# The Complexity of Perfect Zero-Knowledge

Lance Fortnow*
MIT Math Dept.†
Cambridge, MA 02139

**Abstract**

A *Perfect Zero-Knowledge* interactive proof system convinces a verifier that a string is in a language without revealing any additional knowledge in an information-theoretic sense. We show that for any language that has a perfect zero-knowledge proof system, its complement has a short interactive protocol. This result implies that there are not any perfect zero-knowledge protocols for NP-complete languages unless the polynomial time hierarchy collapses. This paper demonstrates that knowledge complexity can be used to show that a language is easy to prove.

## 1 Introduction

Interactive protocols and zero-knowledge, as described by Goldwasser, Micali and Rackoff [GMR], have in recent years proven themselves to be important models of computation in both complexity and cryptography.

Interactive proof systems are a randomized extension to NP which give us a greater understanding of what an infinitely powerful machine can prove to a probabilistic polynomial one. Recent results about interactive protocols have given us an idea of what languages may be efficiently provable in this way.

Zero-knowledge interactive protocols give us a good way to determine which languages can be efficiently proven without giving away any details of the proof. This model consists of an infinitely powerful prover trying to convince a polynomial time verifier that a string is in a certain language. Zero-knowledge requires that the verifier not learn any information useful to him as a polytime machine. Goldreich, Micali and Wigderson [GMW] show that if one way functions exist then all languages in NP have zero-knowledge proofs. However, their proof relies on the fact that the verifier has limited power and is unable to invert these one-way functions. A stronger notion is that of perfect zero-knowledge (PZK) which requires that the verifier not learn any additional information no matter how powerful he may be. There are many languages not known to be in BPP or NP$\cap$co-NP, such as graph isomorphism [GMW], which have perfect zero-knowledge proof systems.

Our main theorem says that for any language which has a perfect zero-knowledge protocol, its complement has a single round interactive protocol. Thus PZK$\subseteq$co-AM, where AM is the class accepted by one-round Arthur-Merlin games as described by Babai [B]. Our result holds in the weaker case where we only require that the verifier which follows the protocol will not learn any additional information.

Combining our main theorem with a result of Boppana, Hastad and Zachos [BHZ], we get that NP-complete languages do not have perfect zero-knowledge proof systems unless the polynomial time hierarchy collapses to the second level. Thus it is unlikely that the result of [GMW] that NP has zero-knowledge proof systems will extend to perfect zero-knowledge.

Our proof makes use of an approximate upper bound protocol that is of independent interest and that may be useful in completely different contexts. This is in contrast to an approximate lower bound protocol used in [S, B, GS].

The results in this paper do not depend on any unproven cryptographic assumptions.

*The author is supported in part by an Office of Naval Research fellowship, NSF Grant DCR-8602062 and Air Force Grant AFOSR-86-0078.

†Much of this work was done while the author was at the University of California at Berkeley.

## 2 Notation and Definitions

Let $P$, the prover, be a probabilistic infinite power Turing machine and $V$, the verifier, be a probabilistic polynomial time machine that share the same input and can communicate with each other. Let $P \leftrightarrow V$ denote the interaction between $P$ and $V$. $P \leftrightarrow V(x)$ accepts if after the interaction, $V$ accepts. $V$'s *view of the conversation* between $P$ and $V$ consists of all the messages between $P$ and $V$ and the random coin tosses of $V$.

$P$ and $V$ form an interactive protocol for a language $L$ if:

1. If $x \in L$ then $\Pr(P \leftrightarrow V(x) \text{ accepts}) \geq \frac{2}{3}$
2. If $x \notin L$ then $\forall P^* \Pr(P^* \leftrightarrow V(x) \text{ accepts}) \leq \frac{1}{3}$

A *round* of an interactive protocol is a message from the verifier to the prover followed by a message from the prover to the verifier. AM is the class of languages accepted by bounded round interactive protocols.

Messages in a conversation will be described by

$$\beta_1, \alpha_1, \beta_2, \ldots, \beta_k, \alpha_k$$

where the $\alpha_i$ are messages from the prover to the verifier at round $i$ and the $\beta_i$ are messages from the verifier to the prover.

$r$ will be used for the random coin tosses of the verifier.

Formally, we think of $P$ as a function from the input and the conversation so far to a probability distribution of messages. We put no restrictions on the complexity of this function other than requiring the lengths of the messages to be bounded in size by a polynomial in the size of the input. This paper will use the informal term, *probabilistic infinite power*, to describe the complexity of the prover.

Let IP be the class of all languages that are *efficiently provable*, i.e., accepted by an interactive protocol.

