Tight Security for Key-Alternating Ciphers with Correlated Sub-Keys

Stefano Tessaro and Xihu Zhang

University of Washington, Seattle, USA
{tessaro,xihu}@cs.washington.edu

Abstract. A substantial effort has been devoted to proving optimal bounds for the security of key-alternating ciphers with independent sub-keys in the random permutation model (e.g., Chen and Steinberger, EUROCRYPT '14; Hoang and Tessaro, CRYPTO '16). While common in the study of multi-round constructions, the assumption that sub-keys are truly independent is not realistic, as these are generally highly correlated and generated from shorter keys.

In this paper, we show the existence of non-trivial distributions of limited independence for which a t-round key-alternating cipher achieves optimal security. Our work is a natural continuation of the work of Chen et al. (CRYPTO '14) which considered the case of $t = 2$ when all-subkeys are identical. Here, we show that key-alternating ciphers remain secure for a large class of $(t-1)$-wise and $(t-2)$-wise independent distribution of sub-keys.

Our proofs proceed by generalizations of the so-called Sum-Capture Theorem, which we prove using Fourier-analytic techniques.

Keywords: Provable Security, Key-alternating Ciphers

1 Introduction

Key-alternating ciphers (KACs) alternate the application of fixed, invertible, and key-independent permutations $P_1, \ldots, P_t$ on the n-bit strings with xor-ing $t+1$ n-bit sub-keys $s_0, s_1, \ldots, s_t$, i.e., the output of the KAC on input x and sub-keys $\boldsymbol{s} = (s_0, s_1, \ldots, s_t)$ is

$$\mathrm{KAC}_{\boldsymbol{s}}(x) = s_t + P_t(s_{t-1} + P_{t-1}(\cdots P_2(s_1 + P_1(s_0 + x))\cdots)) \ ,$$

where + denotes the bit-wise xor. Several modern block cipher designs are KACs – these include in particular Substitution-Permutation Networks (SPNs), like AES [10], PRESENT [3] and LED [14].

Most theoretical analyses of KACs [13,4,22,18,6,9,16] have proved their security as a (strong) pseudorandom permutation in a model where the permutations $P_1, \ldots, P_t$ are randomly and independently chosen, and can be queried by the adversary. Moreover, the sub-keys $\boldsymbol{s} = (s_0, s_1, \ldots, s_t)$ are also chosen

independently.[1] These results show that the number of queries q (to the keyed construction, as well as to the permutations) needed to break the construction is roughly $q = N^{t/(t+1)}$ (where $N = 2^n$), which has been shown to be optimal.

THIS PAPER: SECURITY WITH CORRELATED SUB-KEYS. Real sub-keys are however *not* independent, and are generated from a shorter key using a specific *key schedule* algorithm. However, very little progress has been made in understanding when such key schedules are secure, and independence assumptions are common even in cryptanalysis. In this paper, we therefore ask the following question:

> *For which distributions of sub-keys can we still obtain optimal security against $q = N^{t/(t+1)}$ queries?*

We note that this question was partially addressed by Dunkelman *et al.* [11] for $t = 1$ and later by Chen *et al.* [5], who proved such bounds for the case where $t = 2$, and the subkeys satisfy the constraint $s_0 = s_1 = s_2$.[2] Here, we consider the extension of their work beyond three rounds.

We also stress that our goal is not that of finding practical key schedules which are comparable to those used in actual block cipher designs. Rather, we aim for a broader understanding of correlated key schedules, and when they preserve optimal security. We also point out that with respect to our current state of knowledge, even modest savings in randomness to generate the keys are not known for multi-round KACs.

REDUCING KEY DEPENDENCE FOR ARBITRARY ROUNDS. As our first contribution, we show that for *any* t-round KAC with $t + 1$ subkeys, there are key schedules that merely depend on $t-1$ independent and uniform keys that achieve $q = \Omega(N^{t/(t+1)})$ security. This generalizes the result for $t = 2$ proved by Chen et al. [5] to multi-round instantiations.

We give a general sufficient condition on key distributions for $\boldsymbol{s}$ that achieve optimal security – specifically our condition considers distributions where the $t + 1$ subkeys $\boldsymbol{s}$ for the t-round KAC are a linear function of a vector $\boldsymbol{k}$ of $t - 1$ "master" keys, denoted as $\boldsymbol{s} = A\boldsymbol{k}$, in which we view each master key and subkey as an element of the field $\mathbb{F}_{2^n}$. The sufficient conditions for the key schedules are, in particular, as follows:

1. Any $t - 2$ rows of A forms a matrix of rank $t - 2$.
2. For any t rows of A,
   - the t rows form a matrix of rank $t - 1$.
   - there exists a linear combination of the t rows such that it gives zero vector and there are two neighboring rows with non-zero coefficients.

[1] In fact, Chen and Steinberger [6] already noted that their result holds in the case where the underlying subkeys are t-wise independent. The tight concrete bound proved by Hoang and Tessaro [16] also extends to t-wise independent setting.

[2] Actually, Chen *et al.* [5] also addressed reducing the number of keys and permutations in parallel. They showed that a 2-round KAC is secure against $q = \Omega(N^{2/3})$ queries when instantiated by a single permutation and a single key with a key schedule built over a linear orthomorphism.

For example, a suitable and natural key schedule that satisfies our condition is the one where $\boldsymbol{s}$ is from the $(t-1)$-wise independent distribution obtained by evaluating a random polynomial of degree $t-2$ at $t+1$ distinct points over $\mathbb{F}_{2^n}$. In fact, while our condition on key schedules is more restrictive than $(t-2)$-wise independence, it still allows for simple key schedules for small rounds (e.g. $t=3$ and $t=4$) that do not require field multiplication, which may be considered an expensive operation, i.e., for $t=3$, we show that one can set $\boldsymbol{s}=(k_0,k_0,k_1,k_1)$ to have $q=\Omega(N^{3/4})$. For $t=4$, we set $\boldsymbol{s}=(k_0,k_1,k_2,k_0+k_1,k_1+k_2)$ to have $q=\Omega(N^{4/5})$.

Less independence for more rounds. Of course, we would like to understand whether even more randomness can be saved. We make progress by saving n more bits for a sufficiently large number of rounds. Again, we give a general condition on distributions characterized by linear functions mapping $t-2$ n-bit keys $\boldsymbol{k}$ to $t+1$ keys $\boldsymbol{s}$, i.e., $\boldsymbol{s}=A\boldsymbol{k}$. For any linear mapping A satisfying the property that each $t-2$ rows of A have rank $t-2$, our security proof shows, for $t>5$, a bound that gives security strictly better than $q=\Omega(N^{(t-1)/t})$ and for $t\geqslant 8$, we achieve $q=\Omega(N^{t/(t+1)})$ security. Again, one particular instantiation is obtained by evaluating a random polynomial of degree $t-3$ at $t+1$ distinct points over $\mathbb{F}_{2^n}$.

How far can we go? The end question is of course whether we can push our results even further. Ideally, it would be possible to use a single-key schedule (as in Chen *et al.*) for an arbitrary number of rounds. However, as we explain below, the classical approach to prove security for limited independence is via so-called "sum-capture theorems" [2,23]. In the paper below, we show that the sum-capture theorem necessary to study the trivial key schedule beyond two rounds is not true. This, of course, does *not* mean that the resulting construction is insecure, but improving beyond the results of this paper would require substantially new counting techniques. (See Section 4.3)

Other related works. Another aspect of theoretical analyses over KACs is to reduce the number of random permutations used in the construction. Recently, Wu *et al.* [24] showed that for a three round KAC instantiated with four uniform and independent subkeys and a single random permutation is secure against $q=\Omega(N^{3/4})$ adversarial queries. Dutta [12] considered minimizing the tweakable KAC by reducing the number of random permutations and proves the security of $q=\Omega(N^{2/3})$ for the 2-round tweakable KAC by Cogliati *et al.* [7] and 4-round tweakable KAC by Cogliati and Seurin [8].

1.1 Technical Overview

Our paper follows the well-established paradigm of proving security of key-alternating ciphers based on the *expectation method* by Hoang and Tessaro [16], combined with generalizations of sum-capture theorems as proposed by Chen *et al.* [5].

Chain-based analyses. The core of existing analyses proceeds by identifying a set of *bad* transcripts which contains so-called *chains* – these are transcripts

where the adversary has made direct queries to $P_1, P_2, \ldots, P_t$, and/or to the construction, which are linked together by the chosen subkeys. In the ideal world, such bad transcript would likely become inconsistent with the real world. i.e., the probability of obtaining the bad transcript from the real world can be zero. Formally, we represent a transcript as $\tau = (\mathcal{Q}_E, \mathcal{Q}_1, \ldots, \mathcal{Q}_t, \boldsymbol{k})$, where $\mathcal{Q}_E$ contains queries to the construction, and $\mathcal{Q}_i$'s are the queries to the individual permutations. Further, $\boldsymbol{k}$ are the keys from which the actual sub-keys $\boldsymbol{s} = (s_0, s_1, \ldots, s_t)$ are generated. (As our statements are independent of whether such queries occurred in the forward or in the backward direction, and of their order, we think of the transcript as being made of sets of input-output pairs.) We say that such a τ is *bad* if the subkeys $(s_0, s_1, \ldots, s_t)$ are such that there exist queries $(u_{t+1}, v_0) \in \mathcal{Q}_E, (u_1, v_1) \in \mathcal{Q}_1, \ldots, (u_t, v_t) \in \mathcal{Q}_t$ which constitutes a chain, i.e., if there exists an index i, such that for all $j \in \{0, \ldots, t\}$ satisfying $j \neq i$, one has $v_j + u_{j+1} = s_j$, then we say they form the i-th type of chain. If the sub-keys $\boldsymbol{s}$ are independent and uniform, then the number of chains is at most $(t+1) \cdot q^{t+1}$ (by a simple union bound over all types of chain), and thus, the probability that the transcript is bad is at most $O((t+1)q^{t+1}/N^t)$. (Note that every chain definition only involved t subkeys.)

Handling limited independence. This argument however does not work if $\boldsymbol{s}$ is generated (say) from $(t-1)$-wise independent and uniform n-bit keys, as we can expect (at best) to prove $O((t+1)q^{t+1}/N^{t-1})$. We resolve this by considering a generalized version of the sum-capture quantity which allows us to give tighter bound over the number of chains, namely we define

$$\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) := \left| \left\{ (v_0, (u_1, v_1), \ldots, (u_{t-1}, v_{t-1}), u_t) \in V_0 \times \mathcal{Q}_1 \times \ldots \times \mathcal{Q}_{t-1} \times U_t \ : \ \sum_{i=0}^{t-1} c_i (v_i + u_{i+1}) = 0 \right\} \right| \quad (1)$$

where $V_0, U_t \subseteq \{0,1\}^n$ and the coefficients $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ are field elements of $\mathbb{F}_{2^n}$. A bound on this quantity can be used to bound the number of chains in a non-trivial fashion, as long as the coefficients arising are compatible with the underlying method to generate the sub-keys and satisfy certain conditions (which in turn will give our characterization of which distributions actually give the desired optimal security).

Concretely, when the linear coefficients $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ satisfies the condition that there is an index $0 \leqslant \mathsf{idx} < t-1$ such that $c_{\mathsf{idx}} \neq 0$ and $c_{\mathsf{idx}+1} \neq 0$, we prove the tight bound $\mu_{\boldsymbol{c}} = \Theta(q^{t+1}/N)$ using Fourier Analysis techniques.

Reducing key dependencies further. To obtain our results for construction with subkeys generated from $t-2$ independent and uniform keys, we need to upper bound an even more restrictive version of the above sum-capture quantity where two linear constraints are imposed, i.e.,

$$\mu_{\boldsymbol{c},\boldsymbol{d}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) := \left| \left\{ (v_0, (u_1, v_1), \ldots, (u_{t-1}, v_{t-1}), u_t) \in V_0 \times \mathcal{Q}_1 \times \ldots \times \mathcal{Q}_{t-1} \times U_t \ : \right. \right. \\ \left. \left. \sum_{i=0}^{t-1} c_i(v_i + u_{i+1}) = 0 \, , \sum_{i=0}^{t-1} d_i(v_i + u_{i+1}) = 0 \right\} \right| \quad (2)$$

For the 2-constraint case, we in particular look at the coefficients $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ and $\boldsymbol{d} = (d_0, \ldots, d_{t-1})$ that characterize the underlying subkeys generated via the linear key schedule being $(t-2)$-wise independent and uniform. We then show that, with the subkeys generated from $t-2$ uniform and independent n-bit keys via a linear key schedule:

- for $t > 5$, the t-round KAC is secure against $q = \omega(N^{\frac{t-1}{t}})$ queries.
- for $t \geqslant 8$, the t-round KAC has tight security bound (i.e., $q = \Omega(N^{\frac{t}{t+1}})$)

Given that (2) is a natural generalization of its one constraint counterpart, it is tempting to conclude that upper-bounding (2) is not harder than upper-bounding (1). However, as the number of constraints becomes two, we stress that the problem of upper-bounding (2) is much harder. Moreover, the tightness of upper-bounding (1) crucially relies on a particular step which was referred to as the "Cauchy-Schwartz trick" [2,23,5], which does not seem to apply here. We bypass this limitation by introducing a novel representation for the upper bound of (2) as the 2-norm of a matrix. In particular, one can interpret the Cauchy-Schwartz trick upper bound as essentially a special case of the matrix norm bound in which each row and each column of the matrix contains at most one non-zero entry. Then we use the matrix Frobenius norm which is easier to compute for bounding the matrix 2-norm. Though our current technique only proves tight security bound for $t \geqslant 8$, we believe that the matrix 2-norm is the right characterization and one can extend the tightness result to $t \geqslant 4$ via a better tool to derive the 2-norm bound, as the usage of Frobenius norm is, in most cases, not tight[3].

