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 (Hyperplane)**.** For any integer $0 \leq d < n$, a subset H of K^n is a *hyperplane of dimension* d if there exist parameters $c_{i,j} \in K$ such that

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

Definition 1.2 (Flag)**.** A *flag* in K^n is a sequence $K^n \supset H_{n-1} \supset ... \supset H_1 \supset H_0$ where H_i is a hyperplane of dimension i for $i = 0, ..., n-1$.

Remark 1.3. This definition excludes "degenerated" hyperplanes. For instance, with $K = \mathbb{R}$ and $n = 2$, a hyperplane of dimension 1 is a nonverticale line.

Proposition 1.4. *Let d be a positive integer smaller than n, H a hyperplane of dimention $d-1$ parameterised by the collection $\{c_{i,j}\}$, and G a hyperplane of dimension d parameterised 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

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

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 $k = 0, ..., d-1$. □

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 1.4, 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 1.5 (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, b_1 and b_2 uniquely determine a flag $A \in B \subset C$.*

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

Proof. The points 2 and 2 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 1.6 (Non-Linear Matrix-Style Postprocessing)**.** We consider the postprocessing function $C^{\text{POST}} : K^7 \to K^3$ defined by

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

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 1.7 (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

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

where

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

2 Proof of Collision Resistance

2.1 Strategy

Blabla,

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

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

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

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

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

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, 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 $C = C'$.

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 $B \cap B' \neq \emptyset$ 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 $B \cap B' \neq \emptyset$ 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.

1. First assume that $B = B'$ and $C \neq C'$. Then, $y = y'$, $z \neq z'$ and one of z ou z' is non-zero. Assume it is z'. The event $\mathbf{F_1}$ becomes

   $$yz' = yz + (c - c'),$$

   But $c \neq c'$. Indeed, by the completion property, $B \cap B' \neq \emptyset$, $B = B'$ and $c = c'$ would imply $C = C'$. Therefore $\mathbf{F_1}$ is satisfied if and only if $y \neq 0$ and $z' = z + (c - c')/y$. But z and z' are two random variables uniformely distributed over $\mathbb{F}_{2^n}$, and independent since $C \neq C'$, so we have

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

   Now assume $\mathbf{F_1}$ is true. Then, we have $z - z' = (c' - c)/y$, and $\mathbf{F_2}$ is satisfied if and only if $y(b_2 - b_2') = (c' - c)$. But y is a random variables uniformely distributed over $\mathbb{F}_{2^n}$. Assuming $\mathbf{F_1}$ adds the condition $y \neq 0$. Therefore, $\Pr[\mathbf{F_2} \mid \mathbf{F_1}] \leq 1/(2^n - 1) \leq 1/(2^{n-1})$. Finally, in the case $B = B'$ and $C \neq C'$, we have

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

2. The same argument in the case $B \neq B'$ and $C = C'$ holds

$$\Pr[\mathbf{F_1} \wedge \mathbf{F_2}] \leq \frac{1}{2^{2(n-1)}}.$$

3. Assume now that $B \neq B'$ and $C \neq C'$. This implies that y, y', z and z' are mutualy independant random variables. 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

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

□

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.