

Constant-Round Zero-Knowledge Proof Systems for NP

Sotnikov Dmitry

The School of Computer Science
Tel Aviv University

March 9, 2009

Overview

1. Lower bound for interactive CZK proofs
2. 5-round ZK proof system for NP
3. 7-round ZK proof system for NP

Lower Bound

Fact

Jonathan Katz shows in his article “Which Languages Have 4-Round Zero-Knowledge Proofs?” that 4-round black-box CZK proofs, even with imperfect completeness, exist only for languages whose complement is in MA. This result is unconditional, and holds independent of any cryptographic assumption one might make.

Lower Bound

Fact

Other than the fact that the bound holds only with respect to black-box simulation, this result is essentially the best one could hope for:

- ▶ *Assuming the polynomial hierarchy does not collapse, this result indicates that 5-round CZK proofs for NP are optimal in the round complexity.*
- ▶ *The result applies only for proofs, but not arguments.*

MA Definition

Definition

$L \in MA$ if there exists a probabilistic polynomial-time verifier V , a non-negative function s , and a polynomial p such that the following hold for all sufficiently-long x :

- ▶ If $x \in L$ then there exists a string ω (that can be sent by Merlin) such that $\Pr[V(x, \omega) = 1] \geq s(|x|) + \frac{1}{p(|x|)}$
- ▶ If $x \notin L$ then for all ω (sent by a cheating Merlin) it holds that $\Pr[V(x, \omega) = 1] \leq s(|x|)$

Interactive Proof System

Definition

A pair of interactive machines $\langle P, V \rangle$ is called an interactive proof system for a language L if machine V is polynomial-time and the following two conditions hold with respect to some negligible function $\nu(\cdot)$:

- ▶ **Completeness:** For every $x \in L$,
 $\Pr[\langle P, V \rangle(x) = 1] \geq 1 - \nu(|x|)$
- ▶ **Soundness:** For every $x \notin L$, and every interactive machine B , $\Pr[\langle B, V \rangle(x) = 1] \leq \nu(|x|)$

In case that the soundness condition is required to hold only with respect to a computationally bounded prover, $\langle P, V \rangle$ is called an interactive argument system.

Zero-Knowledge

Definitions

Let $\langle P, V \rangle$ be an interactive proof system for a language L . We say that $\langle P, V \rangle$ is zero-knowledge, if for every probabilistic polynomial-time interactive machine V^* there exist a probabilistic polynomial-time algorithm S^{V^*} such that the ensembles $\{view_{V^*}^P(x)\}_{x \in L}$ and $\{S^{V^*}(x)\}_{x \in L}$ are computationally indistinguishable.

Example of ZK Blum's Protocol

Example

Common Input: A graph $G = (V, E)$ with $n := |V|$

Auxiliary Input to Prover: A Hamiltonian Cycle, $C \subset E$, in G .

- ($\hat{p}1$): Pick a random permutation π of vertices V and commit (using a perfect binding commitment) to the adjacency matrix of the $\pi(G)$.
- ($\hat{v}1$): Send a random chosen bit $\sigma \in \{0, 1\}$.
- ($\hat{p}2$): If $\sigma = 1$, send π to the verifier along with the revealing (i.e., preimages) of all commitments. Otherwise, reveal only the commitments to entries $(\pi(i), \pi(j))$ with $(i, j) \in C$. In both cases also supply the corresponding decommitments.
- ($\hat{v}2$): If the decommitment is valid in regard to the bit σ accept otherwise reject.

Blum's Protocol Analysis

Fact

- ▶ *In class we saw that Blum's protocol is an interactive proof with perfect completeness and $\frac{1}{2}$ error soundness.*
- ▶ *We saw a simulator S that tries to guess σ 's value, and it succeeds with probability $\frac{1}{2}$ (the probability of success based on the computational hiding of prover's commitment).*

Sequential Repetition

Reminder:

In class we saw that sequential repetition of Blum's protocol can decrease the soundness error, but it increases the number of rounds.

