ZAPs and Non-Interactive Witness Indistinguishability from Indistinguishability Obfuscation

Nir Bitansky[1,⋆] and Omer Paneth[2,⋆⋆]

[1] MIT, USA
[2] Boston University, USA

Abstract. We present new constructions of two-message and one-message witness-indistinguishable proofs (ZAPs and NIWIs). This includes:

- ZAPs (or, equivalently, non-interactive zero-knowledge in the common random string model) from indistinguishability obfuscation and one-way functions.
- NIWIs from indistinguishability obfuscation and one-way permutations.

The previous construction of ZAPs [Dwork and Naor, FOCS 00] was based on trapdoor permutations. The two previous NIWI constructions were based either on ZAPs and a derandomization-type complexity assumption [Barak, Ong, and Vadhan CRYPTO 03], or on a specific number theoretic assumption in bilinear groups [Groth, Sahai, and Ostrovsky, CRYPTO 06].

1 Introduction

Zero-knowledge proofs [GMR89] and their feasibility for **NP** [GMW91] are fundamental to modern cryptography, allowing to prove any **NP** statement while guaranteeing total privacy of the witness. One of the main pillars on which this exquisite guarantee relies is interaction. Minimizing interaction from zero-knowledge protocols has drawn significant efforts. Since even two-message zero-knowledge, without any setup assumptions, is impossible [GO94], these efforts have either focused on non-interactive zero-knowledge (NIZK) in the trusted *common random string* (or common reference string) model [BFM88], or on achieving non-interactive systems with weaker security guarantees. One notable relaxation is that of *witness indistinguishability* (WI), guaranteeing that the proof does not reveal which one of many witnesses is used.

Following this direction, Dwork and Naor [DN07] show that, unlike zero-knowledge, two-message WI, or as they term *ZAPs*, can be achieved without any

⋆ Part of this work was done while at Tel Aviv university and supported by an IBM Ph.D. Fellowship and the Check Point Institute for Information Security.

⋆⋆ Supported by the Simons award for graduate students in theoretical computer science and an NSF Algorithmic foundations grant 1218461.

Y. Dodis and J.B. Nielsen (Eds.): TCC 2015, Part II, LNCS 9015, pp. 401–427, 2015.

setup. Specifically, they show that assuming one-way functions, the existence of ZAPs in the plain model is equivalent to that of NIZKs in the common random string model. The latter were constructed by Feige, Lapidot, and Shamir from trapdoor permutations [FLS99]. Furthermore, in ZAPs, the verifier's message is a random string that can be used for multiple proofs, giving rise to a completely non-interactive WI system where the first message is fixed non-uniformly. This provided evidence that diverging from the strong notion of zero-knowledge may very well allow to remove interaction altogether.

Barak, Ong, and Vadhan [BOV07] then constructed completely non-interactive WI (NIWI), without any non-uniformity, by derandomizing the ZAP verifier under the assumption that $\mathbf{Dtime}(2^{O(n)})$ has a problem of non-deterministic circuit complexity $2^{\Omega(n)}$. Groth, Ostrovsky, and Sahai [GOS12] subsequently constructed NIWIs under the decision linear assumption on bilinear groups.

ZAPs and NIWIs have found a great number of applications in cryptography, but only few candidate constructions, from specific assumptions, are known. In particular, both [DN07] and [BOV07] rely on trapdoor permutations, for which there are very few candidates, all based on factoring-related assumptions. Alternatively, the [GOS12] construction is based on a specific assumption on bilinear groups. Different constructions from different computational assumptions are still sought after. In this work, we provide new constructions of ZAPs and NIWIs under a rather different type of assumption: *indistinguishability obfuscation*.

Indistinguishability Obfuscation. The goal of obfuscation is to make code unintelligible while preserving its functionality. It has been long considered to be a holy grail of cryptography, with diverse and far reaching applications. Up until recently, there were no candidates obfuscators, except for very restricted classes of programs, and in fact, some classes were shown to be unobfuscatable under the natural virtual black-box notion [BGI$^+$01]. This dramatically changed with the work by Garg et al. [GGH$^+$13b] who proposed a candidate construction of general-purpose obfuscators, based on graded multilinear encodings [GGH13a], and conjectured that it satisfies the seemingly weak notion of indistinguishability obfuscation (iO) [BGI$^+$01], for which no impossibility results are known. This notion only requires that it is hard to distinguish an obfuscation of C_0 from an obfuscation of C_1, for any two circuits C_0 and C_1 of the same size that compute the exact same function.

Perhaps surprisingly, iO has been shown to have variety of powerful applications, such as functional encryption, deniable encryption, two-message multiparty computation [GGH$^+$13b, SW14, GGHR14], and many more. Still, some basic primitives have so far evaded the long arms of iO, including collision-resistant hashing, fully-homomorphic encryption, and also trapdoor permutations, which as said above, are essential in generic ZAP and NIWI constructions.

1.1 Results

We provide new constructions of ZAPs and NIWIs based on iO.

Our first result is a construction of ZAPs (or NIZKs):

Theorem 1.1 (informal). *Assuming indistinguishability obfuscation for a certain family of polysize circuits and one-way functions, there exist ZAPs in the plain model and NIZKs in the common random string model, for every language in* **NP**.[1]

The new ZAP can, in particular, be plugged into the result of Barak, Ong, and Vadhan [BOV07] to obtain a NIWI for all of **NP**, assuming in addition the existence of a language in $\mathbf{Dtime}(2^{O(n)})$ with circuit complexity $2^{\Omega(n)}$. We give a new construction of NIWIs based on indistinguishability obfuscation and one-way permutations.

Theorem 1.2 (informal). *Assuming indistinguishability obfuscation for a certain family of polysize circuits and one-way permutations, there exist NIWI proofs, for every language in* **NP**.

As explained below, in our construction of NIWI, one-way permutations are used to construct a *dense* non-interactive commitment scheme. We show that such commitments are somewhat inherent (see more details below).

Comparison to Previous Constructions. Sahai and Waters [SW14] constructed, from iO, non-adaptive NIZK arguments in the common reference string model; these are insufficient to obtain ZAPs (let alone, NIWIs).

The assumptions that we rely on for either NIWIs, or NIZKs in the random string model (or equivalently, ZAPs) are incomparable to previously used assumptions. Our main assumption is iO, which is not formally known to be either weaker or stronger. While perhaps not weaker, iO does seem to be of different nature than the previous assumptions. Indeed, previous constructions are based on primitives with an *exact* combinatorial or algebraic structure, such as trapdoor permutations [DN07, BOV07], or bilinear maps in appropriate groups [GOS12]. Finding candidates adhering to such exact structures has proven to be challenging, and such candidates remain scarce. In contrast, iO has candidates based on *noisy* graded encodings [GGH$^+$13b], which by now already have several proposed instantiations [GGH13a, CLT13, GGH14]. Future constructions of iO may be based on primitives with even less algebraic structure.

While our construction of NIZKs relies solely on iO and one-way functions, the NIWI construction also requires (certifiable) one-way permutations, which are already a rather structured object, with only few more candidates than trapdoor permutations (based on the hardness of discrete logs). We find that the main appeal of the suggested NIWI construction, compared to previous ones, is its relative simplicity.

1.2 Techniques

We now overview the techniques behind our constructions. We start with the construction of ZAPs, and then move on to the NIWI construction.

[1] The assumption of one-way functions can be replaced with assuming $\mathbf{NP} \neq \mathbf{coRP}$ [KMN$^+$14].

ZAPs. Our main technical contribution towards obtaining ZAPs is a construction of *invariant signatures in the common random string model*, a concept presented by Goldwasser and Ostrovsky [GO92]. We then apply a series of generic transformations from the literature: in the common random string model, invariant signatures imply NIZKs [GO92], which imply ZAPs [DN07] (in the plain model). As a secondary contribution, we also give a full description and proof of the first transformation, previously sketched in [GO92]. Details follow.

Invariant Signatures. Invariant signatures, introduced by Goldwasser and Ostrovsky [GO92], are digital signatures where all valid signatures of any message are either identical, or share a common property. Concretely, we say that a signature scheme is invariant if there is some efficiently computable property P of signatures such that for any message m^* and any verification key vk there is a unique value $P_{\mathsf{vk}}(m^*)$ such that $P(\sigma) = P_{\mathsf{vk}}(m^*)$ for any valid signature σ with respect to vk. Furthermore, it is required that for every message m^*, for an honestly generated verification key (sampled independently of m^*), the property value $P_{\mathsf{vk}}(m^*)$ is pseudo-random, even given the verification key and a signature oracle on messages $m \neq m^*$. Like in [GO92], we consider a relaxed notion of invariant signatures in the common random string model (CRS). Here we require that the property value P of valid signatures is unique for every verification key vk, with overwhelming probability over the choice of the CRS. Pseudo-randomness of P_{vk} should hold even given the CRS.[2]

Before explaining how we construct invariant signatures, let us first motivate them by recalling how they are utilized in the construction of NIZKs.

NIZKs from Invariant Signatures. Goldwasser and Ostrovsky gave a transformation from invariant signatures to NIZKs [GO92]. Their transformation is based on the construction of Feige, Lapidot and Shamir [FLS99] of NIZKs in the *hidden-bits model*. In this model, a random hidden string is available to the prover but is hidden from the verifier. The prover can reveal to the verifier specific bits of the hidden string in the locations of its choice, but it cannot change the value of these bits. [FLS99] also show how to compile a NIZK in the hidden-bits model into a NIZK in the CRS model assuming trapdoor permutations. [GO92] gave a different compilation technique based on invariant signatures. Next, we describe such a compilation following the same high-level idea as [GO92, BGRV09].

We interpret the CRS (available to both prover and verifier) as containing a CRS for the invariant signature, as well as a sequence of messages $\{m_i\}$ and one-time pad bits $\{s_i\}$ where every (m_i, s_i) will be used to obtain a single hidden bit b_i. The prover will sample keys $(\mathsf{sk}, \mathsf{vk})$ for the invariant signature and send the verification key vk to the verifier as part of the proof. The hidden bit b_i is then defined as the bit $P_{\mathsf{vk}}(m_i)$, the property value of the message m_i, XORed with

[2] In the original definition of [GO92], pseudo-randomness is also required for messages m^* sampled adaptively after the verification key. While we do not achieve such *adaptive* pseudo-randomness, the above *selective* pseudo-randomness will suffice for our purpose.

the the one-time pad bit s_i. To reveal the bit b_i, the prover sends to the verifier a signature σ_i on m_i. The verifier can compute b_i by computing $P(\sigma_i) = P_{\mathsf{vk}}(m_i)$.

The uniqueness of the signature guarantees that the prover cannot affect the value of the hidden bits. Note that the prover can always affect the value of b_i by choosing the verification key vk. However, since the length of vk is bounded, this issue can be overcome via soundness amplification [FLS99, GO92]). Another problem is that the prover might affect the distribution of the hidden bits by choosing a verification key such that $P_{\mathsf{vk}}(\cdot)$ is unbalanced. In [GO92, BGRV09], this is addressed by certifying the fact that $P_{\mathsf{vk}}(\cdot)$ is balanced, using a similar approach to [BY96]. In our construction, we guarantee that the hidden bits are uniformly distributed by simply XORing $P_{\mathsf{vk}}(m_i)$ with the random one-time pad bit s_i. Finally, the pseudo-randomness of P_{vk} guarantees that the bits not revealed by the prover remain hidden from the verifier.

