldreich, O., Impagliazzo, R., Rudich, S., Sahai, A., Vadhan, S.P., Yang, K.: On the (im)possibility of obfuscating programs. J. ACM **59**(2), 6 (2012)

[BGI14] Boyle, E., Goldwasser, S., Ivan, I.: Functional signatures and pseudorandom functions. In: Krawczyk, H. (ed.) PKC 2014. LNCS, vol. 8383, pp. 501–519. Springer, Heidelberg (2014). doi:10.1007/978-3-642-54631-0_29

[BGL+15] Bitansky, N., Garg, S., Lin, H., Pass, R., Telang, S.: Succinct randomized encodings and their applications. In: Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14–17, 2015, pp. 439–448 (2015)

[BIN97] Bellare, M., Impagliazzo, R., Naor, M.: Does parallel repetition lower the error in computationally sound protocols? In: 38th Annual Symposium on Foundations of Computer Science, FOCS 1997, Miami Beach, Florida, USA, October 19–22, 1997, pp. 374–383 (1997)

[BL04] Barak, B., Lindell, Y.: Strict polynomial-time in simulation and extraction. SIAM J. Comput. **33**(4), 738–818 (2004)

[Blu87] Blum, M.: How to prove a theorem so no one else can claim it. In: Proceedings of the International Congress of Mathematicians, pp. 1444–1451 (1987)

[BLV06] Barak, B., Lindell, Y., Vadhan, S.P.: Lower bounds for non-black-box zero knowledge. J. Comput. Syst. Sci. **72**(2), 321–391 (2006)

[BM14] Brzuska, C., Mittelbach, A.: Indistinguishability obfuscation versus multi-bit point obfuscation with auxiliary input. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014. LNCS, vol. 8874, pp. 142–161. Springer, Heidelberg (2014). doi:10.1007/978-3-662-45608-8_8

[BR93] Bellare, M., Rogaway, P.: Random oracles are practical: a paradigm for designing efficient protocols. In: ACM Conference on Computer and Communications Security, pp. 62–73 (1993)

[BS16] Bellare, M., Stepanovs, I.: Point-function obfuscation: a framework and generic constructions. In: Kushilevitz, E., Malkin, T. (eds.) TCC 2016. LNCS, vol. 9563, pp. 565–594. Springer, Heidelberg (2016). doi:10.1007/978-3-662-49099-0_21

[BW13] Boneh, D., Waters, B.: Constrained pseudorandom functions and their applications. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013. LNCS, vol. 8270, pp. 280–300. Springer, Heidelberg (2013). doi:10.1007/978-3-642-42045-0_15

[Can97] Canetti, R.: Towards realizing random oracles: hash functions that hide all partial information. In: Kaliski, B.S. (ed.) CRYPTO 1997. LNCS, vol. 1294, pp. 455–469. Springer, Heidelberg (1997). doi:10.1007/BFb0052255

[CCR15] Canetti, R., Chen, Y., Reyzin, L.: On the correlation intractability of obfuscated pseudorandom functions. IACR Cryptology ePrint Archive, 2015:334 (2015)

[CD08] Canetti, R., Dakdouk, R.R.: Obfuscating point functions with multibit output. In: Smart, N. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 489–508. Springer, Heidelberg (2008). doi:10.1007/978-3-540-78967-3_28

[CGH04] Canetti, R., Goldreich, O., Halevi, S.: The random oracle methodology, revisited. J. ACM **51**(4), 557–594 (2004)

[CKPR02] Canetti, R., Kilian, J., Petrank, E., Rosen, A.: Black-box concurrent zero-knowledge requires (almost) logarithmically many rounds. SIAM J. Comput. **32**(1), 1–47 (2002)

[DNRS99] Dwork, C., Naor, M., Reingold, O., Stockmeyer, L.J.: Magic functions. In: FOCS, pp. 523–534 (1999)

[DRV12] Dodis, Y., Ristenpart, T., Vadhan, S.: Randomness condensers for efficiently samplable, seed-dependent sources. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 618–635. Springer, Heidelberg (2012). doi:10.1007/978-3-642-28914-9_35

[FS86] Fiat, A., Shamir, A.: How to prove yourself: practical solutions to identification and signature problems. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 186–194. Springer, Heidelberg (1987). doi:10.1007/3-540-47721-7_12

[GGH+13] Garg, S., Gentry, C., Halevi, S., Raykova, M., Sahai, A., Waters, B.: Candidate indistinguishability obfuscation and functional encryption for all circuits. In: 54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26–29 October, 2013, Berkeley, CA, USA, pp. 40–49 (2013)

