J. Cryptology (2003) 16: 95–116
DOI: 10.1007/s00145-002-0113-0

Journal of
CRYPTOLOGY

Trading Help for Interaction in Statistical Zero-Knowledge Proofs*

Michael Ben-Or and Dan Gutfreund
School of Computer Science and Engineering,
Hebrew University, Jerusalem, Israel
{benor,danig}@cs.huji.ac.il

Communicated by Oded Goldreich

Received May 2001 and revised April 2002
Online publication 3 March 2003

Abstract. We define interactive and non-interactive statistical zero-knowledge proofs with (limited) help, as proofs that can be almost perfectly simulated, where the prover and the verifier share a reference string that is computed by a probabilistic polynomial-time trusted third party that receives as input the statement to be proven (i.e. the input to the protocol). We compare these models with the standard interactive and non-interactive SZK models, trying to understand when this form of help can replace the interaction between the prover and the verifier and vice versa. We show that every promise problem that has an SZK protocol with help also has one without help. As for the opposite, we show non-interactive SZK proofs with help for natural languages for which only interactive SZK proofs are known. In order to achieve that, we introduce a complete problem for the class of promise problems that have non-interactive SZK proofs with help.

Key words. Zero-knowledge, Non-interactive zero-knowledge, Graph isomorphism.

1. Introduction

When zero-knowledge proofs were first introduced by Goldwasser et al. [9], it seemed that interaction plays a crucial role in those proof systems. Indeed zero-knowledge was shown to exist only for languages in *BPP* in the most straightforward non-interactive model [5]. Blum et al. showed, however, that if we change the model, then non-interactive zero-knowledge can be achieved for languages not known to be trivial [3]. In their model (called the random reference string model) they assume that both prover and verifier are dealt a truly random (uniformly distributed) string, called the reference string. The

* This research was supported in part by the Leibniz Center, a US–Israel Binational research grant, and an EU Information Technologies grant (IST-FP5).

actual proof consists of one message sent from the prover to the verifier, and then the verifier decides whether to accept or reject according to this message, the input and the reference string.

Just like their interactive counterparts, non-interactive zero-knowledge proofs come in three flavors: perfect, statistical and computational zero-knowledge. While extensive study of computational non-interactive zero-knowledge (and some study of perfect non-interactive zero-knowledge) dates back to the late 80s, the study of statistical non-interactive zero-knowledge was initiated only recently by De Santis et al. [4], who showed a complete problem for the class of promise problems that have non-interactive statistical zero-knowledge proofs (hereafter denoted *NISZK*). They were followed by Goldreich et al. [7], who gave more complete problems for this class, as well as conditions under which the class *NISZK* equals the class of promise problems that have (interactive) statistical zero-knowledge proofs (hereafter denoted *SZK*). We do not know if these conditions are true, and the question of whether interaction is necessary for statistical zero-knowledge (i.e. whether *NISZK* equals *SZK*) is still an open one.

In this paper we suggest an alternative resource to interaction in the form of limited help (limited in terms of the computational complexity of the helper) that is given in advance, both to the prover and the verifier. Specifically, we assume that the prover and the verifier have access to a shared string, which is the output of a probabilistic polynomial-time Turing Machine that is given as input the statement to be proven (i.e. the input to the protocol). We stress that the assumption that the prover and the verifier have access to a reference string that is dealt by a trusted third party also appears in the model of [3], that is, in their model it is assumed that the reference string is uniformly distributed. The main difference in our model is that the reference string may depend on the input to the protocol.

First, we show that interaction can always replace (limited) help. That is, every promise problem that has a statistical zero-knowledge proof with help, whether interactive or not, also has a (interactive) statistical zero-knowledge proof without help. As for the converse, we show that in some cases help can replace interaction.[1] We do that by first introducing a complete problem for the class of promise problems that have non-interactive statistical zero-knowledge proofs with help (hereafter denoted $NISZK|_{\mathrm{h}}$). Then, armed with this complete problem, we show non-interactive statistical zero-knowledge proofs with help for two natural problems (Graph Isomorphism and Graph Non-Isomorphism) which are not known to be in *NISZK*.

Our results should be seen in the context of the general question of what resources are necessary for statistical zero-knowledge proofs to work, and, more specifically, whether $SZK = NISZK$. We hope that our work will help to shed light on these fundamental questions (refer to Section 5 for further discussion).

In [2] we also considered non-interactive *computational* zero-knowledge proofs with help. We showed that under the assumption that one-way functions exist, every language in *NP* has such a proof system with perfect completeness and perfect soundness. Furthermore, in this proof system the dealer need not have any access to the input to the protocol. Thus the only difference between this protocol and the standard NIZK model

[1] In [2] the authors claimed that help can always replace interaction. Unfortunately, due to a gap that was found in the proof, we retract here from this claim.

is the fact that the dealer may induce a distribution which is not uniform. In this paper we only consider the statistical zero-knowledge setting.

2. Definitions and Notations

2.1. *Probability Distributions*

Throughout this paper we consider distributions with "succinct" description, i.e. distributions produced by circuits (with multiple output gates) when feeding them a uniformly chosen input. We write C when we refer both to the circuit itself and to the distribution it induces. We denote by $\mathrm{Im}(C)$ the image set (or the range) of the function that the circuit C computes.

Given a distribution X, $x \leftarrow X$ denotes that x is a sample taken from X.

For a set S, $s \in_{\mathrm{R}} S$ denotes that s is a sample taken from the uniform distribution over S.

Recall the definition of Shannon's entropy.

Definition 1. Let X be a random variable, we define the entropy of X, $H(X)$, to be

$$H(X) = \sum_x \Pr(X = x) \log(1/\Pr(X = x)) = E_{x \leftarrow X}[\log(1/\Pr(X = x))].$$

We define measures of distance (or similarity) between distributions. The first measure is the statistical difference.

Definition 2. Let X and Y be two distributions (or random variables) on a discrete space D. The statistical difference between X and Y, denoted as $\|X - Y\|$, is

$$\begin{aligned} \|X - Y\| &= MAX_{S \subseteq D} |\Pr(X \in S) - \Pr(Y \in S)| \\ &= \frac{1}{2} \sum_{d \in D} |\Pr(X = d) - \Pr(Y = d)|. \end{aligned}$$

The second measure is the Kullback–Liebler distance, or the relative entropy.

Definition 3. Let X and Y be two distributions on a finite space D. The relative entropy (or the Kullback–Liebler distance) between X and Y is

$$KL(X \mid Y) = E_{x \leftarrow X}\left[\log \frac{\Pr(X = x)}{\Pr(Y = x)}\right].$$

We denote by $H_2(p)$ the binary entropy function, which is the entropy of a 0-1 random variable with expectation p. $KL_2(p, q)$ denotes the relative entropy between two 0-1 random variables with expectations p and q.

We also define measures of "disjointness" between distributions.

Definition 4. We say that a (ordered) pair of distributions, (C_0, C_1), is α-disjoint if for $x \leftarrow C_0$, $\Pr(x \in \mathrm{Im}(C_1)) \leq 1 - \alpha$. We say that the pair (C_0, C_1) is mutually α-disjoint if (C_0, C_1) and (C_1, C_0) are α-disjoint.[2]

2.2. *Promise Problems*

A promise problem Π is defined by two disjoint sets $\Pi_Y, \Pi_N \subseteq \{0, 1\}^*$. We say that the elements of Π_Y are the YES instances of Π and the elements of Π_N are the NO instances.

For two promise problems Γ and Π we define their AND to be the following promise problem:

$$AND(\Gamma, \Pi)_Y = \{(x, y): x \in \Gamma_Y \text{ and } y \in \Pi_Y\},$$
$$AND(\Gamma, \Pi)_N = \{(x, y): x \in \Gamma_N \text{ or } y \in \Pi_N\}.$$

2.3. *Interaction between Turing Machines*

In what follows we will be talking about interaction between probabilistic Turing machines (TM) and the outcome of their interaction. Typically the outcome will be "accept" or "reject". Given two such TM, M_1 and M_2, and an input x to the protocol between them, $(M_1, M_2)(x)$ denotes the random variable of the possible outcome values of the protocol, and $\langle M_1, M_2\rangle(x)$ denotes the distribution over all the possible conversation transcripts (where the probability space is over the random coins of M_1 and M_2). We also denote by $\langle M_1, M_2\rangle(x)_i$ the distribution over the prefix of the conversation transcripts up to the ith message.

2.4. *Complexity Classes*

Recall that μ is a negligible function, if for every $c > 0$ and sufficiently large n, $\mu(n) < 1/n^c$.

We now state the definition of (an honest verifier) statistical zero-knowledge proof system [9].