While (2) remains to be a promising candidate to consider for saving two keys, we show that for $t = 3$, i.e., for the 3-round KAC with identical subkey and independent permutations, the quantity of (2) is lower bounded by q^3/N with good probability. Hence, a sum capture quantity with highly non-trivial characterizations or an alternative proof strategy for the 3-round KAC is needed to obtain the desired $q = \Omega(N^{3/4})$ security bound.

Good transcript analysis. As we have bounded the probability of a transcript being bad, we move to understand the remaining transcripts which we consider as good. We rely on the expectation method proposed by Hoang and

[3] In fact, the Frobenius norm and 2-norm can have up to $\sqrt{N}$ multiplicative gap for $N \times N$ matrix (e.g. the identity matrix), and we believe that a large gap exists in our Frobenius norm bound. However, to get a better 2-norm bound, it requires a much better understanding to our defined matrix for analyzing (2) than we do.

Tessaro [16], which is a generalization of the H-coefficient method [6,21]. In the expectation method, the final security upper bound is

$$\mathsf{Security\ bound} \leqslant \mathbb{E}_{X_1}[g(X_1)] + \mathsf{Pr}[X_1 \text{ is bad}]$$

in which X_1 is the random variable representing the transcript generated from the adversary interacting with the ideal world, and $g : \mathcal{T} \to [0, +\infty)$ is a non-negative function such that $g(\tau)$ upper bounds the real-world-ideal-world probability ratio of any good transcript τ. The goal is find a function $g : \mathcal{T} \to [0, +\infty)$ so that the value of $\mathbb{E}_{X_1}[g(X_1)]$ is minimized.

It is tempting to believe that the subkeys are needed to be at least t-wise independent and uniform when applying the techniques in [16] to achieve the tight security bound for the good transcripts. However, surprisingly, we show (in Section 5) that as long as the underlying subkeys $\boldsymbol{s} = (s_0, \ldots, s_t)$ are $(t-2)$-wise independent and uniform, we can pick a non-negative function g so that

$$\mathbb{E}_{X_1}[g(X_1)] \leqslant O(q^{t+1}/N^t) \,.$$

Therefore, as long as the t-round KAC has a key schedule that gives $(t-2)$-wise independent and uniform subkeys, our result on the good transcript analysis can be applied as black-box.

1.2 Paper Organization

In Section 2 we define some basic notations and indistinguishability framework. In Section 3 we give the main theorems and show tight security for classes of t-round KAC. In Section 4 we analyze the sum capture quantity for upper-bounding the number of bad transcripts. Then we provide analysis for good transcripts in Section 5 and wrap up proof of theorems in Section 6. Finally we provide conclusions and open problems in Section 7.

2 Preliminaries

NOTATIONS. For a finite set S, we write $x \xleftarrow{\$} S$ to denote that x receives a uniformly sampled value from S. For an algorithm A, we write $y \leftarrow A(x_1, \ldots; r)$ to denote that A takes $x_1, \ldots$ as inputs and runs with the randomness r and assigning the output to y. We let $y \xleftarrow{\$} A(x_1, \ldots,)$ be that A, given the inputs, is executed over a randomly chosen r and the resulting value is assigned to y.

We use $\mathbb{F}_p$ to denote a finite field of size p. For any two elements $u, v \in \{0,1\}^n$, we use $\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i$ to denote the inner product of u and v, where u_i, v_i are the i-th bit of u, v respectively. For any number $1 \leqslant b \leqslant a$, we write $a^{(b)} = a(a-1)\cdots(a-b+1)$ and take $a^{(0)} = 1$ by convention. In all the following, for any two elements $u, v \in \{0,1\}^n$, we take $u + v$ and uv as the field addition and multiplication in $\mathbb{F}_{2^n}$ respectively, in which $u + v$ is implemented as the bit-wise xor over $\{0,1\}^n$. For a fixed n, we write $N = 2^n$. For any vector u and matrix A, we write $u^\top$ and $A^\top$ as their transpose.

PRP SECURITY OF BLOCK CIPHERS. We study the security of the Key Alternating Cipher in the random permutation model. Let $E : \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ be a blockcipher, which is constructed over a set of independent, random permutations $\boldsymbol{P} = (P_1, P_2, \ldots, P_t)$. Let $\mathcal{A}$ be an adversary, the strong PRP advantage of $\mathcal{A}$ is defined as

$$\mathsf{Adv}^{\pm\mathsf{prp}}_{E[\boldsymbol{P}]}(\mathcal{A}) := \mathsf{Pr}[K \xleftarrow{\$} \mathcal{K} : \ \mathcal{A}^{E[\boldsymbol{P}],\boldsymbol{P}} = 1] - \mathsf{Pr}[\mathcal{A}^{P_0,\boldsymbol{P}} = 1]$$

in which P_0 is a random permutation independent of $\boldsymbol{P}$, and "$\pm$" denotes that the adversary $\mathcal{A}$ can query the oracles in both forward direction and backward direction.

INDISTINGUISHABILITY FRAMEWORK. We consider a computationally unbounded distinguisher $\mathcal{A}$ interacting with two systems $\mathbf{S}_0$ and $\mathbf{S}_1$. The interaction between $\mathcal{A}$ and $\mathbf{S}_b$ (where $b \in \{0,1\}$) defines a transcript $\tau = ((u_1, v_1), \ldots, (u_q, v_q))$ that records the q pairs of queries/replies $\mathcal{A}$ made to/received from the system $\mathbf{S}_b$. Let X_b be the random variable representing the transcript, then the goal is to upper bound the following statistical distance

$$\Delta(X_0, X_1) = \sum_{\tau} \max\{0, \mathsf{Pr}[X_1 = \tau] - \mathsf{Pr}[X_0 = \tau]\} \ .$$

FORMULATING SYSTEMS. We follow [19] to describe the system behavior of $\mathbf{S}$ by associating every possible transcript $\tau = ((u_1, v_1), \ldots, (u_q, v_q))$ with a value $\mathsf{p}_{\mathbf{S}}(\tau) \in [0,1]$. One can interpret $\mathsf{p}_{\mathbf{S}}(\tau)$ as the probability that, if the queries $u_1, \ldots, u_q$ in τ are asked sequentially, $\mathbf{S}$ answers with $v_1, \ldots, v_q$ respectively. Note that $\mathsf{p}_{\mathbf{S}}(\cdot)$ is defined only by the underlying system $\mathbf{S}$ and is hence independent of any distinguisher. We also note that $\mathsf{p}_{\mathbf{S}}(\cdot)$ is not a probability distribution over the transcripts, as the sum over all $\mathsf{p}_{\mathbf{S}}(\tau)$ does not necessarily give one.

Since the distinguisher is computationally unbounded, it is sufficient to consider deterministic distinguishers only. Fix any deterministic distinguisher $\mathcal{A}$, let X denote the transcript distribution of $\mathcal{A}$ interacting with $\mathbf{S}$, then it holds that $\mathsf{Pr}[X = \tau] \in \{0, \mathsf{p}_{\mathbf{S}}(\tau)\}$ for any τ because, either $\mathcal{A}$ issues the queries $u_1, \ldots, u_q$ when given the answers $v_1, \ldots, v_q$, leading to $\mathsf{Pr}[X = \tau] = \mathsf{p}_{\mathbf{S}}(\tau)$, or it does not, resulting in $\mathsf{Pr}[X = \tau] = 0$.

Let $\mathcal{T}$ be the set of transcripts τ that has $\mathsf{Pr}[X_1 = \tau] > 0$. Further noting that $\mathsf{Pr}[X_0 = \tau] = \mathsf{p}_{\mathbf{S}_0}(\tau)$ if $\tau \in \mathcal{T}$, we can rewrite the statistical distance as

$$\Delta(X_0, X_1) = \sum_{\tau} \max\{0, \mathsf{p}_{\mathbf{S}_1}(\tau) - \mathsf{p}_{\mathbf{S}_0}(\tau)\} = \sum_{\tau} \mathsf{p}_{S_1}(\tau) \cdot \max\left\{0, 1 - \frac{\mathsf{p}_{\mathbf{S}_0}(\tau)}{\mathsf{p}_{\mathbf{S}_1}(\tau)}\right\} \ .$$

THE EXPECTATION METHOD. In this part we review the expectation method proposed by [16], which is developed based on the H-coefficient method [6,21]. In the H-coefficient method, the set of transcript $\mathcal{T}$ is partitioned into $\mathcal{T}_{\mathsf{good}}$ and $\mathcal{T}_{\mathsf{bad}}$ so that for any $\tau \in \mathcal{T}_{\mathsf{good}}$, $\mathsf{p}_{\mathbf{S}_0}(\tau)/\mathsf{p}_{\mathbf{S}_1}(\tau) \geqslant 1 - \varepsilon$ for some carefully chosen parameter ε. Then, an upper bound of the advantage directly follows. i.e.,

$$\Delta(X_0, X_1) \leqslant \varepsilon + \mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}] \ .$$

However, instead of giving a uniform bound over all good transcripts, we can associate each τ with a non-negative value $g(\tau)$ so that $\mathsf{p}_{\mathbf{S}_0}(\tau)/\mathsf{p}_{\mathbf{S}_1}(\tau) \geqslant 1-g(\tau)$ for every $\tau \in \mathcal{T}_{\mathsf{good}}$. Hence we can instead, derive the upper bound as

$$\Delta(X_0, X_1) \leqslant \sum_{\tau \in \mathcal{T}_{\mathsf{good}}} \mathsf{p}_{\mathbf{S}_1}(\tau) \cdot g(\tau) + \sum_{\tau \in \mathcal{T}_{\mathsf{bad}}} \mathsf{p}_{\mathbf{S}_1}(\tau) \leqslant \mathbb{E}_{X_1}[g(X_1)] + \sum_{\tau \in \mathcal{T}_{\mathsf{bad}}} \mathsf{p}_{\mathbf{S}_1}(\tau) \ ,$$

where we can take the expectation over all $\tau \in \mathcal{T}$ by the fact that $g(\cdot)$ is non-negative. Therefore, we have the following lemma.

Lemma 1 (The expectation method). *If there exists a partition of $\mathcal{T} = \mathcal{T}_{\mathsf{good}} \sqcup \mathcal{T}_{\mathsf{bad}}$, and a function $g : \mathcal{T} \to [0, +\infty)$ such that for any $\tau \in \mathcal{T}_{\mathsf{good}}$, $\mathsf{p}_{\mathbf{S}_0}(\tau)/\mathsf{p}_{\mathbf{S}_1}(\tau) \geqslant 1 - g(\tau)$, then*

$$\Delta(X_0, X_1) \leqslant \mathbb{E}_{X_1}[g(X_1)] + \mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}] \ .$$

3 Main Results

We consider the PRP security of t-round Key Alternating Cipher (KAC) that is built on t random permutations $\boldsymbol{P} = (P_1, \ldots, P_t)$ over $\{0,1\}^n$ and $t+1$ subkeys $(s_0, \ldots, s_t)$ in which $s_i \in \{0,1\}^n$. The t-round KAC, when given input $M \in \{0,1\}^n$, outputs

$$s_t + P_t(s_{t-1} + P_{t-1}(\cdots P_1(s_0 + M)\cdots)) \ .$$

The subkeys are generated from the master key denoted as $(k_0, \ldots, k_w)$ in which k_i are sampled from $\{0,1\}^n$ uniformly and independently. Therefore, the length of the master key is $(w+1)n$ bits. Here we consider only linear key schedule algorithms, which can be represented as a matrix A over $\mathbb{F}_{2^n}$. We define the column vectors $\boldsymbol{s} = (s_0, \ldots, s_t)^\top$ and $\boldsymbol{k} = (k_0, \ldots, k_w)^\top$ in which we naturally take each n-bit string as an element in $\mathbb{F}_{2^n}$ and use $\boldsymbol{s} \leftarrow A\boldsymbol{k}$ to denote the key-scheduling process.

The case of A being an identity matrix of size $(t+1) \times (t+1)$ has been well studied, i.e. it was proved in [6,16] that, when the subkeys $s_0, \ldots, s_t$ are independent and uniform and the permutations $P_1, \ldots, P_t$ are independent, any adversary needs at least $q = \Omega(N^{t/(t+1)})$ queries to achieve constant distinguishing advantage. Here we consider the case in which the permutations are independent but the subkeys are correlated and are generated via linear key schedules from $t-1$ independent n-bit keys (considered Theorem 1) or $t-2$ independent n-bit keys (Theorem 2).

We starts with providing security bound of t-round KAC for a class of key schedules that generate $t+1$ subkeys from $t-1$ independent keys.

Theorem 1. *For the t-round KAC constructed over t random permutations $\boldsymbol{P} = (P_1, \ldots, P_t)$, let the key of KAC be $\boldsymbol{k} = (k_0, k_1, \ldots, k_{t-2})^\top$ in which k_i's are independently uniformly sampled from $\mathbb{F}_{2^n}$. Let subkeys $\boldsymbol{s} = (s_0, s_1, \ldots, s_t)^\top$ be derived by $\boldsymbol{s} \leftarrow A\boldsymbol{k}$ in which A is a $(t+1) \times (t-1)$ matrix over $\mathbb{F}_{2^n}$, with the rows denoted as $A_0, \ldots, A_t$, such that*

1. *Any $t-2$ rows of A forms a matrix of rank $t-2$.*
2. *For any $I \subseteq \{0,\ldots,t\}$, $|I| = t$, then the row vectors $(A_\ell)_{\ell\in I}$ satisfy that*
   - *$(A_\ell)_{\ell\in I}$ forms a matrix of rank $t-1$.*
   - *there exists values $(c_\ell)_{\ell\in I}$ such that $\sum_{\ell\in I} c_\ell A_\ell = \mathbf{0}$ and there are two indices $\mathsf{idx}_1, \mathsf{idx}_2 \in I$ satisfying $\mathsf{idx}_1 - \mathsf{idx}_2 \in \{1, t\}$ and $c_{\mathsf{idx}_1}, c_{\mathsf{idx}_2}$ are both non-zero.*

Then for any adversary $\mathcal{A}$ that issues at most q queries to $\mathsf{KAC}, P_1, \ldots, P_t$, where $9(t+2)n \leqslant q \leqslant N/4$,

$$\mathsf{Adv}^{\pm\mathsf{prp}}_{\mathsf{KAC}[\boldsymbol{P}]}(\mathcal{A}) \leqslant (t^2+t+1)\cdot\frac{4q^{t+1}}{N^t} + 3(t+1)\sqrt{\frac{q^{2t-1}(t+2)n}{N^{2t-2}}}\,.$$

First, we give a key schedule that gives $(t-1)$-wise independent and uniform subkeys for arbitrary t-round KAC.

