The Hidden Subgroup Problem and Non-interactive Perfect Zero-Knowledge Proofs*

Abner F. B. Costa[1], **Henrique Hepp**[1], **Murilo V. G. da Silva**[1], **Leandro M. Zatesko**[2]

[1]Department of Informatics, Federal University of Paraná, Brazil

[2]Academic Department of Informatics, Federal University of Technology — Paraná, Brazil

{afbcosta,hhepp,murilo}@inf.ufpr.br, zatesko@utfpr.edu.br

Abstract. *The Hidden Subgroup Problem (HSP) generalises many problems that are candidates to be* NP*-intermediate. It was shown that the decision version of HSP belongs to the zero-knowledge complexity class* HVPZK *and that, if the size of the group is known, it also belongs to* NISZK*. We show that whenever we can sample uniformly at random elements of the group and of a set, with the same size of the group, that contains the image of the function that hides the subgroup, the problem is in* NIPZK_1 *(i.e.* NIPZK *with perfect completeness). As a second contribution, we show that* NIPZK_1 *has a complete promise problem that is a restricted version of a complete promise problem for the* NIPZK *class.*

1. Introduction

The Hidden Subgroup Problem (HSP) generalises many problems suspected to be NP-intermediate, such as Factorisation [Shor 1994], Graph Isomorphism [Jozsa 2001], and Unique Shortest Vector in Lattices [Regev 2004]. The most commonly found definition in the literature for HSP is the one introduced by [Babai and Szemerédi 1984] in a black-box context. Determining whether a hidden subgroup is the trivial subgroup or not is a well-studied definition for a decision version of HSP [Ettinger et al. 2004, Hayashi et al. 2008, Sdroievski et al. 2019] and it is the one used in this paper, although it is not the only decision version (technical details and further definitions are discussed in the sequel). This decision version (dHSP) was shown to be in the zero-knowledge class HVPZK and, if the size of the group is known, it also belongs to the class NISZK [Sdroievski et al. 2019].

We investigate restrictions under which dHSP has a non-interactive perfect zero-knowledge protocol with perfect completeness (i.e. is in the class NIPZK_1). We show that this holds whenever we can sample uniformly at random elements of the group and of a set, with the same size of the group, that contains the image of the function that hides the subgroup (Thm. 1). Other restrictions are also shown (Cor. 2 and 3).

The class NIPZK was defined by [Malka 2008], who showed a complete problem for the class. He also implied that with a certain restriction, this problem might become complete for NIPZK_1, but he did not explicitly prove this assertion, since that was not the point being addressed. Another contribution of our paper is this proof, in Sect. 3.

Figure 1 depicts the known relationships between the zero-knowledge complexity classes and the restrictions of HSP. To distinguish the different decision versions mentioned, we use dHSP for the general case, $\text{dHSP}_{|G|}$ for the restricted case when the size of the group is known, and dHSP_{Im} for our restricted case.

*Partially supported by CAPES (grant 001) and CNPQ (grant 420079/2021-1).

$$\begin{tikzcd}
& \mathsf{NIPZK}_1 \arrow[r] & \mathsf{NIPZK} \arrow[rr] \arrow[dr] & & \mathsf{NISZK} \arrow[dr] & \\
& & & \mathsf{PZK} \arrow[dr] & & \mathsf{SZK} \\
\mathrm{dHSP}_{Im} \arrow[r] \arrow[uur, red] & \mathrm{dHSP}_{|G|} \arrow[r] \arrow[uurrr, bend left] & \mathrm{dHSP} \arrow[rr] & & \mathsf{HVPZK} \arrow[ur] &
\end{tikzcd}$$

Figure 1. In the figure, an arrow A → B represents: that A $\subseteq$ B if A and B are classes; that A $\in$ B, if A is a problem and B is a class; that A is a restriction of B, if both A and B are problems. Our result is highlighted in red.

We assume that the reader is familiar with basic topics in group theory and interactive proofs, for which we refer to [Herstein 1991, Arora and Barak 2009], respectively.