The notation for describing protocols follows:

$P$: These are computations performed by the prover that can not be seen by the verifier. The prover has probabilistic infinite time to make these computations.

$P \rightarrow V$: This is a message from the prover to the verifier.

$V$: These are computations performed by the verifier that cannot be seen by the prover. These computations must be performed in probabilistic polynomial time.

$V \rightarrow P$: This is a message from the verifier to the prover.

Let $M$ be a simulator for a view of the conversation between $P$ and $V$. $M$ is a probabilistic polynomial time machine that will output a conversation between $P$ and $V$ including the random coin tosses of $V$. Thus each run of $M$ will produce:

$$r, \beta_1, \alpha_1, \beta_2, \ldots, \beta_k, \alpha_k$$

Let $P \leftrightarrow V[x]$ denote the distribution of views of conversations between $P$ and $V$. $M[x]$ denotes the distribution of views of conversations created by running $M$ on $x$.

Let $A[x]$ and $B[x]$ be two distributions of strings. $A[x]$ and $B[x]$ are *statistically close* if for any subset of strings $\mathcal{S}$,

$$\left| \sum_{y \in S} \Pr_{A[x]}(y) - \sum_{y \in S} \Pr_{B[x]}(y) \right| < \frac{1}{q(|x|)}$$

for all polynomials $q$ with $|x|$ large enough. Let $J$ be a probabilistic polynomial time machine that outputs either 0 or 1. $A[x]$ and $B[x]$ are *polytime indistinguishable* if for any $J$,

$$|\Pr(J(A[x]) = 1) - \Pr(J(B[x]) = 1)| < \frac{1}{r(|x|)}$$

for all polynomials $r$ with $|x|$ large enough. $J(A[x])$ is the output of $J$ when run on a string chosen from $A[x]$. Note that if $A[x]$ and $B[x]$ are statistically close then they are polytime indistinguishable.

$P \leftrightarrow V$ is *Zero-Knowledge* (ZK) if for any verifier $V^*$ there is a $M_{V^*}$ such that $(\forall x \in L) P \leftrightarrow V^*[x]$ and $M_{V^*}[x]$ are polytime indistinguishable.

$P \leftrightarrow V$ is *Perfect Zero-Knowledge* (PZK) if for any verifier $V^*$ there is a $M_{V^*}$ such that $(\forall x \in L) P \leftrightarrow V^*[x] = M_{V^*}[x]$.

$P \leftrightarrow V$ is *Almost Perfect Zero-Knowledge* (APZK) if for any verifier $V^*$ there is a $M_{V^*}$ such that $(\forall x \in L) P \leftrightarrow V^*[x]$ and $M_{V^*}[x]$ are statistically close.

Interactive Protocols and Zero-Knowledge were introduced in [GMR]. Perfect Zero-Knowledge was described originally in [GMW]. The class AM was introduced by Babai [B] as the class of languages that have one round interactive protocols where $V$'s message consists exactly of his coin tosses. This was shown to be equivalent to the definition used above by [GS] and [B].

Note that ZK$\supseteq$APZK$\supseteq$PZK. The inclusions are not known to be proper but this paper gives good evidence that ZK$\neq$APZK.

The results in this paper only require a weaker version of zero-knowledge: a simulator only need exist for the given $P$ and $V$ and not necessarily for any $V^*$. For the rest of this paper we will assume we are in this weaker model and $M = M_V$ is the simulator for $P$ and $V$.

# 3 Related Results

Our result shows that for any language $L$ with an almost perfect zero-knowledge protocol, there exists a bounded round interactive protocol for its complement $\overline{L}$. We can then apply several earlier results about bounded round interactive protocols.

Goldwasser and Sipser [GS] have shown that for any language that has an interactive protocol in $Q$ rounds, there is an Arthur-Merlin protocol in $Q + 2$ rounds for that language. Arthur-Merlin protocols are similar to interactive protocols except that the verifier's messages are just random coin tosses. Babai [B] showed that any bounded round Arthur-Merlin protocol is equivalent to a one round Arthur-Merlin protocol. This is just the class AM. Babai also shows that AM$\subseteq NP^R$ for a random oracle R and also that AM$\subseteq \Pi_2^p$. Sipser pointed out that AM is contained in non-uniform NP.

Boppana, Hastad, and Zachos [BHZ] have shown that if co-NP had bounded round interactive proofs then the whole polynomial time hierarchy would be in AM implying that the polynomial time hierarchy collapses to the second level.

Subsequent to our result, Aiello and Hastad [AH] have shown, using similar techniques, that any almost perfect zero-knowledge protocol can be done by a bounded round interactive protocol. This result is a nice complement to our result which describes the complexity of the complement of perfect zero-knowledge languages. Combining the two results we have that any language with a perfect zero-knowledge proof system is in non-uniform NP$\cap$co-NP.

