NP-Hardness of Approximating Meta-Complexity: A Cryptographic Approach

Jingyi Lyu

IIIS, THU

2024/12/13

1. Introduction
2. Preliminaries
3. Hardness from Oracle Witness Encryption
4. Witness Encryption in Oracle World
5. Discussion
6. References

Meta-Complexity

Meta-complexity is about measuring complexity of computing complexity: How hard is it to compute -

- The minimum size of boolean circuit computing function $f : \{0,1\}^n \to \{0,1\}$? (Circuit complexity)
- The minimum description length of Turing machine that outputs string $x \in \{0,1\}^*$? (Kolmogorov complexity)

Applications of Meta-Complexity

Meta-complexity problems have a variety of applications in computational complexity, cryptography, etc. For example:

- Worst-case to average-case reductions for problems in **PH**; [Hir21]
- Equivalence between one-way functions and average-case hardness of computing K^t; [LP20]
- Learning constant-depth circuits with parity gates in quasi-polynomial time. [CIKK16]

Our Goal

Cryptography can benefit from meta-complexity (basing essential primitives on meta-complexity problems).

Reversely, how can meta-complexity investigation benefit from cryptography?

We will present how ideas of **witness encryption** can be used to *unconditionally* prove NP-hardness of approximating **MOCSP**, a variant problem of computing circuit complexity. [HIR23]

Witness Encryption

Let $L \in \mathbf{NP}$. That is, polynomial-time algorithm $V_L(x, w) : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$ exists such that

$$x \in L \iff \exists w, V_L(x, w) = 1$$

A **witness encryption scheme** for L consists of polynomial-time algorithms $(\mathrm{Encrypt}, \mathrm{Decrypt})$ with syntax

- $\mathrm{Encrypt}(1^\lambda, x, b; r)$ takes as input security parameter 1^λ, instance (of L) x, message bit $b \in \{0,1\}$, randomness r, outputs ciphertext c;
- $\mathrm{Decrypt}(1^\lambda, c, x, w)$ takes as input security parameter, instance, ciphertext and witness w, outputs message bit b.

Witness Encryption

The following conditions hold:

(Correctness) For any $\lambda \in \mathbb{N}, b \in \{0,1\}, x \in L$ and witness w such that $V_L(x,w) = 1$, then

$$\Pr_r \left[\mathrm{Decrypt}(\mathrm{Encrypt}(1^\lambda, x, b; r), x, w) = b\right] = 1$$

((S,ϵ)-Security) For every $\lambda \in \mathbb{N}$, every size-$S(\lambda)$ circuit A and any $x \notin L$,

$$\left| \Pr_r \left[A(\mathrm{Encrypt}(1^\lambda, x, 0; r)) = 1\right] - \Pr_r \left[A(\mathrm{Encrypt}(1^\lambda, x, 1; r)) = 1\right] \right| < \epsilon(\lambda)$$

Minimum Oracle Circuit Size Problem (MOCSP)

Assume truth table of function $f : \{0,1\}^n \to \{0,1\}$ and oracle function $\mathcal{O} : \{0,1\}^{O(n)} \to \{0,1\}$ are given.

Compute $\mathrm{CC}_\delta^{\mathcal{O}}(f)$, which is defined by the minimum size of boolean circuit (with oracle access to $\mathcal{O}$) that computes f correctly on $1-\delta$ fraction of inputs. When $\delta = 0$ we simplify the notation to $\mathrm{CC}^{\mathcal{O}}$.

Oracle access can be characterized as "oracle gate" with fan-in $O(n)$ with computes $\mathcal{O}$.

Approximation (Gap) version of MOCSP

To characterize hardness of approximating MOCSP, we define promise problem $\mathrm{GapMOCSP}[s_1(n), s_2(n), \delta_1(n), \delta_2(n)]$ by

- $(f, \mathcal{O})$ is YES instance: $\mathrm{CC}^{\mathcal{O}}_{\delta_1(n)} \leq s_1(n)$;
- $(f, \mathcal{O})$ is NO instance: $\mathrm{CC}^{\mathcal{O}}_{\delta_2(n)} \geq s_2(n)$.