Consider a family of groups $\mathcal{B} = \{B_n\}_{n \geq 1}$ such that: the elements of B_n are uniquely represented by words of length $\text{poly}\,(n)$; inverse, product, and identity testing operations of each $\mathcal{B}_n$ are computed in $\text{poly}\,(n)$ time, denoting by e the identity element of B_n. The formal definition of HSP that we consider is the one proposed by [Sdroievski et al. 2019], in which we are *given* a positive integer n (in unary) and a boolean circuit C_f that takes encodings of elements of a group $G \subseteq B_n$ as input and returns an output of m bits, being m a positive integer. We assume that C_f computes a function f that *hides* a subgroup H in G, i.e. $f(a) = f(b)$ if and only if $aH = bH$ for all $a, b \in G$. The *goal* of the problem HSP is to output a generating set of H. In dHSP, as defined below, the *goal* is to decide if H is trivial, i.e. $f(a) = f(b)$ if and only if $a = b$.

dHSP (for a family of groups $\mathcal{B} = \{B_n\}_{n \geq 1}$ as above)

Given: $(0^n, T, C_f)$, where T is the generating set of a group $G \subseteq B_n$, with $|T| = \text{poly}(n)$, and C_f is a $\text{poly}(n)$-size circuit that takes encodings of elements of B_n and returns m-bit strings, for some $m \in \mathbb{Z}_{\geq 0}$, so that C_f computes a function f which hides a subgroup H in G;

decide: *positive instances:* $\text{dHSP}_Y = \{(0^n, T, C_f) : |H| = 1\}$;
negative instances: $\text{dHSP}_N = \{(0^n, T, C_f) : |H| \geq 2\}$;

promised that f hides a subgroup $H \subseteq G$.

The protocol (P, V) is said to be *non-interactive* (NI) if the prover P and the verifier V share a common reference string r and the first and only message between P and V is sent by the prover. We say that a protocol (P, V) has *perfect zero-knowledge* (PZK) if, in addition to the *efficiency*, *completeness* and *soundness* conditions, there is a probabilistic polynomial-time simulator S such that on all positive instances x, the simulator S outputs `fail` with probability at most $1/2$. Additionally, the random variable $\tilde{S}(x)$, describing the distribution of $S(x)$ conditioned on S not failing, and $\langle P, V \rangle(x)$, the view of verifier on (P, V), are identically distributed. Similarly, a protocol has *statistical zero-knowledge* (SZK) if $\tilde{S}(x)$ and $\langle P, V \rangle(x)$ are statistically indistinguishable for all positive instances. A verifier V is a *honest verifier* (HV) if it does not deviate from the protocol.

The class of problems with a *interactive perfect zero-knowledge protocol with a honest verifier* is called HVPZK; with a *non-interactive statistical zero-knowledge protocol* is NISZK; with a *non-interactive perfect zero-knowledge protocol* is NIPZK. The class NIPZK with *perfect completeness* is NIPZK_1.

2. Decision version of HSP and zero-knowledge complexity classes

Even though dHSP is not known to be in NP, the problem is in coNP, since for $|H| \geq 2$, we can use an element $h \neq e \in H$ as a certificate for a negative instance. To verify it, we check if $f(h) = f(e)$. It is shown in [Sdroievski et al. 2019] a perfect zero-knowledge protocol with honest verifier for dHSP, establishing that dHSP $\in \mathsf{HVPZK}$. Furthermore, [Sdroievski et al. 2019] showed that if we know the size of the group G, as it is the case for permutation groups [Seress 2003], then there is a polynomial Karp reduction from dHSP to the Entropy Approximation Problem (EA), a complete promise problem for NISZK.

We observe that the $Im(f)$, the image of the function f, has at most $|G|$ elements, therefore C_f has at most $|G|$ different possible outputs. We define dHSP$_{Im}$ as a restriction of dHSP where we can sample uniformly at random elements of G and of a set $A_f \subseteq \{\{0,1\}^m\}$ such that $Im(f) \subseteq A_f$ and $|A_f| = |G|$. Although this restriction may seem artificial, notice that it holds for important cases, such as when $G = \mathbb{Z}_n$ with modular addition or multiplication. Corollary 2 further explores when this restriction can be met.

Theorem 1. *dHSP$_{Im}$* $\in \mathsf{NIPZK}_1$.

Proof. Let $r \in A_f$, chosen uniformly at random be the common reference string and $B_r = \{g \in G : f(g) = r\}$. The prover P samples $g \in B_r$ uniformly at random and sends it to the verifier V, which accepts if $f(g) = r$ and rejects otherwise.