The naive approach:

Protocol:

- ▶ Prover concatenates all the $comm(\pi_i(G))$ commitments from the $(\hat{p}1)$ rounds and sends them together to verifier.
- ▶ Verifier sends a random string $\sigma = \sigma_1, \sigma_2, \dots, \sigma_n$.
- ▶ Prover can open the $comm(\pi_i(G))$ commitment with respect to σ_i value.

Sequential Repetition

The naive approach:

Simulator:

- ▶ Simulator S commits commitment of dummy values.
- ▶ After that he sees the string σ and “rewinds” the verifier back to step 1. He commits a new commitment that is consistent with respect to σ' 's value.

But this does not work because σ may depend on the commitment that the verifier got from the prover. The commitment may change due to change of σ , and the probability that the prover guesses σ is negligible.

So What Can We do?

Idea:

We make the verifier choose σ before he sees the commitment, and make him give a commitment on this σ to the prover.

Attention

Because we want to get an interactive proof system and not an interactive argument system we need the verifier's commitment scheme to be perfect-hiding scheme.

5-round Protocol

Definition

Protocol's Input: Common and prover's auxiliary inputs same to Blum's protocol.

Step (P1): Send first message m for perfectly hiding commitment scheme

Step (V1): Commit to random $\sigma \in \{0, 1\}^n$ by the perfect hiding scheme.

Step (P2): Send a concatenation of n commitments $comm(\pi_i(G))$ where π_i is a random permutation. ($c_i := comm(\pi_i(G))$).

5-round Protocol

Definition

Step (V2): Decommit to σ .

Step (P3): If the verifier decommitment is valid then decommit c_i with respect to σ_i , otherwise halt.

Step (V3): If the provers decommitment is valid with respect to σ then accept otherwise reject.

5-round Protocol

Theorem

The 5-round protocol is an interactive proof system.

Proof.

- ▶ Completeness: Clearly, the verifier always accepts a common input in HC .
- ▶ Soundness: If the graph is non-Hamiltonian the probability that all the commitments will be opened legally is : $(1 - \frac{1}{2} + \nu)^n < (\frac{3}{4})^n$ and for big enough n the probability is negligible. (ν is the probability to open the commitment to non-initialized value).



Simulator

Definition

Input: A graph $G = (V, E)$ with $n := |V|$

Initialization: The simulator M^* selects a random tape r for V^* .

Step (S1): Send first message m for perfectly hiding commitment scheme, and get from V^* the commitment to $\sigma \in \{0, 1\}^n$.

Step (S2): Generate n random commitments to dummy values and send to V^* . Get from V^* the decommitment to σ . If V^* failed to open the commitment output **ABORT**.

Simulator

Definition

- Step (S3):** Rewind V^* to the step after σ 's commitment and send the n commitments for the graph G permutation or Hamiltonian cycle with respect to the σ he sees. Get from V^* the decommitment to σ . (Repeat this step until V^* succeeds to decommit the commitment)
- Step (S4):** Decommit the commitments on the graphs, and output the verifier's "accept view".

Run Time Analysis

Simple case:

- ▶ If the verifier always reveals correctly the commitment made in Step (V1) the simulator performed the Step (S3) only once.
- ▶ Consider now a more complex case in which verifier reveals correctly only with probability $\frac{1}{3}$
 - ▶ The probability to open the commitment correctly in Step (S1) is approximately $\frac{1}{3}$
 - ▶ The probability to open the commitment correctly in Step (S3) is $p \approx \frac{1}{3}$ (because the computational secrecy of the prover's commitment)

It follows that the expected number of times Step (S3) is invoked when running the simulator is $\frac{1}{3} \cdot \frac{1}{p} \approx 1$.

Run Time Analysis

General case:

- ▶ Let $q = q(G, r)$ denote the probability that program V^* , on input graph G and random tape r , correctly reveals the commitment made in Step (V1), after receiving random commitments to dummy values.
- ▶ Likewise we denote by $p = p(G, r)$ the probability that V^* , correctly reveals the commitment made in Step (V1), after receiving random commitments to values that satisfies σ .
- ▶ As before the difference between q and p is negligible.