Brassard and Crépeau [BC] have shown perfect zero-knowledge for SAT using a different model for interactive protocols where the prover is a polynomial time machine that knows a satisfying assignment. Our result about perfect zero-knowledge relies on the ability of the prover to have infinite power and thus does not apply to Brassard and Crépeau's model.

# 4 Showing Sets are Large and Small

In this paper, we will need protocols to show sets are large and small. We do both using Carter-Wegman Universal Hash Functions [CW].

Let $\Sigma = \{0,1\}$. Suppose $\mathcal{S} \subseteq \Sigma^N, 0^N \notin \mathcal{S}$. Let $F$ be a random binary $b \times N$ matrix. Let $f : \Sigma^N \to \Sigma^b$ be the function defined by $f(x) = Fx$ using regular matrix multiplication modulo two. We can think of $f$ in terms of linear algebra over the field of two elements. $f$ is distributed evenly over all possible linear functions from $n$-dimensional space to $b$-dimensional space.

**Lemma 1 (Vector Independence)** *Suppose $x_1, x_2, \ldots, x_k \in \Sigma^N$ are linearly independent vectors over the field of two elements. Then $f(x_1), f(x_2), \ldots, f(x_k)$ are independently and uniformly distributed over $\Sigma^b$.*

**Proof** Since $x_1, x_2, \ldots, x_k$ are linearly independent, we can extend to a basis. Pick $b_{k+1}, b_{k+2}, \ldots, b_N$ to complete the basis. Let $T$ be the transformation matrix from this new basis to the canonical basis of $\Sigma^N$. Then the matrix

$B = FT$ describes the function from the new basis to the canonical basis of $\Sigma^b$. Since $T$ is an invertible matrix, there is a one-to-one correspondence between $B$ and $F$. Thus $B$ is distributed uniformly over all possible binary $b \times N$ matrices. $f(x_j)$ is just the $j$-th column of $B$. Thus each $f(x_j)$ is independently distributed over $\Sigma^b$. □

If $|\mathcal{S}| \gg 2^b$ then $f_{\mathcal{S}}$ is likely to be onto most of $\Sigma^b$ and most elements of $\Sigma^b$ will have many preimages.

If $|\mathcal{S}| \ll 2^b$ then the range of $f_{\mathcal{S}}$ is a small subset of $\Sigma^b$ and most elements of $f_{\mathcal{S}}(\mathcal{S})$ will have only one inverse in $\mathcal{S}$.

**Lower Bound Protocol**

If $\mathcal{S}$ is recognizable in polynomial time we use the following protocol to show $\mathcal{S}$ is large:

$V$: Pick $\ell$ independent random hash functions $f_1, \ldots, f_\ell : \Sigma^N \to \Sigma^b$ and $\ell^2$ points $z_1, \ldots, z_{\ell^2} \in \Sigma^b$

$V \to P$: $f_1, \ldots, f_\ell, z_1, \ldots, z_{\ell^2}$

$P \to V$: $x$

$V$: Accept if $x \in \mathcal{S}$ and $f_i(x) = z_j$ for some $i, j$, $1 \le i \le \ell$ and $1 \le j \le \ell^2$

If $\mathcal{S}$ is much smaller than $2^b$ then it is unlikely for there to be any $x$ such that $f_i(x) = z_j$. However if $\mathcal{S}$ is large then there are likely to be many $x$ so a prover should have no trouble exhibiting such an $x$ that $V$ can verify in polynomial time.

**Lemma 2 (Lower Bound)** [GS] *Using the above protocol with a given $N, b, d > 0$ and $\ell > \max\{b, 8\}$*

1. *If $|\mathcal{S}| \ge 2^b$ then* $\Pr(P \leftrightarrow V \text{ accepts}) \ge 1 - 2^{-\frac{\ell}{8}}$
2. *If $|\mathcal{S}| \le \frac{2^b}{d}$ then* $\Pr(P^* \leftrightarrow V \text{ accepts}) \le \frac{\ell^3}{d}$ *for any $P^*$*

**Upper Bound Protocol**

If $V$ has a random element $s \in \mathcal{S}$ that is not known by $P$ then the following protocol is used to show $\mathcal{S}$ is small:

$V$: Pick a random $N \times b$ matrix $F$

$V \to P$: $F, f(s) = Fs$

$P \to V$: $s$

If $\mathcal{S}$ is small then $s$ is likely to be the only element of $\mathcal{S}$ that maps to $f(s)$; thus $P$ can find $s$. If $\mathcal{S}$ is large then many elements of $\mathcal{S}$ map to $f(s)$, and because $s$ is a random element of $\mathcal{S}$, the prover will have no way of determining which element of $\mathcal{S}$ $V$ has.

