Efficient String-Commitment from Weak Bit-Commitment

Kai-Min Chung[⋆1], Feng-Hao Liu[⋆⋆2], Chi-Jen Lu[3], Bo-Yin Yang[3]

[1] School of Engineering & Applied Sciences, Harvard University, Cambridge MA, USA, kmchung@fas.harvard.edu

[2] Department of Computer Science, Brown University, Providence RI, USA, fenghao@cs.brown.edu

[3] Institute of Information Science, Academia Sinica, Taipei, Taiwan, cjlu@iis.sinica.edu.tw, by@crypto.tw

Abstract. We study security amplification for commitment schemes and improve the efficiency of black-box security amplification in the computational setting, where the security holds against PPT active adversaries. We show that $\omega(\log s)$ black-box calls to a weak bit-commitment scheme with constant security is sufficient to construct a commitment scheme with standard negligible security, where s denotes the security parameter and $\omega(\log s)$ denotes any super-logarithmic function of s. Furthermore, the resulting scheme is a string commitment scheme that can commit to $O(\log s)$-bit strings. This improves on previous work of Damgård et al. [DKS99] and Halevi and Rabin [HR08], whose transformations require $\omega(\log^2 s)$ black-box calls to commit a single bit.
As a byproduct of our analysis, we also improve the efficiency of security amplification for message authentication codes, digital signatures, and pseudorandom functions studied in [DIJK09]. This is from an improvement of the "Chernoff-type Theorems" of dynamic weakly-verifiable puzzles of [DIJK09].

1 Introduction

1.1 Security Amplification for Commitment Schemes

Security amplification for weak cryptographic primitives is a basic question that has been studied since the seminal work of Yao [Yao82]. This question has been extensively studied in recent years for a variety of primitives in various settings. To name a few, amplification has been studied for encryption schemes [DNR04,HR05], commitment schemes [DKS99,Wul07,HR08], oblivious transfer [DKS99,Wul07], message authentication codes (MACs), digital signatures, and pseudorandom functions (PRFs) [DIJK09]. Some

⋆ Supported by US-Israel BSF grant 2006060 and NSF grant CNS-0831289.
⋆⋆ Supported by NSF grant CNS-0347661 and CNS-0831293

of these works consider information-theoretic security (e.g., [DKS99]), and others consider computational security. The various security properties of primitives present different interactive settings, for example, commitment schemes are more interactive than encryption schemes, and the chosen-message-attack for MACs introduces a different type of interaction. Proving amplification results tend to be more challenging in an interactive and computational setting.

In this paper, we continue the study of security amplification for commitment schemes, which was previously studied in [DKS99,Wul07,HR08]. We focus on black-box security amplification in the computational setting, where the security holds against probabilistic polynomial time (PPT) active adversaries. Namely, the starting point is a (weak) bit-commitment scheme Com_0 that is *p-hiding* in the sense that no PPT adversarial receiver, who may deviate from the prescribed protocol arbitrarily, can guess the committed bit correctly with probability better than $(1+p)/2$, and *q-binding* in the sense that no PPT adversarial sender can open in two ways with probability better than q, and the goal is to transform Com_0 to a *secure* commitment scheme Com that makes black-box calls to Com_0 and achieves negligible security for both properties.

Previous works focus on feasibility results. Namely, for what values of p and q is the security amplification achievable. In the information-theoretic setting (i.e., the security holds for unbounded adversaries), Damgård, Kilian and Salvail [DKS99] showed that a black-box transformation is possible if and only if $p+q \leq 1-1/\mathrm{poly}(s)$, where s is the security parameter. Halevi and Rabin [HR08] analyzed the transformation of [DKS99] in the computational setting and proved that a black-box transformation is possible whenever $p+q \leq 1-1/\mathrm{polylog}(s)$. Recently and independent of our work, Holenstein and Schoenebeck [HS10] improved the result to optimal. They showed that in the computational setting, black-box security amplification is achievable if and only if $p+q \leq 1-1/\mathrm{poly}(s)$.

However, the existing transformations are not very efficient. To measure the efficiency, let us consider the number of black-box calls to Com_0 that Com makes when p and q are constants with $p+q<1$. Note that the number of black-box calls affects not only the communication complexity, but also the round complexity of the resulting protocol, because in the computational setting, each black-box call needs to be done sequentially.[4]

[4] In general, the commit stage of can consist of multiple rounds. If the black-box calls are done in parallel, one can show by modifying the counter example of Bellare, Impagliazzo, and Naor [BIN97] for interactive arguments that the security may not be amplified at all.

All existing solutions requires $\omega(\log^2 s)$ black-box calls to securely commit a single bit. At a high level, the reason is that they amplify the hiding and binding property *separately*. Amplifying each property from constant to negligible seems to require $\omega(\log s)$ black-box calls, which is the case of the existing constructions and results in $\omega(\log^2 s)$ black-box calls in total. On the other hand, the existing constructions give bit commitment schemes, but there are applications that require *string* commitment schemes. Since it requires $\omega(\log s)$ black-box calls to amplify the security anyway, perhaps we can obtain a string commitment scheme instead of just committing to a single bit, which also improves efficiency in terms of the *rate*, i.e., the number of black-box calls per committed bit. These motivate us to ask the following question.

> **Main question:** How many black-box calls does it require to amplify a (weak) bit commitment scheme with constant security to one with negligible security? What is the length of the string that the resulting Com can commit to, and what is the achievable rate?

Our Results. We give a transformation that amplify a (weak) bit commitment scheme with constant security to a $O(\log s)$-bit string commitment scheme with negligible security using only $\omega(\log s)$ black-box calls, where $O(\log s)$ (resp., $\omega(\log s)$) denotes any $O(\log s)$ (resp., $\omega(\log s)$) function of the security parameter s. In terms of rate, we achieve $\omega(1)$ black-box calls per committed bit. A summary of our result and existing results can be found in Figure 1.

