Combining Outputs of a Random Permutation: New Constructions and Tight Security Bounds by Fourier Analysis

Itai Dinur

Ben-Gurion University and Georgetown University

Abstract. We consider constructions that combine outputs of a single permutation $\pi : \{0,1\}^n \rightarrow \{0,1\}^n$ using a public function. These are popular constructions for achieving security beyond the birthday bound when implementing a pseudorandom function using a block cipher (i.e., a pseudorandom permutation). One of the best-known constructions (denoted SXoP$[2,n]$) XORs the outputs of 2 domain-separated calls to π. Modeling π as a uniformly chosen permutation, several previous works proved a tight information-theoretic indistinguishability bound for SXoP$[2,n]$ of about $q/2^n$, where q is the number of queries. However, tight bounds are unknown for the generalized variant (denoted SXoP$[r,n]$) which XORs the outputs of $r \geq 2$ domain-separated calls to a uniform permutation.
In this paper, we obtain two results. Our first result improves the known bounds for SXoP$[r,n]$ for all (constant) $r \geq 3$ (assuming $q \leq O(2^n/r)$ is not too large) in both the single-user and multi-user settings. In particular, for $r = 3$, our bound is about $\sqrt{u}q_{\max}/2^{2.5n}$ (where u is the number of users and $q_{\max}$ is the maximal number of queries per user), improving the best-known previous result by a factor of at least 2^n.
For odd r, our bounds are tight for $q > 2^{n/2}$, as they match known attacks. For even r, we prove that our single-user bounds are tight by providing matching attacks.
Our second and main result is divided into two parts. First, we devise a family of constructions that output n bits by efficiently combining outputs of 2 calls to a permutation on $\{0,1\}^n$, and achieve multi-user security of about $\sqrt{u}q_{\max}/2^{1.5n}$. Then, inspired by the CENC construction of Iwata [FSE'06], we further extend this family to output $2n$ bits by efficiently combining outputs of 3 calls to a permutation on $\{0,1\}^n$. The extended construction has similar multi-user security of $\sqrt{u}q_{\max}/2^{1.5n}$.
The new single-user ($u = 1$) bounds of $q/2^{1.5n}$ for both families should be contrasted with the previously best-known bounds of $q/2^n$, obtained by the comparable constructions of SXoP$[2,n]$ and CENC.
All of our bounds are proved by Fourier analysis, extending the provable security toolkit in this domain in multiple ways.

1 Introduction

Many efficient implementations of pseudorandom functions today use block ciphers, which are pseudorandom permutations that only achieve security up to

the birthday bound of $q = 2^{n/2}$ queries (where n is the block length). Since the security of many cryptosystems (such as encryption modes, MAC algorithms and authenticated encryption schemes) is based on pseudorandom functions, beyond-birthday bound security has become a popular research area, initiated in papers by Bellare, Krovetz, and Rogaway [2], and by Hall, Wagner, Kelsey, and Schneier [18].

1.1 XORing Permutation Outputs

One of the best-known constructions for achieving security beyond the birthday bound XORs the outputs of 2 permutations calls. This constructions has two main variants. The first variant, denoted XoP$[2, n]$ (XOR of Permutations), uses two permutations $\pi_1, \pi_2 : \{0,1\}^n \mapsto \{0,1\}^n$ to define XoP$[2, n]_{\pi_1,\pi_2} : \{0,1\}^n \mapsto \{0,1\}^n$ by XoP$[2, n]_{\pi_1,\pi_2}(i) = \pi_1(i) \oplus \pi_2(i)$. In practice, π_1 and π_2 are implemented using a block cipher, instantiated with independent keys. The second variant, denoted SXoP$[2, n]$, uses 2 domain-separated calls to a single permutation $\pi : \{0,1\}^n \mapsto \{0,1\}^n$ to define SXoP$[2, n]_\pi : \{0,1\}^{n-1} \mapsto \{0,1\}^n$ by SXoP$[2, n]_\pi(i) = \pi(0\|i) \oplus \pi(1\|i)$ (where $\|$ denotes concatenation). As in the first variant, π is implemented using a block cipher. However, in information-theoretic security proofs, the block ciphers in both variants are replaced by idealized random permutations.

The second variant is more efficient in the sense that it only requires a single key. Yet, the advantage of the first variant is that it achieves better concrete security in idealized models.

Generalizations. Natural generalizations of the above variants XOR the outputs $r \geq 2$ permutations calls. The aim of these generalizations is to obtain even better security bounds.

In this paper, we are mainly interested in a generalization of the second variant, denoted SXoP$[r, n]$. It uses $r \geq 2$ domain-separated calls to a single permutation $\pi : \{0,1\}^n \mapsto \{0,1\}^n$ to define SXoP$[r, n]_\pi : \{0,1\}^{n-\lceil \log r \rceil} \mapsto \{0,1\}^n$ by SXoP$[r, n]_\pi(i) = \pi(0\|i) \oplus \pi(1\|i) \oplus \ldots \oplus \pi(r-1\|i)$.

Previous results. Both variants have been analyzed in the idealized model by numerous papers in both the single-user and multi-user settings. The first variant (XoP) that uses independent permutations (and its generalized version) was analyzed in [7,8,9,10,11,13,22,23,25,26]. A tight security bound for XoP and its generalization was derived in [12] (also see [14] for XoP$[2, n]$), and further extended to the multi-user setting.

Works that analyzed the second variant SXoP (and its generalization) include [1,3,5,9,11,13,19,25,26]. In particular, for SXoP$[2, n]$ a security bound of about $\frac{q}{2^n}$ was proved in [9,11,13]. This bound is tight as it is matched by a simple attack that checks whether the element 0 is output. The bound was extended to give a tight bound in the multi-user setting in [3,19].

For the more general scheme SXoP$[r, n]$ with $r \geq 3$, tight bounds are unknown. The particular case of $r = 3$ was analyzed by Bhattacharya and Nandi in [5], deriving a bound of about $\frac{\sqrt{uq_{\max}}}{2^n}$ in the multi-user setting (where u is the number of users and $q_{\max}$ is the maximal number of queries per user).

Remark 1. In practice, each permutation is instantiated with a keyed block cipher. In such computational settings, one needs to add an additional term (or terms) to the bounds derived above which take into account the optimal advantage in distinguishing the underlying block cipher (or block ciphers) from a uniformly chosen permutation (or permutations).

Remark 2. The restriction that the PRF should not be called with more than $q_{\max}$ queries implies that the key should be rotated every $q_{\max}$ invocations in practice. For the schemes we consider, there is a trivial attack on a single user that achieves constant advantage by querying the PRF on the entire domain. Thus, such a restriction is necessary if one desires security beyond 2^n queries (per all users) in the multi-user setting.

1.2 Iwata's PRF construction

At FSE 2006 [20], Iwata introduced CENC, which is a beyond-birthday bound secure mode of operation. Since its introduction, CENC has been very influential and it is currently considered for practical use as part of the DNDK-GCM mode [17]. CENC is built from a PRF, $\mathrm{F}[w, n] : \{0,1\}^{n-\lceil \log(w+1) \rceil} \mapsto \{0,1\}^{wn}$ using an underlying permutation $\pi : \{0,1\}^n \mapsto \{0,1\}^n$ and defined as

$$\mathrm{F}[w, n]_\pi(i) = (\pi(0\|i) \oplus \pi(1\|i))\|(\pi(0\|i) \oplus \pi(2\|i))\| \ldots \|(\pi(0\|i) \oplus \pi(w\|i)).$$

Thus, in order to generate wn bits of output, F only makes $w + 1$ calls to π, whereas SXoP$[2, n]$ makes $2w$ calls.

When modeling π as an ideal permutation, [4,9,21] proved that F[w,n] has an indistinguishability advantage upper bound of about $\frac{w^2 q}{2^n}$.

1.3 Our Results

In this paper, we obtain two results.

Result 1 - analysis of SXoP$[r, n]$. We improve known bounds for SXoP$[r, n]$ for all (constant) $r \geq 3$ (assuming $q \leq O(2^n/r)$ is not too large).

For odd r, we derive a bound of about $\frac{q}{2^{n(r-0.5)}}$ is the single-user setting and $\frac{\sqrt{u}q_{\max}}{2^{n(r-0.5)}}$ in the multi-user setting. In particular, for $q = 3$, our bound $\frac{\sqrt{u}q_{\max}}{2^{2.5n}}$ improves the best-known previous one of $\frac{\sqrt{uq_{\max}}}{2^n}$ obtained in [5] by a factor of at least 2^n. Our bounds for odd r are tight up to a constant factor (for $q \geq 2^{n/2}$), as they match attacks published by Patarin [27]. This includes the multi-user setting, where our bounds are matched by the simple generalization of the attacks

of Patarin, which applies the single-user attack independently to each user and outputs a majority vote over the answers.

For even r, we prove a bound of about $\frac{q}{2^{nr/2}}$ in the single-user setting and an additional (slightly more complicated) bound of about $\min\left(\frac{\sqrt{uq_{\max}}}{2^{n(r/2-1/2)}}, \frac{uq_{\max}}{2^{nr/2}}\right)$ in the multi-user setting. Furthermore, we prove that our single-user bounds are tight by providing matching attacks, which improve the ones of [27]. The bound for even r in multi-user setting is obtained by combining two different bounds, and we conjecture that it is not tight in all settings. We leave the problem of improving this bound (or devising a matching attack) to future work.

Interestingly, our results show (for example) that SXoP$[3,n]$ (with a tight bound of $\frac{q}{2^{2.5n}}$) is provably more secure than SXoP$[4,n]$ (with a tight bound of $\frac{q}{2^{2n}}$). More generally, for odd $r \geq 3$, SXoP$[r,n]$ (with a bound of $\frac{q}{2^{n(r-0.5)}}$) is provably more secure than SXoP$[2r-2,n]$ (with a bound of $\frac{q}{2^{n(r-1)}}$). Intuitively, the reason for this gap is that for odd r every element in $\{0,1\}^n$ output by SXoP$[r,n]$ is marginally uniformly distributed, while for even r it is not.

Result 2 - definition and analysis of LXoP$[L,n]$ and LXoP$[L,2,n]$.

LXoP*[L,n]*. We propose a family of constructions that output n bits by publicly combining outputs of 2 calls to a single permutation on $\{0,1\}^n$, and achieve multi-user security of about $\frac{\sqrt{u}q_{\max}}{2^{1.5n}}$ (as long as $q_{\max} \leq O(2^n)$ is not too large). Hence, these constructions are provably secure up to $u = o(2^n)$ users for $q_{\max} \geq \Omega(2^n)$. Our (single-user) bound of $\frac{q}{2^{1.5n}}$ improves upon the best previous bound of $\frac{q}{2^n}$ for a construction with similar parameters (obtained for SXoP$[2,n]$).

Our construction family is parameterized by a public linear orthomorphism $L : \{0,1\}^n \to \{0,1\}^n$, which is an invertible linear transformation such that $L'(x) = x \oplus L(x)$ is itself a permutation. The construction is denoted LXoP$[L,n]$ and defined as LXoP$[L,n]_\pi(i) = \pi(0\|i) \oplus L(\pi(1\|i))$, where $i \in \{0,1\}^{n-1}$.

It is easy to show that our bound $\frac{\sqrt{u}q_{\max}}{2^{1.5n}}$ is tight assuming $q \geq 2^{n/2}$ by similar attacks to the ones of [27]. Note that the bound we obtain is of the same order as the tight bound for XoP$[2,n]$, obtained in [12,14].

Importantly, there are many linear orthomorphisms $L : \{0,1\}^n \to \{0,1\}^n$ with the desired properties which are very simple and easy to implement in practice. One example is $L(x^{(1)}, x^{(2)}) = (x^{(2)}, x^{(1)} \oplus x^{(2)})$, where $x^{(1)}, x^{(2)} \in \{0,1\}^{n/2}$. Another example that may be more efficient to implement in hardware is $L(x) = (x \ggg 1) \oplus (x_1, 0, \ldots, 0)$, i.e., cyclically rotate x by 1 bit to the right and XOR the first bit of x (denoted x_1) to the first bit of the result. Yet another example is doubling in the field $\mathbb{F}_{2^n}$. More details about linear orthomorphisms over $\mathbb{F}_2^n$ can be found in [16].

Intuitively, the main reason that such constructions have a high security level is that (unlike SXoP$[2,n]$), every element generated by LXoP$[L,n]$ is marginally uniform in $\{0,1\}^n$. Indeed, let $x \in \{0,1\}^n$ be such an element and write it as $x = y \oplus L(z)$, where $y, z \in \{0,1\}^n$ are drawn uniformly without replacement. Then, fixing any $a \in \{0,1\}^n$, the equality $x = a$ is equivalent to $y \oplus L(z) = a$. If

$y, z \in \{0,1\}^n$ were drawn uniformly and independently, then since L is invertible, the equation $y \oplus L(z) = a$ would have exactly 2^n solutions. However, since y, z are drawn uniformly without replacement, we subtract the solutions that satisfy $y = z$, and as L is an orthomorphism, the equation $y \oplus L(y) = a$ has exactly one solution. Consequently, for any $a \in \{0,1\}^n$, the equation $y \oplus L(z) = a$ has exactly $2^n - 1$ solutions, namely, $x = y \oplus L(z)$ is uniformly distributed.

We remark that the use of linear orthomorphisms in cryptography (particularly in design of block ciphers) is not new. See [6] and references therein for examples. Hence, the main novelty of this work with respect to the LXoP$[L, n]$ family (and its generalization below) is in the security proof, rather than the actual design.

LXoP*[L,2,n]*. After analyzing LXoP$[L, n]$, we extend the construction to obtain better efficiency by outputting $2n$ bits via 3 calls to the underlying permutation. Specifically, we define LXoP$[L, 2, n] : \{0,1\}^{n-2} \mapsto \{0,1\}^{2n}$ as

$$\text{LXoP}[L, 2, n]_\pi(i) = (\pi(0\|i) \oplus L(\pi(1\|i))) \,\|\, (\pi(1\|i) \oplus L(\pi(2\|i)))\,.$$

We prove that LXoP$[L, 2, n]$ offers similar security to LXoP$[L, n]$ in both the single-user and multi-user settings, given that L is a linear orthomorphism. Compared to Iwata's PRF [20], F$[2, n]$, the indistinguishability bound is improved from about $\frac{q}{2^n}$ to $\frac{q}{2^{3n/2}}$ (in the single-user setting), while having comparable parameters.