Definition 5. A statistical zero-knowledge proof system for a promise problem Π, is defined by a computationally unbounded interactive TM P (the prover), a probabilistic polynomial-time interactive TM V (the verifier), a probabilistic polynomial-time TM S (the simulator), and a negligible function μ. A proof system for the membership of an input x in Π, consists of messages that are exchanged between P and V, where the last message is V's random coins. After the interaction ends, V decides whether to accept or reject. The following should hold:

1. (completeness) if $x \in \Pi_Y$, then $\Pr((P, V)(x) = accept) > 2/3$,
2. (soundness) if $x \in \Pi_N$, then for every prover's strategy P^*, $\Pr((P^*, V)(x) = accept) < 1/3$,

[2] Note that if (C_0, C_1) are 1-disjoint, then they are completely disjoint.

3. (zero-knowledge) if $x \in \Pi_Y$, then the statistical difference (see Definition 2) between the following two distributions is bounded by $\mu(|x|)$:
   (a) $\langle P, V\rangle(x)$,
   (b) $S(x)$.

Remark. We only give here the definition of honest verifier statistical zero-knowledge and we limit the discussion to this model due to a result by Goldreich et al. [6], who showed that this model is equivalent to the dishonest verifier model.

We now define non-interactive statistical zero-knowledge with a random reference string [3].

Definition 6. A non-interactive statistical zero-knowledge proof system with a random reference string for a promise problem Π, is defined by a computationally unbounded TM P (the prover), a probabilistic polynomial-time TM V (the verifier), a probabilistic polynomial-time TM S (the simulator), a polynomial q, and a negligible function μ. On an input x both P and V have access to a shared random reference string σ, where $\sigma \in_R \{0, 1\}^{q(|x|)}$. The actual proof consists of one message that is sent from P to V, and then V based on x, σ and this message either accepts or rejects. The following should hold:

1. (completeness) if $x \in \Pi_Y$, then $\Pr(V(x, \sigma, P(x, \sigma)) = accept) > 2/3$,
2. (soundness) if $x \in \Pi_N$, then for every prover's strategy P^*, $\Pr(V(x, \sigma, P^*(x, \sigma)) = accept) < 1/3$,
3. (zero-knowledge) if $x \in \Pi_Y$, then the following two distributions have a statistical difference that is bounded by $\mu(|x|)$:
   (a) $(\sigma, P(x, \sigma))$,
   (b) $S(x)$.

We define statistical zero-knowledge proofs *with help* as in Definition 5, with the only difference that in addition to the input, the prover and the verifier have access to a shared random string which is distributed according to a polynomial-time samplable distribution that may depend on the input.

Definition 7. A statistical zero-knowledge proof system with help for a promise problem Π, is defined by a computationally unbounded interactive TM P (the prover), a probabilistic polynomial-time interactive TM V (the verifier), a probabilistic polynomial-time TM S (the simulator), a probabilistic polynomial-time TM D (the dealer), and a negligible function μ. On an input x, P and V have access to a shared reference string $D(x)$. The proof system consists of messages that are exchanged between P and V, where the last message is V's random coins. After the interaction ends, V decides, according to the input, the reference string and the conversation, whether to accept or reject. The following should hold:

1. (completeness) if $x \in \Pi_Y$, then $\Pr((D, P, V)(x) = accept) > 2/3$,
2. (soundness) if $x \in \Pi_N$, then for every prover's strategy P^*, $\Pr((D, P^*, V)(x) = accept) < 1/3$,

3. (zero-knowledge) if $x \in \Pi_Y$, then the statistical difference between the following two distributions is bounded by $\mu(|x|)$:
   (a) $\langle D, P, V\rangle(x)$,
   (b) $S(x)$.

Definition 8. Non-interactive statistical zero-knowledge with help is defined exactly as non-interactive statistical zero-knowledge with a random reference string, with the only difference that we replace the uniformly distributed reference string σ with the reference string $D(x)$, where x is the input to the protocol, and D is a probabilistic polynomial-time TM.

Remark. Note that Definition 6 is a special case of Definition 8, where $D(x)$ is defined to be $\sigma \in_R \{0,1\}^{q(|x|)}$.

We denote by *SZK* the class of promise problems that have statistical zero-knowledge proofs, by $SZK|_h$ the class of promise problems that have statistical zero-knowledge proofs with help, by *NISZK* the class of promise problems that have non-interactive statistical zero-knowledge proofs in the random reference string model, and by $NISZK|_h$ the class of promise problems that have non-interactive statistical zero-knowledge proofs with help.

2.5. *Complete Problems*

The following two promise problems are complete for the class *SZK* [12], [8]:

Definition 9. Statistical Difference (SD) is the following promise problem:

$$SD_Y = \{(C_0, C_1): \|C_0 - C_1\| > 2/3\},$$
$$SD_N = \{(C_0, C_1): \|C_0 - C_1\| < 1/3\},$$

where (C_0, C_1) is a pair of distributions with "succinct" description.

Definition 10. Entropy Difference (ED) is the following promise problem:

$$ED_Y = \{(C_0, C_1): H(C_0) > H(C_1) + 1\},$$
$$ED_N = \{(C_0, C_1): H(C_0) < H(C_1) - 1\},$$

where (C_0, C_1) is a pair of distributions with "succinct" description.

3. Interaction Can Always Replace Help

In this section we generalize [2], by showing that if we have interaction in hand, then polynomial-time help is not needed. That is, we show that every promise problem that has a statistical zero-knowledge proof (whether interactive or not) with help, also has such a (interactive) proof with no help (i.e. an SZK proof in the traditional model).

Theorem 11. $SZK|_{\mathrm{h}} = SZK$.

Our proof is based on the reduction of Goldreich and Vadhan [8] from every promise problem in *SZK* to the *SZK*-complete problem *ED*.

Intuition. We consider a statistical zero-knowledge proof for a promise problem Π and its corresponding simulator S. We look at the output of the simulator as describing the moves of a virtual prover and a virtual verifier. Following Aiello and Hastad [1], we consider a cheating strategy for a real prover P_S, called the simulation-based prover, which tries to imitate the behavior of the virtual prover. Informally, P_S determines its messages based on the same distribution as the virtual prover, conditioned only on past messages. Let us compare the distributions over the output of the simulator, and the output of the interaction between P_S and the real verifier. On YES instances these two distributions should be relatively close. This is because the behavior of P_S is similar to the behavior of the honest prover, and the output of the simulator is similar to the output of the conversation between the honest prover and the verifier. On NO instances on the other hand, if the simulator outputs "accept" with high probability (we can easily modify the simulator to ensure that), then there must be a fundamental difference between the output of the simulator and the output of the interaction between P_S and the real verifier, because the latter cannot be accepting with high probability. Aiello and Hastad considered the relative entropy between the two distributions as a measure of similarity. They showed how to write this relative entropy as a simple expression involving entropies of prefixes of the simulator's output. Goldreich and Vadhan gave upper and lower bounds for this expression in cases of YES instances and NO instances, respectively. They then showed how to derive from that a reduction from Π to *ED*.

Ideally, we would like to show a similar result for $SZK|_{\mathrm{h}}$, by directly applying this reduction on promise problems in this class. However, in order to do so, we need to change the protocol with the dealer to one without a dealer. One option is to let the verifier play the dealer's role. This will indeed change the proof system into an interactive proof with two players. However, the zero-knowledge condition is not guaranteed anymore as the verifier can "see" now the dealer's private coins. Another option is to let the prover play the dealer's role, this can clearly affect the soundness condition, as the help is not given now by a trusted third party. However, deviations from the dealer's protocol can be tested in zero-knowledge, and it is that approach that we take. Specifically, we consider the simulation-based prover P_S as in [8], and we let it send the first message instead of the dealer (according to the distribution of the virtual dealer). In order to assure that the soundness condition still holds, we add a test to check whether the prover behaves "like" the dealer. We know that if it behaves completely different, then the test will detect it. Otherwise, the soundness condition still holds (maybe with a slight loss) and we can apply the arguments of [8].

Notation. Let (D, P, V) be a statistical zero-knowledge proof system with help for a promise problem Π, and let S be its corresponding simulator. We assume that on input of length n, the verifier tosses $l = l(n)$ coins. Including the dealer's message, $2r$ messages are exchanged ($r = r(n)$), where the verifier is the first to send a message after the dealer,

and each message is of length l. We also assume that the last message is the contents of the verifier's random tape.