Non-polynomial running time:

The problem is that $p - q < \nu \not\leftrightarrow \frac{q}{p} < \text{poly}(n)$ because for $p = 2^{-n}$ and $q = 2^{-\frac{n}{2}}$ the difference is negligible and yet $\frac{q}{p}$ is not bounded by $\text{poly}(n)$.

Simulator Modification

The modified simulator definition

- ▶ Denote the modified simulator by M^{**} .
- ▶ We add an intermediate Step, denoted by (S2.5), to be performed only if the simulator did not halt in Step (S2).
- ▶ The purpose of Step (S2.5) is to provide a good estimate of q . The estimate is computed by repeating the experiment of (S2) but with the same dummy value until a fixed (polynomial in n let denote it by $F(n)$) number of correct V^* – *reveals* are encountered.
- ▶ The number of trials is not necessarily a polynomial but is rather $F(n)/q$, on the average. Denote it by $R(n)$.

Simulator Modification

The modified simulator output changes

- ▶ Step (S3) of the simulator is modified by adding a bound on the number of times it is performed, and if none of these executions yields a correct V^* -reveal then the simulator outputs a special symbol indicating time-out.
- ▶ Specifically, Step (S2) will be performed at most $F(n)/\bar{q}$ times, (that equal to $R(n)$), where \bar{q} is the estimate to q computed in Step (S2.5)
- ▶ In addition, we modify the simulator so that if the verifier ever reveals a correct opening of the commitment that different from the one recorded in Step (S2) then the simulator halts outputting a special symbol indicating ambiguity.

Modified Simulator Running Time Analysis

Theorem

The running time of the modified simulator is polynomial in n .

Proof.

Denote by $p_i(\cdot)$ the polynomial bound on the work required in order to perform Step (S_i) single execution. So we have the expected running time of the simulator:

$$\begin{aligned} & p_1(n) + (1 - q) \cdot p_2(n) + \\ & q \cdot (p_2(n) + R(n) \cdot p_{2.5}(n) + R(n) \cdot p_3(n) + p_4(n)) \leq \\ & p_1(n) + p_2(n) + F(n) \cdot p_{2.5}(n) + F(n) \cdot p_2(n) + p_4(n) = \text{poly}(n) \end{aligned}$$

□

Modified Simulator Outputs Distribution

Theorem

The probability that the modified simulator outputs the time-out symbol is a negligible function of n .

Proof.

Let $\Delta(G, r)$ be the probability that, on a graph G and coin tosses r , the modified simulator outputs a special time-out symbol. Then

$$\begin{aligned}\Delta(G, r) &= q(G, r) \cdot \sum_{i \geq 1} \text{Prob}(\lfloor 1/q \rfloor = i) \cdot (1 - p(G, r))^{i \cdot F(n)} \\ &< q(G, r) \cdot (\text{Prob}(\frac{q(G, r)}{q} = \Theta(1)) \cdot (1 - p(G, r))^{\frac{F(n)}{q(G, r)}} \\ &\quad + \text{Prob}(\frac{q(G, r)}{q} \neq \Theta(1))) \\ &< q(G, r) \cdot (1 - p(G, r))^{F(n)} + \frac{1}{2^n}\end{aligned}$$



Proof (continue)

Proof.

- ▶ We now show that $\Delta(G, r)$ is a negligible function of n .
- ▶ Assume, to the contrary, that there exist a polynomial $P(\cdot)$, an infinite sequence of graphs $\{G_n\}$ (with $|G_n| = n$), and an infinite sequence of random tapes $\{r_n\}$, such that $\Delta(G_n, r_n) > \frac{1}{P(n)}$.
- ▶ It follows that for each such n we have $q(G_n, r_n) > \frac{1}{P(n)}$.
- ▶ We consider two cases:



Proof (continue)

Proof.

