

Combinatorial Codes for Detection of Algebraic Manipulation

Serge Fehr

CWI Amsterdam The Netherlands

Joint work with

Ronald Cramer (CWI/Leiden University) Carles Padró (UPC Barcelona)









$\mathcal{G} \ni x \quad \leadsto$



























(CWI) Centrum voor Wiskunde en Informatica



One-time-pad encryption:

- $c = x \oplus k$ perfectly hides x
- $\tilde{c} = c \oplus \delta$ decrypts to $x \oplus \delta$.

Linear secret-sharing scheme:

- ullet shares of non-qualified players perfectly hide secret x
- incorrect shares enforce reconstruction of $\tilde{x} \neq x$, where (due to linearity) adversary knows/controls $\delta = \tilde{x} - x$: apply reconstruction to the differences of all the shares

Algebraic-Manipulation Detection



Algebraic-manipulation detection (AMD) code:





Algebraic-manipulation detection (AMD) code:





Algebraic-manipulation detection (AMD) code:



Important: We want this to work without secret key !

Note: Only care about detection (correction is impossible).

Application



Recall: In a secret sharing scheme

- \bullet too few shares \leadsto no info on the shared secret
- sufficiently many shares \rightsquigarrow secret can be reconstructed,

Application



Recall: In a secret sharing scheme

- too few shares \rightsquigarrow no info on the shared secret
- sufficiently many shares ~> secret can be reconstructed, but: incorrect shares ~> incorrect secret is reconstructed



Recall: In a secret sharing scheme

- too few shares \rightsquigarrow no info on the shared secret
- sufficiently many shares ~> secret can be reconstructed, but: incorrect shares ~> incorrect secret is reconstructed

Whereas in a robust secret sharing scheme

• fraud in the reconstruction phase is detected



Recall: In a secret sharing scheme

- too few shares \rightsquigarrow no info on the shared secret
- sufficiently many shares ~> secret can be reconstructed, but: incorrect shares ~> incorrect secret is reconstructed

Whereas in a robust secret sharing scheme

• fraud in the reconstruction phase is detected

An AMD code allows to make any linear secret sharing scheme robust: share x = E(s), and decode the reconstructed value \tilde{x} .

Note: Also works for dishonest majority (where VSS techniques fail).

Robust secret sharing has applications to e.g. secure and private storage and secure message transmission.



Definition An (m, n)-AMD code is given by

- probabilistic encoding map $E : S \to G$ (where S = set of cardinality m and G = group of order n)
- deterministic decoding function $D: \mathcal{G} \to \mathcal{S} \cup \{\bot\}$

such that D(E(s)) = s with probability 1 for any $s \in S$.



Definition An (m, n)-AMD code is given by

- probabilistic encoding map $E : S \to G$ (where S = set of cardinality m and G = group of order n)
- deterministic decoding function $D: \mathcal{G} \to \mathcal{S} \cup \{\bot\}$

such that D(E(s)) = s with probability 1 for any $s \in S$.

Definition An AMD code is weakly ε -secure if for any $\delta \in G$ and for random $s \in S$: $Prob\left[D(E(s) + \delta) \notin \{s, \bot\}\right] \leq \varepsilon$.

Definition An AMD code is *strongly* ε -secure if for any $\delta \in G$ and for **any** $s \in S$: $Prob \left[D(E(s) + \delta) \notin \{s, \bot\} \right] \leq \varepsilon$.

Centrum voor Wiskunde en Informatica



Cabello, Padró and Sáez, 2002:

Let \mathbb{F} be a finite field of odd order q. Then

is a weakly 1/q-secure (q, q^2) -AMD code.



Cabello, Padró and Sáez, 2002:

Let \mathbb{F} be a finite field of odd order q. Then

is a weakly 1/q-secure (q, q^2) -AMD code.

Proof. To be accepted, $0 = \tilde{s}^2 - \tilde{p} = (s + \partial s)^2 - (s^2 + \partial p) = 2s\partial s + \partial s^2 - \partial p$ and thus $s = (\partial p - \partial s^2)/2\partial s$, which happens with probability $1/|\mathbb{F}|$. \Box



Cabello, Padró and Sáez, 2002:

Let \mathbb{F} be a finite field of odd order q. Then

is a weakly 1/q-secure (q, q^2) -AMD code.

