Detecting General Algebraic Manipulation Attacks

Kim Ramchen

Department of Computing and Information Systems
The University of Melbourne
kim.ramchen@unimelb.edu.au

Abstract

Algebraic manipulation detection codes are a class of error detecting codes which have found numerous applications in cryptography. In this paper we extend these codes to defeat general algebraic attacks - we call such codes general algebraic manipulation detection (GAMD) codes. Positive results are shown for the existence of GAMDs for the families of tampering functions corresponding to point additions and affine functions over a finite field. Compared to non-malleable codes, we demonstrate both positive and negative results regarding the existence of GAMDs for arbitrary families of tampering functions.

1 Introduction

Fault injection attacks are a class of attacks involve the deliberate introduction of errors into the circuity or memory modules of a cryptographic device in attempt to deduce some secret state. Algebraic manipulation detection codes [CDF+08] are a class of error detecting codes that can thwart such attacks when the class of induced faults corresponds to additions on code-words over a finite space. More precisely let s be a message supplied by an adversary, and suppose c, an element of an abelian group $\mathcal{G}$, is the corresponding code-word. If any $\Delta \in \mathcal{G}$ it holds that $c + \Delta$ decodes to s' for any $s' \neq s$, with probability bounded by ϵ, the scheme is said to be an AMD code with error probability with ϵ.

Even though AMD codes provide an elegant, keyless alternative to the widely used message authentication codes for robust transmission over an error-prone channel, they cannot defeat some types of powerful adversaries. Suppose that an AMD code is used to protect the output of a one time pad scheme. Let $\mathcal{E}(K \oplus M)$ be the output on ciphertext $c = K \oplus M$. If it happens that $\mathcal{E}$ possesses a linear homomorphism ϕ, then we have $\Delta M \circ_\phi c = \Delta M \circ_\phi \mathcal{E}(K \oplus M) = \mathcal{E}(K \oplus (M \oplus \Delta M)) = \mathcal{E}(K \oplus M')$, where M' is the message to be substituted. It is therefore desirable to consider a more powerful adversarial model in which an attacker can choose, in addition to the source message, a tampering function F

from a rich class of tampering functions $\mathcal{F}$. In this work, we consider precisely this model, when the class $\mathcal{F}$ corresponds to algebraic functions over some finite field $\mathbb{F}_q$ corresponding to the co-domain of the AMD code. We call such a code a generalised algebraic manipulation detection code (GAMD code). Following previous works on algebraic manipulation detection, we distinguish the case when the source message is assumed to be uniformly distributed over the message space, from the usual (which provides tampering detection with bounded error probability for any message). These are called weak generalised algebraic manipulation detection (weak GAMD) and generalised algebraic manipulation detection (GAMD) respectively.

1.1 Our Contributions

We formally introduce the model of generalised algebraic manipulation detection, in which tamperings corresponding to algebraic functions over the ambient field of the encoding function. In this model we review the previous constructions for manipulation detection against point additions. We show that such constructions translate directly to our new model, leading to direct instantiations of weak GAMDs and GAMDs for this class. Additionally we present a new construction for weak GAMDs in the case of encoding over $\mathbb{F}_2$ based upon the probabilistic method, leading to the following result (we actually construct a GAMD for a more general class of tampering functions, this is discussed in Section 3.1.1)

Claim 1. *Let n be a power of two. There exists a n^{c-1}-GAMD against the class of point additions on $\mathbb{F}_n$ with rate $c - o(1)$, for any constant $0 < c < 1$.*

We also consider attacks corresponding to the class of affine functions. Such attacks were considered by Aggarwal et al. [ADL14] who showed the existence of non-malleable codes for this class. We show that their constructions imply corresponding weak GAMD codes with constant rate and low error-probability. We present a black-box transformation of any weak GAMD to a GAMD. This construction is quite efficient, implying in view of the above results, the existence of GAMDs with constant rate and low error probability for the classes of point additions and affine functions respectively. Compared to the celebrated non-malleable codes [DPW10] we also establish some separations. Our first result is negative and states that there exists a class of tampering functions for which non-malleable codes but not GAMD codes exist. This may be summarised by

Claim 2. *There exists a family of tampering functions for which non-malleable codes exist with constant rate and negligible simulation error but ϵ-GAMD codes with constant rate do not exist, for any choice of non-negligible ϵ.*

Our second result is a positive one and states that for any non-malleable code there exists a class of tampering functions which violates malleability, but for which an efficient GAMD code exists, leading to

Claim 3. *For any non-malleable code $\mathcal{C}$ there exists a family of tampering functions such that $\mathcal{C}$ is non-malleable with respect to this family but there exists a GAMD for this family with constant rate and negligible error probability.*

2 Preliminaries

We describe the preliminary tools and definitions to be used throughout this paper. We begin firstly by reviewing non-malleable codes [DPW10], secondly by stating some combinatorial results and finally, in Section 2.3, by stating our generalisation of classical algebraic manipulation detection codes [CPS02, DKRS06, CDF+08].

2.1 Non-Malleable Codes

We recall the notion of non-malleable codes for a class of tampering functions. Informally a non-malleable code is one which guarantees that after decoding either the original message is recovered or the message that is recovered is completely "unrelated" to the original.