Constructing Invariant Signatures. The starting point of our construction is the selectively-secure signature scheme of Sahai-Waters based on iO and one-way functions [SW14]. The basic idea behind the construction is as follows. The secret signing key is simply a key K for a pseudo-random function PRF_K, and a signature σ on message m is simply $\sigma = \mathsf{PRF}_K(m)$. The public verification key is an obfuscation $\tilde{C} \leftarrow i\mathcal{O}(C_K)$ of a circuit C_K that given any m returns $y_m = f(\mathsf{PRF}_K(m))$ for some one-way function f. Verification of any σ for m is simply done by computing $f(\sigma)$ and comparing to the value y_m output by $\tilde{C}_K$.

Sahai and Waters show, based on the indistinguishability obfuscation guarantee, that their scheme is selectively-secure; namely, it is impossible to forge a signature for any preselected message m^*, even given a signature oracle on other messages. The idea is to consider an alternative to the the circuit C_K that computes the same function, but while only "knowing" y_{m^*}, and without actually "knowing" the preimage $\mathsf{PRF}_K(m^*)$. This is achieved using their elegant *puncturing technique*. Specifically, instead of using any PRF family, they use a *puncturable PRF*. In such PRFs, it is possible to puncture a given key K at an arbitrary point m^* in the domain of the function. The punctured function $\mathsf{PRF}_{K_{m^*}}$, with punctured key K_{m^*}, preserves functionality at any other point, but hides any information on the point $\mathsf{PRF}_K(m^*)$; namely, the value $\mathsf{PRF}_K(m^*)$ is pseudo-random, even given (m^*, K_{m^*}). Such puncturable PRFs follow from the GGM [GGM86] construction [BW13, BGI14, KPTZ13].

Using a puncturable PRF in the implementation of C_K, it can be shown that if a forger succeeds in finding a preimage of $y_{m^*} = f(\mathsf{PRF}_K(m^*))$, it would also succeed had we provided it with an obfuscation of the alternative circuit $C_{K_{m^*}, y_{m^*}}$. The circuit $C_{K_{m^*}, y_{m^*}}$ computes the same function as C_K, but in a different way: it only has the punctured key K_{m^*}, and has the value y_{m^*} directly hardwired into it, so that it does not have to evaluate the PRF in order to compute it. Thus, the fact that the forger still succeeds follows by the guarantee of indistinguishability obfuscation. However, now by the pseudo-randomness guarantee at the punctured point m^*, we know that $\mathsf{PRF}_K(m^*)$ is pseudo random, even given the circuit $C_{K_{m^*}, y_{m^*}}$, and thus the forger can be used to invert the one-way function f.

We next observe that the Sahai-Waters signature scheme can be made invariant as follows. To get uniqueness, we can use an injective one-way function f instead of an arbitrary one-way function. Indeed, this guarantees that for any, even malicious, verification key $\tilde{C}$ and any message m, the value $y^* = \tilde{C}(m)$ has a unique preimage under f that will be accepted in verification. To get (selective) pseudo-randomness, rather than just (selective) unforgeability, we can define the property P to extract a Goldreich-Levin hardcore bit [GL89] from the unique signature with respect to a fixed seed put in the verification key; this preserves uniqueness.

The above solution requires the extra assumption that injective one-way functions exist. We show that a more significant modification of the SW scheme allows to rely on any one-way function. Unlike the solution above that did not rely on a CRS, the new solution will (as explained before, this is still sufficient for our purposes). The basic idea is the following. Imagine we had at our expense a non-interactive perfectly-binding commitment scheme Com. We could then augment the circuit C_K to output, instead of a one-way function $f(\mathsf{PRF}_K(m^*))$, a commitment $c_{m^*} = \mathsf{Com}(b; r)$, to plaintext $b = \mathsf{PRF}_K(m^*)$, where the randomness r is derived, say, by applying another $\mathsf{PRF}_{K'}$ to m^* (or simply setting $(b, r) = \mathsf{PRF}_K(m^*)$). A signature would then include the plaintext underlying the commitment $\mathsf{PRF}_K(m^*)$ and the randomness $r = \mathsf{PRF}_{K'}(m^*)$. The unique property P will simply be the plaintext b.

Indeed, uniqueness will now follow by the perfect binding of the commitment, and pseudo-randomness of b will follow using a similar puncturing argument to the one above, coupled with the hiding of the commitment Com. However, non-interactive perfectly-binding commitments are only known based on injective one-way functions [Blu81], which may take us back to square one. Here the CRS comes to our aid. We can use Naor's [Nao91] two-message statistically-binding commitment scheme, where the first receiver message is simply a random string that can be put in the CRS; indeed, this commitment can be based on any one-way function.

NIWIs. The first stepping stone in our NIWI construction is a natural idea suggested by Niu et al. [NLLT14], where it is described using the terminology of witness encryption [GGSW13]. In witness encryption, anyone can encrypt a message m under a public candidate instance x for some **NP** language $\mathcal{L}$ (x is thought of as the public key); if $x \in \mathcal{L}$, anyone holding a corresponding witness w can decrypt the encrypted $\mathsf{Enc}_x(m)$; however, if $x \notin \mathcal{L}$, the encryption is semantically secure; namely, $\mathsf{Enc}_x(m)$ is computationally indistinguishable from $\mathsf{Enc}_x(m')$ for any two messages m, m'. Such a scheme can be easily constructed from any indistinguishability obfuscator (as we shall soon see).

Given a witness encryption scheme, Niu et al. suggest the following candidate for a NIWI. Given $(x, w) \in \mathcal{R}_{\mathcal{L}}$, a proof that $x \in \mathcal{L}$ is simply an indistinguishability obfuscation $\tilde{D} \leftarrow i\mathcal{O}(D_{x,w})$ of the witness decryption circuit $D_{x,w}$ that given a witness encryption $\mathsf{Enc}_x(m)$, decrypts it with the witness w and outputs m. Verification is done by running the circuit $\tilde{D}$ on an encryption $\mathsf{Enc}_x(m)$ of a

random $m \leftarrow \{0,1\}^n$, and testing whether it successfully decrypts m. Indeed, if $x \notin \mathcal{L}$, $\tilde{D}$ fails with overwhelming probability due to semantic security.

What about witness indistinguishability? at first it seems that regardless of which witness w is used by $D_{x,w}$, it has the same functionality, since any witness can be used for decryption. Thus, WI should follow by the iO guarantee. However, this argument is flawed—while, for valid (honestly generated) witness encryptions $\mathsf{Enc}_x(m)$, $D_{x,w}$ behaves the same regardless of the witness, *this might not be true for maliciously generated encryptions.*

To illustrate this consider a witness encryption scheme implemented using indistinguishability obfuscation, where $\mathsf{Enc}_x(m)$ consists of an obfuscation $\tilde{E} \leftarrow i\mathcal{O}(E_x^m)$ for the circuit E_x^m that given as input a proper witness $w \in \mathcal{R}_{\mathcal{L}}(x)$ outputs m and otherwise $\perp$. For $x \notin \mathcal{L}$, such a circuit always returns $\perp$, regardless of m, and thus semantic security follows from iO. However, if we instantiate the above candidate NIWI with this witness encryption scheme, the result will be completely insecure. A malicious verifier may obfuscate an arbitrary circuit, instead of a proper circuit E_x^m, and distinguish between different witnesses. Taken to the extreme, it could just obfuscate the identity, and recover from $\tilde{D}$ the entire witness w.

Fixing the NIWI Using ZAPs. the above problem can be resolved by requiring that the malicious verifier proves to $\tilde{D}$ that its witness encryption is indeed a proper encryption of some plaintext with some randomness. However, to maintain soundness, this should be done while keeping the one-wayness of m. To achieve this, we rely on ZAPs, and the Feige-Shamir trapdoor paradigm [FS89]: the prover will hard-code into $\tilde{D}$ a random first message for a ZAP, and the verifier will prove to $\tilde{D}$ that either $\tilde{E}$ was generated properly or that some "trapdoor" statement is true. In order to assure that the verifier's encryption is proper, the trapdoor statement is usually chosen such that it is true, but it is hard for the verifier to find a witness for it, for example, stating that a random sting is in the image of a one-way permutation. However, in our setting, since the ZAP is not a proof of knowledge, such a trapdoor statement is insufficient.

In our protocol, we do not rely on the fact that the trapdoor statement is hard to prove, but rather we aim to design a trapdoor statement such that, if true, certifies the validity of the witness encryption $\tilde{E}$. The problem is that such certification cannot use the encryption's randomness or the plaintext m as a witness, otherwise we cannot argue that the ZAP hides m. The key observation is that it is enough to certify that the encryption $\tilde{E}$ behaves in the same way on any two potential witnesses. To implement this idea, the trapdoor statement will include a pair of perfectly binding commitments $\mathsf{c}_1, \mathsf{c}_2$ chosen by the prover (the honest prover commits to all-zero strings). The statement asserts that there is a pair of candidate witnesses w_1, w_2 such that $\mathsf{c}_1, \mathsf{c}_2$ are commitments to w_1, w_2 and the verifier's encryption $\tilde{E}$ decrypts to the same value when decrypted with either w_1 or w_2.

Let us describe the intuition behind the proof of security. To prove soundness, we rely on the fact that $\mathsf{c}_1, \mathsf{c}_2$ are are computed using a *dense* commitment scheme, where every string has some valid decommitment. Assume there exists

an accepting proof $(\tilde{D}, \mathsf{c}_1, \mathsf{c}_2)$ for a false statement $x \notin \mathcal{L}$, meaning that $\tilde{D}$ manages to invert a witness encryption $\tilde{E}$ for a random message, given also the ZAP described above. We show that, $\tilde{D}$ must break the semantic security of the witness encryption. Indeed, letting w_1, w_2 be the plaintexts underlying $\mathsf{c}_1, \mathsf{c}_2$, we note that, since $x \notin \mathcal{L}$, decrypting with either one results in the same value $\perp$. Therefore, the trapdoor statement is true, and we could have used it to compute a ZAP proof π, without compromising the semantic security of $\tilde{E}$. Since $\tilde{D}$ cannot tell the difference between the two ZAP proofs, it would still invert $\tilde{E}$, and thus violate semantic security. We note that for the above argument to go through, we rely on the fact that the ZAP guarantees witness-indistinguishability against *non-uniform verifiers*; indeed, the reduction described above gets a non-uniform advice: the decommitment information for the commitments $\mathsf{c}_1, \mathsf{c}_2$.

To show that the proof is WI, consider any instance $x \in \mathcal{L}$ with two valid witnesses w_1, w_2. We go through several hybrid experiments. We start by using the hiding property of the commitment to replace c_1 and c_2 with commitments to w_1 and w_2, instead of all-zero strings. Now if the verifier generates an encryption $\tilde{E}$ with a valid ZAP proof π it follows from the the soundness of the ZAP and the binding of the commitment that either $\tilde{E}$ is a valid witness encryption, or $\tilde{E}$ decrypts to the same value when decrypted with either w_1 or w_2. In any case, the witness decryption circuits D_{x,w_1} and D_{x,w_2} agree on the input $(\tilde{E}, \pi)$. By the iO guarantee, the obfuscated decryption circuits are thus indistinguishable, and we can replace one with the other.