Proof. To be accepted, $0 = \tilde{s}^2 - \tilde{p} = (s + \partial s)^2 - (s^2 + \partial p) = 2s\partial s + \partial s^2 - \partial p$ and thus $s = (\partial p - \partial s^2)/2\partial s$, which happens with probability $1/|\mathbb{F}|$. \Box

Similarly

Centrum voor Wiskunde en Informatica





Theorem Let q and $p = q^2 + q + 1$ both be primes. Then, there exists a 1-difference-set $V \subset \mathbb{Z}_p$ of size q + 1.



Theorem Let q and $p = q^2 + q + 1$ both be primes. Then, there exists a 1-difference-set $V \subset \mathbb{Z}_p$ of size q + 1.

Ogata and Kurosawa, 1996: Gives rise to a weakly 1/(q+1)-secure (q+1, p)-AMD code:

E:	V	\rightarrow	${\cal G}$	$D: \mathcal{G}$	\rightarrow	$V \cup \{\bot\}$
	s	\mapsto	s	\widetilde{s}	\mapsto	\tilde{s} if $\tilde{s} \in V$, else \perp



Theorem Let q and $p = q^2 + q + 1$ both be primes. Then, there exists a 1-difference-set $V \subset \mathbb{Z}_p$ of size q + 1.

Ogata and Kurosawa, 1996:

Gives rise to a weakly 1/(q+1)-secure (q+1, p)-AMD code:

E:	V	\rightarrow	${\cal G}$	$D: \mathcal{G} \rightarrow V \cup \{\bot\}$
	s	\mapsto	s	$ ilde{s} \hspace{0.2cm} \mapsto \hspace{0.2cm} ilde{s} \hspace{0.2cm} ext{if} \hspace{0.1cm} ilde{s} \in V, \hspace{0.1cm} ext{else} \hspace{0.1cm} oldsymbol{oldsymbol{eta}}$

Proof. For any $\delta \in \mathcal{G}$, there is only one $s \in V$ with $s + \delta \in V$. \Box



- Introduction, definition, examples etc.
- The combinatorics of AMD codes
- Measuring the quality of AMD codes, and lower bounds
- A simple construction based on authentication codes
- A construction based on error correcting codes
- A near-optimal polynomial-based construction



Definition $V \subset \mathcal{G}$ is a *t*-bounded-difference-set, if $\forall 0 \neq \delta \in \mathcal{G}$: $\delta = v - w$ for **at most** *t* pairs $v, w \in V$.



Definition $V \subset \mathcal{G}$ is a *t*-bounded-difference-set, if $\forall 0 \neq \delta \in \mathcal{G}$: $\delta = v - w$ for **at most** *t* pairs $v, w \in V$.

Theorem If $V \subset \mathcal{G}$ is a *t*-bounded-difference-set, then

 $E: V \ni s \mapsto s \in \mathcal{G}$

is a (deterministic) weakly t/|V|-secure AMD code.



Definition $V \subset \mathcal{G}$ is a *t*-bounded-difference-set, if $\forall 0 \neq \delta \in \mathcal{G}$: $\delta = v - w$ for **at most** *t* pairs $v, w \in V$.

Theorem If $V \subset \mathcal{G}$ is a *t*-bounded-difference-set, then

 $E: V \ni s \mapsto s \in \mathcal{G}$

is a (deterministic) weakly t/|V|-secure AMD code. And, if $E : S \to G$ is a **deterministic** weakly ε -secure AMD code, then $V = E(S) \subset G$ is a $\varepsilon |V|$ -bounded-difference-set.

Combinatorics of Strongly Secure Codes



Definition $V_1, \ldots, V_m \subset \mathcal{G}$ is a $(t_1...t_m)$ -differential-structure, if

- V_i 's are non-empty and disjoint, and
- $\forall i \in \{1, \ldots, m\}, 0 \neq \delta \in \mathcal{G}: \left| (V_i + \delta) \cap \bigcup_{i \neq i} V_j \right| \leq t_i.$

Combinatorics of Strongly Secure Codes



Definition $V_1, \ldots, V_m \subset \mathcal{G}$ is a $(t_1...t_m)$ -differential-structure, if

- V_i 's are non-empty and disjoint, and
- $\forall i \in \{1, \dots, m\}, 0 \neq \delta \in \mathcal{G}: \left| (V_i + \delta) \cap \bigcup_{j \neq i} V_j \right| \leq t_i.$

Theorem If $V_1, ..., V_m \subset \mathcal{G}$ is a $(t_1...t_m)$ -differential-structure, then

 $E: \{1, \dots, m\} \ni s \stackrel{\hat{s} \cdot V_s}{\longmapsto} \hat{s} \in \mathcal{G}$

is a strongly ε -secure AMD code with $\varepsilon = \max_i t_i / |V_i|$.