**Lemma 3 (Upper Bound)** *Using the above protocol with a given $N, b > 0$ and $d > 7$*

1. *If $|\mathcal{S}| \le \frac{2^b}{d}$ then* $\Pr(P \leftrightarrow V \text{ accepts}) \ge 1 - \frac{1}{d}$
2. *If $|\mathcal{S}| > 8d2^b$ then* $\Pr(P^* \leftrightarrow V \text{ accepts}) \le \frac{1}{d}$ *for any $P^*$*

**Proof** Let $A$ be the random variable equal to the number of $x \neq s$ in $\mathcal{S}$ such that $f(x) = f(s)$. Let $\mathcal{S}' = \mathcal{S} - \{s\}$. Let $A_x$ be the indicator random variable equal to one if $f(x) = f(s)$, zero otherwise. Then

$$E(A) = E(\sum_{x \in \mathcal{S}'} A_x) = \sum_{x \in \mathcal{S}'} E(A_x) = \sum_{x \in \mathcal{S}'} 2^{-b} = \frac{|\mathcal{S}| - 1}{2^b}$$

If $|\mathcal{S}| \le \frac{2^b}{d}$ then $E(A) \le \frac{1}{d}$. If $f(s)$ has only $s$ as an inverse in $|\mathcal{S}|$ then $P$ with his infinite power will be able to determine $s$. Thus $\Pr(P \leftrightarrow V \text{ rejects}) \le \Pr(A \ge 1) \le E(A) \le \frac{1}{d}$ since $A$ is an integral random variable.

Suppose $|\mathcal{S}| > 8d2^b$. We can assume $|\mathcal{S}| = 8d2^b + 1$ without increasing the probability of acceptance. Then $E(A) = 8d$. Since $P^*$ has no idea what $s$ $V$ has, $P^*$ can only have a $\frac{1}{A+1}$ probability of predicting the $s$ that $V$ has. We will show that there is a high probability that $A$ is large. To show this we look at the variance of $A$.

Given $x, y, s$ all distinct and $y \neq x \oplus s$ then $x, y, s$ are linearly independent. Then by the Vector Independence Lemma $f(x), f(y), f(s)$ are independently distributed over $\Sigma^b$. It then follows that $A_x$ and $A_y$ are independent random variables and their covariance is zero.

The covariance of any two indicator random variables is never greater then the expected value of one of them.

$$VAR(A) = \sum_{x,y \in \mathcal{S}'} COV(A_x, A_y) = \sum_{x \in \mathcal{S}'} (COV(A_x, A_x) + COV(A_x, A_{x \oplus s})) \leq \sum_{x \in \mathcal{S}'} 2E(A_x) \leq 16d$$

It's possible that $x \oplus s \notin \mathcal{S}$ but this would only decrease the variance. Using Chebyshev's inequality we get:

$$\Pr(A < 2d) \leq \Pr(|A - 8d| \geq 6d) \leq \frac{VAR(A)}{36d^2} \leq \frac{16d}{36d^2} \leq \frac{1}{2d}$$

So with probability at most $\frac{1}{2d}$ $A$ is small enough that $P^*$ can determine $s$ easily; otherwise $P^*$ has at most $\frac{1}{2d}$ chance of guessing $s$, so in total $P^*$ has at most a $\frac{1}{d}$ chance of determining $s$. □

**Comparison Protocol**

Suppose we had two sets $\mathcal{S}_1, \mathcal{S}_2 \subseteq \Sigma^N$ and wanted to show that $|\mathcal{S}_1| \gg |\mathcal{S}_2|$. If $\mathcal{S}_1$ is polynomial time testable and $V$ has a random element $s_2$ of $\mathcal{S}_2$, then the following is a protocol to show $|\mathcal{S}_1| \gg |\mathcal{S}_2|$:

$P \rightarrow V$: $b' \leq N$

$P \rightarrow V$: Use lower bound protocol on $\mathcal{S}_1$ with $b = b', \ell = 8nN$

$P \rightarrow V$: Use upper bound protocol on $\mathcal{S}_2$ with $b = b' - 3n, s = s_2$

**Lemma 4 (Comparison)** *Using the above protocol*

1. *If* $|\mathcal{S}_1| \geq 2^{4n+1}|\mathcal{S}_2|$ *then* $\Pr(P \leftrightarrow V$ *accepts*$) \geq 1 - 2^{1-n}$
2. *If* $|\mathcal{S}_1| \leq 2^{n-4}|\mathcal{S}_2|$ *then* $\Pr(P^* \leftrightarrow V$ *accepts*$) \leq n^3 N^3 2^{9-n}$ *for any* $P^*$