LXoP$[L, w, n]$. One can further extend LXoP to output wn bits via $w + 1$ permutation calls, similarly to Iwata's PRF. Specifically, define

$$\text{LXoP}[L, w, n]_\pi(i) = (\pi(0\|i) \oplus L(\pi(1\|i))) \,\|\, \ldots \,\|\, (\pi(w-1\|i) \oplus L(\pi(w\|i)))\,,$$

where $i \in \{0,1\}^{n-\lceil \log(w+1) \rceil}$. To achieve high security, we require that the iterated invertible linear function L^j has no short cycles of length up to w, namely for every $x \in \{0,1\}^n$ such that $x \neq 0$ and $1 \leq j \leq w$, $x \oplus L^j(x) \neq 0$. Such efficient functions L are easy to build (e.g., from linear-feedback shift registers).

While it is not difficult (albeit somewhat technical) to extend our security analysis of LXoP$[L, 2, n]$ to LXoP$[L, w, n]$ for very small values of $w > 2$, the analysis for general w is more involved and we leave it to future work.

We remark that a different variant of LXoP$[L, w, n]_\pi(i)$ defines the j'th output block (for $j = 1, \ldots, w$) as $L^j(\pi(0\|i)) \oplus \pi(j\|i)$. However, this variant seems to be inferior to the one above in terms of both security (for large w) and efficiency, since the computations of $L^j(\pi(0\|i))$ for different values of j are more difficult to parallelize.

Implications of our results. Since tight bounds are known for SXoP$[r, n]$ with $r = 2$, we focus on $r \geq 3$, aiming for a very high security level at the expense of more permutation calls. Our analysis is therefore mostly of theoretical interest, although it could be practically meaningful when using a block cipher with a short block length n for which the security of SXoP$[2, n]$ is insufficient.

On the other hand, the LXoP$[L, n]$ and LXoP$[L, 2, n]$ constructions combine the efficiency of SXoP$[2, n]$ and CENC (respectively) with the very high security level of XoP$[2, n]$. In our context, efficiency is mainly measured by (1) the number of random permutations (block cipher keys), (2) the number of permutation calls per one PRF call, and 3) the number of bits output in one PRF call. We further argue above that the PRFs can be implemented in practice with little overhead. Consequently, we believe that these PRF constructions are of practical interest.

1.4 Technical Overview

Similarly to the previous works [12,14,15], we prove our results by Fourier analysis. We start by elaborating on the techniques of [12,14] that are relevant to this paper.

Previous techniques [12,14]. First, the distinguishing advantage of the adversary is bounded by the statistical distance between the distribution generated by the analyzed construction and the uniform distribution. Consider a sample from a distribution generated by the analyzed construction, which is over $\mathbb{F}_2^{q \times n}$ (i.e., composed of q elements in $\{0,1\}^n$). The statistical distance of this distribution from the uniform distribution can be bounded in the "Fourier domain" by bounding the bias (i.e., Fourier coefficient) of each of the 2^{qn} possible masks (i.e., linear equations over $\mathbb{F}_2$) applied to the bits of the sample.

In [12,14], the task of bounding the Fourier coefficients for the distribution function generated by the XoP construction was reduced to the task of bounding the Fourier coefficients for the distribution generated by the underlying primitive, namely, a random permutation. This reduction was based on the fact that XORing together samples generated by independent random permutations corresponds to a convolution operation, which is simple multiplication in the Fourier domain.

Considering k elements (for any $1 \leq k \leq q$) drawn uniformly without replacement, the proof of [12] used bounds on two quantities of Fourier coefficients on masks that involve all of these k elements (called level-k coefficients).

1. The maximal level-k Fourier coefficient in absolute value.
2. The level-k Fourier weight, which is equal to the sum of squares of all level-k Fourier coefficients.

Our techniques. We would like to use a similar approach to bound the distinguishing advantage of the adversary against the SXoP and LXoP constructions. However, unlike the XoP construction, these do not involve XORing together independent permutations. Therefore, the step that reduces the analysis to bounding the Fourier coefficients of a random permutation via convolution is no longer applicable.

Nevertheless, we prove that the Fourier coefficients of the distribution generated by the SXoP and LXoP constructions are, in fact, structured subsets of the Fourier coefficients of a random permutation.

For example, denote by $x \in \{0,1\}^n$ a single element of a sample generated by SXoP$[2,n]$. Consider a mask involving a single element $\alpha \in \{0,1\}^n \neq 0$ (i.e., a mask of level 1), and assume we wish to analyze the bias of the linear equation $\alpha_1 x_1 \oplus \ldots \oplus \alpha_n x_n$. Since x is generated by SXoP$[2,n]$, we can write $x = y \oplus z$, where $y, z \in \{0,1\}^n$ are generated by a random permutation. The above linear equation can therefore be written as $\alpha_1(y_1 \oplus z_1) \oplus \ldots \oplus \alpha_n(y_n \oplus z_n) = (\alpha_1 y_1 \oplus \ldots \oplus \alpha_n y_n) \oplus (\alpha_1 z_1 \oplus \ldots \oplus \alpha_n z_n)$, whose bias is exactly the Fourier coefficient of a random permutation on the level-2 symmetric mask $(\alpha, \alpha) \in \{0,1\}^{2n}$.

In general, level-k Fourier coefficients of the distribution generated by SXoP$[r,n]$ correspond to symmetric level-(rk) Fourier coefficients of a random permutation. One can similarly prove that level-k Fourier coefficients of the distribution generated by LXoP$[L,n]$ correspond to level-$2k$ Fourier coefficients of a random permutation (with a certain structure that depends on L). A similar property also holds for LXoP$[L,2,n]$. Therefore, we can use the two bounds above on the Fourier coefficients of a random permutation to analyze the distributions generated by the SXoP and LXoP constructions.

Framework for bounding Fourier weight of sampling without replacement on structured subsets of masks. Unfortunately, using the general level-k bounds naively is not sufficient to obtain tight indistinguishability bounds for the constructions we analyze, particularly for LXoP. Essentially, the general level-k bound on the weight (i.e., the second bound) is tight for dense subsets of masks that contain (a large fraction of) all level-k masks. However, the subsets we need to analyze are structured and very sparse.

As a result, in this paper we develop a framework that allows to bound the Fourier weight of the sampling without replacement density function (normalized distribution function) on structured subsets of masks. The framework takes into account the particular structure of the subset and significantly improves the naive bounds for the constructions we analyze.

Technically, the framework uses a (known) recursive formula for calculating the Fourier coefficient on any single mask α as a sum of Fourier coefficients on lower-level masks, derived from α. We show how to manipulate the formula to collectively analyze the Fourier weight of a subset of masks that have a common structure, determined by the construction we analyze. Specifically, each recursive call bounds the weight of an increasingly denser subset of masks (of a lower level), and we apply the general bounds only at the leaves of the recursion tree, where they are closer to being tight. The power and generality of this framework is demonstrated by applying it to obtain tight indistinguishability bounds for all constructions we analyze in this paper. A notable exception to the above is the SXoP$[r,n]$ construction with even r, whose analysis requires an additional central technical contribution, summarized below.

Mixed $\mathcal{L}^1$ and $\mathcal{L}^2$ bounds. For the SXoP$[r,n]$ construction with even r the above strategy is not sufficient to obtain tight indistinguishability bounds. Essentially, this is because of a quadratic loss of the standard Cauchy-Schwarz inequality that bounds the statistical distance ($\mathcal{L}^1$ distance) of the analyzed distribution

to the uniform distribution using the $\mathcal{L}^2$ distance. In order to overcome this loss, we bound the statistical distance by a mixture of $\mathcal{L}^1$ and $\mathcal{L}^2$ bounds using the Fourier decomposition of the distribution (density) function. While such mixed bounds have been used before in a hybrid argument (e.g., in [11]), we stress that our mixed bounds are purely analytical in the sense that the "hybrids" that we use do not necessarily correspond to actual distributions, but rather to a Fourier decomposition of the density function.

An additional advantage of this technique is that it allows to *lower bound* the statistical distance (i.e. analyze the optimal attack) in the Fourier domain using the reverse triangle inequality. Indeed, the optimal attack on $\mathrm{SXoP}[r, n]$ reveals itself during the analysis of level-1 Fourier coefficients. This attack simply checks whether there is a 0 element of $\{0,1\}^n$ in the sample. We note that combinatorial analysis of the attack for arbitrary even $r \geq 4$ is less straightforward.

1.5 Paper Structure

The rest of this paper is organized as follows. Next, in Section 2, we describe preliminaries. In Section 3 we develop our framework for bounding the Fourier weight of sampling without replacement on structured subsets of masks. In Section 4 we prove our results regarding the SXoP construction, while in Section 5 and Section 6 we analyze the LXoP construction and its variants.

2 Preliminaries

In this section we describe preliminaries. Unless stated otherwise, missing proofs are found in Appendix A.

For a positive integer m (i.e., $m \in \mathbb{Z}^{\geq 1}$), denote $[m] = \{1, 2, \ldots, m\}$. For $m_1, m_2 \in \mathbb{Z}$ such that $m_1 \leq m_2$, denote $[m_1, m_2] = \{m_1, m_1 + 1, \ldots, m_2\}$. For a set $\mathcal{A}$, denote its size by $|\mathcal{A}|$. For any integer $k > 0$ and a real number t, define the falling factorial as $(t)_k = t(t-1)\ldots(t-(k-1))$. Further define $(t)_0 = 1$.

Let $n, m \in \mathbb{Z}^{\geq 1}$ such that $n \geq m$. Then, $(\frac{n}{m})^m \leq \binom{n}{m} \leq (\frac{e \cdot n}{m})^m$.

Let x be an element (from an arbitrary domain) and let $m \in \mathbb{Z}^{\geq 1}$. Define $x^{\circ m} = \underbrace{(x, \ldots, x)}_{m \text{ times}}$ to be the sequence of m repetitions of x. For a sequence $(x_1, \ldots, x_k)$, define $(x_1, \ldots, x_k)^{\odot m} = ((x_1)^{\circ m}, \ldots, (x_k)^{\circ m})$.

Let $m \in \mathbb{Z}^{\geq 1}$. We denote the sequence of elements $(x_1, \ldots, x_m)$ by $x_{1..m}$. Similarly, the sequence of elements $(x^{(1)}, \ldots, x^{(m)})$ is denoted by $x^{1..m}$. Furthermore, for $m_1, m_2 \in \mathbb{Z}^{\geq 1}$, denote the sequence of $m_1 m_2$ elements $(x_1^{(1)}, \ldots, x_1^{(m_2)}, \ldots, x_{m_1}^{(1)}, \ldots, x_{m_1}^{(m_2)})$ by $x_{1..m_1}^{1..m_2}$.

Let $\mathbb{F}$ be a field and $v \in \mathbb{F}^{k_1 \times k_2}$ a matrix of elements in $\mathbb{F}$. We index the elements of v in a natural way, namely, for $i \in [k_1]$, $v_i \in \mathbb{F}^{k_2}$ is the i'th row of v and for $j \in [k_2]$, $v_{i,j} \in \mathbb{F}$ is its j'th entry.

For two (row) vectors $v, u \in \mathbb{F}^k$, we denote by $\langle u, v \rangle_{\mathbb{F}} = u \cdot v^{\mathrm{T}} = \sum_{i \in [k]} u_i v_i$ their inner product (where v^{T} is the transpose of v and addition and multi-

plication are over $\mathbb{F}$). Similarly, for matrices $v, u \in \mathbb{F}^{k_1 \times k_2}$, define $\langle u, v \rangle_{\mathbb{F}} = \sum_{i \in [k_1]} u_i \cdot (v_i)^{\mathrm{T}} = \sum_{(i,j) \in [k_1] \times [k_2]} u_{i,j} v_{i,j}$.

In this paper, we typically deal with matrices $x \in \mathbb{F}_2^{k \times n}$, where n is considered a parameter and k may vary. We denote $N = 2^n$.

Let $L \in \mathbb{F}_2^{n \times n}$. Denote by L^{T} the transpose of L. Further, let $x \in \mathbb{F}_2^{k \times n}$. We define $L(x) \in \mathbb{F}_2^{k \times n}$ by $L(x)_i = x_i \cdot L$ for $i \in [k]$ (where we view x_i as a row vector in $\mathbb{F}_2^n$, multiplied with L).

Define $\mathbb{1}$ as the 0\1 indicator function that takes as input a predicate.

Asymptotic notation. While all of our results are fully explicit, we sometimes use standard asymptotic notation to give intuition about the bounds we obtain. In particular, we use the notation $O_r(\cdot)$ and $\Omega_r(\cdot)$ that suppress arbitrary functions of r (for SXoP$[r, n]$ we think of it as a small constant).

2.1 Probability

Definition 1 (Density function). *A (probability) density function on $\mathbb{F}_2^{q \times n}$ is a nonnegative function $\varphi : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}^{\geq 0}$ satisfying $\mathrm{E}_{x \in \mathbb{F}_2^{q \times n}}[\varphi(x)] = 1$, where $x \in \mathbb{F}_2^{q \times n}$ is uniformly chosen.*

We write $x \sim \varphi$ to denote that x is a sample drawn from the associated probability distribution, defined by $\Pr_{x \sim \varphi}[x = y] = \frac{\varphi(y)}{2^{qn}}$ for every $y \in \mathbb{F}_2^{q \times n}$. Specifically, the uniform probability density function over $\mathbb{F}_2^{q \times n}$ is the constant function 1, denoted by $\mathbf{1}_{qn}$.

Let $\mathcal{A} \subseteq \mathbb{F}_2^{q \times n}$. We write $x \sim \mathcal{A}$ to denote that x is selected uniformly at random from $\mathcal{A}$.

Proposition 1 ([24], Fact 1.21). *If $\varphi : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}^{\geq 0}$ is a density function and $f : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}$, then $\mathrm{E}_{x \sim \varphi}[f(x)] = \mathrm{E}_{x \sim \mathbb{F}_2^{q \times n}}[\varphi(x) f(x)]$.*

Definition 2 (Statistical distance). *The statistical distance between two density functions $\varphi, \psi : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}^{\geq 0}$ is $\mathrm{SD}(\varphi, \psi) = \frac{1}{2} \mathrm{E}_{x \sim \mathbb{F}_2^{q \times n}} |\varphi(x) - \psi(x)|$.*

2.2 Fourier Analysis

We define the Fourier-Walsh expansion of functions on the Boolean cube, adapted to our setting, and state the basic results that we will use. These results are mostly taken from [24].

Definition 3 (Fourier expansion). *Given $\alpha \in \mathbb{F}_2^{q \times n}$, define $\chi_\alpha : \mathbb{F}_2^{q \times n} \mapsto \{-1, 1\}$ by*

$$\chi_\alpha(x) = (-1)^{\langle \alpha, x \rangle_{\mathbb{F}_2}} = \prod_{i \in [q]} (-1)^{\langle \alpha_i, x_i \rangle_{\mathbb{F}_2}} = \prod_{i \in [q], j \in [n]} (-1)^{\alpha_{i,j} \cdot x_{i,j}}.$$

The set $\{\chi_\alpha\}_{\alpha\in\mathbb{F}_2^{q\times n}}$ is an orthonormal basis for the set of functions $\{f \mid f\colon \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}\}$, with respect to the normalized inner product $\frac{1}{|\mathbb{F}_2^{q\times n}|}\langle f,g\rangle_{\mathbb{R}} = \mathrm{E}_{x\sim\mathbb{F}_2^{q\times n}}[f(x)g(x)]$. Hence, each $\{f \mid f\colon \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}\}$ can be decomposed to $f = \sum_{\alpha\in\mathbb{F}_2^{q\times n}} \widehat{f}(\alpha)\chi_\alpha$, where $\widehat{f}(\alpha) = \mathrm{E}[\chi_\alpha f]$, and in particular $\widehat{f}(0) = \mathrm{E}[f]$.