Combinatorics of Strongly Secure Codes



Definition $V_1, \ldots, V_m \subset \mathcal{G}$ is a $(t_1...t_m)$ -differential-structure, if

- V_i 's are non-empty and disjoint, and
- $\forall i \in \{1, \ldots, m\}, 0 \neq \delta \in \mathcal{G}: \left| (V_i + \delta) \cap \bigcup_{i \neq i} V_j \right| \leq t_i.$

Theorem If $V_1, ..., V_m \subset \mathcal{G}$ is a $(t_1...t_m)$ -differential-structure, then

$$E: \{1, \ldots, m\} \ni s \stackrel{\hat{s} \cdot V_s}{\longmapsto} \hat{s} \in \mathcal{G}$$

is a strongly ε -secure AMD code with $\varepsilon = \max_i t_i / |V_i|$.

And, any strongly-secure AMD code **with uniform selection** implies a corresponding differential-structure.

Definition An AMD code is *with uniform selection* if for any $s \in S$, the encoding E(s) is random over its possible values.

Centrum voor Wiskunde en Informatica



- Introduction, definition, examples etc.
- The combinatorics of AMD codes
- Measuring the quality of AMD codes, and lower bounds
- A simple construction based on authentication codes
- A construction based on error correcting codes
- A near-optimal polynomial-based construction



Let $E: S \to G$ be (m, n)-AMD code. Recall: m = |S|, n = |G|. Clearly, $n \ge m$. We want n to be as close to m as possible.

Could measure this by the rate $\rho := \log(m) / \log(n)$.

More handy:

Definition The *tag size* is $\varpi := \log(n) - \log(m)$. (The number of bits added to the source *s*.)



Theorem The tag size of a weakly/strongly $2^{-\kappa}$ -secure (m, n)-AMD code is bounded by

$$arpi \geq \kappa - 2/m$$
 resp. $arpi \geq 2\kappa - 2/m$





Theorem The tag size of a weakly/strongly $2^{-\kappa}$ -secure (m, n)-AMD code is bounded by

 $\varpi \geq \kappa - 2/m$ resp. $\varpi \geq 2\kappa - 2/m$

Proof. Choose $\delta \in \mathcal{G}$ at random. For a random $s \in S$:

$$2^{-\kappa} \ge Prob\left[E(s) + \delta \in \bigcup_{s' \neq s} D^{-1}(s')\right] = \frac{\left|\bigcup_{s' \neq s} D^{-1}(s')\right|}{n} \ge \frac{m-1}{n}$$



Theorem The tag size of a weakly/strongly $2^{-\kappa}$ -secure (m, n)-AMD code is bounded by

 $\varpi \geq \kappa - 2/m$ resp. $\varpi \geq 2\kappa - 2/m$

Proof. Choose $\delta \in \mathcal{G}$ at random. For a random $s \in S$:

$$2^{-\kappa} \ge Prob\left[E(s) + \delta \in \bigcup_{s' \neq s} D^{-1}(s')\right] = \frac{\left|\bigcup_{s' \neq s} D^{-1}(s')\right|}{n} \ge \frac{m-1}{n}$$

And thus

$$\varpi = \log(n) - \log(m) = \log \frac{n}{m-1} \frac{m-1}{m} = \log \underbrace{\frac{n}{m-1}}_{\geq 2^{\kappa}} + \log(1 - \frac{1}{m}) \ge \kappa - \frac{2}{m}.$$

Similarly for strongly secure AMD code.

🛚 Centrum voor Wiskunde en Informatica



Cabello-Padró-Sáez constructions:

The weakly/strongly 1/q-secure AMD codes

 $E: s \mapsto (s, s^2)$ resp. $E: s \stackrel{r \cdot \mathbb{F}}{\longrightarrow} (s, r, s \cdot r)$ have tag size $\log(q)$ resp. $2\log(q)$.



