Refined Analysis of Bounds Related to Linear and Differential Cryptanalysis for the AES

Liam Keliher

Department of Mathematics and Computer Science,
Mount Allison University,**
Sackville, New Brunswick, Canada
lkeliher@mta.ca

Abstract. The best upper bounds on the maximum expected linear probability (MELP) and the maximum expected differential probability (MEDP) for the AES, due to Park et al. [23], are 1.075×2^{-106} and 1.144×2^{-111}, respectively, for $T \geq 4$ rounds. These values are simply the 4th powers of the best upper bounds on the MELP and MEDP for $T = 2$ [3, 23]. In our analysis we first derive nontrivial *lower* bounds on the 2-round MELP and MEDP, thereby trapping each value in a small interval; this demonstrates that the best 2-round upper bounds are quite good. We then prove that these same 2-round upper bounds are not tight—and therefore neither are the corresponding upper bounds for $T \geq 4$. Finally, we show how a modified version of the KMT2 algorithm (or its dual, KMT2-DC), due to Keliher et al. (see [8]), can potentially improve any existing upper bound on the MELP (or MEDP) for any SPN. We use the modified version of KMT2 to improve the upper bound on the AES MELP to 1.778×2^{-107}, for $T \geq 8$.

Keywords: AES, Rijndael, SPN, provable security, linear cryptanalysis, differential cryptanalysis, MELP, MEDP, KMT2, KMT2-DC.

1 Introduction

During the past few years, several papers have appeared dealing with the provable security of substitution-permutation network (SPN) block ciphers against linear and differential cryptanalysis [3, 6, 7, 9, 10, 11, 12, 22, 23, 24]. Most of these results have been applied to the Advanced Encryption Standard (AES) [5]—each new result has demonstrated greater provable security against one or both of these attacks.

Exhibiting provable security against linear and differential cryptanalysis requires proving that the maximum expected linear probability (MELP) and the maximum expected differential probability (MEDP), respectively, are small over

** This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC).

H. Dobbertin, V. Rijmen, A. Sowa (Eds.): AES 2004, LNCS 3373, pp. 42–57, 2005.

Table 1. Previous upper bounds on the MELP and MEDP for the AES

MELP	MEDP	Range of rounds
2^{-24} [7]	2^{-24} [7]	$T \geq 2$
2^{-75} [9]	2^{-75} [10]	$T \geq 7$
$2^{-92.4}$ [11, 12]	$2^{-95.1}$ [12]	$T \geq 9$
2^{-96} [24]	2^{-96} [24]	$T \geq 4$
1.06×2^{-96} [22]	1.06×2^{-96} [22]	$T \geq 4$
1.075×2^{-106} [23]	1.144×2^{-111} [23]	$T \geq 4$

T core cipher rounds (typically $T = R - 1$ or $T = R - 2$, where R is the total number of rounds). Since exact computation of these values often appears to be infeasible, researchers have focused on bounds. A sufficiently small upper bound corresponds to a data complexity that is prohibitively large, since the data complexity is proportional to the inverse of the corresponding MELP / MEDP [19, 20]. Note that bounds often appear in pairs—one each for the MELP and MEDP—because of the well-known duality between linear and differential cryptanalysis [1, 17]. Table 1 summarizes the upper bounds that have been derived for the AES prior to the current paper. [1] [2]

The best upper bounds in Table 1 (last row), due to Park et al., are in fact the 4th powers of the best upper bounds on the MELP and MEDP for $T = 2$, namely $\frac{48,193,441}{2^{52}} \approx 1.44 \times 2^{-27}$ and $\frac{79}{2^{34}} \approx 1.23 \times 2^{-28}$, respectively [3, 23]. In fact, Park et al. show that the 4th power of *any* upper bound on the 2-round MELP / MEDP for the AES is an upper bound on the MELP / MEDP for $T \geq 4$ (this also follows from the work of Sano et al. [24]). Therefore the 2-round MELP and MEDP are important values for analyzing the resistance of the AES to linear and differential cryptanalysis.

In this paper we first derive nontrivial *lower* bounds on the 2-round MELP and MEDP for the AES, namely 1.638×2^{-28} and 1.656×2^{-29}, respectively, thereby trapping each value in a small interval; this demonstrates that the best 2-round upper bounds are quite good.[3] Second, we prove that these same 2-round upper bounds are not tight—and therefore neither are the corresponding upper bounds for $T \geq 4$. Third, we show how a modified version of the KMT2 algorithm (or its dual, KMT2-DC), due to Keliher et al. (see [8]), can potentially improve any existing upper bound on the MELP (or MEDP) for any SPN. We use the modified version of KMT2 to improve the upper bound on the AES MELP

[1] The results in [7] were not applied to the AES, but the values in the first row of Table 1 are the upper bounds that would have resulted.

[2] The almost identical bounds in [24] and [22] were apparently obtained independently.

[3] After the presentation of this paper, we learned that the same lower bounds had previously been obtained by Chun et al. [3].

to 1.778×2^{-107}, for $T \geq 8$. (The KMT2 / KMT2-DC algorithm computes upper bounds on the MELP / MEDP "from scratch"; the modification involves incorporating existing upper bounds that are superior to those computed directly by KMT2 / KMT2-DC in order to refine the former.)

1.1 The Advanced Encryption Standard (AES)

The *Advanced Encryption Standard* (AES) is a U.S. block cipher standard selected in 2000 after an open submission and evaluation process. The AES is the SPN *Rijndael*, designed by Joan Daemen and Vincent Rijmen [5]. A single AES round (minus the subkey mixing) is depicted in Figure 1.

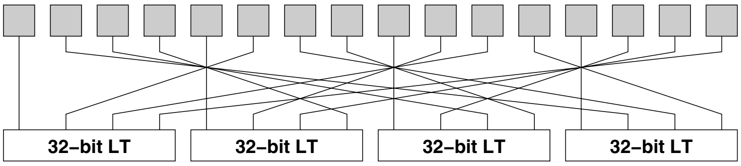

Fig. 1. One AES round

The AES has a block size of 128 bits. The substitution stage consists of 16 identical 8×8 s-boxes (the same s-box is used in all rounds). The linear transformation comprises two steps: a byte permutation, and the parallel application of four copies of a maximally diffusive 32-bit linear transformation (see Remark 4). The number of rounds varies according to the key length as follows: 128-bit key ⇒ 10 rounds, 192-bit key ⇒ 12 rounds, 256-bit key ⇒ 14 rounds.

1.2 Assumption of Independent Subkeys

In analyzing the resistance of block ciphers to linear and differential cryptanalysis, it is standard to assume that each subkey is chosen independently and uniformly from the set of all possible subkeys.[4] We adopt this approach. Because of the complexities introduced by most key schedules, the values relevant to linear and differential cryptanalysis are rarely calculated for the true distribution of subkeys—this remains an interesting and largely unexplored area of study.

2 Linear and Differential Cryptanalysis

Linear and differential cryptanalysis are generally considered to be the most powerful attacks on block ciphers. Linear cryptanalysis, due to Matsui [16], is a known-plaintext attack that exploits the existence of relatively large *expected*

[4] Some authors use AES* to denote the AES modified by this assumption.

linear probability (ELP) values over T core cipher rounds. Differential cryptanalysis, due to Biham and Shamir [2], is a chosen-plaintext attack that exploits the existence of relatively large *expected differential probability* (EDP) values over T core rounds. Typical values of interest are $T = R - 1$ and $T = R - 2$.

