SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Construction of MDS block diffusion matrices for block ciphers and hash functions

ZHAO RuoXin[1,2*], ZHANG Rui[1*], LI YongQiang[1*] & WU BaoFeng[1*]

[1]*State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093*, China;*
[2]*University of Chinese Academy of Sciences, Beijing* 100101*, China*

Appendix A Preliminaries

Appendix A.1 Linear Algebra

In this paper, $\mathcal{M}_{s\times t}(F)$ usually denotes the set consisting of all the $(s \times t)$ matrices over the field F.

Let $\mathbb{F}_q$ (or $GF(q)$) be the finite field with q elements where q is a prime power and V be an n-dimensional $\mathbb{F}_q$-linear space. A mapping $L : V \to V$ is an $\mathbb{F}_q$-linear transformation over V if $L(u\alpha + v\beta) = uL(\alpha) + vL(\beta)$ for every $u, v \in \mathbb{F}_q$ and every $\alpha, \beta \in V$. From linear algebra, we know that there exists a bijection between the set consisting of all the $\mathbb{F}_q$-linear transformation over V and the set $\mathcal{M}_{n\times n}(\mathbb{F}_q)$ under a fixed basis of V. Furthermore, if we regard the two sets as two algebras, the bijection is an algebra isomorphism. Thus, in this paper, we identify every $\mathbb{F}_q$-linear transformation over V with a matrix in $\mathcal{M}_{n\times n}(\mathbb{F}_q)$. Specifically, if L is a matrix in $\mathcal{M}_{n\times n}(\mathbb{F}_q)$, the mapping which maps every row vector $\boldsymbol{x} \in \mathbb{F}_q^n$ to $\boldsymbol{x}L$ is an $\mathbb{F}_q$-linear transformation over $\mathbb{F}_q^n$ uniquely determined by L.

Let $\mathbb{F}_{q^n}$ be an extension of $\mathbb{F}_q$. Then $\mathbb{F}_{q^n}$ is an n-dimensional $\mathbb{F}_q$-linear space. Notice that multiplication with an element in $\mathbb{F}_{q^n}$ is a special $\mathbb{F}_q$-linear transformation over $\mathbb{F}_{q^n}$. More precisely, for every $\alpha \in \mathbb{F}_{q^n}$, the mapping $f : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ defined as $f(x) = \alpha x$ is an $\mathbb{F}_q$-linear transformation over $\mathbb{F}_{q^n}$.

For every vector $\boldsymbol{x} \in \mathbb{F}_q^m$, the Hamming weight of $\boldsymbol{x}$ is defined as the number of non-zero coordinates of $\boldsymbol{x}$ and is denoted by $w_H(\boldsymbol{x})$. Suppose $\boldsymbol{y} \in \mathbb{F}_q^{bn}$ for some positive integers b and n. We divide $\boldsymbol{y}$ into n segments, namely, $\boldsymbol{y} = (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n)$ where $\boldsymbol{y}_i \in \mathbb{F}_q^b$, $i = 1, \cdots, n$. Then each $\boldsymbol{y}_i$ is called a bundle of $\boldsymbol{y}$. The bundle weight of $\boldsymbol{y}$ is defined as the number of non-zero bundles of $\boldsymbol{y}$ and is denoted by $w_b(\boldsymbol{y})$. If $\boldsymbol{z} \in \mathbb{F}_q^{bn}$ is another vector, the bundle distance between $\boldsymbol{y}$ and $\boldsymbol{z}$ is defined as $w_b(\boldsymbol{y} - \boldsymbol{z})$ and denoted by $d_b(\boldsymbol{y}, \boldsymbol{z})$. Note that $w_b(\boldsymbol{y})$ and $w_H(\boldsymbol{y})$ are distinct in most cases.

Suppose $\boldsymbol{x} \in \mathbb{F}_q^{bn}$ be a row vector and $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_q)$ is a matrix. Let $\boldsymbol{y} = \boldsymbol{x}L$ be the image of $\boldsymbol{x}$ under the linear transformation L. From matrix theory, it is convenient to express the multiplication of $\boldsymbol{x}$ and L if we divide $\boldsymbol{x}$ into bundles and divide L into blocks. That is, we may write $\boldsymbol{x} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)$ where $\boldsymbol{x}_i \in \mathbb{F}_q^b$, $i = 1, \cdots, n$ and

$$L = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ L_{n,1} & L_{n,2} & \cdots & L_{n,n} \end{pmatrix}$$

where $L_{i,j} \in \mathcal{M}_{b\times b}(\mathbb{F}_q)$ is a matrix, $i, j = 1 \cdots, n$. Then $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2 \cdots, \boldsymbol{y}_n) = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)L$ where each $\boldsymbol{y}_j \in \mathbb{F}_q^b$ and $\boldsymbol{y}_j = \sum_{i=1}^n \boldsymbol{x}_i L_{i,j}$, $j = 1, \cdots, n$. Techniques dealing with block matrices play an important role in this paper.

As mentioned in Section 1, minimal polynomials of matrices play a significant role in this paper. So we are presenting some knowledge about minimal polynomials. Let F and E be two fields such that $F \subseteq E$ or $E \subseteq F$. For a square matrix $A \in \mathcal{M}_{b\times b}(E)$, a polynomial $f(x) \in F[x]$ is called an annihilator polynomial of A in $F[x]$ if $f(A) = O_b$ where O_b is the zero matrix in $\mathcal{M}_{b\times b}(F)$. For example, from Hamilton-Cayley theorem, we know that the characteristic polynomial of A is an annihilator polynomial of A in $E[x]$. A polynomial $g(x) \in F[x]$ is called the minimal polynomial of A in $F[x]$ if $g(x)$ is the monic annihilator polynomial of A in $F[x]$ with the lowest degree. The minimal polynomial of A is usually denoted by $m_A(x)$. In fact, the minimal polynomial of a matrix has some relation to its annihilator polynomials.

* Corresponding author (email: zhaoruoxin@iie.ac.cn, r-zhang@iie.ac.cn, liyongqiang@iie.ac.cn, wubaofeng@iie.ac.cn)

Proposition 1 ([7]). The minimal polynomial of a matrix A divides all the annihilator polynomials of it.

For two matrices $A, B \in \mathcal{M}_{b\times b}(F)$, we say that A is similar to B (or B is similar to A) if there exists a nonsingular matrix $P \in \mathcal{M}_{b\times b}(F)$ such that $P^{-1}AP = B$. An elementary property of minimal polynomial is stated in the following lemma.

Proposition 2 ([7]). Two similar matrices have the same minimal polynomial.

In matrix theory, there is a proposition useful to us.

Proposition 3 ([1]). The minimal polynomial and the characteristic polynomial of a matrix over a field F have the same irreducible factors in $F[x]$.

According to Proposition 1 and 3, we may test the factors of the characteristic polynomial of A one by one to seek the minimal polynomial of A. But along with the increase of the degree of characteristic polynomial, the amount of computation for this approach will skyrocket. Thus, for those matrices with large sizes, we need other methods to compute their minimal polynomials. For example, the following proposition brings us an effective approach.

Proposition 4 ([1]). Let $A : V \to V$ be linear. Suppose $W_1, \cdots, W_k$ are subspaces of V such that $V = W_1 + \cdots + W_k$, $A(W_i) \subseteq W_i$ for all i, and the restriction of A to W_i has minimal polynomial $m_i(x)$. Then the minimal polynomial of A on V is $\mathrm{lcm}(m_1, \cdots, m_k)$.