Corollary 1. *For $t < 2^n$, pick distinct elements $\alpha_0, \ldots, \alpha_t \in \mathbb{F}_{2^n}$, and let subkey $s_i = F(\alpha_i)$ in which $F(X) = \sum_{j=0}^{t-2} k_j \cdot X^j$, then an adversary needs $\Omega(N^{t/(t+1)})$ queries to achieve constant distinguishing advantage.*

Corollary 1 directly follows from the fact that A is a Vandermonde matrix so that every $t-1$ rows of A forms a full-rank sub-matrix. Hence, any t rows of A are linear dependent with the coefficients $(c_\ell)_{\ell\in I}$ satisfying $c_\ell \neq 0$ for all ℓ.

Note that by letting $t = 2$ in Corollary 1, our result implies the optimal security bound of 2-round KAC with identical subkeys and independent permutations proven by Chen *et al.* [5].

Though it is implied in the theorem statement that we need the subkeys being $(t-2)$-wise independent and uniform, for small round t, we still can obtain some simple key schedules that achieve the optimal bound for q while do not require any field multiplication operations, which may be considered an expensive operation in key-scheduling.

Corollary 2. *Let the 3-round KAC be with key schedule*

$$\boldsymbol{s} = (k_0, k_0, k_1, k_1)$$

in which k_0, k_1 are two independently uniform n-bit keys, then an adversary needs $\Omega(N^{3/4})$ queries to achieve constant distinguishing advantage.

Corollary 3. *Let the 4-round KAC be with key schedule*

$$\boldsymbol{s} = (k_0, k_1, k_2, k_0+k_1, k_1+k_2)$$

in which k_0, k_1, k_2 are three independently uniform n-bit keys, then an adversary needs $\Omega(N^{4/5})$ queries to achieve constant distinguishing advantage.

One can check that the subkeys in Corollary 2 (respectively Corollary 3) are 1-wise (pairwise) independent and uniform, and any t rows forms a sub-matrix

Table 1. $q = \Omega(N^{\lambda})$ for constant security bound in Theorem 2.

t	3	4	5	6	7	8	9	10	$\cdots$
$\lambda = \log_N q$	0.571	0.720	0.800	0.842	0.870	0.889	0.9	0.909	$\cdots$
$t/(t+1)$	0.750	0.800	0.833	0.857	0.875	0.889	0.9	0.909	$\cdots$

of rank $t-1$ with the coefficients $(c_\ell)_{\ell \in I}$ satisfying the given conditions via Gaussian elimination.

As Theorem 1 gives tight bound for all t, one may optimistically expect similar results can be proved with ease when saving one more key. However, for the t-round KAC with subkeys generated from $t-2$ keys, we are only able to make partial progress and prove the following theorem that only implies tight security for $t \geqslant 8$.

Theorem 2. *For the t-round* KAC *constructed over t random permutations $\boldsymbol{P} = (P_1, \ldots, P_t)$, let the key of KAC be $\boldsymbol{k} = (k_0, k_1, \ldots, k_{t-3})^\top$ in which k_i's are independently and uniformly sampled from $\mathbb{F}_{2^n}$. Let subkeys $\boldsymbol{s} = (s_0, s_1, \ldots, s_t)^\top$ be derived by $\boldsymbol{s} = A\boldsymbol{k}$ in which A is a $(t+1) \times (t-2)$ matrix over $\mathbb{F}_{2^n}$ such that any $t-2$ rows of A forms a matrix of rank $t-2$. Then for any adversary $\mathcal{A}$ that issues at most $(t+2)nN^{2/3} \leqslant q \leqslant N/4$ queries to* KAC, $P_1, \ldots, P_t$,

$$\mathsf{Adv}^{\pm\mathsf{prp}}_{\mathsf{KAC}[\boldsymbol{P}]}(\mathcal{A}) \leqslant (t^2+2t) \cdot \frac{(5q)^{t+1}}{N^t} + (t+1)^2 \cdot \frac{(3q)^{2t-2.5}}{N^{2t-4}} \ .$$

Table 1 summarizes the order of q that leads the security bound to $\Omega(1)$. We can observe that, initially Theorem 2 does not give good bound for $t \leqslant 7$. From $t \geqslant 5$, the bound starts getting better than $q = \Omega(N^{(t-1)/t})$ which can be obtained by instantiating a $(t-1)$-round KAC from the provided $t-2$ keys and applying Theorem 1. When $t \geqslant 8$, the bound achieves the optimal $q = \Omega(N^{t/(t+1)})$. The tightness results for $t \leqslant 7$ are left open.

A feasible instantiation of Theorem 2 is to let the subkeys be the evaluations at $t+1$ distinct points of a degree $t-3$ polynomial. Then the following corollary holds.

Corollary 4. *For $8 \leqslant t < 2^n$, pick distinct elements $\alpha_0, \ldots, \alpha_t \in \mathbb{F}_{2^n}$, and let subkey $s_i = F(\alpha_i)$ in which $F(X) = \sum_{j=0}^{t-3} k_j \cdot X^j$, then an adversary needs $\Omega(N^{t/(t+1)})$ queries to achieve constant distinguishing advantage.*

Proof framework. We will use the expectation method (i.e. Lemma 1) to show both theorems. Given the query record $\boldsymbol{\mathcal{Q}} = (\mathcal{Q}_E, \mathcal{Q}_1, \ldots, \mathcal{Q}_t)$, we will be generous and allow the adversary $\mathcal{A}$ to see the key $\boldsymbol{k}$ after making all the queries. Therefore, we let the transcript $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k})$ by attaching $\boldsymbol{k}$ to the end of $\boldsymbol{\mathcal{Q}}$. In the ideal world, we sample and attach a dummy key $\boldsymbol{k}$ to $\boldsymbol{\mathcal{Q}}$. Here we define the set of bad transcript for the t-round KAC.

Definition 1 (Bad transcripts). *For a t-round* KAC*, we say a transcript $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k})$ is bad if*

$$\boldsymbol{k} \in \mathsf{Badkey}_{\boldsymbol{\mathcal{Q}}} = \bigcup_{i=0}^{t} \mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},i}$$

in which for every i,

$$\begin{aligned}\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},i} := \{\boldsymbol{k} :\ & \boldsymbol{s} \leftarrow \mathsf{KeySchedule}(\boldsymbol{k}), \textit{there exists } (u_{t+1}, v_0) \in \mathcal{Q}_E, \\ & (u_1, v_1) \in \mathcal{Q}_1, \ldots, (u_t, v_t) \in \mathcal{Q}_t \\ & \textit{s.t. for all } 0 \leqslant j \leqslant t,\ j \neq i,\ v_j + s_j = u_{j+1}\}\ ,\end{aligned}$$

otherwise we say τ is good. We use $\mathcal{T}_{\mathsf{good}}$ to denote the set of all good transcripts and $\mathcal{T}_{\mathsf{bad}}$ to denote the set of all bad transcripts. Hence $\mathcal{T} = \mathcal{T}_{\mathsf{good}} \sqcup \mathcal{T}_{\mathsf{bad}}$.

Then, we break the analysis into the bad transcript case and the good transcript case. We will use the generalized sum capture quantity in Section 4 as an upper bound for the bad transcripts. We analyze the good transcripts in Section 5. The final proof of theorems will be presented in Section 6.

More fine-grained security. In the above theorems, we use q to be the uniform upper bound over all kinds of queries. However, we note that our proof technique also provides bounds when the number of cipher queries q_e and the number of permutation queries q_p are separated. We provide the bounds in the full version for both theorems.

4 Generalized Sum Capture Quantity for KAC

In [5] Chen *et al.* considered minimizing the 2-round KAC, where they proved a variant of "sum-capture" results [2,15,1,17,23]. The results are often stated that, for a randomly chosen set A of size q, the quantity

$$\mu(A) := \max_{\substack{X,Y \subseteq \mathbb{Z}_2^n \\ |X|=|Y|=q}} |\{(a,x,y) \in A \times X \times Y : a = x + y\}| \tag{3}$$

is close to its expected value q^3/N (when A, X, Y are all chosen at random) with high probability. In the 2-round KAC with identical key schedule, the sum-capture quantity is defined as

$$\mu(\mathcal{Q}) := \max_{\substack{X,Y \subseteq \mathbb{Z}_2^n \\ |X|=|Y|=q}} |\{(x,(u,v),y) \in X \times \mathcal{Q} \times Y : x + u = v + y\}| \tag{4}$$

where one can view the query transcript $\mathcal{Q}$ that derived from the interaction of an adversary $\mathcal{A}$ with the permutation, equivalently as the set A in (3) defined by $A = \{u + v \mid (u,v) \in \mathcal{Q}\}$.

However, both (3) and (4) consider only a single random permutation with a single linear constraints. To generalize the sum capture quantity so that we can

handle the KAC that saves more keys, we consider the sum capture quantity that involves $(t-1)$ independently random permutations and $r \in \{1,2\}$ linear constraints over $\mathbb{F}_{2^n}$ for the t-round KAC with a linear key schedule.

For the $r = 1$ case, we are able to prove the tight bounds of sum capture quantity for any choice of linear constraint, leading to a feasible set of key schedule that enables saving two keys for arbitrary t-round KAC with tight security. However, as we increase the number of constraints to $r = 2$, the problem becomes more complicated and we do not have sophisticated technique to give a tight bound or handle arbitrary linear constraints. We are only able prove a loose upper bound for the linear-constraints that characterizes the underlying subkeys being $(t-2)$-wise independent, leading to partial result for saving three keys of t-round KAC.

Fourier Analysis. To prove the bounds, we will rely on the tool of Fourier analysis. In this part we define some notations for the Fourier analysis over $\{0,1\}^m$. Given a function $f : \{0,1\}^m \to \mathbb{R}$, the Fourier coefficient of f with $\alpha \in \{0,1\}^m$ is defined as

$$\hat{f}(\alpha) := \frac{1}{2^m} \sum_{x \in \{0,1\}^m} f(x)(-1)^{\langle \alpha, x \rangle} .$$

Then we have

$$f(x) = \sum_{\alpha \in \{0,1\}^m} \hat{f}(\alpha)(-1)^{\langle \alpha, x \rangle} . \tag{5}$$

For any set $S \subseteq \{0,1\}^m$, we let $\mathbb{1}_S : \{0,1\}^m \to \{0,1\}$ be the 0/1 indicator function of S. Then the following properties hold for $\mathbb{1}_S$:

$$\widehat{\mathbb{1}_S}(0) = \frac{|S|}{2^m} = \sum_{\alpha \in \{0,1\}^m} \widehat{\mathbb{1}_S}(\alpha)^2 , \tag{6}$$

$$\forall \alpha \in \{0,1\}^m : \ |\widehat{\mathbb{1}_S}(\alpha)| \leqslant \widehat{\mathbb{1}_S}(0) = \frac{|S|}{2^m} \ . \tag{7}$$

4.1 1-constraint Sum Capture Quantity

We let 1-constraint sum capture quantity be associated with a vector of coefficients $\boldsymbol{c} = (c_0, c_1, \ldots, c_{t-1})$, as

$$\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) := \left| \left\{ (v_0, (u_1, v_1), \ldots, (u_{t-1}, v_{t-1}), u_t) \in V_0 \times \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_{t-1} \times U_t \ : \ \sum_{j=0}^{t-1} c_j (v_j + u_{j+1}) = 0 \right\} \right| .$$

Lemma 2. *Let $t \geqslant 2$. Let $P_1, \ldots, P_{t-1}$ be $t-1$ independent uniformly random permutations of $\{0,1\}^n$, and let $\mathcal{A}$ be a probabilistic algorithm that makes adaptive queries to $P_1, \ldots, P_{t-1}$. Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}$ be the query transcripts of $P_1, \ldots, P_{t-1}$ interacting with $\mathcal{A}$. Let $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ be any coefficients so that there exists an index $0 \leqslant \mathsf{idx} < t-1$ satisfying $c_{\mathsf{idx}} \neq 0$ and $c_{\mathsf{idx}+1} \neq 0$, then for any $\mathcal{A}$ that makes at most q queries to each permutations,*

$$\mathsf{Pr}_{P_1,\ldots,P_{t-1}}\Bigg[\exists V_0, U_t \subseteq \mathbb{F}_{2^n}, |V_0| = |U_t| = q, \\ \mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \geqslant \frac{3q^{t+1}}{N} + 3q^{t-1/2}\sqrt{(t+2)n}\Bigg] \leqslant \frac{2t}{N^t} \,.$$

We let $\Phi(\mathcal{Q}_i) := \max_{\alpha\neq 0, \beta \neq 0} N^2 |\widehat{\mathbb{1}_{Q_i}}(\alpha, \beta)|$ for the query records $\mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}$.

To show Lemma 2, we will first rely on the following Lemma 3, which states the upper bound in terms of $\Phi(\mathcal{Q}_i)$ we just defined. Then we will apply the later stated Lemma 4 by Chen *et al.* [5] that provides an upper bound for the $\Phi(\mathcal{Q}_i)$ term to conclude the proof.