The remainder of this section deals with background concepts related to linear and differential cryptanalysis of SPNs. We use N to denote the block size, n to denote the s-box input/output size, and M to denote the number of s-boxes per round (so $M = \frac{N}{n}$). We assume that the same linear transformation and sequence of s-boxes are used in each round (the s-boxes within a round may or many not be identical). It is easy to generalize to the situation in which the linear transformation and s-boxes differ from round to round.

2.1 Linear and Differential Probability

Definition 1. *Let* $B : \{0,1\}^d \rightarrow \{0,1\}^d$*, and let* $\mathbf{a}, \mathbf{b}, \Delta\mathbf{x}, \Delta\mathbf{y} \in \{0,1\}^d$ *be fixed. If* $\mathbf{X} \in \{0,1\}^d$ *is a uniformly distributed random variable, then the* linear probability $LP(\mathbf{a},\mathbf{b})$ *and the* differential probability $DP(\Delta\mathbf{x},\Delta\mathbf{y})$ *are defined as*

$$LP(\mathbf{a},\mathbf{b}) = \left(2 \cdot \mathrm{Prob}_{\mathbf{X}}\left\{\mathbf{a} \bullet \mathbf{X} = \mathbf{b} \bullet B(\mathbf{X})\right\} - 1\right)^2$$
$$DP(\Delta\mathbf{x},\Delta\mathbf{y}) = \mathrm{Prob}_{\mathbf{X}}\left\{B(\mathbf{X}) \oplus B(\mathbf{X} \oplus \Delta\mathbf{x}) = \Delta\mathbf{y}\right\}.$$

If B is parameterized by a key, $\mathbf{k}$*, we write* $LP(\mathbf{a},\mathbf{b};\mathbf{k})$ *and* $DP(\Delta\mathbf{x},\Delta\mathbf{y};\mathbf{k})$*, respectively, and the* expected linear probability $ELP(\mathbf{a},\mathbf{b})$ *and* expected differential probability *are defined as*

$$ELP(\mathbf{a},\mathbf{b}) = E_{\mathbf{K}}\left[LP(\mathbf{a},\mathbf{b};\mathbf{K})\right]$$
$$EDP(\Delta\mathbf{x},\Delta\mathbf{y}) = E_{\mathbf{K}}\left[DP(\Delta\mathbf{x},\Delta\mathbf{y};\mathbf{K})\right],$$

where $\mathbf{K}$ *is a random variable uniformly distributed over the space of keys.*

We view LP, ELP, DP, or EDP values as entries in a $2^d \times 2^d$ table in the obvious way. The values $\mathbf{a}$ and $\mathbf{b}$ in Definition 1 are called input and output *masks,* and the values $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ are called input and output *differences.* For our purposes, the mapping B in Definition 1 will be bijective, and will be an s-box, a single encryption round, or a sequence of consecutive encryption rounds.

Lemma 1. *Let* $B : \{0,1\}^d \rightarrow \{0,1\}^d$ *be bijective, and let* $\mathbf{a}, \mathbf{b}, \Delta\mathbf{x}, \Delta\mathbf{y} \in \{0,1\}^d$*. Then*

$$\sum_{\mathbf{u} \in \{0,1\}^d} LP(\mathbf{a},\mathbf{u}) \;=\; \sum_{\mathbf{u} \in \{0,1\}^d} LP(\mathbf{u},\mathbf{b}) \;=\; 1 \tag{1}$$
$$\sum_{\Delta\mathbf{u} \in \{0,1\}^d} DP(\Delta\mathbf{x},\Delta\mathbf{u}) \;=\; \sum_{\Delta\mathbf{u} \in \{0,1\}^d} DP(\Delta\mathbf{u},\Delta\mathbf{y}) \;=\; 1\,. \tag{2}$$

Proof. The proof of (1) derives directly from Parseval's Theorem [18]. The proof of (2) is trivial.

Remark 1. In what follows, terms such as "first round" and "last round" are relative to the T rounds under consideration. Single-variable superscripts refer to individual rounds, e.g., $LP^t(\mathbf{a},\mathbf{b};\mathbf{k}^t)$ and $ELP^t(\mathbf{a},\mathbf{b})$ are LP and ELP values, respectively, for round t $(1 \le t \le T)$. Superscripts of the form $[i \dots j]$ (with $i < j$) refer to a sequence of consecutive rounds viewed as a single unit, e.g., $EDP^{[1\dots 3]}(\Delta\mathbf{x},\Delta\mathbf{y})$ is an EDP value over rounds $1\dots 3$.

2.2 Provable Security (MELP and MEDP)

Given $T \ge 2$ core rounds under consideration, the critical value for linear cryptanalysis is the *maximum expected linear probability* (MELP)[5]:

$$MELP \;=\; \max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N\setminus\mathbf{0}} ELP^{[1\dots T]}(\mathbf{a},\mathbf{b})\,. \tag{3}$$

The critical value for differential cryptanalysis is the *maximum expected differential probability* (MEDP):

$$MEDP \;=\; \max_{\Delta\mathbf{x},\Delta\mathbf{y}\in\{0,1\}^N\setminus\mathbf{0}} EDP^{[1\dots T]}(\Delta\mathbf{x},\Delta\mathbf{y})\,. \tag{4}$$

For linear cryptanalysis / differential cryptanalysis, the data complexity of an attack with a given probability of success is proportional to the inverse of the MELP / MEDP. Therefore *provable security* can be claimed if this value is sufficiently small that the corresponding data complexity is prohibitive [19, 20].

2.3 Linear and Differential Characteristics

In general, for $T \ge 2$, it appears to be infeasible to compute the MELP or MEDP exactly for most SPNs. A traditional method of approximation involves the use of *characteristics.*

Definition 2. *A* linear characteristic / differential characteristic *for rounds* $1\dots T$ *is a* $(T{+}1)$*-tuple of* N*-bit masks / differences,* $\Omega = \langle \mathbf{a}^1,\mathbf{a}^2,\dots,\mathbf{a}^T,\mathbf{a}^{T+1}\rangle$ / $\Omega = \langle \Delta\mathbf{x}^1,\Delta\mathbf{x}^2,\dots,\Delta\mathbf{x}^T,\Delta\mathbf{x}^{T+1}\rangle$*; we view* $\mathbf{a}^t$ / $\Delta\mathbf{x}^t$ *and* $\mathbf{a}^{t+1}$ / $\Delta\mathbf{x}^{t+1}$ *as input and output masks / differences, respectively, for round* t $(1 \le t \le T)$*. The corresponding* expected linear characteristic probability *(ELCP) /* expected differential characteristic probability *(EDCP) is defined as*

$$ELCP^{[1\dots T]}(\Omega) = \prod_{t=1}^{T} ELP^t(\mathbf{a}^t,\mathbf{a}^{t+1}) \quad /$$

$$EDCP^{[1\dots T]}(\Omega) = \prod_{t=1}^{T} EDP^t(\Delta\mathbf{x}^t,\Delta\mathbf{x}^{t+1})\,.$$

[5] A number of papers (including some by the author) use *maximum average linear hull probability* (MALHP), but MELP is more consistent with other related terminology.