In Proposition 4, $\mathrm{lcm}(m_1, \cdots, m_k)$ denotes the least common multiple of $m_1, \cdots, m_k$. In [1], Proposition 4 leads to an constructive algorithm for computing the minimal polynomial of any square matrix $A \in \mathcal{M}_{n\times n}(E)$. Pick any column vector $v \neq \mathbf{0}$ in $V = E^n$ and consider the sequence of vectors $\{v, Av, A^2v, \cdots\}$. They span a subspace of V that is denoted by W, so $W = \{f(A)v : f(x) \in E[x]\}$. The nice feature of W is that $A(W) \subseteq W$, so A makes sense as a linear operator on W. To determine the minimal polynomial of A on W, find the smallest positive integer d such that the vectors $v, Av, \cdots, A^dv$ are linearly dependent. Since $v, Av, \cdots, A^{d-1}v$ are linearly independent, the linear relation

$$b_{d-1}A^{d-1}v + \cdots b_1Av + b_0v = \mathbf{0}$$

with $b_i \in E$, $i = 1, \cdots, d-1$ implies that $b_i = 0$, $i = 1, \cdots, d-1$. Hence for every nonzero polynomial $f(x) \in E[x]$ with degree less than d, $f(A)v \neq 0$, and then $f(A) \neq O_n$ as an operator on W where O_n denotes the zero matrix in $\mathcal{M}_{n\times n}(E)$, which means the minimal polynomial of A acting on W has degree at least d. There is a linear dependence relation on the set $v, Av, \cdots, A^dv$, and the coefficient of A^dv in the relation must be nonzero since the other vectors are linearly independent. We can make the coefficient of A^dv to be 1, say

$$A^dv + c_{d-1}A^{d-1}v + \cdots + c_1Av + c_0v = \mathbf{0}$$

where $c_i \in E$, $i = 0, 1, \cdots, d-1$. This tells us the polynomial

$$m(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1x + c_0$$

satisfies $m(A)v = \mathbf{0}$, so for every $f(x) \in E[x]$ we have $m(A)f(A)v = f(A)m(A)v = f(A)\mathbf{0} = \mathbf{0}$. Since every element in W is $f(A)v$ for some $f(x)$, so $m(A)$ annihilates all the elements in W. Thus $m(A)$ is just the minimal polynomial of A acting on W. Incidentally, this also shows $\dim W = d$ and W has basis $v, Av, \cdots, A^{d-1}v$. Set $W_1 = W$ and $m_1(x) = m(x)$. If $W_1 \neq V$, pick a column vector $v_2 \notin W_1$ and run through the same argument for the subspace W_2 of V spanned by the vectors $\{v_2, Av_2, A^2v_2, \cdots\}$ to get a minimal polynomial $m_2(x)$ for A on W_2. Since $v_2 \notin W_1$, $\dim(W_1 + W_2) > \dim W_1$. If $W_1 + W_2 \neq V$, proceed this procedure. Since V is finite-dimensional, eventually we will get a sequence of subspaces $W_1, W_2, \cdots, W_k$ where $A(W_i) \subseteq W_i$ for $i = 1, \cdots, k$ and $W_1 + \cdots + W_k = V$. Then the minimal polynomial of A on V is the least common multiple of $m_1(x), \cdots, m_k(x)$ from Proposition 4.

Appendix A.2 Diffusion Layers

The diffusion layers in block ciphers and hash functions are essentially $\mathbb{F}_2$-linear transformations, so sometime we just call them linear transformations or just diffusion matrices.

The branch numbers of diffusion layers are defined as following.

Definition 1 (Differential Branch Number). Let $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ be a diffusion matrix for certain positive integers b and n. The differential branch number of L is defined as

$$\mathcal{B}_d(L) = \min_{\boldsymbol{x}\in\mathbb{F}_2^{bn}, \boldsymbol{x}\neq\mathbf{0}} \{w_b(\boldsymbol{x}) + w_b(L(\boldsymbol{x}))\},$$

where each bundle of vectors in $\mathbb{F}_2^{bn}$ is in $\mathbb{F}_2^b$ and $L(\boldsymbol{x}) = \boldsymbol{x}L$ if we write $\boldsymbol{x}$ as a row vector in $\mathbb{F}_2^{bn}$.

Definition 2 (Linear Branch Number). Let $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ be a diffusion matrix for certain positive integers b and n. The linear branch number of L is defined as

$$\mathcal{B}_l(L) = \min_{\boldsymbol{x}\in\mathbb{F}_2^{bn}, \boldsymbol{x}\neq\mathbf{0}} \{w_b(\boldsymbol{x}) + w_b(L^T(\boldsymbol{x}))\},$$

where each bundle of vectors in $\mathbb{F}_2^{bn}$ is in $\mathbb{F}_2^b$ and L^T is the transposition of L and $L^T(\boldsymbol{x}) = \boldsymbol{x}L^T$ if we write $\boldsymbol{x}$ as a row vector in $\mathbb{F}_2^{bn}$.

The larger the branch numbers are, the stronger the diffusion layer is against differential and linear cryptanalyses. In fact, a diffusion matrix L can be related to a $\mathbb{F}_2$-linear code (refer to [2]). According to Singleton bound (see [5]), for the diffusion matrix L described in Definition 1, $\mathcal{B}_d(L) \leqslant n + 1$ and $\mathcal{B}_l(L) \leqslant n + 1$. The diffusion layers attaining this bound are called MDS and they are the optimal primitives in cryptosystems.

Definition 3 (MDS Diffusion Layer)**.** Let $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ be a diffusion matrix for certain positive integers b and n where b is the length of bundles. Then L is called a MDS diffusion layer if $\mathcal{B}_d(L) = n + 1$.

Let us recall some previous results. The proofs of these results are similar to those about the ordinary MDS linear codes. The result of [2] may be redescribed as the following proposition.

Proposition 5. Let $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ be a diffusion matrix for certain positive integers b and n where b is the length of bundles. Suppose L is divided into n^2 blocks such that

$$L = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ L_{n,1} & L_{n,2} & \cdots & L_{n,n} \end{pmatrix}$$

where $L_{i,j} \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$, $i, j = 1, \cdots, n$. Then L is MDS if and only if every submatrix of L consisting of some of these blocks is nonsingular.

Proposition 6 ([3])**.** A linear diffusion layer D has a maximum differential branch number if and only if it has a maximum linear branch number.

Appendix B Our Strategy for Constructing MDS Block Diffusion Matrices

In this section, we present our method for constructions a sort of MDS diffusion layers.

First of all, we state a lemma about block matrices that is often treated as an exercise in the textbooks of matrix theory. Because it is not very trivial and for completeness of this paper, we give the proof of this lemma in Appendix A.

Lemma 1. Let F be a field, $L \in \mathcal{M}_{bn\times bn}(F)$ be a block matrix for some positive integers b and n such that

$$L = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ L_{n,1} & L_{n,2} & \cdots & L_{n,n} \end{pmatrix}$$

where $L_{i,j} \in \mathcal{M}_{b\times b}(F)$, $i, j = 1, \cdots, n$ and they commute pairwise. Then

$$\det(L) = \det\left(\sum (-1)^{\tau(i_1 i_2 \cdots i_n)+\tau(j_1 j_2 \cdots j_n)} L_{i_1,j_1} L_{i_2,j_2} \cdots L_{i_n,j_n}\right), \tag{1}$$

where $\det(L)$ denotes the determinant of L, the sum on the right side consists of all the products of n blocks having distinct row indices and distinct column indices and a sign, $\tau(i_1 i_2 \cdots i_n)$ denotes the number of inverse-ordered pairs in the permutation $(i_1 i_2 \cdots i_n)$ where an inverse-ordered pair is a pair whose number on the left side is larger than its number on the right side. In other words, if we let

$$\det_s(L) = \sum (-1)^{\tau(i_1 i_2 \cdots i_n)+\tau(j_1 j_2 \cdots j_n)} L_{i_1,j_1} L_{i_2,j_2} \cdots L_{i_n,j_n}, \tag{2}$$

which is the determinant of the block matrix L if we regard all of its blocks $L_{i,j}$, $i, j = 1, \cdots, n$ as entries and regard L as a $(n \times n)$ matrix (we call $\det_s(L)$ the symbolic determinant of L), then $\det(L) = \det(\det_s(L))$.

Remark 1. Another expression of the determinant of L often arises in many papers, that is

$$\det(L) = \det\left(\sum_{\sigma\in S_n} (-1)^{\tau(\sigma(1)\sigma(2)\cdots\sigma(n))} L_{1,\sigma(1)} L_{2,\sigma(2)} \cdots L_{n,\sigma(n)}\right), \tag{17}$$

where S_n is the symmetric group on n elements. It is easy to see that equation (17) is just a special case of equation (1), because in equation (17) the permutation of row indices is $(12\cdots n)$ and $\tau(12\cdots n) = 0$.