Definition 1 (Non-Malleable Code [DPW10])**.** *Let $\mathcal{F}$ be a family of tampering functions. For each $F \in \mathcal{F}$ and $s \in \{0,1\}^k$, define the tampering experiment*

$$\mathsf{Tamper}_s^F =: \left\{ \begin{array}{c} c \leftarrow \mathsf{Enc}(s), \tilde{c} \leftarrow F(c), \tilde{s} = \mathsf{Dec}(c) \\ \textit{Output } \tilde{s}. \end{array} \right\}$$

defining a random variable over the randomness of the encoding function Enc*. Say that a coding scheme* $(\mathsf{Enc}, \mathsf{Dec})$ *is non-malleable w.r.t. $\mathcal{F}$ if for each $F \in \mathcal{F}$, there exists a distribution D_F over $\{0,1\}^k \cup \{\bot, \mathsf{same}^*\}$, such that, for all $s \in \{0,1\}^k$, we have:*

$$\mathsf{Tamper}_s^F \approx \left\{ \begin{array}{c} \tilde{s} \leftarrow D_F \\ \textit{Output } s \textit{ if } \tilde{s} = \mathsf{same}^*, \textit{ and } \tilde{s} \textit{ otherwise.} \end{array} \right\}$$

and D_f is efficiently samplable given oracle access to $F(\cdot)$.

Let $\mathcal{F}_{\mathsf{bit}}$ be the family of tampering functions that tamper every bit of a code-word of length n independently. Formally, $\mathcal{F}_{\mathsf{bit}}$ contains all functions $f : \{0,1\}^n \to \{0,1\}^n$ defined by n functions $f_i : \{0,1\} \to \{0,1\}$, namely $f(c_1, \ldots, c_n) = f_1(c_1), \ldots, f(c_n)$. Each f_i is an affine function on $\mathbb{Z}_2$. We require the following proposition proved by [DPW10], concerning the existence of non-malleable codes against the family of bit-wise independent tampering functions with constant rate and negligible simulation error.

Lemma 4 (Theorem 4.2 [DPW10])**.** *For any $\delta > 0$ and $n \in \mathbb{N}$ there exist non-malleable codes w.r.t the family $\mathcal{F}_{\mathsf{bit}}$, with block length n, message size $k \geq (.811 - \delta)n$ and simulation error $2^{-\Omega(n)}$. Moreover there is an efficient procedure which, given k and n, outputs a description of such a code with probability $1 - 2^{-\Omega(n)}$.*

2.2 Combinatorial Tools

We describe some combinatorial tools used in our constructions of GAMDs.

Definition 2 (Trace [CDN15]). *Let K and L be fields. Suppose that L is separable over K and $n := [L : K] > \infty$. Fix some algebraic closure $\bar{L}$ of L. Let $\sigma_1, \ldots, \sigma_n$ be the distinct K-embeddings of L into $\bar{L}$. The trace map $\mathrm{Tr}_{L/K}$ for each $x \in L$ is:*

$$\mathrm{Tr}_{L/K}(x) = \sum_{i=1}^{n} \sigma_i(x) \in K$$

Definition 3 (Difference Set [CD06]). *Let $(\mathcal{G}, +)$ be an additive abelian group of order v. A subset $D \subseteq \mathcal{G}$ is a (v, c, λ)-external difference set if $|D| = c$ and every non-zero element of $\mathcal{G}$ has exactly λ representations as a difference $d - d'$ for $d, d' \in D$. If every non-zero element of $\mathcal{G}$ has at most λ representations $d - d'$, say that D is a (v, c, λ)-bounded difference set.*

Definition 4 (Authentication Code [Sti90]). *Let $\mathcal{S}$ be a set of source states, $\mathcal{K}$ a set of authentication keys and $\mathcal{A}$ be a mapping $\mathcal{A} : \mathcal{S} \times \mathcal{K} \to \mathcal{T}$ where $\mathcal{T}$ is a set of tags. Let Π be a probability distribution on $\mathcal{K}$. The probability of a successful substitution attack, with respect to family of substitution functions $\mathcal{F}$, is*

$$p_{\mathcal{F}}^{\mathsf{sub}} =: \max_{F \in \mathcal{F}, s \neq s' \in \mathcal{S}} \Pr_{K \leftarrow \Pi}[F(A(s, K)) = A(s', K)].$$

Lemma 5 (Schwartz-Zippel). *Let K be a field and let $P \in K[x_1, \ldots, x_n]$ where $(x_i)_{1 \leq i \leq n}$ are indeterminates. Let $S \subseteq K$ be a finite set and let $(u_i)_{1 \leq i \leq n}$ be selected independently and uniformly at random in K. Then*

$$\Pr[P(u_1, \ldots, u_n) = 0] \leq \frac{\deg(P)}{|S|}$$

2.3 Generalised Algebraic Manipulation Detection Codes

In this section we define a code which is a generalisation of the classical algebraic manipulation detection coding schemes. The main difference is simply that we allow manipulation functions be general algebraic functions modulo a prime, rather than the restriction to point additions on its additive group considered by [CPS02, CDF+08].