Lemma 3. *Fix any $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ such that $c_{\mathsf{idx}} \neq 0$ and $c_{\mathsf{idx}+1} \neq 0$ for some index $0 \leqslant \mathsf{idx} < t-1$, then for any subsets V_0, U_t with $|V_0| = |U_t| = q$,*

$$\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \leqslant \frac{q^{t+1}}{N} + q^{t-1}\Phi(\mathcal{Q}_{\mathsf{idx}+1}) \,.$$

Proof. The very first step is to write $\mu_{\boldsymbol{c}}$ as a sum over indicator functions, then we will perform Fourier transform over each indicator functions. The key point is that, even though the summation will be over many terms and Fourier coefficients, we can eliminate most of the summation term and simplify the equality so that it only sums over a single Fourier coefficient terms.

Here we sum over the indicator functions.

$$\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) = \sum_{v_0} \sum_{u_1, v_1} \cdots \sum_{u_{t-1}, v_{t-1}} \sum_{u_t} \mathbb{1}_{V_0}(v_0)\mathbb{1}_{\mathcal{Q}_1}(u_1, v_1) \cdots \\ \cdots \mathbb{1}_{\mathcal{Q}_{t-1}}(u_{t-1}, v_{t-1}) \cdot \mathbb{1}_{U_t}(u_t) \cdot \mathbb{1}_{\mathsf{Eq}}\left(0, \sum_{j=0}^{t-1} c_j(v_j + u_{j+1})\right)$$

in which $\mathbb{1}_{\mathsf{Eq}}(x, y)$ is the equality indicator function so that $\mathbb{1}_{\mathsf{Eq}}(x, y) = 1$ if and only if $x = y$. Note that for the equality indicator function, we can perform Fourier transformation and get

$$\mathbb{1}_{\mathsf{Eq}}(x, y) = \sum_{\alpha, \beta} \widehat{\mathbb{1}_{\mathsf{Eq}}}(\alpha, \beta) \cdot (-1)^{\langle \alpha, x\rangle + \langle \beta, y\rangle} = \frac{1}{N} \cdot \sum_{\alpha} (-1)^{\langle \alpha, x+y\rangle},$$

in which we use the fact that

$$\widehat{\mathbb{1}_{\mathsf{Eq}}}(\alpha, \beta) = \begin{cases} 1/N & \text{if } \alpha = \beta \\ 0 & o.w. \end{cases}$$

We expand each indicator function using Fourier transform and continue the calculation.

$$
\begin{aligned}
&\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \\
&= \sum_{\substack{v_0, u_1, v_1, \cdots \\ u_{t-1}, v_{t-1}, u_t}} \left(\sum_{\beta_0} \widehat{\mathbb{1}_{V_0}}(\beta_0)(-1)^{\langle \beta_0, v_0 \rangle} \right) \cdot \left(\sum_{\alpha_1, \beta_1} \widehat{\mathbb{1}_{\mathcal{Q}_1}}(\alpha_1, \beta_1)(-1)^{\langle \alpha_1, u_1 \rangle + \langle \beta_1, v_1 \rangle} \right) \cdot \\
&\quad \cdots \left(\sum_{\alpha_{t-1}, \beta_{t-1}} \widehat{\mathbb{1}_{\mathcal{Q}_{t-1}}}(\alpha_{t-1}, \beta_{t-1})(-1)^{\langle \alpha_{t-1}, u_{t-1} \rangle + \langle \beta_{t-1}, v_{t-1} \rangle} \right) \\
&\quad \cdot \left(\sum_{\alpha_t} \widehat{\mathbb{1}_{U_t}}(\alpha_t)(-1)^{\langle \alpha_t, u_t \rangle} \right) \cdot \frac{1}{N} \left(\sum_{\gamma} (-1)^{\langle \gamma, \sum_{j=0}^{t-1} c_j (v_j + u_{j+1}) \rangle} \right) .
\end{aligned}
$$

Here, notice that all Fourier coefficients only depend on the variables αs, βs and γ, so we can expand the multiplication and change the order of summation, and we obtain the following

$$
\begin{aligned}
&\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \\
&= \frac{1}{N} \sum_{\beta_0} \sum_{\alpha_1, \beta_1} \cdots \sum_{\alpha_{t-1}, \beta_{t-1}} \sum_{\alpha_t} \sum_{\gamma} \widehat{\mathbb{1}_{V_0}}(\beta_0) \widehat{\mathbb{1}_{Q_1}}(\alpha_1, \beta_1) \cdots \widehat{\mathbb{1}_{Q_{t-1}}}(\alpha_{t-1}, \beta_{t-1}) \widehat{\mathbb{1}_{U_t}}(\alpha_t) \\
&\quad \cdot \sum_{v_0} \sum_{u_1, v_1} \cdots \sum_{u_{t-1}, v_{t-1}} \sum_{u_t} (-1)^{\langle \beta_0, v_0 \rangle} (-1)^{\langle \alpha_1, u_1 \rangle + \langle \beta_1, v_1 \rangle} \ldots \\
&\quad \cdots (-1)^{\langle \alpha_{t-1}, u_{t-1} \rangle + \langle \beta_{t-1}, v_{t-1} \rangle} \cdot (-1)^{\langle \alpha_t, u_t \rangle} \cdot (-1)^{\langle \gamma, \sum_{j=0}^{t-1} c_j (v_j + u_{j+1}) \rangle} \\
&= \frac{1}{N} \sum_{\beta_0} \sum_{\alpha_1, \beta_1} \cdots \sum_{\alpha_{t-1}, \beta_{t-1}} \sum_{\alpha_t} \sum_{\gamma} \widehat{\mathbb{1}_{V_0}}(\beta_0) \widehat{\mathbb{1}_{Q_1}}(\alpha_1, \beta_1) \cdots \widehat{\mathbb{1}_{Q_{t-1}}}(\alpha_{t-1}, \beta_{t-1}) \\
&\quad \cdot \widehat{\mathbb{1}_{U_t}}(\alpha_t) \cdot \left(\sum_{v_0} (-1)^{\langle \beta_0, v_0 \rangle + \langle \gamma, c_0 v_0 \rangle} \right) \cdot \left(\sum_{u_1} (-1)^{\langle \alpha_1, u_1 \rangle + \langle \gamma, c_0 u_1 \rangle} \right) \\
&\quad \left(\sum_{v_1} (-1)^{\langle \beta_1, v_1 \rangle + \langle \gamma, c_1 v_1 \rangle} \right) \cdots \left(\sum_{u_t} (-1)^{\langle \alpha_t, u_t \rangle + \langle \gamma, c_{t-1} u_t \rangle} \right)
\end{aligned}
$$

The last equality is simply grouping the inner products that share the same u, v terms together. Note that the field multiplication of $c \cdot x$ can be represented as a matrix A_c[4] that applies to an n-dimensional vector x over $\mathbb{F}_2$. If $c = 0$, then $A_c = O$ where we use O to denote an all zero matrix, otherwise A_c is a full-rank matrix. Taking the summation over the v_0 term as an example, we rewrite the

[4] Since we are taking the natural field interpretation over $\{0,1\}^n$, in which the field addition is the bit-wise xor operation, we have the i-th column of A_c defined as the n-dimension vector representation of field element $c \cdot \nu_i$, in which ν_i is the field element that has the corresponding representation to be a basis vector with the i-th position being one and the rest positions being zero.

$\langle \gamma, c_0 v_0 \rangle$ term as $\langle \gamma, c_0 v_0 \rangle = \gamma^\top A_{c_0} v_0 = (A_{c_0}^\top \gamma)^\top v_0 = \langle A_{c_0}^\top \gamma, v_0 \rangle$ where $A_{c_0}^\top$ is the transpose of A_{c_0}. So we get

$$
\begin{aligned}
&\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \\
&= \frac{1}{N} \cdot \sum_{\beta_0} \sum_{\alpha_1, \beta_1} \cdots \sum_{\alpha_{t-1}, \beta_{t-1}} \sum_{\alpha_t} \sum_{\gamma} \widehat{\mathbb{1}_{V_0}}(\beta_0) \cdots \widehat{\mathbb{1}_{U_t}}(\alpha_t) \cdot \left(\sum_{v_0} (-1)^{\langle \beta_0 + A_{c_0}^\top \gamma, v_0 \rangle} \right) \\
&\quad \cdot \left(\sum_{u_1} (-1)^{\langle \alpha_1 + A_{c_0}^\top \gamma, u_1 \rangle} \right) \left(\sum_{v_1} (-1)^{\langle \beta_1 + A_{c_1}^\top \gamma, v_1 \rangle} \right) \cdots \\
&\quad \left(\sum_{v_{t-1}} (-1)^{\langle \beta_{t-1} + A_{c_{t-1}}^\top \gamma, v_{t-1} \rangle} \right) \left(\sum_{u_t} (-1)^{\langle \alpha_t + A_{c_{t-1}}^\top \gamma, u_t \rangle} \right) .
\end{aligned}
$$

It is known that $\sum_{x \in \{0,1\}^n} (-1)^{\langle \alpha, x \rangle} = N$ if and only if $\alpha = 0$, otherwise it equals zero. So we are only interested in the case in which the fourier coefficients gives non-zero summation. And we observe that the set of interesting coefficients can be expressed in terms of γ, i.e., for all $i \in \{0, \ldots, t-1\}$: $\alpha_{i+1} = \beta_i = A_{c_i}^\top \gamma$. Hence the equality calculation can be greatly simplified as

$$
\begin{aligned}
&\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \\
&= N^{2t-1} \sum_{\gamma} \widehat{\mathbb{1}_{V_0}}(A_{c_0}^\top \gamma) \widehat{\mathbb{1}_{Q_1}}(A_{c_0}^\top \gamma, A_{c_1}^\top \gamma) \cdots \widehat{\mathbb{1}_{Q_{t-1}}}(A_{c_{t-2}}^\top \gamma, A_{c_{t-1}}^\top \gamma) \widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^\top \gamma) \\
&= \frac{q^{t+1}}{N} + N^{2t-1} \sum_{\gamma \neq 0} \widehat{\mathbb{1}_{V_0}}(A_{c_0}^\top \gamma) \widehat{\mathbb{1}_{Q_1}}(A_{c_0}^\top \gamma, A_{c_1}^\top \gamma) \cdots \widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^\top \gamma) \\
&\leqslant \frac{q^{t+1}}{N} + N^{2t-1} \sum_{\gamma \neq 0} |\widehat{\mathbb{1}_{V_0}}(A_{c_0}^\top \gamma)| \cdot |\widehat{\mathbb{1}_{Q_1}}(A_{c_0}^\top \gamma, A_{c_1}^\top \gamma)| \cdots |\widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^\top \gamma)| .
\end{aligned}
$$

Next, we let

$$
\begin{aligned}
\mathsf{left} &:= \text{min of } i \text{ such that } c_i \neq 0 \\
\mathsf{right} &:= \text{max of } i \text{ such that } c_i \neq 0
\end{aligned}
$$

To proceed with the calculation, case discussion over $(\mathsf{left}, \mathsf{right})$ is needed, here we consider the case of $\mathsf{left} = 0$ and $\mathsf{right} = t-1$ (i.e., $c_0 \neq 0$ and $c_{t-1} \neq 0$). The other cases give the same upper bound and we left them to the full version. Therefore, we obtain

$$
\begin{aligned}
&\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) - \frac{q^{t+1}}{N} \\
&\leqslant N^{2t-1} \sum_{\gamma \neq 0} |\widehat{\mathbb{1}_{V_0}}(A_{c_0}^\top \gamma)| \cdot |\widehat{\mathbb{1}_{Q_1}}(A_{c_0}^\top \gamma, A_{c_1}^\top \gamma)| \cdots |\widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^\top \gamma)| \\
&\leqslant N^{2t-3} \sum_{\gamma \neq 0} |\widehat{\mathbb{1}_{V_0}}(A_{c_0}^\top \gamma)| \cdot \left(\frac{q}{N^2} \right)^{t-2} \cdot \Phi(\mathcal{Q}_{\mathsf{idx}+1}) \cdot |\widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^\top \gamma)|
\end{aligned}
$$

$$= q^{t-2}N\Phi(\mathcal{Q}_{\mathsf{idx}+1}) \cdot \sum_{\gamma \neq 0} |\widehat{\mathbb{1}_{V_0}}(A_{c_0}^{\top}\gamma)| \cdot |\widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^{\top}\gamma)| \leqslant q^{t-1}\Phi(\mathcal{Q}_{\mathsf{idx}+1}) \,. \tag{8}$$

Note that we have $N^2|\widehat{\mathbb{1}_{\mathcal{Q}_{\mathsf{idx}+1}}}(A_{c_{\mathsf{idx}}}\gamma, A_{c_{\mathsf{idx}+1}}\gamma)| \leqslant \Phi(\mathcal{Q}_{\mathsf{idx}+1})$ for any $\gamma \neq 0$ given the condition that $c_{\mathsf{idx}} \neq 0$ and $c_{\mathsf{idx}+1} \neq 0$. We also used the fact of (7) that, for any α, β, $|\widehat{\mathbb{1}_{\mathcal{Q}_i}}(\alpha, \beta)| \leqslant q/N^2$. The last step of inequality holds because by (6) we have $\sum_\gamma \widehat{\mathbb{1}_{V_0}}(A_{c_0}^{\top}\gamma)^2 = \sum_\gamma \widehat{\mathbb{1}_{U_t}}(A_{c_{t-1}}^{\top}\gamma)^2 = q/N$, so we can apply Cauchy-Schwartz inequality to obtain the result. This exact inequality step ensures the tight bound and was dubbed the *Cauchy-Schwartz trick* used in [2,23,5].

So we proved Lemma 3. □

Now the remaining step is to upper bound $\Phi(\mathcal{Q}_{\mathsf{idx}+1})$. Here we apply the following lemma, which has essentially the same proof of Lemma 6 proved by Chen *et al.* in [5], with the only adjustment of changing their parameter δ into $\delta = \sqrt{(12 \ln N)/q}$.

