A Compression Function Exploiting Discrete Geometry : Generalization to Higher Dimensions

---

Benjamin Wesolowski

Supervised by
Dimitar Jetchev

Laboratory for Cryptologic Algorithms,
EPFL, Fall 2012

Abstract

Abstract...

Introduction

Introduction...

1 Construction of the Compression Function

1.1 Incidence-Based Preprocessing

(here will come explanations about why we are doing this...) Let K be a field, and n a positive integer.

Definition 1.1 (d-dimentional plane)**.** For any integer $0 \leq d < n$, a subset H of K^n is a *plane of dimension* d if there exist and independent set of vectors $\{v_i\}_{i=1}^d$ such that

$$H = \left\{ \sum_{i=1}^{d} \alpha_i v_i \mid \alpha_1, ..., \alpha_d \in K, \sum_{i=1}^{d} \alpha_i = 1 \right\}.$$

Definition 1.2 (Flag)**.** A *flag* in K^n is a sequence $K^n \supsetneq H_k \supsetneq ... \supsetneq H_1 \supsetneq H_0$ where H_i is a plane for $i = 0, ..., k$.

Definition 1.3 (Maximal flag)**.** A flag $K^n \supsetneq H_k \supsetneq ... \supsetneq H_1 \supsetneq H_0$ is *maximal* if $k = n - 1$. Consequently, H_i is a plane of dimension i for any $i = 0, ..., n - 1$.

Remark 1.4. In the following, we are only interested in planes of the form

$$H = \{(x_1, x_2, ..., x_n) \in K^n \mid x_j = c_{1,j}x_1 + ...c_{d,j}x_d + c_{0,j}, \forall j = d + 1, ..., n\}.$$

for some parameters $c_{i,j} \in K$. Therefore we will simply say *plane* to refer to such planes.

Proposition 1.5. *Let d be a positive integer smaller than n, H a plane of dimention $d - 1$ parameterized by the collection $\{c_{i,j}\}$, and G a plane of dimension d parameterized by $\{b_{i,j}\}$. Then, $H \subset G$ if and only if $c_{k,j} = b_{k,j} + b_{i,j}c_{k,i}$ for $j = d + 1, ..., n$ and $k = 0, ..., d - 1$.*

Proof. We proceed by successive equivalences. One has $H \subset G$ if and only if for any $x = (x_1, ..., x_n) \in H$, $x \in G$. But any point in H (resp. in G) is uniquely determined by its first $d - 1$ (resp. d) coordinates, and all the following ones are given by the formula $x_j = c_{1,j}x_1 + ...c_{d-1,j}x_{d-1} + c_{0,j}$ (resp. $x_j = b_{1,j}x_1 + ...b_{d,j}x_{d-1} + b_{0,j}$). Therefore, $H \subset G$ if and only if $\forall x_1, ..., x_{d-1} \in K$, $\forall j = d + 1, ..., n$, we have

$$\begin{aligned} &b_{1,j}x_1 + ... + b_{d-1,j}x_{d-1} + b_{d,j}(\overbrace{c_{1,d}x_1 + ... + c_{d-1,d}x_{d-1} + c_{0,d}}^{\text{corresponding to } x_d}) + b_{0,j} \\ &= c_{1,j}x_1 + ... + c_{d-1,j}x_{d-1} + c_{0,j}. \end{aligned}$$

This equality is equivalent to

$$(b_{0,j} + b_{d,j}c_{0,d} - c_{0,j}) + \sum_{k=1}^{d-1}(b_{k,j} + b_{d,j}c_{k,d} - c_{k,j})x_k = 0. \quad (1)$$

But for a given j, equality 1 stands for all $x_1, ..., x_{d-1} \in K$ if and only if $b_{k,j} + b_{d,j}c_{k,d} - c_{k,j} = 0$ for all $k = 0, ..., d-1$. Indeed, choosing $x_1 = ... = x_{d-1} = 0$ gives $b_{0,j} + b_{d,j}c_{0,d} - c_{0,j} = 0$, and choosing $x_k = 1$ while all the other x's are set to 0 gives $b_{k,j} + b_{d,j}c_{k,d} - c_{k,j} = 0$.
Finally, we have that $H \subset G$ if and only if equality 1 is satisfied for all $x_1, ..., x_{d-1} \in K$ and $j = d+1, ..., n$, which is true if and only if $c_{k,j} = b_{k,j} + b_{i,j}c_{k,i}$ for all $j = d+1, ..., n$ and $k = 0, ..., d-1$. □

Maximal flags in the 3-dimensional space Let Q be a point, L a line and P a plane in K^3, parameterized by

$$\begin{aligned} P &= \{(x_1, x_2, c_{1,3}x_1 + c_{2,3}x_2 + c_{0,3}) \mid x_1, x_2 \in K\}, \\ D &= \{(x_1, b_{1,2}x_1 + b_{0,2}, b_{1,3}x_1 + b_{0,3}) \mid x_1 \in K\}, \\ Q &= (a_{0,1}, a_{0,2}, a_{0,3}). \end{aligned}$$

Applying proposition 1.5, we obtain that (Q, L, P) is a flag (i.e. $Q \in L \subset P$) if and only if

$$\begin{aligned} b_{1,3} &= c_{1,3} + c_{2,3}b_{1,2}, \\ b_{0,3} &= c_{0,3} + c_{2,3}b_{0,2}, \\ a_{0,2} &= b_{0,2} + b_{1,2}a_{0,1}, \\ a_{0,3} &= b_{0,3} + b_{1,3}a_{0,1}. \end{aligned}$$