A Note on Statistical Soundness. At first glance, it may seem that our reliance on computational primitives such as witness encryption implies that the resulting system is only computationally sound; we stress, however, that soundness is statistical.[3] To cheat in our protocol, the (unbounded) prover must produce a proof consisting of a *small* circuit (allegedly an obfuscated witness description circuit). The soundness of the system is based on the fact that this computationally-bounded circuit cannot break the security of the underling primitives. Indeed, the computational assumptions imply that such a circuit simply does not exist.

Additionally, we note that the soundness we get is statistical and not *perfect* as in [BOV07, GOS12]. In the language of [BOV07], we get **MA** proofs rather than **NP** proofs (for all languages in **NP**).

On the Necessity of Dense Commitments. The NIWI construction described above can be based on a non-interactive commitment scheme satisfying the following properties. First, it is computationally hiding. Second, it is statistically binding, but only against *honest* committers; namely, honestly generated commitments can only be opened to a single value. Finally, the commitment is *dense*; that is, every string in the range of the commitment, can be opened to at least one value (commitments that are not generated honestly may potentially

[3] In fact, any single message argument system that is sound against non-uniform provers must be statistically sound, as accepting proofs for false statements may be hardwired to the prover.

be opened to more than one value). We observe that such dense commitments are somewhat necessary. Specifically, using NIWIs, we can transform any non-interactive statistically binding commitment into a commitment satisfying the above three requirements (note that statistically-binding commitments can be constructed from any injective one-way function [Blu81].)

The basic idea is to commit twice to the same value and add a NIWI proof that one of the two commitments were honestly generated. A valid opening of this commitment would consist of an opening of any one of the two underlying commitments. If the NIWI is not accepting, the committed value is set arbitrarily to zero.[4] The hiding of the new commitment follows from that of the original commitment, together with the witness-indistinguishability of the NIWI. Binding, for honestly generated commitments (where the two underlying plaintexts are identical), follows from the binding property of the original commitment. Finally, the fact that the new commitment is dense follows directly from the soundness of the NIWI.

We note that it may still be possible that NIWIs, and in particular, dense commitments as above, can be based on iO and any one-way function.

Organization. In Section 2, we present the basic definitions used in the paper. In Section 3, we define and construct invariant signatures. In Section 4, we describe the Goldwasser-Ostrovsky transformation from invariant signatures to NIZKs. Section 5 describes the NIWI construction.

2 Definitions

2.1 Non-Interactive Zero-Knowledge

Definition 2.1. *A pair of PPT algorithms $(\mathcal{P}, \mathcal{V})$ is a NIZK proof in the CRS model if they satisfy the following properties:*

1. *Completeness: there exists a polynomial r denoting the length of the common random string such that for every $(x, w) \in \mathcal{R}_{\mathcal{L}}$ we have that:*

$$\Pr_{\mathcal{P},\mathsf{crs}\leftarrow\{0,1\}^{r(|x|)}}[\mathcal{V}(x,\mathsf{crs},\pi)=1 : \pi \leftarrow \mathcal{P}(x,w,\mathsf{crs})] = 1 \ .$$

2. *Soundness: for every $x \notin \mathcal{L}$ we have that:*

$$\Pr_{\mathsf{crs}\leftarrow\{0,1\}^{r(|x|)}}[\exists \pi : \mathcal{V}(x,\mathsf{crs},\pi)=1] < 2^{-|x|} \ .$$

3. *Zero-Knowledge: there exists a PPT algorithm $\mathcal{S}$ such that:*

$$\left\{(\mathsf{crs}, \mathcal{P}(x,w,\mathsf{crs})) : \mathsf{crs} \leftarrow U_{r(|x|)}\right\}_{(x,w)\in\mathcal{R}_{\mathcal{L}}} \approx_c \left\{\mathcal{S}(x)\right\}_{(x,w)\in\mathcal{R}_{\mathcal{L}}}$$

[4] Here we assume NIWI with perfect soundness. In particular we assume that the verification procedure of the NIWI is deterministic. Dense commitments satisfying a slightly more involved definition can be constructed from NIWI with only statistical soundness.

Remark 2.1. Definition 2.1 considers only non-adaptive soundness and zero-knowledge. Additionally, zero-knowledge is not guaranteed when multiple statement are proven with respect to the same CRS. We note that any NIZK proof system for **NP** can be transformed into a system that does not have these disadvantages assuming only OWFs [FLS99].

2.2 ZAPs

ZAPs [DN07] are two-message public-coin witness-indistinguishable proofs, defined as follows.

Definition 2.2. *A pair of algorithms $(\mathcal{P}, \mathcal{V})$, where $\mathcal{P}$ is PPT and $\mathcal{V}$ is (deterministic) polytime, is a ZAP for an* **NP** *relation $\mathcal{R}_{\mathcal{L}}$ if it satisfies:*

1. *Completeness: there exists a polynomial r such that for every $(x, w) \in \mathcal{R}_{\mathcal{L}}$,*

$$\Pr_{\mathcal{P}, \mathsf{r} \leftarrow \{0,1\}^{r(|x|)}} [\mathcal{V}(x, \pi, \mathsf{r}) = 1 : \pi \leftarrow \mathcal{P}(x, w, \mathsf{r})] = 1 \ .$$

2. *Adaptive soundness: for every malicious prover $\mathcal{P}^*$ and every $n \in \mathbb{N}$:*

$$\Pr_{\mathsf{r} \leftarrow \{0,1\}^{r(n)}} \left[\exists \begin{matrix} x \in \{0,1\}^n \setminus \mathcal{L} \\ \pi \in \{0,1\}^* \end{matrix} : \mathcal{V}(x, \pi, \mathsf{r}) = 1 \right] \leq 2^{-n} \ .$$

3. *Witness indistinguishability: for any sequence $\mathcal{I} = \{(x, w_1, w_2) : w_1, w_2 \in \mathcal{R}_{\mathcal{L}}(x)\}$ and any first-message sequence $\mathcal{R} = \left\{\mathsf{r}_{x,w_1,w_2} \in \{0,1\}^{r(|x|)} : (x, w_1, w_2) \in \mathcal{I}\right\}$:*

$$\{\pi_1 \leftarrow \mathcal{P}(x, w_1, \mathsf{r}_{x,w_1,w_2})\}_{(x,w_1,w_2) \in \mathcal{I}} \approx_c \{\pi_2 \leftarrow \mathcal{P}(x, w_2, \mathsf{r}_{x,w_1,w_2})\}_{(x,w_1,w_2) \in \mathcal{I}} \ .$$

2.3 NIWIs

NIWIs [BOV07] are completely non-interactive witness-indistinguishable proofs.

Definition 2.3. *A pair of PPT algorithms $(\mathcal{P}, \mathcal{V})$ is a NIWI for an* **NP** *relation $\mathcal{R}_{\mathcal{L}}$ if it satisfies:*

1. *Completeness: for every $(x, w) \in \mathcal{R}_{\mathcal{L}}$,*

$$\Pr_{\mathcal{P}}[\mathcal{V}(x, \pi) = 1 : \pi \leftarrow \mathcal{P}(x, w)] = 1 \ .$$

2. *Soundness: there exists a negligible function μ, such that for every $x \notin \mathcal{L}$ and $\pi \in \{0,1\}^*$:*

$$\Pr_{\mathcal{V}}[\mathcal{V}(x, \pi) = 1] \leq \mu(|x|) \ .$$

3. *Witness indistinguishability: for any sequence $\mathcal{I} = \{(x, w_1, w_2) : w_1, w_2 \in \mathcal{R}_{\mathcal{L}}(x)\}$:*

$$\{\pi_1 : \pi_1 \leftarrow \mathcal{P}(x, w_1)\}_{(x,w_1,w_2) \in \mathcal{I}} \approx_c \{\pi_2 : \pi_2 \leftarrow \mathcal{P}(x, w_2)\}_{(x,w_1,w_2) \in \mathcal{I}} .$$

2.4 Indistinguishability Obfuscation

Indistinguishability obfuscation (iO) was introduced in [BGI+01] and given a candidate construction in [GGH+13b], and subsequently in [BR13, BGK+13].

Definition 2.4 (Indistinguishability Obfuscation [BGI+01]). *A PPT algorithm $i\mathcal{O}$ is said to be an* indistinguishability obfuscator *for a collection of polysize circuits $\mathcal{C} = \bigcup_{n\in\mathbb{N}} \mathcal{C}_n$, if it satisfies:*

1. *Functionality: For any $C \in \mathcal{C}$,*

$$\Pr_{i\mathcal{O}}[\forall x : i\mathcal{O}(C)(x) = C(x)] = 1 \ .$$

2. *Indistinguishability: For any polysize distinguisher $\mathcal{D}$ there negligible function μ, such that for any $n \in \mathbb{N}$ and $C_1, C_2 \in \mathcal{C}_n$ of the same size and functionality*

$$\left|\Pr_{i\mathcal{O}}[\mathcal{D}(i\mathcal{O}(C_1)) = 1] - \Pr_{i\mathcal{O}}[\mathcal{D}(i\mathcal{O}(C_2)) = 1]\right| \leq \mu(n) \ .$$

3 Invariant Signatures from Indistinguishability Obfuscation

In this section, we recall the definition of invariant signatures [GO92] in the common random string (CRS) model and construct them based on iO.

Roughly, invariant signatures are digital signatures where valid signatures of any message are either identical, or share a common property. More accurately, there is an efficiently computable property P of signatures such that for any message m^* and any verification key vk there is a unique value $P_{\mathsf{vk}}(m^*)$ such that $P(\sigma) = P_{\mathsf{vk}}(m^*)$ for any valid signature σ with respect to vk. Furthermore, it is required that for every message m^*, for an honestly generated verification key (sampled independently of m^*), the property value $P_{\mathsf{vk}}(m^*)$ is pseudo-random, even given the verification key and a signature oracle on messages $m \neq m^*$. Like in [GO92], we consider a relaxed notion of invariant signatures in the common random string model (CRS). Here the property value P is unique for every verification key vk, with overwhelming probability over the choice of the CRS, and pseudo-randomness of P_{vk} should hold even given the CRS. (In the original definition of [GO92], pseudo-randomness is also required for messages m^* sampled adaptively after the verification key. While we do not achieve such *adaptive* pseudo-randomness, the above *selective* pseudo-randomness will suffice for our purpose.)