Each element in $\{\chi_\alpha\}_{\alpha\in\mathbb{F}_2^{q\times n}}$ is called a *character*. We refer to α as a *mask*, and to $\widehat{f}(\alpha)$ as the *Fourier coefficient of f on α*. To distinguish the domain of characters from the input domain, we write it as $\widehat{\mathbb{F}}_2^{q\times n}$, hence $f(x) = \sum_{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}} \widehat{f}(\alpha)\chi_\alpha(x)$. For a mask $\alpha \in \widehat{\mathbb{F}}_2^{q\times n}$, define

$$\mathcal{NZ}_\alpha = \{i\colon \alpha_i \neq 0\} \text{ and } \#\alpha = |\mathcal{NZ}_\alpha|.$$

We call $\#\alpha$ the *level* of α, and $\widehat{f}(\alpha)$ is a Fourier coefficient of level $\#\alpha$.

For integer parameters $n \geq 1$ and $0 \leq k_0 \leq k_1$, we define the sets of masks $\mathcal{M}^n_{=k_0,k_1} = \{\alpha \in \widehat{\mathbb{F}}_2^{k_1\times n}\colon \#\alpha = k_0\}$, and $\mathcal{M}^n_{\geq k_0,k_1} = \{\alpha \in \widehat{\mathbb{F}}_2^{k_1\times n}\colon \#\alpha \geq k_0\}$.

Definition 4 (Fourier weight and maximal magnitude). *For a function $f : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}$, we define the Fourier weight of f at level k to be*

$$\mathrm{W}^{=k}[f] = \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \#\alpha=k}} \widehat{f}(\alpha)^2 = \sum_{\alpha\in\mathcal{M}^n_{=k,q}} \widehat{f}(\alpha)^2.$$

The maximal magnitude of a level-k Fourier coefficient of f is

$$\mathrm{M}^{=k}[f] = \max_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \#\alpha=k}} \{|\widehat{f}(\alpha)|\} = \max_{\alpha\in\mathcal{M}^n_{=k,q}} \{|\widehat{f}(\alpha)|\}.$$

Proposition 2 ([24], Proposition 1.13 – variance). *The variance of $f : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}$ is $\mathrm{Var}[f] = \mathrm{E}[f^2] - \mathrm{E}[f]^2 = \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \alpha\neq 0}} \widehat{f}(\alpha)^2 = \sum_{k=1}^{q} \mathrm{W}^{=k}[f]$.*

Proposition 3 (Bidirectional bounds on statistical distance from uniform by $\mathcal{L}^1$ and $\mathcal{L}^2$ distances). *Let $\varphi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ be a density function. Let $\mathcal{S} \subset \widehat{\mathbb{F}}_2^{q\times n}$ be any set of masks, which does not contain the zero mask. Let $\overline{\mathcal{S}} = \widehat{\mathbb{F}}_2^{q\times n}\backslash\{\mathcal{S}\cup\{0\}\}$ be the complementary set of masks (not including the zero mask). Then*

$$-\sqrt{\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)^2} \leq 2\,\mathrm{SD}(\varphi,\mathbf{1}_{qn}) - \mathrm{E}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{S}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \leq \sqrt{\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)^2}.$$

In particular, for $\mathcal{S} = \emptyset$, we obtain $\mathrm{SD}(\varphi,\mathbf{1}_{qn}) \leq \frac{1}{2}\sqrt{\mathrm{Var}[\varphi]}$

We state an additional basic result regarding variance.

Proposition 4 ([12], Proposition 6 – Variance of independent samples). *Let $\varphi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ be a density function. Let $u \geq 1$ be an integer and let $\varphi^{\times u} : \mathbb{F}_2^{(qu)\times n} \mapsto \mathbb{R}^{\geq 0}$ be the density function obtained by concatenating u independent samples drawn from φ. Then,*

$$\mathrm{Var}[\varphi^{\times u}] \leq 2u\,\mathrm{Var}[\varphi], \text{ assuming } u\,\mathrm{Var}[\varphi] \leq \tfrac{1}{2}.$$

2.3 Cryptographic Preliminaries

Standard definitions of adversary's advantage and the optimal advantage against a PRF (in the single-user and multi-user settings) are given in Appendix A.

Bounding the optimal advantage using Fourier analysis. In this paper we will consider keyed families of functions of the form $H : \mathcal{K} \times \{0,1\}^m \mapsto \{0,1\}^n$ with the property that the output distribution is independent of the queries of the adversary over $\{0,1\}^m$. Thus, we ignore these queries and focus on analyzing the output distribution (density function) generated by H. Given that the adversary makes q queries to H, we may denote the density function generated by H as $\varphi_H^{n,q} : \mathbb{F}_2^{q \times n} \to \mathbb{R}^{\geq 0}$.

By well-known properties of the statistical distance, the advantage of the optimal distinguisher against H is equal to the statistical distance of φ_H from uniform, namely,

$$\mathrm{Opt}_H^{\mathrm{prf}}(q) = \mathrm{SD}(\varphi_H^{n,q}, \mathbf{1}_{qn}). \tag{1}$$

In the multi-user setting, an adversary against H obtains a sample of $(\varphi_H^{n,q_{\max}})^{\times u} : \mathbb{F}_2^{(q_{\max} u) \times n} \mapsto \mathbb{R}^{\geq 0}$, where $(\varphi_H^{n,q_{\max}})^{\times u}$ is the density function obtained by concatenating u independent samples drawn from $\varphi_H^{n,q_{\max}}$. Here, $q_{\max}$ is the (maximal) number of queries per user. Similarly to the single-user setting, in the multi-user setting we have

$$\mathrm{Opt}_{H,u}^{\text{mu-prf}}(q_{max}) = \mathrm{SD}((\varphi_H^{n,q_{\max}})^{\times u}, \mathbf{1}_{u q_{\max} n}). \tag{2}$$

In this paper, we mostly bound the optimal advantage by bounding $\mathrm{Var}[\varphi_H^{n,q}]$ using the following basic result.

Proposition 5 (Bounds on advantage using variance). *Assume that the output distribution generated by $H : \mathcal{K} \times \{0,1\}^m \mapsto \{0,1\}^n$ is independent of the queries of the adversary. Denote by $\varphi_H^{n,q} : \mathbb{F}_2^{q \times n} \to \mathbb{R}^{\geq 0}$ the density function generated by H. Then,*

$$\mathrm{Opt}_H^{\mathrm{prf}}(q) \leq \frac{1}{2}\sqrt{\mathrm{Var}[\varphi_H^{n,q}]}, \text{ and } \mathrm{Opt}_{H,u}^{\text{mu-prf}}(q_{max}) \leq \frac{1}{\sqrt{2}}\sqrt{u\,\mathrm{Var}[\varphi_H^{n,q_{\max}}]},$$

assuming $u\,\mathrm{Var}[\varphi_H^{n,q_{\max}}] \leq \frac{1}{2}$, or equivalently, $\frac{1}{\sqrt{2}}\sqrt{u\,\mathrm{Var}[\varphi_H^{n,q_{\max}}]} \leq \frac{1}{2}$.

Symmetric properties. In addition to the output distribution being independent of the queries of the adversary, all the functions $H : \mathcal{K} \times \{0,1\}^m \mapsto \{0,1\}^n$ we analyze in this paper are symmetric in the following sense: if $x \sim \varphi_H^{n,q}$, then for every set of k distinct indices $\{i_1, i_2, \ldots, i_k\} \subseteq [q]$, $(x_{i_1}, \ldots, x_{i_k})$ are k elements that are marginally sampled from $\varphi_H^{n,k}$, namely, $(x_{i_1}, \ldots, x_{i_k}) \sim \varphi_H^{n,k}$.

Therefore, for $1 \le k \le q$, we have $\mathrm{M}^{=k}[\varphi_H^{n,q}] = \mathrm{M}^{=k}[\varphi_H^{n,k}]$ and

$$\mathrm{W}^{=k}[\varphi_H^{n,q}] = \sum_{\alpha \in \mathcal{M}^n_{=k,q}} \widehat{\varphi}_H^{n,q}(\alpha)^2 = \sum_{\{i_1,\ldots,i_k\} \subseteq [q] \text{ distinct}} \sum_{\substack{\beta \in \mathbb{F}_2^{k \times n} \\ \mathcal{NZ}_\beta = \{i_1,\ldots,i_k\}}} \widehat{\varphi}_H^{n,k}(\beta)^2$$
$$= \sum_{\{i_1,\ldots,i_k\} \subseteq [q] \text{ distinct}} \mathrm{W}^{=k}[\varphi_H^{n,k}] = \tbinom{q}{k} \mathrm{W}^{=k}[\varphi_H^{n,k}].$$

These symmetric properties are repeatedly used throughout the paper (often without explicitly referring to them). Another result on symmetric functions (which we do not explicitly use) is given in Appendix B.

Sampling without replacement. We define the density function of sampling without replacement.

Definition 5 (Density function of sampling without replacement). *For positive integers n, q such that $1 \le q \le 2^n$, let $\mu_{n,q} : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}^{\ge 0}$ be the density function associated with the process of uniformly sampling q elements from $\mathbb{F}_2^n$ without replacement. Specifically, for $x \in \mathbb{F}_2^{q \times n}$,*

$$\mu_{n,q}(x) = \begin{cases} \frac{N^q}{(N)_q} & \text{if } x_i \ne x_j \text{ for all } i,j \in [q] \ (i \ne j), \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, define $\mu_{n,0}$ to be the constant 1.

The SXoP$[r,n]$ construction. Let $\mathrm{Perm}(n)$ be the set of all permutations on $\{0,1\}^n$ (i.e., the set of all $\pi : \{0,1\}^n \mapsto \{0,1\}^n$). For positive integers r, n such that $r \ge 2$, define the family of functions $\mathrm{SXoP}[r,n] : (\mathrm{Perm}(n)) \times \{0,1\}^{n-\lceil \log r \rceil} \mapsto \{0,1\}^n$ by $\mathrm{SXoP}[r,n]_\pi(i) = \pi(0\|i) \oplus \pi(1\|i) \oplus \ldots \oplus \pi(r-1\|i)$, where in $\pi(j\|i)$, $j \in \{0,1\}^{\lceil \log r \rceil}$ is encoded in binary for $j = 0, \ldots, r-1$, and $\|$ denotes concatenation. We will be interested in bounding $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q)$ as a function of the parameters r, n, q (and deriving similar bounds in the multi-user setting). By symmetry of the randomly chosen permutation π, an adversary against $\mathrm{SXoP}[r,n]$ obtains the XOR of r samples, each containing q elements of $\{0,1\}^n$, where all rq elements are chosen uniformly without replacement (regardless of the actual queries).

Let $\nu^{(r)}_{n,q} : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}^{\ge 0}$ denote the density function of a sample generated by the $\mathrm{SXoP}[r,n]$ construction.

The LXoP$[L,n]$ and LXoP$[L,2,n]$ constructions. Let $L \in \mathbb{F}_2^{n \times n}$ be an invertible matrix. Define the family of functions $\mathrm{LXoP}[L,n] : (\mathrm{Perm}(n)) \times \{0,1\}^{n-1} \mapsto \{0,1\}^n$ by $\mathrm{LXoP}[L,n]_\pi(i) = \pi(0\|i) \oplus L(\pi(1\|i))$.

Moreover, define the family of functions $\mathrm{LXoP}[L,2,n] : (\mathrm{Perm}(n)) \times \{0,1\}^{n-2} \mapsto \{0,1\}^{2n}$ by $\mathrm{LXoP}[L,2,n]_\pi(i) = (\pi(0\|i) \oplus L(\pi(1\|i))) \,\|\, (\pi(1\|i) \oplus L(\pi(2\|i)))$.

We will be interested in bounding $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{LXoP}[L,n]}(q)$ and $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{LXoP}[L,2,n]}(q)$ as a function of the parameters n, q (and deriving similar bounds in the multi-user setting). As in the case of $\mathrm{SXoP}[r,n]$, the distributions generated by $\mathrm{LXoP}[L,n]$ and $\mathrm{LXoP}[L,2,n]$ are independent of the queries of the adversary. Let $\xi^{(L)}_{n,q} : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ and $\xi^{(L)}_{n,2,q} : \mathbb{F}_2^{q\times 2n} \mapsto \mathbb{R}^{\geq 0}$ denote the density functions of samples generated by the $\mathrm{LXoP}[L,n]$ and $\mathrm{LXoP}[L,2,n]$ constructions, respectively.

2.4 Fourier Properties of $\mu_{n,k}$

We list several results about Fourier properties of $\mu_{n,k}$, mostly taken from [12,14].

Proposition 6 ([12], Proposition 12 – Permuting elements preserves Fourier coefficients). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$. Let $\pi : [k] \mapsto [k]$ be a permutation and define $\alpha^\pi \in \widehat{\mathbb{F}}_2^{k\times n}$ by $(\alpha^\pi)_i = \alpha_{\pi(i)}$ for $i \in [k]$. Then, $\widehat{\mu}_{n,k}(\alpha^\pi) = \widehat{\mu}_{n,k}(\alpha)$.*

Proposition 6 is repeatedly (and implicitly) used throughout the paper.

Proposition 7. *For any $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ such that $\oplus_{i\in[k]}\alpha_i \neq 0$ we have $\widehat{\mu}_{n,k}(\alpha) = 0$.*

The following is a recursive formula for $\widehat{\mu}_{n,k}(\alpha)$ (proved in Appendix A).

Proposition 8 ([14], Section 4 – recursive formula for $\widehat{\mu}_{n,k}(\alpha)$). *For parameters $k_1 \geq k_0 \geq 2$, let $\alpha \in \widehat{\mathbb{F}}_2^{k_1\times n}$ have $\#\alpha = |\mathcal{NZ}_\alpha| = k_0$. Then for any $j \in \mathcal{NZ}_\alpha = \{i \in [k_1] \colon \alpha_i \neq 0\}$,*

$$\widehat{\mu}_{n,k_1}(\alpha) = -\frac{1}{N-k_0+1}\sum_{i\in\mathcal{NZ}_\alpha\setminus\{j\}} \widehat{\mu}_{n,k_1}(\alpha^{\oplus(j,i)}),$$

where $\alpha^{\oplus(j,i)} \in \widehat{\mathbb{F}}_2^{k_1\times n}$(for $i \neq j$) is defined as $(\alpha^{\oplus(j,i)})_\ell = \begin{cases} 0 & \text{if } \ell = j, \\ \alpha_i \oplus \alpha_j & \text{if } \ell = i, \\ \alpha_\ell & \text{if } \ell \notin \{i,j\}. \end{cases}$

Note that $\#\alpha^{\oplus(j,i)} = k_0 - 1$ if $\alpha_i \oplus \alpha_j \neq 0$ and $\#\alpha^{\oplus(j,i)} = k_0 - 2$ if $\alpha_i \oplus \alpha_j = 0$.