Lemma 4. *Assuming that $9(t+2)n \leqslant q \leqslant N/2$. Fix an adversary making q queries to a random permutation P. Let Q denote the transcript of interaction of $\mathcal{A}$ with P. Then for any $\alpha, \beta \in \mathbb{F}_{2^n}$,*

$$\mathsf{Pr}_{P,\omega}\left[\Phi(Q) \geqslant \frac{2q^2}{N} + 3\sqrt{(t+2)nq}\right] \leqslant \frac{2}{N^t} \,,$$

in which the probability is taken over the random permutation P and the random coins ω used by $\mathcal{A}$.

Plugging in the inequality we get

$$\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \leqslant \frac{q^{t+1}}{N} + q^{t-1}\Phi(Q_{\mathsf{idx}+1}) \leqslant \frac{3q^{t+1}}{N} + 3q^{t-1/2}\sqrt{(t+2)n}$$

with probability at least $1 - \frac{2t}{N^t}$. Hence we proved Lemma 2.

Tightness of Lemma 2. We examine the tightness of 1-constraints sum capture quantity in two aspects. One is, given the $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ in which there exists two neighboring c_i, c_{i+1} so that $c_i \neq 0, c_{i+1} \neq 0$, whether the upper bound is tight or not.

We first give the following proposition showing that, if there exists neighboring coefficients $c_i \neq 0$ and $c_{i+1} \neq 0$, then for moderately large q (e.g. $q > N^{2/3}$), $\mu_{\boldsymbol{c}} \geqslant q^{t+1}/2N$ with high probability. We left the detailed proof to the full version.

Proposition 1. *Let q be any positive integer of power of two. Fix any $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ such that there exists an index $0 \leqslant i < t-1$ satisfying $c_i \neq 0$ and $c_{i+1} \neq 0$, then there is an explicit algorithm $\mathcal{A}$ that makes at most q queries to each of $P_1, \ldots, P_{t-1}$, and $V_0, U_t \subseteq \mathbb{F}_{2^n}$ that have $|V_0| = |U_t| = q$, so that*

$$\mathsf{Pr}\left[\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \geqslant \frac{q^{t+1}}{2N}\right] \geqslant 1 - \frac{N}{q} \cdot e^{-q^2/8N} \,.$$

The following proposition, which is complementary to Proposition 1, states that, if $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ satisfies that for any $0 \leqslant i < t-1$, either $c_i = 0$ or $c_{i+1} = 0$, then $\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t)$ can achieve up to q^t, which is larger than q^{t+1}/N. We left the proof to the full version.

Proposition 2. *Let q be any positive integer of power of two. Fix any $\boldsymbol{c} = (c_0, \ldots, c_{t-1})$ such that for any $0 \leqslant i < t-1$, either $c_i = 0$ or $c_{i+1} = 0$, there is an explicit algorithm $\mathcal{A}$ that makes at most q queries to each of $P_1, \ldots, P_{t-1}$, and $V_0, U_t \subseteq \mathbb{F}_{2^n}$ that have $|V_0| = |U_t| = q$, so that*

$$\mu_{\boldsymbol{c}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \geqslant q^t \ .$$

4.2 2-constraints Sum Capture Quantity

Now we move to consider the sum capture quantity in which the number of constraints $r = 2$. We let the 2-constraint sum capture quantity be associated with two vector of coefficients $\boldsymbol{c} = (c_0, c_1, \ldots, c_{t-1})$ and $\boldsymbol{d} = (d_0, d_1, \ldots, d_{t-1})$, as

$$\begin{aligned}\mu_{\boldsymbol{c},\boldsymbol{d}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) := \\ |\{(v_0, (u_1, v_1), \ldots, (u_{t-1}, v_{t-1}), u_t) \in V_0 \times \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_{t-1} \times U_t \ : \\ \sum_{j=0}^{t-1} c_j(v_j + u_{j+1}) = 0, \ \sum_{j=0}^{t-1} d_j(v_j + u_{j+1}) = 0\}| \ . \quad (9)\end{aligned}$$

Though the 2-constraint sum capture quantity is a natural generalization of the 1-constraint case, we note that adding only one more constraint makes proving the tightest upper bound of (9) much harder. Here we only focus on giving bounds over the sum capture quantity with a specific class of coefficients $\boldsymbol{c}, \boldsymbol{d}$ that can be derived from the $(t-2)$-wise independently uniform subkeys. We obtain a bound that gives the tightest KAC security for $t \geqslant 8$. However, for $t < 5$, our 2-constraint upper bound is even worse than a reduction-based bound. While it is interesting to investigate whether our bound can be improved, for $t = 3$, in particular, we show that the above sum capture quantity is lower-bounded by $\Omega(q^3/N)$ and hence cannot be used to prove $q = \Omega(N^{3/4})$ for the 3-round KAC with identical subkeys.

We prove upper bounds for the class of linear constraint coefficients $\boldsymbol{c} = (c_0, \ldots, c_{t-1}), \boldsymbol{d} = (d_0, \ldots, d_{t-1})$ with the property that $c_0 = d_{t-1} = 1$, $c_{t-1} = d_0 = 0$, and for all $i \in \{1, \ldots, t-2\}$, $c_i \neq 0, d_i \neq 0$, and for all $i, j \in \{1, \ldots, t-2\}$ such that $i \neq j$, $c_i d_i^{-1} \neq c_j d_j^{-1}$. We justify that $\boldsymbol{c}, \boldsymbol{d}$ corresponds to the linear key schedule from $t-2$ independent keys that gives $(t-2)$-wise independently uniform subkeys.

JUSTIFICATION. We use $s_0, \ldots, s_{t-1}$ to denote the subkeys. Given the subkeys are generated linearly from $t-2$ independent keys and are $(t-2)$-wise independently uniform, the middle $t-2$ subkeys $s_1, \ldots, s_{t-2}$ uniquely fix the original master keys and hence the first subkey s_0 and the last subkeys s_{t-1} can be uniquely

determined as a linear combination of $s_1, \ldots, s_{t-2}$. i.e.,

$$s_0 = \sum_{i=1}^{t-2} c_i s_i \ , \quad s_{t-1} = \sum_{i=1}^{t-2} d_i s_i \ .$$

Note that all c_i, d_i should be non-zero because otherwise we can obtain a linear combination $t-2$ subkeys that sum to zero, breaking the $(t-2)$-wise independence. Further, we show by contradiction, if there exists i, j such that $i \neq j$ and $c_i d_i^{-1} = c_j d_j^{-1}$, then we pick the set of subkeys $\{s_0, s_{t-1}\} \cup \{s_k \mid 1 \leqslant k \leqslant t-2 \wedge k \notin \{i, j\}\}$ and we have

$$s_0 + c_i d_i^{-1} s_{t-1} = \sum_{k \notin \{0,i,j,t\}} (c_i d_i^{-1} d_k + c_k) s_k$$

which is a linear dependence among $t-2$ subkeys. Thus all $c_i d_i^{-1}$ must be distinct.

Then we have the following lemma for the 2-constraints sum capture quantity.

Lemma 5. *Let $t \geqslant 3$. Let $P_1, \ldots, P_{t-1}$ be $t-1$ independent uniformly random permutations of $\{0,1\}^n$, and let $\mathcal{A}$ be a probabilistic algorithm that makes adaptive queries to $P_1, \ldots, P_{t-1}$. Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}$ be the query transcripts of $P_1, \ldots, P_{t-1}$ interacting with $\mathcal{A}$. Let coefficients $\boldsymbol{c}, \boldsymbol{d}$ be defined as above, then for any $\mathcal{A}$ that makes at most $q \geqslant (t+2)nN^{2/3}$ queries to each permutations,*

$$\mathsf{Pr}_{P_1,\ldots,P_{t-1}}\Bigg[\exists V_0, U_t \subseteq \mathbb{F}_{2^n}, |V_0| = |U_t| = q, \\ \mu_{\boldsymbol{c},\boldsymbol{d}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \geqslant \frac{q^{t+1}}{N^2} + t \cdot \frac{(3q)^{2t-3}}{N^{t-2}} + \frac{(3q)^{2t-2.5}}{N^{t-2}}\Bigg] \leqslant \frac{2t}{N^t} \ .$$

Discussion. Note that when $t \geqslant 5$, the security bound starts getting better than the $t-1$ round KAC bound $q = \Omega(N^{\frac{t-1}{t}})$. For $t \geqslant 8$, the security bound achieves optimal security of $q = \Omega(N^{\frac{t}{t+1}})$.

As in the case of 1-constraint, we will prove an upper bound of $\mu_{\boldsymbol{c},\boldsymbol{d}}$ conditioning on $\Phi(\mathcal{Q}_i)$ being small for all i.

Lemma 6. *Fix $\boldsymbol{c}, \boldsymbol{d}$ defined as in Lemma 5, then conditioning on $\Phi(\mathcal{Q}_i) \leqslant 9q^2/N$ for all $1 \leqslant i \leqslant t-1$, it holds that for any subsets $V_0, U_t \subseteq \mathbb{F}_{2^n}$ with $|V_0| = |U_t| = q$,*

$$\mu_{\boldsymbol{c},\boldsymbol{d}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) \leqslant \frac{q^{t+1}}{N^2} + t \cdot \frac{(3q)^{2t-3}}{N^{t-2}} + \frac{(3q)^{2t-2.5}}{N^{t-2}} \ .$$

Proof. The initial calculation steps are similar to the 1-constraint case. We directly give the calculation result and left the details in the full version.

$$\mu_{\boldsymbol{c},\boldsymbol{d}}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t) = \\ N^{2t-2} \sum_{\alpha,\beta} \widehat{\mathbb{1}_{V_0}}(\theta_0) \widehat{\mathbb{1}_{\mathcal{Q}_1}}(\theta_0, \theta_1) \widehat{\mathbb{1}_{\mathcal{Q}_2}}(\theta_1, \theta_2) \cdots \widehat{\mathbb{1}_{\mathcal{Q}_{t-1}}}(\theta_{t-2}, \theta_{t-1}) \widehat{\mathbb{1}_{U_t}}(\theta_{t-1})$$

in which

$$\theta_0 = \alpha, \quad \theta_{t-1} = \beta,$$
$$\forall i \in \{1, \ldots, t-2\} : \theta_i = A_{c_i}^\top \alpha + A_{d_i}^\top \beta \, .$$

We write $\mathsf{Coeff} = \{\theta_0, \theta_1, \ldots, \theta_{t-1}\}$. Here we partition the summation into three cases and discuss the set of (α, β) assignments that falls into each cases.

1. At least two θs in Coeff are zero.
2. Exactly one θ in Coeff is zero.
3. None of the θs in Coeff is zero.

The following claim shows that, if case one happens, then all coefficients θ are zero.

Claim 1 *If two θs in Coeff are zero, then $\alpha = \beta = 0$.*

Proof. If $\theta_0 = \alpha = \beta = \theta_{t-1} = 0$, then the claim is trivial. If $\alpha = \theta_0 = \theta_i = 0$ for some i with $1 \leqslant i \leqslant t-2$, then given $\theta_i = A_{c_i}^\top \alpha + A_{d_i}^\top \beta = A_{d_i}^\top \beta$ and $A_{d_i}^\top$ is full-rank (because $d_i \neq 0$), we can infer that $\beta = 0$. Similarly we can infer $\alpha = 0$ if $\beta = \theta_{t-1} = \theta_i = 0$ for some i with $1 \leqslant i \leqslant t-2$. Now, if $\theta_i = \theta_j = 0$ for some i, j such that $1 \leqslant i, j \leqslant t-2$ and $i \neq j$. Then the choice of (α, β) must satisfy

$$\begin{cases} A_{c_i}^\top \alpha + A_{d_i}^\top \beta = 0 \\ A_{c_j}^\top \alpha + A_{d_j}^\top \beta = 0 \end{cases}$$

implying $A_{d_{i+1}^{-1} c_{i+1}}^\top \alpha = (A_{d_{i+1}}^\top)^{-1} A_{c_{i+1}}^\top \alpha = \beta = (A_{d_{j+1}}^\top)^{-1} A_{c_{j+1}}^\top \alpha = A_{d_{j+1}^{-1} c_{j+1}}^\top \alpha$. Hence

$$\left(A_{d_{i+1}^{-1} c_{i+1}}^\top + A_{d_{j+1}^{-1} c_{j+1}}^\top\right) \alpha = \left(A_{d_{i+1}^{-1} c_{i+1} + d_{j+1}^{-1} c_{j+1}}^\top\right) \alpha = 0 \, .$$

Here α can be non-zero only if $d_{i+1}^{-1} c_{i+1} = d_{j+1}^{-1} c_{j+1}$. However, this is impossible as we have justified from the $(t-2)$-wise independently uniform property of subkeys. □

Let μ_1, μ_2, μ_3 corresponds to summation for (α, β) that corresponds to case one, two, three, respectively.

Proposition 3.

$$\mu_1 = \frac{q^{t+1}}{N^2}$$

Proof. Since case one only happens when $\alpha = \beta = 0$, we have $\theta_i = 0$ for all i. Therefore, a direct calculation using the fact that $\widehat{\mathbb{1}_{V_0}}(0) = \widehat{\mathbb{1}_{U_t}}(0) = q/N$ and $\widehat{\mathbb{1}_{\mathcal{Q}_i}}(0,0) = q/N^2$ proves the bound. □

Proposition 4.

$$\mu_2 \leqslant \frac{t \cdot (3q)^{2t-3}}{N^{t-2}}$$

We note that the proof of Proposition 4 can be derived via a moderate tweak from the proof of 1-constraint sum capture quantity upper bound (i.e., Lemma 2), we left the complete proof to the full version.