Let $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ be a diffusion matrix for certain positive integers b and n where b is the length of bundles, and L be divided into n^2 blocks such that

$$L = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\ L_{2,1} & L_{2,2} & \cdots & L_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ L_{n,1} & L_{n,2} & \cdots & L_{n,n} \end{pmatrix}$$

where $L_{i,j} \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$, $i, j = 1, \cdots, n$. From Proposition 5, in order to determine whether L is MDS, we need to check the determinants of all the submatrices of L composed of some of these blocks. If one wants to calculate these determinants by Lemma 1, all of the blocks $L_{i,j}$, $i, j = 1, 2, \cdots, n$, need to commute pairwise. But pairwise commutativity is such a high requirement that most sets of matrices cannot meet it. Therefore, in this paper, we focus on a specific sort of matrices whose blocks are all polynomials of certain block. In detail, we only consider such a situation when each block $L_{i,j}$, $i, j = 1, 2, \cdots, n$ of the diffusion matrices is a polynomial in certain $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$. In [8] and [9], M. Sajadieh et al. and S. Wu et al. already talked about this situation. In comparison with their strategies, ours has such advantages:

• They just discussed the recursive diffusion layers, while we consider more general diffusion layers including recursive ones.

• They just found out the conditions for MDS diffusion layers but didn't point out how to construct the block A (denoted by L in their papers) generically, while we explicitly figure out the conditions for MDS diffusion matrices as well as A itself.

• We use some techniques to increase the efficiency of search algorithms for MDS matrices.

As mentioned in the previous paragraphs, we will focus on the block diffusion layers whose blocks are all polynomials in certain block $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$. Let each block $L_{i,j} = f_{i,j}(A)$, where $f_{i,j}(x) \in \mathbb{F}_2[x]$, $i, j = 1, 2, \cdots, n$. In this paper, we call the polynomial matrix

$$\begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x])$$

the external matrix of L. Although every entry of external matrices is a polynomial in $\mathbb{F}_2[x]$, actually we only need to consider a small part of $\mathbb{F}_2[x]$ because A has a minimal polynomial. In detail, for every polynomial $f(x) \in \mathbb{F}_2[x]$, we can get the remainder $r(x)$ of it by modulo $m_A(x) \in \mathbb{F}_2[x]$. Then obviously $f(A) = r(A)$. Thus it is enough to consider the entries whose degree is smaller than $\deg(m_A)$. For the sake of MDS property, we need to check every determinant of the submatrices consisting of some of the blocks of the diffusion matrix. Of course, we can conventionally calculate these determinants. Alternatively, we may calculate them with Lemma 1 because all of the blocks are obviously pairwise commutative. For example, to calculate the determinant of L, we may calculate the symbolic determinant $\det_s(L)$ firstly. And then calculate $\det(\det_s(L))$. Note that $\det_s(L)$ is also a polynomial in A. Actually, the symbolic determinant of every submatrix consisting of those blocks is a polynomial in A. This is a clue for us. Actually, if we know the minimal polynomial of A, there is a more efficient technique to determine whether such submatrices are nonsingular. Let us give the following lemma.

Lemma 2. Let F be a field, $A \in \mathcal{M}_{b\times b}(F)$, $m_A(x)$ be the minimal polynomial of A in the polynomial ring $F[x]$, $g(x) \in F[x]$. Then $\det(g(A)) \neq 0$ if and only if $\mathrm{GCD}(g(x), m_A(x)) = 1$, where $\mathrm{GCD}(g(x), m_A(x))$ denotes the greatest common divisor of $g(x)$ and $m_A(x)$.

Proof. To begin with, if the greatest common divisor of a family of polynomials is equal to 1, we say they are coprime.

Suppose $g(A)$ is nonsingular. Assume $\mathrm{GCD}(g(x), m_A(x)) = d(x)$ and $\deg(d) > 1$. Then there exists $u(x), v(x) \in F[x]$ such that $g(x) = u(x)d(x)$, $m_A(x) = v(x)d(x)$. Consequently, we have $g(A) = u(A)d(A)$, $m_A(A) = v(A)d(A)$. From $O_b = m_A(A) = v(A)d(A)$ and $\deg(v) < \deg(m_A)$, we get $v(A) \neq O_b$. And then we get $d(A)$ is singular, otherwise $v(A)$ would be equal to O_b. Because $g(A) = u(A)d(A)$ and $g(A)$ is nonsingular, we get $d(A)$ is nonsingular. A contradiction! Thus, $\mathrm{GCD}(g(x), m_A(x))$ must be 1.

Conversely, suppose $\mathrm{GCD}(g(x), m_A(x)) = 1$. Then there exists $u(x), v(x) \in F[x]$ such that $g(x)u(x) + m_A(x)v(x) = 1$. If we assign $x = A$, we get $g(A)u(A) = I_b$. Thus $g(A)$ is nonsingular.

As mentioned before, the symbolic determinant of every submatrix consisting of those blocks of the diffusion matrix L is a polynomial in A. From Lemma 2, instead of calculating the determinant of every such submatrix, we present a new technique: to determine whether such a submatrix is nonsingular, we merely need to calculate the symbolic determinant and check whether the symbolic determinant (treated as a polynomial in x) is coprime with $m_A(x)$. It is faster than calculating the determinants of diffusion matrices directly. However, one should note that in order to exploit this technique, we need to know the minimal polynomial of A in advance. Remember that our goal is to definitely obtain a series of MDS diffusion layers which requires us to clearly figure out the building block A. Thus we need a matrix $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ and its minimal polynomial $m_A(x) \in \mathbb{F}_2[x]$ as well.

To show our guideline and avoid getting into the complicated situation too early, we merely consider a special case in this section, namely, when the minimal polynomial of the block A is irreducible in $\mathbb{F}_2[x]$, and leave the general case to Section 4.

When the minimal polynomial of block A is irreducible in $\mathbb{F}_2[x]$, it is obvious that a polynomial $f(x) \in \mathbb{F}_2[x]$ is coprime with $m_A(x)$ if and only if $f(x) \not\equiv 0 \mod m_A(x)$. Hence, we get an easier way to check whether a polynomial in A is nonsingular: for a polynomial $f(x) \in \mathbb{F}_2[x]$, $f(A)$ is nonsingular if and only if $f(x) \not\equiv 0 \mod m_A(x)$. From the above statement, for a given block $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ whose minimal polynomial $m_A(x)$ is irreducible in $\mathbb{F}_2[x]$, the external matrices of MDS diffusion matrices L are just determined by $m_A(x)$ but not by A itself or b. More Specifically, if $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ and $A' \in \mathcal{M}_{b'\times b'}(\mathbb{F}_2)$ are two blocks having the same irreducible minimal polynomial and $H(x) \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x])$ is an polynomial matrix, $H(A)$ is MDS if and only if $H(A')$ is MDS.

Appendix C Our Strategy for a More Generalized Case

In Section 3, we assumed that the minimal polynomial of block A in $\mathbb{F}_2[x]$ is irreducible. But, in fact, it is not true for most matrices in $\mathcal{M}_{b\times b}(\mathbb{F}_2)$ because matrix rings contain zero divisors (refer to [6], page 573). How many matrices in $\mathcal{M}_{b\times b}(\mathbb{F}_2)$ are there whose minimal polynomials are irreducible in $\mathbb{F}_2[x]$? The quantity is related to the finite field $\mathbb{F}_{2^b}$ because every element in $\mathbb{F}_{2^b}$ can be represented by a matrix in $\mathcal{M}_{b\times b}(\mathbb{F}_2)$ (refer to [4], Chapter 2, Section 5). Let us give the following lemma.

Lemma 3. Let $A \in \mathcal{M}_{b\times b}(\mathbb{F}_q)$, $m_A(x)$ be the minimal polynomial of A in $\mathbb{F}_q[x]$. Then $m_A(x)$ is irreducible in $\mathbb{F}_q[x]$ if and only if $A \in \mathbb{F}_{q^b}$.

Proof. Let $f(x)$ denote the characteristic polynomial of A. Then $\deg(f) = b$ where $\deg(f)$ denotes the degree of $f(x)$.

