# Characterization of EME with Linear Mixing

Nilanjan Datta and Mridul Nandi

Cryptology Research Group
Applied Statistics Unit
Indian Statistical Institute
203, B.T. Road, Kolkata, India 700108
nilanjan_isi_jrf@yahoo.com, mridul.nandi@gmail.com

**Abstract.** Encrypt-Mix-Encrypt is a type of SPRP based construction, where a masked plaintext is encrypted in ECB mode of, then a non-linear mixing is performed and then again an encryption is performed in ECB mode which is masked to produce the ciphertext. Using the property of the binary field, the authors proved that the construction is not SPRP secure if the mixing used is linear. In this paper, we observe that relaxing the mixing operation to some specific efficient linear mixing provides the PRP property of the construction. Moreover choosing a linear mixing that gives the online property is not a difficult task. We can use this fact to construct an efficient Online PRP using Encrypt-Mix-Encrypt type of construction with the mix operation being a linear online mixing, making the construction efficient and online. We also show that the construction with linear mixing doesn't provide SPRP security even if we perform all the operations in a prime field instead of binary field. Thus, we fully characterize EME with linear mixing.

**Keywords:** EME, PRP, OLPRP, Lmix, OLmix.

## 1 Introduction

EME [3] is a block-cipher mode of operation, that turns an $n$-bit block cipher into a tweakable enciphering scheme that acts on strings of $mn$ bits, where $m \in [1..n]$. The mode is parallelizable, but as serial-efficient as the non-parallelizable mode CMC [7]. EME algorithm entails two layers of ECB encryption and a non-linear mixing in between. EME is proved to provide SPRP [4] security in the standard, provable security model assuming that the underling block cipher is SPRP secure. The authors showed that, if the construction uses liner mix, instead of non-linear mixing then an SPRP attack can be mounted using the properties of an binary field.

### 1.1 Motivation and Our Contribution.

Motivation of this paper is to fully characterise EME with linear mixing (instead of non-linear mixing). We know there exist attacks against the SPRP security of EME with linear mixing but we don't have any idea about the PRP [4]

or OLPRP security of the construction. In this paper, we observe that some specific efficient linear mixing (respectively Online linear mixing) still retains the PRP (resp. OLPRP) property of the construction. In fact, we showed that, the necessary as well as sufficient condition for EME with linear mixing to have the PRP and OLPRP security, is to use a particular type of linear mixing : Lmix and OLmix respectively. Moreover, we prove that any other linear mixing except Lmix and OLmix, doesn't provide the PRP and OLPRP security respectively. Thus, we fully characterize EME with linear mixing. We can use this fact to construct an efficient PRP (or OLPRP) using Encrypt-Mix-Encrypt type of construction with linear (online) mixing. Moreover, we prove that EME with any linear mixing doesn't provide SPRP security, not just in the binary field (As shown by the authors of EME) but also in any other field.

### 1.2 Application of our Result.

Modern cryptography community put a lot of efforts of designing different Authenticated Encryptions. Lack of being standardized of this notion motivates to make a call for CEASER standard of Authenticated Encryption [8, 9, 10]. For that, various new authenticated encryption designs have been proposed recently. We know that, for constructing an Authenticated Encryption, it is not necessary to be a SPRP construction, rather a PRP construction (not length preserving) with properly generated tag, is good enough to give an Authenticated Encryption.

So, using our result one can efficiently construct an authenticated Encryption using EME with Lmix and then properly generating the tag. For online secure Authenticated Encryption, one can use EME with OLmix as the basis of the construction. Such constructions are very efficient as it doesn't use any non-linear functions and the mode is parallelizable. Infact some of the existing constructions like ELmE [2] and COPA [1] uses this kind of structure (EME with Olmix) as the underlying structure to make the construction online, fully pipelined implementable.

### 1.3 Outline of the paper

In this paper we have observed that if we use linear mixing with certain properties in EME, then the privacy security of the construction still holds. After providing basic preliminaries in section 2, we define two types of linear mixing - online linear mixing and full linear mixing that can be used in the mixing part of EME. Then in section 3, using the well known Interpolation technique and Patarin's H Coefficient [5] method, we prove that those two mixing provides PRP and OLPRP security of EME respectively. Finally in section 4, we show the SPRP attack against the generalized EME construction with linear mixing that works for any fields. Finally we conclude along with some possible future works.

## 2 Preliminaries

### 2.1 Security Measures

Given an adversary $A$ (w.o.l.g. throughout the paper we assume a **deterministic adversary**) and two functions $f$ and $g$, we define the distinguishing advantage of $A$ distinguishing $f$ from $g$. More formally,

$$\mathbf{Adv}_g^f(A) = \mathsf{Pr}[A^f = 1] - \mathsf{Pr}[A^g = 1].$$

In this paper, we give a particularly strong definition of privacy, one asserting indistinguishability from random strings. Consider an adversary $A$ who has access of one of two types of oracles: a "real" encryption oracle or an "ideal" encryption oracle. A real encryption oracle, $F_K$, takes as input $M$ and returns $C = F_K(M)$. Whereas an ideal encryption oracle \$ returns a random string $R$ with $\|R\| = \|M\|$, for every fresh message $M$. Given an adversary $A$ and an encryption scheme $F$, we define the **prp-advantage** of $A$ by the distinguishing advantage of $A$ distinguishing $F$ from \$. More formally,

$$\mathbf{Adv}_F^{\mathrm{prp}}(A) := \mathbf{Adv}_F^{\$}(A) = \mathsf{Pr}_K[A^{F_K} = 1] - \mathsf{Pr}_{\$}[A^{\$} = 1].$$

Similarly, we define the **olprp-advantage** of $A$ for which the the ideal online encryption oracle $\$_{ol}$ responses random string keeping the online property. The online privacy advantage of an adversary $A$ against $F$ is defined as $\mathbf{Adv}_F^{\mathrm{olprp}}(A) := \mathbf{Adv}_F^{\$_{ol}}(A)$.