Therefore we define, for a block-length of n,

$$\begin{aligned} C_3^{\mathrm{PRE}} : K^6 &\longrightarrow K^3 \\ (a_1, b_1, b_2, c_1, c_2, c_3) &\longmapsto (c_1, c_2, c_3), \\ C_2^{\mathrm{PRE}} : K^6 &\longrightarrow K^4 \\ (a_1, b_1, b_2, c_1, c_2, c_3) &\longmapsto (b_1, b_2, c_1 + c_2b_1, c_3 + c_2b_2), \\ C_1^{\mathrm{PRE}} : K^6 &\longrightarrow K^3 \\ (a_1, b_1, b_2, c_1, c_2, c_3) &\longmapsto (a_1, b_2 + b_1a_1, (c_3 + c_2b_2) + (c_1 + c_2b_1)a_1), \end{aligned}$$

and $C^{\mathrm{PRE}} = C_1^{\mathrm{PRE}} \times C_2^{\mathrm{PRE}} \times C_3^{\mathrm{PRE}}$. A triple in $K^3 \times K^4 \times K^3$ belongs to $C^{\mathrm{PRE}}(K^6)$ if and only if it parameterizes a flag in K^3.

Proposition 1.6 (Completion property)**.** *Let $A = (a_1, a_2, a_3)$ define a point, $B = (b_1, b_2, b_3, b_4)$ a line and $C = (c_1, c_2, c_3)$ a plane in K^3. We have*

1. *if $A \in B$, then A, B and c_2 uniquely determine a flag $A \in B \subset C$;*

2. *if $B \subset C$, then B, C and a_1 uniquely determine a flag $A \in B \subset C$;*

3. *if $A \in C$, then A, C and b_1 uniquely determine a flag $A \in B \subset C$.*

We say that C^{PRE} has the completion property*.*

Proof. The points 2 and 3 are clear from the formulas for C^{PRE}. So consider that $A \in B$, and c_2 are given. We need to find c_1 and c_3 such that $A \in B \subset C$. But we have $b_3 = c_1 + c_2 b_1$ and $b_4 = c_3 + c_2 b_2$, so we can derive $c_1 = b_3 - c_2 b_1$ and $c_3 = b_4 - c_2 b_2$. □

1.2 Construction of a New Compression Function

Construction 1.7 (Non-Linear Matrix-Style Postprocessing)**.** We consider the postprocessing function $C^{\mathrm{POST}} : K^7 \to K^3$ defined by

$$C^{\mathrm{POST}}(a, b_1, b_2, c, x, y, z) = \mathfrak{M} \cdot (a \;\; b_1 \;\; b_2 \;\; c \;\; x \;\; y \;\; z \;\; yz \;\; xz \;\; xy)^t,$$

where $\mathfrak{M}$ is a matrix over K with coefficients $(\omega_{ij})_{i,j}$, and columns $(W_j)_j$ satisfying the following conditions :

C1. $\det(W_5 \;\; W_9 \;\; W_{10}) \neq 0$, $\det(W_6 \;\; W_8 \;\; W_{10}) \neq 0$, $\det(W_7 \;\; W_8 \;\; W_9) \neq 0$,

C2. $W_1 = W_9 + W_{10}$, $W_2 = W_8$, $W_3 = W_{10}$, $W_4 = W_8 + W_9$.

C3. $W_5 = W_8$, $W_6 = W_9$, $W_7 = W_{10}$.

We now set $K = \mathbb{F}_{2^n}$.

Construction 1.8 (3-dimentional Szemerédi-Trotter Compression Function)**.** Let $f_{\mathcal{X}}, f_{\mathcal{Z}} : \mathbb{F}_{2^n}^3 \to \mathbb{F}_{2^n}$ and $f_{\mathcal{Y}} : \mathbb{F}_{2^n}^4 \to \mathbb{F}_{2^n}$ be three distinct and independently sampled PuRFs. We define $H^{f_{\mathcal{X}}, f_{\mathcal{Y}}, f_{\mathcal{Z}}} : \mathbb{F}_{2^n}^6 \to \mathbb{F}_{2^n}^3$ as

$$H^{f_{\mathcal{X}}, f_{\mathcal{Y}}, f_{\mathcal{Z}}}(a, b_1, b_2, c_1, c, c_3) = C^{\mathrm{POST}}(a, b_1, b_2, c, x, y, z),$$

where

$$(x, y, z) = (f_{\mathcal{X}} \times f_{\mathcal{Y}} \times f_{\mathcal{Z}}) \left(C^{\mathrm{PRE}}(a, b_1, b_2, c_1, c, c_3) \right).$$

2 Proof of Collision Resistance

2.1 Strategy

Blabla,

$$\mathrm{coll}(\mathcal{Q}) \to \underbrace{\left(\mathrm{coll}_I(\mathcal{Q}) \wedge \neg \mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})\right)}_{\mathbf{E_1}} \vee \underbrace{\left(\neg \mathrm{coll}_{II}(\mathcal{Q}) \wedge \mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})\right)}_{\mathbf{E_2}} \vee \underbrace{\mathrm{coll}_{II}(\mathcal{Q})}_{\mathbf{E_3}}.$$

Figure 1: The game Exp^{B_Σ}.