Previously it's proved that $\mathrm{GapMOCSP}$ is **NP**-hard when gap between s_1, s_2 is in $\mathrm{poly}(n)$, while we are going to show the case with larger approximation factor, i.e., gap between s_1, s_2 being $2^{\Omega(n)}$.

Exact Cover

Apparently, building secure witness encryption requires hard language L in **NP**. We hereby define **exact cover problem** as: given ground set $[n]$ and subsets $X_1, \cdots, X_m \subset [n]$, compute whether there exists index set $\mathcal{I}$, such that

$$\bigcup_{j \in \mathcal{I}} X_j = [n], \quad X_i \cap X_j = \emptyset \ (\forall i, j \in \mathcal{I}, i \neq j)$$

It is proved that exact cover problem is **NP**-Complete. [Kar72]

Main Theorem

Suppose $L \in \mathbf{NP}$, oracle family $\mathscr{O}_\lambda$ exists such that

- $\mathcal{O} : \{0,1\}^{O(\lambda)} \to \{0,1\}$ is sampleable from $\mathcal{U}(\mathscr{O}_\lambda)$ in $\mathrm{poly}(2^\lambda)$ time;
- Witness encryption scheme w.r.t. $\mathcal{O}$ and L exists, that is $(2^{\Omega(\lambda)}, 2^{-\Omega(\lambda)})$-secure with probability $1 - 2^{-\Omega(\lambda)}$ over $\mathcal{O}$,

then for any $0 < \epsilon < 0.3$, there is polynomial-time *randomized* reduction from L to $\mathrm{GapMOCSP}[2^{\epsilon n}, 2^{0.3n}, 0, \frac{1}{2} - 2^{-0.3n}]$.

Intuition

For any function f, imagine some oracle $\mathcal{O}_{\mathrm{ct}}$ consisting with truth table of f, but encrypted by the witness encryption scheme.

- If the instance used in encryption is YES instance, it should be easy to recover f from the encrypted truth table;
- Otherwise, the ciphertext cannot provide significant information of f, so computing f with oracle circuit is hard as computing with raw boolean circuit, which should be hard for most f.

The encryption scheme relies on $\mathcal{O} \leftarrow \mathcal{U}(\mathscr{O})$. If we concatenate $\mathcal{O}$ with $\mathcal{O}_{\mathrm{ct}}$, it should be good choice for reduction output.

Specification of Parameters

Let x be instance of L (as the input of reduction). Assume $|x| = N$. There should exist large enough k such that

- Oracles in family $\mathscr{O}$ requires input length at most $k\lambda$;
- $\mathrm{Encrypt}, \mathrm{Decrypt}$ runs in N^k time;
- Witness length for YES instance is at most N^k;
- Encryption scheme is $(2^{\lambda/k}, 2^{-\lambda/k})$-secure (w.h.p. over oracles).

Let $n = \lceil (3k \log N)/\epsilon \rceil$, $\lambda = 10kn$.

Reduction

Given x, we produce $(f, \mathcal{O}_{\mathrm{r}})$ as instance of $\mathrm{GapMOCSP}$.

1. Truth table of $f : \{0,1\}^n \to \{0,1\}$ is uniformly random over $\{0,1\}^{2^n}$.
2. Sample $\mathcal{O} \leftarrow \mathcal{U}(\mathscr{O}_\lambda)$.
3. For $i \in \{0,1\}^n$, let $c_i \in \{0,1\}^{N^k}$ be encryption of $f(i)$, i.e., $c_i = \mathrm{Encrypt}^{\mathcal{O}}(1^\lambda, x, f(i); r_i)$, where r_i are fresh random bits.
4. Create oracle $\mathcal{O}_{\mathrm{ct}}$ that stores all ciphertexts: For $i \in \{0,1\}^n, j \in \{0,1\}^{\lceil k \log N \rceil}$, $\mathcal{O}_{\mathrm{ct}}(i,j)$ is j-th bit of c_i.
5. Create $\mathcal{O}_{\mathrm{r}} : \{0,1\}^{1+k\lambda} \to \{0,1\}$ that concatenates $\mathcal{O}$ and $\mathcal{O}_{\mathrm{ct}}$. Queries to different parts are distinguished by first bit.