Case 1: For infinitely many of these n 's, it holds that $p(G_n, r_n) \geq q(G_n, r_n)/2$. In such a case we get for these n 's:

$$\Delta(G_n, r_n) \leq (1 - p(G_n, r_n))^{F(n)/q(G_n, r_n)} \leq \left(1 - \frac{q(G_n, r_n)}{2}\right)^{F(n)/q(G_n, r_n)} < 2^{-\text{poly}(n)/2} \text{ which}$$

contradicts our hypothesis that $\Delta(G_n, r_n) > \frac{1}{P(n)}$.

Case 2: For infinitely many of these n 's, it holds that $p(G_n, r_n) < q(G_n, r_n)/2$. It follows that for these n 's we have $|q(G_n, r_n) - p(G_n, r_n)| > P(n)/2$ which leads to contradiction of the computational secrecy of the commitment scheme (used by the prover).

Hence, contradiction follows in both cases.



Modified Simulator Outputs Distribution

Theorem

The probability that the modified simulator outputs the ambiguity symbol is a negligible function of n .

Proof.

- ▶ The problem to use the standard argument is that the number of times Step (S2) is performed is not bounded by a polynomial, only the expected number of times is bounded by a polynomial.
- ▶ The problem is easily resolved by disregarding the executions of the modified simulator in which Step (S3) is performed too many times.



Modified Simulator Outputs Distribution

Proof.

- ▶ Assume by contradiction that the ambiguity symbol is output with probability at least $1/P(n)$, for a polynomial $P(\cdot)$ and an infinite sequence of graphs.
- ▶ Then, we can truncate the execution of M^{**} in which Step (S3) is performed more than $2T(n) \cdot P(n)$ times, where $T(\cdot)$ denotes the expected running time of M^{**} .
- ▶ By averaging argument ($Pr(|X| > a) \leq \frac{E(|X|)}{a}$) it follows that also in these truncated executions M^{**} outputs an ambiguity symbol with non-negligible probability (at least $1/2P(n)$).
- ▶ Contradiction now follows using the standard techniques.



Indistinguishability of Outputs

Theorem

The ensemble $\{M^(G)\}_{G \in HC}$ is computational indistinguishable from ensemble $\{(P, V^*)(G)\}_{G \in HC}$, where $(P, V^*)(G)$ denotes the output of V^* after an interaction with the prover on common input.*

Indistinguishability of Outputs

Proof.

- ▶ Here we have the same problem like in the previous proof, so we can try the same solution by truncating the rare executions of M^* which are too long.
- ▶ Assume that there exists an efficient algorithm A that can distinguish between the distributions with gap $\epsilon(G)$.
- ▶ $|\text{Prob}[A(M^*(G)) = 1] - \text{Prob}[A((P, V^*)(G)) = 1]| = \epsilon(G)$, and that $\epsilon(G) \geq 1/P(n)$ for a polynomial $P(\cdot)$ and an infinite sequence of graphs $\{G_n : n \in S\}$ (with $|G_n| = n$).
- ▶ Defining a predicate R so that $R(y) = 1$ if y is an interaction-transcript in which the verifier correctly reveals the commitment made by verifier and $R(y) = 0$ otherwise, we consider two cases



Indistinguishability of Outputs

Case 1:

- ▶ For infinitely many $n \in S$, it holds that $\text{Prob}(R((P, V^*)(G_n)) = 1) \geq \epsilon(G_n)/3$.
- ▶ On these G'_n s, it is guaranteed that the expected number of times that Step (S3) is performed is at most $3/\epsilon(G_n) < 3P(n)$.
- ▶ Hence, the probability of M^* runs in which Step (S3) is repeated more than $T(n) := 6P(n)^2$ is at most $1/2P(n)$.
- ▶ It follows that algorithm A still distinguishes, with gap at least $\epsilon(G_n) - 1/2P(n)$, the output of the truncated M^* from the real interaction with the prover.
- ▶ At this point, we can apply the standard techniques.