Let $i \leftarrow 0$, ctr $\leftarrow 0$;
while i < q **do**
  $i \leftarrow i + 1$;
  $p_i \leftarrow \mathcal{A}(\mathrm{ctr})$;
  if $0 \leq \sum_{j=1}^{i} p_j \leq B_\Sigma$ **then**
    With propability p_i, return `true`;
  end if
end while
return `false`.

Output lines If A is an $\mathcal{X}$-query, and B and C are two preceding queries such that $A \in B \subset C$ with $y = f_{\mathcal{Y}}(B)$ and $z = f_{\mathcal{Z}}(C)$, then the evaluation of the compression function at $(a, b_1, b_2, c, c_2, c_3)$, i.e. $C^{\mathrm{POST}}(a, b_1, b_2, c, x, y, z)$, lies on the line

$$\mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a} = \{P_{\mathcal{X}}(b_1, b_2, c, y, z, a) + xQ_{\mathcal{X}}(y, z) \mid x \in K\}.$$

where

$$P_{\mathcal{X}}(b_1, b_2, c, y, z, a) = aW_1 + b_1W_2 + b_2W_3 + cW_4 + yW_6 + zW_7 + yzW_8,$$
$$Q_{\mathcal{X}}(y, z) = W_5 + zW_9 + yW_{10}.$$

Setting $x = f_{\mathcal{X}}(A)$ leads to the output point, which can therefore be interpreted as a random point on the line $\mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a}$. In order to truly get a random point, the line must not be degenerated ; this is guaranteed by the condition $\det(W_5 \;\; W_9 \;\; W_{10}) \neq 0$ which implies

$$Q_{\mathcal{X}}(y, z) = (W_5 \;\; W_9 \;\; W_{10}) \cdot (1 \;\; z \;\; y)^t \neq 0.$$

We define similarly the output lines for $\mathcal{Y}$-queries and $\mathcal{Z}$-queries, with the non-degeneracy conditions $\det(W_6 \;\; W_8 \;\; W_{10}) \neq 0$ and $\det(W_7 \;\; W_8 \;\; W_9) \neq 0$.

2.2 Bounding $\Pr[\mathbf{E_1}]$

We need the following proposition, proven in [1].

Proposition 2.1. *In the game* Exp^{B_Σ} *given in figure 1, the advantage of any adversary* $\mathcal{A}$ *is at most* B_Σ.

Lemma 2.2. *Let* $\mathcal{Q}$ *be any query list and* κ *any positive integer. Then,*

$$\Pr[\mathbf{E_1}] = \Pr[\mathrm{coll}_I(\mathcal{Q}) \wedge \neg \mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})] \leq \frac{\kappa Y}{2^n}.$$

Proof. First notice that

$$\Pr[\mathrm{coll}_I(\mathcal{Q}) \wedge \neg \mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})] \leq \Pr[\mathrm{coll}_I(\mathcal{Q}) \mid \neg \mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})].$$

For any positive integer $i \leq$, let n_i denote the number of pairs of preceding queries compatible with the i-th query. Every such pair defines a line from which the answer of the i-th query determines a point to be added to the yield set. If we assume $\mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})$, any of those lines cannot contain more than κ previous yield point, so the probability to hit a previous yield point on a given line is at most $\kappa/2^n$. Thus an union bound on the n_i lines bounds by $n_i\kappa/2^n$ the probability for the i-th query to give rise to a collision. We can now apply proposition 2.1 with $B_\Sigma = \kappa Y/2^n$, since $\sum_{i=1}^{3q} n_i \leq Y$. □

2.3 Bounding $\Pr[\mathbf{E_3}]$

Lemma 2.3. *Let $\mathcal{Q}$ be any query list and γ any positive integer. Then,*

$$\Pr[\mathbf{E_3}] = \Pr[\mathrm{coll}_{II}(\mathcal{Q})] \leq 3\frac{q\gamma}{2^n} + 3\left(\frac{q^4}{2^{2n-3}}\right)^\gamma.$$

Proof. We bound the probability of an $\mathcal{X}$-query A causing a collision. Suppose $B \subset C$ and $B' \subset C'$ are two distinct couples of preceding queries such that $A \in B \subset C$ and $A \in B' \subset C'$. Let

$$\begin{aligned} x &= f_{\mathcal{X}}(A), \\ y &= f_{\mathcal{Y}}(B), \\ z &= f_{\mathcal{Y}}(C), \\ y' &= f_{\mathcal{Z}}(B'), \\ z' &= f_{\mathcal{Z}}(C'). \end{aligned}$$

The $\mathcal{X}$-query A causes a collision if and only if

$$\begin{aligned} &(b_1 - b_1')W_2 + (b_2 - b_2')W_3 + (c - c')W_4 \\ &+ (y - y')W_6 + (z - z')W_7 + (yz - y'z')W_8 \\ &= x\big((z - z')W_9 + (y - y')W_{10}\big) \end{aligned}$$

By the conditions on the W_is, this is equivalent to

$$\begin{aligned} &\big((b_1 - b_1') + (c - c') + (yz - y'z')\big)W_8 \\ &+\big((c - c') + (y - y')\big)W_9 \\ &+\big((b_2 - b_2') + (z - z')\big)W_{10} \\ &= x((z - z')W_9 + (y - y')W_{10}) \end{aligned}$$

