

# Affine-evasive Sets Modulo a Prime

Divesh Aggarwal\*

June 10, 2014

## Abstract

In this work, we describe a simple and efficient construction of a large subset  $S$  of  $\mathbb{F}_p$ , where  $p$  is a prime, such that the set  $A(S)$  for any non-identity affine map  $A$  over  $\mathbb{F}_p$  has small intersection with  $S$ .

Such sets, called affine-evasive sets, were defined and constructed in [ADL14] as the central step in the construction of non-malleable codes against affine tampering over  $\mathbb{F}_p$ , for a prime  $p$ . This was then used to obtain efficient non-malleable codes against split-state tampering.

Our result resolves one of the two main open questions in [ADL14]. It improves the rate of non-malleable codes against affine tampering over  $\mathbb{F}_p$  from  $\log \log p$  to a constant, and consequently the rate for non-malleable codes against split-state tampering for  $n$ -bit messages is improved from  $n^6 \log^7 n$  to  $n^6$ .

---

\*Department of Computer Science, New York University. Email: [divesha@cs.nyu.edu](mailto:divesha@cs.nyu.edu).

# 1 Introduction

**Non-malleable Codes (NMCs).** NMCs were introduced in [DPW10] as a beautiful relaxation of error-correction and error-detection codes. Informally, given a tampering family  $\mathcal{F}$ , an NMC  $(\text{Enc}, \text{Dec})$  against  $\mathcal{F}$  encodes a given message  $m$  into a codeword  $c \leftarrow \text{Enc}(m)$  in a way that, if the adversary modifies  $m$  to  $c' = f(c)$  for some  $f \in \mathcal{F}$ , then the message  $m' = \text{Dec}(c')$  is either the original message  $m$ , or a completely “unrelated value”. As has been shown by the recent progress [DPW10, LL12, DKO13, ADL14, FMVW13, FMNV14, CG14a, CG14b] NMCs aim to handle a much larger class of tampering functions  $\mathcal{F}$  than traditional error-correcting or error-detecting codes, at the expense of potentially allowing the attacker to replace a given message  $x$  by an unrelated message  $x'$ . NMCs are useful in situations where changing  $x$  to an unrelated  $x'$  is not useful for the attacker (for example, when  $x$  is the secret key for a signature scheme.)

**Split-State Model.** NMCs do not exist for the class of all functions  $\mathcal{F}_{\text{all}}$ . In particular, it does not include functions of the form  $f(c) := \text{Enc}(h(\text{Dec}(c)))$ , since  $\text{Dec}(f(\text{Enc}(m))) = h(m)$  is clearly related to  $m$ . One of the largest and practically relevant tampering families for which we can construct NMCs is the so-called split-state tampering family where the codeword is split into two parts  $c_1 \| c_2$ , and the adversary is only allowed to tamper with  $c_1, c_2$  independently to get  $f_1(c_1) \| f_2(c_2)$ . A lot of the aforementioned results [LL12, DKO13, ADL14, CG14b, FMNV14] have studied NMCs against split-state tampering. [ADL14] gave the first (and the only one so far) information-theoretically secure construction in the split-state model from  $n$ -bit messages to  $n^7 \log^7 n$ -bit codewords (i.e., code rate  $n^6 \log^7 n$ ). The security proof of this scheme relied on an amazing property of the inner-product function modulo a prime, that was proved using results from additive combinatorics.

**Affine-evasive Sets and Our Result.** One of the crucial steps in the construction of [ADL14] was the construction of NMC against affine tampering modulo  $p$ . This was achieved by constructing an affine-evasive set of size  $p^{1/\log \log p}$  modulo a prime  $p$ . It was asked as an open question whether there exists an affine-evasive set of size  $p^{\Theta(1)}$ , which will imply constant rate NMC against affine-tampering and rate  $n^6$  NMC against split-state tampering.<sup>1</sup> We resolve this question in the affirmative by giving an affine-evasive set of size  $\Theta(\frac{p^{1/4}}{\log p})$ .

## 2 Explicit Construction

For any set  $S \subset \mathbb{Z}$ , let  $aS + b = \{as + b | s \in S\}$ . By  $S \pmod p \subseteq \mathbb{F}_p$ , we denote the set of values of  $S$  modulo  $p$ .

We first define an affine-evasive set  $S \subseteq \mathbb{F}_p$ .