The simulation-based prover. We describe now the behavior of the simulation-based prover P_S. For an odd i ($1 \leq i \leq 2r$), given a conversation prefix $\gamma \in \{0,1\}^{(i-1)l}$, the next message of P_S (i.e. the ith message of the protocol) will be:

1. If the probability that $S(x)$ outputs a conversation with the prefix γ is 0, then P_S sends a dummy message, say 0^l.
2. Otherwise, P_S replies with the same conditional probability as the virtual prover. That is, it sends $\beta \in \{0,1\}^l$ with probability

$$\Pr[S(x)_i = \gamma\beta \mid S(x)_{i-1} = \gamma].$$

In particular, P_S sends the first message instead of the dealer, and this message is distributed exactly as the first string in the output of the simulator, which is the simulation of the dealer.

Rewriting the relative entropy between $S(x)$ and $\langle P_S, V\rangle(x)$. The following lemma shows how we can express the relative entropy between $S(x)$ and $\langle P_S, V\rangle(x)$ as differences between the entropies of prefixes of the simulator's output. Refer to [8] for the proof.

Lemma 12. $\forall x \in \Pi_Y \cup \Pi_N$, $KL(S(x) \mid \langle P_S, V\rangle(x)) = l - \sum_{i=1}^{r}[H(S(x)_{2i}) - H(S(x)_{2i-1})]$.

Bounding $KL(S \mid \langle P_S, V\rangle)$. We now bound $KL(S \mid \langle P_S, V\rangle)$ on YES instances and NO instances. For YES instances, we use the following lemma that was proved in [8].

Lemma 13. *Let $x \in \Pi_Y$, and let $\varepsilon = \|S(x) - \langle P, V\rangle(x)\|$, then*

$$KL(S(x) \mid \langle P_S, V\rangle(x)) \leq 3r^2 l\varepsilon + 2r H_2(\varepsilon).$$

Next we bound $KL(S \mid \langle P_S, V\rangle)$ on NO instances. We will need the following intuitive lemma. It states that statistical zero-knowledge proofs with help are robust under small deviations from the dealer's protocol. The proof of this lemma is easy and we defer it to Appendix B.

Lemma 14. *Let $s = s(n)$ be the soundness error of the proof system (D, P, V). For $\varepsilon = \varepsilon(n)$, let D' be an arbitrary dealer for which $\|D(x) - D'(x)\| \leq \varepsilon(|x|)$. Denote by $s' = s'(n)$ the soundness error of the protocol in which D is replaced by D'. The following holds: $s' \leq s + 2\varepsilon$.*

We can now state the bound for NO instances. We confine ourselves to the case when the distribution over the first string in the output of the simulator (and hence the first message of P_S) is statistically close to the dealer's distribution.

Lemma 15. *For $x \in \Pi_N$, let p denote the probability that $S(x)$ outputs an accepting transcript. Suppose that $\|D(x) - S(x)_1\| \leq q_1$. Denote by $q_2 = q_2(|x|)$ the soundness of the protocol. Let $q = 2q_1 + q_2$, and suppose that $p \geq q$. Then*

$$KL(S(x) \mid \langle P_S, V\rangle(x)) \geq KL_2(p, q).$$

Proof. Recall that $P_S(x)_1$ and $S(x)_1$ are identically distributed. So by Lemma 14, the soundness error of the protocol $(P_S, V)(x)$ is not more than $2q_1 + q_2 = q$. Let q' be the probability that $(P_S, V)(x)$ accepts, clearly, $q' \leq q$. Define the function $f\colon \{0,1\}^{2rl} \to \{0,1\}$ as follows: $f(\gamma) = 1$ if γ is an accepting transcript and 0 otherwise. Then by Facts A4 and A3 (in Appendix A), we get

$$\begin{aligned} KL(S(x) \mid \langle P_S, V\rangle(x)) &\geq KL(f(S(x)) \mid f(\langle P_S, V\rangle(x))) \\ &= KL_2(p, q') \geq KL_2(p, q). \end{aligned}$$

□

The reduction. Let us assume without loss of generality that the soundness and completeness errors of the protocol for Π are bounded by $(2rl)^{-2}/2$, and the simulator deviation is bounded by $(2rl)^{-2}/2$.

We modify the proof system such that 0^{2rl} is an accepting transcript, and modify the simulator always to output accepting transcripts (by possibly substituting the output with 0^{2rl}). The resulting proof system has soundness error at most $2^{-l} + (2rl)^{-2}/2$ (recall that the last message consists of the verifier's random coins), and the simulator deviation is at most $(2rl)^{-2}$.

We are now ready to present the reduction. It will map Π to the *AND* of two promise problems regarding pairs of distributions with succinct description. Thus, an instance x is mapped into two pairs of distributions, as follows:

1. (a) $X_{1,x}$ is the cross product of the distributions $S(x)_2, S(x)_4, \ldots, S(x)_{2r}$.
   (b) $Y_{1,x}$ is the cross product of the distributions $S(x)_1, S(x)_3, \ldots, S(x)_{2r-1}$ and the uniform distribution over $\{0,1\}^{l(|x|)-2}$.
2. (a) $X_{2,x}$ is the distribution $D(x)$.
   (b) $Y_{2,x}$ is the distribution $S(x)_1$ (i.e. the simulation of the dealer's message).

Motivation. The statistical difference between $X_{2,x}$ and $Y_{2,x}$ measures how much P_S deviates from the dealer's protocol. If this deviation is small enough (and this can be tested in zero-knowledge), then Lemma 15 can be applied. Assuming that this is the case, we look now at $(X_{1,x}, Y_{1,x})$. By Lemma 12, without the last $l-2$ bits of $Y_{1,x}$, the difference between their entropies would be exactly $KL(S(x) \mid \langle P_S, V\rangle(x)) - l$. Thus, the addition of uniformly distributed $l-2$ bits brings this difference to within 2 of $KL(S(x) \mid \langle P_S, V\rangle(x))$. Lemmas 13 and 15 give us upper and lower bounds for this term in case of YES and NO instances, respectively. By setting the right parameters, we can have this term below 1 or above 3 for YES and NO instances, respectively. Thus we get that the difference in entropies, between $X_{1,x}$ and $Y_{1,x}$, is below -1 or above 1 as required by the definition of *ED*.

We now prove this formally.

Lemma 16. *For $x \in \Pi_Y$, $H(X_{1,x}) > H(Y_{1,x}) + 1$ and $\|X_{2,x} - Y_{2,x}\| < (2rl)^{-2}$.*

Proof. The fact that the simulator's deviation is bounded by $(2rl)^{-2}$ implies that $\|D(x) - S(x)_1\| < (2rl)^{-2}$, which in particular establishes the second part.

Next, assume without loss of generality that $rl > 128$ (by padding messages with extra bits if necessary). Define ε to be the simulator's deviation ($\varepsilon < (2rl)^{-2}$). By Lemmas 12 and 13, we have

$$\begin{aligned} H(Y_{1,x}) - H(X_{1,x}) &= \left(l - 2 + \sum_{i=1}^{r} H(S(x)_{2i-1})\right) - \left(\sum_{i=1}^{r} H(S(x)_{2i})\right) \\ &= KL(S(x) \mid \langle P_S, V\rangle(x)) - 2 \\ &\leq 3r^2 l\varepsilon + 2r H_2(\varepsilon) - 2 < -1, \end{aligned}$$

where the last inequality uses $H_2(\varepsilon) \leq \sqrt{\varepsilon}/4$ (since $\varepsilon < 2^{-14}$) and $\sqrt{\varepsilon}/4 < 1/8r$. □

Lemma 17. *For $x \in \Pi_N$, either $H(Y_{1,x}) > H(X_{1,x}) + 1$ or $\|X_{2,x} - Y_{2,x}\| > (2rl)^{-1}$.*

Proof. If $\|D(x) - S(x)_1\| > (2rl)^{-1}$, we are done. Otherwise, the simulation-based prover that sends the first message instead of the dealer does not deviate from the dealer's protocol by more than $(2rl)^{-1}$. Therefore, assuming (without loss of generality) that $2^{-l} + (2rl)^{-2}/2 + (rl)^{-1} < 0.1$, we can apply Lemma 15 with $q_1 = (2rl)^{-1}$, $q_2 = 2^{-l} + (2rl)^{-2}/2$ and $p = 1$, and together with Lemma 12 we get,

$$\begin{aligned} H(Y_{1,x}) - H(X_{1,x}) &= \left(l - 2 + \sum_{i=1}^{r} H(S(x)_{2i-1})\right) - \left(\sum_{i=1}^{r} H(S(x)_{2i})\right) \\ &= KL(S(x) \mid \langle P_S, V\rangle(x)) - 2 \\ &\geq KL_2(1, 0.1) - 2 \\ &= \log 10 - 2 > 1. \end{aligned}$$

□

Claim 18. $\Pi \in SZK$.