	Efficiency (constants p, q)			Feasibility
Work	Number of black-box calls	Length of committed string	Rate	Applicable range of parameters
[HR08]	$\omega(\log^2 s)$	1	$\omega(\log^2 s)$	$p + q < 1 - 1/\mathrm{poly}\log(s)$
[HS10]	$\omega(\log^2 s)$	1	$\omega(\log^2 s)$	$p + q < 1 - 1/\mathrm{poly}(s)$
Ours	$\omega(\log s)$	$O(\log s)$	$\omega(1)$	$p + q < 1 - 1/\mathrm{poly}\log(s)$
Ours + [HS10]	$\omega(\log s)$	$O(\log s)$	$\omega(1)$	$p + q < 1 - 1/\mathrm{poly}(s)$

Fig. 1. Summary of results on security amplification for commitment schemes in the computational setting. Efficiency measures the cost of amplifying commitment schemes from constant security to negligible security. Feasibility refers to the parameter range that security amplification is possible.

To bypass the $\omega(\log^2 s)$ barrier of the previous transformations, we use error-correcting codes and randomness extractors to amplify both

hiding and binding properties *simultaneously*. To analyze our construction, we model the security of commitment schemes as (the hardness of) solving "two-phase" (interactive) puzzle systems, and study the hardness of solving at least r out of n puzzles. Our result on puzzle systems also applies to the dynamic weakly-verifiable puzzle systems of [DIJK09], and hence improves the efficiency of security amplification for MACs, digital signatures, and PRFs.

Due to the space limit, we focus on presenting our results on security amplification of commitment schemes. We discuss our results of puzzle systems in the following section, and defer the detailed definitions and proofs to the full version of this paper.

1.2 Puzzle Systems and Security Amplification for Other Primitives

Informally, in a puzzle system, there is a puzzle generator generates a puzzle and there is a solver trying to solve the puzzle. At a high level, puzzle systems provide a nice way to abstract the security property of cryptographic protocols – the hardness of solving a puzzle models the hardness for an adversary to break the security. Previously, Canetti, Halevi, and Steiner [CHS05] define *weakly-verifierable puzzle systems* to capture the CAPTCHA scenario, and Dodis, Impagliazzo, Jaiswal, and Kabanets [DIJK09] generalized the model to *dynamic weakly-verifiable puzzle systems* to capture the security of MACs, digital signatures, and PRFs. In this paper, we introduce *two-phase puzzle systems*, which also generalize the model of [CHS05], to capture both hiding and binding properties of commitment schemes.

One natural way to achieve hardness/security amplification is via repetition. Suppose solving a puzzle is δ-hard in the sense that no efficient solver S can successfully solve a puzzle with probability higher than δ. If successfully solving different puzzles were independent events, then successfully solving n puzzles should be δ^n-hard. However, since a solver can correlate his answers to different puzzles, the events are not independent and the hardness bound may not hold. In the literature, there are various *(parallel) repetition theorems* for aforementioned puzzle systems saying that the hardness bounds match that of independent events and/or that the hardness is amplified in an exponential rate, which are useful to deduce security amplification results [CHS05,IJK07,DIJK09,Jut10]. In general, hardness amplification results for one puzzle systems do not imply the same results for another puzzle systems. Furthermore, for general interactive protocols, which can be viewed as "interactive puzzle sys-

tems," there are counter examples (under reasonable assumptions) showing that the hardness may not be amplified at all under *parallel* repetition [BIN97,PW07].

Previous Results. For weakly-verifiable puzzle systems, Canetti, Halevi, and Steiner [CHS05] prove a tight *Direct Product Theorem*, saying that solving n puzzles is δ^n-hard[5] if solving a single puzzle is δ-hard, and Impagliazzo, Jaiswal, and Kabanets [IJK07] prove a more general *Chernoff-type Theorem*, saying that solving at least $(1.1) \cdot \delta \cdot n$ out of n puzzles is $2^{-\Omega(\delta \cdot n)}$-hard if solving a single puzzle is δ-hard. The bound of [IJK07] was recently improved by Jutla [Jut10] to nearly optimal. Dodis, Impagliazzo, Jaiswal, and Kabanets [DIJK09] extend the Chernoff-type Theorem to dynamic weakly-verifiable puzzle systems, and use it to achieve security amplification for MACs, digital signatures, and PRFs. However, the proof techniques of [IJK07,DIJK09,Jut10] seem not applicable to the two-phase puzzle systems.

To analyze their transformations for security amplification for commitment schemes, Halevi and Rabin [HR08] prove a Hardness Degradation Theorem for two-phase puzzle systems (without formally defining the model), saying that solving at least one out of n puzzles is $(1-(1-\delta)^n)$-hard if solving a single puzzle is δ-hard (matching the bound for independent events).

Our Results. We show that the three types of hardness results (Direct Product, Chernoff-type, Hardness Degradation) actually hold for the three aforementioned puzzle systems (weakly-verifiable puzzles, dynamic weakly-verifiable puzzles, and two-phase puzzles.) We establish a *Full-Spectrum Amplification Theorem*, which essentially says that the hardness of solving at least r puzzles out of n puzzles matches the bound of independent events if solving a single puzzle is δ-hard for some constant δ. Note that such a bound is optimal, since a solver can always solve each puzzle independently. A summary of our results and previous results can be found in Figure 2.

We prove the Full-Spectrum Amplification Theorem by a single reduction algorithm that is applicable to all three puzzle systems. The reduction algorithm can be viewed as a generalization of the reduction algorithm of Canetti, Halevi, and Steiner [CHS05].

As a consequence, our improvement on the Chernoff-type Theorem for dynamic weakly verifiable puzzle systems of Dodis et al. [DIJK09] implies

[5] We omit the negligible slackness in this informal discussion.