Proposition 9 (Recursive bound for $\widehat{\mu}_{n,k}(\alpha)^2$). *For parameters $k_1 \geq k_0 \geq 2$, let $\alpha \in \widehat{\mathbb{F}}_2^{k_1\times n}$ have $\#\alpha = |\mathcal{NZ}_\alpha| = k_0$. Then for any $j \in \mathcal{NZ}_\alpha = \{i \in [k_1] \colon \alpha_i \neq 0\}$,*

$$\widehat{\mu}_{n,k_1}(\alpha)^2 \leq \frac{k_0-1}{(N-k_0+1)^2}\sum_{i\in\mathcal{NZ}_\alpha\setminus\{j\}} \widehat{\mu}_{n,k_1}(\alpha^{\oplus(j,i)})^2.$$

Proof. By Proposition 8 and the Cauchy–Schwarz inequality,

$$\widehat{\mu}_{n,k_1}(\alpha)^2 = \left(-\tfrac{1}{N-k_0+1}\sum_{i\in\mathcal{NZ}_\alpha\setminus\{j\}} \widehat{\mu}_{n,k_1}(\alpha^{\oplus(j,i)})\right)^2 \leq \tfrac{k_0-1}{(N-k_0+1)^2}\sum_{i\in\mathcal{NZ}_\alpha\setminus\{j\}} \widehat{\mu}_{n,k_1}(\alpha^{\oplus(j,i)})^2.$$

■

Lemma 1 ([14], Lemma 4.1 – Bound on magnitude of level-k Fourier coefficients). *Let $k_1 \geq k_0$ and $0 \leq k_0 \leq N/2$. Then, $\mathrm{M}^{=k_0}[\mu_{n,k_1}] \leq \frac{1}{\sqrt{\binom{N}{k_0}}}$.*

A slightly stronger bound than above was also proved in Lemma 1 of [12], but we give the simpler proof of [14] in Appendix A.

Lemma 2 ([12], Lemma 2 – Bound on level-k Fourier weight). For $1 \leq k \leq N/2$, $\mathrm{W}^{=k}[\mu_{n,k}] \leq \left(\frac{k}{N-k}\right)^{k/2}$.

Proposition 10. *Let $k_1 \geq k_0$ and $2 \leq k_0 < N/2$ for k_0 even. Let $\alpha \in \widehat{\mathbb{F}}_2^{k_1 \times n}$ have $\#\alpha = k_0$. Assume that $\alpha_i = \alpha_j$ for all $i, j \in [k_1]$ such that $\alpha_i, \alpha_j \neq 0$ (i.e., $i, j \in \mathcal{NZ}_\alpha$). Then, $\widehat{\mu}_{n,k_1}(\alpha) = (-1)^{k_0/2} \frac{k_0-1}{N-1} \frac{k_0-3}{N-3} \cdots \frac{1}{N-(k_0-1)}$. Moreover, $\frac{1}{\sqrt{k_0 \binom{N}{k_0}}} \leq |\widehat{\mu}_{n,k_1}(\alpha)| \leq \frac{1}{\sqrt{\binom{N}{k_0}}}$.*

3 Framework for Bounding the Weight of $\mu_{n,k}$ on Structured Subsets

In this section we describe our framework. Unless stated otherwise, missing proofs are found in Appendix C. We begin with a motivating example.

Let $r \geq 2$ and $k \geq 1$ be parameters where $rk \leq O(N)$. Suppose we want to upper bound the expression $\sum_{\substack{\alpha \in \widehat{\mathbb{F}}_2^{k \times n} \\ \#\alpha = k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 = \sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2$, where $\alpha^{\odot r} = (\alpha_1^{\circ r}, \ldots, \alpha_k^{\circ r})$. As $\#\alpha^{\odot r} = r\#\alpha = rk$, apply Lemma 1 and obtain

$$\sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 \leq N^k (\mathrm{M}^{=rk}[\mu_{n,rk}])^2 \leq N^k \frac{1}{\binom{N}{rk}}. \tag{3}$$

Another option to bound the expression is to use Lemma 2 and deduce

$$\sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 \leq \sum_{\beta \in \mathcal{M}^n_{=rk,rk}} \widehat{\mu}_{n,rk}(\beta)^2 = \mathrm{W}^{=rk}[\mu_{n,rk}] \leq \left(\frac{rk}{N-rk}\right)^{rk/2}. \tag{4}$$

The above bounds are generally far from tight, as they make little use of the structure of masks we sum over. To improve the bounds, for every $\alpha \in \mathcal{M}^n_{=k,k}$ apply Proposition 9 to $\alpha^{\odot r}$ (with $k_1 = k_0 = rk$ and $j = rk$), obtaining

$$\begin{aligned} \sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 &\leq \frac{rk-1}{(N-rk+1)^2} \sum_{\alpha \in \mathcal{M}^n_{=k,k}} \sum_{i=1}^{rk-1} \widehat{\mu}_{n,rk}((\alpha^{\odot r})^{\oplus(rk,i)})^2 \\ &= \frac{rk-1}{(N-rk+1)^2} \sum_{i=1}^{rk-1} \sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}((\alpha^{\odot r})^{\oplus(rk,i)})^2. \end{aligned} \tag{5}$$

Fix $i \in [rk-1]$. We analyze the term $\sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}((\alpha^{\odot r})^{\oplus(rk,i)})^2$, assuming $r \geq 3$. As $(\alpha^{\odot r})^{\oplus(rk,i)}$ changes only entries rk and i of $\alpha^{\odot r}$, then given $i \in [rk-1]$,

$(\alpha^{\odot r})^{\oplus(rk,i)}$ fully determines α (and $\alpha^{\odot r}$). Indeed, for every $\ell \in [k]$, α_ℓ still appears in at least one entry of $(\alpha^{\odot r})^{\oplus(rk,i)}$. This property does not hold for $r = 2$, since for $i = 2k-1$, $(\alpha^{\odot 2})^{\oplus(2k,2k-1)}$ is independent of α_k.

In other words, given $i \in [rk-1]$ the i'th operation in Proposition 9 applied to $\alpha^{\odot r}$ (whose outcome is $(\alpha^{\odot r})^{\oplus(rk,i)}$) is invertible for $r \geq 3$. Partition all $\alpha \in \mathcal{M}^n_{=k,k}$ into two sets according to the non-zero index set of $(\alpha^{\odot r})^{\oplus(rk,i)}$ by defining $\mathcal{S}_{i,0} = \{\alpha \in \mathcal{M}^n_{=k,k} \colon \mathcal{NZ}_{(\alpha^{\odot r})^{\oplus(rk,i)}} = [rk]\backslash\{rk,i\}\}$ and $\mathcal{S}_{i,1} = \{\alpha \in \mathcal{M}^n_{=k,k} \colon \mathcal{NZ}_{(\alpha^{\odot r})^{\oplus(rk,i)}} = [rk]\backslash\{rk\}\}$. As $\#(\alpha^{\odot r})^{\oplus(rk,i)} \in \{rk-1, rk-2\}$,

$$\begin{aligned}
&\sum_{\alpha\in\mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}((\alpha^{\odot r})^{\oplus(rk,i)})^2 \\
&= \sum_{\alpha\in\mathcal{S}_{i,0}} \widehat{\mu}_{n,rk}((\alpha^{\odot r})^{\oplus(rk,i)})^2 + \sum_{\alpha\in\mathcal{S}_{i,1}} \widehat{\mu}_{n,rk}((\alpha^{\odot r})^{\oplus(rk,i)})^2 \\
&\leq \sum_{\beta\in\mathcal{M}^n_{=rk-2,rk-2}} \widehat{\mu}_{n,rk-2}(\beta)^2 + \sum_{\beta\in\mathcal{M}^n_{=rk-1,rk-1}} \widehat{\mu}_{n,rk-1}(\beta)^2 \\
&= \mathrm{W}^{=rk-2}[\mu_{n,rk-2}] + \mathrm{W}^{=rk-1}[\mu_{n,rk-1}].
\end{aligned}$$

The above inequality crucially uses two properties: (1) for each $i \in [rk-1]$, in the (multi) set $\{(\alpha^{\odot r})^{\oplus(rk,i)} \colon \alpha \in \mathcal{M}^n_{=k,k}\}$ each mask appears only once due to the invertibility of $(\alpha^{\odot r})^{\oplus(rk,i)}$, and (2) all masks in each of $\{(\alpha^{\odot r})^{\oplus(rk,i)} \colon \alpha \in \mathcal{S}_{i,0}\}$ and $\{(\alpha^{\odot r})^{\oplus(rk,i)} \colon \alpha \in \mathcal{S}_{i,1}\}$ have the same set of zero indices (which is trimmed). Thus, the right-hand side sums over the squared Fourier coefficients of a superset of the trimmed left-hand side masks. Combining with (5),

$$\begin{aligned}
&\sum_{\alpha\in\mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 \leq \tfrac{rk-1}{(N-rk+1)^2} \sum_{i=1}^{rk-1} (\mathrm{W}^{=rk-2}[\mu_{n,rk-2}] + \mathrm{W}^{=rk-1}[\mu_{n,rk-1}]) \\
&= (\tfrac{rk-1}{N-rk+1})^2 (\mathrm{W}^{=rk-2}[\mu_{n,rk-2}] + \mathrm{W}^{=rk-1}[\mu_{n,rk-1}]) \\
&\leq (\tfrac{rk-1}{N-rk+1})^2 ((\tfrac{rk-2}{N-rk+2})^{(rk-2)/2} + (\tfrac{rk-1}{N-rk+1})^{(rk-1)/2}) \leq 2(\tfrac{rk-1}{N-rk+1})^{(rk/2)+1},
\end{aligned}$$

where the penultimate inequality is by Lemma 2. Comparing this bound to (4), we get a significant improvement by a factor of about $2(\frac{rk}{N-rk})$. More generally, when the masks initially have level k_0, then the improvement over the straightforward application of Lemma 2 is by a factor of about $2(\frac{k_0}{N-k_0})$.

Recursion. We can obtain improved bounds by applying Proposition 9 recursively to each of the $2(rk-1)$ sets $\{(\alpha^{\odot r})^{\oplus(rk,i)} \colon \alpha \in \mathcal{S}_{i,j}\}_{i\in[rk-1],j\in\{0,1\}}$. The outcome is a recursion tree and we apply Lemma 2 only at the leaves. Next, we generalize the above analysis.

General Framework

We consider the following initial setting.

Setting 1 *Let $k' > 0$ be an integer parameter. Let $\mathcal{S}$ be a set of strings. Let $T : \mathcal{S} \to \widehat{\mathbb{F}}_2^{k' \times n}$ be a mapping such that the following two restrictions hold:*
(a1) *T is injective, i.e., for any $\alpha, \beta \in \mathcal{S}$ such that $\alpha \neq \beta$, $T(\alpha) \neq T(\beta)$, and*
(a2) *there is a common non-zero index subset $\mathcal{N} \subseteq [k']$ such that for every $\alpha \in \mathcal{S}$, $\mathcal{NZ}_{T(\alpha)} = \mathcal{N}$, i.e., for any $\ell \in [k']$, $T(\alpha)_\ell \neq 0$ if and only if $\ell \in \mathcal{N}$.*

The restrictions correspond to the two crucial properties that allow to apply Lemma 2. Specifically, restriction (**a2**) implies that all $\alpha \in \mathcal{S}$ have $\#T(\alpha) = |\mathcal{N}|$.

Our goal is to bound $\sum_{\alpha \in \mathcal{S}} \widehat{\mu}_{n,k'}(T(\alpha))^2$.[1] We start from the initial mask set $\{T(\alpha) \colon \alpha \in \mathcal{S}\}$, and invoke recursive calls of Proposition 9, where Lemma 2 is applied at the leaves of the recursion tree.

Let $\beta \in \widehat{\mathbb{F}}_2^{k' \times n}$ be a mask. Consider the operation $\beta^{\oplus(j,i)}$ for $j \in \mathcal{NZ}_\beta$ and $i \in \mathcal{NZ}_\beta \backslash \{j\}$. The formula of Proposition 9 applied to β includes $|\mathcal{NZ}_\beta| - 1$ such operations, where j is fixed and i ranges over all $\mathcal{NZ}_\beta \backslash \{j\}$. Thus, we call index j the *primary index*, while we call each $i \in \mathcal{NZ}_\beta \backslash \{j\}$ a *secondary index*.

Each recursive node v at depth $d \geq 0$ is labeled by a recursion stack which consists of the sequence of d secondary indices $i_1, \ldots, i_d \in [k']$ for the recursive calls up to this node, and a sequence of bits $b_1, \ldots, b_d \in \{0, 1\}$. For $d' \in [d]$, bit $b_{d'}$ specifies whether the outcome of the XOR operation at index $i_{d'}$ was zero or not. These bits keep track of the set $\mathcal{N}$ that evolves during the recursion.

We will assume that there is a *primary index selector*, or PIS, which is an application-dependent procedure that selects the next primary index (denoted j_{d+1}) for the invocation of Proposition 9. The input to the PIS includes the recursion stack $v = (i_1, \ldots, i_d, b_1, \ldots, b_d)$. Initially, the recursion stack is empty, and thus the first primary index is fixed by the PIS implementation.[2] We remark that the PIS also depends on the initial parameters of Setting 1, $(\mathcal{S}, T)$. However, $(\mathcal{S}, T)$ are assumed to be fixed and hardcoded inside the PIS.

Fixing a PIS implementation *pis*, we define a recursive procedure up to depth $d_{\max}$ (called $\mathrm{calcW}_{pis,d_{\max}}$) for upper bounding the weight $\sum_{\alpha \in \mathcal{S}} \widehat{\mu}_{n,k'}(T(\alpha))^2$.

Definition of calcW. The procedure $\mathrm{calcW}_{pis,d_{\max}}$ obtains 5 parameters:
(1) (current) recursion depth d,
(2) stack trace $v = (i_1, \ldots, i_d, b_1, \ldots, b_d)$,
(3) set $\mathcal{S}_v$,
(4) mapping $T_v : \mathcal{S}_v \to \widehat{\mathbb{F}}_2^{k' \times n}$, and
(5) set $\mathcal{N}_v \subset [k']$ such that for all $\alpha \in \mathcal{S}_v$, $\mathcal{NZ}_{T_v(\alpha)} = \mathcal{N}_v$ ($T_v(\alpha)_i \neq 0 \Leftrightarrow i \in \mathcal{N}_v$).