Proof. Let m be the description length of the pair of circuits $(X_{2,x}, Y_{2,x})$. Clearly, $2r(|x|)l(|x|)$ is polynomial in m, let $p(m)$ denote this polynomial. Define the promise problem,

$$\begin{aligned} SD'_Y &= \left\{(C_0, C_1)\colon \|C_0 - C_1\| < \frac{1}{p(m)^2}\right\}, \\ SD'_N &= \left\{(C_0, C_1)\colon \|C_0 - C_1\| > \frac{1}{p(m)}\right\}, \end{aligned}$$

where C_0, C_1 are distributions with succinct description and m is their description length.

By the Polarization Lemma of [13] and the closure of *SZK* under complementation [11], $SD' \in SZK$. By Lemmas 16 and 17, the reduction above is polynomial-time many–one from Π to $AND(ED, SD')$. Since *SZK* is closed under *AND* (and polynomial-time many–one reductions), we conclude that $\Pi \in SZK$. □

Since Π is an arbitrary problem in $SZK|_{\mathrm{h}}$, Theorem 11 follows.

An alternative proof. Theorem 11 has an alternative proof that was suggested to us by the anonymous referee. It is somewhat more technically involved but its underlying idea is very appealing. We give here a rough sketch of this proof. We refer the reader to [8] for the tools and terminology that we use below.

The idea is to replace the dealer by a protocol that P and V execute. The outcome of this protocol can be either V accepts/rejects or V outputs a string y. If the latter happens, then P and V continue the original protocol with y as a reference string. The protocol has the properties that if P and V are honest, then either (with probability close to $1/2$) V accepts or it outputs y that is distributed statistically close to the dealer's distribution. On the other hand, no matter how P behaves, y cannot hit with high probability any fixed set that the dealer does not hit with high probability (unless there is high probability that V rejects). Furthermore, the protocol has a simulation that is statistically close to the real conversation.[3] It can be shown that with these properties the new protocol has arbitrarily small (with a security parameter) completeness error, soundness error that is bounded by a constant, and a simulation that has negligible statistical deviation from the real conversation.

The protocol is based on the sample generation and test protocols of [8], and it has a similar structure to the protocol for *ED*. First, we take the original protocol and run it many times in parallel (and take a majority vote for acceptance). The effect of this is that the new dealer's (the one that outputs many reference strings) distribution, D', is statistically close to a "flat" distribution. This does not change the zero-knowledge property of the protocol, because the verifier is honest. Then P and V execute the sample generation protocol to obtain a sample, w, from D' that is not too "heavy" (i.e. the number of its pre-images is not too large). Next, they execute the hash-based "complementary sampling" to obtain a pair (y, r) such that r is a pre-image of w, and y is not too "light" (i.e. the number of its pre-images is not too small). Finally, V flips a coin and then either (with probability $1/2$) P and V continue the original protocol with y as the reference string, or they execute the sample test protocol on y (and the distribution D').[4] The proof techniques are similar to those of [8] (although the claims and the choice of parameters are somewhat different).

[3] We actually need something stronger, and that is the strong zero-knowledge property of the sample generation protocol from [8].

[4] The reason that we cannot run the sample test protocol and then continue the original proof with y as a reference string is that a pre-image of y is revealed to V during the sample test protocol. This may affect the zero-knowledge property, because in the original protocol, V does not see a pre-image of the reference string.

4. The Class $NISZK|_{\mathrm{h}}$

4.1. *A Complete Problem*

In this section we show a complete problem for the class $NISZK|_{\mathrm{h}}$.

Definition 19. Image Intersection Density (IID) is the following promise problem:

$$IID_{\mathrm{Y}} = \left\{(D_0, D_1): \|D_0 - D_1\| < \frac{1}{n^2}\right\},$$

$$IID_{\mathrm{N}} = \left\{(D_0, D_1): \text{ the pair } (D_0, D_1) \text{ is } \left(1 - \frac{1}{n^2}\right)\text{-disjoint}\right\},$$

where (D_0, D_1) is a pair of distributions with succinct description, and n is their description length.

Note that the *NISZK*-complete problem *ID* [4] is a special case of *IID*,[5] where D_0 is taken to be the uniform distribution over $\{0, 1\}^n$. On the other hand, *IID* is a restriction of the *SZK*-complete problem $\overline{SD}$ (that is, for NO instances, the distributions are "disjoint" and not only statistically far).

The main technical step towards proving that *IID* is complete for the class $NISZK|_{\mathrm{h}}$ is the following lemma. It can be viewed as an analogue to the Polarization Lemma of [13], in which the disjointness property is maintained. We defer the proof to Appendix C.

Lemma 20. *There is a polynomial-time procedure that takes a pair of distributions* (C_0, C_1) *with description length* n, *and outputs a pair of distributions* (D_0, D_1) *such that*

$$\|C_0 - C_1\| < \frac{1}{n^2} \Rightarrow \|D_0 - D_1\| < 2^{-n},$$

$$(C_0, C_1) \textit{ is } \left(1 - \frac{1}{n^2}\right)\textit{-disjoint} \Rightarrow (D_0, D_1) \textit{ is } (1 - 2^{-n})\textit{-disjoint}.$$

We can now show that *IID* is complete for $NISZK|_{\mathrm{h}}$. The proof is based on ideas from [4].

Theorem 21. $IID \in NISZK|_{\mathrm{h}}$.

Proof. Given a pair of circuits (C_0, C_1) with description length n, we first apply the procedure from Lemma 20 to get a pair (D_0, D_1) which either has statistical difference

[5] This is not exactly accurate. In [4] the bound, in the definition of *ID*, on the statistical difference and the disjointness, is given by some negligible function. However, setting a particular function is problematic. The reason is that the reduction to *ID* uses a negligible function that is specific to the *NISZK* problem of the instance on which we apply the reduction. This function is arbitrary and is possibly larger than the fixed negligible function in the definition of *ID*. We overcome this by setting the bounds in the definition of *IID* to be an inverse of a fixed polynomial. We can do that because we have Lemma 20.

at most 2^{-n} or is $(1 - 2^{-n})$-disjoint. Let l be the number of input gates to D_0 and D_1 (without loss of generality they have the same number of input gates).

The proof system.
Common input: (D_0, D_1).
Shared reference string: $\sigma = D_0(r)$, where $r \in_R \{0, 1\}^l$.
The protocol:

1. P sends $r' \in_R D_1^{-1}(\sigma)$ (where $D_1^{-1}(\sigma) = \{r\colon\ D_1(r) = \sigma\}$).
2. V accepts if and only if $D_1(r') = \sigma$.

Completeness. If $(C_0, C_1) \in IID_Y$, then $\|D_0 - D_1\| < 2^{-n}$. Hence, by the definition of statistical difference, the probability that $x \leftarrow D_0$ is in $\mathrm{Im}(D_1)$ is at least $1 - 2^{-n}$. Therefore, the probability (over r) that the (computationally unbounded) prover will be able to find r' such that $D_1(r') = \sigma$ is at least $1 - 2^{-n}$.

Soundness. If $(C_0, C_1) \in IID_N$, then the pair (D_0, D_1) is $(1 - 2^{-n})$-disjoint. That is, the probability that σ is not in the image of D_1 is at least $1 - 2^{-n}$. We know that when this happens, there is no r' such that $D_1(r') = \sigma$, and V will reject.

Simulation. S: Choose $r \in_R \{0, 1\}^l$, output $(D_1(r), r)$.

If $(C_0, C_1) \in IID_Y$, then the distributions D_0 and D_1 have statistical difference at most 2^{-n}. Therefore, the statistical difference between $\sigma = D_0(r)$ (where $r \in_R \{0, 1\}^l$) and the first part of the simulator's output is 2^{-n} at the most. Given the first part, the second part, both in the protocol and the simulator, is selected at random from the set $\{r'\colon\ r' \in D_1^{-1}(\sigma)\}$, and thus has the same (conditional) distribution. □

Theorem 22. *IID is NISZK$|_h$-hard.*

Proof. Let (D, P, V) be an NISZK protocol with help for a promise problem Π, with exponentially vanishing completeness and soundness errors (this can be easily achieved by parallel repetitions). Let S be the simulator for the protocol. Define μ to be the negligible function bounding the statistical difference between $\langle D, P, V\rangle$ and the output of S, when they are given an input in Π_Y.

The reduction. For an input x, define the following pair of distributions:

1. D_0: $D(x)$.
2. D_1: run the simulator S on x to obtain (σ, p), if $V(x, \sigma, p) =$ "*accept*" output σ, otherwise output a special symbol $\perp$ (that D never outputs).