	Weakly Verifiable	Dynamic Weakly Verifiable	Two-Phase
Direct Product	[CHS05]	[DIJK09], Ours	[HR08]
Chernoff-type	[IJK07,Jut10], Ours	[DIJK09], Ours	Ours, [HS10]
Hardness Degradation	[HR08]	Ours*	[HR08]
Full-Spectrum	Ours, [HS10]	Ours*	Ours, [HS10]

Fig. 2. Summary of results on different types of puzzle systems. "Ours" means that either we obtain new results or we improve bounds over previous ones. The work of [HS10] and our work are independent works. (*): Our hardness degradation and full-spectrum results only hold for a variant of the dynamic weakly verifiable puzzle systems (see the full version of this paper for details).

improvement on the efficiency of security amplification for MACs, digital signatures, and PRFs.

Historical Notes. The work of Holenstein and Schoenebeck [HS10] and our work were done independently, but have significant overlap. We briefly compare the results and make some historical notes as follows. For security amplification for commitment schemes, both works improve the result of Halevi and Rabin [HR08], but in complementary ways. Holenstein and Schoenebeck shows a feasibility result saying that security amplification is possible if and only if $p+q \leq 1-1/\mathrm{poly}(s)$. We improve the efficiency of the transformation, saying that only $\omega(\log s)$ black-box calls is sufficient to amplify security from constant to negligible and the resulting commitment scheme can commit to a $O(\log s)$-bit string. The constructions in both work are very different. As shown in the figure 1, the two results can be combined to obtain both improvements simultaneously.

For puzzle systems, Holenstein and Schoenebeck [HS10] present essentially the same idea and reduction algorithm as in our work. However, they have a cleaner way to deal with the parameters, and hence their result holds for every δ as opposed to constant δ in our result. Also, they consider more general "monotone combining functions" in addition to the threshold functions considered in our work. On the other hand, the application to efficiency improvement of security amplification for MACs, digital signatures, and PRFs was pointed out by us.

2 Preliminaries

All log's are base 2. s is the security parameter, and $\mathsf{ngl} = \mathsf{ngl}(s)$ denotes a negligible function of the security parameter, i.e. $s^{-\log(s)}$. We use U_n to denote uniform distribution over n-bit strings. We identify $\{0,1\}$ with $\mathbb{F}_2$, the finite field of size 2. If $x, y \in \{0,1\}^n$ are vectors in $\mathbb{F}_2^n$, then $x \oplus y \in$

$\{0,1\}^n$ denotes their sum, (i.e. bit-wise xor) and $x \cdot y \stackrel{\text{def}}{=} \sum_i x_i y_i \in \{0,1\}$ denotes their inner product.

We review the facts we need about error-correcting codes. The lemma below says that a short random linear code has good minimum distance with overwhelming probability. It can be proved by standard probabilistic methods, and we omit the proof. The constants in the lemma are actually small.

Definition 1. *The* Hamming distance *of two strings x and y is the number of coordinates i such that $x_i \neq y_i$. Let $C : \{0,1\}^n \to \{0,1\}^{n'}$ be a code. The* minimum distance *of C is the minimum Hamming distance over all parts of codewords $C(x)$ and $C(y)$ such that $x \neq y$.*

Lemma 1. *There exist universal constants d_0, d_1 such that the following holds. Let k be a positive integer, and $\gamma, \delta \in [0,1]$ be numbers such that $\gamma > d_0 \cdot \delta \log(1/\delta)$. Let n be an integer such that $n > d_1 \cdot k/\delta$. Let $C : \{0,1\}^n \to \{0,1\}^{(1+\gamma)n}$ be a random linear code defined by $C(m) = (m, Am)$, where $A \in \{0,1\}^{\gamma n \times n}$ is a random 0-1 matrix. Then C has minimum distance at least $\delta \cdot n$ with probability at least $1 - 2^{-k}/2$.*

3 Definitions and Main Theorems

3.1 Commitment Schemes

In this section, we formally define commitment schemes and present our main theorem. We consider a standard model where the communication is over the classical noiseless channel and the decommitment is non-interactive [Gol01,HR08].

Definition 2 (Commitment Scheme). *A* **commitment scheme** *is an interactive protocol* $\mathrm{Com} = (S, R)$ *with the following properties:*

1. *Scheme* Com *consists of two stages: a* **commit stage** *and a* **reveal stage**. *In both stages, the* **sender** *S and the* **receiver** *R receive a security parameter 1^s as common input.*
2. *At the beginning of the commit stage, sender S receives a private input $v \in \{0,1\}^t$, which denotes the string to which S is supposed to commit. The commitment stage results in a joint output, which we call the* **commitment** *$x = \mathrm{output}((S(v), R)(1^s))$, and a private output for S, which we call the* **decommitment** *string $d = \mathrm{output}_S((S(v), R)(1^s))$. Without loss of generality, x can be taken to be the full transcript of the interaction between S and R, and d to be the private coin tosses of S.*

3. *In the reveal stage, sender S sends the pair (v, d), where d is the decommitment string for string v. Receiver R accepts or rejects based on v, d, x.*
4. *Both sender S and receiver R are efficient, i.e., both run in probabilistic polynomial time in the security parameter s.*
5. *R will always accept with probability $1-\mathsf{ngl}$ if both the sender S and the receiver R follow their prescribed strategy. If R accepts with probability 1, we say* Com *has* **perfect correctness**.
6. *When the commit string v is just a bit in $\{0,1\}$, we call* Com *a* **bit-commitment scheme**. *Otherwise, we call* Com *a t-***bit string-commitment scheme**.

Remark 1. The assumption of non-interactive reveal stage is essential in both our work and the previous work [HR08]. This assumption can be made without loss of generality as long as no additional property (e.g., if the sender wants to decommit in a zero-knowledge manner) is required, because in the reveal stage, the sender S can send his coin tosses to the receiver R, who can check the consistency and simulate the protocol. On the other hand, the assumption of perfect correctness can be relaxed to $(1 - \mathsf{ngl})$-correctness in both works.