Definition 3.1 (Invariant Signatures in the CRS Model). *A triple of polytime algorithms $(\mathsf{Gen}, \mathsf{Sign}, \mathsf{Ver})$, where Gen is randomized, is a digital signature scheme with invariant signatures and selective security in the CRS model if it satisfies the following properties:*

1. *Syntax and completeness: There exists a polynomial r such that for every security parameter $n \in \mathbb{N}$, and for every message $m \in \{0,1\}^*$ we have that:*

$$\Pr_{\mathsf{crs}\leftarrow U_{r(n)}}[\mathsf{Ver}_{\mathsf{vk}}(\mathsf{crs}, m, \sigma) = 1 : \sigma \leftarrow \mathsf{Sign}_{\mathsf{sk}}(m), (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs})] = 1 \ .$$

2. *Uniqueness: There exists a deterministic, efficiently computable, predicate* $\overline{P : \{0,1\}^* \to \{0,1\}}$*, and a negligible function* μ *such that:*

$$\Pr_{\mathsf{crs} \leftarrow U_{r(n)}}[\exists m, \mathsf{vk}, \sigma_1, \sigma_2 : P(\sigma_1) \neq P(\sigma_2) \wedge \mathsf{Ver}_{\mathsf{vk}}(\mathsf{crs}, m, \sigma_1) = \mathsf{Ver}_{\mathsf{vk}}(\mathsf{crs}, m, \sigma_2) = 1] \leq \mu(n) \ .$$

3. *Pseudo-randomness: For every poly-size adversary* A*, there exists a negligible function* μ *such that for every security parameter* $n \in \mathbb{N}$*, and for every message* $m \in \{0,1\}^n$*:*

$$\left| \begin{array}{l} \Pr_{\mathsf{A},\mathsf{crs} \leftarrow U_{r(n)}}[\mathsf{A}^{\mathsf{Sign}^*_{\mathsf{sk},m}}(\mathsf{crs}, \mathsf{vk}, m, P(\mathsf{Sign}_{\mathsf{sk}}(m))) = 1 : (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs})] - \\ \Pr_{\mathsf{A},\mathsf{crs} \leftarrow U_{r(n)}, b \leftarrow U_1}[\mathsf{A}^{\mathsf{Sign}^*_{\mathsf{sk},m}}(\mathsf{crs}, \mathsf{vk}, m, b) = 1 : (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs})] \end{array} \right| \leq \mu(n) \ ,$$

where $\mathsf{Sign}^*_{\mathsf{sk},m}$ *is an oracle that is identical to* $\mathsf{Sign}_{\mathsf{sk}}$ *excpet that on input* m *it outputs* $\perp$*.*

Remark 3.1 (Unforgeability). We do not explicitly require that the signature scheme is unforgeable against selective attackers. Unforgeability is, in fact, implied by uniqueness and the pseudo-randomness properties. In particular, if an adversary can forge a signature σ on a message m, it can compute $P(\sigma)$ and break pseudo-randomness.

Sahai and Waters construct digital signature scheme with based on iO and one-way functions [SW14]. As outline in the introduction, we observe that a modification of their construction is also invariant assuming also *injective* one-way functions.

Theorem 3.1 (follows from [SW14]). *Assuming indistinguishability obfuscation and injective OWFs, there exists a selectively secure invariant signature scheme.*

We show that, in the CRS model, we can in fact construct selectively secure invariant signatures based on iO and *any* one-way function.

Theorem 3.2. *Assuming indistinguishability obfuscation and one-way functions, there exists a selectively-secure invariant signature scheme in the CRS model.*

Like the Shahi-Waters construction, the construction here relies on their punctured program paradigm. We next define puncturable pseudo-random functions, a central tool in our construction, and then move to describe the construction.

3.1 Puncturable PRFs

We consider a simple case of the puncturable PRFs where any PRF might be punctured at a single point. The definition is formulated as in [SW14].

Definition 3.2 (Puncturable PRFs). *Let ℓ, m be polynomially bounded length functions. An efficiently computable family of functions*

$$\mathcal{PRF} = \left\{ \mathsf{PRF}_K : \{0,1\}^{m(n)} \to \{0,1\}^{\ell(n)} \;\middle|\; K \in \{0,1\}^n, n \in \mathbb{N} \right\} \ ,$$

associated with an efficient (probabilistic) key sampler $\mathcal{K}_{\mathcal{PRF}}$, is a puncturable PRF if there exists a puncturing algorithm Punc *that takes as input a key $K \in \{0,1\}^n$, and a point x^*, and outputs a punctured key K_{x^*}, so that the following conditions are satisfied:*

1. *Functionality is preserved under puncturing: For every $x^* \in \{0,1\}^{\ell(n)}$,*

$$\Pr_{K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n)} [\forall x \neq x^* : \mathsf{PRF}_K(x) = \mathsf{PRF}_{K_{x^*}}(x) \mid K_{x^*} = \mathsf{Punc}(K, x^*)] = 1 \ .$$

2. *Indistinguishability at punctured points: The following ensembles are computationally indistinguishable:*
   - $\{x^*, K_{x^*}, \mathsf{PRF}_K(x^*) \mid K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n), K_{x^*} = \mathsf{Punc}(K, x^*)\}_{x^* \in \{0,1\}^{m(n)}, n \in \mathbb{N}}$
   - $\left\{x^*, K_{x^*}, u \;\middle|\; K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n), K_{x^*} = \mathsf{Punc}(K, x^*), u \leftarrow \{0,1\}^{\ell(n)}\right\}_{x^* \in \{0,1\}^{m(n)}, n \in \mathbb{N}}$
     .

To be explicit, we include x^* in the distribution; throughout, we shall assume for simplicity that a punctured key K_{x^*} includes x^* in the clear. As shown in [BGI14, BW13, KPTZ13], the GGM [GGM86] PRF yield puncturable PRFs as defined above.

3.2 Invariant Signatures Construction

We now present the details of our construction, an overview is given in the introduction. We shall rely on the following primitives:

- A two-message statistically binding commitment Com with a random first message based on any one-way function [Nao91]. We denote by C, S the polynomials such that $\mathsf{Com}_s(b; r) \in \{0,1\}^{C(n)}$ is a commitment to a bit b, and where the first commitment message is $s \in \{0,1\}^{S(n)}$, and the randomness is $r \in \{0,1\}^n$.
- A family of puncturable PRFs $\mathcal{PRF} = \{\mathsf{PRF}_K\}$ from $\{0,1\}^n$ to $\{0,1\}^{n+1}$ associated with a key sampler $\mathcal{K}_{\mathcal{PRF}}$ and a puncturing algorithm Punc.
- An indistinguishability obfuscator $i\mathcal{O}$.

Construction 3.3 (A selectively-secure invariant signature)

The CRS. *The CRS consists of a random first message s for* Com*. Throughout the construction, one may identify the notation* crs *with that of s.*

The algorithm Gen. *Given the CRS s,* Gen *samples a PRF key $K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n)$.* Gen *sets* $\mathsf{sk} = K$ *and sets* $\mathsf{vk} = i\mathcal{O}([C_{s,K}]_\ell)$ *where $[C_K]_\ell$ is a "commitment to pseudo-random property" circuit $C_{s,K}$, given by Figure 1, padded up to length the maximum size ℓ of the circuits given in Figures 1,2.*

Hardwired: CRS containing a first message $s \in \{0,1\}^{S(n)}$ for Com and a PRF key $K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n)$.
Input: Message $m \in \{0,1\}^n$.
Output: Obtain $(b', r') = \mathsf{PRF}_K(m)$ and output $\mathsf{Com}_s(b', r')$.

Fig. 1. The "commitment to pseudo-random property" circuit $C_{s,K}$

Hardwired:
1. CRS containing a first message $s \in \{0,1\}^{S(n)}$ for Com.
2. Punctured PRF key $K_{m^*} = \mathsf{Punc}(K, m^*)$ where $K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n)$.
3. Commitment $\mathsf{c}^* \in \{0,1\}^{C(n)}$.

Input: Message $m \in \{0,1\}^n$.
Output:
1. If $m = m^*$, output c^*.
2. Else, obtain $(b', r') = \mathsf{PRF}_{K_{m^*}}(m)$ and output $\mathsf{Com}_s(b', r')$.

Fig. 2. The circuit $C_{s,K_{m^*},\mathsf{c}^*}$

The algorithm Sign. *Given the secret key K and a message m output $(b', r') = \mathsf{PRF}_K(m)$.*

The algorithm Ver. *Given the obfuscated circuit vk, the CRS s, a message m and a signature $\sigma = (b, r)$, obtain $\mathsf{c} = \mathsf{vk}(m)$. Output 1 if $\mathsf{c} = \mathsf{Com}_s(b; r)$. Otherwise, output 0.*

Proposition 3.1. *Construction 3.3 is a selectively-secure invariant signature scheme in the CRS model.*

Proof. It is straightforward to verify the completeness of the construction. Next we prove the uniqueness and pseudo-randomness properties.

Uniqueness. For a signature $\sigma = (b, r)$ let $P(\sigma)$ be the a predicate that outputs b. Let G be the event, over the choice of the CRS crs that there exist a message $m \in \{0,1\}^n$, a verification key vk and a pair of signatures $\sigma_1 = (b_1, r_1), \sigma_2 = (b_2, r_2)$ such that:

$$P(\sigma_1) \neq P(\sigma_2) \wedge \mathsf{Ver}_{\mathsf{vk}}(\mathsf{crs}, m, \sigma_1) = \mathsf{Ver}_{\mathsf{vk}}(\mathsf{crs}, m, \sigma_2) = 1 \ .$$

Equivalently, $b_1 \neq b_2$ and $\mathsf{Com}_s(b_1; r_1) = \mathsf{Com}_s(b_2; r_2) = \mathsf{vk}(m)$. Since with overwhelming probability Com_s is perfectly binding property of Com, it holds that $\Pr_{\mathsf{crs} \leftarrow U_{r(n)}}[G] \leq \mathrm{negl}(n)$, as required.

Pseudo-Randomness. Fix any polysize adversary A, and for every message $m \in \{0,1\}^n$, let $p_0(m)$ denote the probability that it outputs 1 given the unique property b of any signature (b, r) on m:

$$p_0(m) = \Pr\left[\mathsf{A}^{\mathsf{Sign}^*_{\mathsf{sk},m}(\cdot)}(\mathsf{crs},\mathsf{vk},m,P(\mathsf{Sign}_{\mathsf{sk}}(m))) = 1 : \begin{array}{c}\mathsf{crs} \leftarrow U_{r(n)} \\ (\mathsf{sk},\mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs})\end{array}\right]$$

$$= \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b) = 1 : \begin{array}{c} s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ (b,r) = \mathsf{PRF}_K(m) \\ \mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K}]_\ell)\end{array}\right] ,$$

where $\mathsf{Sign}^*_{\mathsf{sk},m}(\cdot) \equiv \mathsf{PRF}^*_{K,m}(\cdot)$ is an oracle that is identical to $\mathsf{Sign}_{\mathsf{sk}}(\cdot) \equiv \mathsf{PRF}_K(\cdot)$, except that on input m it outputs $\perp$.

Consider an alternative experiment where vk is chosen to be an obfuscation of the circuit C_{s,K_m,c^*}, rather than $C_{s,K}$, where $\mathsf{c}^* = \mathsf{Com}_s(b,r)$, and (b,r) are computed as before. Let $p_1(m)$ denote the probability that A outputs 1 in this augmented experiment:

$$p_1(m) = \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b) = 1 : \begin{array}{c} s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ K_m \leftarrow \mathsf{Punc}(K,m) \\ (b,r) = \mathsf{PRF}_K(m) \\ \boxed{\mathsf{c}^* = \mathsf{Com}_s(b,r)} \\ \boxed{\mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K_m,\mathsf{c}^*}]_\ell)}\end{array}\right] .$$

Since the circuits $C_{s,K}$ and C_{s,K_m,c^*} are equivalent, it follows from the security of $i\mathcal{O}$ that the circuits $i\mathcal{O}([C_{s,K}]_\ell)$ and $i\mathcal{O}([C_{s,K_m,\mathsf{c}^*}]_\ell)$ are computationally indistinguishable, and therefore

$$|p_0(m) - p_1(m)| < \mathrm{negl}(n) \ .$$

Next, consider another experiment where the signature (b,r) is chosen uniformly at random, instead of being set to $\mathsf{PRF}_K(m)$. We denote by $p_2(m)$ be the probability that A outputs 1 in this experiment:

$$p_2(m) = \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b) = 1 : \begin{array}{c} s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ K_m \leftarrow \mathsf{Punc}(K,m) \\ \boxed{(b,r) = U_{n+1}} \\ \mathsf{c}^* = \mathsf{Com}_s(b,r) \\ \mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K_m,\mathsf{c}^*}]_\ell)\end{array}\right] .$$

By the indistinguishability at punctured points property of $\mathcal{PRF}$:

$$|p_1(m) - p_2(m)| \leq \mathrm{negl}(n) \ ;$$

Indeed, to distinguish between $K_m, \mathsf{PRF}_K(m)$ and $K_m, U_{|m|+1})$, a distinguisher can perfectly emulate A, by answering its oracle queries $m' \neq m$ using the punctured key K_m.