First, we show that if $x \in \Pi_Y$, then $(D_0, D_1) \in IID_Y$. Let $n = |x|$, and let m be the description length of D_0 and D_1. Clearly, $m = poly(n)$. The statistical difference between the first part of the simulator's output and the real reference string is bounded by $\mu(n)$. Also, $\Pr((D, P, V)(x) =$ "*accept*"$) > 1 - 2^{-n}$. Therefore, with probability $1 - 2^{-n} - \mu(n)$, D_1 will output the first part of the simulator's output, and thus $\|D_0 - D_1\| < 2\mu(n) + 2^{-n}$ (which for large enough n is less than $1/m^2$).

Next, we show that if $x \in \Pi_N$, then $(D_0, D_1) \in IID_N$.

Define: $T = \{\sigma\colon \sigma \in \mathrm{Im}(D_0)$, and $\exists p$ s.t $V(x, \sigma, p) =$ "*accept*"$\}$.

That is, T is the set of all reference strings for which there exist a proof that convinces V. Since the soundness error is bounded by 2^{-n}, $\Pr(D(x) \in T) < 2^{-n}$. As D_1 only outputs $\sigma \in T$ or $\perp$, and $D_0 = D(x)$ outputs a real reference string, the probability that a sample from D_0 will be in $\mathrm{Im}(D_1)$ is at most 2^{-n}. So for large enough n, the pair (D_0, D_1) is $(1 - 1/m^2)$-disjoint. □

4.2. *Cases Where Help Can Replace Interaction*

In this section we show evidence that help may replace interaction in some cases. We show that two languages, not known to be in *NISZK*, are in $NISZK|_h$. The languages that we consider are Graph Isomorphism (GI) and Graph Non-Isomorphism (GNI). We stress that the claims here apply to a wider class of languages that have similar properties, however, for simplicity and clarity we focus on these two languages.

Let S_n be the set of permutations over n elements. Given a (n-vertex) graph G and a permutation $\pi \in S_n$, the graph $\pi(G)$ is the graph G with the nodes permuted according to π.

Definition 23. $GI = \{(G_0, G_1)\colon \exists \pi \in S_n$ s.t. $\pi(G_0) = G_1\}$.

The language GNI is defined to be the complement of GI.

Theorem 24. $GI \in NISZK|_h$.

Proof. We show a reduction from *GI* to *IID*. Suppose that we have a way to sample uniformly from the set of strings that encode permutations in S_n (and we ignore for now the technical difficulty of how to achieve that). Given a pair of graphs (G_0, G_1), we construct a pair of circuits (C_0, C_1) as follows: C_0 (and C_1) receive as inputs permutations in S_n, where n is the number of nodes in G_0 (and G_1 without loss of generality). On an input $\pi \in S_n$, the circuit C_0 (resp. C_1) outputs $\pi(G_0)$ (resp. $\pi(G_1)$).

Clearly, if $(G_0, G_1) \in GI$, and the inputs are uniformly distributed over S_n, then $\|C_0 - C_1\| = 0$. On the other hand, if $(G_0, G_1) \notin GI$, then $\|C_0 - C_1\| = 1$, because otherwise, by composing permutations we get an isomorphism from G_0 to G_1.

Going back to the issue of choosing permutations uniformly, we define the number of input gates to C_0 and C_1 to be $m = n^3\lceil \log(n) \rceil$. We can view the inputs to the circuits as n blocks of n^2 numbers between 1 and n. We say that a string $s \in \{0, 1\}^m$ encodes the permutation $\pi \in S_n$, if for every $1 \le i \le n$, $\pi(i)$ is the first number in the ith block that is not equal to $\pi(1), \ldots, \pi(i-1)$. It is easy to verify that with probability 2^{-n} at the most, a string chosen uniformly from $\{0, 1\}^m$ does not encode a permutation in S_n. Also, every permutation has the same probability to come up. We now change C_0 and C_1 to output special (different) symbols $\perp_0$ and $\perp_1$, respectively, when the input does not encode a permutation in S_n, and to continue as before otherwise. Then if $(G_0, G_1) \notin GI$, we still have that $\|C_0 - C_1\| = 1$ and $(C_0, C_1) \in IID_N$. On the other hand, for $(G_0, G_1) \in GI$, on $\mathrm{Im}(C_0) \cap \mathrm{Im}(C_1)$ the circuits induce the same

distribution, and only with probability at most 2^{-n} is their output not in the intersection. Hence, $\|C_0 - C_1\| < 2^{-n}$ and $(C_0, C_1) \in IID_Y$. □

Theorem 25. *GNI* $\in$ *NISZK*$|_h$.

Proof. By the reduction above, *GNI* reduces to $\overline{IID}$. Unfortunately, we do not know a reduction from $\overline{IID}$ to *IID* (this would clearly finish the proof). However, we observe that the mapping in the proof of Theorem 24 is to a special case of $\overline{IID}$. We will show that this special case can be reduced to *IID*.

As before, we ignore for now the technical difficulty of uniformly sampling permutations (we will deal with this issue at the end of the proof), and assume that we have the perfect case described in the first part of the proof of Theorem 24. Note that the reduction above maps pairs of graphs to pairs of circuits that induce distributions that are either identical or completely disjoint. Furthermore, if (G_0, G_1) are isomorphic, the distributions are uniform over the domain of all the graphs that are isomorphic to G_0 (and G_1). We will use these properties to reduce *GNI* to *IID*.

The reduction. Sahai and Vadhan [13], showed a reduction from the promise problem SD to its complement. We show that this reduction also works in our case. Given a pair of graphs, (G_0, G_1), apply the reduction from the proof of Theorem 24 (assuming that we have the perfect sampling case from the first part of that proof) to get a pair of circuits (C_0, C_1) such that

$$(G_0, G_1) \in GNI \quad \Longrightarrow \quad \|C_0 - C_1\| = 1,$$
$$(G_0, G_1) \notin GNI \quad \Longrightarrow \quad \|C_0 - C_1\| = 0.$$

Let n be the number of nodes in G_0 and G_1 (without loss of generality they have the same number of nodes), let $q = \log(n!)$, let l be the number of output gates of C_0 and C_1, and let $m = n^3\lceil q \rceil^2$.

Define a new circuit $\bar{C}$: $\{0,1\}^m \times S_n^m \longrightarrow \{0,1\}^{ml}$:

$$\bar{C}(\bar{b}, \bar{\pi}) = (C_{b_1}(\pi_1), \ldots, C_{b_m}(\pi_m)).$$

We sample $\bar{C}$ by choosing, uniformly and independently, m bits, $b_1, \ldots, b_m$, and m permutations over n elements, $\pi_1, \ldots, \pi_m$, and computing m graphs, where the ith graph is $\pi_i(G_{b_i})$.

Let H be a 2-universal family of hash functions from $\{0,1\}^m \times S_n^m \times \{0,1\}^{ml}$ to $T = \{0,1\}^{\lceil (q+1)m - 2\Delta - n \rceil}$, where $\Delta = m/n$.

We can now describe the new circuits:

D_0: Let $(\bar{b}, \bar{\pi}) \in_R \{0,1\}^m \times S_n^m$, $\bar{y} \leftarrow \bar{C}$ and $h \in_R H$. Output $(\bar{C}(\bar{b}, \bar{\pi}), \bar{b}, h, h(\bar{b}, \bar{\pi}, \bar{y}))$.

D_1: Let $(\bar{b}, \bar{\pi}) \in_R \{0,1\}^m \times S_n^m$, $h \in_R H$ and $t \in_R T$. Output $(\bar{C}(\bar{b}, \bar{\pi}), \bar{b}, h, t)$.

Note. In the definition of D_0, $\bar{C}(\bar{b}, \bar{\pi})$ and $\bar{y}$ are two independent samples from the same distribution.

We claim that if $(G_0, G_1) \in GNI$, then (D_0, D_1) are statistically close, and if $(G_0, G_1) \notin GNI$, then (D_0, D_1) are almost disjoint.