We proceed to define the hiding and binding properties of commitment schemes. To facilitate the presentation of our results and analysis, we are precise about the adversary's running time in the definition and define the binding property in terms of binding games.

Definition 3 (p-hiding against time T). *A commitment scheme* $\mathrm{Com} = (S, R)$ *is* p-hiding against uniform time T *if for every probabilistic time T cheating receiver R^*, the distributions* $(\mathrm{view}_{R^*}(S(U_t), R^*), U_t)$ *and* $(\mathrm{view}_{R^*}(S(U_t), R^*), U'_t)$ *are p-indistinguishable for time T, where U'_t is an i.i.d. copy of U_t. That is, for every probabilistic time T distinguisher D,*

$$|\Pr[D(\mathrm{view}_{R^*}(S(U_t), R^*), U_t) = 1] - \Pr[D(\mathrm{view}_{R^*}(S(U_t), R^*), U'_t) = 1]| \leq p/2$$

We say Com *is* p**-hiding** *if* $\mathrm{Com}(1^s)$ *is p-hiding against time s^c for every constant c, and sufficiently large security parameter s.*

We remark that the hiding property above is defined as the indistinguishability for *random values*, which does not generally imply the standard semantic security for the hiding property. Nevertheless, it is easy to transform a commitment scheme Com with the above hiding property to one with standard semantic security – one can use Com to commit to a random string $r \in_R \{0,1\}^t$, and use r as a one-time pad to hide the actual string v that the sender wants to commit to.

Remark 2. For bit-commitment schemes, p-hiding is equivalent to saying that the receiver can guess the committed bit with probability at most $1/2 + p/2$. Formally, for every time T predictor P,

$$\Pr[P(\text{view}_{R^*}(S(U_1), R^*)) = U_1)] \leq 1/2 + p/2.$$

Definition 4 (Binding Game). *The* **binding game** *for a commitment scheme* $\text{Com} = (S, R)$ *is played between a honest receiver* R*, and* (S^*, F)*, a cheating sender* S^* *with a decommitment finder* F*. The game consists of two stages:*

1. *In the commit stage,* S^* *interacts with* R *to produce a view* $\text{view}_{S^*}(S^*, R)$*.*
2. *In the decommitment finding stage,* F *gets the view* $\text{view}_{S^*}(S^*, R)$*, and produces two decommitment strings* (s, d) *and* (s', d')*.*

(S^*, F) **succeeds** *if in the reveal stage,* R *accepts both decommitment strings* (s, d) *and* (s', d')*.*

Definition 5 (q-binding against time T). *A commitment scheme* $\text{Com} = (S, R)$ *is* q**-binding against time** T*, if in the binding game, for every time* T *pair* (S^*, F)*, the success probability of* (S^*, F) *is at most* q*. We say* Com *is* q**-binding** *if* $\text{Com}(1^s)$ *is* q*-binding against time* s^c *for every constant* c*, and sufficiently large security parameter* s*.*

Definition 6 (security of commitment schemes). *A commitment scheme* Com *is* (p, q)**-secure (against time** T**)** *if* Com *is* p*-hiding and* q*-binding (against time* T*).* Com *is* **secure** *if* $\text{Com}(1^s)$ *is* (s^{-c}, s^{-c})*-secure for every constant* c*, and sufficiently large security parameter* s*.*

We proceed to state our main result on efficient security amplification for commitment schemes. The following theorem says that we can securely commit a $O(\log s)$-bit string using only $\omega(\log s)$ black-box call to a weak commitment scheme Com_0 with constant hiding and binding properties.

Theorem 1. *Let* $p, q \in (0, 1)$ *be constants with* $p + q < 1$*. Suppose there exists a* (p, q)*-secure bit commitment scheme* Com_0*. Then for every* $t(s) = O(\log s)$*,* $n(s) = \omega(t + \log s)$ *where* s *is the security parameter, there exists a secure* t*-bit string-commitment scheme* Com *that makes only* n *black-box call to* Com_0*.*

4 Efficient Security Amplification for Commitment Schemes

In this section, we present our result on efficient black-box security amplification for commitment schemes in the computational setting, where

the security holds against PPT active adversaries. We start by reviewing the previous construction of Halevi and Rabin [HR08], and then discuss its limitation and our improvement. The construction in [HR08] uses the following two transformations, each of which improves one property significantly without hurting the other property too much.

- **Secret-sharing transformation.** Let Com_0 be a bit commitment scheme, and $n \in \mathbb{N}$ be a parameter. The transformation gives a bit commitment scheme $\mathrm{Com} = (S, R)$ as follows. To commit a bit $b \in \{0,1\}$ to R, S generates random $b_1, b_2, \ldots b_n \in \{0,1\}$ such that $\bigoplus_{i\in[n]} b_i = b$, i.e. a secret sharing of b, and then uses Com_0 to sequentially commit to each b_i to R.
  Intuitively, this transformation improves the hiding property, since an adversarial R^* needs to learn all bits $b_1, \ldots, b_n$ to recover b, but hurts the binding property, since an adversarial S^* only needs to cheat on any single bit b_i to decommit in two ways. Indeed, Halevi and Rabin proved that if Com_0 is (p, q)-secure, then Com is $(p^n, 1-(1-q)^n)$-secure.[6]
- **Repetition transformation.** Let Com_0 be a bit commitment scheme, and $n \in \mathbb{N}$ be a parameter. The transformation gives a bit commitment scheme $\mathrm{Com} = (S, R)$ as follows. To commit a bit $b \in \{0,1\}$ to R, S sequentially uses Com_0 n times to commit to the same bit b to R.
  Intuitively, this transformation improves the binding property, since an adversarial S^* needs to cheat on all copies of Com_0 to decommit in two ways, but hurts the hiding property, since an adversarial R^* can learn the bit b from any copy of the commitments. Indeed, Halevi and Rabin proved that if Com_0 is (p, q)-secure, then Com is $(1-(1-p)^n, q^n)$-secure.