A more strong definition of security is given by the *sprp* notion which is similar to the *prp* notion except that the adversary has the power of a accessing decryption oracle as well. Formally,

$$\mathbf{Adv}_F^{\mathrm{sprp}}(A) = \mathsf{Pr}_K[A^{F_K, F_K^{-1}} = 1] - \mathsf{Pr}_{\$}[A^{\$,\$^{-1}} = 1].$$

### 2.2 View and *A*-realizable

We define view of a deterministic adversary $A$ interacting with an oracle $\mathcal{O}$ by a tuple $\tau(A^{\mathcal{O}}) := (Q_1, R_1, \ldots, Q_q, R_q)$ where $Q_i$ is the $i^{\text{th}}$ query and $R_i$ is the response by $\mathcal{O}$. It is also called $\mathcal{O}$-view. A tuple $\tau = (Q_1, R_1, \ldots, Q_q, R_q)$ is called $A$-realizable if it makes query $Q_i$ after obtaining all previous responses $R_1, \ldots, R_{i-1}$. As $A$ is assumed to be deterministic, given $R_1, \ldots, R_q$, there is an unique $q$-tuple $Q_1, \ldots, Q_q$ for which the combined tuple is $A$-realizable. Now we describe the popular coefficient H-technique which can be used to bound distinguish advantage. Suppose $f$ and $g$ are two oracles and $V$ denotes all possible $A$-realizable views while $A$ interacts with $f$ or $g$ (they have same input and output space).

**Lemma 1 (Coefficient H Technique [5, 6]).** *If $\forall v \in V_{good} \subseteq V$ (as defined above), $\mathsf{Pr}[\tau(A^g) = v] \geq (1-\epsilon)\mathsf{Pr}[\tau(A^f) = v]$, then the distinguishing advantage $\mathbf{Adv}_g^f(A)$ of $A$ is at most $\epsilon + \mathsf{Pr}[\tau(A^f) \notin V_{good}]$.*

*Proof.* Here we calculate the advantage of any adversary $\mathcal{A}$ distinguishing two random online functions $f$ and $g$.

$$\begin{aligned}\mathbf{Adv}_g^f(\mathcal{A}) &= \mathsf{Pr}[\mathcal{A}^f = 1] - \mathsf{Pr}[\mathcal{A}^g = 1] \\ &= \sum_{v \in V} (\mathsf{Pr}[\tau(\mathcal{A}^f) = v] - \mathsf{Pr}[\tau(\mathcal{A}^g) = v]) \\ &= \sum_{v \in V \bigcap V_{good}} (\mathsf{Pr}[\tau(\mathcal{A}^f) = v] - \mathsf{Pr}[\tau(\mathcal{A}^g) = v]) \\ &\quad + \sum_{v \in V \setminus V_{good}} (\mathsf{Pr}[\tau(\mathcal{A}^f) = v] - \mathsf{Pr}[\tau(\mathcal{A}^g) = v]) \\ &= \epsilon + \mathsf{Pr}[\tau(\mathcal{A}^f) \notin V_{good}]\end{aligned}$$

□

### 2.3 Consistent Collision Relations for a Linear Function

Assume, $\mathbb{B} = \{0,1\}^n$. An $\ell$-tuple $x \in \mathbb{B}^\ell$ is denoted by $(x[1], x[2], \ldots, x[\ell])$. We call $\ell := \|x\|$ block-length of $x$. For $0 \le a \le b < \ell$ we denote $x[a..b] := (x[a], x[a+1], \ldots, x[b])$, $x[..b] = x[1..b]$.

Suppose $X = X[1..r_1]$ is a $r_1$-tuple of variables of $\mathbb{B}$ and $\mathcal{L} : \mathbb{B}^{r_1} \to \mathbb{B}^{r_2}$ is a linear function. We denote $Y = \mathcal{L}(X)$ which is an $r_2$-tuple of variables from $\mathbb{B}$. Let $\gamma_1$ and $\gamma_2$ are two equivalence relations defined on the sets respectively $[1..r_1]$ and $[1..r_2]$. Let $X^{\gamma_1}$ denote the tuple of variables which satisfies the collision relation $\gamma_1$ by replacing identical variables by the variable which occurred with minimum index. We say that $(\gamma_1, \gamma_2)$ is consistent with $\mathcal{L}$ if $\mathcal{L}_i(X^\gamma) \equiv \mathcal{L}_j(X^{\gamma_1})$ if and only if $i$ and $j$ are related in $\gamma_2$. Clearly, given any $\gamma_1$ and $\mathcal{L}$ there is exactly one $\gamma_2$ for which $(\gamma_1, \gamma_2)$ is consistent with $\mathcal{L}$. We write $\gamma_1 \Rightarrow_{\mathcal{L}} \gamma_2$.