Remark 2. For most SPNs, it is feasible to compute the values $ELP^t(\mathbf{a}^t, \mathbf{a}^{t+1})$ / $EDP^t(\Delta\mathbf{x}^t, \Delta\mathbf{x}^{t+1})$ (see Section 2.5), and therefore to compute $ELCP^{[1\ldots T]}(\Omega)$ / $EDCP^{[1\ldots T]}(\Omega)$.

Using the Best Characteristic (Practical Security). A *best* linear / differential characteristic is one that maximizes $ELCP^{[1\ldots T]}(\Omega)$ / $EDCP^{[1\ldots T]}(\Omega)$ (a best characteristic is not necessarily unique) . There are well-known (and relatively efficient) algorithms for finding best characteristics [17, 21]. Denote the best linear / differential characteristic by $\hat{\Omega}_L = \langle \hat{\mathbf{a}}^1, \hat{\mathbf{a}}^2, \ldots, \hat{\mathbf{a}}^T, \hat{\mathbf{a}}^{T+1} \rangle$ / $\hat{\Omega}_D = \langle \Delta\hat{\mathbf{x}}^1, \Delta\hat{\mathbf{x}}^2, \ldots, \Delta\hat{\mathbf{x}}^T, \Delta\hat{\mathbf{x}}^{T+1} \rangle$. The data complexity of linear / differential cryptanalysis is often estimated by assuming that

$$MELP = ELP^{[1\ldots T]}(\hat{\mathbf{a}}^1, \hat{\mathbf{a}}^{T+1}) \approx ELCP^{[1\ldots T]}(\hat{\Omega}_L) \ / \tag{5}$$

$$MEDP = EDP^{[1\ldots T]}(\Delta\hat{\mathbf{x}}^1, \Delta\hat{\mathbf{x}}^{T+1}) \approx EDCP^{[1\ldots T]}(\hat{\Omega}_D)\,. \tag{6}$$

If the resulting data complexity is prohibitive, the cipher is *practically secure* [13].

2.4 Linear Hulls and Differentials

The concept of *linear hulls* is due to Nyberg [19]. The counterpart for differential cryptanalysis is the concept of *differentials,* due to Lai et al. [14].

Definition 3. *If $T \geq 2$ and $\mathbf{a}, \mathbf{b} \in \{0,1\}^N$ / $\Delta\mathbf{x}, \Delta\mathbf{y} \in \{0,1\}^N$, then the corresponding linear hull / differential, denoted $ALH(\mathbf{a}, \mathbf{b})$[6] / $DIFF(\Delta\mathbf{x}, \Delta\mathbf{y})$, is the set of all linear / differential characteristics for rounds $1 \ldots T$ having $\mathbf{a}$ / $\Delta\mathbf{x}$ as the first mask / difference and $\mathbf{b}$ / $\Delta\mathbf{y}$ as the last mask / difference, i.e., all linear / differential characteristics of the form*

$$\Omega = \langle \mathbf{a}, \mathbf{a}^2, \mathbf{a}^3, \ldots, \mathbf{a}^T, \mathbf{b} \rangle \quad / \quad \Omega = \langle \Delta\mathbf{x}, \Delta\mathbf{x}^2, \Delta\mathbf{x}^3, \ldots, \Delta\mathbf{x}^T, \Delta\mathbf{y} \rangle\,.$$

Theorem 1 ([19, 14]). *Let $\mathbf{a}, \mathbf{b} \in \{0,1\}^N$. Then*

$$ELP^{[1\ldots T]}(\mathbf{a}, \mathbf{b}) = \sum_{\Omega \in ALH(\mathbf{a},\mathbf{b})} ELCP^{[1\ldots T]}(\Omega)$$

$$EDP^{[1\ldots T]}(\Delta\mathbf{x}, \Delta\mathbf{y}) = \sum_{\Omega \in DIFF(\Delta\mathbf{x},\Delta\mathbf{y})} EDCP^{[1\ldots T]}(\Omega)\,.$$

It follows from Theorem 1 that the approximation in (5) / (6) does not hold in general, since $ELP^{[1\ldots T]}(\mathbf{a}, \mathbf{b})$ / $EDP^{[1\ldots T]}(\Delta\mathbf{x}, \Delta\mathbf{y})$ is seen to be the sum of (a large number of) terms $ELCP^{[1\ldots T]}(\Omega)$ / $EDCP^{[1\ldots T]}(\Omega)$, and therefore, in general, the ELCP / EDCP of any characteristic will be strictly *less than* the corresponding ELP / EDP value. Further, the MELP / MEDP may not be equal to (i.e., may be strictly greater than) the ELP / EDP associated with any best characteristic. This situation may result in an overestimation of the data complexity—beneficial for an attacker, but problematic for a cipher designer.

[6] Nyberg originally used *approximate linear hull,* hence the abbreviation ALH.

2.5 Active S-Boxes and Branch Numbers

Let $\mathcal{L}$ denote the SPN linear transformation represented as an invertible $N \times N$ binary matrix, i.e., if $\mathbf{x}, \mathbf{y} \in \{0,1\}^N$ are the input and output, respectively, for the linear transformation, then $\mathbf{y} = \mathcal{L}\mathbf{x}$ (view $\mathbf{x}$ and $\mathbf{y}$ as column vectors).

Lemma 2 ([5]). *If $\mathbf{a} \in \{0,1\}^N$ is a mask applied to the inputs of $\mathcal{L}$, then there is a unique corresponding mask $\mathbf{b} \in \{0,1\}^N$ applied to the outputs, i.e., there is a mask $\mathbf{b}$ such that for all $\mathbf{x} \in \{0,1\}^N$, $\mathbf{a} \bullet \mathbf{x} = \mathbf{b} \bullet (\mathcal{L}\mathbf{x})$. The relationship between $\mathbf{a}$ and $\mathbf{b}$ is given by $\mathbf{a} = \mathcal{L}'\mathbf{b}$, where $\mathcal{L}'$ is the matrix transpose of $\mathcal{L}$.*

If $\Delta\mathbf{x} \in \{0,1\}^N$ is an input difference for $\mathcal{L}$, then $\Delta\mathbf{y} = \mathcal{L}(\Delta\mathbf{x})$ is the unique corresponding output difference, i.e., $\mathcal{L}(\mathbf{x}) \oplus \mathcal{L}(\mathbf{x} \oplus \Delta\mathbf{x}) = \Delta\mathbf{y}$ for all $\mathbf{x} \in \{0,1\}^N$.

