Optimal algebraic manipulation detection codes and difference families

Douglas R. Stinson

David R. Cheriton School of Computer Science, University of Waterloo

Southeastern Conference, Boca Raton, March 9, 2017

This talk is based on joint work with Bill Martin and Maura Paterson

Algebraic Manipulation Detection Codes

Algebraic Manipulation Detection Codes (AMD codes) were first defined in

Cramer, Dodis, Fehr, Padró, Wichs. *Detection of algebraic manipulation with applications to robust secret sharing and fuzzy extractors.* Eurocrypt 2008.

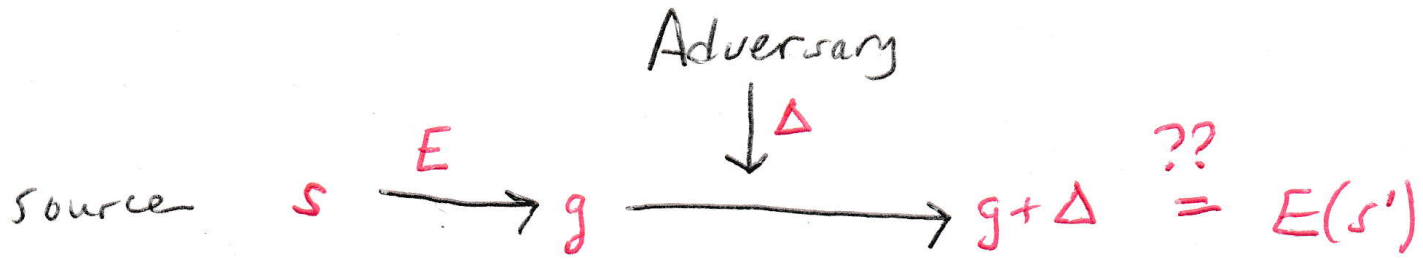
The purpose of this talk is to explore links between AMD codes and certain kinds of difference families, as discussed in

Maura B. Paterson, Douglas R. Stinson. *Combinatorial Characterizations of Algebraic Manipulation Detection Codes Involving Generalized Difference Families*. *Discrete Mathematics* **339** (2016), 2891–2906.

In particular, we present some recent results on the new problem of strong external difference families.

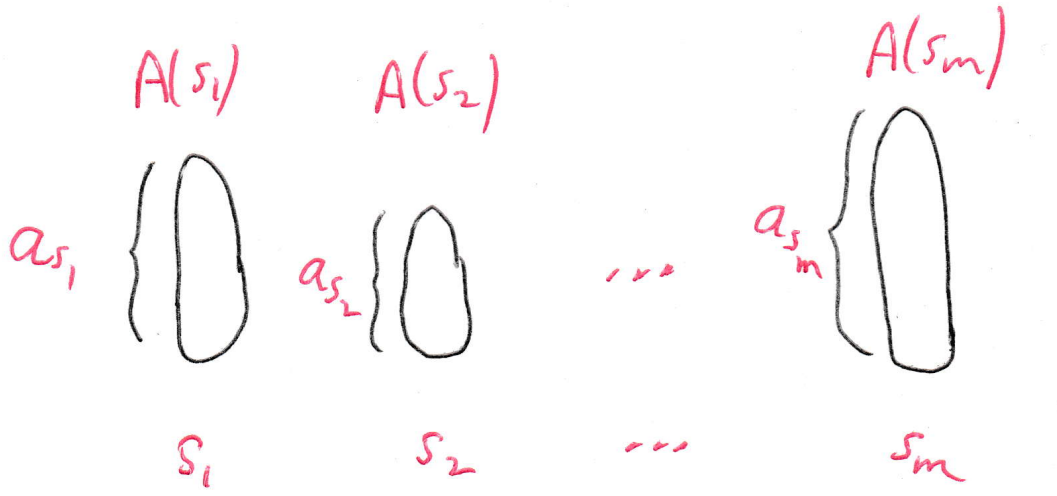
Notation

- Source space $\mathcal{S}$ with $|\mathcal{S}| = m$.
- Encoded message space G (additive Abelian group of order n).
- Encoding rule E (possibly randomized) maps a source $s \in \mathcal{S}$ to some $g \in G$.
- The adversary chooses $\Delta \in G \setminus \{0\}$.
- Then the source s is chosen uniformly at random by the encoder.
- s is encoded into some g by E.
- The adversary wins if and only if $g + \Delta$ is an encoding of s' for some $s' \neq s$.