Suppose $m_A(x)$ be irreducible in $\mathbb{F}_q[x]$. Let $\deg(m_A) = d$ and $\mathbb{F}_q(\alpha)$ denote the smallest extension field of $\mathbb{F}_q$ that includes α. Then, from field theory, we get $\mathbb{F}_q(A) \cong \mathbb{F}_q[x]/< m_A(x) > = \mathbb{F}_{q^d}$ where "$\cong$" means "isomorphic to" and $< m_A(x) >$ is the ideal generated by $m_A(x)$. Because $f(x)$ is the characteristic polynomial of A, $f(A) = O_b$ from Hamilton-Cayley theorem. Then $m_A(x) \mid f(x)$ from Proposition 1. Meanwhile, according to Proposition 3, $f(x)$ is necessarily a power of $m_A(x)$. Consequently, $d \mid b$. And it is followed by $\mathbb{F}_{q^d} \subseteq \mathbb{F}_{q^b}$. Thus, $A \in \mathbb{F}_{q^d} \subseteq \mathbb{F}_{q^b}$.

Conversely, suppose $A \in \mathbb{F}_{q^b}$. Assume $m_A(x)$ is reducible in $\mathbb{F}_q[x]$. Then there exists two polynomials $u(x), v(x) \in \mathbb{F}_q[x]$ such that $m_A(x) = u(x)v(x)$ and $\deg(u) < \deg(m_A)$ and $\deg(u) < \deg(m_A)$. Because $m_A(x)$ is the minimal polynomial of A, $u(A) \neq 0$ and $v(A) \neq 0$. But $u(A)v(A) = m_A(A) = 0$. It is a contradiction since any field does not contain zero divisors. Thus $m_A(x)$ must be irreducible in $\mathbb{F}_q[x]$.

From Lemma 3, the number of matrices in $\mathcal{M}_{b\times b}(\mathbb{F}_q)$ whose minimal polynomials are irreducible in $\mathbb{F}_q[x]$ is q^b, while the cardinality of $\mathcal{M}_{b\times b}(\mathbb{F}_q)$ is q^{b^2}. The proportion is too small. If we only focus on those blocks whose minimal polynomials are irreducible in $\mathbb{F}_2[x]$, we will miss a large amount of MDS candidates. Therefore, in this section, we will remove this condition and consider a more generalized case: the minimal polynomial of block A is reducible in $\mathbb{F}_2[x]$. In this case, we may get the standard factorization of $m_A(x)$ by Berlekamp's algorithm ([4]). Suppose $m_A(x) = p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$ is the standard factorization of $m_A(x)$ where $p_1(x), \cdots, p_s(x)$ are distinct irreducible polynomials in $\mathbb{F}_2[x]$. From Lemma 2, for a polynomial $f(x) \in \mathbb{F}_2[x]$, $f(A)$ is nonsingular if and only if $\mathrm{GCD}(f(x), m_A(x)) = 1$. Then, in this case, $\mathrm{GCD}(f(x), m_A(x)) = 1$ if and only if $\mathrm{GCD}(f(x), p_i(x)) = 1$ for $i = 1, \cdots, s$. Moreover, because each $p_i(x)$ is irreducible in $\mathbb{F}_2[x]$, we only need to check whether $f(x) \equiv 0 \mod p_i(x)$ for $i = 1, \cdots, s$. From an algebraic viewpoint, for a polynomial $f(x) \in \mathbb{F}_2[x]$, $f(A)$ is a nonsingular matrix in $\mathcal{M}_{b\times b}(\mathbb{F}_2)$ if and only if $f(x)$ is an invertible element in the fields $\mathbb{F}_2[x]/< p_i(x) >$ for $i = 1, \cdots, s$. Therefore, for an external matrix

$$H(x) = \begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}\left(\mathbb{F}_2[x]/< m_A(x) >\right),$$

if we want to argue whether $H(A)$ is an MDS block matrix with block size $(b \times b)$, what we need to do is just to regard $H(x)$ as a matrix over $\mathbb{F}_2[x]/< p_i(x) >$, $i = 1, \cdots, s$ and check whether it is MDS over these field respectively. From the above statement and Lemma 2, we have the following theorem.

Theorem 1. Let $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ with the minimal polynomial $m_A(x) \in \mathbb{F}_2[x]$. Suppose $m_A(x)$ has the standard factorization $m_A(x) = p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$ where $p_1(x), \cdots, p_s(x)$ are distinct irreducible polynomials in $\mathbb{F}_2[x]$ and $e_1, \cdots, e_s$ are positive integers. Let

$$H(x) = \begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x]).$$

Then the following four statements are equivalent.

1. $H(A)$ is an MDS block matrix with block size $(b \times b)$.
2. Every minor determinant of $H(x)$ is coprime with $m_A(x)$.
3. Every minor determinant of $H(x)$ is coprime with $p_i(x)$ for $i = 1, \cdots, n$.
4. Every minor determinant of $H(x)$ is not congruent to 0 modulo $p_i(x)$ for $i = 1, \cdots, n$.

According to the above discussion and statements in Section 3, for a given block $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$, the external matrices of MDS diffusion matrices $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ are absolutely determined by $m_A(x)$ but not by A itself or b, no matter whether $m_A(x)$ is reducible or not. We formally state this conclusion in the following theorem.

Theorem 2. Let $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ and $A' \in \mathcal{M}_{b'\times b'}(\mathbb{F}_2)$ be two matrices having the same minimal polynomial. Suppose

$$H(x) = \begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x]).$$

Then $H(A)$ is MDS if and only if $H(A')$ is MDS.

For a matrix $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ and a polynomial matrix

$$H(x) = \begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x]),$$

let us think about two kinds of elementary operations on $H(x)$, namely, interchanging two rows (columns) and multiplying a row (column) with a polynomial $g(x) \in \mathbb{F}_2[x]$ coprime to $m_A(x)$. If we obtain another polynomial matrix $H'(x) \in$

$\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$ from $H(x)$ via interchanging two rows (columns) of $H(x)$, according to the properties of determinants, we know every minor determinant $D'(x)$ of $H'(x)$ is equal to certain minor determinant $D(x)$ of $H(x)$ multiplied by 1 or -1. If $D(x)$ is coprime to $m_A(x)$, $D'(x)$ is coprime to $m_A(x)$ obviously. Thus interchanging two rows (columns) of $H(x)$ does not change MDS property of $H(A)$. Likewise, if we multiply a row (column) of $H(x)$ with a polynomial $g(x) \in \mathbb{F}_2[x]$ coprime to $m_A(x)$, every minor determinant $D'(x)$ of obtained polynomial matrix $H'(x)$ will be equal to certain minor determinant $D(x)$ of $H(x)$ multiplied by $g(x)$. If $D(x)$ is coprime to $m_A(x)$, $D'(x) = D(x)g(x)$ must be coprime to $m_A(x)$ since $g(x)$ is also coprime to $m_A(x)$. So multiplying a row (column) with a polynomial $g(x) \in \mathbb{F}_2[x]$ coprime to $m_A(x)$ does not change MDS property of $H(A)$ either. By contrast, the third kind of elementary operation on matrices, namely, adding a row (column) multiplied by a polynomial to another row (column) cannot retain MDS property of $H(x)$, because it might make some entries become zero. However, we discover another operation on $H(x)$ that is slightly similar to the third kind of elementary operation mentioned above and can retain MDS property of $H(A)$. Let us clarify this kind of operation in the following theorem.

Theorem 3. Let A be a matrix in $\mathcal{M}_{b\times b}(\mathbb{F}_2)$ with the minimal polynomial $m_A(x) \in \mathbb{F}_2[x]$. Suppose the standard factorization of $m_A(x)$ in $\mathbb{F}_2[x]$ is $m_A(x) = p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$, where $p_1(x), \cdots, p_s(x)$ are distinct irreducible polynomials in $\mathbb{F}_2[x]$ and $e_1, \cdots, e_s$ are positive integers. Let $g(x) = p_1(x)p_2(x)\cdots p_s(x)$. Let

$$H(x) = \begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x])$$

and

$$H'(x) = H(x) + g(x)\begin{pmatrix} h_{1,1}(x) & h_{1,2}(x) & \cdots & h_{1,n}(x) \\ h_{2,1}(x) & h_{2,2}(x) & \cdots & h_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ h_{n,1}(x) & h_{n,2}(x) & \cdots & h_{n,n}(x) \end{pmatrix},$$

where $h_{i,j}(x) \in \mathbb{F}_2[x]$ for $i = 1, \cdots, n$, $j = 1, \cdots, n$. Then $H(A)$ is MDS if and only if $H'(A)$ is MDS.