**Proof** Let $d = 2^n$.

1. Pick $b' = \lfloor \log |\mathcal{S}_1| \rfloor$. Then each protocol accepts with probability $\geq 1 - 2^{-n}$, so that both will accept with probability $\geq 1 - 2^{1-n}$ by the upper and lower bound lemmas.
2. There are two cases depending on what $b'$ $P$ chooses
   (a) If $b' \geq \lceil \log |\mathcal{S}_1| \rceil - n$ then by the lower bound lemma the probability of $V$ accepting is $\leq n^3 N^3 2^{9-n}$
   (b) If $b' < \lceil \log |\mathcal{S}_1| \rceil + n$ then by the hypothesis $b' - 3n < \lfloor \log |\mathcal{S}_2| \rfloor - n - 3$ and by the upper bound lemma the probability of $V$ accepting is $\leq 2^{-n}$. □

Using Carter-Wegman Hashing to show a set is large was introduced by Sipser [S] and used extensively in [S, B, GS]. To the author's knowledge this paper is the first use of an interactive protocol to show a set is small.

# 5 Main Theorem

We will start with a simple version of the theorem:

**Theorem 1** *Let $L$ be a language with a perfect zero-knowledge interactive protocol. Then there exists an interactive protocol accepting $\overline{L}$.*

## 5.1 Structure of Proof

We are given a prover and verifier ($P$ and $V$) for the language $L$, and a simulator $M$ that produces views of conversations between $P$ and $V$ and the random coin tosses of $V$. Note that one can simulate the computation of $V$ in polynomial time, checking, for example, whether or not $V$ accepts. On $x \in L$, $M$ produces a view of a conversation from exactly the same probability distribution as when $P$ and $V$ run on $x$. The key idea of the proof of the theorem is to notice what $M$ might do on $x \notin L$. There is no requirement in the definition of perfect zero-knowledge on what $M$ does on $x \notin L$; however there are three possibilities:

1. $M$ will produce "garbage", something that clearly is not a randomly selected member of $P \leftrightarrow V[x]$.
2. $M$ will produce views of conversations that cause $V$ to reject most of the time.
3. $M$ will produce a simulation that looks valid and causes $V$ to accept. It may not be possible in polynomial time to tell this view from one created by $P$ and $V$ when $x \in L$. However, $M$ must be producing views of conversations from a distribution quite different from the distribution of views between $P$ and $V$, since in the real views $V$ is likely to reject.

We will create a new prover and verifier, $P'$ and $V'$ that will determine if one of the three cases occur. $V'$ will run $M$ and get a view of a conversation between $P$ and $V$ and $r$, the random coin tosses of $V$. $V'$ will check that this view is valid and that $V$ halts accepting. If the view fails this test then it falls in cases 1 or 2 so $V'$ knows that $x \notin L$ and $V'$ accepts. Otherwise $V'$ will send to $P'$ some initial segment of the conversation. $P'$ will then convince $V'$ that the conversation came from a bad distribution by "predicting" $r$ better than $P'$ could have done from a good distribution.

## 5.2 An Example: Graph Isomorphism

Graph isomorphism is a well studied problem that is clearly in NP but not known to be in co-NP or BPP. A perfect zero-knowledge proof of graph isomorphism was presented in [GMW]. We will show how our theorem converts this perfect zero-knowledge protocol to an interactive protocol for graph non-isomorphism. This protocol for graph non-isomorphism is identical to the graph non-isomorphism protocol described in [GMW]; our proof, however, shows that the similarity between the two protocols is not coincidental.

Let $\pi : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ be a permutation of the vertices of a graph. For a graph $G = (V, E)$ let $(\pi(v_1), \pi(v_2)) \in \pi(E) \Leftrightarrow (v_1, v_2) \in E$. Let $\pi(G) = (V, \pi(E))$.

Two graphs $G_1$ and $G_2$ are *isomorphic* if there exists a permutation $\pi$ such that $\pi(G_1) = G_2$. A perfect zero-knowledge protocol for graph isomorphism suggested by [GMW] works as follows:

$P$: Generate random permutation $\pi$ and computes $G = \pi(G_1)$

$P \rightarrow V$: $G$

$V \rightarrow P$: $i = 1$ or 2 chosen at random

$P \rightarrow V$: $\pi'$ chosen at random such that $\pi'(G_i) = G$

If $G_1 \cong G_2$ then $G$ will be a permutation of both $G_1$ and $G_2$ and $P$ will always be able to find a $\pi'$. If $G_1 \ncong G_2$ then $G$ cannot be a permutation of both $G_1$ and $G_2$, so at least half of the time $V$ will choose an $i$ such that no $\pi'$ exists. Thus we have an interactive protocol for graph isomorphism. This protocol also is perfect zero-knowledge.