Since C_f is a $\mathrm{poly}(n)$-size circuit, the verifier V can compute $f(g)$ efficiently. If $|H| = 1$, then V always accepts, achieving the completeness property. Now, we show the soundness property. If $|H| \geq 2$ then $|Im(f)|/|G| \leq 1/2$. Since $r \in A_f$ is chosen uniformly at random, the probability that $r \in Im(f)$ is at most $1/2$. Hence, the probability that there exists $g \in G$ that can be sent from P to V such that $f(g) = r$ is at most $1/2$.

To achieve the zero-knowledge property, let S be the simulator which chooses $g' \in G$ uniformly at random and computes $r' = f(g')$. Since $|H| = 1$, the function f is a bijection, obtain $r' \in A_f$ also uniformly at random. Therefore, the transcripts $\langle C_f, r', g' \rangle$ of the simulator S and the transcripts $\langle C_f, r, g \rangle$ of the protocol are identically distributed whenever $|H| = 1$. □

Corollaries 2 and 3 follows immediately from Theorem 1.

Corollary 2. *If we can sample uniformly at random elements of G, and if A_f is the set of the first $|G|$ strings in lexicographic order, then dHSP* $\in \mathsf{NIPZK}_1$. □

Corollary 3. *If we can sample uniformly at random elements of G, and $m = \log|G|$, then dHSP* $\in \mathsf{NIPZK}_1$. □

3. A complete problem for NIPZK_1

The class NIPZK was defined by [Malka 2008], who also presented a promise problem which is complete for the class. He also implied that a restricted version of this problem is also complete for the class NIPZK_1, although he did not explicitly prove this assertion, since this was not his aim at that point. As a second contribution of our paper, we show, by following a similar structure of the proof of Malka, that the result holds indeed.

Below we define the problem *Uniform-or-Small* (US), where U_m denotes the uniform distribution on all m-bit strings, and $sup(X) = \{y \in \{0,1\}^m : \Pr[X = y] \neq 0\}$ is the *support* of X. This problem was already studied by [Dixon et al. 2020], who showed that there is an oracle A relative to which US is not in the probabilistic class SBP.

US

Given: a $\mathrm{poly}(n)$-size circuit $C : \{0,1\}^n \to \{0,1\}^m$ encoding a distribution X;
decide: *positive instances:* $\mathrm{US}_Y = \{X : \Delta(X, U_m) = 0\}$;
negative instances: $\mathrm{US}_N = \{X : |sup(X)| \leq 2^m/3\}$.
promised that one of the cases hold.

Theorem 4. *US is complete for* NIPZK_1.

Proof that US $\in \mathsf{NIPZK}_1$. Let $r \in \{0,1\}^m$ chosen uniformly at random, be the common reference string, and $B_r = \{\pi \in \{0,1\}^n : X(\pi) = r\}$. The prover P samples $\pi \in B_r$ uniformly at random and sends it to the verifier V, which accepts if $X(\pi) = r$ and rejects otherwise.

Since C is a $\mathrm{poly}(n)$-size circuit, the verifier V can compute $X(\pi)$ efficiently. If $X \in \mathrm{US}_Y$, then V always accepts, achieving the completeness property. Now, we show the soundness property. if $X \in \mathrm{US}_N$ then, since $|sup(X)| \leq 2^m/3$ and r is uniformly random, the probability that $r \in sup(X)$ is at most $1/3$. Therefore, the probability that there is some $\pi \in \{0,1\}^n$ with $X(\pi) = r$ that can be sent by P to V is at most $1/3$.

To achieve the zero-knowledge property, let $X \in \mathrm{US}_Y$ and S be the simulator which chooses $\pi' \in \{0,1\}^n$ uniformly at random and computes $r' = X(\pi')$. When $X \in \mathrm{US}_Y$, we obtain $r' \in \{0,1\}^m$ uniformly at random. Therefore, the transcripts $\langle X, r', \pi' \rangle$ of the simulator S and the transcripts $\langle X, r, \pi \rangle$ of the protocol are identically distributed whenever $X \in \mathrm{US}_Y$. □

Proof that US is NIPZK_1*-hard.* Let $\Pi = \langle \Pi_Y, \Pi_N \rangle$ be a NIPZK_1 problem, and (P, V) be a non-interactive protocol for Π with perfect completeness and soundness error at most $1/3$. We denote r as the common reference string, choosed uniformly at random, with size $|r| = |x|^c$ for some $c \in \mathbb{N}$ and every $x \in \Pi_Y \cup \Pi_N$. Let S be a simulator for $\langle P, V \rangle$ and d be a constant such that S uses no more than $|x|^d$ random bits for every input x.