Justification

We need to prove the following:

- **Validity:** Output of reduction has size $\mathrm{poly}(N)$ and given in $\mathrm{poly}(N)$-time. (This is easy to verify.)
- **Completeness:** Output from YES instance produces YES instance w.h.p.
- **Soundness:** Output from NO instance produces NO instance w.h.p.

Completeness

If $x \in L$, then $C^{\mathcal{O}_\mathrm{r}}$ computes $f(i)$ by:

1. Make N^k queries to $\mathcal{O}_\mathrm{ct}$ to determine ciphertext c_i of $f(i)$.
2. Output $\mathrm{Decrypt}^{\mathcal{O}_\lambda}(1^\lambda, c_i, x, w)$.

With w hard-wired into circuit, the size of $C^{\mathcal{O}_\mathrm{r}}$ is at most $N^{2k} \le 2^{\epsilon n}$, so $(f, \mathcal{O}_\mathrm{r})$ is YES instance.

Soundness: Our Plan

If $x \notin L$, we aim to show that w.h.p. over randomness of reduction, for every $2^{0.3n}$-size oracle circuit C,

$$\mathrm{correct}(C^{\mathcal{O}_{\mathrm{r}}}, f) = \Pr_{i \leftarrow \{0,1\}^n}[C^{\mathcal{O}_{\mathrm{r}}} = f(i)] \leq \frac{1}{2} + 2^{-0.3n}$$

We build this result in a two-step:

- First, imagine a "hardest case" where $\mathcal{O}_{\mathrm{r}}$ is independent of f and estimate the failure probability (such that $\mathrm{correct}$ exceeds the bound).
- Second, move gradually to the actual $\mathcal{O}_{\mathrm{r}}$ produced by reduction. Security property guarantees the failure probability doesn't grow too much.

Soundness: The Hybrid Argument

First fix $\mathcal{O} \leftarrow \mathcal{U}(\mathscr{O}_\lambda)$. With probability $1 - o(1)$ our encryption scheme stays secure in this oracle setting.

With natural correspondence between $\{0, 1, \cdots, 2^n - 1\}$ and $\{0, 1\}^n$, we define $2^n + 1$ hybrid worlds $\mathrm{Hyb}_0, \cdots, \mathrm{Hyb}_{2^n}$, while in Hyb_g, production of $\mathcal{O}_{\mathrm{ct}}$ is totally same except for $i < g$, where c_i are replaced by encryption to random bit.

Then we compose new $\mathcal{O}_{\mathrm{ct}}$ with $\mathcal{O}$ to give $\mathcal{O}_{\mathrm{r}}$ in different hybrid worlds.

Soundness: The Hybrid Argument

Apparently, Hyb_0 is the reduction output, while Hyb_{2^n} is the "hardest world" where $\mathcal{O}_r$ is irrelevant to f.

We define a notion of "hardness" to hybrid worlds, which is

$$\mathrm{adv}_g(f) = \max_C \left\{ \Pr_{\mathcal{O}_\mathrm{r} \leftarrow \mathrm{Hyb}_g} \left[\mathrm{correct}(C^{\mathcal{O}_\mathrm{r}}, f) \geq \frac{1}{2} + 2^{-0.3n} \right] \right\}$$

where the maximum is taken over all C with size $\leq 2^{0.3n}$. We want to prove $\mathrm{adv}_0(f)$ is negligibly small for overwhelming portion of f.