Indistinguishability of Outputs

Case 2:

- ▶ For infinitely many $n \in S$, it holds that $\text{Prob}(R((P, V^*)(G_n)) = 1) < \epsilon(G_n)/3$.
- ▶ It follows that, on these G'_n s, with probability at least $1 - \epsilon(G_n)/3$, the interaction of V^* with the real prover is suspended at Step (V2).
- ▶ There are two sub-cases to consider:

Indistinguishability of Outputs

Sub-case 1:

- ▶ The simulator halts in Step (S2) with probability at most $1 - \epsilon(G_n)/2$.
- ▶ Thus, there is a gap, of at least $\epsilon(G_n)/6$ between the probability that V^* correctly reveals its commitments when interacting with the prover and the probability that V^* correctly reveals its commitments when “interacting” with the simulator.
- ▶ So V^* can be used to distinguish the commitments to dummy values (as produced by the simulator) from commitment to the Hamiltonian graph (as produced by the prover), in contradiction to the computational secrecy of the prover's commitment scheme.

Indistinguishability of Outputs

Sub-case 2:

- ▶ The simulator halts in Step (S2) with probability at least $1 - \epsilon(G_n)/2$.
- ▶ This means that both the real and the simulated interactions are suspended with probability at least $1 - \epsilon(G_n)/2$. Hence, algorithm A must distinguish such suspended interactions with gap at least $\epsilon(G_n)/2$.
- ▶ It follows that algorithm A distinguishes commitments to dummy values (as produced by the simulator) from commitment to the Hamiltonian graph (as produced by the prover), in contradiction to the computational secrecy of the prover's commitment scheme.

End of the Proof

Proof.

Since in all cases we reached contradiction to the computational secrecy of the prover's commitment, the claim follows. \square

7-round Protocol

Definition

Common and Auxiliary Inputs: The common and prover's auxiliary input same to Blum's protocol.

Additional parameter: A super-logarithmic function $k(n)$.

Stage 1:

Step (P1): Send first message for perfectly hiding commitment scheme

Step (V1): Commit to random

$\sigma, \{\sigma_i^0\}_{i=1}^k, \{\sigma_i^1\}_{i=1}^k$ s.t. $\sigma_i^0 \oplus \sigma_i^1 = \sigma$
for all i . Where $\sigma \in \{0, 1\}^n$

Step (P2): Send a random k -bit string

$r = r_1, r_2, \dots, r_k$.

Step (V2): Decommit to $\sigma_1^{r_1}, \dots, \sigma_k^{r_k}$.

7-round Protocol

Definition

Stage 2: Engage in the 3-round protocol for Hamiltonian cycle using σ as a challenge:

- (p1): Produce first prover message for HC protocol (as in $(\hat{p}1)$)
- (v1): Decommit to σ and to $\{\sigma_i^{1-r_i}\}_{i=1}^k$
- (p2): Answer σ with second prover message of HC protocol (as in $(\hat{p}2)$)
- (v2): Accept if and only if all corresponding conditions hold (as in $(\hat{v}2)$)

Simulator

Definition

Step (S1): Randomly generate $(P1)$ and obtain
 $(V1) = V^*(G, (P1); s)$.

Step (S2): Randomly generate $(P2)$ and obtain
 $(V2) = V^*(G, (P1), (P2); s)$.

1. If $(V2) \neq \mathbf{ABORT}$, proceed to Step (S3).
2. If $(V2) = \mathbf{ABORT}$, output
 $\langle (P1), (V1), \mathbf{ABORT} \rangle$ and stop.

Simulator

Definition

Step (S3): For $j=1,2,\dots$

1. Randomly generate $(P2)_j$ and obtain $(V2)_j = V^*(G, (P1), (P2)_j; s)$.
2. If $(V2)_j \neq \mathbf{ABORT}$, proceed to Step (S4)
3. If $(V2)_j = \mathbf{ABORT}$ continue

Step (S4):

1. If $(P2) = (P2)_j$, output \perp and stop.
2. If $(P2) \neq (P2)_j$,
 $\exists i \in \{1, \dots, k\}. [(P2)]_i \neq [(P2)_j]_i$. So we have
 $\sigma = \sigma_i^{[(P2)]_i} \oplus \sigma_i^{[(P2)_j]_i}$.