*Example 1.* If $\gamma_1 = \{\{1,3\},\{2\},\{4,6\},\{5\}\}$ for $r_1 = 6$, then we write $X^{\gamma_1} = (X_1, X_2, X_1, X_4, X_5, X_4)$. Let $\mathcal{L}$ map into three variables (i.e., $r_2 = 3$ such that $\mathcal{L}_1 = X_1 + X_2 + X_3 + X_6$, $\mathcal{L}_2 = X_4 + X_5 + X_6$ and $\mathcal{L}_3 = X_2 + X_4$ then $\mathcal{L}_1(X^{\gamma_1}) = \mathcal{L}_3(X^{\gamma_1}) = X_2 + X_4$ and $\mathcal{L}_2(X^{\gamma_1}) = X_5$ (we work it here in binary field). So $\gamma_1 \Rightarrow_{\mathcal{L}} \gamma_2$ where $\gamma_2 = \{\{1,3\},\{2\}\}$.

**Lemma 2.** *[Number of Solutions for Consistent relations] Let $(\gamma_1, \gamma_2)$ be consistent with $\mathcal{L} : \mathbb{B}^{r_1} \to \mathbb{B}^{r_2}$ then*

$$|\{X : Coll(X) = \gamma_1, Coll(\mathcal{L}(X)) = \gamma_2\}| \ge 2^{ns_1} \times (1 - \frac{s^2}{2^{n+1}})$$

*where $s_1$ and $s_2$ denote the number of equivalence classes of $\gamma_1$ and $\gamma_2$ respectively and $s = s_1 + s_2$.*

**Proof.** Let $Y = \mathcal{L}(X)$. Because of consistency, for all related $i, j$ in $\gamma_2$, $Y_i = Y_j$. There may be additional equality which must be avoided. For all unrelated pair $(i,j)$ in $\gamma_2$ we must choose $X$ in a manner such that $Y_i \ne Y_j$ and similarly for

all unrelated pair $(i, j)$ in $\gamma_1$ we have $X_i \neq X_j$. Due to consistency, any one can happen for at most $2^{n(s_1-1)}$ many $X$'s as $\mathcal{L}_i(X^{\gamma_1}) = \mathcal{L}_j(X^{\gamma_1})$ gives a non-trivial equation. So the result follows as we have at most $\binom{s}{2}$ such equalities. □

## 3 $\mathsf{EME}_{\mathsf{mix}}$ : A General Form of $EME$ with Linear Mixing

The construction $\mathsf{EME}_{\mathsf{mix}}$ is a variant of $EME$, where linear mixing function $\mathsf{mix}$ is used instead of non-linear mixing. Other than the mixing part, the construction remains same as $EME$. For a message $M$ of length is $l$, the construction is described below. Here $E_K : \mathbb{F} \to \mathbb{F}$ is a block cipher that maps a field element to another. Here possible choices of $\mathbb{F}$ is $GF(2^n)$ or $GF(p^n)$ where $p$ is a prime. Although we don't have any blockcipher with $\mathbb{F} = GF(p^n)$ but we keep it for theoretical importance.

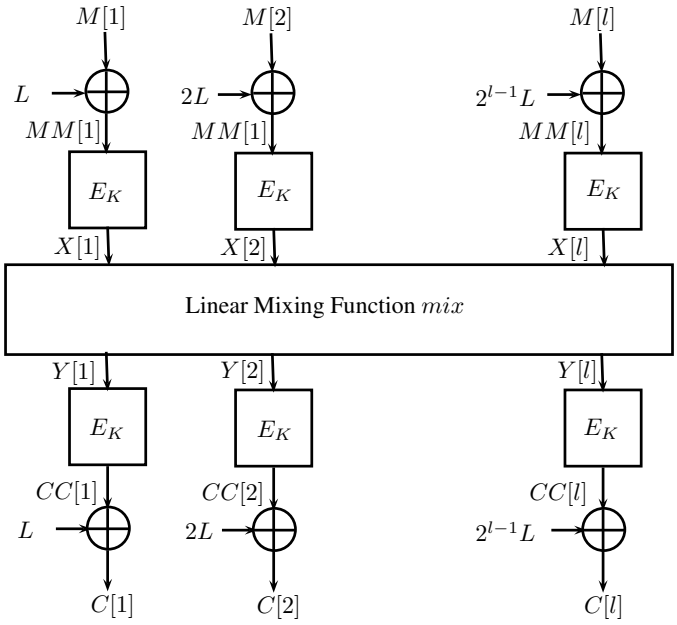


**Fig. 3.1.** (Encrypt-Mix-Encrypt Construction

- Layer-1 (Input Masking Layer): $MM[j] = M[j] + 2^{j-1}L$, $1 \leq j \leq l$.
- Layer-2 (1st Encrypt Layer) : $X[j] = E_K(MM[j])$, $1 \leq j \leq l$.
- Layer-3 (Linear Mix Layer) : $Y = \mathsf{mix}(X)$.
- Layer-4 (2nd Encrypt Layer) : $CC[j] = E_K(Y[j])$, $1 \leq j \leq l$.
- Layer-5 (Output Masking Layer) : $C[j] = CC[j] + 2^{j-1}L$, $1 \leq j \leq l$.

Now, the main focus is on the linear mix function $\mathsf{mix}$. The linear function $\mathsf{mix}$ is obtained via a matrix multiplication as given below :

$$Y = \mathsf{mix}(X) = \big(B\big).\ \big(X\big)$$

$$\begin{pmatrix} Y[1] \\ \cdot \\ \cdot \\ Y[l] \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \cdot & \cdot & \cdots & \cdot \\ b_{l1} & b_{l2} & \cdots & b_{ll} \end{pmatrix} \begin{pmatrix} X[1] \\ \cdot \\ \cdot \\ X[l] \end{pmatrix}$$

where $B$ is an $(l \times l)$ invertible matrix. The invertibility property is required to ensure that the decryption is possible. The inverse matrix of $B$ is denoted by $A_{(l \times l)}$

We are typically interested in two types of mixing functions - Online linear mixing function $\mathsf{OLmix}$ and full linear mixing function $\mathsf{Lmix}$, described below :

1. **Lmix**: It is the $\mathsf{mix}$ function for which $B$ is a full matrix i.e. all the entries of matrix $B$ is non-zero.