It follows from Lemma 2 that if $\mathbf{a}^t$ / $\Delta\mathbf{x}^t$ and $\mathbf{a}^{t+1}$ / $\Delta\mathbf{x}^{t+1}$ are input and output masks / differences for round t, then the resulting input and output masks / differences for the *substitution stage* of round t are $\mathbf{a}^t$ / $\Delta\mathbf{x}^t$ and $\mathbf{b}^t = \mathcal{L}'\mathbf{a}^{t+1}$ / $\Delta\mathbf{y}^t = \mathcal{L}^{-1}(\Delta\mathbf{x}^{t+1})$. Further, $\mathbf{a}^t$ / $\Delta\mathbf{x}^t$ and $\mathbf{b}^t$ / $\Delta\mathbf{y}^t$ can be naturally partitioned into input and output masks / differences for each s-box in round t. Enumerate the s-boxes from left to right as $S_1^t, S_2^t, \ldots, S_M^t$, and let the input and output masks / differences for S_m^t be denoted $\mathbf{a}_m^t$ / $\Delta\mathbf{x}_m^t$ and $\mathbf{b}_m^t$ / $\Delta\mathbf{y}_m^t$ $(1 \le m \le M)$. Then from Matsui's Piling-up Lemma [16],

$$ELP^t(\mathbf{a}^t, \mathbf{a}^{t+1}) = \prod_{m=1}^{M} LP^{S_m^t}(\mathbf{a}_m^t, \mathbf{b}_m^t) \tag{7}$$

$$EDP^t(\Delta\mathbf{x}^t, \Delta\mathbf{x}^{t+1}) = \prod_{m=1}^{M} DP^{S_m^t}(\Delta\mathbf{x}_m^t, \Delta\mathbf{y}_m^t) \tag{8}$$

(here the superscript S_m^t has the obvious meaning).

Definition 4 ([25]). *Let Ω be a T-round linear / differential characteristic for rounds $1 \ldots T$. Then Ω is called* consistent *if, for each s-box in rounds $1 \ldots T$, the input and output masks / differences determined by Ω for that s-box are either both zero or both nonzero.*

Definition 5 ([1]). *Given a consistent linear / differential characteristic, any s-box for which the resulting input and output masks / differences are nonzero is called* linearly / differentially active *(or just* active, *when the context is clear).*

For the remainder of this paper, we limit our consideration to consistent characteristics.

Definition 6. *Given a linear / differential characteristic, let $\mathbf{v} \in \{0,1\}^N$ be the input or output mask / difference for the substitution stage of round t. Then the active s-boxes in round t can be determined from $\mathbf{v}$ (without knowing the corresponding output or input mask / difference). We define $\gamma_\mathbf{v}$ to be the M-bit vector that encodes this pattern of active s-boxes: $\gamma_\mathbf{v} = \gamma_1\gamma_2 \ldots \gamma_M$, where $\gamma_i = 1$ if the i^{th} s-box is active, and $\gamma_i = 0$ otherwise, for $1 \le i \le M$.*

Definition 7. *Let $\gamma, \hat{\gamma} \in \{0,1\}^M$. Then*

$$W_l[\gamma, \hat{\gamma}] = \#\left\{\mathbf{y} \in \{0,1\}^N : \gamma_{\mathbf{x}} = \gamma,\, \gamma_{\mathbf{y}} = \hat{\gamma}, \quad \text{where } \mathbf{x} = \mathcal{L}'\mathbf{y}\right\}$$
$$W_d[\gamma, \hat{\gamma}] = \#\left\{\Delta\mathbf{x} \in \{0,1\}^N : \gamma_{\Delta\mathbf{x}} = \gamma,\, \gamma_{\Delta\mathbf{y}} = \hat{\gamma}, \quad \text{where } \Delta\mathbf{y} = \mathcal{L}(\Delta\mathbf{x})\right\}.$$

Remark 3. Informally, the value $W_l[\gamma, \hat{\gamma}]$ / $W_d[\gamma, \hat{\gamma}]$ represents the number of ways the linear transformation can "connect" a pattern of active s-boxes in one round (γ) to a pattern of active s-boxes in the next round ($\hat{\gamma}$).

The diffusive power of a linear transformation is its ability to force some minimum number of s-boxes to be active over a sequence of rounds. This is quantified in the following definition.

Definition 8 ([5]). *The* linear / differential branch number*, $\mathcal{B}_l$ / $\mathcal{B}_d$, of an SPN linear transformation is the minimum number of linearly / differentially active s-boxes in two consecutive rounds for any nonzero characteristic:*

$$\mathcal{B}_l = \min\left\{wt(\gamma_{\mathbf{x}}) + wt(\gamma_{\mathbf{y}}) : \mathbf{y} \in \{0,1\}^N \setminus \mathbf{0}, \;\; \mathbf{x} = \mathcal{L}'\mathbf{y}\right\} \quad /$$
$$\mathcal{B}_d = \min\left\{wt(\gamma_{\Delta\mathbf{x}}) + wt(\gamma_{\Delta\mathbf{y}}) : \Delta\mathbf{x} \in \{0,1\}^N \setminus \mathbf{0}, \;\; \Delta\mathbf{y} = \mathcal{L}(\Delta\mathbf{x})\right\}.$$

Remark 4. It is trivial to show that $2 \le \mathcal{B}_l, \mathcal{B}_d \le (M+1)$. Hong et al. [6] prove that $\mathcal{B}_l = (M+1)$ if and only if $\mathcal{B}_d = (M+1)$; in this case, the linear transformation is called *maximally diffusive.*

3 General Analysis of 2-Round MELP / MEDP

In this section we analyze the 2-round MELP and MEDP for a general SPN, stating results that will be useful in later sections. We focus primarily on the MELP, as the development for the MEDP is essentially parallel (we point out the significant differences in Section 3.1).

Without loss of generality, assume that the linear transformation is omitted from round 2. Therefore the active s-boxes in round 2 can be determined directly from an output mask for round 2, without applying Lemma 2.

Let $\mathbf{a}, \mathbf{b} \in \{0,1\}^N \backslash \mathbf{0}$ be input and output masks, respectively, for round 1 and round 2, and let $f = wt(\gamma_{\mathbf{a}})$, $\ell = wt(\gamma_{\mathbf{b}})$. Enumerate the active s-boxes in round 1 as $S_1^1, S_2^1, \ldots, S_f^1$, and enumerate the active s-boxes in round 2 as $S_1^2, S_2^2, \ldots, S_\ell^2$. Let $\boldsymbol{\alpha}_i$ be the input mask for S_i^1 (derived from $\mathbf{a}$), for $1 \le i \le f$, and let $\boldsymbol{\beta}_j$ be the output mask for S_j^2 (derived from $\mathbf{b}$), for $1 \le j \le \ell$. The characteristics in $ALH(\mathbf{a}, \mathbf{b})$ have the form $\langle \mathbf{a}, \mathbf{y}, \mathbf{b} \rangle$; enumerate the distinct "middle" masks as $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_W$, where $W = W_l[\gamma_{\mathbf{a}}, \gamma_{\mathbf{b}}]$. The $\mathbf{y}_w$ are input masks for round 2; denote the corresponding output masks for the *substitution stage* of round 1 as $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_W$ ($\mathbf{x}_w$ and $\mathbf{y}_w$ are related as in the first part of Lemma 2). For a given $\mathbf{x}_w$ ($1 \le w \le W$), let $\boldsymbol{\chi}_{(w,i)}$ be the output mask for S_i^1 ($1 \le i \le f$), and for the corresponding $\mathbf{y}_w$, let $\boldsymbol{\upsilon}_{(w,j)}$ be the input mask for S_j^2 $(1 \le j \le \ell)$. It follows from Theorem 1, Definition 2, and (7) that

$$ELP^{[1\ldots 2]}(\mathbf{a},\mathbf{b}) \;=\; \sum_{w=1}^{W} \left(\prod_{i=1}^{f} LP^{S_i^1}(\boldsymbol{\alpha}_i, \boldsymbol{\chi}_{(w,i)}) \cdot \prod_{j=1}^{\ell} LP^{S_j^2}(\boldsymbol{\upsilon}_{(w,j)}, \boldsymbol{\beta}_j) \right) . \quad (9)$$

It is useful to consider the set of vectors (of length $f+\ell$) of the form

$$V_w = \left\langle \boldsymbol{\chi}_{(w,1)},\ \boldsymbol{\chi}_{(w,2)},\ \ldots,\ \boldsymbol{\chi}_{(w,f)},\ \boldsymbol{\upsilon}_{(w,1)},\ \boldsymbol{\upsilon}_{(w,2)},\ \ldots,\ \boldsymbol{\upsilon}_{(w,\ell)} \right\rangle , \qquad (10)$$

for $1 \le w \le W$. Each coordinate of V_w is an element of $\{0,1\}^n \setminus \mathbf{0}$ (recall that n is the s-box input/output size).