Definition 5. *Let p be a prime and n be a positive integer. Let $\mathcal{G} = (\mathbb{F}_p)^n$ and let $\mathcal{F}$ be a family of algebraic tampering functions on $\mathcal{G}$.*[1] *Let $\mathcal{S}$ be a set of symbols. Let $\mathcal{E} : \mathcal{S} \to \mathcal{G}$ be a probabilistic encoding and $\mathcal{D} : \mathcal{G} \to \mathcal{S} \cup \{\perp\}$ be a deterministic decoding procedure such that $\Pr_{\mathcal{E}}[\mathcal{D}(\mathcal{E}(s)) = s] = 1$ for all $s \in \mathcal{S}$.*

- *The tuple $(\mathcal{E}, \mathcal{D})$ is an ϵ-generalised algebraic manipulation detection (GAMD) code if $\forall s \in \mathcal{S}, \forall F \in \mathcal{F}$ $\Pr_{\mathcal{E}}[\mathcal{D}(F(\mathcal{E}(s))) \notin \{s, \perp\}] \leq \epsilon$.*
- *The tuple $(\mathcal{E}, \mathcal{D})$ is a weak ϵ-generalised algebraic manipulation detection code if $\forall F \in \mathcal{F}$ $\Pr_{\mathcal{E}, s \in_R \mathcal{S}}[\mathcal{D}(F(\mathcal{E}(s))) \notin \{s, \perp\}] \leq \epsilon$.*

The (information) rate of a GAMD code is defined as $r = \frac{\log_2 |\mathcal{S}|}{\log_2 |\mathcal{G}|}$.

[1] Recall that a function in n variables is algebraic iff it is the root of a polynomial equation in n+1 variables.

2.3.1 Families of Tampering Functions

In this paper we consider two classes of tampering functions on a GAMD $(\mathcal{E}, \mathcal{D})$ with co-domain $\mathcal{G} = \mathbb{F}_p^n$ for some prime p and positive integer n.

- **Point Additions:** let $\mathcal{F}_{\mathsf{add}} = \{F_\Delta\}_{\Delta \in \mathcal{G}}$ where $F_\Delta := x \mapsto x + \Delta$ over $\mathcal{G}$.
- **Affine Functions:** let $\mathcal{F}_{\mathsf{aff}} = \{F_{(a,b)}\}_{a,b \in \mathcal{G}}$ where $F_{(a,b)} := x \mapsto ax + b$ over $\mathcal{G}$.

3 Constructions

In this section we review some constructions for GAMD codes against the class of tampering functions corresponding to point additions and also affine functions. Our results show that efficient GAMDs (i.e, one ones with constant rate and low error probability) exist for these classes. Specifically for the class of point additions, we present two constructions of GAMDs based upon different sets. Our first can be seen as a specific instantiation of the AMD codes implicit in Section 4.1 [CPS02]. Our second which is based upon the probabilistic method, achieves only slightly worse parameters, while allowing the construction of GAMDs for a considerably broader class of functions.

3.1 Point Additions

We begin with some auxiliary lemmas, our objective is to explicitly construct a difference set via the template described in Section 4.1 [CPS02], which involved constructing a difference set in $\mathbb{F}_{q^n} \times \mathbb{F}_{q^r}$ from any surjective map $\phi : \mathbb{F}_{q^n} \to \mathbb{F}_{q^r}$. Our specific construction, which involves using the Trace function (Definition 1) for ϕ, is described in Lemma 8. Using this construction we can build a weak-GAMD with rate $1 - o(1)$ and arbitrarily low error probability, described in Lemma 9.

Proposition 6. *Suppose that p is a prime and l and k are positive integers such that $k|l$. Then $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(\alpha) = \sum_{i=0}^{\frac{l}{k}-1} \alpha^{p^{ki}}$.*

Proof. See Appendix. □

Corollary 7. *Suppose that p is a prime and l and k are positive integers such that $k|l$. Then $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(\alpha)$ is a surjective, $\mathbb{F}_{p^k}$-linear map.*

Proof. See Appendix. □

Lemma 8. *Let p be an odd prime and l and k be positive integers such that $k|l$. Let $(\mathcal{G}, +)$ be the product of groups, $\mathbb{F}_{p^l} \times \mathbb{F}_{p^k}$ under addition. Define*

$$D = \{(\alpha, \mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(\alpha^2) : \alpha \in \mathbb{F}_{p^l}\} \subseteq \mathcal{G}$$

Then D is a (p^{l+k}, p^l, p^{l-k})-external difference set.

Proof. See Appendix. □

Lemma 9. *For a prime p and positive integer n let $\mathcal{G} = \mathbb{F}_p^n$. Then there exists a deterministic weak (p^{-1})-GAMD code with respect to the family of point additions, $\mathcal{F}_{\mathsf{add}}$, on $\mathcal{G}$, with efficient encoding and decoding procedure and rate $1 - o(1)$.*

Proof. Note $\mathcal{G} \sim (\mathbb{F}_{p^n}, +)$. By Lemma 8 we know that for any $n > 1$ there exists a (p^n, p^{n-1}, p^{n-2})-external difference set $D \subseteq \mathcal{G}$. Let $\mathcal{E}(\mathcal{S}) = D$ and consider the quantity $p_\Delta := \Pr_{s \in_R S}[F_\Delta(\mathcal{E}(s)) \notin \{s, \perp\}]$. Since $\mathcal{E}$ is deterministic, and s is chosen uniformly at random, $p_\Delta = \frac{\#\{s' \in S : \mathcal{E}(s') - \mathcal{E}(s) = \Delta\}}{|\mathcal{S}|}$. Thus for each $\Delta \in \mathcal{G}$, since $\mathcal{E}(\mathcal{S})$ is a (p^n, p^{n-1}, p^{n-2})-difference set, $p_\Delta \leq \frac{p^{n-2}}{p^{n-1}} = \frac{1}{p}$. The rate of $\mathcal{E}$ is $\frac{\log |D|}{\log |\mathcal{G}|} = 1 - n^{-1} = 1 - o(1)$, as required. □

3.1.1 A New Construction

We note that so far the constructions of GAMD codes against the class of point additions have followed a similar recipe to the constructions of AMD codes presented in [CPS02, CDF+08]. In this section we present a new construction for this class based upon the probabilistic method.