Initially, $\mathcal{S}, T$ are defined by Setting 1, and thus $d = 0$, $v = \mathrm{NULL}$, $\mathcal{S}_v = \mathcal{S}$, $T_v = T$ and $\mathcal{N}_v = \mathcal{N}$. In most (but not all) of our applications, $\mathcal{N}_v = [k']$, as the level of all masks $T(\alpha)$ for $\alpha \in \mathcal{S}$ will be k'.

[1] Note that $\sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2$ analyzed in the motivating example is a special case with $k' = rk$, $\mathcal{S} = \mathcal{M}^n_{=k,k}$ and $T(\alpha) = T_{r,k}(\alpha) = \alpha^{\odot r}$ (here $\mathcal{N} = [rk]$). Since (**a1**), (**a2**) hold we could apply Lemma 2 to derive (4).

[2] For example, for $T(\alpha) = \alpha^{\odot r}$ analyzed above we initially set $j_1 = rk$.

$\mathrm{calcW}_{pis,d_{\max}}(d, v = (i_1, \ldots, i_d, b_1, \ldots, b_d), \mathcal{S}_v, T_v, \mathcal{N}_v)$

1. $k'_v \leftarrow |\mathcal{N}_v|$.
2. If $d = d_{\max}$, return $(\frac{k'_v}{N-k'_v})^{k'_v/2}$.
3. $j \leftarrow pis(v)$.
4. $W \leftarrow 0$.
5. For all $i \in \mathcal{N}_v \backslash \{j\}$:
   (a) $v_{i,0} \leftarrow (i_1, \ldots, i_d, i, b_1, \ldots, b_d, 0)$, $v_{i,1} \leftarrow (i_1, \ldots, i_d, i, b_1, \ldots, b_d, 1)$.
   (b) Define $T_{v_i}(\alpha) = T_v(\alpha)^{\oplus(j,i)}$ for every $\alpha \in \mathcal{S}_v$.
   (c) $T_{v_{i,0}} \leftarrow T_{v_i}$, $T_{v_{i,1}} \leftarrow T_{v_i}$.
   (d) $\mathcal{S}_{v_{i,0}} \leftarrow \{\alpha \in \mathcal{S}_v \colon T_{v_i}(\alpha)_i = 0\}$, $\mathcal{S}_{v_{i,1}} \leftarrow \{\alpha \in \mathcal{S}_v \colon T_{v_i}(\alpha)_i \neq 0\}$.
   (e) $\mathcal{N}_{v_{i,0}} \leftarrow \mathcal{N}_v \backslash \{i, j\}$, $\mathcal{N}_{v_{i,1}} \leftarrow \mathcal{N}_v \backslash \{j\}$.
   (f) $W \leftarrow W + \mathrm{calcW}_{pis,d_{\max}}(d+1, v_{i,0}, \mathcal{S}_{v_{i,0}}, T_{v_{i,0}}, \mathcal{N}_{v_{i,0}})$,
   $W \leftarrow W + \mathrm{calcW}_{pis,d_{\max}}(d+1, v_{i,1}, \mathcal{S}_{v_{i,1}}, T_{v_{i,1}}, \mathcal{N}_{v_{i,1}})$.
6. Return $\frac{k'_v - 1}{(N-k'_v+1)^2} W$.

Thus, calcW implements the recursive invocation of Proposition 9, where Lemma 2 is applied at the leaves in the second step. As in the motivating example, for each $i \in \mathcal{N}_v \backslash \{j\}$ we need two recursive calls, since the non-zero index set of each mask $T_v(\alpha)^{\oplus(j,i)}$ can be either $\mathcal{N}_v \backslash \{j\}$ or $\mathcal{N}_v \backslash \{i, j\}$ (and this index set must be consistent in each call to calcW).

Remark 3. Assume that $\mathcal{S}, T$ and *pis* are fixed. Since the output of *pis* only depends on the recursion stack (but not on specific masks), the primary indices are uniquely defined by the recursion stack v, even though v does not include them explicitly. More generally, the 4 parameters $d, \mathcal{S}_v, T_v, \mathcal{N}_v$ of calcW are uniquely determined by v. The only reason we explicitly include them as parameters of calcW is to simplify its description.

Applicability of calcW. The correctness of calcW will rely on the assumption that the two restrictions of Setting 1 hold at all internal nodes, as they will be crucial for applying Lemma 2 at the leaves. Since for $b \in \{0, 1\}$ each set $\mathcal{S}_{v_{i,b}}$ is defined in correspondence with the non-zero index set $\mathcal{N}_{v_{i,b}}$, restriction (**a2**) indeed holds at all internal nodes. However, restriction (**a1**) may not hold recursively, and it requires special treatment depending on the specific application. We formalize the corresponding conditions in the following definition.

Definition 6 (Applicability of calcW). calcW *is applicable up to depth* $d_{\max}$ *with parameters* $(\mathcal{S}, T)$ *and a PIS pis, if the following conditions hold:*
*(**b1**) the pair* $(\mathcal{S}, T)$ *satisfies the restrictions of Setting 1 with* $2d_{\max} < |\mathcal{N}|$ *(*$\mathcal{N}$ *is defined in Setting 1), and*
*(**b2**) considering the recursion tree with root* $\mathrm{calcW}_{pis,d_{\max}}(0, (NULL), \mathcal{S}, T, \mathcal{N})$*: for every node* v *at depth at most* $d_{\max} - 1$ *such that* $j = pis(v)$*, for all* $i \in \mathcal{N}_v \backslash \{j\}$ *and* $\alpha \in \mathcal{S}_v$*,* $T_v(\alpha)$ *can be (uniquely) recovered from* $T_{v_i}(\alpha) = T_v(\alpha)^{\oplus(j,i)}$.

Restriction (**b2**) needs to holds only up to depth $d_{\max}-1$ as Proposition 9 is not used at the leaves. Before formally analyzing calcW, we simplify the second condition of Definition 6. This simplification will be useful in applications.

Proposition 11 (Sufficient condition for applicability of calcW). *Given $(\mathcal{S},T)$ and a PIS pis, assume that*
*(**c1**) the pair $(\mathcal{S},T)$ satisfies the restrictions of Setting 1 with $2d_{\max}<|\mathcal{N}|$ ($\mathcal{N}$ is defined in Setting 1), and*
*(**c2**) considering the recursion tree with root $\mathrm{calcW}_{pis,d_{\max}}(0,(NULL),\mathcal{S},T,\mathcal{N})$: for every node v at depth at most $d_{\max}-1$ such that $j=pis(v)$, for all $i\in\mathcal{N}_v\backslash\{j\}$ and $\alpha\in\mathcal{S}_v$, $T_v(\alpha)_j$ can be recovered from $T_{v_i}(\alpha)$.*
Then, $\mathrm{calcW}_{d_{\max}}$ is applicable up to depth $d_{\max}$ with $\mathcal{S},T$ and pis.

Proof. Condition (**b1**) of Definition 6 holds by assumption. We prove condition (**b2**). Fix a node v of depth at most $d_{\max}-1$ and let $\alpha\in\mathcal{S}_v$. According to Definition 6, we need to prove that $T_v(\alpha)$ can be recovered from $T_{v_i}(\alpha)=T_v(\alpha)^{\oplus(j,i)}$ (for all $i\in\mathcal{N}_v\backslash\{j\}$). Since only entries j and i are modified in $T_v(\alpha)$ by the mapping T_{v_i}, it is sufficient to prove that both $T_v(\alpha)_j$ and $T_v(\alpha)_i$ can be computed from $T_{v_i}(\alpha)$. By assumption, $T_v(\alpha)_j$ can be recovered from $T_{v_i}(\alpha)$. Moreover, since $T_{v_i}(\alpha)_i=(T_v(\alpha)^{\oplus(j,i)})_i=T_v(\alpha)_i\oplus T_v(\alpha)_j$, then $T_v(\alpha)_i=T_{v_i}(\alpha)_i\oplus T_v(\alpha)_j$ can also be recovered from $T_{v_i}(\alpha)$. Hence both conditions of Definition 6 hold. ■

The following definition will be useful in applications.

Definition 7 (Unaltered index). *An index $\ell'\in[k']$ is called unaltered at a node $v=(i_1,\ldots,i_d,b_1,\ldots,b_d)$ if ℓ' has not been selected as primary or secondary index. Namely, $\ell'\neq j_{d'}$ and $\ell'\neq i_{d'}$ for all $d'\in[d]$.*

The definition is motivated by the simple property that if ℓ' is unaltered at node v, then for any $\alpha\in\mathcal{S}_v$, $T_v(\alpha)_{\ell'}=T(\alpha)_{\ell'}$ (where T is the initial mapping at the root). This property holds since the mappings T_{v_i} at any node v only modify the entries of the primary index j and secondary index i.

Denote by $\mathcal{U}_v$ the set of all unaltered indices at node v. At the root node v, $\mathcal{U}_v=[k']$. Since every child of any node v has one primary and one secondary index, a node at depth d has $|\mathcal{U}_v|\geq k'-2d$.

Analysis of calcW. The main result regarding calcW is given below.

Lemma 3. *Assume that* calcW *is applicable up to depth $d=d_{\max}\geq 0$ with parameters $(\mathcal{S},T)$ (where $T:\mathcal{S}\to\widehat{\mathbb{F}}_2^{k'\times n}$) and a PIS, pis. Assume further that initially $\#T(\alpha)=k_0\leq k'$ satisfies $2d<k_0\leq N/8$ for all $\alpha\in\mathcal{S}$. Then,*

$$\sum_{\alpha\in\mathcal{S}}\widehat{\mu}_{n,k'}(T(\alpha))^2\leq 2^d\frac{(k_0)^{2d}(k_0-2d)^{k_0/2-d}}{(N-k_0)^{k_0/2+d}}\leq 2^d\left(\frac{k_0}{N-k_0}\right)^{k_0/2+d}.$$

The proof of Lemma 3 is given in Appendix C. Note that the improvement over the naive application of Lemma 2 is by a factor of about $2^d(\frac{k_0}{N-k_0})^d$. This emphasizes the importance of defining a PIS that allows applying calcW up to a large depth d. Appendix C also describes possible variants of calcW.

4 Indistinguishability Upper and Lower Bounds for SXoP$[r, n]$

In this section, we analyze the SXoP$[r, n]$ construction, proving the main theorem below. Unless stated otherwise, missing proofs are found in Appendix D.

Theorem 1. *Assume that $rq \le N/8$ and $N \ge 2^{13}r$. The following bounds (depending on r) hold.*

$$\underline{\text{Odd } r \ge 3.} \qquad \begin{aligned} \mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) &\le 2^{r-1}r^r\frac{q}{N^{r-0.5}} \le O_r\left(\frac{q}{N^{r-0.5}}\right), \text{ and} \\ \mathrm{Opt}^{\text{mu-prf}}_{\mathrm{SXoP}[r,n],u}(q_{max}) &\le 2^{r-0.5}r^r\frac{\sqrt{u}q_{\max}}{N^{r-0.5}} \le O_r\left(\frac{\sqrt{u}q_{\max}}{N^{r-0.5}}\right), \end{aligned} \tag{6}$$

where the second inequality also requires $2^{r-0.5}r^r\frac{\sqrt{u}q_{\max}}{N^{r-0.5}} \le \frac{1}{2}$.

$$\underline{r = 2.} \qquad \mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[2,n]}(q) \le \frac{5q}{N} \le O\left(\frac{q}{N}\right). \tag{7}$$

$$\underline{\text{Even } r \ge 4.} \qquad \begin{aligned} \mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) &\le 2r^{r/2}\frac{q}{N^{r/2}} \le O_r\left(\frac{q}{N^{r/2}}\right), \text{ and} \\ \mathrm{Opt}^{\text{mu-prf}}_{\mathrm{SXoP}[r,n],u}(q_{max}) &\le \min\left(r^{r/2}\frac{\sqrt{uq_{\max}}}{N^{r/2-1/2}}, 2r^{r/2}\frac{uq_{\max}}{N^{r/2}}\right) \\ &\le \min\left(O_r\left(\frac{\sqrt{uq_{\max}}}{N^{r/2-1/2}}\right), O_r\left(\frac{uq_{\max}}{N^{r/2}}\right)\right), \end{aligned} \tag{8}$$

where the first bound on $\mathrm{Opt}^{\text{mu-prf}}_{\mathrm{SXoP}[r,n],u}(q_{max})$ also requires $r^{r/2}\frac{\sqrt{uq_{\max}}}{N^{r/2-1/2}} \le \frac{1}{2}$.

$$\underline{\text{Lower bound for even } r \ge 4.} \ \mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) \ge 2^{-1}e^{-r/2}r^{(r-1)/2}\frac{q}{N^{r/2}} \ge \Omega_r\left(\frac{q}{N^{r/2}}\right). \tag{9}$$

Note that for $r \ge 4$, the theorem proves matching upper and lower single-user bounds of $\Theta_r\left(\frac{q}{N^{r/2}}\right)$. The bound for $r = 2$ and both bounds for odd $r \ge 3$ are tight, as they are matched by attacks described in previous works.

The proof relies on the following three lemmas (proved in the remainder of this section) regarding the density function $\nu^{(r)}_{n,k}$, generated by SXoP$[r, n]$.

Lemma 4 ($\mathcal{L}^1$ bidirectional bounds on $\widehat{\nu}^{(r)}_{n,k}$ for even r). *Assuming $k \le N/4$ and r is even, $\frac{3k}{2\sqrt{r\binom{N}{r}}} \le \mathrm{E}_{x\sim\mathbb{F}_2^{k\times n}}\left|\sum_{\alpha\in\mathcal{M}^n_{=1,k}}\widehat{\nu}^{(r)}_{n,k}(\alpha)\chi_\alpha(x)\right| \le \frac{2k}{\sqrt{\binom{N}{r}}}$.*

Lemma 5 (Variance and weight bounds for $\nu^{(2)}_{n,q}$). *Assume that $N \ge 100$ and $q \le N/16$. Then, $\sum_{k=2}^{q}\mathrm{W}^{=k}[\nu^{(2)}_{n,q}] \le \frac{18q^2}{N^2}$, and $\mathrm{Var}[\nu^{(2)}_{n,q}] \le \frac{4q}{N}$.*

Lemma 6 (Variance and weight bounds for $\nu^{(r)}_{n,q}$ with $r \ge 3$). *Assume that $N \ge 2^{13}r$, $rq \le N/8$. Then, for odd $r \ge 3$, $\mathrm{Var}[\nu^{(r)}_{n,q}] \le 2^{2r}r^{2r}\frac{q^2}{N^{2r-1}}$. For even $r \ge 4$, $\sum_{k=2}^{q}\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \le 2^{2r+1}r^{2r}\frac{q^2}{N^{2r-2}}$, and $\mathrm{Var}[\nu^{(r)}_{n,q}] \le 2r^r\frac{q}{N^{r-1}}$.*

Proof overview of Theorem 1. The proof of Theorem 1 is in Appendix D. Most upper bounds follow directly from the variance bounds of Lemma 5 and Lemma 6, combined with Proposition 5. The more interesting proofs for even r use Lemma 4 with Proposition 3, as summarized below for $r \geq 4$.