Lemma 3. *Given $\mathbf{a},\mathbf{b} \in \{0,1\}^N \setminus \mathbf{0}$ that satisfy $wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) = \mathcal{B}_l$, let $W = W_l[\gamma_{\mathbf{a}},\gamma_{\mathbf{b}}]$, $f = wt(\gamma_{\mathbf{a}})$, $\ell = wt(\gamma_{\mathbf{b}})$, and let $\boldsymbol{\chi}_{(w,i)}$, $\boldsymbol{\upsilon}_{(w,j)}$ be defined as above. Then for fixed i $(1 \le i \le f)$, the values $\boldsymbol{\chi}_{(1,i)},\ldots,\boldsymbol{\chi}_{(W,i)}$ are distinct, and for fixed j $(1 \le j \le \ell)$, the values $\boldsymbol{\upsilon}_{(1,j)},\ldots,\boldsymbol{\upsilon}_{(W,j)}$ are distinct. In other words, for the set of vectors $\{V_w\}_{w=1}^W$, all the values in any one position are distinct.*

Proof. Contained in the proof of Theorem 1 in [23].

Remark 5. Clearly if $wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) = \mathcal{B}_l$, then $W_l[\gamma_{\mathbf{a}},\gamma_{\mathbf{b}}] \le (2^n - 1)$. Further, the values $\boldsymbol{\chi}_{(w,i)}$ and $\boldsymbol{\upsilon}_{(w,j)}$ depend only on $\gamma_{\mathbf{a}}$ and $\gamma_{\mathbf{b}}$, not on the specific values of $\mathbf{a}$ and $\mathbf{b}$.

Lemma 4. *Given $\mathbf{a},\mathbf{b} \in \{0,1\}^N \setminus \mathbf{0}$ that satisfy $wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) > \mathcal{B}_l$, let $W = W_l[\gamma_{\mathbf{a}},\gamma_{\mathbf{b}}]$, $f = wt(\gamma_{\mathbf{a}})$, $\ell = wt(\gamma_{\mathbf{b}})$, and let $\boldsymbol{\chi}_{(w,i)}$, $\boldsymbol{\upsilon}_{(w,j)}$ be defined as above. Consider the vectors V_w in (10). Select any $(f+\ell-\mathcal{B}_l)$ vector positions, and fix a value in $\{0,1\}^n \setminus \mathbf{0}$ for each position. Now form the subset of $\{V_w\}_{w=1}^W$ consisting of those V_w that contain the selected fixed values in the specified positions—denote this subset by $\mathcal{V}$. Then for each of the $\mathcal{B}_l$ vector positions whose values were not fixed, all the values in that position are distinct as we range over $\mathcal{V}$.*

Proof. Contained in the proof of Theorem 1 in [23].

Remark 6. It follows that the number of vectors in $\mathcal{V}$ is at most $(2^n - 1)$. The vectors in $\mathcal{V}$ depend on $\gamma_{\mathbf{a}}$ and $\gamma_{\mathbf{b}}$ (not on the specific values of $\mathbf{a}$ and $\mathbf{b}$), and also on the choice of vector positions to be assigned fixed values, together with the particular fixed values chosen for those positions.

Definition 9. *For $T = 2$ core SPN rounds, a* $\mathcal{B}_l$-list *is a set of vectors, each of which has length $\mathcal{B}_l$, that has been derived in one of two ways:*

1. *by selecting any $\mathbf{a},\mathbf{b} \in \{0,1\}^N \setminus \mathbf{0}$ satisfying $wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) = \mathcal{B}_l$ and forming the set $\{V_w\}_{w=1}^W$, as in Lemma 3;*

2. by selecting any $\mathbf{a}, \mathbf{b} \in \{0,1\}^N \setminus \mathbf{0}$ satisfying $wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) > \mathcal{B}_l$, forming any set $\mathcal{V}$ according to Lemma 4, and then shortening each vector in $\mathcal{V}$ to length $\mathcal{B}_l$ by removing those positions that were assigned fixed values.

Let $\mathcal{B}_l$-LIST$^{(i)}$ denote the set of all $\mathcal{B}_l$-lists that are formed by Option i above, for $i = 1, 2$, and let

$$\mathcal{B}_l\text{-}\mathit{LIST} \;=\; \mathcal{B}_l\text{-}\mathit{LIST}^{(1)} \,\cup\, \mathcal{B}_l\text{-}\mathit{LIST}^{(2)}\,.$$

For any $\mathcal{Z} \in \mathcal{B}_l$-LIST, let $\delta(\mathcal{Z})$ denote the number of vectors in $\mathcal{Z}$.

It follows from Remarks 5 and 6 that $\delta(\mathcal{Z}) \le (2^n - 1)$ for any $\mathcal{Z} \in \mathcal{B}_l$-LIST. For any vector $\mathbf{z} = \langle \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \ldots, \boldsymbol{\zeta}_{\mathcal{B}_l} \rangle$ in any $\mathcal{B}_l$-list $\mathcal{Z}$, each $\boldsymbol{\zeta}_i$ is either an output mask for a particular s-box in round 1, or an input mask for a particular s-box in round 2. In the former case, let $\boldsymbol{\alpha}_i$ denote a nonzero *input* mask for the same s-box, and let $LP^*(\boldsymbol{\alpha}_i, \boldsymbol{\zeta}_i) = LP(\boldsymbol{\alpha}_i, \boldsymbol{\zeta}_i)$. In the latter case, let $\boldsymbol{\alpha}_i$ denote a nonzero *output* mask for the same s-box, and let $LP^*(\boldsymbol{\alpha}_i, \boldsymbol{\zeta}_i) = LP(\boldsymbol{\zeta}_i, \boldsymbol{\alpha}_i)$; here $\boldsymbol{\alpha}_i$ is playing the role of one of the $\boldsymbol{\beta}_j$ used earlier, e.g., in (9), and $LP^*(\cdot,\cdot)$ is the *transpose* of the s-box LP table. (For simplicity, the specific s-box is now implicit in the notation.)

Definition 10. *Let $\mathcal{Z} \in \mathcal{B}_l$-LIST. Then*

$$\sigma(\mathcal{Z}) \;\overset{\text{def}}{=}\; \max_{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{\mathcal{B}_l} \in \{0,1\}^n \setminus \mathbf{0}} \left(\sum_{\langle \boldsymbol{\zeta}_1,\ldots,\boldsymbol{\zeta}_{\mathcal{B}_l} \rangle \in \mathcal{Z}} \; \prod_{i=1}^{\mathcal{B}_l} LP^*(\boldsymbol{\alpha}_i, \boldsymbol{\zeta}_i) \right).$$

The following two theorems are central to this paper.