[GGM86] Goldreich, O., Goldwasser, S., Micali, S.: How to construct random functions. J. ACM **33**(4), 792–807 (1986)

[GK96] Goldreich, O., Krawczyk, H.: On the composition of zero-knowledge proof systems. SIAM J. Comput. **25**(1), 169–192 (1996)

[GK03] Goldwasser, S., Kalai, Y.T.: On the (in)security of the fiat-shamir paradigm. In: FOCS, pp. 102–113 (2003)

[GK05] Goldwasser, S., Kalai, Y.T.: On the impossibility of obfuscation with auxiliary input. In: FOCS, pp. 553–562 (2005)

[GK16] Goldwasser, S., Tauman Kalai, Y.: Cryptographic assumptions: a position paper. In: Kushilevitz, E., Malkin, T. (eds.) TCC 2016. LNCS, vol. 9562, pp. 505–522. Springer, Heidelberg (2016). doi:10.1007/978-3-662-49096-9_21

[GLSW14] Gentry, C., Lewko, A.B., Sahai, A., Waters, B.: Indistinguishability obfuscation from the multilinear subgroup elimination assumption. IACR Cryptology ePrint Archive 2014:309 (2014)

[GMR89] Goldwasser, S., Micali, S., Rackoff, C.: The knowledge complexity of interactive proof systems. SIAM J. Comput. **18**(1), 186–208 (1989)

[GO94] Goldreich, O., Oren, Y.: Definitions and properties of zero-knowledge proof systems. J. Cryptol. **7**(1), 1–32 (1994)

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

[HILL99] Håstad, J., Impagliazzo, R., Levin, L.A., Luby, M.: A pseudorandom generator from any one-way function. SIAM J. Comput. **28**(4), 1364–1396 (1999)

[HT98] Hada, S., Tanaka, T.: On the existence of 3-round zero-knowledge protocols. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 408–423. Springer, Heidelberg (1998). doi:10.1007/BFb0055744

[KPR98] Kilian, J., Petrank, E., Rackoff, C.: Lower bounds for zero knowledge on the internet. In: 39th Annual Symposium on Foundations of Computer Science, FOCS 1998, November 8–11, 1998, Palo Alto, California, USA, pp. 484–492 (1998)

[KPTZ13] Kiayias, A., Papadopoulos, S., Triandopoulos, N., Zacharias, T.: Delegatable pseudorandom functions and applications. In: ACM CCS, pp. 669–684 (2013)

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

[MNPS04] Malkhi, D., Nisan, N., Pinkas, B., Sella, Y.: Fairplay - secure two-party computation system. In: USENIX Security Symposium, pp. 287–302 (2004)

[MV16] Mittelbach, A., Venturi, D.: Fiat–shamir for highly sound protocols is instantiable. In: Zikas, V., Prisco, R. (eds.) SCN 2016. LNCS, vol. 9841, pp. 198–215. Springer, Cham (2016). doi:10.1007/978-3-319-44618-9_11

[Nao03] Naor, M.: On cryptographic assumptions and challenges. In: Boneh, D. (ed.) CRYPTO 2003. LNCS, vol. 2729, pp. 96–109. Springer, Heidelberg (2003). doi:10.1007/978-3-540-45146-4_6

[OO98] Ohta, K., Okamoto, T.: On concrete security treatment of signatures derived from identification. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 354–369. Springer, Heidelberg (1998). doi:10.1007/BFb0055741

[PS96] Pointcheval, D., Stern, J.: Security proofs for signature schemes. In: Maurer, U. (ed.) EUROCRYPT 1996. LNCS, vol. 1070, pp. 387–398. Springer, Heidelberg (1996). doi:10.1007/3-540-68339-9_33

[Rey01] Reyzin, L.: Zero-Knowledge with Public Keys. Ph.D. thesis, MIT (2001)

[Ros00] Rosen, A.: A note on the round-complexity of concurrent zero-knowledge. In: Bellare, M. (ed.) CRYPTO 2000. LNCS, vol. 1880, pp. 451–468. Springer, Heidelberg (2000). doi:10.1007/3-540-44598-6_28

[RR97] Razborov, A.A., Rudich, S.: Natural proofs. J. Comput. Syst. Sci. **55**(1), 24–35 (1997)

[RRR16] Reingold, O., Rothblum, G.N., Rothblum, R.D.: Constant-round interactive proofs for delegating computation. In: Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18–21, 2016, pp. 49–62 (2016)

[SW14] Sahai, A., Waters, B.: How to use indistinguishability obfuscation: deniable encryption, and more. In: STOC, pp. 475–484 (2014)