Proof. Obviously, the transformation from $H(x)$ to $H'(x)$ is invertible. So we only need to prove MDS property of $H(A)$ implies MDS property of $H'(A)$.

Suppose $H(A)$ is MDS. According to Theorem 1, what we need to do is to prove every minor determinant of $H'(x)$ is coprime to $p_i(x)$ for $i = 1, \cdots, s$. Without loss of generality, let us think of a square submatrix $M'(x)$ of $H'(x)$ obtained by choosing the i-th rows for $i = 1, \cdots, m$ and the j-th columns for $j = 1, \cdots, m$ of $H'(x)$, where m is a positive integer and $m \leqslant n$. Then

$$\det(M'(x)) = \begin{vmatrix} f_{1,1}(x) + g(x)h_{1,1}(x) & f_{1,2}(x) + g(x)h_{1,2}(x) & \cdots & f_{1,m}(x) + g(x)h_{1,m}(x) \\ f_{2,1}(x) + g(x)h_{2,1}(x) & f_{2,2}(x) + g(x)h_{2,2}(x) & \cdots & f_{2,m}(x) + g(x)h_{2,m}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{m,1}(x) + g(x)h_{m,1}(x) & f_{m,2}(x) + g(x)h_{m,2}(x) & \cdots & f_{m,m}(x) + g(x)h_{m,m}(x) \end{vmatrix}.$$

According to the properties of determinant, we may write $\det(M'(x))$ as the sum of a series of m-order determinants. More specifically, for every $j = 1, \cdots, m$, we can split the j-th column of $\det(M'(x))$ into two columns

$$\begin{pmatrix} f_{1,j}(x) \\ f_{2,j}(x) \\ \vdots \\ f_{m,j}(x) \end{pmatrix} \text{ and } \begin{pmatrix} g(x)h_{1,j}(x) \\ g(x)h_{2,j}(x) \\ \vdots \\ g(x)h_{m,j}(x) \end{pmatrix}.$$

Finally, $\det(M'(x))$ can be written as the sum of 2^m determinants. For instance, one of these determinants is

$$\begin{vmatrix} g(x)h_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,m}(x) \\ g(x)h_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,m}(x) \\ \cdots & \cdots & \cdots & \cdots \\ g(x)h_{m,1}(x) & f_{m,2}(x) & \cdots & f_{m,m}(x) \end{vmatrix} = g(x)\begin{vmatrix} h_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,m}(x) \\ h_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,m}(x) \\ \cdots & \cdots & \cdots & \cdots \\ h_{m,1}(x) & f_{m,2}(x) & \cdots & f_{m,m}(x) \end{vmatrix}.$$

Obviously, all of these 2^m determinants are multiples of $g(x)$ expect one, namely,

$$\begin{vmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,m}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,m}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{m,1}(x) & f_{m,2}(x) & \cdots & f_{m,m}(x) \end{vmatrix}.$$

Let $M(x)$ denote the matrix of this determinant. Then $\det(M'(x)) = \det(M(x)) + g(x)q(x)$, where $q(x) \in \mathbb{F}_2[x]$. Since $H(A)$ is MDS, $\det(M(x))$ is coprime to $p_i(x)$ for $i = 1, \cdots, s$. Thus, $\det(M'(x))$ is also coprime to $p_i(x)$ for $i = 1, \cdots, s$ because $g(x)$ is a multiple of $p_i(x)$ for $i = 1, \cdots, s$. Similarly, every minor determinant of $H'(x)$ is coprime to $p_i(x)$ for $i = 1, \cdots, s$. Therefore, $H'(A)$ is MDS according to Theorem 1.

Remark 2. Theorem 3 gives us an approach to construct new MDS diffusion matrices from a fixed MDS matrix. Obviously, it is useless to the case when $m_A(x)$ only has simple factors (including the case when $m_A(x)$ is irreducible in $\mathbb{F}_2[x]$). In this case, $g(x) = m_A(x)$. Then for every entry $f_{i,j}(x)$ of $H(x)$, $f_{i,j}(A) + g(A)h_{i,j}(A)$ has no difference from $f_{i,j}(A)$. However, it does make sense when $m_A(x)$ has multiple factors. In this case, the approach coming from Theorem 3 can give us at least $2^l - 1$ extra options for every entry of the external matrix of an MDS block matrix, where $l = \deg(m_A) - \deg(g)$. In detail, if $H(A)$ is MDS, for every entry $f_{i,j}(x)$ of the external matrix of $H(A)$, we may randomly pick a polynomial $h_{i,j} \in \mathbb{F}_2[x]$ such that $\deg(\mathrm{h_{i,j}}) < \deg(\mathrm{m_A}) - \deg(\mathrm{g})$ and substitute $f_{i,j}(x) + g(x)h_{i,j}(x)$ for $f_{i,j}(x)$. This kind of operation on $H(x)$ does not alter MDS property of $H(A)$ and it has further advantage. Let us define a binary relation γ on the set $\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$. For two matrices $H(x)$ and $H'(x)$ in $\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$, $H(x)$ is γ-related to $H'(x)$ if there exists a matrix

$$P(x) = \begin{pmatrix} h_{1,1}(x) & h_{1,2}(x) & \cdots & h_{1,n}(x) \\ h_{2,1}(x) & h_{2,2}(x) & \cdots & h_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ h_{n,1}(x) & h_{n,2}(x) & \cdots & h_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x])$$

such that $H'(x) = H(x) + g(x)P(x)$. It is easy to verify γ is an equivalence relation (a relation holding reflexivity, symmetry and transitivity). So we can partition $\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$ into equivalence classes by γ. And according to Theorem 3, if one polynomial matrix $H(x)$ makes $H(A)$ MDS, every polynomial matrix $H'(x)$ γ-related to $H(x)$ makes $H'(A)$ MDS. In other words, MDS property is an invariant on every equivalence class obtained from γ. Besides, from the definition of the relation γ, it is not hard to calculate the numbers of equivalence classes obtained from it. We merely need to consider the entries of external matrices as the residues with respect to modulo $g(x)$. Thus, there are totally $2^{\deg(g)\cdot n^2}$ equivalence classes obtained from γ. If we restrict the degree of entries of external matrices less than $\deg(m_A)$ (regard every entry as a residue modulo $m_A(x)$), the cardinality of every equivalence class is $2^{l\cdot n^2}$ where $l = \deg(m_A) - \deg(g)$. Therefore, if we want to search for all the external matrices of MDS matrices (or a part of them), we may only take the representatives of the equivalence classes obtained form γ into account. And this approach will greatly reduce the amount of search when $2^{l\cdot n^2}$ is large. More importantly, according to the definition of γ, it is convenient to generate a whole equivalence class from a representative of it. Another important metric for diffusion matrices is implementation efficiency. Fundamentally, the less nonzero entries a diffusion matrix has, the more efficient its implementation is. As the Hamming weight of a sequence, we may extend the notion of Hamming weight to matrices. For a matrix $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$, we define its Hamming weight $w_H(L)$ as the number of its nonzero entries. Then for a matrix $A \in \mathcal{M}_{b\times b}(\mathbb{F}_2)$ and a polynomial matrix

$$H(x) = \begin{pmatrix} f_{1,1}(x) & f_{1,2}(x) & \cdots & f_{1,n}(x) \\ f_{2,1}(x) & f_{2,2}(x) & \cdots & f_{2,n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_{n,1}(x) & f_{n,2}(x) & \cdots & f_{n,n}(x) \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x]),$$

obviously $w_H(H(A)) = \sum_{i,j=1}^{n} w_H(f_{i,j}(A))$. So, if we need a diffusion matrix with Hamming weight as small as possible, we may manage to reduce the Hamming weight of each block $f_{i,j}(A)$ respectively for $i, j = 1, \cdots, n$. With respect to a equivalence class obtained from γ, we can choose the one with the smallest Hamming weight from

$$\{f_{i,j}(A) + g(A)h_{i,j}(A) \mid h_{i,j}(x) \in \mathbb{F}_2[x], \deg(h) < \deg(m_A) - \deg(g)\}$$

for each (i, j), and then we will get the most efficient diffusion matrix in this equivalence class.