Observe that $\det(W_5\ W_9\ W_{10}) \neq 0$ together with the condition that $W_5 = W_8$ implies that W_8, W_9 and W_{10} are linearly independent. Hence, A causes a collision if and only if

$$\begin{aligned}(b_1 - b_1') + (c - c') + (yz - y'z') &= 0,\\ (c - c') + (y - y') &= x(z - z'),\\ (b_2 - b_2') + (z - z') &= x(y - y').\end{aligned}$$

We cannot have simultaneously $y = y'$ and $z = z'$. Indeed, this would imply $c = c'$, $b_1 = b_1'$ and $b_2 = b_2'$. But by the completion property, A, b_1, b_2 and c uniquely determine a flag, so $B = B'$ and $C = C'$. Thus, there is at most one solution for x that would cause a collision, and a prerequisite is that $(B \subset C)$ and $(B' \subset C')$ satisfy

$$(b_1 - b_1') + (c - c') + (yz - y'z') = 0, \tag{2}$$

and

$$(y - y')\big((c - c') + (y - y')\big) = (z - z')\big((b_2 - b_2') + (z - z')\big). \tag{3}$$

Moreover, if $B = B'$, then $b_2 = b_2'$ and $y = y'$ so we also have $c = c'$. By the completion property, A, B and c uniquely determines a flag so $C = C'$, a contradiction, so $B \neq B'$. By the same token, $C \neq C'$. If $(B \subset C)$ and $(B' \subset C')$, satisfy all those conditions, we say they are *quadratically compatible*. In order to bound properly the probability of $\mathbf{E_3}$, we need to introduce an auxiliary event dealing with such pairs. For any positive integer γ, let $\mathrm{bad}_{\mathcal{X}:\mathrm{quad}[\gamma]}(\mathcal{Q})$ be the event that more than γ pairs $(B \subset C)$ and $(B' \subset C')$ in $\mathcal{Q}$ satisfying $B \neq B'$, $C \neq C'$, $B \cap B' \neq \emptyset$ and either $y \neq y'$ or $z \neq z'$ are quadratically compatible.

Lemma 2.4. *Let $(B \subset C)$ and $(B' \subset C')$ be two distinct pairs such that $B \neq B'$, $C \neq C'$, $B \cap B' \neq \emptyset$ and either $y \neq y'$ or $z \neq z'$. Then, the probability that $(B \subset C)$ and $(B' \subset C')$ are quadratically compatible is lower than $1/2^{2n-3}$.*

Proof. We denote by $\mathbf{F_1}$ the event that equation 2 is satisfied, and by $\mathbf{F_2}$ the event that equation 3 is satisfied. The fact that $B \neq B'$ and $C \neq C'$ implies that y, y', z and z' are mutualy independant random variables. This depends on the fact that $f_{\mathcal{Y}}$ and $f_{\mathcal{Z}}$ are two independent PuRFs. Suppose $y' \neq 0$. Then, $\mathbf{F_1}$ becomes

$$z' = \frac{y}{y'}z + \frac{(b_1 - b_1') + (c - c')}{y'},$$

and is satisfied with probability $1/2^n$. Then,

$$\Pr[\mathbf{F_1}] \leq \Pr[\mathbf{F_1} \mid y' \neq 0] + \Pr[y' = 0] \leq \frac{1}{2^{n-1}}.$$

Assume that $\mathbf{F_1}$ is true. Suppose $y' \neq 0$. We have

$$z - z' = (1 - \frac{y}{y'})z - \frac{(b_1 - b_1') + (c - c')}{y'}$$

Suppose furthermore that $y \neq y'$. Then, given y and y', $z - z'$ is a random variable uniformely distributed over $\mathbb{F}_{2^n}$, and

$$(z-z')^2 + (z-z')(b_2 - b_2') = (y-y')\big((c-c') + (y-y')\big)$$

admits at most 2 solutions for $z - z'$. Therefore,

$$\Pr[\mathbf{F_2} \mid \mathbf{F_1} \wedge y \neq y' \wedge y' \neq 0] \leq \frac{1}{2^{n-1}}.$$

Thus,

$$\begin{aligned}\Pr[\mathbf{F_2} \mid \mathbf{F_1}] &\leq \Pr[\mathbf{F_2} \mid \mathbf{F_1} \wedge y \neq y' \wedge y' \neq 0] + \Pr[y = y' \vee y' = 0 \mid \mathbf{F_1}] \\ &\leq \frac{1}{2^{n-1}} + \frac{1}{2^{n-1}} \\ &= \frac{1}{2^{n-2}}\end{aligned}$$

Finally, we have

$$\Pr[\mathbf{F_1} \wedge \mathbf{F_2}] = \Pr[\mathbf{F_1}] \Pr[\mathbf{F_2} \mid \mathbf{F_1}] \leq \frac{1}{2^{n-1}} \cdot \frac{1}{2^{n-2}} = \frac{1}{2^{2n-3}}.$$

□

Lemma 2.5. *Let $\mathcal{Q}$ be generated adaptatively. For any positive integer γ,*

$$\Pr[\mathrm{bad}_{\mathcal{X}:\mathrm{quad}[\gamma]}(\mathcal{Q})] \leq \left(\frac{q^4}{2^{2n-3}}\right)^{\gamma}.$$

Proof. There are less than q^4 possible pairs $(B \subset C)$, $(B' \subset C')$, and each of them are quadratically compatible with probability smaller than $1/2^{2n-3}$, so

$$\Pr[\mathrm{bad}_{\mathcal{X}:\mathrm{quad}[\gamma]}(\mathcal{Q})] \leq \binom{q^4}{\gamma} \frac{1}{2^{(2n-3)\gamma}} \leq \frac{q^{4\gamma}}{2^{(2n-3)\gamma}}.$$