We use Proposition 3 with $\mathcal{S} = \mathcal{M}^n_{=1,q} = \{\alpha \in \widehat{\mathbb{F}}_2^{q\times n} \colon \#\alpha = 1\}$. Thus, combining (1) in Section 2 and Proposition 3 we obtain $2\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) \leq \mathrm{E}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)| + \sqrt{\sum_{\alpha\in\mathcal{M}^n_{\geq 2,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)^2}$. By Lemma 4, the first term is bounded by $\frac{2q}{\sqrt{\binom{N}{r}}} \leq O_r(\frac{q}{N^{r/2}})$. By Lemma 6 the second term is bounded by $\sqrt{2^{2r+1}r^{2r}\frac{q^2}{N^{2r-2}}} \leq O_r(\frac{q}{N^{r-1}})$. Summing up the terms (noting that $r/2 < r-1$ as $r \geq 4$), we conclude $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) \leq O_r(\frac{q}{N^{r/2}})$, asymptotically proving the first inequality of (8).

For the other direction, by (1) and Proposition 3, $2\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) \geq \mathrm{E}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)| - \sqrt{\sum_{\alpha\in\mathcal{M}^n_{\geq 2,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)^2}$. By Lemma 4, the first term is lower bounded as $\frac{3q}{2\sqrt{r\binom{N}{r}}} \geq \Omega_r(\frac{q}{N^{r/2}})$. We have already upper bounded the second term above by $O_r(\frac{q}{N^{r-1}})$, and thus the first term dominates the second. This implies $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) \geq \Omega_r(\frac{q}{N^{r/2}})$, asymptotically proving (9).

4.1 Relation Between $\widehat{\nu}^{(r)}_{n,k}$ and $\widehat{\mu}_{n,rk}$

We first establish the connection between $\widehat{\nu}^{(r)}_{n,k}$ and $\widehat{\mu}_{n,rk}$.

Proposition 12 (Relation between $\widehat{\nu}^{(r)}_{n,k}$ and $\widehat{\mu}_{n,rk}$). *For any $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$, $\widehat{\nu}^{(r)}_{n,k}(\alpha) = \widehat{\mu}_{n,rk}(\alpha^{\circ r}) = \widehat{\mu}_{n,rk}(\alpha^{\odot r})$, where $\widehat{\mu}_{n,rk}(\alpha^{\odot r}) = \widehat{\mu}_{n,rk}(\alpha_1^{\circ r}, \ldots, \alpha_k^{\circ r})$.*

Proof. By definition of SXoP$[r,n]$ and Proposition 1, for any $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$

$$\begin{aligned}
\widehat{\nu}^{(r)}_{n,k}(\alpha) &= \mathop{\mathrm{E}}_{x\sim\nu^{(r)}_{n,k}}[\chi_\alpha(x)] = \mathop{\mathrm{E}}_{y^{1..r}_{1..k}\sim\mu_{n,rk}}[\chi_\alpha(\bigoplus_{\ell=1}^{r} y_1^{(\ell)}, \ldots, \bigoplus_{\ell=1}^{r} y_k^{(\ell)})] \\
&= \mathop{\mathrm{E}}_{y^{1..r}_{1..k}\sim\mu_{n,rk}}[\prod_{\substack{i\in[k]\\ \ell\in[r]}} \chi_{\alpha_i}(y_i^{(\ell)})] = \mathop{\mathrm{E}}_{y^{1..r}_{1..k}\sim\mu_{n,rk}}[\prod_{\ell\in[r]} \chi_\alpha(y_{1..k}^{(\ell)})] \\
&= \mathop{\mathrm{E}}_{y^{1..r}_{1..k}\sim\mu_{n,rk}}[\chi_{\alpha^{\circ r}}(y_{1..k}^{(1)}, \ldots, y_{1..k}^{(r)})] = \widehat{\mu}_{n,rk}(\alpha^{\circ r}).
\end{aligned}$$

Finally, $\widehat{\mu}_{n,rk}(\alpha^{\circ r}) = \widehat{\mu}_{n,rk}(\alpha^{\odot r})$ holds by Proposition 6. ■

Lemma 5 is proved in Appendix D. It is based on bounds similar to (3) and (4), proved in the motivating example of Section 3.

4.2 Proof of Lemma 4 and Optimal Adversary

Proof (of Lemma 4). Let $\beta \in \widehat{\mathbb{F}}_2^{k\times n}$ be any fixed mask with $\#\beta = 1$. Also, let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ be a mask with $\#\alpha = 1$. Observe that $\alpha^{\circ r}$ has $\#(\alpha^{\circ r}) = r$ and all its r non-zero elements are equal. Since r is even, by Proposition 10, $\widehat{\mu}_{n,rk}(\alpha^{\circ r}) = \widehat{\mu}_{n,rk}(\beta^{\circ r})$ is independent of the actual non-zero element. Fixing $x \in \mathbb{F}_2^{k\times n}$, and applying Proposition 12 and Proposition 10,

$$\sum_{\alpha\in\mathcal{M}^n_{=1,k}} \widehat{\nu}^{(r)}_{n,k}(\alpha)\chi_\alpha(x) = \sum_{\alpha\in\mathcal{M}^n_{=1,k}} \widehat{\mu}_{n,rk}(\alpha^{\circ r})\chi_\alpha(x) = \widehat{\mu}_{n,rk}(\beta^{\circ r}) \sum_{\alpha\in\mathcal{M}^n_{=1,k}} \chi_\alpha(x). \tag{10}$$

For $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ with $\#\alpha = 1$, let $in(\alpha)$ be the unique index with $\alpha_{in(\alpha)} \neq 0$. Then,

$$\sum_{\alpha\in\mathcal{M}^n_{=1,k}} \chi_\alpha(x) = \sum_{\alpha\in\mathcal{M}^n_{=1,k}} \prod_{i\in[k]} \chi_{\alpha_i}(x_i) = \sum_{\alpha\in\mathcal{M}^n_{=1,k}} \chi_{\alpha_{in(\alpha)}}(x_{in(\alpha)}) = \sum_{i=1}^{k} \sum_{\substack{\gamma\in\widehat{\mathbb{F}}_2^n \\ \gamma\neq 0}} \chi_\gamma(x_i)$$

$$= \sum_{i=1}^{k}\Big(\sum_{\gamma\in\widehat{\mathbb{F}}_2^n} \chi_\gamma(x_i) - \chi_0(x_i)\Big) = N\sum_{i=1}^{k}\Big(\mathop{\mathrm{E}}_{\gamma\sim\widehat{\mathbb{F}}_2^n}[\chi_\gamma(x_i)]\Big) - k = N\sum_{i=1}^{k}(\mathbb{1}(x_i = 0)) - k$$

$$= N|\{i \in [k] \colon x_i = 0\}| - k = N \cdot Z_x - k,$$

where the sixth equality is by orthogonality of the characters (as $\chi_\gamma(x_i) = \chi_{x_i}(\gamma) = \chi_{x_i}(\gamma)\chi_0(\gamma)$), and $Z_x = |\{i \in [k] \colon x_i = 0\}|$. For $x \sim \mathbb{F}_2^{k\times n}$, the random variable Z_x is binomially distributed with number of experiments k and success probability $\frac{1}{N}$, and thus satisfies $\mathrm{E}[Z_x] = \frac{k}{N}$. Hence, $N \cdot Z_x - k = N(Z_x - \mathrm{E}[Z_x])$. Combining with (10),

$$\sum_{\alpha\in\mathcal{M}^n_{=1,k}} \widehat{\nu}^{(r)}_{n,k}(\alpha)\chi_\alpha(x) = N \cdot \widehat{\mu}_{n,rk}(\beta^{\circ r})(Z_x - \mathrm{E}[Z_x]). \tag{11}$$

Hence, the expression we wish to bound satisfies

$$\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}} \Big|\sum_{\alpha\in\mathcal{M}^n_{=1,k}} \widehat{\nu}^{(r)}_{n,k}(\alpha)\chi_\alpha(x)\Big| = N \cdot |\widehat{\mu}_{n,rk}(\beta^{\circ r})| \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}} |Z_x - \mathrm{E}[Z_x]|.$$

By Proposition 10, $\frac{1}{\sqrt{r\binom{N}{r}}} \leq |\widehat{\mu}_{n,rk}(\beta^{\circ r})| \leq \frac{1}{\sqrt{\binom{N}{r}}}$. It remains to prove that $\frac{3k}{2N} \leq \mathrm{E}_x |Z_x - \mathrm{E}[Z_x]| \leq \frac{2k}{N}$. Since $\mathrm{E}[Z_x] = \frac{k}{N}$ satisfies $0 < \mathrm{E}[Z_x] < 1$,

$$\begin{aligned}
&\mathop{\mathrm{E}}_x |Z_x - \mathrm{E}[Z_x]| = \mathop{\mathrm{E}}_x[(\mathrm{E}[Z_x] - Z_x)(\mathbb{1}(\mathrm{E}[Z_x] > Z_x) - \mathbb{1}(Z_x > \mathrm{E}[Z_x]))] \\
&= \mathop{\mathrm{E}}_x[(\mathrm{E}[Z_x] - Z_x)(\mathbb{1}(Z_x = 0) - \mathbb{1}(Z_x > 0))] \\
&= \mathop{\mathrm{E}}_x[(\mathrm{E}[Z_x] - Z_x)(2 \cdot \mathbb{1}(Z_x = 0) - (\mathbb{1}(Z_x = 0) + \mathbb{1}(Z_x > 0)))] \\
&= 2\mathop{\mathrm{E}}_x[(\mathrm{E}[Z_x] - Z_x)\mathbb{1}(Z_x = 0)] - \mathop{\mathrm{E}}_x[\mathrm{E}[Z_x] - Z_x] = 2\tfrac{k}{N}\Pr[Z_x = 0].
\end{aligned} \tag{12}$$

Finally, $\Pr[Z_x = 0] = (1 - \frac{1}{N})^k \leq 1$, and as $k \leq \frac{N}{4}$, $(1 - \frac{1}{N})^k \geq (1 - \frac{1}{N})^{N/4} \geq \frac{3}{4}$. ■

Optimal adversary for SXoP$[r,n]$ with even $r \geq 4$. For $x \in \mathbb{F}_2^{q\times n}$, let $Z_x = |\{i \in [q] \colon x_i = 0\}|$. Define the adversary $A(x) = \mathbb{1}(Z_x = 0)$. We argue that A is optimal. Similarly to the proof of (9) in Theorem 1, its advantage is

$$|\operatorname*{E}_x[(\nu^{(r)}_{n,q}(x) - 1)\mathbb{1}(Z_x = 0)]| = |\operatorname*{E}_x[(\sum_{\alpha\neq 0} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x))\mathbb{1}(Z_x = 0)]|$$

$$\geq |\operatorname*{E}_x[(\sum\nolimits_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x))\mathbb{1}(Z_x = 0)]| - \sqrt{\sum\nolimits_{\alpha\in\mathcal{M}^n_{\geq 2,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)^2}.$$

By (11) in the proof of Lemma 4, there is some C independent of x such that $\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x) = C(Z_x - \mathrm{E}[Z_x])$, and the first term is

$$|C\cdot\operatorname*{E}_x[(Z_x - \mathrm{E}[Z_x])\mathbb{1}(Z_x = 0)]| = \tfrac{1}{2}\operatorname*{E}_x|C(Z_x - \mathrm{E}[Z_x])| = \tfrac{1}{2}\operatorname*{E}_x|\sum\nolimits_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)|,$$

where the two equalities are by (12) and the definition of C above. Thus, ignoring the second term, the advantage of A is lower bounded by $\frac{1}{2}\operatorname{E}_x|\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)|$, which dominates the optimal advantage by the proof of Theorem 1 (the second term is negligible). Hence, A is optimal up to a negligible factor.

4.3 Application of Main Framework and Proof of Lemma 6

We apply our main framework and use it to prove Lemma 6.

Proposition 13. *Assume that $rk \leq N/8$. Define $c_{rk} = 0$ if rk is even and $c_{rk} = \frac{1}{2}$ if rk is odd (i.e., $c_{rk} = \frac{rk \bmod 2}{2}$). Then, for any $r \geq 3$*

$$\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \leq \binom{q}{k} 2^{(r-2)k/2 + c_{rk}} \left(\frac{rk}{N - rk}\right)^{(r-1)k + c_{rk}}.$$

Proof. Applying Proposition 12,

$$\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] = \tbinom{q}{k}\mathrm{W}^{=k}[\nu^{(r)}_{n,k}] = \tbinom{q}{k}\sum_{\alpha\in\mathcal{M}^n_{=k,k}} \widehat{\nu}^{(r)}_{n,k}(\alpha)^2 = \tbinom{q}{k}\sum_{\alpha\in\mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2. \tag{13}$$

We would like to bound $\sum_{\alpha\in\mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2$ using Lemma 3. We start by introducing several definitions referring to Setting 1 and then define the PIS, *pis*. First, define $\mathcal{S} = \mathcal{M}^n_{=k,k} = \{\alpha \in \widehat{\mathbb{F}}_2^{k\times n} \colon \#\alpha = k\}$, and $T(\alpha) = T_{r,k}(\alpha) = \alpha^{\odot r} = ((\alpha_1)^{\circ r}, \ldots, (\alpha_k)^{\circ r})$ (using notation of Lemma 3, here $k_0 = k' = rk$).

Given an index $\ell \in [k]$, for all $\ell' \in [(\ell-1)r + 1, \ell r]$, $T(\alpha)_{\ell'} = (\alpha^{\odot r})_{\ell'} = \alpha_\ell$. Thus, for a recursion node v, if $\ell' \in [(\ell-1)r+1, \ell r]$ is unaltered by Definition 7, then $T_v(\alpha)_{\ell'} = T(\alpha)_{\ell'} = \alpha_\ell$ for every $\alpha \in \mathcal{S}_v$. We call an index $\ell \in [k]$ redundant (for a node v) if at least 3 of the r indices in $[(\ell-1)r + 1, \ell r]$ are unaltered.

Given a recursion node v, let $\ell \in [k]$ be the largest redundant index. The PIS *pis* selects as primary index the largest unaltered index $j \in [(\ell-1)r + 1, \ell r]$.