Cabello-Padró-Sáez constructions:

The weakly/strongly 1/q-secure AMD codes

 $E: s \mapsto (s, s^2) \quad \text{resp.} \quad E: s \stackrel{r \cdot \mathbb{F}}{\longmapsto} (s, r, s \cdot r)$ have tag size $\log(q)$ resp. $2\log(q)$.

Ogata-Kurosawa construction:

The weakly 1/(q+1)-secure AMD code

 $E: V \ni s \mapsto s \in \mathcal{G} = \mathbb{Z}_p$

where |V| = q + 1 with $p = q^2 + q + 1$, has tag size

 $\log(p) - \log(q+1) = \log(q^2 + q + 1) - \log(q+1) = \log(q + \frac{1}{q+1}).$





All three constructions are (essentially) optimal...

CWI) Centrum voor Wiskunde en Informatica



All three constructions are (essentially) optimal...

... but they are all **not scalable**: $|\mathcal{S}| = 1/\varepsilon = 1/q$.

May want to choose $|S| = 2^{\ell}$ and $\varepsilon = 2^{-\kappa}$ independently.

Example: $\kappa = 128$ and $\ell = 1$ MB. Then lower bound dictates $\varpi \ge 128$ for weak security whereas example codes have $\varpi = 8 \cdot 2^{20} = 8 \cdot 388 \cdot 608$.



All three constructions are (essentially) optimal...

... but they are all **not scalable**: $|S| = 1/\varepsilon = 1/q$.

May want to choose $|\mathcal{S}| = 2^{\ell}$ and $\varepsilon = 2^{-\kappa}$ independently.

Example: $\kappa = 128$ and $\ell = 1$ MB. Then lower bound dictates $\varpi \ge 128$ for weak security whereas example codes have $\varpi = 8 \cdot 2^{20} = 8388608$.

Definition The *effective tag size* of a family of weakly/strongly secure AMD codes, with respect to κ and ℓ , is

 $\varpi^*(\kappa,\ell) := \min\{n\} - \ell$

where the min is over all weakly/strongly ε -secure (m, n)-AMD codes with $\varepsilon \leq 2^{-\kappa}$ and $m \geq 2^{\ell}$.

Known AMD codes are optimal (wrt. to effective tag size) only for $\ell \approx \kappa$.



- Introduction, definition, examples etc.
- The combinatorics of AMD codes
- / Measuring the quality of AMD codes, and lower bounds
- A simple construction based on authentication codes
- A construction based on error correcting codes
- A near-optimal polynomial-based construction



Theorem Let

- $A:\mathcal{K} imes\mathcal{S} o\mathcal{T}$ an A-code with substitution probability p_S
- $E': \mathcal{K} \to \mathcal{G}'$ be a **weakly** ε' -secure AMD code.

Then

$$E: \mathcal{S} \longrightarrow \mathcal{S} \times \mathcal{G}' \times \mathcal{T}, \ s \stackrel{k \cdot \mathcal{K}}{\longmapsto} (s, E'(k), A(k, s))$$

is a strongly ε -secure AMD-code with $\varepsilon = \varepsilon' + p_S$.



Theorem Let

- $A:\mathcal{K} imes\mathcal{S} o\mathcal{T}$ an A-code with substitution probability p_S
- $E': \mathcal{K} \to \mathcal{G}'$ be a **weakly** ε' -secure AMD code.

Then

$$E: \mathcal{S} \longrightarrow \mathcal{S} \times \mathcal{G}' \times \mathcal{T}, \ s \stackrel{k \cdot \mathcal{K}}{\longmapsto} \left(s, E'(k), A(k, s)\right)$$

is a strongly ε -secure AMD-code with $\varepsilon = \varepsilon' + p_S$.

Using "good" A-codes: strongly secure AMD-codes with effective tag size $\varpi^*(\kappa, \ell) \approx 4\kappa \ \forall \kappa, \ell$.

Proposition For any such A-code based AMD code: $\varpi^*(\kappa, \ell) \gtrsim 4\kappa$.

Recall: Lower bound would allow $\varpi^*(\kappa, \ell) \approx 2\kappa$.