From Theorem 3 and Remark 2, we give the following corollary.

Corollary 1. Let A be a matrix in $\mathcal{M}_{b\times b}(\mathbb{F}_2)$ with the minimal polynomial $m_A(x) \in \mathbb{F}_2[x]$. Suppose the standard factorization of $m_A(x)$ in $\mathbb{F}_2[x]$ is $m_A(x) = p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$, where $p_1(x), \cdots, p_s(x)$ are distinct irreducible polynomials in $\mathbb{F}_2[x]$ and $e_1, \cdots, e_s$ are positive integers. Let $g(x) = p_1(x)p_2(x)\cdots p_s(x)$. Let γ be the binary relation defined in Remark 2 on the set $\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$. And let $MDSEM_{n\times n}(A)$ denote the set $\{H(x) \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x]) \mid H(A)\ is\ MDS\}$. Then γ is an equivalence relation on the set $\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$ and partition $\mathcal{M}_{n\times n}(\mathbb{F}_2[x])$ into $2^{\deg(g)\cdot n^2}$ equivalence classes. Moreover, $MDSEM_{n\times n}(A)$ is the union of some of the equivalence classes obtained from γ.

In summary of the above statements, now we present Algorithm 1 that can find out all the MDS diffusion matrices L such that

- $L \in \mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ for given parameters b and n;
- L can be divided to n^2 blocks and each block $L_{i,j} \in M_{b\times b}(\mathbb{F}_2)$ is a polynomial in a given block $A \in M_{b\times b}(\mathbb{F}_2)$.

Appendix C.1 Some Experimental Results

We have conducted some experiments and found a series of external matrices of some types of MDS matrices. In this subsection, we merely describe one experiment about searching for recursive MDS block matrices.

We conduct the experiment in this paper with Magma V2.12-16 on a computer with an Intel Core i5-2400@3.10GHz CPU and DDR3, 4GBytes, 665.1MHz RAM. Besides, its operating system is 32-bit Windows 7 Professional.

Algorithm C1 Search for MDS Diffusion Matrices

Require: two integers $b, n \in \mathbb{Z}^+$, a matrix $A \in M_{b\times b}(\mathbb{F}_2)$ together with its minimal polynomial $m(x) \in \mathbb{F}_2[x]$, $m(x)$'s standard factorization $m(x) = p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}$ in $\mathbb{F}_2[x]$, $g(x) = p_1(x)p_2(x)\cdots p_s(x)$.

Ensure: some polynomial matrices, an integer k.

define an integer k and $k \leftarrow 0$;
$d := \deg(g)$;
define a set $PS(d) := \{h(x) \in \mathbb{F}_2[x] \mid \deg(h) < d, \mathrm{GCD}(h(x), p_i(x)) = 1, \forall i = 1, \cdots, s\}$;
define a matrix $L \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x])$ and $L \leftarrow O_n$;
define an integer r and $r \leftarrow 0$;
define $f_i(x) \in \mathbb{F}_2[x]/ < p_i(x) >$ and $f_i(x) \leftarrow 0$, $i = 1, \cdots, s$;
print "The $(n \times n)$-size external matrices of MDS diffusion matrices in $\mathcal{M}_{bn\times bn}(\mathbb{F}_2)$ are:";
for $L \in \mathcal{M}_{n\times n}(PS(d))$ **do**
  for $i = 1, \cdots, s$ **do**
    turn L into $L_i \in \mathcal{M}_{n\times n}(\mathbb{F}_2[x]/ < p_i(x) >)$ by a ring homomorphism $\eta : \mathbb{F}_2[x] \to \mathbb{F}_2[x]/ < p_i(x) >$ such that $\eta(h(x)) = h(x) + < p_i(x) >$;
    $r \leftarrow n$;
    while $r \geqslant 2$ **do**
      define a matrix $B \in \mathcal{M}_{r\times r}(\mathbb{F}_2[x]/ < p_i(x) >)$ and $B \leftarrow O_r$;
      for B runs over all the $(r \times r)-$size submatrices of L_i **do**
        $f_i(x) \leftarrow \det(B)$;
        if $f_i(x) = 0$ **then**
          goto Step 34;
        else
          if $r \geqslant 3$ **then**
            compute B^{-1};
            for $f_i(x)$ runs over all the entries of B^{-1} **do**
              if $f_i(x) = 0$ **then**
                goto Step 34;
              end if
            end for
          end if
        end if
      end for
      $r \leftarrow r - 2$;
    end while
  end for
  print L;
  $k \leftarrow k + 1$;
  switch to the next L;
end for
print "where $x = A$.";
print "There are k such MDS diffusion matrices.".

In the experiment we aim to definitely find recursive MDS block matrices with our strategy described in Section 4. We choose parameters $b = 8$, $n = 4$ and the block

$$A = \begin{pmatrix} 0&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&1 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \end{pmatrix} \in \mathcal{M}_{8\times 8}(\mathbb{F}_2)$$

with the minimal polynomial $m_A(y) = (y^4 + y + 1)^2 \in \mathbb{F}_2[y]$. Then $g(y) = y^4 + y + 1$. We suppose

$$B(y) = \begin{pmatrix} 0&0&0&B_1(y) \\ 1&0&0&B_2(y) \\ 0&1&0&B_3(y) \\ 0&0&1&B_4(y) \end{pmatrix} \in \mathcal{M}_{4\times 4}(\mathbb{F}_2[y])$$

and write $B(y) = [B_1(y), B_2(y), B_3(y), B_4(y)]$ for simplicity. We choose 4 as the iteration number of diffusion layer, then the external matrix of diffusion matrix is $H(y) = B^4(y)$. In this case, we say $H'(y) = (B'(y))^4$ is γ-equivalent to $H(y)$ if $B_i(y) \equiv B'_i(y) \mod g(y)$ for $i = 1, 2, 3, 4$, where $B'(y) = [B'_1(y), B'_2(y), B'_3(y), B'_4(y)]$. According to Theorem 3 and Remark 2, we merely need to investigate the case when $\deg(B_i) < 4$ for $i = 1, \cdots, 4$. The equivalence class containing $H(y)$ consists of all the matrices

$$[B_1(y) + h_1(y)g(y), B_2(y) + h_2(y)g(y), B_3(y) + h_3(y)g(y), B_4(y) + h_4(y)g(y)]^4$$

for any $h_1(y)$,$h_2(y)$,$h_3(y)$,$h_4(y)$ in $\mathbb{F}_2[y]$ whose degrees are less than 4. Therefore, the cardinality of each equivalence class is $(2^4)^4 = 2^{16}$.

After running a series of Magma codes, we get 5820 external matrices of recursive MDS matrices in 9.516 seconds. We list part of the round matrices (namely, $B_i(y)$s) of these external matrices in Appendix B. Note that these matrices are just representatives with respect to γ. Thus, we actually get 5820×2^{16} external matrices of recursive MDS matrices.

From Proposition 2 and Theorem 2, if we assign the matrices similar to A to y in these external matrices, we will get more than 5820×2^{16} recursive MDS matrices easily.

In Remark 2, we talked about the implementation efficiency of diffusion layers. For the specific case of recursive diffusion layer, the implementation efficiency may be measured by the total Hamming weight of recursive coefficients. In this experiment, it is the sum of Hamming weights of $B_1(A), B_2(A), B_3(A), B_4(A)$. One may optimize the implementation efficiency with the method mentioned in Remark 2. But it needs much extra computation. So in practice, we often choose an original block A with low Hamming weight and recursive coefficients containing few monomials in A. For instance, in this experiment we choose an optimal block A that costs only 1 XOR gate. In addition, A^2, A^3 cost 2, 3 XOR gates respectively. After the experiment, we do find some extremely efficient diffusion layers. For example, $[A, I_8, A^2, I_8]$ and $[A, I_8, I_8, A^2]$ only cost 3 XOR gates per round. More importantly, we can substitute any matrix A' in the set $\{PAP^{-1} | P \in \mathcal{M}_{8\times 8}(\mathbb{F}_2)$ *is a permutation matrix*$\}$ for A which retains not only MDS property of diffusion matrices (according to Theorem 2) but also the the number of XOR gates of them (because P^{-1} is a permutation matrix too). With this technique, we get $5820 \times 2^{16} \times (8!)$ recursive MDS matrices.