Consider yet another experiment where, instead of giving A the bit b, we replace it with a random independent bit. We denote by $p_3(m)$ the probability that the adversary outputs 1 in this experiment:

$$p_3(m) = \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,\boxed{b'}) = 1 : \begin{array}{c} \boxed{b' \leftarrow U_1} \\ s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ K_m \leftarrow \mathsf{Punc}(K,m) \\ (b,r) = U_{n+1} \\ \mathsf{c}^* = \mathsf{Com}_s(b,r) \\ \mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K_m,\mathsf{c}^*}]_\ell) \end{array}\right] .$$

Then, by the computational hiding property of Com:

$$|p_2(m) - p_3(m)| \leq \mathrm{negl}(n) \ .$$

We define the probabilities p_4, p_5 in the same way we defined p_1, p_0 respectively, except that in these experiments, A gets a random independent bit b'; that is,

$$p_4(m) = \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b') = 1 : \begin{array}{c} b' \leftarrow U_1 \\ s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ K_m \leftarrow \mathsf{Punc}(K,m) \\ \boxed{(b,r) = \mathsf{PRF}_K(m)} \\ \mathsf{c}^* = \mathsf{Com}_s(b,r) \\ \mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K_m,\mathsf{c}^*}]_\ell) \end{array}\right] ,$$

$$p_5(m) = \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b') = 1 : \begin{array}{c} b' \leftarrow U_1 \\ s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ (b,r) = \mathsf{PRF}_K(m) \\ \boxed{\mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K}]_\ell)} \end{array}\right] ,$$

Following the same arguments as before:

$$|p_3(m) - p_4(m)| \leq \mathrm{negl}(n) \quad , \quad |p_4(m) - p_5(m)| \leq \mathrm{negl}(n) \ ,$$

and overall:

$$|p_0(m) - p_5(m)| \leq \mathrm{negl}(n) \ .$$

Thus, we have shown as required that for every $m \in \{0,1\}^n$:

$$\left| \begin{array}{l} \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b) = 1 : \begin{array}{c} s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ (b,r) = \mathsf{PRF}_K(m) \\ \mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K}]_\ell) \end{array}\right] \\ - \Pr\left[\mathsf{A}^{\mathsf{PRF}^*_{K,m}(\cdot)}(s,\mathsf{vk},m,b') = 1 : \begin{array}{c} b' \leftarrow U_1 \\ s \leftarrow U_{r(n)} \\ K \leftarrow \mathcal{K}_{\mathcal{PRF}}(1^n) \\ (b,r) = \mathsf{PRF}_K(m) \\ \mathsf{vk} \leftarrow i\mathcal{O}([C_{s,K}]_\ell) \end{array}\right] \end{array} \right| \leq \mathrm{negl}(n) \ .$$

4 NIZKs and ZAPs from Invariant Signatures

In this section, we show how to construct NIZKs in the CRS model based on invariant signatures. A construction of ZAPs from NIZKs is given in [DN07]. Feige, Lapidot and Shamir constructed a NIZK proof system that is unconditionally secure in the *hidden-bits model*. They also showed how to transform NIZK in the hidden-bits model to NIZK in the CRS model. Goldwasser and Ostrovsky give a different transformation based on invariant signatures. We present a transformation that follows [GO92], in most parts, and provide a full proof of security.

We start by formally defining NIZK in the hidden-bits model. In this model, a random string crs is sampled as a trusted setup. The prover can read all the bits of crs and reveal a subset of these bits to the verifier, corresponding to indices $\mathcal{I}$. The prover cannot change the bits of crs, and the verifier gets no information about the bits of crs that where not revealed by the prover.

Definition 4.1 (NIZK Proof in the Hidden-Bits Model). *A pair of PPT algorithms $(\mathcal{P}, \mathcal{V})$ is a NIZK proof in the hidden-bits model if it satisfies the following properties:*

1. *Completeness: there exists a polynomial r denoting the length of the hidden random string, such that for every $(x, w) \in \mathcal{R}_{\mathcal{L}}$ we have that:*

$$\Pr_{\mathcal{P},\mathsf{crs}\leftarrow\{0,1\}^{r(|x|)}}[\mathcal{V}(x, \mathsf{crs}|_{\mathcal{I}}, \pi) = 1 : (\pi, \mathcal{I}) \leftarrow \mathcal{P}(x, w, \mathsf{crs})] = 1 \ ,$$

   where $\mathcal{I} \subseteq [r(|x|)]$ and $\mathsf{crs}|_{\mathcal{I}} = \{(i, \mathsf{crs}[i]) : i \in \mathcal{I}\}$.
2. *Soundness: for every $x \notin \mathcal{L}$ we have that:*

$$\Pr_{\mathsf{crs}\leftarrow\{0,1\}^{r(|x|)}}[\exists \pi, \mathcal{I} : \mathcal{V}(x, \mathsf{crs}|_{\mathcal{I}}, \pi) = 1] < 2^{-n} \ .$$

3. *Zero-Knowledge: there exists a PPT algorithm $\mathcal{S}$ such that:*

$$\left\{(\mathsf{crs}|_{\mathcal{I}}, \pi) : \mathsf{crs} \leftarrow U_{r(|x|)}, (\pi, \mathcal{I}) \leftarrow \mathcal{P}(x, w, \mathsf{crs})\right\}_{(x,w)\in\mathcal{R}_{\mathcal{L}}} \approx_c \{\mathcal{S}(x)\}_{(x,w)\in\mathcal{R}_{\mathcal{L}}} \ .$$

Next we construct a NIZK proof in the CRS model.

Construction 4.1 (NIZK in the CRS Model). *We make use of the following primitives:*

- *A selectively secure invariant signature scheme $(\mathsf{Gen}, \mathsf{Sign}, \mathsf{Ver})$ with an invariant predicate P. For security parameter n, let $r_\sigma = r_\sigma(n)$ be the length of the CRS, and let $k = k(n)$ be the length of the verification key.*
- *A NIZK proof system $(\mathcal{P}_{\mathsf{hb}}, \mathcal{V}_{\mathsf{hb}})$ in the hidden-bits model with hidden random string of length $r = r(n)$.*

The NIZK system $(\mathcal{P}, \mathcal{V})$ in the CRS model is defined as follows:

The CRS. *The common random string is of length* $r_\sigma + k \cdot r \cdot (n+1)$. *The first* r_σ *bits of the CRS are interpreted as a CRS for the signature scheme* crs_σ. *We think of the rest of the CRS as divided into* $k \cdot r$ *blocks, each of length* $n+1$. *For every* $i \in [k], j \in [r]$, *we think of the* (i,j)*-th block as divided into a message* $m_{i,j} \in \{0,1\}^n$ *and a one-time pad bit* $s_{i,j} \in \{0,1\}$.

The prover $\mathcal{P}$. *Given* $(x,w) \in \mathcal{R}_\mathcal{L}$, *and the CRS* $(\mathsf{crs}_\sigma, \{m_{i,j}, s_{i,j}\})$ $\mathcal{P}$

1. *samples a pair of keys* $(\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(1^n)$,
2. *computes the strings* $\{\widetilde{\mathsf{crs}}_i \in \{0,1\}^r : i \in [k]\}$ *such that* $\widetilde{\mathsf{crs}}_i[j] = P(\sigma_{i,j}) \oplus s_{i,j}$ *and* $\sigma_{i,j} = \mathsf{Sign}_{\mathsf{sk}}(m_{i,j})$,
3. *for* $i \in [k]$, *emulates* $\mathcal{P}_{\mathsf{hb}}(x, w, \widetilde{\mathsf{crs}}_i)$ *and obtains a proof string* π_i *and a set of indices* $\mathcal{I}_i$,
4. *outputs a proof that contains the verification key* vk *and the hidden-bits proofs* $\{\pi_i, \Sigma_i : i \in [k]\}$, *where* $\Sigma_i = \{(j, \sigma_{i,j}) : j \in \mathcal{I}_i\}$.

The verifier $\mathcal{V}$. *Given* x, *the CRS* $(\mathsf{crs}_\sigma, \{m_{i,j}, s_{i,j}\})$, *and a proof* $(\mathsf{vk}, \{\pi_i, \Sigma_i\})$, $\mathcal{V}$

1. *for every* $i \in [k], j \in [r]$ *such that* $(j, \sigma_{i,j}) \in \Sigma_i$ *verifies that* $\mathsf{Ver}_{\mathsf{vk}}(\mathsf{crs}_\sigma, m_{i,j}, \sigma_{i,j}) = 1$; *otherwise,* $\mathcal{V}$ *rejects,*
2. *for* $i \in [k]$, *computes the set* $\widetilde{\mathsf{crs}}_i(\Sigma_i) = \{(j, P(\sigma_{i,j}) \oplus s_{i,j}) : (j, \sigma_{i,j}) \in \Sigma_i\}$,
3. *for each* $i \in [k]$, *emulates* $\mathcal{V}_{\mathsf{hb}}(x, \mathsf{crs}_i(\Sigma_i), \pi_i)$,
4. *accepts iff the emulation of* $\mathcal{V}_{\mathsf{hb}}$ *accepts for every* i.

Proposition 4.1. *The protocol given by Construction 4.1 is a NIZK proof in the CRS model.*

Proof. The completeness property of $(\mathcal{P}, \mathcal{V})$ follows from the completeness of $(\mathcal{P}_{\mathsf{hb}}, \mathcal{V}_{\mathsf{hb}})$ by construction. Next we prove the soundness and zero-knowledge properties.

Soundness. Fix some $x \in \{0,1\}^n \setminus \mathcal{L}$. Let $(\mathsf{crs}_\sigma, \{m_{i,j}, s_{i,j} : i \in [k], j \in [r]\})$ be uniform random variables describing the content of the CRS. Let $\{\widetilde{\mathsf{crs}}_i : i \in [k]\}$ be the set of hidden random strings for the protocol $(\mathcal{P}_{\mathsf{hb}}, \mathcal{V}_{\mathsf{hb}})$ computed by the honest prover $\mathcal{P}$ from $\{m_{i,j}, s_{i,j}\}$.