Theorem 2. *The 2-round MELP is lower bounded by*

$$\max\left\{ \sigma(\mathcal{Z}) : \mathcal{Z} \in \mathcal{B}_l\text{-}\mathit{LIST}^{(1)} \right\}.$$

Proof. It is easy to see that $\max\left\{ \sigma(\mathcal{Z}) : \mathcal{Z} \in \mathcal{B}_l\text{-LIST}^{\,(1)} \right\}$ is exactly equal to

$$\max_{\substack{\mathbf{a},\mathbf{b} \in \{0,1\}^N \setminus \mathbf{0} \\ wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) = \mathcal{B}_l}} \quad ELP^{[1\ldots 2]}(\mathbf{a}, \mathbf{b})\,,$$

which clearly lower bounds the 2-round MELP (see (3)).

Theorem 3. *The 2-round MELP is upper bounded by*

$$\max\left\{ \sigma(\mathcal{Z}) : \mathcal{Z} \in \mathcal{B}_l\text{-}\mathit{LIST} \right\}.$$

Proof. Let $\mathcal{M} = \max\{\sigma(\mathcal{Z}) : \mathcal{Z} \in \mathcal{B}_l\text{-LIST}\}$. Given the proof of Theorem 2, it suffices to show that

$$\max_{\substack{\mathbf{a},\mathbf{b} \in \{0,1\}^N \setminus \mathbf{0} \\ wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) > \mathcal{B}_l}} \quad ELP^{[1\ldots 2]}(\mathbf{a}, \mathbf{b}) \;\;\le\;\; \mathcal{M}\,.$$

Let $\mathbf{a}, \mathbf{b} \in \{0,1\}^N \setminus \mathbf{0}$ such that $F \overset{\text{def}}{=} wt(\gamma_{\mathbf{a}}) + wt(\gamma_{\mathbf{b}}) - \mathcal{B}_l > 0$. In keeping with Lemma 4, isolate F of the active s-boxes to be assigned fixed output/input masks (fixed output masks for round-1 s-boxes, and fixed input masks for round-2 s-boxes), let these fixed masks be denoted $\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F$, and let the corresponding input/output masks derived from $\mathbf{a}$ and $\mathbf{b}$ be denoted $\overline{\boldsymbol{\alpha}}_1, \ldots, \overline{\boldsymbol{\alpha}}_F$. Let the masks derived from $\mathbf{a}$ and $\mathbf{b}$ for the "non-fixed" s-boxes be denoted $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{\mathcal{B}_l}$. Denote the $\mathcal{B}_l$-list resulting from this setup by $\mathcal{Z}_{\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F}$. Then

$$
\begin{aligned}
& ELP^{[1\ldots 2]}(\mathbf{a}, \mathbf{b}) \\
&= \sum_{\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F \in \{0,1\}^n \setminus \mathbf{0}} \;\; \sum_{\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_{\mathcal{B}_l} \rangle \in \mathcal{Z}_{\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F}} \left(\prod_{i=1}^{F} LP^*(\overline{\boldsymbol{\alpha}}_i, \overline{\boldsymbol{\zeta}}_i) \cdot \prod_{j=1}^{\mathcal{B}_l} LP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j) \right) \\
&= \sum_{\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F \in \{0,1\}^n \setminus \mathbf{0}} \; \prod_{i=1}^{F} LP^*(\overline{\boldsymbol{\alpha}}_i, \overline{\boldsymbol{\zeta}}_i) \left(\sum_{\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_{\mathcal{B}_l} \rangle \in \mathcal{Z}_{\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F}} \; \prod_{j=1}^{\mathcal{B}_l} LP^*(\boldsymbol{\alpha}_j, \boldsymbol{\zeta}_j) \right) \\
&\le \mathcal{M} \left(\sum_{\overline{\boldsymbol{\zeta}}_1, \ldots, \overline{\boldsymbol{\zeta}}_F \in \{0,1\}^n \setminus \mathbf{0}} \; \prod_{i=1}^{F} LP^*(\overline{\boldsymbol{\alpha}}_i, \overline{\boldsymbol{\zeta}}_i) \right) \\
&= \mathcal{M}\,,
\end{aligned}
$$

where the last equality follows easily from (1).

3.1 Considerations Specific to MEDP

Tailoring the above definitions and results to the MEDP is straightforward: $\mathcal{B}_d$ is substituted for $\mathcal{B}_l$, $W_d[\ ]$ for $W_l[\ ]$, "difference" for "mask," DP values for LP values, and $DIFF(\cdot,\cdot)$ for $ALH(\cdot,\cdot)$. As well, the relationship between input and output differences over the linear transformation is via the second part of Lemma 2.

4 Lower Bounding the AES 2-Round MELP / MEDP

For the AES, $\mathcal{B}_l = \mathcal{B}_d = 5$; this is due to the fact that $\mathcal{B}_l = \mathcal{B}_d = 5$ for the 32-bit linear transformation component of the 128-bit AES linear transformation (see Figure 1) [5]. Hereafter we refer to $\mathcal{B}_l$-lists or $\mathcal{B}_d$-lists as 5-lists. As noted earlier, all AES s-boxes are identical. It is not hard to see that computing the MELP / MEDP for 2 AES rounds is equivalent to computing the MELP / MEDP for the "reduced" SPN depicted in Figure 2.

To lower bound the 2-round MELP / MEDP for the AES, we compute the value in Theorem 2 (or its MEDP counterpart) for the SPN in Figure 2. There are 56 pairs $(\gamma, \hat{\gamma}) \in \{0,1\}^4 \times \{0,1\}^4$ for which $wt(\gamma) + wt(\hat{\gamma}) = 5$; enumerate these as $(\gamma_1, \hat{\gamma}_1)$, $(\gamma_2, \hat{\gamma}_2)$, $\ldots$, $(\gamma_{56}, \hat{\gamma}_{56})$. A straightforward computation reveals that the 5-list associated with each pair $(\gamma_s, \hat{\gamma}_s)$ contains exactly $\left(2^8 - 1\right) = 255$

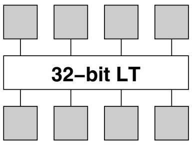

Fig. 2. Reduced 2-round AES

```
LowerBound = 0
For $s = 1$ to 56
    For $1 \leq \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4, \boldsymbol{\alpha}_5 \leq 255$
        Sum = 0
        For $w = 1$ to 255
            Prod = 1
            For $i = 1$ to 5
                Prod = Prod × $XP^*(\boldsymbol{\alpha}_i, D[s, w, i])$
            End For
            Sum = Sum + Prod
        End For
        If (Sum > LowerBound)
            LowerBound = Sum
        End If
    End For
End For
```

Fig. 3. Algorithm for lower bounding the 2-round MELP / MEDP for the AES

vectors. We store all the 5-lists in a 3-dimensional array of bytes, $D[\cdot, \cdot, \cdot]$, of size $56 \times 255 \times 5$, such that $D[s, \cdot, \cdot]$ contains the 5-list for $(\gamma_s, \hat{\gamma}_s)$ in the obvious way. Computing the lower bound on the MELP / MEDP reduces to the algorithm in Figure 3. (We use XP to mean either LP or DP, as appropriate.) The algorithm as presented is computationally intensive; however, we can make use of the fact that we are searching for a maximum to incorporate significant pruning, greatly reducing the running time.