Appendix C.2 Results of the Experiment

$[y^2, y^2, 1, y]$, $[y^2, y^2, y, y^3]$, $[y^2, 1, y, y]$, $[y^2, 1, y^3, 1]$, $[y^2, y, y, y^2]$, $[y^2, y, y, y^3]$, $[y^2, y, y^3, 1]$, $[y^2, y^3, y, 1]$, $[y^2, y^3, y, y]$, $[1, y^2, y^2, y^3]$, $[1, y^2, 1, y^3]$, $[1, y^2, y^3, y]$, $[1, y, y^3, y^2]$, $[1, y^3, y^2, y^2]$, $[1, y^3, 1, y]$, $[1, y^3, y, y^3]$, $[y, 1, y^2, 1]$, $[y, 1, 1, y^2]$, $[y, 1, 1, y^3]$, $[y, y, 1, y^3]$, $[y, y^3, y^2, 1]$, $[y, y^3, y^2, y^3]$, $[y, y^3, y, 1]$, $[y, y^3, y^3, y]$, $[y^3, y^2, y^2, y^3]$, $[y^3, y^2, y^3, y^3]$, $[y^3, 1, y, y]$, $[y^3, y, 1, y^3]$, $[y^3, y, y^3, y^3]$, $[y^3, y^3, 1, y^2]$, $[1, y^3 + y^2 + y + 1, y^3, y^3 + y + 1]$, $[1, y^3 + y^2 + y + 1, 1, y^3 + y^2 + y]$, $[1, y^3 + y^2 + y + 1, 1, y]$, $[1, y^3 + y^2 + y + 1, 1, y^3 + y^2 + 1]$, $[1, y^2 + 1, y^2 + y, y^3 + y^2 + y + 1]$, $[1, y^2 + 1, y^2 + y, y^2 + 1]$, $[1, y^2 + 1, y^2 + y, 1]$, $[1, y^2 + 1, y^3 + y^2 + y + 1, y + 1]$, $[1, y^2 + 1, y^3 + y^2 + y + 1, y^3 + 1]$, $[1, y^2 + 1, y^2 + 1, y^3 + y^2 + y]$, $[1, y^2 + 1, y^2 + 1, y^3 + y^2 + y + 1]$, $[1, y^2 + 1, y^3 + y^2 + 1, y^3 + y^2 + 1]$, $[1, y^2 + 1, y^2, y^3 + 1]$, $[1, y^2 + 1, y + 1, y^3 + y^2]$, $[1, y^2 + 1, y^3 + y, y]$, $[1, y^2 + 1, y^3 + y, y^3 + y^2 + 1]$, $[1, y^2 + 1, y^3 + y + 1, y^2 + y + 1]$, $[1, y^2 + 1, y^3 + y + 1, y^3 + 1]$, $[1, y^2 + 1, y^3, y + 1]$, $[1, y^2 + 1, y^3 + 1, y^2 + 1]$, $[1, y^2 + 1, y^3 + 1, y^3 + y^2 + 1]$, $[1, y^2 + 1, y^3 + 1, y^3]$, $[1, y^2 + 1, 1, y^2 + y]$, $[1, y^2 + 1, 1, y^3 + y^2 + y + 1]$, $[1, y^2 + 1, 1, y^3 + y^2]$, $[1, y^2 + 1, 1, y^2]$, $[1, y^2 + 1, 1, y^2 + y + 1]$, $[1, y^3 + y^2, y^2 + y, y^3 + y^2]$, $[1, y^3 + y^2, y^3 + y^2 + y, y^3 + y^2 + y + 1]$, $[1, y^3 + y^2, y, y^3 + y^2]$, $[1, y^3 + y^2, y, y^2]$, $[1, y^3 + y^2, y, y^3 + 1]$, $[1, y^3 + y^2, y^2 + 1, y^2 + y]$, $[1, y^3 + y^2, y^2 + 1, y^3 + y^2 + y]$, $[1, y^3 + y^2, y^2 + 1, y^3 + y + 1]$, $[1, y^3 + y^2, y^3 + y^2, y^3 + y + 1]$, $[1, y^3 + y^2, y^3 + y^2, y^2 + y + 1]$, $[1, y^3 + y^2, y^3 + y^2 + 1, y]$, $[1, y^3 + y^2, y^3 + y^2 + 1, y^3 + y^2]$, $[1, y^3 + y^2, y^2, y^3 + y^2 + y]$, $[1, y^3 + y^2, y^2, y^3 + y^2]$, $[1, y^3 + y^2, y^2, y^3 + y + 1]$, $[1, y^3 + y^2, y^2 + 1, y^3 + y^2]$, $[1, y^3 + y^2, y^2 + 1, y^3 + y + 1]$, $[1, y^3 + y^2, y + 1, y^3 + y^2 + y + 1]$, $[1, y^3 + y^2, y + 1, y + 1]$, $[1, y^3 + y^2, y + 1, y^3 + y + 1]$, $[1, y^3 + y^2, y^3 + y, y^3 + y^2 + y]$, $[1, y^3 + y^2, y^3 + y, y^2 + y + 1]$, $[1, y^3 + y^2, y^3 + y, y^3 + 1]$, $[1, y^3 + y^2, y^2 + y + 1, y^3 + y^2 + y + 1]$, $[1, y^3 + y^2, y^2 + y + 1, y + 1]$, $[1, y^3 + y^2, y^3, y^3 + y^2 + y + 1]$, $[1, y^3 + y^2, y^3, y^3 + y^2]$, $[1, y^3 + y^2, y^3, y^3 + y + 1]$, $[1, y^3 + y^2, y^3, y^2 + y + 1]$, $[1, y^3 + y^2, y^3 + 1, y^3 + y^2 + y]$, $[1, y^3 + y^2, y^3 + 1, y^3 + y^2 + y + 1]$, $[1, y^3 + y^2, y^3 + 1, y^3 + y^2 + 1]$, $[1, y^3 + y^2, y^3 + 1, y^2]$, $[1, y^3 + y^2, 1, y^2 + 1]$, $[1, y^3 + y^2, 1, y^3 + y^2 + 1]$, $[1, y^3 + y^2, 1, y^2]$, $[1, y^3 + y^2 + 1, y, y^3]$, $[1, y^3 + y^2 + 1, y^3 + y^2 + y + 1, y^2 + 1]$,

$[1, y^3+y^2+1, y^3+y^2+y+1, y^3]$, $[1, y^3+y^2+1, y^3+y^2+y+1, 1]$, $[1, y^3+y^2+1, y^2+1, y^3+y^2+1]$, $[1, y^3+y^2+1, y^3+y^2, y^3+y^2+1]$, $[1, y^3+y^2+1, y^3+y^2+1, y^2+1]$, $[1, y^3+y^2+1, y^2, y^3+y^2+y+1]$, $[1, y^3+y^2+1, y^2, y^3+y+1]$, $[1, y^3+y^2+1, y+1, y^2+y]$, $[1, y^3+y^2+1, y+1, y]$, $[1, y^3+y^2+1, y+1, y+1]$, $[1, y^3+y^2+1, y+1, y^3]$, $[1, y^3+y^2+1, y^3+y, y]$, $[1, y^3+y^2+1, y^3+y, y^2+1]$, $[1, y^3+y^2+1, y^3+y, y^3+y^2+1]$, $[1, y^3+y^2+1, y^3+y, 1]$, $[1, y^3+y^2+1, y^3+y+1, y^3]$, $[1, y^3+y^2+1, 1, y^3+y^2+y]$, $[1, y^3+y^2+1, 1, y^3+y^2+y+1]$, $[1, y^3+y^2+1, 1, y^3+y^2]$, $[1, y^3+y^2+1, 1, y+1]$, $[1, y^3+y^2+1, 1, y^2+y+1]$, $[1, y^2, y^2+y, y^2]$, $[1, y^2, y^2+y, y^3]$, $[1, y^2, y^2+y, 1]$, $[1, y^2, y^3+y^2+y, y^3+y^2+y+1]$, $[1, y^2, y^3+y^2+y, y^2]$, $[1, y^2, y^3+y^2+y, y^3+y+1]$, $[1, y^2, y, y^3+y]$, $[1, y^2, y^3+y^2+y+1, y]$, $[1, y^2, y^2+1, y^3+y^2+y]$.