Intuition. Assume that $\bar{C}$ is uniformly distributed (over some set). This is certainly true when G_0 and G_1 are isomorphic, and close enough for our arguments to work when they are not.[6] Then, each $\bar{z} \in \mathrm{Im}(\bar{C})$ has the same number, w, of pre-images. Thus, given a sample $\bar{z}$, taken from $\bar{C}$, there are w different possibilities for inputs that produced it. Each such input contains two parts, $\bar{b}$ and $\bar{\pi}$. What happens if together with $\bar{z}$ we also reveal the first part $\bar{b}$? When the graphs are non-isomorphic, we know that C_0 and C_1 are disjoint. So $\bar{b}$ is determined by $\bar{z}$ and revealing it does not add any new information. That is, there are still w different possibilities for inputs that produced $\bar{z}$. If, on the other hand, the graphs are isomorphic, then every one of the 2^m values of $\bar{b}$ is possible. So by revealing the one that was actually used, we reduce the number of possible inputs that produced $\bar{z}$ to $w/2^m$. We also know that the size of the image set of $\bar{C}$ is $|\{0,1\}^m \times S_n^m|/w = 2^{(q+1)m}/w$. So when $(G_0, G_1) \in GNI$, given the first three components of D_0 (which are identical to the first three components of D_1), there are exactly $w \cdot 2^{(q+1)m}/w = 2^{(q+1)m}$ possible inputs that h can take in the fourth component. This is much larger than the range of our hashing, $|T| = 2^{(q+1)m-2m/n-n}$. Therefore, by the Leftover Hash Lemma [10], the fourth component of D_0 will be almost uniformly distributed over T, and thus will be statistically close to the fourth component of D_1. On the other hand, when $(G_0, G_1) \notin GNI$, given the first three components, the number of inputs that h can take is $w/2^m \cdot 2^{(q+1)m}/w = 2^{qm}$. This is much smaller than $|T|$, and the fourth component of D_0 can only cover a small fraction of T. The probability that the fourth component of D_1 will fall in this fraction is small, and thus the distributions will be almost disjoint.

The case of $(G_0, G_1) \in GNI$. First we argue that $\bar{C}$ is close to uniform. Let $L = \{0,1\}^l$. For $\bar{z} \in L^m$, let $wt(\bar{z}) = \log(|\{(\bar{b}, \bar{\pi})\colon\ \bar{C}(\bar{b}, \bar{\pi}) = \bar{z}\}|)$ be the weight of $\bar{z}$. That is, the weight of $\bar{z}$ is the logarithm of the size of its pre-image set. Let w be the expected weight of an image of $\bar{C}$, $w = E_{\bar{z} \leftarrow \bar{C}}(wt(\bar{z}))$. Since $\bar{C}$ is composed of many independent and identically distributed random variables, we can apply a Chernoff argument and get that $\Pr_{\bar{z} \leftarrow \bar{C}}(|wt(\bar{z}) - w| > \Delta) < 2^{-\Omega(n)}$. More formally, for $z \in L$, let $wt_0(z) = \log(|\{(b, \pi)\colon\ C_b(\pi) = z\}|)$. Then for $\bar{z} \in L^m$, $wt(\bar{z}) = wt_0(z_1) + \cdots + wt_0(z_m)$. Note that $z_1, \ldots, z_m$ are independent and identically distributed. Furthermore, for any $z \in L$, $0 \le wt_0(z) \le \lceil q \rceil$. So by the Chernoff bound we have

$$\Pr_{\bar{z} \leftarrow \bar{C}}(|wt(\bar{z}) - w| > \Delta) < 2e^{-2\Delta^2/m\lceil q \rceil^2} = 2e^{-2n}.$$

We would like to consider only $\bar{z} \in L^m$ which have weight close to the mean. Therefore, we define the set of "good" inputs, $G = \{(\bar{b}, \bar{\pi})\colon\ |wt(\bar{C}(\bar{b}, \bar{\pi})) - w| \le \Delta\}$. We also define the distribution $\bar{C}'$, that samples uniformly $(\bar{b}, \bar{\pi}) \in G$ and outputs $\bar{C}(\bar{b}, \bar{\pi})$.[7] We

[6] Note that when G_0 and G_1 have automorphism groups of different sizes, then $\bar{C}$ is not uniformly distributed. However, as we will show, it is close to uniform.

[7] $\bar{C}'$ is defined for the sake of argument, it is not constructive because we do not know how to sample uniformly from G. However, it is much easier to prove our claims using this distribution. Furthermore, as we will show, it is enough to prove our claims on this distribution instead of $\bar{C}$, because the two distributions are statistically close.

can now define a new pair of distributions, (D'_0, D'_1), which are similar to (D_0, D_1), with the only difference that they sample inputs uniformly from G instead of $\{0,1\}^m \times S_n^m$:

D'_0: Let $(\bar{b}, \bar{\pi}) \in_R G$, $\bar{y} \leftarrow \bar{C}'$ and $h \in_R H$. Output $(\bar{C}(\bar{b}, \bar{\pi}), \bar{b}, h, h(\bar{b}, \bar{\pi}, \bar{y}))$.
D'_1: Let $(\bar{b}, \bar{\pi}) \in_R G$, $h \in_R H$ and $t \in_R T$. Output $(\bar{C}(\bar{b}, \bar{\pi}), \bar{b}, h, t)$.

We claim that in order to finish the proof, it is enough to prove that D'_0 and D'_1 are statistically close. By the argument above, $|G| \geq (1 - 2^{-\Omega(n)})|\{0,1\}^m \times S_n^m|$. Let $U(D)$ denote the uniform distribution over some set D. Then $\|U(G) - U(\{0,1\}^m \times S_n^m)\| < 2^{-\Omega(n)}$. Hence, by Fact A2 (in Appendix A), $\|\bar{C} - \bar{C}'\| < 2^{-\Omega(n)}$, and therefore $\|D_0 - D'_0\| < 2^{-\Omega(n)}$ and $\|D_1 - D'_1\| < 2^{-\Omega(n)}$. So, by Fact A5, it suffices to prove that $\|D'_0 - D'_1\| < 2^{-\Omega(n)}$.

We now show that D'_0 and D'_1 are indeed statistically close. The first two components of D'_0 and D'_1 are identically distributed. We would like to show that conditioned on the value of those two components, the third and fourth components have small statistical difference. Note that if $(G_0, G_1) \in GNI$, then C_0 and C_1 are completely disjoint. Hence, given the first component, the second is determined. So it is sufficient to prove that conditioned on the first component, the third and fourth have small statistical difference. Given a sample $\bar{z} \leftarrow \bar{C}'$, the size of its pre-image set is $2^{wt(\bar{z})} \geq 2^{w-\Delta}$. Also, any sample from $\bar{C}'$ is chosen with probability at most $2^{w+\Delta}/|G|$. So conditioned on the first component of D'_0, the probability of any triple $(\bar{b}, \bar{r}, \bar{y})$ that h can take as input in the fourth component is at most

$$\left(\frac{1}{2^{w-\Delta}}\right)\left(\frac{2^{w+\Delta}}{|G|}\right) \leq \frac{2^{2\Delta}}{(1 - 2^{-\Omega(n)})(2n!)^m} = \frac{2^{-\Omega(n)}}{|T|}.$$

Recall that in D'_1, conditioned on the first component, the fourth component is uniformly distributed over T. So by the Leftover Hash Lemma [10], conditioned on the first component, the third and fourth components of D'_0 and D'_1 have statistical difference at most $2^{-\Omega(n)}$, and, hence, $\|D'_0 - D'_1\| < 2^{-\Omega(n)}$.

The case of $(G_0, G_1) \notin GNI$. If $(G_0, G_1) \notin GNI$, then C_0 and C_1 are identically distributed. Also, they are uniformly distributed over the domain of all the graphs that are isomorphic to G_0 (and G_1). Since the distribution $\bar{C}$ is composed of many copies of C_0 and C_1, it is also uniformly distributed. Thus, by using the terminology above, the weight of each sample $\bar{z} \leftarrow \bar{C}$ is w, and the size of its pre-image set is 2^w. Let a denote the size of the automorphism group of G_0 (and G_1). Then we know that for a fixed $\bar{b} \in \{0,1\}^m$, $|\{\bar{\pi}\colon \bar{C}(\bar{b}, \bar{\pi}) = \bar{z}\}| = a^m$. Thus, $2^w = 2^m a^m$.

The size of the image set of $\bar{C}$ is $|\{0,1\}^m \times S_n^m|/2^w = 2^{(q+1)m}/2^w$. So given the first two components of D_0, the number of triplets, $(\bar{b}, \bar{\pi}, \bar{y})$, that h can take as input in the fourth component is $a^m(2^{(q+1)m}/2^w) = 2^{qm}$. Fix any $\bar{z}$ which is selected according to $\bar{C}$, also fix $\bar{b} \in \{0,1\}^m$, and $h \in H$. Then, in D_0, conditioned on $\bar{z} = \bar{C}(\bar{b}, \bar{\pi}), \bar{b}$ and h, there are at most $2^{qm} = 2^{-\Omega(n)}|T|$ possible values for $(\bar{b}, \bar{\pi}, \bar{y})$. So the fourth component of D_0, $h(\bar{b}, \bar{\pi}, \bar{y})$, can cover at most a $2^{-\Omega(n)}$ fraction of T. On the other hand, conditioned on any values for the first three components of D_1, the fourth component is uniformly

distributed over T. So the probability that it will fall in the range of the fourth component of D_0 is at most $2^{-\Omega(n)}$.