2. **OLmix**: It is the $\mathsf{mix}$ function where $B$ is a lower triangular matrix with all it's lower triangular entries nonzero.

In the following sections, we'll show that the necessary and sufficient condition for an $\mathsf{EME}_{\mathsf{mix}}$ to obtain full security is to have $\mathsf{Lmix}$ as the $\mathsf{mix}$ function and $\mathsf{OLmix}$ to obtain the online security.

## 4 Security of $\mathsf{EME}_{\mathsf{Lmix}}$ and $\mathsf{EME}_{\mathsf{OLmix}}$

The main two results that we have proved in this paper is given below.

**Theorem 1.** *The distinguishing advantage of an adversary $\mathcal{D}$, distinguishing $\mathsf{EME}_{\mathsf{OLmix}}$ ($\mathsf{EME}$ with linear mix property 2) encryption and decryption from \$, which is capable of making at most $q$ distinct queries, is bounded by,*

$$\mathbf{Adv}^{\mathrm{olprp}}_{\mathsf{EME}_{\mathsf{OLmix}}}(D) \ \leq \ \frac{5s^2}{2^n}$$

*where $s$ is the no. of distinct non-prefixed blocked queried.*

**Theorem 2.** *The distinguishing advantage of an adversary $\mathcal{D}$, distinguishing $\mathsf{EME}_{\mathsf{Lmix}}$ ($EME$ with linear mix property 1) encryption from \$, which is capable of making at most $q$ distinct queries, is bounded by,*

$$\mathbf{Adv}^{\mathrm{prp}}_{\mathsf{EME}_{\mathsf{Lmix}}}(D) \ \leq \ \frac{5\sigma^2}{2^n}$$

*where $\sigma$ is total length of all the $q$ queries, made by the adversary.*

**Proof.** We prove theorem 1 by applying proposition 1 and 3 in Patarin's H-coefficient technique and theorem 2 by using proposition 2 and 4 in Patarin's H-Coefficient technique. □

**Proof Idea.** Here we give the proof idea of theorem 1. The basic idea is as follows : We first define good online views, where two ciphertext blocks are equal iff their corresponding messages upto that block is identical. We'll show that in case of ideal online cipher, generating such a good online view has very high probability. We call $L$ is Valid, if the value of $L$ ensures the collision relation of $M$ and $MM$ and collision relation of $C$ and $CC$ are same. We prove that, with very high probability $L$ is valid. As permutations preserve collision relations, a valid $L$ ensures that the collision relation of $M$ and $X$ are same and collision relation of $C$ and $Y$ are same. Then, using consistent collision relations for linear functions, we show that with very high probability, the collision relation of $Y$ and $\mathsf{OLmix}(X)$ is same. Thus for a fixed valid $L$, the conditional interpolation probability is very high for the construction.

Now, applying the two results : High probability of obtaining a good online view and high interpolation probability of $\mathsf{EME}_{\mathsf{OLmix}}$ for good online views, in Patarin's H-Coefficient technique, we have the Online PRP security of $\mathsf{EME}_{\mathsf{OLmix}}$.

We use exactly similar idea for the theorem 2.

### 4.1 Notation Setup and Definitions

Let $A$ be an adversary which makes $q$ queries $M_i$ and obtains responses $C_i$, $1 \leq i \leq q$. We denote $\|M_i\| = \|C_i\| = l_i$. Let $\sigma = \sum_{i=1}^{q} l_i$. The first part follows directly from the coefficient H technique (see Lemma 1 and following Propositions 1 and 2. For this, we first need to define a set of good views $V_{good}$ which would be applied in the proposition. We denote $(M_1, \ldots, M_q)$ by $\tau_{in}$. We assume that all $M_i$'s are distinct.

**Definition 1 (Good Online Views).** *A ciphertext tuple $\tau_{out} = (C_1, \ldots, C_q)$ (also the complete view $\tau = (\tau_{in}, \tau_{out})$) is called* **good online** *view (belongs to $\tau_{good}$) w.r.t. $\tau_{in}$ if $(\tau_{in}, \tau_{out})$ is an online view (i.e., it must be realized by an online cipher, see section 2) and the following condition hold:*

$$C_i[j] = C_{i'}[j] \Rightarrow M_i[..j] = M_{i'}[..j]$$

The condition says that we can have collision of ciphertext blocks in a position only if they are ciphertexts of two messages with same prefixes up to that block.

**Definition 2 (Good Views).** *A ciphertext tuple $\tau_{out} = (C_1, \ldots, C_q)$ (also the complete view $\tau = (\tau_{in}, \tau_{out})$) is called* **good** *view (belongs to $\tau_{good}$) w.r.t. $\tau_{in}$ if $(\tau_{in}, \tau_{out})$ is a view and the following condition hold:*

$$C_i[j] \neq C_{i'}[j]$$

The condition says that we don't have any collision in ciphertext blocks.

**Collision Relation.** Consider the following relations, defined on a good view $\tau$. Let $\gamma_1$ and $\gamma_2$ are collision relations defined on the set $\{(i,j) : i \leq q,\ j \leq l_i\}$. A pair $((i,j),(i',j'))$ is related in $\gamma_1$ if $j = j'$ and $M_i[j] = M_{i'}[j]$. On the other hand, the pair $((i,j),(i',j'))$ is related in $\gamma_2$ if $j = j'$ and $C_i[j] = C_{i'}[j]$. All

other pairs are unrelated. Let the no. of equivalence class of $\gamma_i$ be $s_i$, $i = 1, 2$. Note that $s_2 = s$, the number of prefixes of $M_i$ containing at least one message block.