Notation (cont.)

- For $s \in \mathcal{S}$, we call $A(s) \subseteq G$ the set of valid encodings of s.
- Note that $A(s) \cap A(s') = \emptyset$ if $s \neq s'$.
- Then $\mathcal{A} = \{A(s) : s \in \mathcal{S}\}$ is the set of all valid encodings.
- Let $a_s = |A(s)|$.
- Then $a = \sum_{s \in \mathcal{S}} a_s$ is the total number of valid encodings.
- k-uniform: $a_s = k$ for all $s \in \mathcal{S}$
- k-regular: k-uniform, and in addition the encoding of s is chosen uniformly from the elements of $A(s)$
- 1-regular is the same as deterministic encoding



Weak AMD codes

Definition 1

Suppose $|\mathcal{S}| = m$, $|G| = n$, $\mathcal{A} \subseteq G$ and $E : \mathcal{S} \to A$. Then $(\mathcal{S}, G, \mathcal{A}, E)$ is a weak (m, n, ϵ)-AMD code if the adversary wins the following game with probability at most ϵ:

1. The adversary chooses $\Delta \in G \setminus \{0\}$.
2. The source s is chosen uniformly at random by the encoder.
3. s is encoded into $g \in A(s)$ by the function E.
4. The adversary wins if and only if $g + \Delta \in A(s')$ for some $s' \neq s$.

An Example of a Weak AMD Code

Example 2

$\mathcal{S} = \{s_1, s_2, s_3\}$, $G = \mathbb{Z}_7$.
$A(s_1) = \{0\}$, $A(s_2) = \{1\}$, $A(s_3) = \{3\}$.

This is a deterministic (i.e., 1-regular) weak $\left(3, 7, \frac{1}{3}\right)$-AMD code.

Choose any $\Delta \in \mathbb{Z}_7$, $\Delta \neq 0$. The substitution $g \mapsto g + \Delta$ will yield a valid encoding for exactly one of the three possible valid encodings g.

$\Delta = 1$ works iff $g = 0$, $\Delta = 2$ works iff $g = 1$, $\Delta = 3$ works iff $g = 0$, etc.

Another Example of a Weak AMD Code

Example 3

$\mathcal{S} = \{s_1, s_2\}$, $G = \mathbb{Z}_9$.
$A(s_1) = \{0, 1\}$, $A(s_2) = \{2, 4\}$.
For a given source s, choose the encoding $E(s)$ uniformly at random from $A(s)$.

This is a 2-regular weak $\left(2, 9, \frac{1}{4}\right)$-AMD code.

Choose any $\Delta \in \mathbb{Z}_9$, $\Delta \neq 0$. The substitution $g \mapsto g + \Delta$ will yield a valid encoding for exactly one of the four possible valid encodings g.

$\Delta = 1$ works iff $g = 1$, $\Delta = 2$ works iff $g = 0$, $\Delta = 3$ works iff $g = 1$, etc.

The Random Strategy for Weak AMD Codes

Suppose Δ is chosen uniformly at random from $G \setminus \{0\}$. For any $g \in A(s)$ and for a random Δ, the probability that the adversary wins is $(a - a_s)/(n-1)$. The overall success probability is

$$\begin{aligned}
&\sum_s \mathbf{Pr}[s] \sum_{g \in A(s)} \left(\mathbf{Pr}[E(s) = g] \times \frac{a - a_s}{n-1} \right) \\
&= \sum_s \left(\mathbf{Pr}[s] \times \frac{a - a_s}{n-1} \right) \\
&= \frac{a}{n-1} - \sum_s \frac{a_s}{m(n-1)} \quad (\mathbf{Pr}[s] = 1/m \text{ for all } s) \\
&= \frac{a}{n-1} - \frac{a}{m(n-1)} \\
&= \frac{a(m-1)}{m(n-1)}.
\end{aligned}$$

A weak AMD code is R-optimal if the random strategy is optimal.