Centrum voor Wiskunde en Informatica



- Introduction, definition, examples etc.
- The combinatorics of AMD codes
- / Measuring the quality of AMD codes, and lower bounds
- A simple construction based on authentication codes
- A construction based on error correcting codes
- A near-optimal polynomial-based construction



Let $\mathcal{C} \subset \mathbb{F}^n$ be a (not necessarily linear) error correcting code, and $\mathbb{C} : \mathbb{F}^k \to \mathcal{C}$ the encoding function.

Then

$$E: \mathbb{F}^k \to \mathbb{F}^k \times \mathbb{Z}_n \times \mathbb{F}, \ \mathbf{s} \xrightarrow{x \in \mathbb{Z}_n} \left(\mathbf{s}, x, [\mathbf{C}(\mathbf{s})]_x \right)$$

is a strongly ε -secure AMD code with ε as follows:

- Extend \mathcal{C} to multiset $\operatorname{cl}(\mathcal{C}) = \{(c_t, c_{t+1}, ..., c_{t-1}) \mid \mathbf{c} \in \mathcal{C}, t \in \mathbb{Z}_n\}.$
- If $cl(\mathcal{C})$ contains doubles, then $\varepsilon = 1$ (i.e., no security).
- Else, let M be the max number of occurences of an entry within $\mathbf{c} \mathbf{c}'$, quantified over $\mathbf{c} \neq \mathbf{c}' \in \mathrm{cl}(\mathcal{C})$. Then $\varepsilon = M/n$.



- Introduction, definition, examples etc.
- The combinatorics of AMD codes
- / Measuring the quality of AMD codes, and lower bounds
- A simple construction based on authentication codes
- A construction based on error correcting codes
- A near-optimal polynomial-based construction

A Poly-Based Construction - First Try



Consider the code

 $\mathcal{C} = \{ \mathbf{c}_{f(X)} \,|\, f(X) \in \mathbb{F}[X] \text{ with } \deg f(X) \le k \}$

where

 $\mathbf{c}_{f(X)} = \left(f(x)\right)_{x \in \mathbb{F}}$

Obviously

$$\operatorname{rot}_t(c_{f(X)}) = c_{f(X+t)} \in \mathcal{C}$$

and thus gives **no** (good) AMD code.



 $f_{\mathbf{s}}(X) = 0 + s_1 X + \dots + s_d X^d + 0 \cdot X^{d+1} + X^{d+2} \in \mathbb{F}[X].$

Theorem The resulting AMD code

 $E: \mathbb{F}^d \to \mathbb{F}^d \times \mathbb{F} \times \mathbb{F}, \ \mathbf{s} \stackrel{x \cdot \mathbb{F}}{\longmapsto} (\mathbf{s}, x, f_{\mathbf{s}}(x) = s_1 x + \dots + s_d x^d + x^{d+2})$ is strongly ε -secure with $\varepsilon = (d+1)/q$.

Proof.



 $f_{\mathbf{s}}(X) = 0 + s_1 X + \dots + s_d X^d + 0 \cdot X^{d+1} + X^{d+2} \in \mathbb{F}[X].$

Theorem The resulting AMD code

 $E: \mathbb{F}^d \to \mathbb{F}^d \times \mathbb{F} \times \mathbb{F}, \ \mathbf{s} \stackrel{x \cdot \mathbb{F}}{\longmapsto} (\mathbf{s}, x, f_{\mathbf{s}}(x) = s_1 x + \dots + s_d x^d + x^{d+2})$ is strongly ε -secure with $\varepsilon = (d+1)/q$.

Proof. • Attacker may transform $(\mathbf{s}, x, f_{\mathbf{s}}(x))$ to $(\mathbf{s}', x + \partial x, f_{\mathbf{s}}(x) + \partial e)$.



 $f_{\mathbf{s}}(X) = 0 + s_1 X + \dots + s_d X^d + 0 \cdot X^{d+1} + X^{d+2} \in \mathbb{F}[X].$

Theorem The resulting AMD code

 $E: \mathbb{F}^d \to \mathbb{F}^d \times \mathbb{F} \times \mathbb{F}, \ \mathbf{s} \stackrel{x \cdot \mathbb{F}}{\longmapsto} (\mathbf{s}, x, f_{\mathbf{s}}(x) = s_1 x + \dots + s_d x^d + x^{d+2})$ is strongly ε -secure with $\varepsilon = (d+1)/q$.