**Lemma 3.** *The collision relations defined as above is consistent with* OLmix *and* Lmix.

**Proof.** Let $\mathbf{Y} = (Y_1 := \mathsf{OLmix}(X_1), \ldots, Y_q := \mathsf{OLmix}(X_q))$. Since the view is good online, $C_i[j] = C_{i'}[j]$ can happen if $M_i[..j] = M_{i'}[..j]$. In this case, clearly, $Y_i[j] = Y_{i'}[j]$. Now for any other pair $((i, j), (i', j'))$, it is easy to see that OLmix function leads to a non-trivial equation $\mathsf{OLmix}_j(X_i^{\gamma_1}) = \mathsf{OLmix}_{j'}(X_{i'}^{\gamma_1})$ i.e.

$$\sum_{k=1}^{j} b_{jk} X_i[k] \;=\; \sum_{k=1}^{j'} b_{j'k} X_{i'}[k]$$

.

Similarly, it is easy to see that, for Lmix, any pair $((i, j), (i', j'))$, Lmix function leads to a non-trivial equation $\mathsf{Lmix}_j(X_i^{\gamma_1}) = \mathsf{Lmix}_{j'}(X_{i'}^{\gamma_1})$. Hence the consistency holds. □

### 4.2 High Probability of obtaining Good and Good Online Views

**Proposition 1 (Obtaining a Good Online View has high probability).**

$$\mathsf{Pr}[\tau(A^{\$^{ol}}) \notin V_{good,ol}] \;\leq\; \frac{s_2^2}{2^n}.$$

**Proof.** According to the definition, an online view is not a good view if $\exists i, j, i', j'$ with $C_i[j] = C_{i'}[j']$, where $M_i[..j] \neq M_{i'}[..j']$. Suppose $i < i'$ or $i = i', j < j'$. Then $C_i[j]$ is computed by $M_i[..j]$ before the computation of $C_{i'}[j']$. As $M_i[..j] \neq M_{i'}[..j']$, the outcome of $C_{i'}[j']$ is random and fresh from $C_i[j]$. So, the probability that $C_i[j]$ takes the previously computed fixed value $C_i[j]$ is $\frac{1}{2^n}$. As at most $\binom{s_2}{2}$ pairs are there, the probability that $\tau(A^{\$^{ol}}) \notin V_{good}$ is at most $\frac{s_2^2}{2^n}$.

**Proposition 2 (Obtaining a Good View has high probability).**

$$\mathsf{Pr}[\tau(A^{\$}) \notin V_{good}] \;\leq\; \frac{\sigma^2}{2^n}.$$

**Proof.** According to the definition, a view is not a good view if $\exists (i, j) \neq (i', j')$ with $C_i[j] = C_{i'}[j']$. Suppose $i < i'$ or $i = i', j < j'$. Then $C_i[j]$ is computed by $M_i[..j]$ before the computation of $C_{i'}[j']$. As $M_i[..j] \neq M_{i'}[..j']$, the outcome of $C_{i'}[j']$ is random and fresh from $C_i[j]$. So, the probability that $C_i[j]$ takes the previously computed fixed value $C_i[j]$ is $\frac{1}{2^n}$. As at most $\binom{\sigma}{2}$ pairs are there, the probability that $\tau(A^{\$}) \notin V_{good}$ is at most $\frac{\sigma^2}{2^n}$. □

### 4.3 High Interpolation Probabilty of $\mathsf{EME}_{\mathsf{Lmix}}$ and $\mathsf{EME}_{\mathsf{OLmix}}$

In this section, we prove that the interpolation probability is high for $ELE_{OLmix}$.

**Proposition 3 (High interpolation probability of $\mathsf{EME}_{\mathsf{OLmix}}$).**

$\forall \tau \in V_{good,ol}$,

$$\mathsf{Pr}[\tau(A^{\mathsf{EME}^{\Pi,\mathbf{L}}_{OLmix}}) = \tau] \geq (1 - \tfrac{4s_2^2}{2^n}) \times \mathsf{Pr}[\tau(A^{\$^{ol}}) = \tau].$$

*Note that* $\mathsf{Pr}[\tau(A^{\$^{ol}}) = \tau] = 2^{-ns}$ *where* $s$ *denotes the number of non-empty prefixes of* $M_i$, $1 \leq i \leq q$ *as for every different prefixes,* $\$^{ol}$ *assigns an independent and uniform ciphertext blocks.*

**Proof.** As adversary is deterministic, we restrict to those good views which can be obtained by $A$. Hence the probability $\mathsf{Pr}[\tau(A^{\mathsf{EME}}) = \tau]$ is same as

$$\mathsf{Pr}[\mathsf{EME}^{\Pi,\mathbf{L}}_{OLmix}(M_i) = C_i, 1 \leq i \leq q].$$

Before computing interpolation probability we denote all intermediate variables while computing $\mathsf{EME}^{L,\pi}_{OLmix}(M_i) = C_i$. Let for all $i$ and $j$ whenever defined

1. $MM_i[j] = L \cdot 2^{j-1} + M_i[j]$
2. $\Pi(MM_i[j]) = X_i[j]$,
3. $\mathsf{OLmix}(X_i) = Y_i$
4. $CC_i[j] = L \cdot 2^{j-1} + C_i[j]$