□

We can now continue our bounding of $\Pr[\mathbf{E_3}]$. Given $\neg\mathrm{bad}_{\mathcal{X}:\mathrm{quad}[\gamma]}(\mathcal{Q})$, there are at most γ possible pairs $(B \subset C)$ and $(B' \subset C')$ that could give rise to a collision. Given such a pair, there is at most one x satisfying

$$\begin{aligned}(c-c') + (y-y') &= x(z-z'), \\ (b_2 - b_2') + (z-z') &= x(y-y').\end{aligned}$$

The probability of x hitting the solution is $1/2^n$, so the probability for an $\mathcal{X}$-query to cause a collision of type II is lower than $3q\gamma/2^n$. The same argument stands for $\mathcal{Y}$ and $\mathcal{Z}$-queries, and we obtain the result

$$\Pr[\mathbf{E_3}] \leq 3\frac{q\gamma}{2^n} + 3\left(\frac{q^4}{2^{2n-3}}\right)^{\gamma}.$$

□

2.4 Mats, bunches and constellations

We will now introduce a few notions that will be useful to bound the probability of $\mathbf{E_2}$.

Let $B \subset C$ be two compatible $\mathcal{Y}$ and $\mathcal{Z}$-queries, resulting in $y = f_{\mathcal{Y}}(B)$ and $z = f_{\mathcal{Z}}(C)$. By the completion property, for each $a \in \mathbb{F}_{2^n}$, there is a unique point (a, a_2, a_3) such that $A \in B \subset C$ forms a flag. Then, querying $f_{\mathcal{X}}(A)$ results in a new yield point on the line $\mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a}$, whose slope is determined by y and z, and independent of a. Thus, ranging over all possible $a \in \mathbb{F}_{2^n}$ creates a set of parallel lines. We call this set a *mat* :

$$\mathcal{M}_{\mathcal{X}:b_1,b_2,c,y,z} = \{\mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a} \mid a \in \mathbb{F}_{2^n}\}.$$

The notions of mat corresponding to $\mathcal{Y}$ and $\mathcal{Z}$ output lines are analogous. The following results are only prooved for the first kind of mat, but the other cases are similar.

Proposition 2.6. *All the lines of a mat are disctinct.*

Proof. Observe from the equation of a line that for $a \neq a'$, we have $\mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a} = \mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a'}$ if and only if W_1 and $W_5 + zW_9 + yW_{10}$ are colinear. But $W_1 = W_9 + W_{10}$, $W_5 = W_8$, so this would mean that $W_9 + W_{10}$ and $W_8 + zW_9 + yW_{10}$ are colinear. This is impossible since W_8, W_9 and W_{10} are linearly independent. □

Since the lines of a mat are parallel, the proposition actualy imply that the intersection of any two lines is empty.

We now define the notion of *bunch*. Given a point A, the bunch of A with respect to the set $\mathcal{Q}$ of preceeding queries is

$$\mathcal{B}_{\mathcal{X}:A} = \{\mathcal{L}_{\mathcal{X}:b_1,b_2,c,y,z,a} \mid (B, y), (C, z) \in \mathcal{Q} \wedge A \in B \subset C\}.$$

The i-th query $x = f_{\mathcal{X}}(A)$ adds to the yield set $\mathcal{Q}_{i-1}$ a single point for each line in the bunch. We refer to this set of points in a bunch as a *constellation*, denoted by

$$\mathcal{C}_{\mathcal{X}:A} = \{H^{f_{\mathcal{X}},f_{\mathcal{Y}},f_{\mathcal{Z}}}(a, b_1, b_2, c_1, c, c_3) \mid (B, y), (C, z) \in \mathcal{Q}_{i-1} \wedge A \in B \subset C\}.$$

We say that a set of constellations is collinear if one can pick a point from each constellation such that all the chosen points are collinear.

2.5 Bounding $\Pr[\mathbf{E_2}]$

We now focus on the event $\mathbf{E_2} = \neg\mathrm{coll}_{II}(\mathcal{Q}) \wedge \mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})$. In order to estimate the amount collinearity within the yield set, we will look at the probability of too much collinearity occuring in a single constellation, and the probability of too many constellations being collinear. For positive integers λ, μ and k,

let $\mathrm{bad}_{\mathrm{int}[\lambda]}(\mathcal{Q})$ be the event that there exists a constellation having a set of more than λ collinear points, and define $\mathrm{bad}_{\mathrm{ext}[k]}(\mathcal{Q})$ to be the event that there exists a line ℓ passing through more than k constellations whose bunches do not contain ℓ. Let $\mathrm{bad}_{\mathrm{pm}[\mu]}(\mathcal{Q})$ be the event that $\mathcal{Q}$ results in more than μ parallel mats.