Using the above algorithm, the lower bound on the 2-round MELP for the AES is 1.638×2^{-28}. Since the best upper bound is $\frac{48,193,441}{2^{52}} \approx 1.44 \times 2^{-27}$ [3, 23], the 2-round MELP is now known almost exactly. The lower bound on the 2-round MEDP is 1.656×2^{-29}, and since the best upper bound is $\frac{79}{2^{34}} \approx 1.23 \times 2^{-28}$ [3, 23], the 2-round MEDP is also now known almost exactly.

These lower bounds are important in light of the fact, stated earlier, that the 4th power of any upper bound on the 2-round MELP / MEDP for the AES (including the exact value) is an upper bound on the MELP / MEDP for $T \geq 4$ [23, 24]. This is how Park et al. obtain the upper bounds in the last row of Table 1. However, by constraining the 2-round MELP / MEDP as above, we see that this approach is essentially exhausted (for the AES).

5 Best AES 2-Round Upper Bounds Not Tight

In this section we show that the best upper bounds on the 2-round MELP and MEDP for the AES are not tight. First, we state the rationale behind the current bounds, and then we explain why they cannot be attained.

It is well known that all the nontrivial rows and columns of the LP / DP table for the AES s-box have the same distribution of values, given in Table 2 for the LP table and Table 3 for the DP table (ρ_i is a distinct value, and ϕ_i is the frequency with which it occurs) [11, 23].

Table 2. Distribution of LP values for the AES s-box

i	1	2	3	4	5	6	7	8	9
ρ_i	$\left(\frac{8}{64}\right)^2$	$\left(\frac{7}{64}\right)^2$	$\left(\frac{6}{64}\right)^2$	$\left(\frac{5}{64}\right)^2$	$\left(\frac{4}{64}\right)^2$	$\left(\frac{3}{64}\right)^2$	$\left(\frac{2}{64}\right)^2$	$\left(\frac{1}{64}\right)^2$	0
ϕ_i	5	16	36	24	34	40	36	48	17

Table 3. Distribution of DP values for the AES s-box

i	1	2	3
ρ_i	$\frac{1}{64}$	$\frac{1}{128}$	0
ϕ_i	1	126	129

We again use XP to mean either LP or DP, as appropriate. Consider the upper bound given in Theorem 3 (or its MEDP counterpart). Let $\mathcal{Z} \in$ 5-LIST. Adapting Theorem 1 in [23] to our notation, $\sigma(\mathcal{Z})$ equals the maximum value possible for a 5-list if, for some $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_5 \in \{0,1\}^8 \setminus \mathbf{0}$, the following two conditions are satisfied:

1. For every $\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_5 \rangle \in \mathcal{Z}$,
$$XP^*(\boldsymbol{\alpha}_1, \boldsymbol{\zeta}_1) = XP^*(\boldsymbol{\alpha}_2, \boldsymbol{\zeta}_2) = \cdots = XP^*(\boldsymbol{\alpha}_5, \boldsymbol{\zeta}_5) .$$
2. No nonzero XP value is omitted. In other words, for each $i \in \{1 \ldots 5\}$, as we range over all $\langle \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_5 \rangle \in \mathcal{Z}$ the values $XP^*(\boldsymbol{\alpha}_i, \boldsymbol{\zeta}_i)$ include every nonzero value in the XP table row or column indexed by $\boldsymbol{\alpha}_i$. (Obviously a necessary condition for the omission of a nonzero XP value is that $\delta(\mathcal{Z}) < 255$.)

Using the notation of Table 2 / Table 3, the maximum possible value for $\sigma(\mathcal{Z})$ is $\sum \rho_i^5\, \phi_i$ — this is exactly the best upper bound on the 2-round MELP /

MEDP [3, 23]. However, we have determined that this "worst-case" situation never occurs.

In brief, we systematically generated all 5-lists in 5-LIST$^{(2)}$ (it is clear from Section 4 that we don't need to consider the elements of 5-LIST$^{(1)}$). We observed that $\delta(\mathcal{Z}) < 255$ for all $\mathcal{Z} \in$ 5-LIST$^{(2)}$ (specifically, $251 \leq \delta(\mathcal{Z}) \leq 254$). For each 5-list generated, we ran an algorithm which ascertained that there do not exist masks $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_5 \in \{0,1\}^8 \setminus \mathbf{0}$ satisfying both Condition 1 and Condition 2 above. We used aggressive pruning to avoid iterating through all possible values of the $\boldsymbol{\alpha}_i$ (approximately 2^{40}) for any 5-list.

6 Modified Version of KMT2 Algorithm

The KMT2 algorithm (resp. its dual, KMT2-DC), due to Keliher, Meijer, and Tavares [8, 11], is a general algorithm that can be used to compute an upper bound on the MELP (resp. MEDP) for each value of $T \geq 2$ for any SPN. For a fixed value $T \geq 2$, and for all nonzero patterns of active s-boxes in the first and last rounds given by γ and $\hat{\gamma}$, KMT2 (resp. KMT2-DC) computes a value $UB^{[1\ldots T]}(\gamma, \hat{\gamma})$ such that for all $\mathbf{a}, \mathbf{b} \in \{0,1\}^N \setminus \mathbf{0}$, if $\gamma_{\mathbf{a}} = \gamma$ and $\gamma_{\mathbf{b}} = \hat{\gamma}$, then $ELP^{[1\ldots T]}(\mathbf{a}, \mathbf{b}) \leq UB^{[1\ldots T]}(\gamma, \hat{\gamma})$ (resp. $EDP^{[1\ldots T]}(\mathbf{a}, \mathbf{b}) \leq UB^{[1\ldots T]}(\gamma, \hat{\gamma})$). The values in the third row of Table 1 are from KMT2 and KMT2-DC.

The KMT2 / KMT2-DC algorithm works recursively, i.e., for $T \geq 3$, the values $UB^{[1\ldots T]}(\gamma, \hat{\gamma})$ depend on the values $UB^{[1\ldots (T-1)]}(\gamma, \hat{\gamma})$. This allows for a very simple improvement: For any $T \geq 2$, suppose that B is known to be an upper bound on the MELP / MEDP for that value of T (from some external source of information). Then if $B < UB^{[1\ldots T]}(\gamma, \hat{\gamma})$, replace $UB^{[1\ldots T]}(\gamma, \hat{\gamma})$ with B before proceeding to the computations for $T + 1$. In other words, enhance

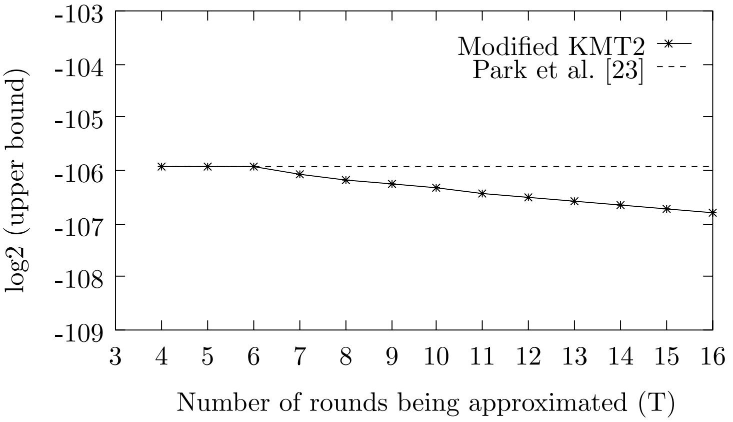

Fig. 4. Results from modified KMT2 for AES

KMT2 / KMT2-DC by incorporating other upper bounds when those bounds are superior to the values determined directly by the algorithm.