The simulator $M$ works as follows:

$M$ generates $\pi$ and $i$ at random and computes $G = \pi(G_i)$, then outputs the following view of a conversation:

$r$: $i$

$P \rightarrow V$: $G$

$V \rightarrow P$: $i$

$P \rightarrow V$: $\pi$

It is easy to verify that when $G_1 \cong G_2$, $M$ produces exactly the same distribution of views of conversations as $P$ and $V$. Notice what happens when $G_1 \ncong G_2$. The output of $M$ always causes $V$ to accept. Thus when $G_1 \ncong G_2$, $M$ must produce views of conversations from a very different distribution from what $P$ and $V$ produce. In fact whenever $G_1 \ncong G_2$, one can always predict $r = i$ from the $G$ produced by $M$. This leads to a new interactive protocol between a new prover and verifier, $P'$ and $V'$, for graph non-isomorphism as follows:

$V'$: Generate $\pi$ and $i$ at random and compute $G = \pi(G_i)$

$V' \rightarrow P'$: $G$

$P' \rightarrow V'$: $i$

## 5.3 The Protocol for $\bar{L}$

We are given a prover and verifier, $P$ and $V$ for a language $L$ and a simulator $M$ that exactly simulates views of conversations between $P$ and $V$ when $x \in L$. Let $n = |x|$ and let $k$ be the number of rounds of the protocol which is bounded by a polynomial in $n$. We can decrease the probability of error in the protocol between $P$ and $V$ to $2^{-n^t}$ for any constant $t$ by the standard trick of running the protocol several times in parallel and having $V$ accept if the majority of individual protocols accept. This new protocol is still perfect zero-knowledge—we just run the simulator in parallel. Note that we make use of the fact that we only need a simulator for the real verifier $V$. In general, it is not known whether perfect zero-knowledge protocols remain perfect zero-knowledge when run in parallel.

Thus we can assume:

1. If $x \in L$ then $\Pr(P \leftrightarrow V(x) \text{ accepts}) \geq 1 - 2^{-6kn}$
2. If $x \notin L$ then $\forall P^* \Pr(P^* \leftrightarrow V(x) \text{ accepts}) \leq 2^{-6kn}$

For the sake of the comparison protocol, let us require that $V$ immediately rejects if all its coin tosses are zero. Since this will happen with an exponentially small probability it will not affect the correctness of the protocol. The protocol remains perfect zero-knowledge by having the simulator output no conversation if the verifier's coins are all zero.

A protocol between a new prover and verifier, $P'$ and $V'$, works as follows:

$V'$: Run $M$ and get $r, \beta_1, \alpha_1, \ldots, \beta_k, \alpha_k$. $V'$ now checks two things:

1. Check that the conversation is *valid*, i.e., that $r, \alpha_1, \ldots, \alpha_k$ will cause $V$ to say $\beta_1, \ldots, \beta_k$.
2. Check that the conversation causes $V$ to accept.

If either of these tests fail then $V'$ can be very sure that $x \notin L$ so $V'$ quits now and accepts. Otherwise $V'$ continues.
Let $j = 1$.

$V' \rightarrow P'$: $\beta_j, \alpha_j$

$P' \rightarrow V'$: Look at the sets $\mathcal{R}_1$ and $\mathcal{R}_2$ as defined below. If $|\mathcal{R}_1| \geq 2^{4n+1}|\mathcal{R}_2|$ then use the comparison protocol described in section 4 to show $|\mathcal{R}_1| \gg |\mathcal{R}_2|$. Otherwise let $j = j + 1$. If $j \leq k$ tell $V'$ to TRY NEXT ROUND, otherwise GIVE UP.

$\mathcal{R}_1$ can be thought of as all the possible random strings of $V$ after round $j$ of the protocol. $\mathcal{R}_2$ are the possible random strings of $V$ generated by $M$. More formally:

Let $\mathcal{R}$ be the set of all possible coin tosses of $V$.
Let $\mathcal{R}_1 = \{R \in \mathcal{R} | R \text{ and } \alpha_1, \ldots, \alpha_j \text{ cause } V \text{ to say } \beta_1, \ldots, \beta_j\}$.
Let $\mathcal{R}_2 = \{R \in \mathcal{R} | M \text{ can output } R, \beta_1, \alpha_1, \ldots, \beta_j, \alpha_j \text{ part of a valid, accepting conversation}\}$

Note that $\mathcal{R}_2 \subseteq \mathcal{R}_1$ and if $x \in L$ then $\mathcal{R}_2 \approx \mathcal{R}_1$. Also note that $\mathcal{R}_1$ is independent of $\alpha_j$.