External Difference Families

Ogata, Kurosawa, Stinson, Saido. *New combinatorial designs and their applications to authentication codes and secret sharing schemes,* Discrete Mathematics, 2004.

Definition 4 ((n, m, k, λ)-EDF)

An (n, m, k, λ)-external difference family is a set $A_1, A_2 \ldots, A_m$ of m disjoint k-subsets of additive abelian group G (with $|G| = n$) such that each nonzero element of G occurs λ times as a difference between elements in different subsets.

Necessary condition: $\lambda(n-1) = k^2 m(m-1)$.

An $(n, m, 1, \lambda)$-EDF is just an (n, m, λ)-difference set.

Levenshtein defined a closely related concept, difference systems of sets, in 1971.

Examples of External Difference Families

Example 5

- $(7, 3, 1, 1)$-EDF: $\{0\}, \{1\}, \{3\} \subseteq \mathbb{Z}_7$
- $(9, 2, 2, 1)$-EDF: $\{0, 1\}, \{2, 4\} \subseteq \mathbb{Z}_9$
- $(19, 3, 3, 3)$-EDF: $\{1, 7, 11\}, \{4, 9, 6\}, \{16, 17, 5\} \subseteq \mathbb{Z}_{19}$

Theorem 6 (Tonchev, 2003)

Suppose that $q = 2u\ell + 1$ is a prime power, where u and ℓ are odd. Then there exists a $(q, u, \ell, (q - 2\ell - 1)/4)$-EDF in $\mathbb{F}_q$.

Proof.

Let $\alpha \in \mathbb{F}_q$ be a primitive element. Let C be the subgroup of ${\mathbb{F}_q}^*$ having order u and index 2ℓ. The ℓ cosets $\alpha^{2i}C$ $(0 \leq i \leq \ell - 1)$ form the EDF. □

R-Optimal k-regular Weak AMD Codes

A weak AMD code is R-optimal if the random strategy is optimal.

A k-regular R-optimal AMD code has $a = km$, so

$$\epsilon = \frac{a(m-1)}{m(n-1)} = \frac{k(m-1)}{(n-1)}.$$

Theorem 7

A k-regular R-optimal weak (m, n, ϵ)-AMD code is equivalent to an (n, m, k, λ)-EDF.

Generalized External Difference Families

Definition 8 ((n, m, λ)-GEDF)

An (n, m, λ)-generalized external difference family is a set $A_1, A_2 \ldots, A_m$ of m disjoint subsets of additive abelian group G (with $|G| = n$) such that each nonzero element of G occurs λ times as a difference between elements in different subsets.

The only difference between a GEDF and an EDF is that the subsets in a GEDF do not have to all be the same size.

Example 9

- $(13, 2, 1)$-GEDF: $\{0, 1\}, \{2, 4, 6\} \subseteq \mathbb{Z}_{13}$
- $(11, 3, 1)$-GEDF: $\{0\}, \{1\}, \{3, 5\} \subseteq \mathbb{Z}_{11}$

GEDFs might not yield R-Optimal Weak AMD Codes

Example 10

$\mathcal{S} = \{s_1, s_2, s_3\}$, $G = \mathbb{Z}_{11}$.
$A(s_1) = \{0\}$, $A(s_2) = \{1\}$, $A(s_3) = \{3, 5\}$.

Using the random strategy, an adversary wins with probability

$$\frac{a(m-1)}{m(n-1)} = \frac{4(3-1)}{3(11-1)} = \frac{4}{15}.$$

However, this is not optimal. For example, the choice of $\Delta = 1$ works whenever $g = 0$, which wins with probability $1/3$.

Partitioned External Difference Families

Wa have seen that EDFs yield R-optimal weak AMD codes, and GEDFs do not necessarily yield R-optimal weak AMD codes.