Definition 6. *For even integer n denote by I_n, the subset of permutations on n objects consisting of involutions with no fixed points.*

Lemma 10. *Let $\mathcal{G}$ be an abelian group of order n where n is even. Let $0 \leq c < 1$ be arbitrary. Let $I'_n \subset I_n$ be of polynomial size. Then there exists a subset $S \subset \mathcal{G}$ and maps $\mathcal{E} : [|S|] \to \mathcal{G}$ and $\mathcal{D} : \mathcal{G} \to [|S|]$ which define a weak n^{c-1}-GAMD with respect to the set I'_n. The rate is $c - o(1)$.*

Proof. We show that, by analogy to the affine-evasive set case (see Section 3.2 following) for any positive constants $0 \leq \gamma < \nu < 1$, there exists a set $S \subset \mathcal{G}$ for which $|S| \leq \gamma|\mathcal{G}|$ and $|S \cap F(S)| \leq \nu|S|$ hold for any $F \in I'_n$. Taking $\mathcal{S} = [|S|]$, $\nu = n^{c-1}$, $\gamma = n^{(c-1)-o(1)}$ and $\mathcal{E}$ and $\mathcal{D}$ as in the statement of the lemma, yields a code with error probability n^{c-1} and rate $\frac{\log \gamma n}{\log n} = c - o(1)$ (see also Lemma 13, Section 3.2). We will demonstrate the existence of S via a probabilistic argument.

Consider the set S defined by sampling each element of $\mathcal{G}$ independently with probability γ. Clearly the size of S, N_0, has a Binomial distribution with parameters (n, γ). We now analyse the set of the intersection $S \cap F(S)$, where $F \in I'_n$ is arbitrary. Observe that each such F induces a matching on $\mathcal{G}$ given by $(x, F(x)) : x < F(x)$. Moreover, since F contains no fixed points, each such pair occurs independently with probability γ^2. Thus $N_1 := |S \cap F(S)|/2$ follows a Binomial distribution, with parameters $(\frac{n}{2}, \gamma^2)$. Now by applying Chernoff bounds, if ϵ is such that $\frac{\nu}{\gamma} > 1 - \epsilon$ then

$$Pr[N_0 \leq n\gamma(1-\epsilon)] \leq e^{\frac{-n\gamma\epsilon^2}{2}}$$
$$Pr[N_1 \geq \frac{\nu n\gamma(1-\epsilon)}{2}] \leq e^{\frac{-n\gamma^2(\nu(1-\epsilon)-\gamma)^2}{6}}$$

Secondly, applying a union bound over all $F \in I'_n$, we have

$$\Pr_S[|S| \geq n\gamma(1-\epsilon) \cap |S \cap F(S)| \leq \nu n\gamma(1-\epsilon) \text{ for all } F \in I'_n] \geq 1 - e^{-\frac{n\gamma\epsilon^2}{2}} - |I'_n|e^{-\frac{n\gamma^2(\nu(1-\epsilon)-\gamma)^2}{6}}$$

As $|I'_n|$ is polynomial in n, for large enough n this probability is strictly greater than 0. Thus S exists for which $|S| < n\gamma(1-\epsilon)$ and $|S \cap F(S)| \leq \nu|S|$ for all $F \in I'_n$. □

Corollary 11. *Let $G = (\mathbb{F}_n, +)$ where $n = 2^k$. Then there exists a weak $(n^{-1/2})$-GAMD with respect to the family $\mathcal{F}_{\mathsf{add}}$, with rate $\frac{1}{2} - o(1)$.*

Proof. The family $\mathcal{F}_{\mathsf{add}}$ defines a subset of I_n of order n. Thus $S \subseteq \mathcal{G}$ exists with the properties of Lemma 10, taking $c = 1/2$ yields a $n^{-1/2}$-GAMD with rate $1/2 - o(1)$. In fact, S defines an $(n, \sqrt{n}, 1)$-bounded difference set. □

3.2 Affine Functions

In this section we review known results about non-malleable codes resistant to the class of affine functions [ADL14, Agg15] in the GAMD setting.

Definition 7. *[ADL14] A non-empty set $S \subseteq \mathbb{F}_p$ is said to be (γ, ν)-affine-evasive if $|S| \leq \gamma p$, and for any $(a, b) \in \mathbb{F}_p^2 \backslash \{(1, 0)\}$, we have*

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

Theorem 12. *[Agg15] For any sufficiently large prime p, there exists a set $S \subset \mathbb{Z}_p$ that is $(\Theta(p^{-3/4}/\log p), \Theta(p^{-1/4}\log p))$-affine-evasive. Moreover the i^{th} element of S, S_i, is samplable in polynomial time.*

Lemma 13. *Let p be a large prime. Let $\mathcal{S} = [|S|]$, where S is the set defined in Theorem 12. Let $\mathcal{G} = \mathbb{F}_p$. Then the map $\mathcal{E} : \mathcal{S} \to \mathcal{G}$ given by $E(i) = S_i$ together with the map $\mathcal{D} : \mathcal{G} \to \mathcal{S}$ given by*

$$\mathcal{D}(x) = \begin{cases} i \text{ if } x \in S_i \text{ for some } i \\ \perp \text{ if } x \in \mathbb{F}_p \backslash \{S\} \end{cases}$$

defines a weak ϵ-GAMD with respect to the class of non-trivial affine functions $\mathcal{F}^*_{\mathsf{aff}}$. Here $\epsilon = p^{-1/4}\log p$ and the rate is $\frac{1}{4} - o(1)$.