**Definition 1** A non-empty set  $S \subseteq \mathbb{F}_p$  is said to be  $(\gamma, \nu)$ -affine-evasive if  $|S| \leq \gamma p$ , and for any  $(a, b) \in \mathbb{F}_p^2 \setminus \{(1, 0)\}$ , we have

$$|S \cap (aS + b \pmod p)| \leq \nu |S|.$$

---

<sup>1</sup>Under a plausible conjecture, this will imply constant rate NMC against split-state tampering. See Theorem 4 for more details.

Now we give a construction of an affine-evasive set.

Let  $Q := \{q_1, \dots, q_t\}$  be the set of all primes less than  $\frac{1}{2}p^{1/4}$ . Define  $S \subset \mathbb{F}_p$  as follows:

$$S := \left\{ \frac{1}{q_i} \pmod{p} \mid i \in [t] \right\}. \quad (1)$$

Thus,  $S$  has size  $\Theta(\frac{p}{\log p})$  by the prime number theorem.

**Theorem 1** *For any prime  $p$ , the set  $S$  defined in Equation (1) is  $(\frac{1}{2}p^{-3/4}, \Theta(p^{-1/4} \cdot \log p))$ -affine-evasive.*

*Proof.* Clearly,

$$|S| = t \leq \frac{1}{2}p^{1/4} = \frac{1}{2}p^{-3/4} \cdot p.$$

Fix  $a, b \in \mathbb{F}_p$ , such that  $(a, b) \neq (1, 0)$ . Now, we show that  $|S \cap (aS + b \pmod{p})| \leq 3$ . Assume, on the contrary, that there exist distinct  $\alpha_i \in Q$  for  $i \in \{0, 1, 2, 3\}$  such that  $1/\alpha_i \pmod{p} \in S \cap (aS + b \pmod{p})$ . We have

$$\frac{a}{\alpha_i} + b = \frac{1}{\beta_i} \pmod{p} \text{ for } i = 1, 2, 3, \quad (2)$$

where  $\alpha_i, \beta_i \in Q$  for  $i \in \{0, 1, 2, 3\}$ , and  $\beta_i \neq \beta_j$  for any  $i \neq j$ .

For any  $i$ , if  $\alpha_i = \beta_i$ , then  $b \cdot \alpha_i = 1 - a \pmod{p}$ , which has at most one solution (since we assume  $(a, b) \neq (1, 0)$ ). Thus, without loss of generality, we assume that  $\alpha_i \neq \beta_i$ , for  $i \in \{1, 2, 3\}$ , and  $\alpha_1 < \alpha_2 < \alpha_3$ .

From Equation (2), we have that

$$\frac{\frac{a}{\alpha_1} + b - \frac{a}{\alpha_2} - b}{\frac{a}{\alpha_1} + b - \frac{a}{\alpha_3} - b} = \frac{\frac{1}{\beta_1} - \frac{1}{\beta_2}}{\frac{1}{\beta_1} - \frac{1}{\beta_3}} \pmod{p},$$

which on simplification implies

$$(\beta_3 - \beta_1)(\alpha_2 - \alpha_1)\alpha_3\beta_2 = (\beta_2 - \beta_1)(\alpha_3 - \alpha_1)\alpha_2\beta_3 \pmod{p}.$$

Note that both the left-hand and right-hand side of the above equation takes values between  $\frac{-p}{16}$  and  $\frac{p}{16}$ , and hence the equality holds in  $\mathbb{Z}$  (and not just in  $\mathbb{Z}_p$ ).

$$(\beta_3 - \beta_1)(\alpha_2 - \alpha_1)\alpha_3\beta_2 = (\beta_2 - \beta_1)(\alpha_3 - \alpha_1)\alpha_2\beta_3. \quad (3)$$

By equation 3, we have that  $\alpha_3$  divides  $(\beta_2 - \beta_1)(\alpha_3 - \alpha_1)\alpha_2\beta_3$ . Clearly,  $\alpha_3$  is relatively prime to  $\beta_3$ ,  $\alpha_2$ , and  $\alpha_3 - \alpha_1$ . Therefore,  $\alpha_3$  divides  $(\beta_2 - \beta_1)$ . This implies

$$\alpha_3 \leq |\beta_2 - \beta_1|. \quad (4)$$

Also, from equation 3, we have that  $\beta_2$  divides  $(\beta_2 - \beta_1)(\alpha_3 - \alpha_1)\alpha_2\beta_3$ , which by similar reasoning implies  $\beta_2$  divides  $\alpha_3 - \alpha_1$ . Thus, using that  $\alpha_3 > \alpha_1$ ,

$$0 < \beta_2 \leq \alpha_3 - \alpha_1 < \alpha_3. \quad (5)$$

Similarly, we can obtain  $\beta_1$  divides  $\alpha_3 - \alpha_2$ , which implies

$$0 < \beta_1 \leq \alpha_3 - \alpha_2 < \alpha_3. \quad (6)$$

Equation (5) and (6) together imply that  $|\beta_2 - \beta_1| < \alpha_3$ , which contradicts Equation (4).  $\square$