We define a type of difference family that is "in between" EDFs and GEDFs which still yields R-optimal weak AMD codes.

Definition 11 (Partitioned external difference family)

Let $\mathcal{G}$ be an additive abelian group of order n. An $(n, m; c_1, \ldots, c_\ell; k_1, \ldots, k_\ell; \lambda_1, \ldots, \lambda_\ell)$-partitioned external difference family is a set of $m = \sum_i c_i$ disjoint subsets of $\mathcal{G}$, say $A_1, \ldots, A_m$, such that there are c_h subsets of size k_h, for $1 \leq h \leq \ell$, and the following holds for every h, $1 \leq h \leq \ell$:

$$\bigcup_{\{i:|A_i|=c_h\}} \bigcup_{\{j:j\neq i\}} \{x - y : x \in A_i, y \in A_j\} = \lambda_i(\mathcal{G} \setminus \{0\}).$$

A Construction for PEDFs

Theorem 12

Suppose $A_1, \ldots, A_m$ is a partition of $\mathcal{G}$ (where $|\mathcal{G}| = n$) such that there are c_h subsets of size k_h for $1 \leq h \leq \ell$. Then $A_1, \ldots, A_m$ is an $(n, m; c_1, \ldots, c_\ell; k_1, \ldots, k_\ell; \lambda_1, \ldots, \lambda_\ell)$-PEDF if and only if the subsets of cardinality k_h form an $(n, k_h, c_h k_h - \lambda_h)$-difference family in $\mathcal{G}$, for $1 \leq h \leq \ell$.

Example 13

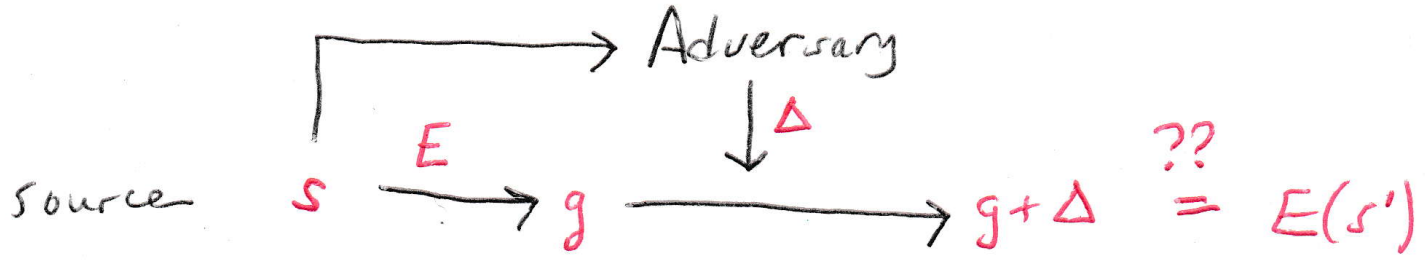
Let $\mathcal{G} = (\mathbb{Z}_{13}, +)$, $A_1 = \{0, 1, 4\}$, $A_2 = \{3, 5, 10\}$, $A_3 = \{2, 6, 7, 9\}$, $A_4 = \{8\}$, $A_5 = \{11\}$, $A_6 = \{12\}$. The two sets of size 3 form a $(13, 2, 3, 1)$-DF; the set of size 4 is a $(13, 1, 4, 1)$-DF; and the three sets of size 1 form a $(13, 3, 1, 0)$-DF. Therefore, $A_1, \ldots, A_6$ is a $(13, 6; 2, 1, 3; 3, 4, 1; 5, 3, 3)$-PEDF.

Strong AMD codes

Definition 14

Suppose $|\mathcal{S}| = m$, $|G| = n$, $\mathcal{A} \subseteq G$ and $E : \mathcal{S} \to A$. Then $(\mathcal{S}, G, \mathcal{A}, E)$ is a strong (m, n, ϵ)-AMD code if the adversary wins the following game with probability at most ϵ (for any choice of $s \in \mathcal{S}$):

1. The source s is given to the adversary.
2. Then the adversary chooses $\Delta \in G \setminus \{0\}$.
3. s is encoded into $g \in A(s)$ by the function E.
4. The adversary wins if and only if $g + \Delta \in A(s')$ for some $s' \neq s$.