Proposition 5.

$$\mu_3 \leqslant \frac{(3q)^{2t-2.5}}{N^{t-2}}$$

Proof (of Proposition 5). We define a $N \times N$ matrix M with each entry labeled by $(\alpha, \beta) \in \mathbb{F}_{2^n} \times \mathbb{F}_{2^n}$ so that

$$M_{\alpha,\beta} = \begin{cases} 0 & \text{if some } \theta \in \mathsf{Coeff} \text{ is } 0 \\ \widehat{\mathbb{1}_{\mathcal{Q}_1}}(\alpha, A_{c_1}^\top \alpha + A_{d_1}^\top \beta) \cdots \widehat{\mathbb{1}_{\mathcal{Q}_{t-1}}}(A_{c_{t-2}}^\top \alpha + A_{d_{t-2}}^\top \beta, \beta) & o.w. \end{cases}$$

Note that M is a $2^n \times 2^n$ matrix. We also define the column vectors $\boldsymbol{v}$, $\boldsymbol{u}$ with each entry labeled by $\alpha \in \mathbb{F}_{2^n}$ so that $\boldsymbol{v}_\alpha = \widehat{\mathbb{1}_{V_0}(\alpha)}$ and $\boldsymbol{u}_\alpha = \widehat{\mathbb{1}_{U_t}(\alpha)}$. Therefore, we can write μ_3 as

$$\mu_3 = N^{2t-2} \sum_{\alpha,\beta \mid M_{\alpha,\beta} \neq 0} \widehat{\mathbb{1}_{V_0}(\alpha)} \cdot M_{\alpha,\beta} \cdot \widehat{\mathbb{1}_{U_t}(\beta)} = N^{2t-2} \boldsymbol{v}^\top M \boldsymbol{u} \ .$$

Noting that the equivalent definition for the matrix 2-norm as

$$\|M\|_2 := \sup_{\|x\|_2 = 1} \|Mx\|_2 = \sup_{\|x\|_2 = 1, \|y\|_2 = 1} y^\top M x \ ,$$

we can use the matrix norm as the upper bound of μ_3, that is

$$\mu_3 = N^{2t-2} \cdot \boldsymbol{v}^\top M \boldsymbol{u} \leqslant N^{2t-2} \|\boldsymbol{v}\|_2 \|M\|_2 \|\boldsymbol{u}\|_2 \ .$$

By (6), we can infer that $\|\boldsymbol{v}\|_2 = \sqrt{\sum_\alpha v_\alpha^2} = \sqrt{\sum_\alpha \widehat{\mathbb{1}_{V_0}(\alpha)^2}} = \sqrt{q/N}$ and $\|\boldsymbol{u}\|_2 = \sqrt{q/N}$. We also use the fact that $\|M\|_2 \leqslant \|M\|_F$ where $\|M\|_F = \sqrt{\sum_{i,j} M_{i,j}^2}$ is the Frobenius norm, then we have

$$\mu_3 \leqslant N^{2t-2} \cdot \sqrt{\frac{q}{N}} \|M\|_2 \sqrt{\frac{q}{N}} \leqslant q N^{2t-3} \|M\|_F = q N^{2t-3} \sqrt{\sum_{\alpha,\beta} M_{\alpha,\beta}^2}$$

where

$$\sum_{\alpha,\beta} M_{\alpha,\beta}^2 = \sum_{\alpha,\beta \mid M_{\alpha,\beta} \neq 0} \widehat{\mathbb{1}_{\mathcal{Q}_1}}(\alpha, A_{c_1}^\top \alpha + A_{d_1}^\top \beta)^2 \cdots \widehat{\mathbb{1}_{\mathcal{Q}_{t-1}}}(A_{c_{t-2}}^\top \alpha + A_{d_{t-2}}^\top \beta, \beta)^2$$

$$\leqslant \sum_{\alpha,\beta \mid M_{\alpha,\beta}\neq 0} \widehat{\mathbb{1}_{\mathcal{Q}_1}}(\alpha, A_{c_1}^\top \alpha + A_{d_1}^\top \beta)^2 \cdot \frac{(3q)^{4(t-2)}}{N^{6(t-2)}}$$
$$\leqslant \frac{(3q)^{4(t-2)}}{N^{6(t-2)}} \sum_{\alpha,\beta} \widehat{\mathbb{1}_{\mathcal{Q}_1}}(\alpha, A_{c_1}^\top \alpha + A_{d_1}^\top \beta)^2 = \frac{(3q)^{4(t-2)}}{N^{6(t-2)}} \cdot \frac{q}{N^2} \leqslant \frac{(3q)^{4t-7}}{N^{6t-10}} .$$

So we get

$$\mu_3 \leqslant qN^{2t-3} \cdot \frac{(3q)^{2t-3.5}}{N^{3t-5}} \leqslant \frac{(3q)^{2t-2.5}}{N^{t-2}} .$$

□

Putting the propositions all together, we have

$$\mu_{\boldsymbol{c},\boldsymbol{d}} = \mu_1 + \mu_2 + \mu_3 \leqslant \frac{q^{t+1}}{N^2} + t \cdot \frac{(3q)^{2t-3}}{N^{t-2}} + \frac{(3q)^{2t-2.5}}{N^{t-2}} .$$

□

4.3 Tightness of 2-constraint Sum Capture Quantity for 3-round KAC

A natural question is whether the upper bound of the 2-constraint sum capture quantity can be improved so that it gives tight security bound for t-round KAC when $t < 7$. In particular, the most interesting case is to prove tight security bound $q = \Omega(N^{3/4})$ for 3-round KAC with identical subkeys, which corresponds to the instantiation in Corollary 4 when $t = 3$. However, for the 3-round KAC with identical key schedule, we show that it is impossible to show the conjectured optimal security bound via upper-bounding the sum capture quantity, as the sum capture quantity for 3-round identical-subkey KAC is lower-bounded by $\Omega(q^3/N)$ with high probability, giving $\mu_{\boldsymbol{c}}/N = \Omega(q^3/N^2)$ instead of the desired q^4/N^3. The sum capture quantity lower bound for 3-round identical-subkey KAC directly follows from the following proposition with $c_1 = d_1 = 1$. We left the proof of proposition to the full version.

Proposition 6. *Let q be any positive integer of power of two. Let $t = 2$ and fix $\boldsymbol{c} = (1, c_1, 0)$, $\boldsymbol{d} = (0, d_1, 1)$ where c_1, d_1 are non-zero, then there exists an explicit algorithm $\mathcal{A}$ that makes at most q queries to each of P_1, P_2 and $V_0, U_3 \subseteq \mathbb{F}_{2^n}$ that have $|V_0| = |U_3| = q$, so that*

$$\mathsf{Pr}[\mu_{\boldsymbol{c},\boldsymbol{d}}(V_0, \mathcal{Q}_1, \mathcal{Q}_2, U_3) \geqslant q^3/2N] \geqslant 1 - \frac{N}{q} \cdot e^{-q^2/8N} .$$

Though Proposition 6 gives a lower bound of $\Omega(q^3/N)$ for the sum capture quantity $\mu_{\boldsymbol{c},\boldsymbol{d}}$, it does not immediately imply a distinguishing attack against the 3-round KAC. This is because the number of bad keys generated by our constructed $\mathcal{A}$ is at most q, so we have $\mathsf{Pr}[\boldsymbol{k} \in \mathsf{Badkey}] \leqslant q/N$. The reason of $\mu_{\boldsymbol{c},\boldsymbol{d}}$ being too large is that a bad key may be counted multiple times in the sum capture quantity. Therefore, we cannot proceed with the sum capture quantity to prove the optimal $q = \Omega(N^{3/4})$ bound for 3-round KAC with identical subkeys if the overcounting cannot be eliminated.

5 Good Transcript Analysis

Our next goal is to obtain upper bounds of $1 - \mathsf{ps}_0(\tau)/\mathsf{ps}_1(\tau)$ for each $\tau \in \mathcal{T}_{\mathsf{good}}$. In particular, we will show the following lemma.

Lemma 7. *If the t-round KAC is instantiated with a key schedule that gives $(t-2)$-wise independently uniform subkeys, then there exists a function $g : \mathcal{T} \to [0, +\infty)$ so that for any $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k}) \in \mathcal{T}_{\mathsf{good}}$,*

$$1 - \frac{\mathsf{ps}_0(\tau)}{\mathsf{ps}_1(\tau)} \leqslant g(\tau) \ ,$$

and for any query records $\boldsymbol{\mathcal{Q}}$,

$$\mathbb{E}_{\boldsymbol{k}}\left[g(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})\right] \leqslant \frac{t^2(4q)^{t+1}}{N^t} \ .$$

To obtain the desired function $g(\cdot)$, we need to understand the ratio $\mathsf{ps}_0(\tau)/\mathsf{ps}_1(\tau)$ first. Given the transcript $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k})$ in which $\boldsymbol{\mathcal{Q}} = (\mathcal{Q}_E, \mathcal{Q}_1, \ldots, \mathcal{Q}_t)$, we write $E \downarrow \mathcal{Q}_E$ to denote that the real-world cipher construction E is consistent with the recorded query $\mathcal{Q}_E$, that is, for each $(x, y) \in \mathcal{Q}_E$, it holds that $E(x) = y$. Similarly, we write $P_i \downarrow \mathcal{Q}_i$ to denote that the permutation P_i is consistent with the recorded query $\mathcal{Q}_i$. Then following [5,16] one can derive that

$$\frac{\mathsf{ps}_0(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})}{\mathsf{ps}_1(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})} = N^{(|\mathcal{Q}_E|)} \cdot \mathsf{Pr}[E_{\boldsymbol{k}} \downarrow \mathcal{Q}_E \mid P_1 \downarrow \mathcal{Q}_1, \ldots, P_t \downarrow \mathcal{Q}_t] \ , \tag{10}$$

where $N^{(|\mathcal{Q}_E|)} = N(N-1)\cdots(N-|\mathcal{Q}_E|+1)$. We provide a proof of (10) in the full version.

To analyze the probability term on the RHS, we need to take the following graph view for KAC, which was originally introduced by Chen and Steinberger in [6].

5.1 Graph Definition and an Useful Lemma

Let G be a graph that consists of vertices which can be divided into $m+1$ layers $L_0, \ldots, L_m$ such that each layer contains exactly N vertices, and edges that can be partition into m sets $\boldsymbol{E} = (E_{(0,1)}, E_{(1,2)}, \ldots, E_{(m-1,m)})$ such that $E_{(i,i+1)}$ forms a partial (but possibly perfect) matching from L_i to L_{i+1}.

We say a vertex $u \in L_i$, where $i < m$, is right-free if no edge connects u to any vertex in L_{i+1}. Analogously, we say a vertex $v \in L_j$, where $j > 0$, is left-free if no edge connects v to any vertex in L_{j-1}.

For any vertex $u \in L_0$ we define the following probabilistic procedure that generates a path $(w_0, w_1, \ldots w_m)$ from u to a vertex in L_m.

- Let $w_0 = u$.
- For i from 1 to m, if w_{i-1} is not right-free and connects to some vertex $w' \in L_i$, then let $w_i = w'$, otherwise let w_i be uniformly sampled from all left-free vertices in L_i.

We write $\mathsf{Pr}[u \to v]$ to denote the probability that the path $(u, w_1, \ldots, w_m)$ satisfies $w_m = v$. In particular, we are interested in the pair of (u, v) such that u is right-free and v is left-free.

For the layered graph G, we let $\mathcal{U}_G(a, b)$, where $a \leqslant b$, be the set of paths that starts at a left-free vertex in L_a and reaches a vertex in L_b. We note that the path in $\mathcal{U}_G(a, b)$ does not necessarily ends in L_b. We write $U_G(a, b) = |\mathcal{U}_G(a, b)|$. Note that $U_G(a, a)$ denotes the total number of left-free vertices in L_a.

Given any $\sigma = ((i_0, i_1), (i_1, i_2), \ldots, (i_{|\sigma|-1}, i_{|\sigma|}))$ in which $i_0 < i_1 < \cdots < i_{|\sigma|}$, we say σ is an interesting (a, b)-segment partition with regard to the index set $\mathcal{I} \subseteq \{0, \ldots, m\}$ if $i_0 = a, i_{|\sigma|} = b$ and for all $1 < j < |\sigma|$ we have $i_j \in \mathcal{I}$. We use $\mathcal{B}_{\mathcal{I}}(a, b)$ to denote the set that contains all interesting (a, b)-segment partition of the set $\mathcal{I}$. Given a layered graph G, we let the interesting indices of G as

$$\mathcal{I}(G) := \{i \in \{0, 1, 2, \ldots, m\} \mid U_G(i, i) > 0\} .$$

Then we are ready to state the following lemma, which is a slightly different variant of the lemma proved by Chen and Steinberger in [6] but with essentially the same proof. We include the proof in the full version.

Lemma 8. *For any graph G defined as above, and any $u \in L_0, v \in L_m$ such that u is right-free and v is left-free, it holds that*

$$\mathsf{Pr}[u \to v] = \frac{1}{N} - \frac{1}{N} \sum_{\sigma \in \mathcal{B}_{\mathcal{I}(G)}(0,m)} (-1)^{|\sigma|} \prod_{h=1}^{|\sigma|} \frac{U_G(i_{h-1}, i_h)}{U_G(i_h, i_h)} .$$

5.2 Graph View of KAC

The KAC can also be interpreted in the graph view. Given a transcript $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k})$ where $\boldsymbol{\mathcal{Q}} = (\mathcal{Q}_E, \mathcal{Q}_1, \ldots, \mathcal{Q}_t)$ and let subkeys $\boldsymbol{s} = (s_0, \ldots, s_t)$ be generated from the key $\boldsymbol{k}$, we define $E_{(2i,2i+1)} := \{(v, v + s_i) \mid v \in L_{2i}\}$ for $i \in \{0, \ldots, t\}$. That is, L_{2i} and L_{2i+1} are connected by the "subkey edges", which corresponds to the step of xoring the subkey s_i in the KAC execution. For $i \in \{1, \ldots, t\}$, we let $E_{(2i-1,2i)} := \{(u, v) \mid (u, v) \in \mathcal{Q}_i\}$. This corresponds to the queries made to the permutation P_i. Now, note that the interesting indices for KAC can only be a subset of $\{0, 2, 4, \ldots, 2t\}$.