Note that $CC$ has been defined through ciphertext and $L$ instead of applying $\Pi$ on $Y$ blocks. Let $\mathbf{MM} = (MM_1, \ldots, MM_q)$ and similarly we define $\mathbf{X}, \mathbf{Y}$ and $\mathbf{CC}$. So, we have $\mathsf{mix}(\mathbf{X}) = \mathbf{Y}$ with the extended definition of $\mathsf{mix}$ which applies mix function for each $X_i$. We call $L$ valid if it computes $(\mathbf{MM}, \mathbf{CC})$ for which only equality among the blocks occurs in $SS_i[j] = SS_{i'}[j]$ where $S_i[j] = S_{i'}[j]$. One can easily show that $L$ is valid with probability at least $(1 - \frac{2s_2^2}{2^n})$ :

**Lemma 4.** $\mathsf{Pr}[L \text{ is valid}] \geq (1 - \epsilon_1)$ *where* $\epsilon_1 = \frac{2s_2^2}{2^n}$.

**Proof.** Consider the bad cases when $L$ is not valid :

**Case 1:** $MM_i[j] = MM_{i'}[j']$ which implies $L = \frac{(M_i[j] + M_{i'}[j'])}{2^{j-1} + 2^{j'-1}}$ which occurs with probabilty $\frac{\binom{s_1}{2}}{2^n}$.
**Case 2:** $MM_i[j] = CC_{i'}[j']$ which implies $L = \frac{(M_i[j] + C_{i'}[j'])}{2^{j-1} + 2^{j'-1}}$ which occurs with probabilty $\frac{s_1 \cdot s_2}{2^n}$.
**Case 3:** $CC_i[j] = MM_{i'}[j']$ which implies $L = \frac{(C_i[j] + C_{i'}[j'])}{2^{j-1} + 2^{j'-1}}$ which occurs with probabilty $\frac{\binom{s_2}{2}}{2^n}$.

Using the union bound, we have $L$ is not valid with probability at most $\frac{2s_2^2}{2^n}$. Hence follows the result. □

Now, since the two collision relations $\gamma_1$ and $\gamma_2$ defined earlier, are consistent with the Online linear mix function, $\mathsf{OLmix}$ (Proved in Lemma 2) we have the following corollary from Lemma 1 :

**Corollary 1.** $\#\{X : \mathit{coll}(X) = \gamma_1, \mathit{coll}(Y) = \gamma_2\} \geq 2^{ns_1}(1 - \frac{2s_2^2}{2^n})$

Now, for a fixed valid $L$, the conditional interpolation probability is

$$\sum_X \frac{\#\pi : \pi(MM) = X, \pi(CC) = Y}{\#\pi} \geq (1 - \frac{2s_2^2}{2^n}) \times 2^{-ns_2}.$$

So by multiplying the probability for validness of $L$ the proof of the proposition completes.

*Remark 1.* Note that, if we define $L$ from $E_K$ then we need to revise the proof of the Proposition 3. The revision is mainly by defining more internal bad events that some of the $\Pi$ inputs is 0 (the inputs are used to generate $L$ value). As this adds notational complexity and does not increase the order of advantage (except the constant factor will increase) we skip it for clarity throughout the paper.

**Proposition 4 (High interpolation probability of $\mathsf{EME}_{\mathsf{Lmix}}$).**

$\forall \tau \in V_{good}$,

$$\mathit{Pr}[\tau(A^{\mathit{EME}^{\Pi,\mathbf{L}}_{\mathit{Lmix}}}) = \tau] \geq (1 - \tfrac{4\sigma^2}{2^n}) \times \mathit{Pr}[\tau(A^{\$}) = \tau].$$

**Proof.** According to the definition of good view, $s_2 = \sigma$. Assuming that, the proof is identical to the previous one.

### 4.4 Neccsity of Lmix (or OLmix) to obtain PRP (or OLPRP) Security of $\mathsf{EME}_{\mathsf{mix}}$

Suppose, the mix function used is not $\mathsf{Lmix}$. Hence at least one entries of $B$ matrix is zero. Let $b_{ij}$ be the entry. Now we have the following equation for $Y[i]$

$$Y[i] = \sum_{k \neq j} b_{ik} X[k]$$

Since $Y[i]$ is indepentdent of $X[j]$, hence any two messages $M_1$ and $M_2$ (of same length) whose all blocks are same except the $j^{th}$ block (all other blocks are same) would yield $C_1[i] = C_2[i]$ - which breaks the full security of the construction. This shows the necessity of $\mathsf{Lmix}$ to obtain $PRP$ Security of the construction.

Similarly, the necessity of $\mathsf{OLmix}$ can be shown inorder to obtain $OLPRP$ Security of the construction. Suppose, the mix function used is not $\mathsf{OLmix}$. Hence at least one lower triangle entries of $B$ matrix is zero. Assume $b_{ij} = 0$, where $j \leq i$. Now we have the following equation for $Y[i]$

$$Y[i] = \sum_{\substack{k \neq j \\ j \leq i}} b_{ik} X[k]$$

Since $Y[i]$ is indepentdent of $X[j]$, hence any two messages $M_1$ and $M_2$ (of same length) whose all blocks are same except the $j^{th}$ block (all other blocks are same) would yield $C_1[i] = C_2[i]$ although $M_1[..i] \neq M_2[..i]$ as they differ in the $j^{th}$ block. Hence it breaks the online security.

## 5 SPRP Attack against $\mathsf{EME}_{\mathsf{mix}}$ for any underlying field

In this section, we show an SPRP attack against this construction. First, we revisit the attack that authors proposed in the paper [3] to prove that the construction with linear mix doesn't provide SPRP security. We give an adversary $A$ that attacks the mode, distinguishing it from a Pseudo random permutation and its inverse using only 4 queries.