### 3 Affine-evasive function and Efficient NMCs

We recall here the definition of affine-evasive functions from [ADL14]. Affine-evasive functions immediately give efficient construction of NMCs against affine-tampering.

**Definition 2** *A surjective function  $h : \mathbb{F}_p \mapsto \mathcal{M} \cup \{\perp\}$  is called  $(\gamma, \delta)$ -affine-evasive if for any  $a, b \in \mathbb{F}_p$  such that  $a \neq 0$ , and  $(a, b) \neq (1, 0)$ , and for any  $m \in \mathcal{M}$ ,*

1.  $\Pr_{U \leftarrow \mathbb{F}_p}(h(aU + b) \neq \perp) \leq \gamma$
2.  $\Pr_{U \leftarrow \mathbb{F}_p}(h(aU + b) \neq \perp \mid h(U) = m) \leq \delta$
3. *A uniformly random  $X$  such that  $h(X) = m$  is efficiently samplable.*

We now mention a result that shows that we can construct an affine-evasive function from an affine-evasive set  $S$ .

**Lemma 1 ([ADL14, Claim 5])** *Let  $S \subseteq \mathbb{F}_p$  be a  $(\gamma, \nu)$ -affine-evasive set with  $\nu \cdot K \leq 1$ , and  $K$  divides  $|S|$ .<sup>2</sup> Furthermore, let  $S$  be ordered such that for any  $i$ , the  $i$ -th element is efficiently computable in  $O(\log p)$ . Then there exists a  $(\gamma, \nu \cdot K)$ -affine-evasive function  $h : \mathbb{F}_p \mapsto \mathcal{M} \cup \{\perp\}$ .*

Note that the above result requires that for any  $i$ , the  $i$ -th element of  $S$  is efficiently computable for some ordering of the set  $S$ . This is not possible for our construction since for our construction this would mean efficiently sampling the  $i$ -th largest prime. However, this requirement was made just to make sure that  $h^{-1}$  is efficiently samplable. We circumvent this problem by giving a slightly modified definition of the affine-evasive function  $h$  in the proof of the following.

**Lemma 2** *There exists an efficiently computable  $(p^{-3/4}, \Theta(K \log p \cdot p^{-1/4}))$ -affine-evasive function  $h : \mathbb{F}_p \mapsto \mathcal{M} \cup \{\perp\}$ .*

*Proof.* Without loss of generality, let  $\mathcal{M} = \{1, \dots, K\}$ , for some integer  $K$ . Let  $S \subseteq \mathbb{F}_p$  be as defined in Section 2. Define  $S_1, \dots, S_K$  to be a partition of  $S$  as follows.

$$S_i := \left\{ s \in S \mid \frac{1}{s} \in \left[ \frac{i-1}{2K} p^{1/4}, \frac{i}{2K} p^{1/4} \right) \right\}. \quad (7)$$

Note that by the construction of  $S$  and the prime number theorem, each  $S_i$  has size at least  $\Theta\left(\frac{p^{1/4}}{K \log p}\right)$ .

Let  $h : \mathbb{F}_p \mapsto \mathcal{M} \cup \{\perp\}$  be defined as follows:

$$h(x) = \begin{cases} i & \text{if } x \in S_i \\ \perp & \text{otherwise.} \end{cases}$$

The statement  $\Pr(h(aU + b) \neq \perp) \leq p^{-3/4}$  is obvious by the definition of  $S$ , and the observation that  $aU + b$  is uniform in  $\mathbb{F}_p$ .

---

<sup>2</sup>The assumption  $K$  divides  $|S|$  is just for simplicity.