Let $d = d_{\max} = \lceil \frac{(r-2)k}{2} \rceil$. We first prove that there is always a redundant index for nodes up to depth $d-1 = \lceil \frac{(r-2)k}{2} \rceil - 1$.

Every recursive call can remove at most 2 unaltered indices. Thus, the number of unaltered indices of each node at depth $d-1$ is at least $k' - 2(d-1) = rk - 2\lceil \frac{(r-2)k}{2} \rceil + 2 \geq rk - ((r-2)k+1) + 2 = 2k+1$. By an averaging argument, there exists $\ell \in [k]$ such that $[(\ell-1)r+1, \ell r]$ contains at least $\lceil \frac{2k+1}{k} \rceil = 3$ unaltered indices. Namely, ℓ is redundant. This proves that pis is well-defined up to depth $d-1 = \lceil \frac{(r-2)k}{2} \rceil - 1$ (at the leaves of depth d we do not invoke pis).

In order to apply Lemma 3, it is sufficient to prove that the two conditions of Proposition 11 hold. Clearly, the pair $(\mathcal{S}, T)$ satisfies the restrictions of Setting 1, and condition (**c1**) holds (note that $2d = 2\lceil \frac{(r-2)k}{2} \rceil \leq (r-2)k + 1 < rk = k_0$).

We now prove condition (**c2**). Specifically, we prove that for a node v such that $j = pis(v)$ and $\alpha \in \mathcal{S}_v$, $T_v(\alpha)_j$ can be computed from $T_{v_i}(\alpha) = T_v(\alpha)^{\oplus(j,i)}$ (where i is a secondary index).

For a node v we select a primary index $j \in [rk]$ such that $T_v(\alpha)_j = \alpha_\ell$ and since ℓ is redundant, $T_v(\alpha)_{\ell'} = \alpha_\ell$ for at least 3 indices $\ell' \in [(\ell-1)r+1, \ell r]$. As $T_{v_i}(\alpha) = T_v(\alpha)^{\oplus(j,i)}$, and $T_v(\alpha)^{\oplus(j,i)}$ modifies 2 entries of $T_v(\alpha)$, then $\alpha_\ell = T_v(\alpha)_j$ still appears at least $3-2=1$ time in $T_{v_i}(\alpha)$. This proves condition (**c2**) as required.

Applying our framework of Lemma 3 (with $d = \lceil \frac{(r-2)k}{2} \rceil = \frac{(r-2)k}{2} + c_{rk}$, $k_0 = k' = rk$), we obtain

$$\sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 \leq 2^d \left(\tfrac{k'}{N-k'}\right)^{k'/2+d} = 2^{(r-2)k/2+c_{rk}} \left(\tfrac{rk}{N-rk}\right)^{(r-1)k+c_{rk}}.$$

Combining with (13) completes the proof. ■

The proof of Lemma 6 (given in Appendix D) uses the bounds on $\mathrm{W}^{=k}[\nu^{(r)}_{n,q}]$ of Proposition 13 and additional simple bounds to analyze several sums of weights. It essentially shows that in all cases the lowest-level weight bound dominates the sum (except for r odd, as for $k=1$, $\mathrm{W}^{=1}[\nu^{(r)}_{n,q}] = 0$).

5 Indistinguishability Bounds for LXoP$[L, n]$

In this section, we state and prove our main theorem regarding LXoP$[L, n]$. Missing proofs are given in Appendix E.

Theorem 2. *Assume that the function $L'(x) = x \oplus L(x)$ is a permutation on $\mathbb{F}_2^n$. Given that $N \geq 2^{10}$ and $q \leq N/16$, $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{LXoP}[L,n]}(q) \leq \frac{4q}{N^{1.5}}$.*

Moreover, assuming $\frac{6\sqrt{u}q_{\max}}{N^{1.5}} \leq \frac{1}{2}$, $\mathrm{Opt}^{\text{mu-prf}}_{\mathrm{LXoP}[L,n],u}(q_{max}) \leq \frac{6\sqrt{u}q_{\max}}{N^{1.5}}$.

The proof uses the following lemma, proved in the remainder of this section.

Lemma 7. *Assume that the function $L'(x) = x \oplus L(x)$ is a permutation on $\mathbb{F}_2^n$. Given that $N \geq 2^{10}$ and $q \leq N/16$, $\mathrm{Var}[\xi^{(L)}_{n,q}] \leq \frac{64q^2}{N^3}$.*

Proof (of Theorem 2). Immediate from Lemma 7 and Proposition 5. ■

5.1 Elementary Results

We establish the connection between the Fourier coefficients of $\xi_{n,k}^{(L)}$ and $\mu_{n,2k}$.

Proposition 14 (Relation between $\widehat{\xi}_{n,k}^{(L)}$ and $\widehat{\mu}_{n,2k}$). *For any $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$,*

$$\widehat{\xi}_{n,k}^{(L)}(\alpha) = \widehat{\mu}_{n,2k}(\alpha, L^{\mathrm{T}}(\alpha)) = \widehat{\mu}_{n,2k}(\alpha_1, L^{\mathrm{T}}(\alpha_1), \ldots, \alpha_k, L^{\mathrm{T}}(\alpha_k)).$$

Proof. By definition of LXoP$[L,n]$ and Proposition 1, for any $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$

$$\widehat{\xi}_{n,k}^{(L)}(\alpha) = \mathop{\mathrm{E}}_{x\sim\xi_{n,k}^{(L)}}[\chi_\alpha(x)] = \mathop{\mathrm{E}}_{y_{1..k}^{1,2}\sim\mu_{n,2k}}[\chi_\alpha(y_1^{(1)} \oplus L(y_1^{(2)}), \ldots, y_k^{(1)} \oplus L(y_k^{(2)}))]$$

$$= \mathop{\mathrm{E}}_{y_{1..k}^{1,2}\sim\mu_{n,2k}}[\chi_{\alpha,\alpha}(y_{1..k}^{(1)}, L(y_{1..k}^{(2)}))] = \mathop{\mathrm{E}}_{y_{1..k}^{1,2}\sim\mu_{n,2k}}[\chi_{\alpha,L^{\mathrm{T}}(\alpha)}(y_{1..k}^{(1)}, y_{1..k}^{(2)})] = \widehat{\mu}_{n,2k}(\alpha, L^{\mathrm{T}}(\alpha)).$$

■

Proposition 15. *Assuming that $L'(x) = x \oplus L(x)$ is a permutation on $\mathbb{F}_2^n$, then* $\mathrm{W}^{=1}[\xi_{n,q}^{(L)}] = 0$.

Proof. By Proposition 14,

$$\mathrm{W}^{=1}[\xi_{n,q}^{(L)}] = \tbinom{q}{1} \sum\nolimits_{\alpha\in\mathcal{M}_{=1,1}^n} \widehat{\xi}_{n,1}^{(L)}(\alpha)^2 = q \sum\nolimits_{\alpha\in\mathcal{M}_{=1,1}^n} \widehat{\mu}_{n,2}(\alpha, L^{\mathrm{T}}(\alpha))^2.$$

For $\alpha \in \widehat{\mathbb{F}}_2^n \neq 0$ we have $\alpha \oplus L^{\mathrm{T}}(\alpha) = (L')^{\mathrm{T}}(\alpha)$. Since L' is a permutation, so is $(L')^{\mathrm{T}}$. Since $(L')^{\mathrm{T}}(0) = 0$, this implies that $(L')^{\mathrm{T}}(\alpha) \neq 0$, hence $\alpha \oplus L^{\mathrm{T}}(\alpha) \neq 0$. By proposition 7, we deduce $\widehat{\mu}_{n,2}(\alpha, L^{\mathrm{T}}(\alpha)) = 0$, implying $\mathrm{W}^{=1}[\xi_{n,q}^{(L)}] = 0$. ■

5.2 Application of Main Framework and Proof of Lemma 7

We use our main framework to prove Lemma 7.

Proposition 16. *Assume that the function $L'(x) = x \oplus L(x)$ is a permutation on $\mathbb{F}_2^n$, and assume that $2 \le k \le N/16$. Define $c_k = 0$ if k is even and $c_k = \frac{1}{2}$ otherwise (i.e. $c_k = \frac{k \bmod 2}{2}$). Then,*

$$\mathrm{W}^{=k}[\xi_{n,q}^{(L)}] \le \binom{q}{k} 2^{3k/2+3c_k} \frac{k^{k+2c_k}(k-2c_k)^{k/2-c_k}}{(N-2k)^{3k/2+c_k}}.$$

Proof. Based on Proposition 14,

$$\mathrm{W}^{=k}[\xi_{n,q}^{(L)}] = \tbinom{q}{k} \sum_{\alpha\in\mathcal{M}_{=k,k}^n} \widehat{\xi}_{n,k}^{(L)}(\alpha)^2 = \tbinom{q}{k} \sum_{\alpha\in\mathcal{M}_{=k,k}^n} \widehat{\mu}_{n,2k}(\alpha, L^{\mathrm{T}}(\alpha))^2. \tag{14}$$

We now use Lemma 3 to upper bound $\sum_{\alpha\in\mathcal{M}_{=k,k}^n} \widehat{\mu}_{n,2k}(\alpha, L^{\mathrm{T}}(\alpha))^2 = \sum_{\alpha\in\mathcal{M}_{=k,k}^n} \widehat{\mu}_{n,2k}(\alpha_1, L^{\mathrm{T}}(\alpha_1), \ldots, \alpha_k, L^{\mathrm{T}}(\alpha_k))^2$.

We start by introducing some definitions referring to Setting 1 and then define the PIS *pis*. Define $\mathcal{S} = \mathcal{M}^n_{=k,k} = \{\alpha \in \widehat{\mathbb{F}}_2^{k\times n} \colon \#\alpha = k\}$, and $T(\alpha) = T_k(\alpha) = (\alpha_1, L^{\mathrm{T}}(\alpha_1), \ldots, \alpha_k, L^{\mathrm{T}}(\alpha_k))$ (using notation of Lemma 3, $k_0 = k' = 2k$).

Given a node v, we say that an index $\ell \in [k]$ is redundant if both $2\ell - 1$ and 2ℓ are unaltered by Definition 7. Note that if ℓ is redundant, then for every $\alpha \in \mathcal{S}_v$, $T_v(\alpha)_{2\ell-1} = \alpha_\ell$ and $T_v(\alpha)_{2\ell} = L^{\mathrm{T}}(\alpha_\ell)$.

At a given recursion node v, the PIS *pis* will select as primary index the largest index $2\ell - 1$ such that $\ell \in [k]$ is redundant.

Let $d = d_{\max} = \lceil k/2 \rceil$. We first prove that there is always a redundant index for nodes up to depth $d - 1 = \lceil k/2 \rceil - 1$. Indeed, every recursive call can remove at most 2 redundant indices, and thus at depth $d - 1$, we have at least $k - 2(d-1) = k - 2\lceil k/2 \rceil + 2 \geq 1$ redundant indices. This proves that *pis* is well-defined up to depth $d = \lceil k/2 \rceil$ (at leaves of depth d we do not invoke *pis*).

In order to apply Lemma 3, we prove that the two conditions of Proposition 11 hold. First, the pair $(\mathcal{S}, T)$ satisfies the restrictions of Setting 1, and condition (**c1**) holds (note that $2d = 2\lceil k/2 \rceil \leq k + 1 < 2k$, as $k \geq 2$).

We now prove condition (**c2**). Namely, for a node v such that $j = pis(v)$ and $\alpha \in \mathcal{S}_v$, we prove that $T_v(\alpha)_j$ can be computed from $T_{v_i}(\alpha) = T_v(\alpha)^{\oplus(j,i)}$ (where i is a secondary index).

For a node v we select as primary index $j = 2\ell - 1$ for $\ell \in [k]$ redundant, and we have $T_v(\alpha)_j = T_v(\alpha)_{2\ell-1} = \alpha_\ell$ and $T_v(\alpha)_{2\ell} = L^{\mathrm{T}}(\alpha_\ell)$.

If $i \neq 2\ell$, then

$$T_{v_i}(\alpha)_{2\ell} = (T_v(\alpha)^{\oplus(2\ell-1,i)})_{2\ell} = T_v(\alpha)_{2\ell} = L^{\mathrm{T}}(\alpha_\ell),$$

and we can compute $L^{-\mathrm{T}}(T_{v_i}(\alpha)_{2\ell}) = \alpha_\ell = T_v(\alpha)_j$. Otherwise, $i = 2\ell$, and

$$T_{v_i}(\alpha)_{2\ell} = (T_v(\alpha)^{\oplus(2\ell-1,2\ell)})_{2\ell} = T_v(\alpha)_{2\ell-1} \oplus T_v(\alpha)_{2\ell} = \alpha_\ell \oplus L^{\mathrm{T}}(\alpha_\ell) = (L')^{\mathrm{T}}(\alpha_\ell).$$

Since $(L')^{\mathrm{T}}$ is an invertible linear transformation, we can compute $(L')^{-\mathrm{T}}(T_{v_i}(\alpha)_{2\ell}) = \alpha_\ell = T_v(\alpha)_j$. This proves condition (**c2**).

Applying our framework of Lemma 3 (with $d = \lceil k/2 \rceil = k/2 + c_k$, $k_0 = k' = 2k$), we obtain

$$\sum_{\alpha \in \mathcal{M}^n_{=k,k}} \widehat{\mu}_{n,2k}(\alpha_1, L^{\mathrm{T}}(\alpha_1), \ldots, \alpha_k, L^{\mathrm{T}}(\alpha_k))^2 \leq 2^d \frac{(k')^{2d}(k'-2d)^{k'/2-d}}{(N-k')^{k'/2+d}}$$

$$= 2^{k/2+c_k} \frac{(2k)^{k+2c_k}(k-2c_k)^{k/2-c_k}}{(N-2k)^{3k/2+c_k}} = 2^{3k/2+3c_k} \frac{k^{k+2c_k}(k-2c_k)^{k/2-c_k}}{(N-2k)^{3k/2+c_k}}.$$

Combining with (14) completes the proof. ■

The proof of Lemma 7 (given in Appendix E) uses Proposition 16 and shows that the bound on $\mathrm{W}^{=2}[\xi^{(L)}_{n,q}]$ dominates $\mathrm{Var}[\xi^{(L)}_{n,q}]$.

6 Indistinguishability Bounds for LXoP$[L, 2, n]$

In this section, we state and prove our main theorem regarding LXoP$[L, 2, n]$. Missing proofs are given in Appendix F.

Throughout this section, we assume that $L'(x) = x \oplus L(x)$ is invertible.

Theorem 3. *Given that* $N \geq 2^{10}$ *and* $q \leq N/32$, $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{LXoP}[L,2,n]}(q) \leq \frac{23q}{N^{1.5}}$. *Moreover, assuming* $\frac{32\sqrt{u}q_{\max}}{N^{1.5}} \leq \frac{1}{2}$, $\mathrm{Opt}^{\text{mu-prf}}_{\mathrm{LXoP}[L,2,n],u}(q_{max}) \leq \frac{32\sqrt{u}q_{\max}}{N^{1.5}}$.