For a fixed query records $\boldsymbol{\mathcal{Q}}$, let $Z_{\boldsymbol{s}}(a, b)$, where $a \leqslant b$, be the total number of paths that connects a vertex in L_a and a vertex in L_b when the subkeys are fixed to $\boldsymbol{s}$. Note that the paths do not necessarily start at L_a or end at L_b. For the ℓ-th cipher query (x_ℓ, y_ℓ), let $\alpha_\ell[\boldsymbol{s}]$ denote the largest possible index of the layer that is reachable from x_ℓ when the subkeys are fixed to be $\boldsymbol{s}$. let $\beta_\ell[\boldsymbol{s}]$ denote the smallest index of the layer than is reachable from y_ℓ. Note that in the good key case, we always have $\alpha_\ell[\boldsymbol{s}] < \beta_\ell[\boldsymbol{s}]$.

Now, to bound the probability $\mathsf{Pr}[E \downarrow \mathcal{Q}_E \mid P_1 \downarrow \mathcal{Q}_1, \ldots, P_t \downarrow \mathcal{Q}_t]$, we analyze the following experiment that can be divided into $|\mathcal{Q}_E|$ stages.

1. Initially, G_0 is defined according to the given transcript $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k})$.

2. For ℓ from 1 to $|\mathcal{Q}_E|$, given $G_{\ell-1}$ is defined, the probabilistic path generating process is run for the ℓ-th query $(x_\ell, y_\ell) \in \mathcal{Q}_E$ over the graph $G_{\ell-1}$, from vertex $x_\ell \in L_0$.
   – If the generated path from x_ℓ does not arrive at y_ℓ, the experiment outputs 0 and aborts.
   – otherwise we first set $G_\ell = G_{\ell-1}$, then we remove all vertices on the path of (x_ℓ, y_ℓ) from G_ℓ. The new graph G_ℓ will have $N - \ell$ vertices in each layer.
3. If $G_{|\mathcal{Q}_E|}$ is successfully defined, the experiment outputs 1.

So we have

$$\frac{\mathsf{ps}_0(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})}{\mathsf{ps}_1(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})} = N^{(|\mathcal{Q}_E|)}\mathsf{Pr}[\mathsf{Exp}(\tau) = 1] = N^{(|\mathcal{Q}_E|)} \prod_{\ell=1}^{|\mathcal{Q}_E|} \mathsf{Pr}[x_\ell \to y_\ell \mid G_{\ell-1}]$$

Now we are ready to state the core lemma that defines the function $g(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})$ and prove it using Lemma 8.

Lemma 9. *For any query records $\boldsymbol{\mathcal{Q}}$ with $q \leqslant N/4$ and subkeys $\boldsymbol{k}$ such that the transcript $\tau = (\boldsymbol{\mathcal{Q}}, \boldsymbol{k}) \in \mathcal{T}_{\mathsf{good}}$,*

$$\frac{\mathsf{ps}_0(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})}{\mathsf{ps}_1(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})} \geqslant 1 - \sum_{\ell=1}^{q} \sum_{1 \leqslant a \leqslant b \leqslant t} \mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}] \sum_{\sigma \in \mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma| \geqslant 2} \prod_{h=1}^{|\sigma|} \frac{Z_{\boldsymbol{s}}(i_{h-1}, i_h)}{N - 2q}$$

in which the set of interesting indices $\mathcal{I}$ of the segment partition set $\mathcal{B}_{\mathcal{I}}$ is defined as $\mathcal{I} = \{0, 2, \ldots, 2t\}$, and $\mathsf{R}_{a,b,\ell}[\boldsymbol{s}] := \mathbb{1}(\alpha_\ell[\boldsymbol{s}] \geqslant a, \beta_\ell[\boldsymbol{s}] \leqslant b)$.

Proof. For the ℓ-th cipher query (x_ℓ, y_ℓ) given the graph support $G_{\ell-1}$, we can define a graph G from $G_{\ell-1}$ that removes all layers L_i for $i < \alpha_\ell[\boldsymbol{s}]$ and L_j for $j > \beta_\ell[\boldsymbol{s}]$. Thus, in the graph G we starts at a right-free vertex $u \in L_0$ and targets a left-free vertex $v \in L_m$, allowing us to apply Lemma 8.

$$\begin{aligned}
&\mathsf{Pr}_G[(x_\ell \to y_\ell) \mid G_{\ell-1}] \\
&= \frac{1}{N - \ell + 1} \left(1 - \sum_{\sigma \in \mathcal{B}_{\mathcal{I}(G)}(0,m)} (-1)^{|\sigma|} \prod_{h=1}^{|\sigma|} \frac{U_G(i_{h-1}, i_h)}{U_G(i_h, i_h)}\right) \\
&= \frac{1}{N - \ell + 1} \left(1 + \frac{U_G(0,m)}{U_G(m,m)} - \sum_{\sigma \in \mathcal{B}_{\mathcal{I}(G)}(0,m),\ |\sigma| \geqslant 2} (-1)^{|\sigma|} \prod_{h=1}^{|\sigma|} \frac{U_G(i_{h-1}, i_h)}{U_G(i_h, i_h)}\right) \\
&\geqslant \frac{1}{N - \ell + 1} \left(1 - \sum_{\sigma \in \mathcal{B}_{\mathcal{I}(G)}(0,m),\ |\sigma| \geqslant 2} \prod_{h=1}^{|\sigma|} \frac{U_G(i_{h-1}, i_h)}{U_G(i_h, i_h)}\right) \\
&\geqslant \frac{1}{N - \ell + 1} \left(1 - \sum_{\sigma \in \mathcal{B}_{\mathcal{I}}(\alpha_\ell[\boldsymbol{s}], \beta_\ell[\boldsymbol{s}]),\ |\sigma| \geqslant 2} \prod_{h=1}^{|\sigma|} \frac{Z_{\boldsymbol{s}}(i_{h-1}, i_h)}{N - 2q}\right) \qquad (11)
\end{aligned}$$

Now we only consider the case where the lower bound (11) $\geqslant 0$ for all ℓ. Otherwise Lemma 9 becomes trivially true. Hence we have

$$\frac{\mathsf{ps}_0(\boldsymbol{\mathcal{Q}},\boldsymbol{k})}{\mathsf{ps}_1(\boldsymbol{\mathcal{Q}},\boldsymbol{k})} \geqslant \prod_{\ell=1}^{q}\left(1-\sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(\alpha_\ell[\boldsymbol{s}],\beta_\ell[\boldsymbol{s}]),\ |\sigma|\geqslant 2}\ \prod_{h=1}^{|\sigma|}\frac{Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N-2q}\right)$$

$$\geqslant 1-\sum_{\ell=1}^{q}\sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(\alpha_\ell[\boldsymbol{s}],\beta_\ell[\boldsymbol{s}]),\ |\sigma|\geqslant 2}\ \prod_{h=1}^{|\sigma|}\frac{Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N-2q} \tag{12}$$

$$\geqslant 1-\sum_{\ell=1}^{q}\sum_{1\leqslant a\leqslant b\leqslant t}\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma|\geqslant 2}\ \prod_{h=1}^{|\sigma|}\frac{Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N-2q} \tag{13}$$

in which (12) is due to $(1-a)(1-b)\geqslant 1-a-b$ for any $a,b\geqslant 0$ and (13) is due to the indicator function R is non-negative and satisfies $\mathsf{R}_{\alpha[\boldsymbol{s}],\beta[\boldsymbol{s}],\ell}[\boldsymbol{s}]=1$. We note that (13) is the exact quantity we pick for $1-g(\boldsymbol{\mathcal{Q}},\boldsymbol{k})$. □

Lemma 10. *If $q\leqslant N/4$, then,*

$$\mathbb{E}_{\boldsymbol{k}}\left(\sum_{\ell=1}^{q}\sum_{1\leqslant a\leqslant b\leqslant t}\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),|\sigma|\geqslant 2}\ \prod_{h=1}^{|\sigma|}\frac{Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N-2q}\right)\leqslant\frac{t^2(4q)^{t+1}}{N^t}\ .$$

Proof. By the sum of expectation and noting that none of $\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b)$ would have $|\sigma|\geqslant 2$ if $a=b$, we have

$$\mathbb{E}_{\boldsymbol{k}}\left(\sum_{\ell=1}^{q}\sum_{1\leqslant a\leqslant b\leqslant t}\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma|\geqslant 2}\ \prod_{h=1}^{|\sigma|}\frac{Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N-2q}\right)$$

$$=\sum_{\ell=1}^{q}\sum_{1\leqslant a<b\leqslant t}\ \sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma|\geqslant 2}\mathbb{E}_{\boldsymbol{s}}\left(\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\cdot\prod_{h=1}^{|\sigma|}\frac{Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N-2q}\right).$$

Hence it is sufficient to derive bounds for each (a,b,σ). Note that for each a,b, $R_{2a-1,2b,j}[\boldsymbol{s}]$ only depends on the subkeys $s_0,\ldots,s_{a-2},s_{b+1},\ldots,s_t$, which are $(a-2+1)+(t-(b+1)+1)=t-b+a-1$ subkeys in total.

Next, given a fixed $\sigma=((i_0,i_1),(i_1,i_2),\ldots,(i_{|\sigma|-1},i_{|\sigma|}))$, we analyze the key dependency for each $Z_{\boldsymbol{s}}(i_{h-1},i_h)$.

1. For $Z_{\boldsymbol{s}}(i_0,i_1)$, note that $i_0=2a-1$ which is odd, and i_1 is even. So $Z_{\boldsymbol{s}}(i_0,i_1)$ $(i_1-i_0-1)/2$ subkeys between L_{i_0} and L_{i_1}.
2. For any (i_{h-1},i_h) where $h>1$, given i_{h-1} is an even number, implying that $L_{i_{h-1}}$ and $L_{i_{h-1}+1}$ are connected by "key-edges", always forming a perfect matching regardless of the subkey choice. Then the equality $Z_s(i_{h-1},i_h)=Z_s(i_{h-1}+1,i_h)$ always holds. And we can see that $Z_s(i_{h-1}+1,i_h)$ only depends on $(i_h-i_{h-1}-2)/2$ subkeys.

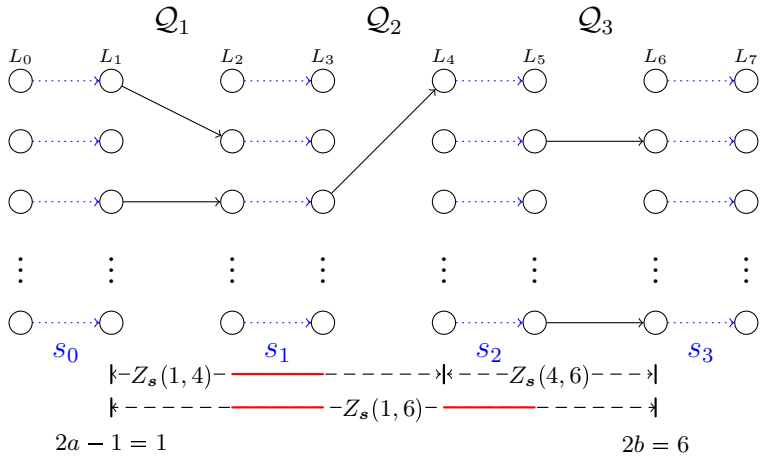

Fig. 1. A 3-round KAC with fixed query records $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3$. The subkeys $\boldsymbol{s} = (s_0, \ldots, s_3)$ are random and to be sampled. The red solid line indicates that the $Z_{\boldsymbol{s}}(\mathsf{left}, \mathsf{right})$ that counts the number of paths from L_{left} to L_{right} depends on the corresponding subkeys. Consider $2a-1=1, 2b=6$, then $\mathsf{R}_{1,6,\ell}[\boldsymbol{s}] = 1$ and depends on $(a-1)+(3-b)=0$ subkeys, because any s_0 allows x_ℓ from L_0 to reach L_1, and y_ℓ from L_7 to L_6. For $\sigma = ((1,6))$, the value of $Z_{\boldsymbol{s}}(1,6)$ depends on two subkeys s_1, s_2. However, if the σ is further paritioned into $((1,4),(4,6))$, then $Z_{\boldsymbol{s}}(1,4)$ depends on s_1 but $Z_{\boldsymbol{s}}(4,6)$ does not depend on any subkeys, because $Z_{\boldsymbol{s}}(4,6) = Z_{\boldsymbol{s}}(5,6) = |\mathcal{Q}_3|$.