**Revisiting SPRP Attack against $\mathsf{EME}_{\mathsf{mix}}$ for binary field :**

1. $A$ queries two messages $M_1 = (M_1[1], M_1[2], M_1[3])$ and $M_2 = (M_2[1], M_1[2], M_1[3])$. Let $C_1 = (C_1[1], C_1[2], C_1[3])$ and $C_2 = (C_2[1], C_2[2], C_2[3])$ are the responses, respectively .

2. Now, $A$ queries two ciphertexts $C_3 = (C_1[1], C_2[2], C_2[3])$ and $C_4 = (C_2[1], C_1[2], C_1[3])$. Let $M_3 = (M_3[1], M_3[2], M_3[3])$ and $M_4 = (M_4[1], M_4[2], M_4[3])$ are the respective responses.

3. If $M_3[1] \neq M_4[1]$ and $\forall j \geq 1,\ M_3[j] = M_4[j]$; then $A$ returns 1 (meaning the real). Else $A$ returns 0 (i.e. the random).

**Logic behind the Attack.** The main observation is that, $\forall i :\ Y_3[i]+Y_4[i] = Y_1[i]+Y_2[i]$ and as the underlying field is binary, hence addition of two elements is zero imply that the elements are same. The authors use this two property to mount the attack. The proof is given below :
$\forall j = 1, 2, 3$ we have,

$$\begin{aligned} X_3[j] + X_4[j] &= (a_{j1}Y_3[1] + a_{j2}Y_3[2] + a_{j3}Y_3[3]) + (a_{j1}Y_4[1] + a_{j2}Y_4[2] + a_{j3}Y_4[3]) \\ &= a_{j1}(Y_3[1] + Y_4[1]) + a_{j1}(Y_3[2] + Y_4[2]) + a_{j1}(Y_3[3] + Y_4[3]) \\ &= a_{j1}(Y_2[1] + Y_1[1]) + a_{j1}(Y_1[2] + Y_2[2]) + a_{j1}(Y_1[3] + Y_2[3]) \\ &= (a_{j1}Y_1[1] + a_{j2}Y_1[2] + a_{j3}Y_1[3]) + (a_{j1}Y_2[1] + a_{j2}Y_2[2] + a_{j3}Y_2[3]) \\ &= X_2[j] + X_1[j] \end{aligned}$$

Note that, in this attack, the authors uses the property of a binary field that the addition of two elements is zero implies both the values are same. Now, the question raise is what about the security of $\mathsf{ELE}_{\mathsf{mix}}$ when the underlying field is not binary. In the following two subsections, we prove that, $\mathsf{ELE}_{\mathsf{mix}}$ doesn't provide SPRP even in non-binary fields.

### 5.1 SPRP Attack against $\mathsf{EME}_{\mathsf{OLmix}}$

Here we show an SPRP attack against the EME construction that uses OLmix linear mixing.

1. $A$ queries four messages $M_1 = (M_1[1], M_1[2], M_1[3])$ and $M_2 = (M_1[1], M_2[2], M_1[3])$. Let $C_1 = (C_1[1], C_1[2], C_1[3])$ and $C_2 = (C_1[1], C_2[2], C_2[3])$ are the responses, respectively .

2. Now, $A$ queries two ciphertexts $C_3 = (C_3[1], C_1[2], C_1[3])$ and $C_4 = (C_3[1], C_2[2], C_2[3])$. Let $M_3 = (M_3[1], M_3[2], M_3[3])$ and $M_4 = (M_4[1], M_4[2], M_4[3])$ are the respective responses.

3. If $M_3[3] = M_4[3]$; then $A$ returns 1 (meaning the real). Else $A$ returns 0 (i.e. the random).

**Main idea why the attack works.** The main idea of the attack is that the value $M_3[3]$ is calculated using $C_1[2]$ and $C_1[3]$ both of whose values are dependent on the value $M_1[2]$. So, the value of $M_3[3]$ should have an influence of the value $M_3[3]$ but the effects of $M_1[2]$ via $C_1[2]$ and $C_1[3]$ cancels each other out implying the value of $M_1[2]$ has no effect in calculating $M_3[3]$. Using this observation, we mount the attack. The formal proof is as follows :

$$\begin{aligned}
X_3[3] &= a_{31}Y_2[1] + a_{32}Y_1[2] + a_{33}Y_1[3] \\
&= a_{31}Y_2[1] + a_{32}(b_{21}X_1[1] + b_{22}X_1[2]) \\
&\quad + a_{33}(b_{31}X_1[1] + b_{32}X_1[2] + b_{33}X_1[3]) \\
&= a_{31}Y_2[1] + (a_{32}b_{21} + a_{33}b_{31})X_1[1] + (a_{32}b_{22} + a_{33}b_{32})X_1[2] \\
&\quad + a_{33}b_{33}X_1[3] \\
&= a_{31}Y_2[1] + (a_{32}b_{21} + a_{33}b_{31})X_1[1] + (a_{32}b_{22} + a_{33}b_{32})X_2[2] \\
&\quad + a_{33}b_{33}X_1[3] \quad (\text{As } a_{32}b_{22} + a_{33}b_{32} = 0) \\
&= a_{31}Y_2[1] + a_{32}Y_2[2] + a_{33}Y_1[3] \\
&= X_4[3]
\end{aligned}$$

### 5.2 SPRP Attack against $\mathsf{EME}_{\mathsf{Lmix}}$.

The SPRP attack against the EME construction that uses Lmix as the linear mixing, is as follows :

1. $A$ queries four messages $M_1 = (M_1[1], M_1[2], M_1[3])$, $M_2 = (M_2[1], M_1[2], M_1[3])$, $M_3 = (M_1[1], M_3[2], M_1[3])$ and $M_4 = (M_2[1], M_3[2], M_1[3])$. Let $C_1 = (C_1[1], C_1[2], C_1[3])$, $C_2 = (C_2[1], C_2[2], C_2[3])$,$C_3 = (C_3[1], C_3[2], C_3[3])$ and $C_4 = (C_4[1], C_4[2], C_4[3])$ are the responses, respectively .