Halevi and Rabin showed that, as long as p and q satisfy $p + q \leq 1 - 1/\mathrm{polylog}(s)$, then given a (p, q)-secure (weak) bit commitment scheme Com_0, one can apply the above two transformations alternately to obtain a secure bit commitment scheme Com. To measure the efficiency, consider the case where both p and q are constants with $p + q < 1$. Since improving either hiding or binding property from constant to negligible requires $\omega(\log s)$ invocations to Com_0, and the above transformations improve two properties *separately*, the construction of Halevi and Rabin requires at least $\omega(\log^2 s)$ black-box calls to Com_0.

[6] We omit the negligible slackness in the informal discussion.

Remark 3. Independent of our work, Holenstein and Schoenebeck [HS10] present a different construction that improves the result of Halevi and Rabin in the following sense.For any (p, q)-secure bit commitment scheme Com_0 with $p + q \leq 1 - 1/\mathrm{poly}(s)$, their construction gives a secure bit commitment scheme Com using $\mathrm{poly}(s)$ black-box calls to Com_0. Their construction uses Valiant's monotone formula for majority [Val84] to improve both properties. However, a closer inspection shows that their construction is equivalent to applying the secret sharing transformation and a variant of repetition transformation (with the same effect on the parameters) alternately. Hence, in terms of the efficiency, their construction also requires at least $\omega(\log^2 s)$ black-box calls to amplify a (p, q)-secure weak commitment scheme with constant p and q to a secure one.

To bypass the $\omega(\log^2 s)$ barrier of the existing constructions, our main idea is to use error-correcting codes and randomness extractors to amplify both hiding and binding properties *simultaneously*. For intuition, we give an informal description of our transformation first. Let us informally use $\mathrm{Com}_0(b)$ to denote a commitment of a bit b, and let $C : \{0,1\}^n \to \{0,1\}^{n'}$ be an error-correcting code with minimum distance at least $\delta \cdot n'$, and $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^t$ a strong randomness extractor. Our transformation uses Com_0, C and Ext to commit to a string $v \in \{0,1\}^t$ as follows (recall that we obtain string commitment schemes as opposed to bit commitment schemes of other existing constructions).

- **Commit Stage**: the sender S samples a message $m \in_R \{0,1\}^n$ uniformly at random, and sequentially commits to each bit of the codeword $C(m)$ using Com_0, which generates commitments $\mathrm{Com}_0(C(m)) \stackrel{\mathrm{def}}{=} (\mathrm{Com}_0(C(m)_1), \ldots, \mathrm{Com}_0(C(m)_{n'}))$. Then S samples a uniform seed $z \in_R \{0,1\}^d$, and sends the seed z with $v \oplus \mathrm{Ext}(m, z)$ to the receiver R. In sum, the commitment is $\mathrm{Com}(v) = (\mathrm{Com}_0(C(m)), z, v \oplus \mathrm{Ext}(m, z))$.
- **Reveal Stage**: the sender S sends the value v, the message m and reveals each committed bit of $C(m)$ to R, who checks consistency and accepts or rejects accordingly.

Intuitively, the binding property is improved because for an adversarial sender S^* to cheat, S^* needs to decommit $C(m)$ into two valid codewords. Since the code C has good minimum distance, S^* needs to successfully cheat on at least $\delta \cdot n'$ committed bits out of n' commit bits. The q-binding property of Com_0 says that, for each committed bit, S^* can cheat with probability at most q. Thus, in expectation, S^* can cheat on only $q \cdot n'$ commit bits. If $q < (0.9)\delta$, the Chernoff bound suggests that S^*

should be able to cheat on at least $\delta \cdot n'$ commit bits with only exponentially small probability in n'. On the other hand, the hiding property is improved because after seeing the commitments of $C(m)$, an adversarial receiver R^* has only partial information about m by the p-hiding property of Com_0. Thus, Ext extracts the remaining (computational) entropy from m, which is used to hide the value v. Ideally, when both p and q are constants, we can set both $n, n' = \omega(\log s)$ and commit to $\Omega(n)$-bit string.

In sum, our efficient security amplification for commitment schemes consists of three steps: given a (p, q)-secure bit commitment scheme Com_0 with constants $p + q < 1$, (1) we first apply the transformations of Halevi and Rabin to obtain a (p', q')-secure bit commitment scheme Com_1 with sufficiently small constants p', q', which costs a constant number of black box calls, (2) we apply the above construction to obtain a (s^{-c}, s^{-c})-secure $O(\log s)$-bit string commitment scheme Com_2, which costs $O(\log s)$ black box calls, and (3) we apply a string version of the transformations of Halevi and Rabin [HR08] to obtain a secure $O(\log s)$-bit string commitment scheme Com_3, which costs $\omega(1)$ black box calls. The number of black-box calls multiply over steps, and hence the resulting Com_3 uses $\omega(\log s)$ black-box calls to Com_0.

We proceed to give a formal description of the above construction and its analysis in Section 4.1, and present a string version of the transformations of Halevi and Rabin used in the third step in Section 4.2.

4.1 Efficient Security Amplification in the Known-Security Setting

In this section, we present a transformation that converts a (p, q)-secure bit commitment scheme Com_0 to a (s^{-c}, s^{-c})-secure $O(\log s)$-bit string commitment scheme Com using $O(\log s)$ black-box calls to Com_0, where c is an arbitrary constant. Our transformation uses error-correcting codes and randomness extractors to amplify both hiding and binding properties *simultaneously*. The transformation requires to use a *systematic* code with good distance and the "Goldreich-Levin" extractor. We will discuss the reason when we prove the security below. A formal description of our transformation can be found in Figure 3.