Soundness: Step 1

We first prove that w.p. $1 - o(1)$ over f, $\mathrm{adv}_{2^n}(f) \leq 2^{-0.4n}$.

The intuition is: For some good C and $\mathcal{O}_\mathrm{r}$ such that $\mathrm{correct}(C^{\mathcal{O}_\mathrm{r}}, f)$ is non-trivially better than $\frac{1}{2}$, we can describe f if we know all inputs where $C^{\mathcal{O}_\mathrm{r}}$ computes incorrectly, by:

- Using reverted output of $C^{\mathcal{O}_\mathrm{r}}$ if it's on the incorrect position;
- Using original output otherwise.

Since the "incorrect positions" can be encoded into short strings, the description length of f can be slightly shorter than trivial $2^n + O(1)$, and we can upper-bound the number of such f by standard counting argument.

Soundness: Step 1

Let $\gamma = 2^{-0.4n}, \kappa = 2^{-0.3n}$. Assume $\mathrm{adv}_{2^n}(f) > \gamma$, so there exists C with size $2^{0.3n}$ such that

$$\Pr_{\mathcal{O}_\mathrm{r} \leftarrow \mathrm{Hyb}_g}\left[\mathrm{correct}(C^{\mathcal{O}_\mathrm{r}}, f) \geq \frac{1}{2} + \kappa\right] > \gamma$$

The set of incorrect positions is of size $(\frac{1}{2} - \kappa)2^n$, so the total description length of f, conditioned on already-known $\mathcal{O}_\mathrm{r}$, is at most

$$|C| + \log \binom{2^n}{\left(\frac{1}{2} - \kappa\right) 2^n} + O(1) \leq 2^n - \Omega(2^{0.4n})$$

Soundness: Step 1

With probability at least γ over $\mathcal{O}_{\mathrm{r}}$, there is a machine M of description length $2^n - \Omega(2^{0.4n})$ that outputs f. By external counting arguments on *probablistic Kolmogorov complexity* from [GKLO22], such f occupies at most portion of $2^{-2^{0.35n}}$.

Taking union bound over all circuits of size within $2^{0.3n}$, we have $\mathrm{adv}_{2^n}(f) \le 2^{-0.4n}$ with probability $1 - o(1)$ over choice of f.

Soundness: Step 2

Next we proceed to proving the gap between $\mathrm{adv}_g(f)$ and $\mathrm{adv}_{g+1}(f)$. We claim that

$$\mathrm{adv}_g(f) \leq \mathrm{adv}_{g+1}(f) + 2^{-\lambda/k}$$

holds for every f and g. To apply security property of witness encryption, we design an adversary A. Given ciphertext c, $A(c)$ attempts to

1. Produce $\mathcal{O}_{\mathrm{ct}}$ as in Hyb_g, but replace c_i by c.
2. Build the complete $\mathcal{O}_{\mathrm{r}}$.
3. Use $O(2^n|C|)$ oracle queries to compute $\mathrm{correct}(C^{\mathcal{O}_{\mathrm{r}}}, f)$, returns 1 if and only if $\mathrm{correct}(C^{\mathcal{O}_{\mathrm{r}}}, f) \geq \frac{1}{2} + 2^{-0.3n}$.

Soundness: Step 2

Define p_b to be probability $A(c) = 1$ when c is encryption of b. By substituting the parameters, we can verify

- Program length of A is at most $2^n + O(1) < 2^{\lambda/k}$ (since it only hard-wires f), and
- Number of queries made by A is at most $2^{\lambda/k}$.

So we apply security property of encryption scheme to get $|p_0 - p_1| \leq 2^{-\lambda/k}$.