Proof. Let $E(\mathcal{S}) = S \subseteq \mathbb{F}_p$. The map $\mathcal{E}$ is efficient, since the i^{th} element of S is samplable in polynomial time. For $F \in \mathcal{F}^*_{\mathsf{aff}}$, define $p_F := \Pr_{s \in_R \mathcal{S}}[F(E(s)) \notin \{s, \perp\}]$. Since $\mathcal{E}$ is deterministic and s is uniformly distributed over $\mathcal{S}$, $p_F = \frac{\#\{s, s' \in \mathcal{S} \ a\mathcal{E}(s)+b=\mathcal{E}(s')\}}{|\mathcal{S}|}$. Then for each $F \equiv x \mapsto ax + b$ we have $p_F \leq \frac{\nu|\mathcal{S}|}{\mathcal{S}} \leq \nu = p^{-1/4}\log p$, by Theorem 12. The rate is $\frac{\log_2|\mathcal{S}|}{\log_2 p} = \frac{\log(p^{1/4}/\log p)}{\log p} = 1/4 - o(1)$, as required. □

3.3 A Weak GAMD to GAMD Transformation

In this section we present a sufficient result for transforming any weak GAMD to a GAMD. Our main result here is Lemma 15, which states that if the classes of tampering functions can be represented by a set of polynomials in one or more variable of bounded degree $d = o(p)$, then any weak GAMD for this family can transformed to a GAMD.

Proposition 14. *Suppose that $(\mathcal{E}', \mathcal{D}')$ is a weak ϵ'-GAMD with respect the a family of algebraic tampering functions $\mathcal{F}$ where $\mathcal{E}' : \mathcal{S}' \to \mathcal{G}'$. Let $\mathcal{A} : \mathcal{S} \times \mathcal{K} \to \mathcal{T}$ be a message authentication code, $\mathcal{K} = \mathcal{S}'$, with maximum substitution probability $p^{\mathsf{sub}}_{\mathcal{F}}$ with respect to family $\mathcal{F}$. Let $\mathcal{G} = \mathcal{S} \times \mathcal{G}' \times \mathcal{T}$. Define $\mathcal{E} : \mathcal{S} \to \mathcal{G}$ by $\mathcal{E}(s) = (s, \mathcal{E}'(k), \mathcal{A}(s, k))$, where $k \in_R \mathcal{K}$. Define $\mathcal{D} : \mathcal{G} \to \mathcal{S} \cup \{\perp\}$ by $\mathcal{D}(s, c', \tau) = s$ iff $\mathcal{D}'(c') \neq \perp$ and $\tau = \mathcal{A}(s, \mathcal{D}'(c'))$. Then $(\mathcal{E}, \mathcal{D})$ is an ϵ-GAMD with respect to $\mathcal{F}$ where $\epsilon = \epsilon' + p^{\mathsf{sub}}_{\mathcal{F}}$.*

Proof. Suppose that $c = (\tilde{s}, \tilde{c'}, \tilde{\tau})$ is a received code-word for source symbol s under key k. Suppose that $s \neq \tilde{s}$. Then $\Pr[\mathcal{D}'(\tilde{c'}) \neq \{k, \perp\}] \leq \epsilon'$ since $(\mathcal{E}', \mathcal{D}')$ is a weak ϵ'-GAMD and k is chosen uniformly at random in $\mathcal{K}$. Moreover, $\Pr[\mathcal{A}(\tilde{s}, k) = \tilde{\tau}] \leq p^{\mathsf{sub}}_{\mathcal{F}}$ since $s \neq \tilde{s}$. Thus the event $\mathcal{D}(c) = \tilde{s}$ occurs with probability at most $\epsilon' + p^{\mathsf{sub}}_{\mathcal{F}}$. The claim follows. □

Lemma 15. *Let p be a prime and d and l be positive integers. Let $\mathcal{P}_{\leq d}$ be the space of all polynomials of total degree at most d. Let $\mathcal{A} : \mathcal{S} \times \mathcal{K} \to \mathcal{T}$ be the message authentication code defined by $\mathcal{A}((s_1, \ldots, s_l), (x, y)) = \sum_{i=1}^{l} s_i x^i + y$. Then $p^{\mathsf{sub}}_{\mathcal{P}_{\leq d}} \leq \frac{ld}{p}$.*