For the AES, we modified KMT2 in this fashion by incorporating the upper bound on the MELP for $T \geq 4$ due to Park et al [23]. The results are plotted in Figure 4. For $T \geq 8$, for example, the upper bound on the MELP is improved to 1.778×2^{-107}. This improvement is slight, but it is an effective "proof of concept." For other ciphers, the modified version of KMT2 / KMT2-DC may yield much more significant improvements over existing upper bounds.

Interestingly, modifying KMT2-DC using the upper bound on the AES MEDP for $T \geq 4$ due to Park et al. yielded no improvement over the existing bound for $T \geq 4$. This appears to be an artifact of the simple distribution of DP values for the AES s-box (LP / DP values play a fundamental role in KMT2 / KMT2-DC).

7 Conclusion

We have carefully analyzed bounds related to linear and differential cryptanalysis for the AES. We present nontrivial lower bounds on the 2-round maximum expected linear probability (MELP) and maximum expected differential probability (MEDP), trapping each value in a small interval. We then prove that the best upper bounds on the 2-round MELP and MEDP are not tight. Finally, we show how a modified version of the KMT2 / KMT2-DC algorithm can potentially improve existing upper bounds on the MELP / MEDP for any SPN, and we use the modified KMT2 algorithm to tighten the upper bound on the AES MELP to 1.778×2^{-107}, for $T \geq 8$.

References

1. E. Biham, *On Matsui's linear cryptanalysis,* Advances in Cryptology—EUROCRYPT'94, LNCS 950, pp. 341–355, Springer-Verlag, 1995.
2. E. Biham and A. Shamir, *Differential cryptanalysis of DES-like cryptosystems,* Advances in Cryptology—CRYPTO'90, LNCS 537, pp. 2–21, Springer-Verlag, 1991.
3. K. Chun, S. Kim, S. Lee, S.H. Sung, S. Yoon, *Differential and linear cryptanalysis for 2-round SPNs,* Information Processing Letters, Vol. 87, pp. 277–282, 2003.
4. J. Daemen, L. Knudsen, and V. Rijmen, *The block cipher* SQUARE, Fast Software Encryption (FSE'97), LNCS 1267, pp. 149–165, Springer-Verlag, 1997.
5. J. Daemen and V. Rijmen, *The Design of Rijndael: AES—The Advanced Encryption Standard,* Springer-Verlag, 2002.
6. S. Hong, S. Lee, J. Lim, J. Sung, and D. Cheon, *Provable security against differential and linear cryptanalysis for the SPN structure,* Fast Software Encryption (FSE 2000), LNCS 1978, pp. 273–283, Springer-Verlag, 2001.
7. J.-S. Kang, S. Hong, S. Lee, O. Yi, C. Park, and J. Lim, *Practical and provable security against differential and linear cryptanalysis for substitution-permutation networks,* ETRI Journal, Vol. 23, No. 4, December 2001.
8. L. Keliher, *Linear cryptanalysis of substitution-permutation networks,* Ph.D. Thesis, Queen's University, Kingston, Canada, 2003.

9. L. Keliher, H. Meijer, and S. Tavares, *New method for upper bounding the maximum average linear hull probability for SPNs,* Advances in Cryptology—EUROCRYPT 2001, LNCS 2045, pp. 420–436, Springer-Verlag, 2001.
10. L. Keliher, H. Meijer, and S. Tavares, *Dual of new method for upper bounding the maximum average linear hull probability for SPNs,* Technical Report, IACR ePrint Archive (`http://eprint.iacr.org`, Paper # 2001/033), 2001.
11. L. Keliher, H. Meijer, and S. Tavares, *Improving the upper bound on the maximum average linear hull probability for Rijndael,* Eighth Annual International Workshop on Selected Areas in Cryptography (SAC 2001), LNCS 2259, pp. 112–128, Springer-Verlag, 2001.
12. L. Keliher, H. Meijer, and S. Tavares, *Completion of computation of improved upper bound on the maximum average linear hull probability for Rijndael,* Technical Report, IACR ePrint Archive (`http://eprint.iacr.org`, Paper # 2004/074), 2004.
13. L. Knudsen, *Practically secure Feistel ciphers,* Fast Software Encryption, LNCS 809, pp. 211–221, Springer-Verlag, 1994.
14. X. Lai, J. Massey, and S. Murphy, *Markov ciphers and differential cryptanalysis,* Advances in Cryptology—EUROCRYPT'91, LNCS 547, pp. 17–38, Springer-Verlag, 1991.
15. C.H. Lim, *CRYPTON: A new 128-bit block cipher,* The First Advanced Encryption Standard Candidate Conference, Proceedings, Ventura, California, August 1998.
16. M. Matsui, *Linear cryptanalysis method for DES cipher,* Advances in Cryptology—EUROCRYPT'93, LNCS 765, pp. 386–397, Springer-Verlag, 1994.
17. M. Matsui, *On correlation between the order of s-boxes and the strength of DES,* Advances in Cryptology—EUROCRYPT'94, LNCS 950, pp. 366–375, Springer-Verlag, 1995.
18. W. Meier and O. Staffelbach, *Nonlinearity criteria for cryptographic functions,* Advances in Cryptology—EUROCRYPT'89, LNCS 434, pp. 549–562, Springer-Verlag, 1990.
19. K. Nyberg, *Linear approximation of block ciphers,* Advances in Cryptology—EUROCRYPT'94, LNCS 950, pp. 439–444, Springer-Verlag, 1995.
20. K. Nyberg and L. Knudsen, *Provable security against a differential attack,* Journal of Cryptology, Vol. 8, No. 1, pp. 27–37, 1995.
21. K. Ohta, S. Moriai, and K. Aoki, *Improving the search algorithm for the best linear expression,* Advances in Cryptology—CRYPTO'95, LNCS 963, pp. 157–170, Springer-Verlag, 1995.
22. S. Park, S.H. Sung, S. Chee, E-J. Yoon, and J. Lim, *On the security of Rijndael-like structures against differential and linear cryptanalysis,* Advances in Cryptology—ASIACRYPT 2002, LNCS 2501, pp. 176–191, Springer-Verlag, 2002.
23. S. Park, S.H. Sung, S. Lee, J. Lim, *Improving the upper bound on the maximum differential and the maximum linear hull probability for SPN structures and AES,* Fast Software Encryption (FSE 2003), LNCS 2887, pp. 247–260, Springer-Verlag, 2003.
24. F. Sano, K. Ohkuma, H. Shimizu, and S. Kawamura, *On the security of nested SPN cipher against the differential and linear cryptanalysis,* IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E86-A, No. 1, pp. 37–46, 2003.
25. S. Vaudenay, *On the security of CS-Cipher,* Fast Software Encryption (FSE'99), LNCS 1636, pp. 260–274, Springer-Verlag, 1999.