$\mathcal{R}_1$ is polynomial time testable. If $x \in L$ then $M$ produces the exact distribution between $P$ and $V$ and thus $r$ is a random element of $\mathcal{R}_2$ which $P'$ doesn't know. In that case we have fulfilled the requirements of the comparison protocol. If $x \notin L$ it is possible that $r$ is not a random element of $\mathcal{R}_2$ but this can only increase the probability of the comparison protocol accepting.

### 5.4 The Protocol Constitutes an Interactive Protocol for $\bar{L}$

To show that this is an interactive protocol for $\overline{L}$, we must show two things:

1. If $(x \in \overline{L})$ then $P' \leftrightarrow V'(x)$ accepts with probability $\geq \frac{2}{3}$
2. If $(x \notin \overline{L})$ then $\forall \hat{P}\ \hat{P} \leftrightarrow V'(x)$ accepts with probability $\leq \frac{1}{3}$

We will prove the second statement first since it is the easier of the two to prove.

2. Suppose $x \in L$. Then $M$ will produce views of conversations from exactly the same distribution as $P$ and $V$. Thus every conversation produced by $M$ will be valid. Assume a prover $P$ is able to convince $V'$ to accept with probability $\geq \frac{1}{3}$. There may be an exponentially small chance that $V$ will reject in this conversation and this will cause $V'$ to accept. If $|\mathcal{R}_1| \leq 2^{n-4}|\mathcal{R}_2|$ on any round then the comparison protocol will accept with a exponentially small probability. Thus we can assume that with probability $> \frac{1}{4}$ that $|\mathcal{R}_1| > 2^{n-4}|\mathcal{R}_2|$ for some round $j$. Since $M$ outputs all possible conversations, $\mathcal{R}_2$ is just the random coin tosses of $V$ which might cause $V$ to accept in the future. So at round $j$ of the protocol, the probability of $V$'s acceptance $< \frac{|\mathcal{R}_2|}{|\mathcal{R}_1|} \leq 2^{4-n}$. Since this happens at least a fourth of the time the probability of $V$'s acceptance in in general is $\leq \frac{3}{4} + 2^{4-n}$ which contradicts the fact that $V$ will accept with probability greater than $1 - 2^{-6kn}$.

1. Suppose to the contrary that $x \notin L$ and the protocol does not work. If $|\mathcal{R}_1| < 2^{4n+1}|\mathcal{R}_2|$ then by the comparison lemma the comparison protocol will fail with at most an exponentially small probability. So $|\mathcal{R}_1| \geq 2^{4n+1}|\mathcal{R}_2|$ at all rounds $j$ with probability at least one fourth. We use this to derive a contradiction by demonstrating that $P \leftrightarrow V$ is not an interactive proof system for $L$ by presenting a prover $P^*$ that will convince $V$ (the original verifier) that $x \in L$ with probability greater than $2^{-6kn}$.

At round $j$ suppose the conversation so far has been $\beta'_1, \alpha'_1, \ldots, \beta'_j$. $P^*$ works as follows:

$P^*$: Run $M$ which outputs $r, \beta_1, \alpha_1, \ldots, \beta_k, \alpha_k$. Check that this is a valid accepting conversation. If not, try again. See if $\beta_1, \alpha_1, \ldots, \beta_j = \beta'_1, \alpha'_1, \ldots, \beta'_j$. If not, try again.

$P^* \rightarrow V$: $\alpha_j$

At round $j$ when $P^*$ has a conversation from $M$ that matches the conversation so far, $\mathcal{R}_1$ is the set of possible random coin tosses of $V$. When $P^*$ says $\alpha_j$, $\mathcal{R}_2$ is the set of coin tosses of $V$ that will still keep $V$ heading towards an accepting path. Since $|\mathcal{R}_2| \geq 2^{5n+1}|\mathcal{R}_1|$, this will happen with probability $\geq 2^{-(5n+1)}$. So after $k$ rounds, $V$ will end up accepting with a probability at least $\frac{1}{4}2^{-(5kn+k)}$ which is higher than the $2^{-6kn}$ maximum accepting probability we assumed for $V$ and any $P^*$, when $x \notin L$. □

Note that $P^*$ may require exponential expected time to complete its part of the protocol but in our model an infinitely powerful $P^*$ is allowed.