Proof. Let F be a fixed polynomial in $\mathcal{P}_{\leq d}$. Let $s \neq s' \in \mathcal{S}$. Consider the polynomial $P(x, y) = F(\sum_{i=1}^{l} s_i x^i + y) - (\sum_{i=1}^{d} s'_i x^i + y)$ in $\mathbb{F}_p[x, y]$. We argue this is a non-zero polynomial as follows. First observe that if $P \equiv 0$, then $\deg(F) = 1$, since otherwise $P(x, y)$ contains a non-trivial power of y. So let $F(u) = a_0 u + a_1$. Then $a_0 = 1$ by a similar argument. Thus $P = \sum_{i=1}^{l} (s_i - s'_i) x^i + a_0$, which is a contradiction since $s \neq s'$ implies there exists i for which $s_i \neq s'_i$. On the other hand the degree of P is at most $\deg(F) \cdot l \leq ld$. Thus by the Schwartz-Zippel Lemma, P has at most ld roots. As $k = (x, y)$ is chosen uniformly in $\mathbb{F}_p^2$, the event $P = 0$ occurs with probability at most $\frac{ld}{p}$. Finally, $P = 0 \Leftrightarrow F(\mathcal{A}(s, k)) = \mathcal{A}(s', k)$, concluding the proof. □

Corollary 16. *For any $n \in \mathbb{N}$ there exists a ϵ-GAMD with of block length n, with respect to the family $\mathcal{F}_{\mathsf{add}}$ where $\epsilon = 2^{-\Omega(n)}$. The rate is $1 - o(1)$.*

Proof. Pick prime p so that $p > 2^n$. By Lemma 9 we can construct $\mathcal{E}'$ so that $\mathcal{E}' : \mathbb{F}_p^2 \to \mathbb{F}_p^3$ has error probability $\frac{1}{p}$. Let $\mathcal{A} : \mathbb{F}_p^{n-4} \times \mathbb{F}_p^2 \to \mathbb{F}_p$ as in Lemma 15. Then as $\deg(F) = 1$ for all $F \in \mathcal{F}_{\mathsf{add}}$, we have $p^{\mathsf{sub}}_{\mathcal{F}_{\mathsf{add}}} \leq \frac{n-4}{p}$ by Lemma 15. The rate of $\mathcal{E}$ is $\frac{n-4}{n} = 1 - o(1)$. The error probability is bounded by $\epsilon = p^{\mathsf{sub}}_{\mathcal{F}_{\mathsf{add}}} + p^{-1} \leq \frac{n-3}{p} = 2^{-\Omega(n)}$. □

We can prove an even stronger result, assuming a mixed alphabet $\mathcal{K} \neq \mathbb{F}_p^2$.

Corollary 17. *For any $n \in \mathbb{N}$ there exists a ϵ-GAMD with of block length n, with respect to the family $\mathcal{F}_{\mathsf{aff}}$ with $\epsilon = 2^{-\Omega(n)}$. The rate is $1 - o(1)$.*

Proof. Pick primes p, p' so that $p > 2^n$ and $\frac{p^2}{2} < p' < p^2$, by Bertrand's postulate. By Lemma 10 we can construct $\mathcal{E}'$ so that $\mathcal{E}' : \mathcal{K} \to \mathbb{F}_{p'}$, $\mathcal{K} \subseteq [p']$ with error probability $\epsilon' = {p'}^{-1/4} \log_2 p' < p^{-1/2+\delta}$ for any $\delta > 0$. Let $\mathcal{A} : \mathbb{F}_p^{n-3} \times \mathbb{F}_p^2 \to \mathbb{F}_p$ as in Lemma 15. Then as $\deg(F) = 1$ for all $F \in \mathcal{F}_{\mathsf{aff}}$, by Lemma 15 we have $p^{\mathsf{sub}}_{\mathcal{F}_{\mathsf{aff}}} \leq \frac{n-3}{p}$. The rate of $\mathcal{E}$ is $\frac{n-3}{n} = 1 - o(1)$. The error probability is bounded by $\epsilon = p^{\mathsf{sub}}_{\mathcal{F}_{\mathsf{aff}}} + \frac{1}{p'} \leq p^{-3/7} + \frac{n-3}{p} = 2^{-\Omega(n)}$. □

4 Separations

In this section we describe some separations regarding non-malleable codes and GAMDs. Although non-malleable codes have already proved a valuable digression from the classical notion of error correction and detection, here we provide evidence that GAMD codes provide a strengthening of classical algebraic manipulation detection distinct to that provided by non-malleable cryptography. Specifically we are able to prove (Theorem 19) that any non-malleable code can be broken by some tampering family for which a GAMD with high rate and low error probability exists. This family actually corresponds to a re-coding functionality in which a code-word is decoded, one is added to the message which is then again encoded, so is a natural candidate for this task. On the negative side, however, we show that for at least one family of tampering functions, non-malleable codes exist but GAMDs do not.

Theorem 18. *There exists a family of tampering functions $\mathcal{F}$ for which for any $n \in \mathbb{N}$, non-malleable codes of block length n exist with constant rate and simulation error $2^{-\Omega(n)}$, but ϵ-GAMD codes with constant rate do not exist, for any choice of non-negligible (in block-length) ϵ.*