Addressing the problem of uniformly sampling permutations. Let C_0 and C_1 be the circuits from the second part of the proof of Theorem 24 (i.e. those that can output $\perp_0$ and $\perp_1$), and let D_0 and D_1 sample from them. Both D_0 and D_1 take many samples from C_0 and C_1. We modify them as follows: if one of these samples is $\perp_0$ or $\perp_1$, D_0 and D_1 output special (different) symbols $\perp_0'$ and $\perp_1'$, respectively. Otherwise they continue as before. The probability that D_0 (resp. D_1) outputs $\perp_0'$ (resp. $\perp_1'$) is bounded by $2m2^{-n} = 2^{-\Omega(n)}$. It is now easy to verify that when $(G_0, G_1) \in GNI$ this change can only increase the statistical difference by $2^{-\Omega(n)}$, and when $(G_0, G_1) \notin GNI$, D_0 and D_1 are even more disjoint. □

5. Conclusion

In this paper we compare two resources that can be used in the construction of statistical zero-knowledge proofs. The first resource is the interaction between the prover and the verifier. The second is (limited) help, given to the prover and the verifier in the form of a shared reference string that is dealt by a trusted (computationally bounded) third party. We checked whether one of these resources can be traded for the other. Indeed, we showed that help can always be replaced by interaction. As for replacing interaction with help, we showed non-interactive statistical zero-knowledge proofs with help, for languages for which only interactive protocols are currently known. In order to do that, we showed a complete problem for the class $NISZK|_{\mathrm{h}}$.

We hope that our results will help to shed light on the *SZK* versus *NISZK* question. By introducing a class that is in between *SZK* and *NISZK*, we break this question into two supposedly easier questions. We emphasize that each one of these questions is interesting in its own right. Specifically, we showed that

$$SZK \supseteq NISZK|_{\mathrm{h}} \supseteq NISZK. \tag{1}$$

If we could show that $SZK = NISZK|_{\mathrm{h}}$ we would get an interesting result that interaction and (limited) help are equivalent and interchangable resources in the construction of statistical zero-knowledge protocols. If we could show that $NISZK|_{\mathrm{h}} = NISZK$, then we would get that languages such as Graph Isomorphism and Graph-Non-Isomorphism are in *NISZK*. On the other hand, if we could show that one of the containments in (1) is strict we would get that $SZK \neq NISZK$.

Acknowledgments

We thank Salil Vadhan and Oded Goldreich for valuable discussions about statistical zero-knowledge, and the anonymous referee for many helpful suggestions. The second author also wishes to express his gratitude to Avi Wigderson who directed him to this topic of research.

Appendix A. Facts about Probability Distributions

We state some useful facts regarding probability distributions (or random variables).

Fact A1. Let X and Y be two random variables ranging over a domain D, and let $\delta = \|X - Y\|$, then

$$|H(X) - H(Y)| \leq (\log|D| - 1)\delta + H_2(\delta).$$

Fact A2. For any two random variables X and Y, and any randomized procedure P,

$$\|X - Y\| \geq \|P(X) - P(Y)\|.$$

Fact A3. For any $0 \leq q' \leq q \leq p \leq 1$, it holds that $KL_2(p, q') \geq KL_2(p, q)$.

Fact A4. For any two random variables X and Y, and any function f,

$$KL(X \mid Y) \geq KL(f(X) \mid f(Y)).$$

Fact A5. For any probability distributions X, Y and Z, $\|X-Y\| \leq \|X-Z\|+\|Z-Y\|$.

Appendix B. Proof of Lemma 14

Let P^* be an arbitrary prover's strategy. For every $x \in \Pi_{\mathrm{N}}$, let $s_x = \Pr[(D, P^*, V)(x) = \text{“}accept\text{”}]$ and $s'_x = \Pr[(D', P^*, V)(x) = \text{“}accept\text{”}]$. Also let q be the polynomial that defines the length of the help string. Denote by $(P^*, V)_y(x)$, the outcome of the protocol between P^* and V, given that x is the input and y is the reference string. Then

$$\begin{aligned}
|s_x - s'_x| &= \Bigg| \sum_{y\in\{0,1\}^{q(|x|)}} \Pr[D(x) = y] \Pr[(P^*, V)_y(x) = \text{“}accept\text{”}] \\
&\qquad - \sum_{y\in\{0,1\}^{q(|x|)}} \Pr[D'(x) = y] \Pr[(P^*, V)_y(x) = \text{“}accept\text{”}] \Bigg| \\
&= \Bigg| \sum_{y\in\{0,1\}^{q(|x|)}} \Pr[(P^*, V)_y(x) = \text{“}accept\text{”}](\Pr[D(x) = y] - \Pr[D'(x) = y)]) \Bigg| \\
&\leq \sum_{y\in\{0,1\}^{q(|x|)}} |\Pr[(P^*, V)_y(x) = \text{“}accept\text{”}](\Pr[D(x) = y] - \Pr[D'(x) = y])| \\
&\leq \sum_{y\in\{0,1\}^{q(|x|)}} |\Pr[D(x) = y] - \Pr[D'(x) = y]| = 2\varepsilon.
\end{aligned}$$

Where the last equality is by the definition of the statistical difference. □

Appendix C. Proof of Lemma 20

Following [13], the procedure in Lemma 20 contains three building blocks, below we define them and prove their properties.

Lemma C6. *Given a pair of circuits* (C_0, C_1) *with n input gates, consider the following pair*:

D_0: *choose* $r \in_R \{0,1\}^n$ *and* $b \in_R \{0,1\}$, *output* $(C_b(r), b)$.
D_1: *choose* $r \in_R \{0,1\}^n$ *and* $b \in_R \{0,1\}$, *output* $(C_b(r), \bar{b})$.

The following holds:

(1) $\|D_0 - D_1\| = \|C_0 - C_1\|$.
(2) *If the pair* (C_0, C_1) *is* α*-disjoint, then* (D_0, D_1) *is mutually* $(\alpha/2)$*-disjoint*.[8]

Proof. (1) Let S be the domain over which C_0 and C_1 are distributed. Then

$$\begin{aligned}\|D_0 - D_1\| &= \frac{1}{2}\sum_{x\in S}\sum_{b\in\{0,1\}} |\tfrac{1}{2}(\Pr(x \leftarrow C_b) - \Pr(x \leftarrow C_{\bar{b}}))| \\ &= \frac{1}{4}\sum_{x\in S}(|\Pr(x \leftarrow C_0) - \Pr(x \leftarrow C_1)| + |\Pr(x \leftarrow C_1) - \Pr(x \leftarrow C_0)|) \\ &= \frac{1}{2}\sum_{x\in S}|\Pr(x \leftarrow C_0) - \Pr(x \leftarrow C_1)| = \|C_0 - C_1\|.\end{aligned}$$

Remark. In [13] a more general statement was proven, which can be used to give an upper bound on the statistical difference between D_0 and D_1. In our specific case, this term is equal to the statistical difference between C_0 and C_1.

(2) Let $S_0 = \mathrm{Im}(C_0)\backslash\mathrm{Im}(C_1)$. By definition, for $x \leftarrow C_0$ and $x' \leftarrow C_1$, $\Pr(x \in S_0) > \alpha$, and $\Pr(x' \in S_0) = 0$. Then for $(x, b) \leftarrow D_0$ and $(x', b') \leftarrow D_1$, $\Pr(x \in S_0 \wedge b = 0) > \alpha/2$ and $\Pr(x' \in S_0 \wedge b' = 0) = 0$. Similarly, $\Pr(x \in S_0 \wedge b = 1) = 0$ and $\Pr(x' \in S_0 \wedge b' = 1) > \alpha/2$. □

The following two lemmas are the main building blocks in the proof of the Polarization Lemma of [13]. We will use their results regarding the statistical difference, and prove useful properties when applied to mutually disjoint distributions.

Lemma C7. *Given a pair of circuits* (C_0, C_1) *with n input gates, and a parameter k, consider the following pair*:

D_0: *choose* $(r_1, \ldots, r_k) \in_R \{0,1\}^{kn}$, *output* $(C_0(r_1), \ldots, C_0(r_k))$.
D_1: *choose* $(r_1, \ldots, r_k) \in_R \{0,1\}^{kn}$, *output* $(C_1(r_1), \ldots, C_1(r_k))$.

[8] See Definition 4 for (mutually) α-disjoint.

The following holds:

(1) $\|D_0 - D_1\| \leq k \cdot \|C_0 - C_1\|$.
(2) *If the pair* (C_0, C_1) *is mutually* α*-disjoint, then* (D_0, D_1) *is mutually* $(1-(1-\alpha)^k)$*-disjoint*.

Proof. (1) This was proved in [13].