The Random Strategy for k-regular Strong AMD Codes

A random strategy for a k-regular strong AMD code has success probability at least $\frac{k(m-1)}{(n-1)}$.

A k-regular strong AMD code is R-optimal if the random strategy is optimal for every possible source s.

Example 15

$\mathcal{S} = \{s_1, s_2\}$, $G = \mathbb{Z}_{10}$.
$A(s_1) = \{0, 1, 2\}$, $A(s_2) = \{3, 6, 9\}$,
For a given source s, choose the encoding $E(s)$ uniformly at random from $A(s)$. This is a 3-regular strong $\left(2, 10, \frac{1}{3}\right)$-AMD code.

Strong External Difference Families

Paterson and Stinson (2016) defined strong EDFs.

Definition 16 ((n, m, k, λ)-SEDF)

An (n, m, k, λ)-strong external difference family is a set $A_1, A_2 \ldots, A_m$ of m disjoint k-subsets of additive abelian group G (with $|G| = n$) such that the following multiset equation holds for every i:

$$\bigcup_{j \neq i} \{x - y : x \in A_i, y \in A_j\} = \lambda(G \setminus \{0\}).$$

Necessary condition: $\lambda(n-1) = k^2(m-1)$.

An (n, m, k, λ)-SEDF is an $(n, m, k, m\lambda)$-EDF.

Theorem 17

A k-regular R-optimal strong (m, n, ϵ)-AMD code is equivalent to an (n, m, k, λ)-SEDF.

SEDFs with $\lambda = 1$

Example 18

Let $\mathcal{G} = (\mathbb{Z}_{k^2+1}, +)$, $A_1 = \{0, 1, \ldots, k-1\}$ and $A_2 = \{k, 2k, \ldots, k^2\}$. This is a $(k^2+1, 2, k, 1)$-SEDF.

Example 19

Let $\mathcal{G} = (\mathbb{Z}_n, +)$ and $A_i = \{i\}$ for $0 \leq i \leq n-1$. This is an $(n, n, 1, 1)$-SEDF.

SEDFs with $\lambda = 1$ (cont.)

Theorem 20 (PS, 2016)

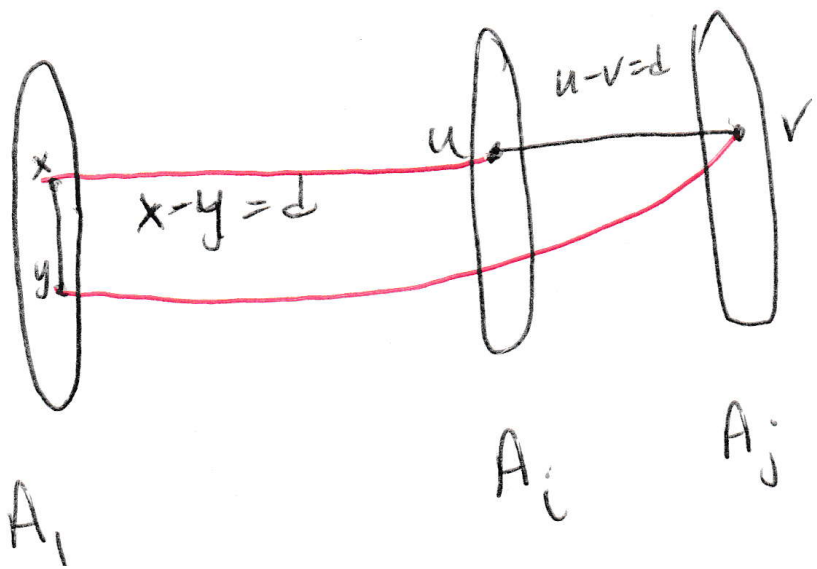
There does not exist an $(n, m, k, 1)$-SEDF with $m \geq 3$ and $k > 1$.

Proof.