Simulator Running Time Analysis

Polynomial run time

- ▶ Let $\tau = \tau(G, (P1), s)$ be the probability that the verifier V^* does not send an **ABORT** message in message (V2). The probability τ is taken over the random choices of message (P2). (So we have the same probability in Steps (S2) and (S3).1)
- ▶ Denote by $p_i(\cdot)$ the polynomial bound on the work required in order to perform Step (S_i) single execution. So we have the expected running time of the simulator:
$$p_1(n) + (1 - \tau) \cdot p_2(n) + \tau \cdot \left(p_2(n) + \frac{1}{\tau} p_3(n) + p_4(n) \right) \leq p_1(n) + p_2(n) + p_3(n) + p_4(n) = poly(n)$$

Simulator Output Distribution

Theorem

Suppose that the commitment used in Step (p1) is computationally hiding. Then the ensemble $\{S^{v^}(G)\}_{G \in HC}$ is computationally indistinguishable from the ensemble $\{\text{view}_{V^*}^P(G)\}_{G \in HC}$.*

Simulator Output Distribution

Proof.

- ▶ To prove the theorem we define a new simulator \hat{S} that obtains a Hamiltonian cycle $C \subset E$ in G (as auxiliary input) and uses it in order to produce real prover messages whenever it reaches the second stage of the protocol.
- ▶ Specifically, when it reaches the second stage, the hybrid simulator checks whether the original simulator S should output \perp (in which case it also does).
- ▶ We claim that the ensemble $\{S^{V^*}(G)\}_{G \in HC}$ and ensemble $\{\hat{S}^{V^*}(G, C)\}_{G \in HC}$ are computationally indistinguishable.
- ▶ We claim that the ensemble $\{\hat{S}^{V^*}(G, C)\}_{G \in HC}$ conditioned on it not being \perp , is identically distributed to the ensemble $\{\text{view}_{V^*}^P(G)\}_{G \in HC}$.



Simulator Output Distribution

Proof.

- ▶ We claim that the ensemble $\{\widehat{S}^{v^*}(G, C)\}_{G \in HC}$ is computationally indistinguishable from the ensemble $\{\widehat{S}^{v^*}(G, C)\}_{G \in HC}$ conditioned on it not being \perp .



Simulator Output Distribution

Theorem

For any $G = (V, E) \in HC$, the probability that $\widehat{S}^{v^}(G, C) = \perp$ is negligible in $|V|$.*

Proof.

- ▶ Let $G \in HC$ with $n := |V|$.
- ▶ $(P2) = r \in \{0, 1\}^k$
- ▶ Let $\widetilde{V}^* = \widetilde{V}^*((P1), s)$ be the “residual” strategy of V^* when $\langle (P1), s \rangle$ are fixed (i.e.,
 $\widetilde{V}^*(G, r) := V^*(G, (P1), r; s)$)
- ▶ Let $\tau = \tau(G, (P1), s)$ be the probability that the verifier V^* does not send an **ABORT** message in (V2) (as before).



Simulator Output Distribution

Proof.

$$\begin{aligned} & Pr_r[\widehat{S}^{\widetilde{V}^*}(G, C) = \perp] \\ &= Pr_r[\widehat{S}^{\widetilde{V}^*}(G, C) = \perp \mid \widehat{S} \text{ reaches } (S3)] \cdot Pr_r[S \text{ reaches } (S3)] \\ &= Pr_r[\widehat{S}^{\widetilde{V}^*}(G, C) = \perp \mid \widehat{S} \text{ reaches } (S3)] \cdot \tau \\ &= Pr_r[(P2) = (P2)_j] \cdot \tau \end{aligned}$$

Since $(P2)$ and $(P2)_j$ are uniformly and independently chosen from $\{0, 1\}^k$, and since the number of $r \in \{0, 1\}^k$ for which $\widetilde{V}^*(G, r)$ is not equal to **ABORT** is precisely $2^k \cdot \tau$, then it holds that $Pr[(P2) = (P2)_j] = 1/(2^k \cdot \tau)$.

$$Pr_r[\widehat{S}^{\widetilde{V}^*}(G, C) = \perp] = \frac{1}{2^k \cdot \tau} \cdot \tau = \frac{1}{2^k}$$