Proof. Let $\mathcal{F} = \mathcal{F}_{\mathsf{bit}}$ be the family of bit-wise independent tampering functions (see Section 2.1). By Lemma 4 we know that for any n there exists a non-malleable code $(\mathsf{Enc}, \mathsf{Dec})$ with block size n, simulation error $2^{-\Omega(n)}$ and rate $\approx .811$. Now suppose that an ϵ-GAMD $(\mathcal{E}, \mathcal{D})$ exists for $\mathcal{F}_{\mathsf{bit}}$ where $\mathcal{E} : \mathcal{S} \to \mathbb{F}_2^n$. Since $|\mathcal{S}| > 1$ we know that there exists distinct code-words $c = \mathsf{Enc}(s), c' = \mathsf{Enc}(s') : c \neq c'$. The the function $F_{c'}(x) = c'$ is contained in $\mathcal{F}_{\mathsf{bit}}$ and is not detectable except with probability at most $\frac{1}{|\mathcal{S}|}$. By assumption $(\mathcal{E}, \mathcal{D})$ is constant rate, so $\epsilon \leq \frac{1}{|\mathcal{S}|} = 2^{-\Omega(n)}$. □

Theorem 19. *For any non-malleable code $\mathcal{C}$ of block length n there exists a family $\mathcal{F}$ of tampering functions $\mathcal{F}$ such that $\mathcal{C}$ is non-malleable with respect to $\mathcal{F}$ but there exists an $(2^{-\Omega(n)})$-GAMD code $\mathcal{C}'$ with respect to $\mathcal{F}$ with rate $r - o(1)$, where r is the rate of $\mathcal{C}$.*

Proof. Let $(\mathsf{Enc}, \mathsf{Dec})$ be a non-malleable code where $\mathsf{Enc} : \mathbb{F}_p^k \to \mathbb{F}_p^n$. We construct GAMD code $(\mathcal{E}, \mathcal{D})$ and family of tampering functions $\mathcal{F}$, as follows. Let $\mathbf{1} = (0, \ldots, 0, 1) \in \mathbb{F}_p^k$ and let F be the function

$$F(c) = \begin{cases} \mathsf{Enc}(\mathsf{Dec}(c) + \mathbf{1}) \text{ if } \mathsf{Dec}(c) \neq \ \perp \\ c \text{ otherwise} \end{cases}$$

and $\mathcal{F} = \{F\}$. Note that F being a polynomial in n variables, is indeed an algebraic function. Let $(\mathcal{E}', \mathcal{D}')$ with $\mathcal{E}' : \mathcal{S} \to \mathbb{F}_p^k$ be an ϵ'-GAMD with respect to the family $\mathcal{F}_{\mathsf{add}}$ of point addition functions on $\mathbb{F}_p^k$. Define $\mathcal{E} : \mathcal{S} \to \mathbb{F}_p^k$ and $\mathcal{D} : \mathbb{F}_p^k \to \mathcal{S}$ by

$$\mathcal{E}(s) = \mathsf{Enc}(\mathcal{E}'(s))$$
$$\mathcal{D}(c) = \begin{cases} \mathcal{D}'(\mathsf{Dec}(c)) & \text{if } \mathsf{Dec}(c) \neq \perp \\ \perp & \text{otherwise} \end{cases}$$

We claim that $(\mathcal{E}, \mathcal{D})$ is an ϵ'-GAMD for $\mathcal{F}$ as follows. We have

$$\begin{aligned} &\Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}(\mathcal{E}(s)) = s] \\ &= \Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}(\mathsf{Enc}(\mathcal{E}'(s))) = s] \\ &= \Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}'(\mathsf{Dec}(\mathsf{Enc}(\mathcal{E}'(s)))) = s] \\ &= \Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}'(\mathcal{E}'(s))) = s] = 1 \end{aligned}$$

by the correctness of non-malleable $(\mathsf{Enc}, \mathsf{Dec})$ and ϵ'-GAMD $(\mathcal{E}', \mathcal{D}')$ respectively. On the other hand for any $s \neq s'$ in $\mathcal{S}$,

$$\begin{aligned} &\Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}(F(\mathcal{E}(s))) = s'] \\ &\leq \Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}'(\mathsf{Dec}(F(\mathcal{E}(s)))) = s'] \\ &\leq \Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}'(\mathsf{Dec}(\mathsf{Enc}((\mathcal{E}(s) + \mathbf{1})))) = s'] \\ &\leq \Pr_{\mathsf{Enc},\, \mathcal{E}'}[\mathcal{D}'((\mathcal{E}(s) + \mathbf{1})) = s'] \leq \epsilon' \end{aligned}$$

as $(\mathcal{E}', \mathcal{D}')$ detects tampering by point additions on $\mathbb{F}_p^k$ with probability at least $1 - \epsilon'$. The rate of $(\mathcal{E}, \mathcal{D})$ is

$$\frac{\log |S|}{\log |\mathcal{C}|} = \frac{\log p^k}{\log |C|} \cdot \frac{\log |S|}{\log |\mathcal{G}'|} = r \cdot (1 - o(1)) = r - o(1)$$

since Lemma 6 implies we can choose $(\mathcal{E}', \mathcal{D}')$ such that $\mathcal{G}' \subseteq \mathbb{F}_p^k$ with rate $1 - o(1)$. □

References

[ADL14] Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. In *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 774–783, New York, NY, USA, 2014. ACM.

[Agg15] Divesh Aggarwal. Affine-evasive sets modulo a prime. *Inf. Process. Lett.*, 115(2):382–385, February 2015.

[CD06] Charles J. Colbourn and Jeffrey H. Dinitz. *Handbook of Combinatorial Designs, Second Edition (Discrete Mathematics and Its Applications)*. Chapman & Hall/CRC, 2006.