Suppose $(A_1, \ldots, A_m)$ is an $(n, m, k, 1)$-SEDF with $m \geq 3$ and $k > 1$. Let $x, y \in A_1$, $x \neq y$. Since $m > 2$, there exists $u \in A_i$, $v \in A_j$ such that $i, j > 1$, $i \neq j$ and $u - v = x - y$. Then $u - x = v - y$. □

$u - v = x - y$

$\Rightarrow u - x = v - y$



SEDFs with $\lambda = 1$ (cont.)

Theorem 21 (PS, 2016)

There exists an $(n, m, k, 1)$-SEDF if and only if $m = 2$ and $n = k^2 + 1$, or $k = 1$ and $m = n$.

Proof.

We only need to consider cases where $m = 2$ or $k = 1$.
Since $\lambda = 1$, we have

$$n - 1 = k^2(m - 1).$$

If $m = 2$, then we must have $n = k^2 + 1$, and the relevant SEDF exists from Example 18.
If $k = 1$, then we must have $m = n$, and the relevant SEDF exists from Example 19. □

Recent Preprints about SEDF on ArXiv

Some Nonexistence Results for Strong External Difference Families Using Character Theory, William J. Martin, Douglas R. Stinson, 20 Oct. 2016

Existence and Non-Existence Results for Strong External Difference Families, Sophie Huczynska, Maura B. Paterson, 17 Nov. 2016

The (n, m, k, λ)-Strong External Difference Family with $m \geq 5$ Exists, Jiejing Wen, Minghui Yang, Keqin Feng, 30 Dec. 2016

New Existence and Nonexistence Results for Strong External Difference Families, Jingjun Bao, Lijun Ji, Ruizhong Wei, Yong Zhang, 30 Dec. 2016

Cyclotomic Construction of Strong External Difference Families in Finite Fields, Jiejing Wen, Minghui Yang, Fangwei Fu, Keqin Feng, 7 Jan. 2017

Construction and nonexistence of strong external difference families, Jonathan Jedwab, Shuxing Li, 28 Jan. 2017

The Group Algebra

Martin and Stinson expressed the SEDF requirements in the group algebra. Let G be a finite abelian group of order v (written multiplicatively) with identity element 1. We seek (pairwise disjoint) subsets $A_1, \ldots, A_m \subseteq G$ with $|A_j| = k$ $(1 \leq j \leq m)$ satisfying

$$\sum_{\ell \neq j} A_j A_\ell^{-1} = \lambda(G - 1) \tag{1}$$

for each $1 \leq j \leq m$. Let $\mathcal{D}$ denote the union of all the sets A_j: in group algebra notation,

$$\mathcal{D} = \sum_{j=1}^{m} A_j.$$

With this, Equation (1) becomes

$$A_j \mathcal{D}^{-1} - A_j A_j^{-1} = \lambda(G - 1). \tag{2}$$

Applying Characters

Lemma 22 (MS, 2016)

Suppose χ is a non-principal character. Then the following holds:

$$\chi(A_j)\overline{\chi(\mathcal{D})} - \chi(A_j)\overline{\chi(A_j)} = -\lambda.$$

Further, if $\chi(\mathcal{D}) \neq 0$, then there exist nonzero real numbers $\alpha_1, \ldots, \alpha_m$ such that $\chi(A_j) = \alpha_j \chi(\mathcal{D})$ for $j = 1, \ldots, m$.

From this lemma, the following theorem can be proven:

Theorem 23 (MS, 2016)

There does not exist a $(v, 3, k, \lambda)$-SEDF for any $v > 3$, and there does not exist a $(v, 4, k, \lambda)$-SEDF for any $v > 4$.

Using additional results about characters, we prove

Theorem 24 (MS, 2016)

If v is prime, $k > 1$ and $m > 2$, then there does not exist a (v, m, k, λ)-SEDF.

Results from Huczynska and Paterson

The following theorem is a generalization of Theorem 20.