(2) Let $S = \mathrm{Im}(C_0) \cap \mathrm{Im}(C_1)$, and $(x_1, \ldots, x_k) \leftarrow D_0$. If for some $1 \leq i \leq k$, $x_i \notin S$, then $\Pr((x_1, \ldots, x_k) \leftarrow D_1) = 0$. That is, $(x_1, \ldots, x_k)$ is in $\mathrm{Im}(D_1)$ only if for every $1 \leq i \leq k$, $x_i \in S$. Thus, the probability that a sample from D_0 is in $\mathrm{Im}(D_1)$ is at most $(1-\alpha)^k$. Similarly, it can be shown that the probability that a sample from D_1 is in $\mathrm{Im}(D_0)$ is at most $(1-\alpha)^k$. By definition, the pair (D_0, D_1) is mutually $(1-(1-\alpha)^k)$-disjoint. □

Lemma C8. *Given a pair of circuits* (C_0, C_1) *with* n *input gates, and a parameter* k*, consider the following pair*:

D_0: *choose* $(b_1, \ldots, b_k) \in_{\mathrm{R}} \{(c_1, \ldots, c_k) \in \{0,1\}^k : c_1 \oplus \cdots \oplus c_k = 0\}$, *and* $(r_1, \ldots, r_k) \in_{\mathrm{R}} \{0,1\}^{kn}$, *output* $(C_{b_1}(r_1), \ldots, C_{b_k}(r_k))$.
D_1: *choose* $(b_1, \ldots, b_k) \in_{\mathrm{R}} \{(c_1, \ldots, c_k) \in \{0,1\}^k : c_1 \oplus \cdots \oplus c_k = 1\}$, *and* $(r_1, \ldots, r_k) \in_{\mathrm{R}} \{0,1\}^{kn}$, *output* $(C_{b_1}(r_1), \ldots, C_{b_k}(r_k))$.

The following holds:

(1) $\|D_0 - D_1\| = \|C_0 - C_1\|^k$.
(2) *If the pair* (C_0, C_1) *is mutually* α*-disjoint, then* (D_0, D_1) *is mutually* α^k*-disjoint*.

Lemma C8 can be easily proven by induction once we have the following:

Claim C9. *Given two pairs of circuits* (C_0, C_1), (C'_0, C'_1) *with* n *and* n' *input gates, respectively, consider the following circuits*:

D_0: *choose* $b \in_{\mathrm{R}} \{0,1\}$, $r \in_{\mathrm{R}} \{0,1\}^n$ *and* $r' \in_{\mathrm{R}} \{0,1\}^{n'}$, *output* $(C_b(r), C'_b(r'))$.
D_1: *choose* $b \in_{\mathrm{R}} \{0,1\}$, $r \in_{\mathrm{R}} \{0,1\}^n$ *and* $r' \in_{\mathrm{R}} \{0,1\}^{n'}$, *output* $(C_b(r), C'_{\bar{b}}(r'))$.

The following holds:

(1) $\|D_0 - D_1\| = \|C_0 - C_1\| \cdot \|C'_0 - C'_1\|$.
(2) *If* (C_0, C_1) *and* (C'_0, C'_1) *are mutually* α*-disjoint and mutually* α'*-disjoint, repectively, then the pair* (D_0, D_1) *is mutually* $(\alpha\alpha')$*-disjoint*.

Proof. (1) This was proved in [13].

(2) Define the following sets, $S_0 = \mathrm{Im}(C_0) \backslash \mathrm{Im}(C_1)$, $S_1 = \mathrm{Im}(C_1) \backslash \mathrm{Im}(C_0)$, $S'_0 = \mathrm{Im}(C'_0) \backslash \mathrm{Im}(C'_1)$ and $S'_1 = \mathrm{Im}(C'_1) \backslash \mathrm{Im}(C'_0)$. Then for $(x, x') \leftarrow D_0$,

$$\begin{aligned}
&\Pr((x,x') \in (S_0, S'_0) \vee (x,x') \in (S_1, S'_1)) \\
&\quad = \Pr((x,x') \in (S_0, S'_0)) + \Pr((x,x') \in (S_1, S'_1)) \\
&\quad \geq \frac{\alpha\alpha'}{2} + \frac{\alpha\alpha'}{2} = \alpha\alpha'.
\end{aligned}$$

On the other hand, for $(y, y') \leftarrow D_1$, $\Pr((y, y') \in (S_0, S_0') \vee (x, x') \in (S_1, S_1')) = 0$. Similarly, we show that $\Pr((x, x') \in (S_0, S_1') \vee (x, x') \in (S_1, S_0')) = 0$ and $\Pr((y, y') \in (S_0, S_1') \vee (y, y') \in (S_1, S_0')) \geq \alpha\alpha'$. Hence the pair (D_0, D_1) is mutually $(\alpha\alpha')$-disjoint. □

Proof of Lemma 20. We are given a pair of circuits (C_0, C_1) with description length n. We may assume that n is large enough for the arguments below to work. We know that the pair (C_0, C_1) either has statistical difference at most $1/n^2$ or is $(1 - 1/n^2)$-disjoint. First, we apply Lemma C6 to obtain a pair which either has statistical difference at most $1/n^2$ or is mutually $\frac{1}{3}$-disjoint (using $\frac{1}{2}(1 - 1/n^2) > \frac{1}{3}$). We can now achieve the amplification by alternating between the tools we developed above, as was done in [13] (actually, one alternation will suffice). We first apply Lemma C7 with $k = 3n$, to obtain a pair which either has statistical difference at most $3/n < \frac{1}{2}$, or is mutually $(1 - 3^{-n})$-disjoint. We then apply Lemma C8 with $k = n$, to obtain a pair which either has statistical difference at most 2^{-n}, or is mutually $(1 - 2^{-n})$-disjoint (using $(1 - 3^{-n})^n > 1 - 2^{-n}$). □

References

[1] William Aiello and Johan Hastad. Statistical zero-knowledge languages can be recognized in two rounds. *Journal of Computer and System Sciences*, 42(3):327–345, 1991.

[2] Michael Ben-Or and Danny Gutfreund. Increasing the power of the dealer in non-interactive zero-knowledge proof systems. In *Proceedings of ASIACRYPT* 2000, LNCS 1976, pages 429–443, Springer-Verlag, Berlin, 2000.

[3] Manuel Blum, Paul Feldman, and Silvio Micali. Non-interactive zero-knowledge and its applications (extended abstract). In *Proceedings of the* 20*th ACM Symposium on the Theory of Computing*, pages 103–112, Chicago, Illinois, 2–4 May 1988.

[4] Alfredo De Santis, Giovanni Di Crescenzo, Giuseppe Persiano, and Moti Yung. Image density is complete for non-interactive SZK. In *Automata, Languages and Programming*, 25*th International Coloquium*, LNCS 1443, pages 784–795, Springer-Verlag, Berlin, 1998.

[5] Oded Goldreich and Yair Oren. Definitions and properties of zero-knowledge proof systems. *Journal of Cryptology*, 7(1):1–32, Winter 1994.

[6] Oded Goldreich, Amit Sahai, and Salil Vadhan. Honest-verifier statistical zero-knowledge equals general statistical zero-knowledge. In *Proceedings of the* 30*th Annual ACM Symposium on the Theory of Computing*, pages 399–408, 1998.

[7] Oded Goldreich, Amit Sahai, and Salil Vadhan. Can statistical zero-knowledge be made non-interactive? or on the relationship of SZK and NISZK. In Michael Wiener, editor, *Advances in Cryptology, CRYPTO* '99, LNCS 1666, pages 467–484, Springer-Verlag, Berlin, 1999.

[8] Oded Goldreich and Salil Vadhan. Comparing entropies in statistical zero-knowledge with application to the structure of SZK. In *Proceedings of the* 14*th Annual IEEE Conference on Computational Complexity*, pages 54–73, 1998.

[9] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal of Computing*, 18(1):186–208, 1989.

[10] Russell Impaliazzo, Leonid A. Levin, and Michael Luby. Pseudo-random generation from one-way functions (extended abstract). In *Proceedings of the* 21*st Annual ACM Symposium on the Theory of Computing*, pages 12–24, 1989.

[11] Tatsuaki Okamoto. On relationships between statistical zero-knowledge proofs. In *Proceedings of the* 28*th Annual Symposium on the Theory of Computing*, pages 649–658, 1996.

[12] Amit Sahai and Salil Vadhan. A complete promise problem for statistical zero-knowledge. In *Proceedings of the* 38*th Annual Symposium on Foundations of Computer Science*, pages 448–457, 1997.

[13] Amit Sahai and Salil Vadhan. Manipulating statistical difference. In Panos Pardalos, Sanguthevar Rajaseekaran, and Jose Rolim, editors, *Proceedings of the DIMACS Workshop on Randomization Methods in Algorithm Design*, pages 251–270, 1998.