Proposition 2.7. *Let $k, \lambda, \mu > 0$ be fixed integer such that $\kappa = k\lambda + \mu$. Then, for arbitrary $\mathcal{Q}$,*

$$\mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q}) \Rightarrow \left(\mathrm{bad}_{\mathrm{int}[\lambda]}(\mathcal{Q}) \vee \mathrm{bad}_{\mathrm{ext}[k]}(\mathcal{Q}) \vee \mathrm{bad}_{\mathrm{pm}[\mu]}(\mathcal{Q})\right).$$

Proof. Suppose that we have

$$\neg\left(\mathrm{bad}_{\mathrm{int}[\lambda]}(\mathcal{Q}) \vee \mathrm{bad}_{\mathrm{ext}[k]}(\mathcal{Q}) \vee \mathrm{bad}_{\mathrm{pm}[\mu]}(\mathcal{Q})\right),$$

and that $\tilde{\kappa}$ points lies on the same line ℓ. Each mat parallel to this line can contribute to at most one point on ℓ, and there are at most μ such mats. Removing these mats, there are at most k constellations collinear to this line, and each of these constellations contains at most λ points on ℓ. Thus, $\tilde{\kappa} \leq k\lambda + \mu = \kappa$, so we have $\neg\mathrm{bad}_{\mathrm{cl}[\kappa]}(\mathcal{Q})$. □

Definition 2.8 (Bouquet)**.** Given a set of queries $\mathcal{Q}$ and a positive integer λ, a *$\mathcal{Y}$-$\mathcal{Z}$-bouquet* of size λ is a set of λ distinct pairs of queries $(B^{(i)}, C^{(i)})$ of $\mathcal{Q}$, for $i = 1, ..., \lambda$, such that there exist an A compatible with all the pairs, in the sense that $A \in B^{(i)} \subset C^{(i)}$ for any $i = 1, ..., \lambda$.
If the intersection of the $B^{(i)}$'s is a single point, then A is uniquely detemined. Otherwise, all the $B^{(i)}$'s are equal and we say that the bouquet is *degenerated*. The *realization at A* of the bouquet is the set of all the points

$$H^{f_{\mathcal{X}}, f_{\mathcal{Y}}, f_{\mathcal{Z}}}(a, b_1^{(i)}, b_2^{(i)}, c_1^{(i)}, c^{(i)}, c_3^{(i)}),$$

for $i = 1, ..., \lambda$. We define similarly the notions of $\mathcal{X}$-$\mathcal{Y}$-bouquets and $\mathcal{X}$-$\mathcal{Z}$-bouquets.

For any positive integer λ, define $\mathrm{bad}_{\mathrm{bou}[\lambda]}(\mathcal{Q})$ to be the event that there is a non-degenerated bouquet of size λ whose realization is a set of λ distinct collinear points. Similarily, define $\mathrm{bad}_{\mathrm{dbou}[\lambda]}(\mathcal{Q})$ to be the event that there is a degenerated bouquet of size λ for which there is a realization consisting in a set of λ distinct collinear points.

Proposition 2.9. *For arbitrary $\mathcal{Q}$ and any integer $\lambda \geq 3$,*

$$\left(\neg\mathrm{coll}_{II}(\mathcal{Q}) \wedge \mathrm{bad}_{\mathrm{int}[\lambda]}(\mathcal{Q})\right) \Rightarrow \left(\mathrm{bad}_{\mathrm{bou}[\lambda]}(\mathcal{Q}) \vee \mathrm{bad}_{\mathrm{dbou}[\lambda]}(\mathcal{Q})\right)$$

Proof. Suppose $\mathrm{bad}_{\mathrm{int}[\lambda]}(\mathcal{Q})$ is caused by λ collinear points in $\mathcal{C}_{\mathcal{X}:A}$, where A is the i-th query. This precisely means that there is a $\mathcal{Y}$-$\mathcal{Z}$-bouquet in $\mathcal{Q}_{i-1} \subseteq \mathcal{Q}$ of size λ whose realization at A is a set of collinear points. The event $\neg\mathrm{coll}_{II}(\mathcal{Q})$ imply that the λ points of the realization are disctinct. Thus, we have either $\mathrm{bad}_{\mathrm{bou}[\lambda]}(\mathcal{Q})$ or $\mathrm{bad}_{\mathrm{dbou}[\lambda]}(\mathcal{Q})$, depending on the degeneracy of the bouquet. □

Proposition 2.10. *Let $x \in \mathbb{F}_{2^n}$, let $\mathcal{B}$ be a non degenerated $\mathcal{Y}$-$\mathcal{Z}$-bouquet such that $A \in \mathbb{F}_{2^n}^3$, the intersection of its lines, satisfy $f_{\mathcal{X}}(A) = x$, and let* $\mathrm{bad}(\mathcal{B})$ *denote the event that $\mathcal{B}$ causes a* $\mathrm{bad}_{\mathrm{bou}[3]}(\mathcal{Q})$*. Then,*

$$\Pr[\mathrm{bad}(\mathcal{B})] \leq ?,$$

Proof. Let $\mathrm{bad}(\mathcal{B})$ denote the event that the bouquet $\mathcal{B}$ causes a $\mathrm{bad}_{\mathrm{bou}[3]}(\mathcal{Q})$. Let $H^{(i)}$ be the output point of the postprocessing function at $(A, B^{(i)}, C^{(i)})$. The points $H^{(1)}$, $H^{(2)}$ and $H^{(3)}$ are collinear if and only if there exists an $\alpha \in \mathbb{F}_{2^n}$ such that $\alpha \neq 0, 1$ and

$$H^{(2)} - H^{(1)} = \alpha(H^{(3)} - H^{(1)}).$$

Since $\mathcal{B}$ is non-degenerated, we can suppose without loss of generality that $B^{(2)} \neq B^{(3)}$; then, because we are considering a PuRF, $y^{(2)}$ and $y^{(3)}$ are independent random variables. Inspecting the formula for the postprocessing in terms of W_is, and considering that W_8, W_9 and W_{10} and linearly independent, we obtain that $\mathrm{bad}(\mathcal{B})$ is equivalent to the existence of an α satisfying the system