Theorem 25 (HP, 2016)

Suppose $\lambda \geq 2$, $m \geq 3$ *and* $k \geq \lambda + 1$. *Then an* (n, m, k, λ)*-SEDF exists only if*

$$\lambda(k-1)(m-2) \leq (\lambda - 1)k(m-1).$$

Theorem 26 (HP, 2016)

For any prime power $q \equiv 1 \bmod 4$, *there exists a* $(q, 2, (q-1)/2, (q-1)/4)$*-SEDF.*

An SEDF with $m > 2$ and $k > 1$

The following is the only known example of an SEDF with $m > 2$ and $k > 1$.

Theorem 27 (WYF 2016,JL 2017)

There exists a $(243, 11, 22, 20)$-SEDF.

Proof.

Let C_0 be the subgroup of $\mathbb{F}_{3^5}^*$ having order 22, and let $C_1, \ldots, C_{10}$ be its cosets. $\{C_0, \ldots, C_{10}\}$ forms the desired SEDF. □

The parameter set $(243, 11, 22, 20)$ is quite special; the following result concerning near-complete SEDF is proven by Jedwab and Li.

Theorem 28 (JL, 2017)

If there exists an (n, m, k, λ)-SEDF with $n = km + 1$, then $(n, m, k, \lambda) = (v, 2, (v-1)/2, (v-1)/4)$ for some $v \equiv 1 \bmod 4$ or $(n, m, k, \lambda) = (243, 11, 22, 20)$.

Nonexistence Results from [BJWZ]

Using character theory, the following results are proven by Bao, Ji, Wei and Zhang.

Theorem 29 (BJWZ, 2016)

If $n = pq$ where p and q are distinct primes, $k > 1$ and $m > 2$, then there does not exist an (n, m, k, λ)-SEDF.

Theorem 30 (BJWZ, 2016)

If $n = p^2$ where p is an odd primes, $k > 1$ and $m > 2$, then there does not exist an (n, m, k, λ)-SEDF in a cyclic group of order p^2.

Theorem 31 (BJWZ, 2016)

If $n = p_1 p_2 \cdots p_s$ where the p_i's are distinct primes, $\gcd(mk, n) - 1$, $k > 1$ and $m > 4$, then there does not exist an (n, m, k, λ)-SEDF.

Existence Results for SEDF with $m = 2$

Theorem 32

The following SEDF with $m = 2$ exist:

1. *a $(k^2 + 1, 2, k, 1)$-SEDF in $\mathbb{Z}_{k^2+1}$ (PS, 2016);*
2. *an $\left(n, 2, \frac{n-1}{2}, \frac{n-1}{2}\right)$-SEDF, whenever an $\left(n, \frac{n-1}{2}, \frac{n-5}{4}, \frac{n-1}{4}\right)$-PDS exists (DHM, 2015) and (HP, 2016);*
3. *a $\left(q, 2, \frac{q-1}{4}, \frac{q-1}{16}\right)$-SEDF in $\mathbb{F}_q$, wherever $q = 16t^2 + 1$ is a prime power (BJWZ, 2016);*
4. *a $\left(q, 2, \frac{q-1}{6}, \frac{q-1}{36}\right)$-SEDF in $\mathbb{F}_q$, wherever $q = 108t^2 + 1$ is a prime power (BJWZ, 2016).*

Definition 33 ((n, k, λ, μ)-PDS)

An (n, k, λ, μ)-partial difference set is a k-subset $D \subseteq G$ (where $|G| = n$) such that every $x \in D \setminus \{0\}$ occurs exactly λ times as a difference of two elements of D and every $x \in G \setminus (D \cup \{0\})$ occurs exactly μ times as a difference of two elements of D.

Nonexistence Results from Jedwab and Li

Jedwab and Li obtain many interesting results using character theory. Here are a couple of examples.

Theorem 34 (JL, 2017)

Let λ be a fixed positive integer. Then for all sufficiently large k, there does not exist a $(n, 5, k, \lambda)$-SEDF and there does not exist a nontrivial $(n, 6, k, \lambda)$-SEDF.

Theorem 35 (JL, 2017)

For $n \leq 10000$ and $m > 2$, there are 70 possible parameter sets for which a non-near-complete (n, m, k, λ)-SEDF with $k > 1$ might exist.

Thank You For Your Attention!