Soundness: Step 2

By definition of adv, we can observe $\mathrm{adv}_g(f) = p_{f(g)}$. Recall that in adv_{g+1}, raw message bit of c_g is uniformly distributed, so $\mathrm{adv}_{g+1} \geq \frac{1}{2}(p_0 + p_1)$; So (no matter $f(g) = 0$ or 1)

$$\mathrm{adv}_g(f) \leq \mathrm{adv}_{g+1}(f) + \frac{1}{2}|p_0 - p_1| \leq \mathrm{adv}_{g+1}(f) + 2^{-\lambda/k}$$

Immediately we can derive

$$\mathrm{adv}_0(f) \leq 2^{-0.4n} + 2^{n-\lambda/k} \leq 2^{-0.3n}$$

So w.h.p. we have $\mathrm{CC}^{\mathcal{O}_{\mathrm{r}}}_{\frac{1}{2}-2^{-0.3n}}(f) > 2^{0.3n}$, i.e., $(f, \mathcal{O}_{\mathrm{r}})$ is NO instance.

Multilinear Map Model

The idea of witness encryption is proposed first in [GGSW13], where they utilized **multilinear map model** and conjectured they have certain notion of security.

Multilinear Map Model

Assume a sequence of cyclic groups $G_1, \cdots, G_{n+1}$ of order p, where we denote identity, generator, group operation by $0, 1, +$ respectively. Define $e_{i,j} : G_i \times G_j \to G_{i+j}$ by

$$e_{i,j}(g_i, g_j) = g_i + g_j$$

And we can naturally generalize this notion into

$$e(g_{i_1}, \cdots, g_{i_k}) = g_{i_1} + \cdots + g_{i_k}$$

where g_{i_j} is from G_{i_j} and the result lies in $G_{i_1+\cdots+i_k}$.

Multilinear Map Model: Using the Hard Language

Assume instance $[n], X_1, \cdots, X_m$ of exact cover is given. We sample $a_1, \cdots, a_n \leftarrow \mathcal{U}(G_1)$ and $r \leftarrow \mathcal{U}(G_n)$. For any subset $X = \{n_{i_1}, \cdots, n_{i_k}\}$, define

$$e(X) = e(a_{n_{i_1}}, \cdots, a_{n_{i_k}}) \in G_{|X|}$$

We denote $e(X)$ by s_X. For any sequence of subset we naturally have

$$e(X_{i_1}, \cdots, X_{i_k}) = e(s_{X_{i_1}}, \cdots, s_{X_{i_k}}) \in G_{|X_{i_1}|+\cdots+|X_{i_k}|}$$

[GGSW13] conjectured that if $[n], X_1, \cdots, X_m$ is NO instance, then

$$(\{s_{X_i}\}, e([n])) \approx_c (\{s_{X_i}\}, r)$$

Multilinear Map Model: Using the Hard Language

Consider the following witness encryption scheme, on the same instance:

- $\mathrm{Encrypt}$ samples $m_0, m_1 \leftarrow \mathcal{U}(G_1)$, then compute $s_{[n+1]} = e(s_{[n]}, m_b) \in G_{n+1}$. On message bit b, output ciphertext $(\{s_{X_i}\}, s_{[n+1]}, m_0, m_1)$.
- $\mathrm{Decrypt}$ parses cipher text into $\{s_{X_i}\}, s^*, m_0, m_1)$. With witness $X_{i_1}, \cdots, X_{i_k}$, it computes $s_{[n]} = e(s_{X_{i_1}}, \cdots, s_{X_{i_k}})$, then check whether $e(s_{[n]}, m_b) = s^*$ over $b = 0, 1$.

If conjecture holds, this encryption scheme indeed has correctness and security.

Introduce the Oracle

To build hardness result of $\mathrm{GapMOCSP}$, we must introduce oracle into the computation domain of encryption scheme. Both $\mathrm{Encrypt}, \mathrm{Decrypt}$ and adversary can perform oracle query.

With appropriate choice of oracle, we can indeed prove that $(\mathrm{Encrypt}, \mathrm{Decrypt})$ is secure against any computationally bounded (bounded number of oracle queries) adversary.