Appendix D Proof of Lemma 1

To prove $\det(L) = \det(\det_s(L))$, we use induction on n.

In case when $n = 1$, we do not need to prove anything.

To clarify the general case, we illustrate the case when $n = 2$ for simplicity. When $n = 2$,

$$L = \begin{pmatrix} L_{1,1} & L_{1,2} \\ L_{2,1} & L_{2,2} \end{pmatrix}. \tag{3}$$

From the multiplication of block matrices, we get

$$\begin{pmatrix} I_b & O_b \\ -L_{2,1} & I_b \end{pmatrix} \begin{pmatrix} I_b & O_b \\ O_b & L_{1,1} \end{pmatrix} \begin{pmatrix} L_{1,1} & L_{1,2} \\ L_{2,1} & L_{2,2} \end{pmatrix} = \begin{pmatrix} L_{1,1} & L_{1,2} \\ O_b & L_{1,1}L_{2,2} - L_{1,2}L_{2,1} \end{pmatrix}, \tag{4}$$

where I_b denotes the identity matrix in $\mathcal{M}_{b\times b}(F)$ and O_b denotes the zero matrix in $\mathcal{M}_{b\times b}(F)$. By calculating the determinants of both sides of equation (4) we get

$$\det(L_{1,1})\det(L) = \det(L_{1,1})\det(L_{1,1}L_{2,2} - L_{1,2}L_{2,1}). \tag{5}$$

If $\det(L_{1,1}) \neq 0$, the equation (1) is proved immediately. If $\det(L_{1,1}) = 0$, we have to use a trick. Specifically, we can regard L as a matrix over the polynomial ring $F[x]$ where x is an indeterminant of F and substitute $xI_b + L_{1,1}$ for $L_{1,1}$ in L. From this viewpoint, equation (5) becomes

$$\det(xI_b + L_{1,1})\det(L) = \det(xI_b + L_{1,1})\det((xI_b + L_{1,1})L_{2,2} - L_{1,2}L_{2,1}). \tag{6}$$

Note that $\det(xI_b + L_{1,1})$, $\det(L)$ and $\det((xI_b + L_{1,1})L_{2,2} - L_{1,2}L_{2,1})$ all become polynomials in $F[x]$ now. Obviously, $\det(xI_b + L_{1,1}) \neq 0$, so

$$\det(L) = \det((xI_b + L_{1,1})L_{2,2} - L_{1,2}L_{2,1}). \tag{7}$$

If two polynomials are equal, their corresponding coefficients are equal too. So if we assign $x = 0$ in the equation (7), we prove the equation (1) in case when $n = 2$.

Now we suppose equation (1) is true for $1, 2, \cdots, n-1$. From the multiplication of block matrices, we get

$$\begin{pmatrix} I_b & O_b & \cdots & O_b \\ -L_{2,1} & I_b & \cdots & O_b \\ \cdots & \cdots & \ddots & \cdots \\ -L_{n,1} & O_b & \cdots & I_b \end{pmatrix} \begin{pmatrix} I_b & O_b & \cdots & O_b \\ O_b & L_{1,1} & \cdots & O_b \\ \cdots & \cdots & \ddots & \cdots \\ O_b & O_b & \cdots & L_{1,1} \end{pmatrix} L = \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\ O_b & & & \\ \vdots & & A & \\ O_b & & & \end{pmatrix}, \tag{8}$$

where $A \in \mathcal{M}_{b(n-1)\times b(n-1)}(F)$. Let U, V, W denote

$$\begin{pmatrix} I_b & O_b & \cdots & O_b \\ -L_{2,1} & I_b & \cdots & O_b \\ \cdots & \cdots & \ddots & \cdots \\ -L_{n,1} & O_b & \cdots & I_b \end{pmatrix}, \begin{pmatrix} I_b & O_b & \cdots & O_b \\ O_b & L_{1,1} & \cdots & O_b \\ \cdots & \cdots & \ddots & \cdots \\ O_b & O_b & \cdots & L_{1,1} \end{pmatrix}, \begin{pmatrix} L_{1,1} & L_{1,2} & \cdots & L_{1,n} \\ O_b & & & \\ \vdots & & A & \\ O_b & & & \end{pmatrix}$$

respectively. Then we can write the equation (8) as $UVL = W$. On one hand, by taking symbolic determinant of both sides of $UVL = W$, we get

$$\det_s(UVL) = \det_s(W). \tag{9}$$

From the definition of symbolic determinant, we can easily get

$$\det_s(UVL) = \det_s(U)\det_s(V)\det_s(L). \tag{10}$$

Thus we have

$$\det_s(U)\det_s(V)\det_s(L) = \det_s(W). \tag{11}$$

By computing the symbolic determinants of the both sides of equation (11), we get

$$L_{1,1}^{n-1}\det_s(L) = L_{1,1}\det_s(A). \tag{12}$$

Then by taking the determinants of both sides of equation (12), we have

$$\det(L_{1,1})^{n-1}\det(det_s(L)) = \det(L_{1,1})\det(\det_s(A)). \tag{13}$$

On the other hand, by directly taking the determinants of both sides of $UVL = W$, we get

$$\det(U)\det(V)\det(L) = \det(UVL) = \det(W). \tag{14}$$

By computing the determinants of both sides of equation (14), we have

$$\det(L_{1,1})^{n-1}\det(L) = \det(L_{1,1})\det(A). \tag{15}$$

From the induction hypothesis, we know $\det(\det_s(A)) = \det(A)$. Thus,

$$\det(L_{1,1})^{n-1}\det(\det_s(L)) = \det(L_{1,1})^{n-1}\det(L). \tag{16}$$

If $\det(L_{1,1}) \neq 0$, we get $\det(\det_s(L)) = \det(L)$, and the equation (1) is proved immediately. If $\det(L_{1,1}) = 0$, we just need to use the same technique as in case when $n = 2$. Specifically, we also regard L as a matrix over the polynomial ring $F[x]$ where x is an indeterminant of F and substitute $xI_b + L_{1,1}$ for $L_{1,1}$ in L. Finally, by assigning $x = 0$, we complete the proof of equation (1).

References

1 Burrow M D. The minimal polynomial of a linear transformation. Amer Math Monthly, 1973, 80: 1129–1131
2 Blaum M, Roth R M. On lowest density MDS codes. IEEE Transactions on Information Theory, 1999, 45(1): 46–59
3 Daemen J, Rijmen V. The Design of Rijndael AES - The Advanced Encryption Standard. Berlin: Springer-Verlag, 2002. 17–19
4 Lidl R, Niederreiter H. Finite Fields. 2nd ed. Cambridge: Cambridge University Press, 1997. 66–69
5 MacWilliams F J, Sloane N J A. The Theory of Error-Correcting Codes. Amsterdam: North-Holland Publishing Company, 1977. 317–331
6 Rotman J J. Advanced Modern Algebra. Upper Saddle River: Prentice Hall, 2003. 182–198
7 Serre D. Matrices Theory and Applications. New York: Springer-Verlag, 2002. 15–30
8 Sajadieh M, Dakhilalian M, Mala H, et al. Recursive diffusion layers for block ciphers and hash functions. In: Proceedings of International Workshop, FSE 2012. Heidelberg: Springer-Verlag, 2012. 385–401
9 Wu S B, Wang M S, Wu W L. Recursive diffusion layers for (lightweight) block ciphers and hash functions. In: Proceedings of International Conference, SAC 2012. Heidelberg: Springer-Verlag, 2013. 355–371