We will show that if Com_0 is a (p, q)-secure bit commitment scheme for small constants p, q, then by setting $n, \ell, t = O(\log s)$, the resulting string commitment scheme is (s^{-c}, s^{-c})-secure for some constant c. Note that both parties in Com run in time polynomial in n, ℓ, t, and the running time of Com_0, which is efficient. Formally, we prove the following theorem.

Transformation $\mathcal{T}(\mathrm{Com}_0, n, \ell, t)$:

- **Inputs.** A bit commitment scheme Com_0, and parameters $n, \ell, t \in \mathbb{N}$.
- **Outputs.** A t-bit string-commitment scheme $\mathrm{Com} = (S, R)$ defined as follows.
- **Commit Stage.** Let $v \in \{0,1\}^t$ be the string to which S is committing to.
  1. R samples a uniformly random matrix $A \leftarrow \{0,1\}^{\ell \times n}$, and sends A to S.
     /* i.e., R selects a random systematic linear code $C(m) \stackrel{\text{def}}{=} (m, Am)$. */
  2. S samples the following uniformly at random: a message $m \leftarrow \{0,1\}^n$ and a matrix $Z \leftarrow \{0,1\}^{t \times n}$.
     /* Z is a random seed for a (strong) randomness extractor $\mathrm{Ext}(m, Z) \stackrel{\text{def}}{=} Zm$.*/
  3. S uses Com_0 to commit to each bit of m and each bit of Am to R sequentially. Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_\ell)$ denote the commitment of each bit respectively.
     /* i.e., S commits to each bit of the codeword $C(m)$. */
  4. S sends $(Z, v \oplus Zm)$ to R, where $v \oplus Zm$ is the bit-wise xor of v and Zm.
     /* i.e., S uses $\mathrm{Ext}(m, Z)$ as a one-time pad to hide the commit string v. */

  In sum, the commitment of v is $(A, \boldsymbol{x}, \boldsymbol{y}, Z, v \oplus Zm)$.
- **Reveal Stage.** S sends v and its coin tosses r to R, and R checks that v and r are consistent with the honest sender's algorithm.

Fig. 3. Our black-box transformation $\mathcal{T}(\mathrm{Com}_0, n, \ell, t)$.

Theorem 2. *The following holds for all sufficiently small constants $p, q \in (0,1)$, and $k = O(\log s)$: Suppose there exists a (p,q)-secure (weak) bit commitment scheme* Com_0*, then there exists a $(2^{-k}, 2^{-k})$-secure $t = \Omega(k)$-bit string-commitment scheme* Com *that makes $O(k)$ black-box calls to* Com_0*. Specifically,* $\mathrm{Com} = \mathcal{T}(\mathrm{Com}_0, n, \ell, t)$ *for appropriate $n, \ell = O(k)$, and $t = \Omega(k)$.*

We formalize the aforementioned intuition to analyze the hiding and binding properties in the below subsections.

Analysis of the Binding Property In this section, we analyze the binding property of our transformation $\mathcal{T}(\mathrm{Com}_0, n, \ell, t)$. We first recall the intuition of why the binding property is improved. Recall that in the construction, the sender S is supposed to commit to each bit of a valid codeword $C(m) = (m, Am)$ using Com_0, where C is a random linear code chosen by the receiver R. By Lemma 1, C has good min-distance $\delta \cdot n$ with overwhelming probability. For an adversarial sender S^* to cheat, S^* needs to decommit the $n + \ell$ commitments into two valid codewords $C(m_1), C(m_2)$, which means that S^* needs to successfully cheat on at least $\delta \cdot n$ commitments out of $n + \ell$ commitments. Intuitively, suppose breaking

the binding property of each commitment were independent events with success probability at most q, and if $\delta \cdot n \geq (1.1) \cdot q \cdot (n + \ell)$, then by Chernoff bounds, the success probability of S^* should be exponentially small in n.

Of course, the events are not independent since S^* has chance to correlate his strategy for different instances. However, breaking the binding property of sequentially committed bits can be modeled as repetition of two-phase puzzle systems, and hence the above intuition can be formalized using the Full-Spectrum Amplification Theorem (appeared in the full version of this paper), which says the success probability of S^* behaves the same as the case of independent events.

Formally, we prove the following lemma, which essentially says that when q is sufficiently smaller than the min-distance of the code, the binding property is amplified in an exponential rate. We formulate the lemma in concrete parameters for preciseness. For intuition, think of $n, \ell = \Theta(k)$, $k = O(\log s)$, $T_0 = \mathrm{poly}(s)$, and $T = s^{\omega(1)}$.

Lemma 2 (Binding). *Let d_0 be the universal constant in Lemma 1. There exist universal constants c_1 such that the following holds. For any $q \in (0,1), n, k, \ell, t, T_0, T \in \mathbb{N}$ satisfying (i) $d_0 \cdot (3q) \cdot \log(1/3q) < 1$, (ii) $2c_1 \cdot k/q \geq n \geq c_1 \cdot k/q$, (iii) $n > \ell \geq d_0 \cdot (3q) \cdot \log(1/3q) \cdot n$, if a bit-commitment scheme $\mathrm{Com}_0 = (S_0, R_0)$ with runtime T_0 is q-binding against time T, then $\mathrm{Com} = \mathcal{T}(\mathrm{Com}_0, n, \ell, t)$ is 2^{-k}-binding against time $T' = T/\mathrm{poly}(2^k, T_0, t)$.*

Analysis of the Hiding Property In this section, we analyze the hiding property of our transformation $\mathcal{T}(\mathrm{Com}_0, n, \ell, t)$. We first recall the intuitive entropy argument of why the hiding property is improved. Recall that in the construction, the sender S samples a random n-bit message m, which contains n bits of entropy. Then S commits to each bit of the codeword $C(m) = (m, Am)$, each of which leaks information about m. Intuitively, if we set the parameters so that there are entropy left in m, S can use randomness extractor to extract a string $\mathrm{Ext}(x, Z)$ that is (pseudo-)random from an adversarial receiver R^*'s point of view, and use it as one-time-pad to hide the commit value v.