We show that Π has a polynomial Karp reduction to US. We shall define a polynomial-time Turing machine that, on input $x \in \Pi_Y \cup \Pi_N$, outputs a circuit $C : \{0,1\}^{|x|^d} \to \{0,1\}^{|x|^c}$ that encodes a distribution X, such that if $x \in \Pi_Y$, then $X \in \mathrm{US}_Y$, and if $x \in \Pi_N$, then $X \in \mathrm{US}_N$. Given x, the circuit C is constructed such that, on input r_S with $|r_S| \leq |x|^d$, emulates the computation of S on x under randomness r_S, yielding the view $\langle x, r, \pi \rangle$, wherein r is the common reference string in (P, V) and π is the message sent from P to V. Then, C outputs r if $V(x, r, \pi) = \texttt{accept}$, and $0^{|x|^c}$ otherwise.

Now, we analyse the reduction. If $x \in \Pi_Y$, then $V(x, r, \pi) = \texttt{accept}$. By construction, C outputs r. Since r is uniformly distributed, X is the uniform distribution. If $x \in \Pi_N$, we show that $|sup(X)| \leq 2^{|x|^c}/3$. By the soundness condition, the probability that the verifier $V(x)$ accepts in $\langle P, V \rangle$ is less than $1/3$. Thus, by construction, in more than $2/3$ cases for all the $2^{|x|^c}$ possible outputs of C we have the string $0^{|x|^c}$. Therefore, $|sup(X)| \leq 2^{|x|^c}/3$. □

Obviously, $|sup(X)| \leq 2^m/3$ in the definition of US could be replaced by $|sup(X)| \leq \beta \cdot 2^m$ for any positive constant $\beta < 1$, and Theorem 4 would also hold. In fact, US was equivalently stated by [Dixon et al. 2020] with $\beta = 1/2$.

References

Arora, S. and Barak, B. (2009). *Computational complexity: a modern approach*. Cambridge University Press.

Babai, L. and Szemerédi, E. (1984). On the complexity of matrix group problems I. In *25th Annual Symposium on Foundations of Computer Science, 1984*, pages 229–240. IEEE.

Dixon, P., Gayen, S., Pavan, A., and Vinodchandran, N. (2020). Perfect zero knowledge: New upperbounds and relativized separations. In *Theory of Cryptography: 18th International Conference, TCC 2020, Durham, NC, USA, November 16–19, 2020, Proceedings, Part I*, pages 684–704. Springer.

Ettinger, M., Høyer, P., and Knill, E. (2004). The quantum query complexity of the hidden subgroup problem is polynomial. *Information Processing Letters*, 91(1):43–48.

Hayashi, M., Kawachi, A., and Kobayashi, H. (2008). Quantum measurements for hidden subgroup problems with optimal sample complexity. *Quantum Information & Computation*, 8(3):345–358.

Herstein, I. N. (1991). *Topics in algebra*. John Wiley & Sons.

Jozsa, R. (2001). Quantum factoring, discrete logarithms, and the hidden subgroup problem. *Computing in science & engineering*, 3(2):34–43.

Malka, L. (2008). How to achieve perfect simulation and a complete problem for non-interactive perfect zero-knowledge. In *Theory of Cryptography: Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19-21, 2008. Proceedings 5*, pages 89–106. Springer.

Regev, O. (2004). Quantum computation and lattice problems. *SIAM Journal on Computing*, 33(3):738–760.

Sdroievski, N. M., da Silva, M. V., and Vignatti, A. L. (2019). The hidden subgroup problem and MKTP. *Theoretical Computer Science*, 795:204–212.

Seress, Á. (2003). *Permutation group algorithms*. Number 152. Cambridge University Press.

Shor, P. W. (1994). Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. IEEE.