We prove that with overwhelming probability over the CRS, there is no proof $(\mathsf{vk}', \{\pi_i', \Sigma_i'\})$ that will make $\mathcal{V}$ accept. The uniqueness property of the signature holds with overwhelming probability over the crs_σ; from hereon, we condition on this event. Fix some $i \in [k]$ and a verification key vk'. Recall that for every $i \in [k], j \in [r]$, we have that $\widetilde{\mathsf{crs}}_i[j] = P(\sigma_{i,j}) \oplus s_{i,j}$. By the uniqueness property of the signature, the value of $P(\sigma_{i,j})$ is determined by the CRS of the signature crs_σ, the verification key vk and the messages $\{m_{i,j}\}$ and is independent of the pad bits $\{s_{i,j}\}$. It follows that $\widetilde{\mathsf{crs}}_i$ is uniformly distributed.

Let $\mathcal{I}_i' \subseteq [r]$ be a set of indices such that Σ_i' is of the form $\Sigma_i' = \left\{(j, \sigma_{i,j}') : j \in \mathcal{I}_i'\right\}$ for some signatures $\left\{\sigma_{i,j}'\right\}$. By the uniqueness property of the signature we have that if Σ_i' contains an element $(j, \sigma_{i,j}')$ such that $\widetilde{\mathsf{crs}}_i[j] \neq P(\sigma_{i,j}) \oplus s_{i,j}$ the verifier $\mathcal{V}$ rejects the proof. Therefore, if $\mathcal{V}$ accepts, it must be that $\widetilde{\mathsf{crs}}_i(\Sigma_i') = \widetilde{\mathsf{crs}}_i|_{\mathcal{I}_i'}$, where:

$$\widetilde{\mathsf{crs}}_i(\Sigma'_i) = \left\{(j, \sigma'_{i,j} \oplus s_{i,j}) : (j, \sigma'_{i,j}) \in \Sigma'_i\right\} \ ,$$
$$\widetilde{\mathsf{crs}}'_i|_{\mathcal{I}'_i} = \left\{(j, \widetilde{\mathsf{crs}}'_i[j]) : j \in \mathcal{I}'_i\right\} \ .$$

It follows that:

$$\Pr_{m_{i,j}, s_{i,j}}[\exists \pi'_i, \Sigma'_i : \mathcal{V}_{\mathsf{hb}}(x, \widetilde{\mathsf{crs}}_i(\Sigma'_i), \pi'_i) = 1] = \Pr_{\widetilde{\mathsf{crs}}_i \leftarrow U_r}[\exists \pi'_i, \mathcal{I}'_i : \mathcal{V}_{\mathsf{hb}}(x, \widetilde{\mathsf{crs}}_i|_{\mathcal{I}'_i}, \pi'_i) = 1] \ .$$

By the soundness of $(\mathcal{P}_{\mathsf{hb}}, \mathcal{V}_{\mathsf{hb}})$ we have that the above probability is at most 2^{-n}. Since this is true independently for every i and since $\mathcal{V}$ accepts iff all k executions of $\mathcal{V}_{\mathsf{hb}}$ accept, we have that:

$$\Pr_{m_{i,j}, s_{i,j}}[\exists \left\{\pi'_i, \Sigma'_i\right\} : \mathcal{V}_{\mathsf{hb}}(x, \left\{m_{i,j}, s_{i,j}\right\}, (\mathsf{vk}', \left\{\pi'_i, \Sigma'_i\right\})) = 1] \leq 2^{-n \cdot k} \ .$$

Since there are at most 2^k verification keys, by a union bound:

$$\Pr_{m_{i,j}, s_{i,j}}[\exists \mathsf{vk}', \left\{\pi'_i, \Sigma'_i\right\} : \mathcal{V}_{\mathsf{hb}}(x, \left\{m_{i,j}, s_{i,j}\right\}, (\mathsf{vk}', \left\{\pi'_i, \Sigma'_i\right\})) = 1] \leq 2^{-n} \ ,$$

as required.

Zero-Knowledge. We start be describing the simulator $\mathcal{S}$.

1. $\mathcal{S}$ is given as input a statement $x \in \mathcal{L}$ of length n.
2. For every $i \in [k]$, execute the simulator $\mathcal{S}_{\mathsf{hb}}$ of the protocol $(\mathcal{P}_{\mathsf{hb}}, \mathcal{V}_{\mathsf{hb}})$ and obtain:
$$(B_i, \pi_i) \leftarrow \mathcal{S}(x) \ ,$$
where $B_i = \{(j, b_{i,j}) : j \in \mathcal{I}_i\}$ for some set $\mathcal{I}_i \subseteq [r]$ and bits $\{b_{i,j}\}$.
3. Sample $\mathsf{crs}_\sigma \leftarrow U_{r_\sigma(n)}$ and $(\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs}_\sigma)$.
4. For every $i \in [k]$, $j \in [r]$ sample $m_{i,j} \leftarrow U_n$
5. For every $i \in [k]$, $j \in [r]$ if $j \notin \mathcal{I}_i$ sample $s_{i,j} \leftarrow U_1$, otherwise set:
$$s_{i,j} = P(\mathsf{Sign}_{\mathsf{sk}}(m_{i,j})) \oplus b_i \ .$$
6. Output the CRS $(\mathsf{crs}_\sigma, \{(m_{i,j}, s_{i,j}) : i \in [k], j \in [r]\})$. Output a simulated proof containing the verification key vk and the simulated hidden-bits proofs $\{\pi_i, \Sigma_i : i \in [k]\}$ where
$$\Sigma_i = \{(j, \mathsf{Sign}_{\mathsf{sk}}(m_{i,j})) : j \in \mathcal{I}_i\} \ .$$

Next we prove that the output of the simulator is indistinguishable from an honestly generated proof. For $0 \leq i \leq k$ consider the experiment H_i where for every $i' \leq i$ the messages and pad bits $\{(m_{i',j}, s_{i',j}) : j \in [r]\}$ are chosen uniformly, and the hidden-bits proof $(\pi_{i'}, \Sigma_{i'})$ is computed following the honest prover strategy, and for every $i < i'$ they are computed according to the simulated strategy as above. For every $(x, w) \in \mathcal{R}_{\mathcal{L}}$ we have that:

$$H_0(x, w) \approx \mathcal{S}(x) \ ,$$
$$H_{k(|x|)}(x, w) \approx \left\{(\mathsf{crs}, \mathcal{P}(x, w, \mathsf{crs})) : m_{i,j} \leftarrow U_{|x|}, s_{i,j} \leftarrow U_1\right\} \ .$$

Therefore, the correctness of the simulation follows from the next claim.

Claim. For every polysize distinguisher D, there exists a negligible function μ such that for every $(x,w) \in \mathcal{R}_{\mathcal{L}}$ and for every $i \in [k(|x|)]$

$$|\Pr[D(H_{i-1}(x,w)) = 1] - \Pr[D(H_i(x,w)) = 1]| \leq \mu(|x|) \ .$$

Proof. For $i \in [k(|x|)]$, consider the experiment H'_i that is defined just like H_i except that instead of sampling $(B_i, \pi_i) \leftarrow \mathcal{S}(x)$, we do the following:

1. Sample a random string $\widetilde{\mathsf{crs}}_i \leftarrow U_{r(|x|)}$.
2. Emulate $\mathcal{P}_{\mathsf{hb}}(x, w, \widetilde{\mathsf{crs}}_i)$ and obtain the proof string π_i and the set of indices $\mathcal{I}_i$.
3. Set $B_i = \widetilde{\mathsf{crs}}_i|_{\mathcal{I}_i} = \{(j, \widetilde{\mathsf{crs}}_i[j]) : j \in \mathcal{I}_i\}$.

By the zero-knowledge property of $(\mathcal{P}_{\mathsf{hb}}, \mathcal{V}_{\mathsf{hb}})$:

$$\left\{(\widetilde{\mathsf{crs}}_i|_{\mathcal{I}_i}, \pi_i) : \widetilde{\mathsf{crs}}_i \leftarrow U_{r(|x|)}, (\pi_i, \mathcal{I}_i) \leftarrow \mathcal{P}_{\mathsf{hb}}(x, w, \widetilde{\mathsf{crs}}_i)\right\}_{(x,w)\in\mathcal{R}_{\mathcal{L}}} \approx_c \{\mathcal{S}(x)\}_{(x,w)\in\mathcal{R}_{\mathcal{L}}} \ ,$$

and therefore, for every $(x,w) \in \mathcal{R}_{\mathcal{L}}$ and for every $i \in [k(|x|)]$:

$$|\Pr[D(H'_i(x,w)) = 1] - \Pr[D(H_i(x,w)) = 1]| \leq \mathrm{negl}(|x|) \ . \tag{1}$$

For every $0 \leq j \leq r(|x|)$ consider the experiment $H'_{i,j}$ that is defined just like H'_i except that for all $j' \leq j$ we set:

$$s_{i,j'} = P(\mathsf{Sign}_{\mathsf{sk}}(m_{i,j'})) \oplus \widetilde{\mathsf{crs}}_i[j'] \ . \tag{2}$$

For $j' > j$ choose $s_{i,j'}$ as in the experiment H'_i. That is, if $j' \in \mathcal{I}_i$ we set $s_{i,j'}$ as in 2 and if $j' \notin \mathcal{I}_i$ we sample $s_{i,j'}$ uniformly.

Note that the output distribution of the experiments $H'_{i,j-1}$ and $H'_{i,j}$ may differ when $j \notin \mathcal{I}_i$. This is due to the fact that conditioned on $j \notin \mathcal{I}_i$, the bit $\widetilde{\mathsf{crs}}_i[j]$ may no longer be uniform. However, based on the pseudo-randomness property of the signature we will show that the experiments are computationally indistinguishable.

Claim. For every polysize distinguisher D, there exists a negligible function μ such that for every $(x,w) \in \mathcal{R}_{\mathcal{L}}$ and for every $i \in [k(|x|)], j \in [r(|x|)]$:

$$\left|\Pr[D(H'_{i,j-1}(x,w)) = 1] - \Pr[D(H'_{i,j}(x,w)) = 1]\right| < \mu(|x|) \ .$$

Proof. Assume towards contradiction that there is a distinguisher D and a polynomial p such that for infinitely many $(x,w) \in \mathcal{R}_{\mathcal{L}}$ there exist $i \in [k(|x|)], j \in [r(|x|)]$ such that:

$$\left|\Pr[D(H'_{i,j-1}(x,w)) = 1] - \Pr[D(H'_{i,j}(x,w)) = 1]\right| > \frac{1}{p(|x|)} \ . \tag{3}$$

We construct a distinguisher $\tilde{D}$ that breaks the pseudo-randomness property of the signature. That is for infinity many values of n:

$$\left| \begin{array}{l} \Pr\left[\begin{array}{l} m \leftarrow U_n, (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs}_\sigma), \mathsf{crs}_\sigma \leftarrow U_{r_\sigma(n)}, b = P(\mathsf{Sign}_{\mathsf{sk}}(m)) : \\ \tilde{D}^{\mathsf{Sign}_{\mathsf{sk}}(\cdot)}(\mathsf{crs}_\sigma, \mathsf{vk}, m, b) = 1 \end{array}\right] - \\ \Pr\left[\begin{array}{l} m \leftarrow U_n, (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs}_\sigma), \mathsf{crs}_\sigma \leftarrow U_{r_\sigma(n)}, b \leftarrow U_1 : \\ \tilde{D}^{\mathsf{Sign}_{\mathsf{sk}}(\cdot)}(\mathsf{crs}_\sigma, \mathsf{vk}, m, b) = 1 \end{array}\right] \end{array} \right| > \frac{1}{p(|x|)} \ , \tag{4}$$

where $\tilde{D}$ never queries its oracle on m. $\tilde{D}$ will have hardcoded $(x, w) \in \mathcal{R}_{\mathcal{L}}$ and i, j for which (3) holds. Then $\tilde{D}(\mathsf{crs}_\sigma, \mathsf{vk}, m, b)$ emulates $H'_{i,j}(x, w)$ with the following modifications:

1. When the experiment $H'_{i,j}(x, w)$ samples crs_σ and vk, D uses its input crs_σ and vk instead.
2. Every time a the emulation needs to sign a message $\tilde{D}$ forwards the message to the signing oracle (note that the experiment $H'_{i,j}(x, w)$ does not use the secret key sk except for signing messages).
3. If $j \notin \mathcal{I}_i$ set $s_{i,j} = b \oplus \widetilde{\mathsf{crs}}_i[j]$.