We argue that it is very hard for R^* to predict the whole message m after he sees the $n + \ell$ commitments, and hence one can apply the Goldreich-Levin theorem to extract pseudo-random bits. This is why our transformation requires to use the Goldreich-Levin extractor. To argue that m is hard to predict from the commitments $(\boldsymbol{x}, \boldsymbol{y})$, we first argue

that m is hard to predict from $\boldsymbol{x}$. We can view predicting n sequentially committed message bits of m from the commitments $\boldsymbol{x}$ as n-fold direct product of a two-phase puzzle system. By Direct Product Theorem of Halevi and Rabin [HR08], the success probability of R^* is at most $((1+p)/2)^n$ (up to a negligible factor). Observing that $\boldsymbol{y}$ contains at most ℓ bits of information about m, the success probability of R^* to predict m from $(\boldsymbol{x}, \boldsymbol{y})$ is at most $2^\ell \cdot ((1+p)/2)^n$. Hence, by the Goldreich-Levin theorem, we can extract $\Omega(\log(2^\ell \cdot ((1+p)/2)^n))$ pseudorandom bits.

Formally, we prove the following lemma, which essentially says that we can extract $\Omega(\log(2^\ell \cdot ((1+p)/2)^n))$ pseudorandom bits. Again, we formulate the lemma in concrete parameters for preciseness, and we use parameter $\alpha = 1 - p$ for clarity. For intuition, think of $n, \ell = \Theta(k)$, $k = O(\log s)$, $T_0 = \mathrm{poly}(s)$, and $T = s^{\omega(1)}$.

Lemma 3 (Hiding). *There exist universal constants c_2 such that the following holds. For every $\alpha \in (0,1), n, k, \ell, t, T_0, T \in \mathbb{N}$ satisfying (i) $2c_2 \cdot k/\alpha \geq n \geq c_2 \cdot k/\alpha$, (ii) $\ell, t \leq \alpha n/12$, if $\mathrm{Com}_0 = (S_0, R_0)$ with runtime T_0 is a $(1-\alpha)$-hiding against time T, then $\mathrm{Com} = \mathcal{T}(\mathrm{Com}_0, n, \ell, t)$ is 2^{-k}-hiding against time $T' = T/\mathrm{poly}(2^k, T_0)$.*

We leave the proofs of Lemma 2 and 3 in the full version of this paper.

Proof of Theorem 2 Theorem 2 follows by applying Lemma 2 and 3 with properly chosen parameters.

Proof. (**of Theorem 2**) We set the parameters n, k, ℓ as follows: $n = \max\{\frac{c_1 k}{q}, \frac{c_2 k}{1-p}\} = \Theta(k)$, $\ell = d_0(3q)\log(3q) \cdot n$, and $t = \frac{(1-p)n}{12} = \Omega(k)$, where c_1, c_2, d_0 are the constants in the Lemma 1, 2, and 3. The theorem follows directly from Lemma 2 and 3.

4.2 Security Amplification for String Commitment Schemes

In this section, we generalize the transformations of Halevi and Rabin [HR08] to the case of string commitment schemes, with the goal of amplifying the (s^{-c}, s^{-c})-secure string commitment scheme obtained from our transformation to achieve negligible security. This is a simpler task, and can be done by applying a secret-sharing transformation first and then a repetition transformation. A formal description of the transformations can be found in Figure 4.

We proceed to analyze the binding and hiding properties of the resulting commitment schemes of the two transformations. For the binding

Secret-sharing $\mathcal{SS}(\mathrm{Com}_0, u)$. Let Com_0 be a t-bit string commitment scheme, and $u \in \mathbb{N}$ be a parameter. The transformation gives a t-bit string commitment scheme $\mathrm{Com} = (S, R)$ as follows. To commit a value $v \in \{0,1\}^t$ to R, S generates random $v_1, v_2, \ldots v_n \in \{0,1\}^t$ such that $v_1 \oplus v_2 \oplus \cdots \oplus v_u = v$, where $\oplus$ denotes the bit-wise xor of v_i's (i.e. a secret sharing of v), and then uses Com_0 to sequentially commit to each v_i to R.

Repetition $\mathcal{R}(\mathrm{Com}_0, u)$. Let Com_0 be a t-bit string commitment scheme, and $u \in \mathbb{N}$ be a parameter. The transformation gives a t-bit string commitment scheme $\mathrm{Com} = (S, R)$ as follows. To commit a value $v \in \{0,1\}^t$ to R, S sequentially uses Com_0 u times to commit to the same value v to R.

Fig. 4. Secret-sharing and repetition transformation for string commitment schemes.

property, the analysis is essentially the same as in [HR08]: for repetition, it requires to break all u commitments of Com_0, and for secret-sharing, it requires to break only 1 out of u commitments of Com_0, which can be modeled as solving corresponding repetition of two-phase puzzles. The Direct Product Theorem and Hardness Degradation Theorem of Halevi and Rabin [HR08] (or our Full-Spectrum Amplification Theorem) imply the following lemma.

Lemma 4 ([HR08]). *Let* Com_0 *be a* t*-bit string-commitment scheme,* $u = u(s) \leq \mathrm{poly}(s)$ *a efficiently computable function, and* $q \in (0,1)$ *a constant. Suppose* Com_0 *is* q*-binding, then* $\mathcal{R}(\mathrm{Com}_0, u)$ *is* $(q^u + \mathsf{ngl})$*-binding, and* $\mathcal{SS}(\mathrm{Com}_0, u)$ *is* $(1 - (1-q)^u + \mathsf{ngl})$*-binding.*

On the other hand, analyzing the hiding property is tricker. For the secret-sharing transformation, we need a string version of XOR Lemma to show that the hiding property is amplified. Maurer and Tessaro [MT09] proved a more general result (Theorem 2 of [MT09]) in the context of system composition, which implies the following lemma.