$$\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)} = \alpha(\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)}) \tag{4}$$
$$\gamma^{(2)} + Y^{(2)} + xZ^{(2)} = \alpha(\gamma^{(3)} + Y^{(3)} + xZ^{(3)}) \tag{5}$$
$$\beta_2^{(2)} + Z^{(2)} + xY^{(2)} = \alpha(\beta_2^{(3)} + Z^{(3)} + xY^{(3)}), \tag{6}$$

where $\beta_j^{(i)} = b_j^{(i)} - b_1^{(1)}$, $\gamma^{(i)} = c^{(i)} - c^{(1)}$, $\zeta^{(i)} = y^{(i)}z^{(i)} - y^{(1)}z^{(1)}$, $Y^{(i)} = y^{(i)} - y^{(1)}$ and $Z^{(i)} = z^{(i)} - z^{(1)}$. Observe that if $\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0$, then the first equation uniquely determines α. We have that,

$$\begin{aligned}\Pr[\mathrm{bad}(\mathcal{B})] = {} & \Pr[\mathrm{bad}(\mathcal{B}) \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0] \Pr[\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0] \\ & + \Pr[\mathrm{bad}(\mathcal{B}) \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0] \Pr[\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0].\end{aligned}$$

But if $\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0$, then we also have $\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)} = 0$, so

$$\begin{aligned}& \Pr[\mathrm{bad}(\mathcal{B}) \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0] \\ & \leq \Pr[\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)} = 0 \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0]\end{aligned}$$

The probability that $y^{(2)}z^{(2)} = y^{(1)}z^{(1)} - \beta_1^{(2)} - \gamma^{(2)}$ knowing that $y^{(3)}z^{(3)} = y^{(1)}z^{(1)} - \beta_1^{(3)} - \gamma^{(3)}$ and $z^{(2)} \neq 0$ is $1/2^n$ since $y^{(3)}$ is independent from all the other RVs. Consider the case where $z^{(2)} = 0$. If $C^{(2)} \neq C^{(3)}$, then $z^{(2)}$ is independent to all the other RVs and

$$\Pr[z^{(2)} = 0 \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0] = \frac{1}{2^n}.$$

If $C^{(2)} = C^{(3)}$, then $z^{(2)} = 0$ implies that $z^{(3)} = 0$. Then conditioning on $y^{(3)}z^{(3)} = y^{(1)}z^{(1)} - \beta_1^{(3)} - \gamma^{(3)}$ and $z^{(2)} = 0$ implies $y^{(1)}z^{(1)} = \beta_1^{(3)} + \gamma^{(3)}$, and

the event $y^{(2)}z^{(2)} = y^{(1)}z^{(1)} - \beta_1^{(2)} - \gamma^{(2)}$ is equivalent to $y^{(1)}z^{(1)} = \beta_1^{(2)} + \gamma^{(2)}$. But we are supposing that $C^{(2)} = C^{(3)}$, so $\gamma^{(2)} = \gamma^{(3)}$, so the event $y^{(2)}z^{(2)} = y^{(1)}z^{(1)} - \beta_1^{(2)} - \gamma^{(2)}$ would imply $\beta_1^{(2)} = \beta_1^{(3)}$, so $b_1^{(2)} = b_1^{(3)}$. By the completion property, A, b_1 and $C^{(2)} = C^{(3)}$ uniquely determines a flag, so $B^{(2)} = B^{(3)}$, a contradiction, so we cannot have $z^{(2)} = 0$. Using the fact that for any events U, V and W,

$$\Pr[U \mid W] \leq \Pr[U \mid W, V] + \Pr[\neg V \mid W],$$

we have proven that

$$\Pr[\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)} = 0 \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0] \leq \frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}}.$$

We now have

$$\Pr[\text{bad}(\mathcal{B})] \leq \Pr[\text{bad}(\mathcal{B}) \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0] + \frac{\Pr[\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0]}{2^{n-1}}.$$

Let us focus on $\Pr[\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0]$. This event is equivalent to $y^{(3)}z^{(3)} = y^{(1)}z^{(1)} - \beta_1^{(3)} - \gamma^{(3)}$. If $z^{(3)} \neq 0$, then $y^{(3)}$ takes the value $(y^{(1)}z^{(1)} - \beta_1^{(3)})/z^{(3)}$ with probability $1/2^n$. Moreover, the probability that $z^{(3)} = 0$ is also $1/2^n$. So $\Pr[\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} = 0] \leq 1/2^{n-1}$. We then have

$$\Pr[\text{bad}(\mathcal{B})] \leq \Pr[\text{bad}(\mathcal{B}) \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0] + \frac{1}{2^{2(n-1)}}.$$

We can now look at $\Pr[\text{bad}(\mathcal{B}) \mid \beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0]$. Suppose $\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)} \neq 0$. Then α is determined by the equation 4,

$$\alpha = \frac{\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)}}{\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)}}.$$

Let $\mathbf{F_1}$ denote the event that equation 5 is satisfied, and $\mathbf{F_2}$ that equation 6 is satisfied. The event $\mathbf{F_1}$ becomes

$$\gamma^{(2)} + Y^{(2)} + xZ^{(2)} = \frac{\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)}}{\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)}} \left(\gamma^{(3)} + Y^{(3)} + xZ^{(3)}\right),$$

and $\mathbf{F_2}$ becomes

$$\beta_2^{(2)} + Z^{(2)} + xY^{(2)} = \frac{\beta_1^{(2)} + \gamma^{(2)} + \zeta^{(2)}}{\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)}} \left(\beta_2^{(3)} + Z^{(3)} + xY^{(3)}\right).$$