Proof. • Attacker may transform $(\mathbf{s}, x, f_{\mathbf{s}}(x))$ to $(\mathbf{s}', x + \partial x, f_{\mathbf{s}}(x) + \partial e)$.

• Define $g(X) := f_{\mathbf{s}'}(X + \partial x) - f_{\mathbf{s}}(X) - \partial e$.



 $f_{\mathbf{s}}(X) = 0 + s_1 X + \dots + s_d X^d + 0 \cdot X^{d+1} + X^{d+2} \in \mathbb{F}[X].$

Theorem The resulting AMD code

 $E: \mathbb{F}^d \to \mathbb{F}^d \times \mathbb{F} \times \mathbb{F}, \ \mathbf{s} \xrightarrow{x \cdot \mathbb{F}} (\mathbf{s}, x, f_{\mathbf{s}}(x) = s_1 x + \dots + s_d x^d + x^{d+2})$ is strongly ε -secure with $\varepsilon = (d+1)/q$.

Proof. • Attacker may transform $(\mathbf{s}, x, f_{\mathbf{s}}(x))$ to $(\mathbf{s}', x + \partial x, f_{\mathbf{s}}(x) + \partial e)$.

- Define $g(X) := f_{\mathbf{s}'}(X + \partial x) f_{\mathbf{s}}(X) \partial e$.
- Gets decoded to $s' \neq s$ if $f_{\mathbf{s}'}(x + \partial x) = f_{\mathbf{s}}(x) + \partial e$, i.e., if g(x) = 0.



 $f_{\mathbf{s}}(X) = 0 + s_1 X + \dots + s_d X^d + 0 \cdot X^{d+1} + X^{d+2} \in \mathbb{F}[X].$

Theorem The resulting AMD code

 $E: \mathbb{F}^d \to \mathbb{F}^d \times \mathbb{F} \times \mathbb{F}, \ \mathbf{s} \stackrel{x \cdot \mathbb{F}}{\longrightarrow} (\mathbf{s}, x, f_{\mathbf{s}}(x) = s_1 x + \dots + s_d x^d + x^{d+2})$ is strongly ε -secure with $\varepsilon = (d+1)/q$.

Proof. • Attacker may transform $(\mathbf{s}, x, f_{\mathbf{s}}(x))$ to $(\mathbf{s}', x + \partial x, f_{\mathbf{s}}(x) + \partial e)$.

- Define $g(X) := f_{\mathbf{s}'}(X + \partial x) f_{\mathbf{s}}(X) \partial e$.
- Gets decoded to $s' \neq s$ if $f_{\mathbf{s}'}(x + \partial x) = f_{\mathbf{s}}(x) + \partial e$, i.e., if g(x) = 0.
- Easy to see: $1 \leq \deg g(X) \leq d$. Thus, $\varepsilon \leq (d+1)/q$.



 $f_{\mathbf{s}}(X) = 0 + s_1 X + \dots + s_d X^d + 0 \cdot X^{d+1} + X^{d+2} \in \mathbb{F}[X].$

Theorem The resulting AMD code

 $E: \mathbb{F}^d \to \mathbb{F}^d \times \mathbb{F} \times \mathbb{F}, \ \mathbf{s} \stackrel{x \cdot \mathbb{F}}{\longmapsto} (\mathbf{s}, x, f_{\mathbf{s}}(x) = s_1 x + \dots + s_d x^d + x^{d+2})$ is strongly ε -secure with $\varepsilon = (d+1)/q$.

Corollary The effective tag size of this AMD code family ranges within $2\kappa + 2\log(\ell) \lesssim \varpi^*(\kappa, \ell) \leq 3\kappa + 3\log(\ell)$

with \approx on the LHS (i.e. near-optimality) if $\ell \approx d(\kappa + \log(d))$ for a $d \in \mathbb{N}$.



- Notion of algebraic-manipulation and of AMD codes
- Combinatorial understanding of AMD codes
- Lower bounds
- Constructions: based on A-codes
 - based on error-correcting codes
 - based on polynomials

Open problem:

Even better constructions

(based on algebraic curves, or on results from combinatorics?)



"Thank you for your attention !!!"