2. Now, $A$ queries two ciphertexts $C_5 = (C_2[1], C_1[2], C_1[3])$ and $C_6 = (C_4[1], C_3[2], C_3[3])$. Let $M_5 = (M_5[1], M_5[2], M_5[3])$ and $M_6 = (M_6[1], M_6[2], M_6[3])$ are the respective responses.

3. If $M_5[3] = M_6[3]$; then $A$ returns 1 (meaning the real). Else $A$ returns 0 (i.e. the random).

**Main Idea behind the attack.** The main idea of the attack is similar to the previous attack. Here, we observe that the value $M_5[3]$ is calculated using the values $C_2[1]$, $C_1[2]$ and $C_1[3]$ each of which are dependent on the value $M_1[2]$. So, the value of $M_5[3]$ should have an influence of the value $M_1[2]$ but the combined effects of $M_1[2]$ via $C_2[1]$, $C_1[2]$ and $C_1[3]$ cancels each other out implying the value of $M_1[2]$ has no effect in calculating $M_5[3]$. Using this observation, we mount the attack. The formal proof is as follows :

$$
\begin{aligned}
X_5[3] &= a_{31}Y_2[1] + a_{32}Y_1[2] + a_{33}Y_1[3] \\
&= a_{31}(b_{11}X_2[1] + b_{12}X_1[2] + b_{13}X_1[3]) + a_{32}(b_{21}X_1[1] + b_{22}X_1[2] + b_{23}X_1[3]) \\
&\quad + a_{33}(b_{31}X_1[1] + b_{32}X_1[2] + b_{33}X_1[3]) \\
&= a_{31}b_{11}X_2[1] + (a_{32}b_{21} + a_{33}b_{31})X_1[1] + (a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32})X_1[2] + \\
&\quad (a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33})X_1[3] \\
&= a_{31}b_{11}X_2[1] + (a_{32}b_{21} + a_{33}b_{31})X_1[1] + (a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32})X_3[2] + \\
&\quad (a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33})X_1[3] \quad (\text{As } a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} = 0\ ) \\
&= a_{31}(b_{11}X_2[1] + b_{12}X_3[2] + b_{13}X_1[3]) + a_{32}(b_{21}X_1[1] + b_{22}X_3[2] + b_{23}X_1[3]) \\
&\quad + a_{33}(b_{31}X_1[1] + b_{32}X_3[2] + b_{33}X_1[3]) \\
&= a_{31}Y_4[1] + a_{32}Y_3[2] + a_{33}Y_3[3] \\
&= X_6[3]
\end{aligned}
$$

## 6 Conclusion

In this paper, we characterized EME with linear mixing. In particular, we showed that, the necessary as well as sufficient condition for EME with linear mixing to have the PRP and OLPRP security, is to use Lmix and OLmix respectively, as the linear mixing. We also proved that EME with linear mixing doesn't provide the SPRP security even if the construction is extended to some non-binary fields. Thus our result is a guideline that, one can have an efficient PRP or OLPRP construction using EME with efficient linear mix having Lmix or OLmix property respectively and for SPRP security based on EME construction the mixing has to be non-linear.

## References


[1] Elena Andreeva, Andrey Bogdanov, Atul Luykx, Bart Mennink, Elmar W. Tischhauser and Kan Yasuda, *Parallelizable (authenticated) online ciphers* (2013), *Asiacrypt*, 2013. Citations in this document: §1.2.

[2] Nilanjan Datta and Mridul Nandi, *Misuse Resistant Parallel Authenticated Encryptions* (2013). URL: http://eprint.iacr.org/2013/767. Citations in this document: §1.2.

[3] Shai Halevi and Phillip Rogaway, *A Tweakable Enciphering Mode* **2729** (2003), 482–499, *CRYPTO*, Lecture Notes in Computer Science, 2003. Citations in this document: §1, §5.
[4] Michael Luby, Charles Rackoff, *How to construct pseudorandom permutations from pseudorandom functions*, SIAM Journal of Computing (1988), 373–386. Citations in this document: §1, §1.1.
[5] Jacques Patarin, *The "Coefficients H" Technique* **5381** (2009), 328–345, *Selected Areas in Cryptography*, Lecture Notes in Computer Science, 2009. Citations in this document: §1.3.
[6] Serge Vaudenay, *Decorrelation: A Theory for Block Cipher Security*, Journal of Cryptology (2003), 249–286.
[7] S. Halevi and P. Rogaway., *A Tweakable Enciphering Mode.* **2729** (2003), *CRYPTO*, Lecture Notes in Computer Science, 2003. Citations in this document: §1.
[8] — (no editor), *CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness*. URL: http://competitions.cr.yp.to/caesar.html.
[9] — (no editor), *DIAC : Directions in Authenticated Ciphers* (2012). URL: http://hyperelliptic.org/DIAC/.
[10] — (no editor), *DIAC : Directions in Authenticated Ciphers* (2013). URL: http://2013.diac.cr.yp.to/.