Also note that the sets of dependent subkeys for $Z_s(i_{h-1}, i_h)$ and $\mathsf{R}_{2a-1,2b,j}[s]$ are disjoint. Putting the results altogether, after fixing (a, b, σ), the total number of subkeys that each expectation term depends on are at most

$$\begin{aligned}
\#\text{dependent subkeys} &= (t-b+a-1) + \frac{i_1 - i_0 - 1}{2} + \sum_{h=2}^{|\sigma|} \left(\frac{i_h - i_{h-1}}{2} - 1 \right) \\
&= (t-b+a-1) + \frac{\sum_{h=1}^{|\sigma|}(i_h - i_{h-1}) - 1}{2} - (|\sigma| - 1) \\
&= t - b + a - 1 + \frac{2b-2a}{2} - |\sigma| + 1 \\
&= t - |\sigma| \leqslant t - 2 \,,
\end{aligned}$$

in which we observe that a summation term of (a, b, σ) depends on fewer subkeys if the size of σ is larger (See Figure 1 for a specific case illustration). Because our construction ensures that any $t-2$ subkeys are independently and uniformly distributed, the random variables in each expectation terms are mutually independent and hence we can break the terms into

$$\mathbb{E}_{\boldsymbol{k}} \left(\sum_{\ell=1}^{q} \sum_{1 \leqslant a \leqslant b \leqslant t} \mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}] \sum_{\sigma \in \mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma| \geqslant 2} \prod_{h=1}^{|\sigma|} \frac{Z_{\boldsymbol{s}}(i_{h-1}, i_h)}{N - 2q} \right)$$

$$\leqslant \sum_{\ell=1}^{q} \sum_{1\leqslant a<b\leqslant t} \sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma|\geqslant 2} \mathbb{E}_{\boldsymbol{s}}\left(\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\cdot\prod_{h=1}^{|\sigma|}\frac{2Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N}\right)$$

$$= \sum_{\ell=1}^{q} \sum_{1\leqslant a<b\leqslant t} \sum_{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),\ |\sigma|\geqslant 2} \mathbb{E}_{\boldsymbol{s}}\left(\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\right)\cdot\prod_{h=1}^{|\sigma|}\mathbb{E}_{s}\left(\frac{2Z_{\boldsymbol{s}}(i_{h-1},i_h)}{N}\right) \tag{14}$$

$$\leqslant \sum_{\ell=1}^{q} \sum_{1\leqslant a<b\leqslant t} \left(\frac{q}{N}\right)^{t-b+a-1} \sum_{\substack{\sigma\in\mathcal{B}_{\mathcal{I}}(2a-1,2b),\\ |\sigma|\geqslant 2}} \left(\frac{2q}{N}\right)^{(i_1-i_0+1)/2} \prod_{h=2}^{|\sigma|}\left(\frac{2q}{N}\right)^{(i_h-i_{h-1})/2} \tag{15}$$

$$\leqslant \sum_{\ell=1}^{q} \sum_{1\leqslant a<b\leqslant t} \left(\frac{q}{N}\right)^{t-b+a-1}\cdot\left(\frac{4q}{N}\right)^{b-a+1} \leqslant t^2\cdot\frac{(4q)^{t+1}}{N^t}\ . \tag{16}$$

In the above calculation, (14) is due to the subkeys are $(t-2)$-wise independent. The first "q/N" term of (15) comes from moving the $\mathbb{E}_{\boldsymbol{s}}\left(\mathsf{R}_{2a-1,2b,\ell}[\boldsymbol{s}]\right)$, and inside the summation the "$2q/N$" terms are the direct calculation upper bound of $\mathbb{E}_s\left(2Z_{\boldsymbol{s}}(i_{h-1},i_h)/N\right)$ for each (i_{h-1},i_h). Finally we have the first inequality of (16) holds because the size of $\mathcal{B}_{\mathcal{I}}(2a-1,2b)$ is upper-bounded by 2^{b-a}, which is absorbed into "$2q/N$" term yielding a "$4q/N$" term. □

6 Concluding the Proof

Given the similarity of proofs for both theorems, we provide the proof of Theorem 1 here and left the proof of Theorem 2 to the full version.

6.1 Proof of Theorem 1

Proof. We partition the set of transcripts $\mathcal{T} = \mathcal{T}_{\mathsf{good}} \sqcup \mathcal{T}_{\mathsf{bad}}$ according to Definition 1. By applying Lemma 1, we have $\Delta(X_0, X_1) \leqslant \mathbb{E}_{X_1}[g(X_1)] + \mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}]$. We start with bounding $\mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}]$.

Claim.

$$\mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}] \leqslant (t+1)\cdot\frac{3q^{t+1}}{N^t} + 3(t+1)\sqrt{\frac{q^{2t-1}(t+2)n}{N^{2t-2}}} + \frac{t(t+1)}{N^t}\ .$$

Proof (of claim). We note that in the system $\mathbf{S}_1$, the set of bad keys $\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}}}$ is defined only by the query records $\boldsymbol{\mathcal{Q}} = (\mathcal{Q}_E, \mathcal{Q}_1, \ldots, \mathcal{Q}_t)$. Therefore, we have

$$\mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}] \leqslant \mathsf{Pr}_{\boldsymbol{\mathcal{Q}}}\left[|\mathsf{Badkey}_{\mathcal{Q}}| > C\right] + \frac{C}{N^{t-1}}\ .$$

To get the size bound for $\mathsf{Badkey}_{\mathcal{Q}}$, we compute the size of $\mathsf{Badkey}_{\mathcal{Q},i}$ for $0 \leqslant i \leqslant t$. Then, we have

$$|\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},0}| \leqslant \mu_{\boldsymbol{c}_0}(V_1, \mathcal{Q}_2, \mathcal{Q}_3, \ldots, \mathcal{Q}_{t-1}, \mathcal{Q}_t, U_{t+1})$$

$$
\begin{aligned}
|\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},1}| &\leqslant \mu_{\boldsymbol{c}_1}(V_2, \mathcal{Q}_3, \mathcal{Q}_4, \ldots, \mathcal{Q}_t, \mathcal{Q}_E, U_1) \\
&\vdots \\
|\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},t-1}| &\leqslant \mu_{\boldsymbol{c}_{t-1}}(V_t, \mathcal{Q}_E, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-2}, U_{t-1}) \\
|\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},t}| &\leqslant \mu_{\boldsymbol{c}_t}(V_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_{t-1}, U_t)\ .
\end{aligned}
$$

where the linear coefficient tuples $\boldsymbol{c}_i$ are given by the condition 2 of Theorem 1 so that there are two neighboring coefficients that are non-zero, and

$$
\begin{aligned}
\forall i \in \{1, \ldots, t\} :\ U_i &= \{u \mid \exists v :\ (u, v) \in \mathcal{Q}_i\}, & V_i &= \{v \mid \exists u :\ (u, v) \in \mathcal{Q}_i\} \\
U_{t+1} &= \{u \mid \exists v :\ (u, v) \in \mathcal{Q}_E\}, & V_0 &= \{v \mid \exists u :\ (u, v) \in \mathcal{Q}_E\}\ .
\end{aligned}
$$

The size of $\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},i}$ is bounded by $\mu_{\boldsymbol{c}_i}$ because any key $\boldsymbol{k} \in \mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},i}$ is uniquely mapped to the subkeys $(s_0, \ldots, s_{i-1}, s_{i+1}, s_t)$ as the linear mapping has rank $t-1$ (stated in condition 2 of Theorem 1).

Now we can apply Lemma 2 to upper bound $\mathsf{Badkey}_{\boldsymbol{\mathcal{Q}},i}$ with high probability. For every i, by letting $C_i = \frac{3q^{t+1}}{N} + 3q^{t-1/2}\sqrt{(t+2)n}$, we obtain that $\mathsf{Pr}_{\mathcal{Q}}[|\mathsf{Badkey}_{\mathcal{Q},i}| > C_i] \leqslant \frac{2}{N^t}$. Therefore, setting $C = \sum_{i=0}^{t} C_i$, we have

$$
\begin{aligned}
\mathsf{Pr}[X_1 \in \mathcal{T}_{\mathsf{bad}}] &\leqslant \sum_{i=0}^{t} \mathsf{Pr}_{\boldsymbol{\mathcal{Q}}}\left[|\mathsf{Badkey}_{\mathcal{Q}_i}| > C_i\right] + \frac{C}{N^{t-1}} \\
&\leqslant \frac{2t(t+1)}{N^t} + (t+1) \cdot \frac{3q^{t+1}}{N^t} + 3(t+1) \cdot \frac{q^{t-1/2}\sqrt{(t+2)n}}{N^{t-1}}
\end{aligned}
$$

Hence we proved the claim □

The next step is to pick a function g and upper bound $\mathbb{E}_{X_1}[g(X_1)]$. Note that by condition 1 of Theorem 1, any $t-2$ rows of key schedule matrix A has rank $t-2$, implying that any subset of $t-2$ subkeys are independent and uniform. Therefore we can apply Lemma 7 and obtain a function g. Noting that X_1 is in the ideal world so $\boldsymbol{k}$ is sampled independently of $\boldsymbol{\mathcal{Q}}$, we have

$$
\mathbb{E}_{X_1}[g(X_1)] = \mathbb{E}_{\boldsymbol{\mathcal{Q}}}\mathbb{E}_{\boldsymbol{k}}[g(\boldsymbol{\mathcal{Q}}, \boldsymbol{k})] \leqslant \mathbb{E}_{\boldsymbol{\mathcal{Q}}}\left[\frac{t^2(4q)^{t+1}}{N^t}\right] = \frac{t^2(4q)^{t+1}}{N^t}\ .
$$

Then by summing up the two quantities and numerical simplifications, the theorem follows. □

7 Conclusion and Open Problems

In this paper, we provided key schedules of limited independence for t-round key-alternating ciphers achieving tight security. We proved that the t-round key-alternating cipher remains tightly secure for a class of $(t-1)$-wise independent sub-key distributions and, when $t \geqslant 8$, for $(t-2)$-wise sub-key distributions.

While, for $3 \leqslant t \leqslant 7$, our result does not extends to $(t-2)$-wise independent sub-key distributions, we expect that a tighter analysis of the matrix 2-norm for

the sum-capture quantity should give a proof for $4 \leqslant t \leqslant 7$. Also, it is interesting to investigate new methods for bounding the bad keys and proving tight security of 3-round key-alternating cipher with identical key schedule. Further, it would be also interesting to study whether the tightness result holds for $(t-3)$-wise distributions or beyond.

Acknowledgements

We thank the anonymous reviewers for sharing many helpful suggestions that improved the paper. Stefano Tessaro and Xihu Zhang were partially supported by NSF grants CNS-1930117 (CAREER), CNS1926324, CNS-2026774, a Sloan Research Fellowship, and a JP Morgan Faculty Award.

References

1. Noga Alon, Tali Kaufman, Michael Krivelevich, and Dana Ron. Testing triangle-freeness in general graphs. In *17th SODA*, pages 279–288. ACM-SIAM, January 2006.
2. László Babai. The fourier transform and equations over finite abelian groups: An introduction to the method of trigonometric sums. Lecture notes, 1989. Available at http://people.cs.uchicago.edu/~laci/reu02/fourier.pdf.
3. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: An ultra-lightweight block cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *CHES 2007*, volume 4727 of *LNCS*, pages 450–466. Springer, Heidelberg, September 2007.
4. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, François-Xavier Standaert, John P. Steinberger, and Elmar Tischhauser. Key-alternating ciphers in a provable setting: Encryption using a small number of public permutations - (extended abstract). In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 45–62. Springer, Heidelberg, April 2012.
5. Shan Chen, Rodolphe Lampe, Jooyoung Lee, Yannick Seurin, and John P. Steinberger. Minimizing the two-round Even-Mansour cipher. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 39–56. Springer, Heidelberg, August 2014.
6. Shan Chen and John P. Steinberger. Tight security bounds for key-alternating ciphers. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 327–350. Springer, Heidelberg, May 2014.
7. Benoit Cogliati, Rodolphe Lampe, and Yannick Seurin. Tweaking Even-Mansour ciphers. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 189–208. Springer, Heidelberg, August 2015.
8. Benoît Cogliati and Yannick Seurin. Beyond-birthday-bound security for tweakable Even-Mansour ciphers with linear tweak and key mixing. In Tetsu Iwata and Jung Hee Cheon, editors, *ASIACRYPT 2015, Part II*, volume 9453 of *LNCS*, pages 134–158. Springer, Heidelberg, November / December 2015.
9. Benoit Cogliati and Yannick Seurin. On the provable security of the iterated Even-Mansour cipher against related-key and chosen-key attacks. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 584–613. Springer, Heidelberg, April 2015.

10. Joan Daemen and Vincent Rijmen. *The design of Rijndael*, volume 2. Springer, 2002.
11. Orr Dunkelman, Nathan Keller, and Adi Shamir. Minimalism in cryptography: The Even-Mansour scheme revisited. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 336–354. Springer, Heidelberg, April 2012.
12. Avijit Dutta. Minimizing the two-round tweakable Even-Mansour cipher. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 601–629. Springer, Heidelberg, December 2020.
13. Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. *Journal of Cryptology*, 10(3):151–162, June 1997.
14. Jian Guo, Thomas Peyrin, Axel Poschmann, and Matthew J. B. Robshaw. The LED block cipher. In Bart Preneel and Tsuyoshi Takagi, editors, *CHES 2011*, volume 6917 of *LNCS*, pages 326–341. Springer, Heidelberg, September / October 2011.
15. Thomas P. Hayes. A large-deviation inequality for vector-valued martingales, 2003. Available at https://www.cs.unm.edu/~hayes/papers/VectorAzuma/VectorAzuma20050726.pdf.
16. Viet Tung Hoang and Stefano Tessaro. Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 3–32. Springer, Heidelberg, August 2016.
17. Eike Kiltz, Krzysztof Pietrzak, and Mario Szegedy. Digital signatures with minimal overhead from indifferentiable random invertible functions. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 571–588. Springer, Heidelberg, August 2013.
18. Rodolphe Lampe, Jacques Patarin, and Yannick Seurin. An asymptotically tight security analysis of the iterated Even-Mansour cipher. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 278–295. Springer, Heidelberg, December 2012.
19. Ueli M. Maurer. Indistinguishability of random systems. In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 110–132. Springer, Heidelberg, April / May 2002.
20. Alessandro Panconesi and Aravind Srinivasan. Fast randomized algorithms for distributed edge coloring (extended abstract). In Norman C. Hutchinson, editor, *11th ACM PODC*, pages 251–262. ACM, August 1992.
21. Jacques Patarin. The "coefficients H" technique (invited talk). In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *SAC 2008*, volume 5381 of *LNCS*, pages 328–345. Springer, Heidelberg, August 2009.
22. John Steinberger. Improved security bounds for key-alternating ciphers via hellinger distance. Cryptology ePrint Archive, Report 2012/481, 2012. http://eprint.iacr.org/2012/481.
23. John P. Steinberger. Counting solutions to additive equations in random sets. *CoRR*, abs/1309.5582, 2013.
24. Yusai Wu, Liqing Yu, Zhenfu Cao, and Xiaolei Dong. Tight security analysis of 3-round key-alternating cipher with a single permutation. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 662–693. Springer, Heidelberg, December 2020.