We have that:

$$\Pr[D(H'_{i,j}(x,w)) = 1] = \Pr\left[\begin{array}{l} m \leftarrow U_n, (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs}_\sigma), \mathsf{crs}_\sigma \leftarrow U_{r_\sigma(n)}, b = P(\mathsf{Sign}_{\mathsf{sk}}(m)) : \\ \tilde{D}^{\mathsf{Sign}_{\mathsf{sk}}(\cdot)}(\mathsf{crs}_\sigma, \mathsf{vk}, m, b) = 1 \end{array}\right] ,$$

$$\Pr[D(H'_{i,j-1}(x,w)) = 1] = \Pr\left[\begin{array}{l} m \leftarrow U_n, (\mathsf{sk}, \mathsf{vk}) \leftarrow \mathsf{Gen}(\mathsf{crs}_\sigma), \mathsf{crs}_\sigma \leftarrow U_{r_\sigma(n)}, b \leftarrow U_1 : \\ \tilde{D}^{\mathsf{Sign}_{\mathsf{sk}}(\cdot)}(\mathsf{crs}_\sigma, \mathsf{vk}, m, b) = 1 \end{array}\right] ,$$

and therefore, (4) follows from (3) and we get a contradiction to the pseudorandomness property of the signature.

The experiment $H'_{i,0}$ is identical to the experiment H'_i by definition. It follows from Claim 4 that for every $(x, w) \in \mathcal{R}_{\mathcal{L}}$ and for every $i \in [k(|x|)]$:

$$\left|\Pr[D(H'_i(x,w)) = 1] - \Pr[D(H'_{i,r(|x|)}(x,w)) = 1]\right| \leq \mathrm{negl}(|x|) \ . \tag{5}$$

Note that for $i \in [k]$, the experiment H_{i-1} is identical to the experiment $H'_{i,r(|x|)}$ except for the order in which the the pad bits $\{s_{i,j}\}$ and the random hidden string $\widetilde{\mathsf{crs}}_i$ are sampled. Specifically, in the experiment H_{i-1}:

1. First sample $m_{i,j} \leftarrow U_{|x|}, s_{i,j} \leftarrow U_1$, for every $j \in [r(|x|)]$.
2. Then compute $\widetilde{\mathsf{crs}}_i$ where $\widetilde{\mathsf{crs}}_i[j] = P(\mathsf{Sign}_{\mathsf{sk}}(m_{i,j})) \oplus s_{i,j}$.

Since in both experiments H_{i-1} and $H'_{i,r(|x|)}$ we have that $\widetilde{\mathsf{crs}}_i$ is uniform and $m_{i,j}, s_{i,j}$ are uniform conditioned on the fact that:

$$\widetilde{\mathsf{crs}}_i[j] = P(\mathsf{Sign}_{\mathsf{sk}}(m_{i,j})) \oplus s_{i,j} \ ,$$

we have that the experiments H_{i-1} and $H'_{i,r(|x|)}$ are identical. Combining this with (1) and (5) we get that for every $(x, w) \in \mathcal{R}_{\mathcal{L}}$ and for every $i \in [k(|x|)]$:

$$|\Pr[D(H_{i-1}(x,w)) = 1] - \Pr[D(H_i(x,w)) = 1]| \leq \mathrm{negl}(|x|) \ ,$$

as required.

5 Non-Interactive Witness-Indistinguishability

In this section, we construct a NIWI proof system based on indistinguishability obfuscation and one-way permutations.

Theorem 5.1. *Assuming iO for* $\mathbf{P/poly}$ *and one-way permutations, there exist NIWI proof for every language in* $\mathbf{NP}$.[5]

We now describe the NIWI system yielding the theorem. A high-level overview of the construction and the main ideas behind it are provided in the introduction.

Primitives and Notation. The construction relies on an indistinguishability obfuscator $i\mathcal{O}$, a ZAP system (that can be constructed from $i\mathcal{O}$ and OWFs as in Section 4), and a non-interactive (one message) statistically binding commitment Com. We require that Com is dense, in the sense that every string of appropriate length is a valid commitment to some message. Such a commitment can be constructed from one-way permutations [Blu81].

Let $\mathcal{L}$ be any $\mathbf{NP}$ language. For every candidate instance $x \in \{0,1\}^n$ and message $m \in \{0,1\}^n$, denote by E_x^m the canonical "witness-encryption" circuit that given any $w \in \mathcal{R}_{\mathcal{L}}(x)$ outputs m and otherwise outputs $\perp$. Let $\mathcal{T}$ be the $\mathbf{NP}$ language containing instances of the form $(x, \mathsf{c}_1, \mathsf{c}_2, \tilde{E})$ where x is candidate instance for $\mathcal{L}$, $\mathsf{c}_1, \mathsf{c}_2$ are commitments, and $\tilde{E}$ is an obfuscation such that at least one of the following conditions holds:

1. $\tilde{E}$ is a valid obfuscation of a witness-encryption circuit. That is, there exist randomness r and a message m such that $\tilde{E} = i\mathcal{O}(E_x^m; r)$.
2. $\tilde{E}$ has the same output on the plaintexts underlying the commitments $\mathsf{c}_1, \mathsf{c}_2$. That is, there exist decommitments (w_1, r_1) and (w_2, r_2) such that:

$$\mathsf{c}_1 = \mathsf{Com}(w_1; r_1) \quad \wedge \quad \mathsf{c}_2 = \mathsf{Com}(w_2; r_2) \quad \wedge \quad \tilde{E}(w_1) = \tilde{E}(w_2) \ .$$

Finally, let $D_{x,w}^{s,\mathsf{c}_1,\mathsf{c}_2}$ be a "witness-decryption" circuit as described in Figure 3.

Hardwired:
1. Instance and witness $(x, w) \in \mathcal{R}_{\mathcal{L}}$,
2. first ZAP message s,
3. commitments $\mathsf{c}_1, \mathsf{c}_2$.

Input:
1. A circuit $\tilde{E}$,
2. second ZAP message π.

Output:
1. Verify that π is a valid proof for the statement $(x, \mathsf{c}_1, \mathsf{c}_2, \tilde{E}) \in \mathcal{T}$ with respect to the first message s. If not, output $\perp$.
2. Output $\tilde{E}(w)$.

Fig. 3. The "witness-decryption" circuit $D_{x,w}^{s,\mathsf{c}_1,\mathsf{c}_2}$

[5] We assume iO for all circuits for simplicity of exposition; naturally, it suffices to have iO for a certain restricted class of circuits that we use in our construction and analysis.

Construction 5.2 (NIWI Proof). *The NIWI system $(\mathcal{P}, \mathcal{V})$ is defined as follows:*

The prover $\mathcal{P}$ *given* $x \in \{0,1\}^n \cap \mathcal{L}$ *and* $w \in \mathcal{R}_{\mathcal{L}}(x)$*:*

1. *Sample a first ZAP message* $s \in \{0,1\}^{\mathrm{poly}(n)}$,
2. *compute a pair of commitments to the all zero string* $\mathsf{c}_1, \mathsf{c}_2 \leftarrow \mathsf{Com}(0^{|w|})$,
3. *compute the obfuscation* $\tilde{D} \leftarrow i\mathcal{O}(D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w})$,
4. *output* $(s, \mathsf{c}_1, \mathsf{c}_2, \tilde{D})$ *as the proof.*

The verifier $\mathcal{V}$ *given* x *and the proof* $(s, \mathsf{c}_1, \mathsf{c}_2, \tilde{D})$*:*

1. *Sample a message* $m \leftarrow \{0,1\}^n$,
2. *compute the obfuscation* $\tilde{E} \leftarrow i\mathcal{O}(E^m_x)$,
3. *compute a proof* π *for the statement* $(\mathsf{c}_1, \mathsf{c}_2, \tilde{E}) \in \mathcal{T}$ *with respect to the first message* s*. Use* m *and the randomness used to compute* $\tilde{E}$ *as a witness for the fact that* $\tilde{E}$ *is a valid obfuscation of a witness-encryption circuit,*
4. *accept if* $m = \tilde{D}(\tilde{E}, \pi)$ *accept, otherwise reject.*

Proposition 5.1. *The protocol given by Construction 5.2 is a NIWI proof.*

Proof. The completeness of the system follows readily from the completeness of the ZAP and the functionality of $i\mathcal{O}$. We focus on proving soundness and then witness-indistinguishability.

Soundness. Assume towards contradiction that there exist a polynomial p such that for infinitely many $x \notin \mathcal{L}$ there exists a proof $(s, \mathsf{c}_1, \mathsf{c}_2, \tilde{D})$ such that:

$$\Pr[\mathcal{V}(x, (s, \mathsf{c}_1, \mathsf{c}_2, \tilde{D})) = 1] \geq \frac{1}{p(|x|)} \ .$$

Let m be the random message sampled by $\mathcal{V}$ in a random execution, and let $(\tilde{E}, \pi)$ be the obfuscation and proof computed by $\mathcal{V}$. By our assumption:

$$\Pr[\tilde{D}(\tilde{E}, \pi) = m] \geq \frac{1}{p(|x|)} \ .$$

Let (w_1, r_1) and (w_2, r_2) be decommitments of $\mathsf{c}_1, \mathsf{c}_2$ respectively (such decommitments exist since Com is dense). Since $x \notin \mathcal{L}$, and by the (perfect) functionality of $i\mathcal{O}$, the circuit $\tilde{E}$ outputs $\perp$ on all inputs. Therefore, the decommitments $(w_1, r_1), (w_2, r_2)$ can be used as a witness for the statement $(\mathsf{c}_1, \mathsf{c}_2, \tilde{E}) \in \mathcal{T}$. Let π' be a proof for the statement $(\mathsf{c}_1, \mathsf{c}_2, \tilde{E}) \in \mathcal{T}$ with respect to the first message s computed using the witness $(w_1, r_1), (w_2, r_2)$. By the witness indistinguishability of the ZAP: $\pi \approx_c \pi'$. Therefore,

$$\Pr[\tilde{D}(\tilde{E}, \pi') = m] \geq \frac{1}{p(|x|)} - \mathrm{negl}(|x|) \ .$$

Let $\tilde{E}'' = i\mathcal{O}(E^{0^n}_x)$. Since the circuits $\tilde{E}$ and $\tilde{E}''$ are of the same size, and since both output $\perp$ on all inputs, it follows from the security of $i\mathcal{O}$ that $\tilde{E} \approx_c \tilde{E}''$. Let π'' be a proof with respect to $\tilde{E}''$ rather than for $\tilde{E}$. Then $(\tilde{E}, \pi') \approx_c$

$(\tilde{E}'', \pi'')$. Indeed, a distinguisher between $(\tilde{E}, \pi'), (\tilde{E}'', \pi'')$ can be reduced to a distinguisher between $\tilde{E}, \tilde{E}''$, since computing π' and π'' does not require the randomness underlying $\tilde{E}_1, \tilde{E}_2$). Hence, it also holds that:

$$\Pr[\tilde{D}(\tilde{E}'', \pi'') = m] \geq \frac{1}{p(|x|)} - \mathrm{negl}(|x|) \ .$$

Since m is uniform in $\{0,1\}^{|x|}$ and in the above experiment the view of $\tilde{D}$ is independent of m, we get a contradiction.