We have that $\mathbf{F_1}$ is equivalent to an event of the form

$$y^{(2)}\left(1 - z^{(2)}h\left(y^{(1)}, z^{(1)}, y^{(3)}, z^{(3)}\right)\right) = g\left(y^{(1)}, z^{(1)}, y^{(3)}, z^{(3)}, z^{(2)}\right),$$

where

$$h = h\left(y^{(1)}, z^{(1)}, y^{(3)}, z^{(3)}\right) = \frac{\gamma^{(3)} + Y^{(3)} + xZ^{(3)}}{\beta_1^{(3)} + \gamma^{(3)} + \zeta^{(3)}},$$

so if $1 - z^{(2)}h \neq 0$, there is a unique solution for $y^{(2)}$ and since it is independent from all the other variables, it occurs with probability $1/2^n$. Moreover, we have $1 = z^{(2)}h$ if and only if $z^{(2)} \neq 0$ and

$$y^{(3)}\left(1 - \frac{z^{(3)}}{z^{(2)}}\right) = \frac{\beta_1^{(3)} + \gamma^{(3)} - y^{(1)}z^{(1)}}{z^{(2)}}.$$

We now consider $B^{(3)} \neq B^{(1)}$, and the case where there is equality is treated similarily considering the equation

$$y^{(3)}\left(1 - \frac{z^{(3)} - z^{(1)}}{z^{(2)}}\right) = \frac{\beta_1^{(3)} + \gamma^{(3)}}{z^{(2)}}.$$

First suppose $C^{(2)} \neq C^{(3)}$. With $B^{(3)} \neq B^{(1)}$, $y^{(3)}$ is independent from $y^{(1)}$ so if $1 - z^{(3)}/z^{(2)} \neq 0$, there is a unique solution for $y^{(3)}$ which is reached with probability $1/2^n$. But $1 - z^{(3)}/z^{(2)} = 0$ if and only if $z^{(3)} = z^{(2)}$, so since $C^{(2)} \neq C^{(3)}$, this occures with probability $1/2^n$. Now suppose $C^{(2)} = C^{(3)}$. Then $z^{(3)} = z^{(2)}$ so the event $1 = z^{(2)}h$ requires that $\beta_1^{(3)} + \gamma^{(3)} - y^{(1)}z^{(1)} = 0$, which occurs with probability lower than $1/2^{n-1}$, considering that $y^{(1)}$ and $z^{(1)}$ are independent. Putting everything together, we obtain

$$\Pr[\mathbf{F_1}] \leq \Pr[\mathbf{F_1} \mid z^{(2)}h \neq 1] + \Pr[z^{(2)}h \neq 1] \leq \frac{1}{2^n} + \frac{1}{2^{n-1}} \leq \frac{1}{2^{n-2}}.$$

We now have

$$\Pr[\text{bad}(\mathcal{B})] \leq \frac{\Pr[\mathbf{F_2} \mid \mathbf{F_1}]}{2^{n-2}} + \frac{1}{2^{2(n-1)}}.$$

We now want to bound $\Pr[\mathbf{F_2} \mid \mathbf{F_1}]$... *to do* □

Lemma 2.11. *For any integer* $\lambda \geq 3$*,*

$$\Pr[\text{bad}_{\text{bou}[\lambda]}(\mathcal{Q})] \leq \Pr[\text{bad}_{\text{bou}[3]}(\mathcal{Q})]^{\lambda-2}.$$

Proof. Let $\mathcal{B} = \{(B^{(i)}, C^{(i)}) \mid i = 1, ..., \lambda\}$ be a non-degenerated $\mathcal{Y}$-$\mathcal{Z}$-bouquet of size λ. Let $A \in \mathbb{F}_{2^n}^3$ be the intersection of its lines. Without loss of generality, $B^{(1)} \neq B^{(2)}$. For $i = 3, ..., \lambda$, we define the non-degenerated bouquet

$$\mathcal{B}_i = \{(B^{(1)}, C^{(1)}), (B^{(2)}, C^{(2)}), (B^{(i)}, C^{(i)})\}.$$

We have that $\text{bad}(\mathcal{B})$ implies $\text{bad}(\mathcal{B}_i)$ for every i. We denote $H^{(i)}$ the output point of the flag $A \in B^{(i)} \subset C^{(i)}$. If we have $\text{bad}(\mathcal{B})$, then $H^{(1)}$ and $H^{(2)}$ lies on a unique line ℓ, and $\text{bad}(\mathcal{B}_i)$ occurs if and only if $H^{(i)}$ lies on ℓ. *Problem : the events* $H^{(i)} \in \ell$ *are not mutually independent, since we can have* $B^{(i)} = B^{(j)}$ *or* $C^{(i)} = C^{(j)}$. □

Conclusion

Conclusion...

References

[1] Dimitar Jetchev, Onur Özen, Martijn Stam, *Collisions are not incidental: a compression function exploiting discrete geometry.* In *Proceedings of the 9th international conference on Theory of Cryptography* (TCC'12), Ronald Cramer (Ed.). Springer-Verlag, Berlin, Heidelberg, 303-320, 2012.