## 6 Extensions and Corollaries

**Theorem 2** *Suppose $P \leftrightarrow V$ is an interactive protocol for a language $L$ and there is a probabilistic polynomial time simulator $M$ such that $M[x]$ is statistically close to $P \leftrightarrow V[x]$. Then there is a single round interactive protocol for the complement of $L$.*

**Idea of Proof** This extends the main theorem in two ways. First, we do not require $M[x] = P \leftrightarrow V[x]$, just that they be statistically close. One can check the proof in the previous section and notice that, with some minor adjustments to the probabilities, statistically close is good enough.

Second, we would like to get a single round protocol for the complement of $L$. Notice that in the protocol given above the number of rounds is dependent on when $P'$ decides to say STOP. To get bounded rounds we must make the following change to the protocol:

$V'$: Run $M$ $k^3$ times independently and get $k^3$ views of conversations; check that each conversation is valid and accepting.

$V' \rightarrow P'$: For $1 \leq i \leq k^3$ send the first $i \bmod k$ rounds of the $i$th conversation.

$P' \rightarrow V'$: Pick any conversation j and show $|\mathcal{R}_1| \gg |\mathcal{R}_2|$ for the view of that conversation.

It is not hard to verify that the above proof still works for this new protocol. Once we have bounded rounds we apply the theorems of [B, GS] which imply that all bounded round protocols can be made into single round protocols.

Some trivial corollaries that follow from results that are described in section 3:

**Corollary 1** *If $L \in APZK$ then $\overline{L} \in AM$.*

**Corollary 2** *If $L$ has an almost perfect zero-knowledge interactive protocol (possibly with an unbounded number of rounds) then $L \in (NP \cap co\text{-}NP)^R$, where $R$ is a random oracle.*

**Corollary 3** *If any NP-complete language has an almost perfect zero-knowledge interactive protocol then the polynomial time hierarchy collapses to the second level.*

**Corollary 4** *If there are one-way functions and the polynomial time hierarchy does not collapse then $NP \subseteq ZK$; but $NP \not\subseteq APZK$, so $ZK \neq APZK$.*

# 7 Open Problems

There are several interesting problems remaining, including:

- What is the relationship between PZK and APZK?
- Are complement of perfect or almost perfect zero-knowledge languages themselves perfect zero-knowledge in any sense?
- Are cryptographic assumptions necessary to show NP has zero-knowledge protocols? Although this paper shows that NP probably doesn't have perfect zero-knowledge proof systems, it is still conceivable that the intractability of SAT is a good enough assumption for a zero-knowledge protocol.

# 8 Acknowledgements

The author would like to express his gratitude to his advisor, Mike Sipser, for his support and encouragement.

The author would also like to thank Mike, Silvio Micali, Oded Goldreich, Joan Feigenbaum, Paul Beame, Eric Schwabe and Su-Ming Wu for their useful comments on this paper.

# References

[AGH] Aiello, W., S. Goldwasser and J. Hastad, "On the power of Interaction", *Proc. 27th FOCS*, 1986, pp.368-379.

[AH] Aiello, W. and J. Hastad, "Statistical Zero-Knowledge Languages can be Recognized in Two Rounds", *JCSS*, to appear. Extended abstract available in *Proc. 28th FOCS*, 1987, pp. 439-448.

[B] Babai, L., "Trading Group Theory for Randomness", *Proc. 17th STOC*, 1985, pp. 421-429.

[BHZ] Boppana, R., J. Hastad and S. Zachos, "Does co-NP Have Short Interactive Proofs?", *IPL*, to appear.

[BC] Brassard, G. and C. Crépeau, "Non-Transitive Transfer of Confidence: A Perfect Zero-Knowledge Interactive Protocol for SAT and Beyond", *Proc. 27th FOCS*, 1986, pp. 188-195.

[CW] Carter, J.L. and M.N. Wegman, "Universal Classes of Hash Functions", *JCSS* **18** 2, 1979, pp.143-154.

[GMW] Goldreich, O., S. Micali and A. Wigderson, "Proofs that Yield Nothing But their Validity and a Methodology of Cryptographic Protocol Design", *Proc. 27th FOCS*, 1986, pp. 174-187.

[GMR] Goldwasser, S., S. Micali and C. Rackoff, "The Knowledge Complexity of Interactive Proof-Systems", *Proc. 17th STOC*, 1985, pp. 291-304.

[GS] Goldwasser, S. and M. Sipser, "Private Coins versus Public Coins in Interactive Proof Systems". In S. Micali, editor, *Randomness and Computation*, Volume 5 of *Advances in Computing Research*, JAI Press, 1987. Extended Abstract available in *Proc. 18th STOC*, 1986, pp. 59-68.

[S] Sipser, M., "A Complexity Theoretic Approach to Randomness", *Proc. 15th STOC*, 1983, pp. 330-335.