Witness Indistinguishability. Let $\mathcal{I} = \{(x, w_1, w_2) : w_1, w_2 \in \mathcal{R}_{\mathcal{L}}(x)\}$, be a sequence of instances $x \in \mathcal{L}$, with two corresponding witnesses w_1, w_2. We show that

$$\left\{(s, \mathsf{c}_1, \mathsf{c}_2, \tilde{D}) \leftarrow \mathcal{P}(x, w_1))\right\}_{(x,w_1,w_2)\in\mathcal{I}} \approx_c \left\{(s, \mathsf{c}_1, \mathsf{c}_2, \tilde{D}) \leftarrow \mathcal{P}(x, w_2))\right\}_{(x,w_1,w_2)\in\mathcal{I}} \ ,$$

by considering a sequence of hybrid distributions.

Hyb_1**:** Here $(\mathsf{c}_1, \mathsf{c}_2, \tilde{D}) \leftarrow \mathcal{P}(x, w_1)$ corresponds to a proof using the first witness w_1.

Hyb_2**:** Here for each $b \in \{0,1\}$, $\mathsf{c}_b \leftarrow \mathsf{Com}(w_b)$ is a commitment to the corresponding witness rather than to the all-zero string. By the computational-hiding of Com, $\mathsf{Hyb}_2 \approx_c \mathsf{Hyb}_1$.

Hyb_3**:** Here the first ZAP message s is sampled conditioned the on the ZAP being absolutely sound; that is, there exists no accepting proof, with respect to s, for any false statement. By the soundness of the ZAP, this holds with overwhelming probability and thus $\mathsf{Hyb}_3 \approx_s \mathsf{Hyb}_2$.

Hyb_4**:** Here instead of sampling $\tilde{D} \leftarrow i\mathcal{O}(D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_1})$ using w_1, it is sampled using w_2, i.e., $\tilde{D} \leftarrow i\mathcal{O}(D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_2})$. To show that $\mathsf{Hyb}_4 \approx_c \mathsf{Hyb}_5$, we show that for any realization of $s, \mathsf{c}_1, \mathsf{c}_2$ (which have the same distribution in $\mathsf{Hyb}_4, \mathsf{Hyb}_5$), the two circuits $D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_1}, D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_2}$ have the exact same functionality and thus, by the iO guarantee, $i\mathcal{O}(D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_1}) \approx_c i\mathcal{O}(D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_2})$. Indeed, for any input $(\tilde{E}, \pi)$ for $D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_b}$, there are two options:

1. π is not a valid proof for the statement $(x, \mathsf{c}_1, \mathsf{c}_2, \tilde{E}) \in \mathcal{T}$ with respect to the first message s. In this case, by the definition of $D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_b}$, it holds that $D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_1}(\tilde{E}, \pi) = D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_2}(\tilde{E}, \pi) = \bot$.
2. π is a valid proof. In this case, by the soundness of the ZAP, $(x, \mathsf{c}_1, \mathsf{c}_2, \tilde{E}) \in \mathcal{T}$. This in turn implies one of two cases
   (a) $\tilde{E}$ is a valid obfuscation $i\mathcal{O}(E^m_x)$, in which case by the definition of E^m_x, and the functionality of $i\mathcal{O}$, $\tilde{E}(w_1) = \tilde{E}(w_2)$.
   (b) $\mathsf{c}_1, \mathsf{c}_2$ can be opened to $\tilde{w}_1, \tilde{w}_2$, such that $\tilde{E}(\tilde{w}_1) = \tilde{E}(\tilde{w}_2)$, in which case by the binding of Com, for both $b \in \{0,1\}$, $w_b = \tilde{w}_b$, and thus also $\tilde{E}(w_1) = \tilde{E}(w_2)$.

So in either case $D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_1}(\tilde{E}, \pi) = D^{s,\mathsf{c}_1,\mathsf{c}_2}_{x,w_2}(\tilde{E}, \pi)$, as required.

Hyb_5**:** Here we remove the requirement that s is sampled conditioned on absolute soundness. Like before, it holds that $\mathsf{Hyb}_5 \approx_s \mathsf{Hyb}_4$ by the soundness of the ZAP.

Hyb_6**:** Here $(\mathsf{c}_1, \mathsf{c}_2, \tilde{D}) \leftarrow \mathcal{P}(x, w_2)$ corresponds to a proof using the first witness w_2. This hybrid differs from Hyb_5 only in that $\mathsf{c}_1, \mathsf{c}_2$ are commitments to all-zero strings rather than to w_1, w_2. Like before, it holds that $\mathsf{Hyb}_6 \approx_c \mathsf{Hyb}_5$ by the computational hiding of the commitment Com.

Remark 5.1 (Relying on relaxed dense commitments). The non-interactive commitment scheme used in the NIWI construction can be somewhat relaxed. Indeed, it suffices to require a non-interactive commitment scheme that is statistically binding, but only against *honest* committers; namely, honestly generated commitments can only be opened to a single value. The commitment should still be *dense* in the sense that every string in the range of the commitment, can be opened to at least one value (commitments that are not generated honestly may potentially be opened to more than one value).

Remark 5.2 (Using witness-encryption generically). We note that we refrain from explicitly defining witness encryption, and in the above construction, directly implement witness encryption using iO (which we anyhow rely on). While we find that thinking about witness encryption in terms of obfuscation is helpful in this context, it is possible to state the construction in terms of generic witness encryption.

Acknowledgements. We thank Ran Canetti for discussions and valuable advice. We thank Rafail Ostrovsky for discussions on the the way that unbalanced properties are dealt with in [GO92], and for referring us to [BGRV09]. We also thank Sanjam Garg for discussing NIZKs based on graded encodings.

References

[BFM88] Blum, M., Feldman, P., Micali, S.: Non-interactive zero-knowledge and its applications (extended abstract). In: STOC, pp. 103–112 (1988)

[BGI+01] Barak, B., Goldreich, O., Impagliazzo, R., Rudich, S., Sahai, A., Vadhan, S.P., Yang, K.: On the (Im)possibility of obfuscating programs. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 1–18. CRYPTO, Heidelberg (2001)

[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)

[BGK+13] Barak, B., Garg, S., Kalai, Y.T., Paneth, O., Sahai, A.: Protecting obfuscation against algebraic attacks. Cryptology ePrint Archive, Report 2013/631 (2013), `http://eprint.iacr.org/`

[BGRV09] Brakerski, Z., Goldwasser, S., Rothblum, G.N., Vaikuntanathan, V.: Weak verifiable random functions. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 558–576. Springer, Heidelberg (2009)

[Blu81] Blum, M.: Coin flipping by telephone. In: Proceedings of the 18th Annual International Cryptology Conference, pp. 11–15 (1981)

[BOV07] Barak, B., Ong, S.J., Vadhan, S.P.: Derandomization in cryptography. SIAM J. Comput. 37(2), 380–400 (2007)

[BR13] Brakerski, Z., Rothblum, G.N.: Virtual black-box obfuscation for all circuits via generic graded encoding. Cryptology ePrint Archive, Report 2013/563 (2013), http://eprint.iacr.org/

[BW13] Boneh, D., Waters, B.: Constrained pseudorandom functions and their applications. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013, Part II. LNCS, vol. 8270, pp. 280–300. Springer, Heidelberg (2013)

[BY96] Bellare, M., Yung, M.: Certifying permutations: Noninteractive zero-knowledge based on any trapdoor permutation. J. Cryptology 9(3), 149–166 (1996)

[CLT13] Coron, J.-S., Lepoint, T., Tibouchi, M.: Practical multilinear maps over the integers. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 476–493. Springer, Heidelberg (2013)

[DN07] Dwork, C., Naor, M.: Zaps and their applications. SIAM J. Comput. 36(6), 1513–1543 (2007)

[FLS99] Feige, U., Lapidot, D., Shamir, A.: Multiple noninteractive zero knowledge proofs under general assumptions. SIAM J. Comput. 29(1), 1–28 (1999)

[FS89] Feige, U., Shamir, A.: Zero knowledge proofs of knowledge in two rounds. In: Brassard, G. (ed.) CRYPTO 1989. LNCS, vol. 435, pp. 526–544. Springer, Heidelberg (1990)

[GGH13a] Garg, S., Gentry, C., Halevi, S.: Candidate multilinear maps from ideal lattices. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 1–17. Springer, Heidelberg (2013)

[GGH$^+$13b] Garg, S., Gentry, C., Halevi, S., Raykova, M., Sahai, A., Waters, B.: Candidate indistinguishability obfuscation and functional encryption for all circuits. In: FOCS (2013)

[GGH14] Gentry, C., Gorbunov, S., Halevi, S.: Graded multilinear maps from lattices. Cryptology ePrint Archive, Report 2014/645 (2014), http://eprint.iacr.org/

[GGHR14] Garg, S., Gentry, C., Halevi, S., Raykova, M.: Two-round secure MPC from indistinguishability obfuscation. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 74–94. Springer, Heidelberg (2014)

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

[GGSW13] Garg, S., Gentry, C., Sahai, A., Waters, B.: Witness encryption and its applications. In: STOC, pp. 467–476 (2013)

[GL89] Goldreich, O., Levin, L.A.: A hard-core predicate for all one-way functions. In: STOC 1989: Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing, pp. 25–32. ACM, New York (1989)

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

[GMW91] Goldreich, O., Micali, S., Wigderson, A.: Proofs that yield nothing but their validity for all languages in np have zero-knowledge proof systems. J. ACM 38(3), 691–729 (1991)

[GO92] Goldwasser, S., Ostrovsky, R.: Invariant signatures and non-interactive zero-knowledge proofs are equivalent. In: Brickell, E.F. (ed.) CRYPTO 1992. LNCS, vol. 740, pp. 228–245. Springer, Heidelberg (1993)

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

[GOS12] Groth, J., Ostrovsky, R., Sahai, A.: New techniques for noninteractive zero-knowledge. J. ACM 59(3), 11 (2012)

[KMN+14] Komargodski, I., Moran, T., Naor, M., Pass, R., Rosen, A., Yogev, E.: One-way functions and (im)perfect obfuscation. IACR Cryptology ePrint Archive, 2014:347 (2014)

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

[Nao91] Naor, M.: Bit commitment using pseudorandomness. J. Cryptology 4(2), 151–158 (1991)

[NLLT14] Niu, Q., Li, H., Liang, B., Tang, F.: One-round witness indistinguishability from indistinguishability obfuscation. IACR Cryptology ePrint Archive, 2014:176 (2014)

[SW14] Sahai, A., Waters, B.: How to use indistinguishability obfuscation: Deniable encryption, and more. In: STOC (2014)