Also, for any  $m \in \mathcal{M}$ , and for any  $(a, b) \neq (1, 0)$ , and  $a \neq 0$ ,

$$\begin{aligned} \Pr(h(aU + b) \neq \perp | h(U) = m) &= \frac{\Pr(aU + b \in S \wedge U \in S_m)}{\Pr(U \in S_m)} \\ &\leq \frac{\Pr(aU + b \in S \wedge U \in S)}{|S_m|/p} \\ &= \frac{p}{|S_m|} \Pr(U \in S \cap (a^{-1}S - ba^{-1}) \pmod{p}) \\ &= \Theta(K \log p \cdot p^{-1/4}). \end{aligned}$$

Also, sampling a uniformly random  $X$  such that  $h(X) = m$  is equivalent to sampling a uniformly random prime  $q$  in the interval

$$I := \left[ \frac{m-1}{2K} p^{1/4}, \frac{m}{2K} p^{1/4} \right)$$

and computing  $1/q \pmod{p}$ . Sampling  $q$  can be done in time polynomial in  $\log p$  by repeatedly sampling a random element in  $I$  until we get a prime. Computing  $1/q \pmod{p}$  can be done efficiently using Extended Euclidean Algorithm.  $\square$

Note that the proof of Lemma 2 is identical to the proof of Lemma 1, except the proof that a uniformly random  $X$  such that  $h(X) = m$  is efficiently samplable for any given  $m$ . Using this and the construction of [ADL14], we get the following results.

**Theorem 2** *There exists an efficient coding scheme  $(\text{Enc}, \text{Dec})$  encoding  $k$ -bit messages to  $\Theta(k + \log(\frac{1}{\varepsilon}))$  that is  $\varepsilon$ -non malleable w.r.t. the family of affine tampering functions  $\mathcal{F}_{\text{aff}}$ .*

**Theorem 3** *There exists an efficient coding scheme  $(\text{Enc}, \text{Dec})$  encoding  $k$ -bit messages to  $\Theta((k + \log(\frac{1}{\varepsilon})^7))$  that is  $\varepsilon$ -non malleable w.r.t. the family of split-state tampering functions  $\mathcal{F}_{\text{split}}$ .*

Also, assuming the following conjecture from [ADL14], our result gives the first NMC with constant rate in the split-state model.

**Conjecture 1 ([ADL14, Conjecture 2])** *There exists absolute constants  $c, c' > 0$  such that the following holds. For any finite field  $\mathbb{F}_p$  of prime order, and any  $n > c'$ , let  $L, R \in \mathbb{F}_p^n$  be uniform, and fix  $f, g : \mathbb{F}_p^n \rightarrow \mathbb{F}_p^n$ . Then*

$$\Delta(\phi_{f,g}(L, R) ; \mathcal{D}) \leq p^{-cn}.$$

**Theorem 4** *Assuming Conjecture 1, there exists an efficient coding scheme  $(\text{Enc}, \text{Dec})$  encoding  $k$ -bit messages to  $\Theta(k + \log(\frac{1}{\varepsilon}))$  that is  $\varepsilon$ -non malleable w.r.t. the family of split-state tampering functions  $\mathcal{F}_{\text{split}}$ .*

## References

- [ADL14] D. Aggarwal, Y. Dodis, and S. Lovett. Non-malleable codes from additive combinatorics. In *STOC*, 2014. To appear.
- [CG14a] M. Cheraghchi and V. Guruswami. Capacity of non-malleable codes. In *Innovations in Theoretical Computer Science*. ACM, 2014. To appear.

- [CG14b] M. Cheraghchi and V. Guruswami. Non-malleable coding against bit-wise and split-state tampering. In *Theory of Cryptography Conference - TCC*. Springer, 2014. To appear.
- [DKO13] Stefan Dziembowski, Tomasz Kazana, and Maciej Obremski. Non-malleable codes from two-source extractors. In *Advances in Cryptology-CRYPTO 2013*. Springer, 2013.
- [DPW10] Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs. Non-malleable codes. In Andrew Chi-Chih Yao, editor, *ICS*, pages 434–452. Tsinghua University Press, 2010.
- [FMNV14] S. Faust, P. Mukherjee, J. Nielsen, and D. Venturi. Continuous non-malleable codes. In *Theory of Cryptography Conference - TCC*. Springer, 2014. To appear.
- [FMVW13] S. Faust, P. Mukherjee, D. Venturi, and D. Wichs. Efficient non-malleable codes and key-derivation for poly-size tampering circuits. *IACR Cryptology ePrint Archive*, 2013.
- [LL12] Feng-Hao Liu and Anna Lysyanskaya. Tamper and leakage resilience in the split-state model. In *Advances in Cryptology-CRYPTO 2012*, pages 517–532. Springer, 2012.