[CDF+08] Ronald Cramer, Yevgeniy Dodis, Serge Fehr, Carles Padró, and Daniel Wichs. Detection of algebraic manipulation with applications to robust secret sharing and fuzzy extractors. In Nigel Smart, editor, *Advances in Cryptology – EUROCRYPT 2008: 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, pages 471–488, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[CDN15] Ronald Cramer, Ivan Bjerre Damgård, and Jesper Buus Nielsen. *Secure Multiparty Computation and Secret Sharing*. Cambridge University Press, New York, NY, USA, 1st edition, 2015.

[CPS02] Sergio Cabello, Carles Padró, and Germán Sáez. Secret sharing schemes with detection of cheaters for a general access structure. *Des. Codes Cryptography*, 25(2):175–188, February 2002.

[DKRS06] Yevgeniy Dodis, Jonathan Katz, Leonid Reyzin, and Adam Smith. Robust fuzzy extractors and authenticated key agreement from close secrets. In *Proceedings of the 26th Annual International Conference on Advances in Cryptology*, CRYPTO'06, pages 232–250, Berlin, Heidelberg, 2006. Springer-Verlag.

[DPW10] Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs. Non-malleable codes. In *ICS*, pages 434–452, 2010.

[Sti90] D. R. Stinson. The combinatorics of authentication and secrecy codes. *J. Cryptol.*, 2(1):23–49, January 1990.

A Proof of Auxiliary Results

Proof of Proposition 6

Proof. It suffices to show that the p^k-power map $\phi_{p^k} : \alpha \to \alpha^{p^k}$ on $\mathbb{F}_{p^l}$ generates $\mathrm{Gal}(\mathbb{F}_{p^l}/\mathbb{F}_{p^k})$. First note that by Lagrange's theorem, $x^{p^k} = x$ for all $x \in \mathbb{F}_{p^k}$, so $\phi_{p^k} : \mathbb{F}_{p^l} \to \mathbb{F}_{p^l}$ fixes $\mathbb{F}_{p^k}$ point-wise. Also ϕ_{p^k} is a field homomorphism and therefore injective. Since $\mathbb{F}_{p^l}$ is finite, it follows that ϕ_{p^k} is also surjective. Therefore $\phi_{p^k} \in \mathrm{Gal}(\mathbb{F}_{p^l}/\mathbb{F}_{p^k})$. Now suppose that the order of ϕ_{p^k} is $t \geq 1$. Then $(\phi_{p^k})^t = \mathrm{id}$. hence $q(x) = x^{p^{tk}} - x$ has at least $|\mathbb{F}_{p^l}| = p^l$ roots. Since $q(x)$ is of degree p^{tk}, we have $p^{tk} \geq p^l$, so $t \geq l/k$. On the other hand $|\mathrm{Gal}(\mathbb{F}_{p^l}/\mathbb{F}_{p^k})| = [\mathbb{F}_{p^l} : \mathbb{F}_{p^k}] = l/k$. It follows that ϕ_{p^k} has order exactly l/k, i.e. it generates $\mathrm{Gal}(\mathbb{F}_{p^l}/\mathbb{F}_{p^k})$, completing the proof. □

Proof of Corollary 7

Proof. For any $c \in \mathbb{F}_{p^k}$, we have $c^{p^k} = c$. Therefore $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(c\alpha) = \sum_{i=0}^{\frac{l}{k}-1}(c\alpha)^{p^{ki}} = \sum_{i=0}^{\frac{l}{k}-1} c^{p^{ki}}\alpha^{p^{ki}} = \sum_{i=0}^{\frac{l}{k}-1}(c)\alpha^{p^{ki}} = c\sum_{i=0}^{\frac{l}{k}-1}\alpha^{p^{ki}} = c\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(\alpha)$, thus $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}$ is $\mathbb{F}_{p^k}$-linear. By $\mathbb{F}_{p^k}$-linearity, surjectivity follows if $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}$ is not identically zero. On the other hand, there exists some element $\alpha \in \mathbb{F}_{p^l}$, whose minimal polynomial $f_\alpha(x)$ over $\mathbb{F}_{p^k}$ has degree $[\mathbb{F}_{p^l} : \mathbb{F}_{p^k}] = l/k$. Since $\deg(\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}) = l/k - 1$, it must be $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(\alpha) \neq 0$, so surjectivity follows. □

Proof of Lemma 8

Proof. Fix $g = (\varepsilon_1, \varepsilon_2) \in \mathcal{G}$. Suppose $d, d' \in D$ satisfy $d - d' = g$. Since $\varepsilon_1 = 0$ implies that d and d' are the same, WLOG $\varepsilon_1 \neq 0$. Write $d = (x, \mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(x^2))$. Then $d' = (x - \varepsilon_1, \mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(x^2)) - \varepsilon_2)$. It follows that $\varepsilon_2 = \mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(x^2)) - \mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}((x-\varepsilon_2)^2))$. $\mathbb{F}_{p^k}$-linearity then implies $\mathrm{Tr}_{\mathbb{F}_{p^l}/\mathbb{F}_{p^k}}(\varepsilon_1(2x - \varepsilon_1)) = \varepsilon_2$. Since $\varepsilon_1 \neq 0$, Corollary 7 implies there are exactly $p^l/p^k = p^{l-k}$ solutions for x, whence p^{l-k} possibilities for $(d, d') : d - d' = g$. Since $|D| = p^l$, it follows that D is a $(p^{l+k}, p^l, \lambda = p^{l-k})$-external difference set. □