Lemma 5 ([MT09]). *Let* Com_0 *be a* t*-bit string-commitment scheme, and* $\mathrm{Com} = \mathcal{SS}(\mathrm{Com}_0, u)$ *with efficiently computable* $u = u(s) \leq \mathrm{poly}(s)$*. If* Com_0 *is* p*-hiding, then* Com *is* $(p^u + \mathsf{ngl})$*-binding.*

We next show that repetition transformation preserves the (negligible) hiding property. This is sufficient for our purpose since we will apply the secret-sharing transformation to amplify the hiding property to negligible before applying the repetition transformation.

Lemma 6. *Let* $\mathrm{Com}_0 = (S_0, R_0)$ *be a* t*-bit string-commitment scheme, and* $\mathrm{Com} = \mathcal{R}(\mathrm{Com}_0, u)$ *with efficiently computable* $u = u(s) \leq \mathrm{poly}(s)$*. If* Com_0 *is* ngl*-hiding, so does* Com*.*

We leave the proof in the full version of this paper.

4.3 Put Things Together

We are ready to present a formal description of our efficient security amplification for commitment schemes (in Figure 5) and prove Theorem 1.

Final Construction.

- **Inputs.** A (p, q)-secure bit commitment scheme Com_0 with $p + q < 1$.
- **Outputs.** A secure t-bit string-commitment scheme Com with $t = O(\log s)$.

1. Apply the transformations of Halevi and Rabin alternately to obtain a (p', q')-secure bit commitment scheme Com_1 with sufficiently small constants p', q'.
2. Apply our transformations $\mathcal{T}(\mathrm{Com}_1, n, \ell, t)$ to obtain a (s^{-c}, s^{-c})-secure t-bit string commitment scheme Com_2, where $n, \ell = O(\log s)$, and c is some constant.
3. Let a be an arbitrary $\omega(1)$ function. Apply $\mathcal{SS}(\mathrm{Com}_2, a)$ to obtain a $(\mathsf{ngl}, a \cdot s^{-c} + \mathsf{ngl})$-secure t-bit string commitment scheme Com_3.
4. Apply $\mathcal{R}(\mathrm{Com}_3, a)$ to obtain a secure t-bit string commitment scheme Com.

Fig. 5. Efficient security amplification of commitment schemes.

Proof (of Theorem 1). The fact that Com is a secure string commitment scheme follows straightforwardly from Theorem 2 and Lemma 4, 5, 6. Observing the Com_1 makes $O(1)$ black-box calls to Com_0, Com_2 makes $O(\log s)$ black-box calls to Com_1, Com_3 makes $\omega(1)$ black-box calls to Com_2, and finally Com makes $\omega(1)$ black-box calls to Com_3, the total number of black-box calls that Com makes to Com_0 is $\omega(\log s)$, as desired.

Acknowledgments

We thank Salil Vadhan for his discussions throughout the whole work of this paper. We also thank Anna Lysyanskaya and Yevgeniy Dodis for sharing their insights.

References

[BIN97] Mihir Bellare, Russell Impagliazzo, and Moni Naor. Does parallel repetition lower the error in computationally sound protocols? In *FOCS*, pages 374–383, 1997.

[CHS05] Ran Canetti, Shai Halevi, and Michael Steiner. Hardness amplification of weakly verifiable puzzles. In *TCC*, pages 17–33, 2005.

[DIJK09] Yevgeniy Dodis, Russell Impagliazzo, Ragesh Jaiswal, and Valentine Kabanets. Security amplification for interactivecryptographic primitives. In *TCC*, pages 128–145, 2009.

[DKS99] Ivan Damgård, Joe Kilian, and Louis Salvail. On the (im)possibility of basing oblivious transfer and bit commitment on weakened security assumptions. In *EUROCRYPT*, pages 56–73, 1999.

[DNR04] Cynthia Dwork, Moni Naor, and Omer Reingold. Immunizing encryption schemes from decryption errors. In *EUROCRYPT*, pages 342–360, 2004.

[Gol01] Oded Goldreich. *Foundations of Cryptography. Basic tools*. Cambridge University Press, 2001.

[HR05] Thomas Holenstein and Renato Renner. One-way secret-key agreement and applications to circuit polarization and immunization of public-key encryption. In *CRYPTO*, pages 478–493, 2005.

[HR08] Shai Halevi and Tal Rabin. Degradation and amplification of computational hardness. In *TCC*, pages 626–643, 2008.

[HS10] Thomas Holenstein and Grant Schoenebeck. General hardness amplification of predicates and puzzles. *CoRR*, abs/1002.3534, 2010.

[IJK07] Russell Impagliazzo, Ragesh Jaiswal, and Valentine Kabanets. Chernoff-type direct product theorems. In *CRYPTO*, pages 500–516, 2007.

[Jut10] Charanjit S. Jutla. Almost optimal bounds for direct product threshold theorem. In Daniele Micciancio, editor, *TCC*, volume 5978 of *Lecture Notes in Computer Science*, pages 37–51. Springer, 2010.

[MT09] Ueli Maurer and Stefano Tessaro. Computational indistinguishability amplification: Tight product theorems for system composition. In Shai Halevi, editor, *Advances in Cryptology — CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 350–368. Springer-Verlag, August 2009.

[PW07] Krzysztof Pietrzak and Douglas Wikström. Parallel repetition of computationally sound protocols revisited. In *TCC*, pages 86–102, 2007.

[Val84] Leslie G. Valiant. Short monotone formulae for the majority function. *J. Algorithms*, 5(3):363–366, 1984.

[Wul07] Jürg Wullschleger. Oblivious-transfer amplification. In *EUROCRYPT*, pages 555–572, 2007.

[Yao82] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *FOCS*, pages 80–91, 1982.