This leads to an unconditional result of $\mathrm{GapMOCSP}$.

Generic Multilinear Map Model

Idea is to obfuscate group elements with permutation.

Imagine there are $n+1$ uniformly random permutations $\sigma_1 : [p] \to G_1, \cdots, \sigma_{n+1} : [p] \to G_{n+1}$. Instead of manipulating group elements directly, we now refer to labels produced by permutations, i.e.,

$$e_{i,j} : [p] \times [p] \to [p], \ (g_i, g_j) \mapsto \sigma_{i+j}^{-1}(\sigma_i(g_i), \sigma_j(g_j))$$

And those generalized notion can be defined similarly.

Generic Multilinear Map Model

We hereby introduce an oracle $\mathcal{O}$. Given security parameter λ, let $p \in (2^{\lambda}, 2^{\lambda+1})$ be a prime, identify $[p]$ with the lexicographically smallest p strings in $\{0,1\}^{\lambda+1}$. Assume certain sequence of permutations $\sigma_1, \cdots, \sigma_n$ is fixed, then we define

$$\mathcal{O} : \{0,1\}^{\lceil \log(n+1) \rceil} \times \{0,1\}^{\lambda+1} \times \{0,1\}^{\lceil \log(n+1) \rceil} \times \{0,1\}^{\lambda+1} \to \{0,1\}^{\lambda+1}$$

by

$$\mathcal{O}(i, g_i, j, g_j) = \sigma_{i+j}^{-1}(\sigma_i(g_i), \sigma_j(g_j))$$

We define the oracle family $\mathscr{O}_\lambda$ consists $\mathcal{O}$ with all possibilities of random permutations.

Security in the Oracle World

We can redefine the encryption scheme, just by replacing group elements with labels, and replacing *e*-evaluation by oracle queries.

The original paper proved the following security statement in the oracle setting:

- There exists $S(\lambda) = 2^{\Omega(\lambda)}, \epsilon(\lambda) = 2^{-\Omega(\lambda)}$ such that, with probability at least $1 - \epsilon(\lambda)$ over $\mathcal{O} \leftarrow \mathcal{U}(\mathscr{O}_\lambda)$, the oracle encryption scheme is (S, ϵ)-secure.

The original proof is tedious without much intriguing idea. We omit it here and revisit it if time permits.

What's More...

This paper [HIR23] discussed hardness of approximating several different meta-complexity problem, as with application of different cryptographic primitives. Those results usually progress significantly in approximation factor, making the hardness result closer to important applications, such as worst-case to average-case reduction.

It is also an inspiring discovery that cryptographic techniques can be used in meta-complexity as such. Subsequent works also arise [Ila23], stepping closer to major breakthrough, for example, **NP**-hardness of original MCSP.

1. Introduction
2. Preliminaries
3. Hardness from Oracle Witness Encryption
4. Witness Encryption in Oracle World
5. Discussion
6. References

References

[CIKK16] Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonima Kolokolova.
Learning algorithms from natural proofs., 2016.

[GGSW13] Sanjam Garg, Craig Gentry, Amit Sahai, and Brent Waters.
Witness encryption and its applications., 2013.

[GKLO22] Halley Goldberg, Valentine Kabanets, Zhenjian Lu, and Igor Carboni Oliveira.
Probabilistic kolmogorov complexity with applications to average-case complexity., 2022.

[Hir21] Shuichi Hirahara.
Average-case hardness of np from exponential worst-case hardness assumptions., 2021.

[HIR23] Yizhi Huang, Rahul Ilango, and Hanlin Ren.
Np-hardness of approximating meta-complexity: A cryptographic approach, 2023.

[Ila23] Rahul Ilango.
Sat reduces to minimum circuit size problem with a random oracle, 2023.

[Kar72] Richard M. Karp.
Learning algorithms from natural proofs., 1972.

[LP20] Yanyi Liu and Rafael Pass.
On one-way functions and kolmogorov complexity., 2020.