The proof is based on the following two lemmas, proved below.

Lemma 8. $\mathrm{W}^{=1}[\xi^{(L)}_{n,2,q}] = \frac{4q}{(N-1)(N-2)^2}$.

Lemma 9. *Given that* $N \geq 2^{10}$ *and* $q \leq N/32$, $\sum_{k=2}^{q} \mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \leq \frac{2^{10.5}q^2}{N^3}$.

Proof (of Theorem 3). By Lemma 8 and Lemma 9, $\mathrm{Var}[\xi^{(L)}_{n,2,q}] = \sum_{k=1}^{q} \mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \leq \frac{4q}{(N-1)(N-2)^2} + \frac{2^{10.5}q^2}{N^3} \leq \frac{2^{11}q^2}{N^3}$, as $N \geq 2^{10}$. The result follows from Proposition 5. ■

6.1 Relation Between $\widehat{\xi}^{(L)}_{n,2,k}$ and $\widehat{\mu}_{n,3k}$ and Proof of Lemma 8

We first establish a connection between the Fourier coefficients of $\xi^{(L)}_{n,2,k}$ and those of $\mu_{n,3k}$. For $\alpha = (\alpha_1, \ldots, \alpha_k) = (\alpha^{(1)}_1, \alpha^{(2)}_1, \ldots \alpha^{(1)}_k, \alpha^{(2)}_k) \in \widehat{\mathbb{F}}_2^{k\times 2n}$, denote

$$t(\alpha) = (\alpha^{(1)}_1, L^{\mathrm{T}}(\alpha^{(2)}_1), \alpha^{(2)}_1 \oplus L^{\mathrm{T}}(\alpha^{(1)}_1) \ldots, \alpha^{(1)}_k, L^{\mathrm{T}}(\alpha^{(2)}_k), \alpha^{(2)}_k \oplus L^{\mathrm{T}}(\alpha^{(1)}_k)) \in \widehat{\mathbb{F}}_2^{3k\times n}.$$

Thus $t(\alpha_i) = (\alpha^{(1)}_i, L^{\mathrm{T}}(\alpha^{(2)}_i), \alpha^{(2)}_i \oplus L^{\mathrm{T}}(\alpha^{(1)}_i))$.

Proposition 17 (Relation between $\widehat{\xi}^{(L)}_{n,2,k}$ and $\widehat{\mu}_{n,3k}$). *For any* $\alpha \in \widehat{\mathbb{F}}_2^{k\times 2n}$, $\widehat{\xi}^{(L)}_{n,2,k}(\alpha) = \widehat{\mu}_{n,3k}(t(\alpha))$.

Proof. By definition of $\mathrm{LXoP}[L,2,n]$ and Proposition 1, for any $\alpha \in \widehat{\mathbb{F}}_2^{k\times 2n}$

$$\begin{aligned}
&\widehat{\xi}^{(L)}_{n,2,k}(\alpha) = \mathop{\mathrm{E}}_{x\sim\xi^{(L)}_{n,2,k}}[\chi_\alpha(x)] \\
&= \mathop{\mathrm{E}}_{y^{1,2,3}_{1..k}\sim\mu_{n,3k}}[\chi_{\alpha^{1,2}_{1..k}}(y^{(1)}_1 \oplus L(y^{(2)}_1), y^{(2)}_1 \oplus L(y^{(3)}_1), \ldots, y^{(1)}_k \oplus L(y^{(2)}_k), y^{(2)}_k \oplus L(y^{(3)}_k))] \\
&= \mathop{\mathrm{E}}_{y^{1,2,3}_{1..k}\sim\mu_{n,3k}}[\chi_{\alpha^{(1)}_{1..k}}(y^{(1)}_{1..k})\chi_{L^{\mathrm{T}}(\alpha^{(1)}_{1..k})\oplus\alpha^{(2)}_{1..k}}(y^{(2)}_{1..k})\chi_{L^{\mathrm{T}}(\alpha^{(2)}_{1..k})}(y^{(3)}_{1..k})] \\
&= \widehat{\mu}_{n,3k}(\alpha^{(1)}_{1..k}, L^{\mathrm{T}}(\alpha^{(1)}_{1..k}) \oplus \alpha^{(2)}_{1..k}, L^{\mathrm{T}}(\alpha^{(2)}_{1..k})) = \widehat{\mu}_{n,3k}(t(\alpha)).
\end{aligned}$$

The proof of Lemma 8 is given in Appendix F. It uses Proposition 17 and simple calculation. ■

6.2 Basic Properties of $t(\alpha)$

We prove basic properties of $t(\alpha)$. For $\alpha \in \widehat{\mathbb{F}}_2^{k\times 2n}$, recall that $\#\alpha = |\{i \in [k] \colon \alpha_i \neq 0\}|$ is the size of the support of α (over elements of $\widehat{\mathbb{F}}_2^{2n}$). On the other hand, for $t(\alpha) \in \widehat{\mathbb{F}}_2^{3k\times n}$, $\#t(\alpha) = |\{i \in [3k] \colon t(\alpha)_i \neq 0\}|$. In addition, note that for $i \in [k]$, $\#t(\alpha_i) \in \{0,1,2,3\}$. In fact, as proved below, $\#t(\alpha_i) \in \{0,2,3\}$.

Proposition 18 (Basic properties of $t(\alpha)$). *For $\alpha = (\alpha^{(1)}, \alpha^{(2)}) \in \widehat{\mathbb{F}}_2^{2n}$ with $\#\alpha = 1$ (i.e., $\alpha \neq 0$), let $t(\alpha) = (\alpha^{(1)}, L^{\mathrm{T}}(\alpha^{(2)}), \alpha^{(2)} \oplus L^{\mathrm{T}}(\alpha^{(1)}))$.*

Then, $\#t(\alpha) \in \{2, 3\}$. Moreover, if $\#t(\alpha) = 2$, denoting $(L^2)^{\mathrm{T}}(\alpha^{(1)}) = L^{\mathrm{T}}(L^{\mathrm{T}}(\alpha^{(1)}))$,

$$t(\alpha) \in \{(0, L^{\mathrm{T}}(\alpha^{(2)}), \alpha^{(2)}), (\alpha^{(1)}, 0, L^{\mathrm{T}}(\alpha^{(1)})), (\alpha^{(1)}, (L^2)^{\mathrm{T}}(\alpha^{(1)}), 0)\}.$$

Proof. We iterate over the 3 possibilities for a zero entry in $t(\alpha)$.

If $\alpha^{(1)} = 0$, then $\alpha^{(2)} \neq 0$. We have $t(\alpha) = (0, L^{\mathrm{T}}(\alpha^{(2)}), \alpha^{(2)})$, where $L^{\mathrm{T}}(\alpha^{(2)}) \neq 0$, as $L^{\mathrm{T}}(0) = 0$, $\alpha^{(2)} \neq 0$ and L^{T} is a permutation. Hence, $\#t(\alpha) = 2$.

Similarly, if $L^{\mathrm{T}}(\alpha^{(2)}) = 0$, then $\alpha^{(2)} = 0$ and hence $\alpha^{(1)} \neq 0$. Therefore, $\#t(\alpha) = \#(\alpha^{(1)}, 0, L^{\mathrm{T}}(\alpha^{(1)})) = 2$.

Finally, if $\alpha^{(2)} \oplus L^{\mathrm{T}}(\alpha^{(1)}) = 0$ then $\alpha^{(2)} = L^{\mathrm{T}}(\alpha^{(1)})$, hence $\alpha^{(1)} \neq 0$. We have $\#t(\alpha) = \#(\alpha^{(1)}, L^{\mathrm{T}}(L^{\mathrm{T}}(\alpha^{(1)})), 0) = \#(\alpha^{(1)}, (L^2)^{\mathrm{T}}(\alpha^{(1)}), 0) = 2$, as $\alpha^{(1)} \neq 0$, $(L^2)^{\mathrm{T}}(0) = 0$ and $(L^2)^{\mathrm{T}}$ is a permutation. ■

We conclude that if $\#\alpha = k$ (for $\alpha \in \widehat{\mathbb{F}}_2^{k \times 2n}$), then $\#t(\alpha) \in [2k, 3k]$. Denote $\#_2\alpha = |\{i \in [k] \colon \#t(\alpha_i) = 2\}|$ and $\#_3\alpha = \#\alpha - \#_2\alpha = |\{i \in [k] \colon \#t(\alpha_i) = 3\}|$. Therefore, if $\#\alpha = k$ and $\#_3\alpha = m$ then $\#t(\alpha) = 2(k - m) + 3m = 2k + m$.

6.3 Application of Main Framework and Proof of Lemma 9

We apply our main framework and use it to prove Lemma 9.

Proposition 19. *Let $2 \leq k \leq q \leq N/32$, and define $c_k = 0$ if k is even and $c_k = \frac{1}{2}$ otherwise (i.e. $c_k = \frac{k \bmod 2}{2}$). Then $\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \leq \binom{q}{k} 2^{7k/2+3c_k} (\frac{k}{N-2k})^{3k/2+c_k}$.*

Overview of the proofs of Proposition 19 and Lemma 9. The proof of Proposition 19 is given in Appendix F. It is a generalization of the analogous proof of Proposition 16 for $\mathrm{LXoP}[L, n]$, but is more technical, as it takes into account the different mask structures according to Proposition 18.

We begin by applying Proposition 17 and Proposition 18 and deducing

$$\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] = \tbinom{q}{k} \mathrm{W}^{=k}[\xi^{(L)}_{n,2,k}] = \tbinom{q}{k} \sum_{\alpha \in \mathcal{M}^{2n}_{=k,k}} \widehat{\xi}^{(L)}_{n,2,k}(\alpha)^2 = \tbinom{q}{k} \sum_{\alpha \in \mathcal{M}^{2n}_{=k,k}} \widehat{\mu}_{n,3k}(t(\alpha))^2$$
$$= \tbinom{q}{k} \sum_{m=0}^{k} \sum_{\alpha \in \mathcal{M}^{2n}_{=k,k}, \#_3\alpha = m} \widehat{\mu}_{n,3k}(t(\alpha))^2 = \tbinom{q}{k} \sum_{m=0}^{k} \sum_{\alpha \in \mathcal{S}^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2, \tag{15}$$

where $\mathcal{S}^{(k,m)} = \{\alpha \in \mathcal{M}^{2n}_{=k,k} \colon \#_3\alpha = m\}$. Fix a pair (k, m). We upper bound $\sum_{\alpha \in \mathcal{S}^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2$ using Lemma 3. According to restriction (**a2**) of Setting 1, we first need to partition the set $\mathcal{S}^{(k,m)}$ into subsets such that the (transformed) masks in each subset, $t(\alpha)$, share the same non-zero entries (over $\widehat{\mathbb{F}}_2^n$).

For every $\alpha \in \mathcal{S}^{(k,m)}$, there are m indices i with $\#t(\alpha_i) = 3$ and $k - m$ indices i with $\#t(\alpha_i) = 2$ (thus $\#t(\alpha) = 3m + 2(k - m) = 2k + m$ for $\alpha \in \mathcal{S}^{(k,m)}$).

By Proposition 18, every $i \in [k]$ with $\#t(\alpha_i) = 2$ has 3 possible structures that determine which 2 of its 3 entries are non-zero over $\widehat{\mathbb{F}}_2^n$ (while every $i \in [k]$ with $\#t(\alpha_i) = 3$ always has 3 non-zero entries). Therefore, $t(\alpha)$ for $\alpha \in \mathcal{S}^{(k,m)}$ has $\binom{k}{m}3^{k-m}$ possible non-zero index sets (with non-zero values over $\widehat{\mathbb{F}}_2^n$).

We thus partition the set $\mathcal{S}^{(k,m)}$ into $\binom{k}{m}3^{k-m}$ subsets, each with a common non-zero index set for $t(\alpha)$. We then apply Lemma 3 to bound the contribution of each subset of $\mathcal{S}^{(k,m)}$ to the total weight very similarly to Proposition 16. The Lemma 3 parameters we use are $d = k/2 + c_k$ (as in Proposition 16) and $k_0 = \#t(\alpha) = 2k + m$ (instead of $k_0 = 2k$ in Proposition 16). Ignoring terms of order $2^{O(k)}$, Lemma 3 bounds the weight for each of the $\binom{k}{m}3^{k-m}$ subsets of $\mathcal{S}^{(k,m)}$ by

$$2^{O(k)}\left(\tfrac{k_0}{N-k_0}\right)^{k_0/2+d} \le 2^{O(k)}\left(\tfrac{2k+m}{N-2k-m}\right)^{3k/2+m/2+c_k}.$$

Since $\binom{k}{m}3^{k-m} \le 4^k \le 2^{O(k)}$, then

$$\begin{aligned}&\sum_{\alpha\in\mathcal{S}^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2\\ &\le \tbinom{k}{m}3^{k-m}2^{O(k)}\left(\tfrac{2k+m}{N-2k-m}\right)^{3k/2+m/2+c_k} \le 2^{O(k)}\left(\tfrac{2k+m}{N-2k-m}\right)^{3k/2+m/2+c_k}.\end{aligned}$$

Combining with (15), the total weight is bounded as

$$\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \le \tbinom{q}{k}\sum_{m=0}^{k}\sum_{\alpha\in\mathcal{S}^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2 \le \tbinom{q}{k}\sum_{m=0}^{k} 2^{O(k)}\left(\tfrac{2k+m}{N-k-m}\right)^{3k/2+m/2+c_k}.$$

Observing that the term with $m = 0$ dominates the sum (as $k \le q$ is bounded), we deduce

$$\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \le \tbinom{q}{k}2^{O(k)}\left(\tfrac{2k}{N-2k}\right)^{3k/2+c_k} \le \tbinom{q}{k}2^{O(k)}\left(\tfrac{k}{N-2k}\right)^{3k/2+c_k},$$

as claimed in Proposition 19 (up to the constant factor hidden in $2^{O(k)}$).

Finally, the proof of Lemma 9 (given in Appendix F) uses Proposition 19 and shows that the bound on $\mathrm{W}^{=2}[\xi^{(L)}_{n,2,q}]$ dominates the sum $\sum_{k=2}^{q}\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}]$.

Acknowledgements. The author was supported by the Israel Science Foundation through grant no. 1903/20 and by a gift to Georgetown University.

References

1. Bellare, M., Impagliazzo, R.: A tool for obtaining tighter security analyses of pseudorandom function based constructions, with applications to PRP to PRF conversion. IACR Cryptol. ePrint Arch. (1999), http://eprint.iacr.org/1999/024
2. Bellare, M., Krovetz, T., Rogaway, P.: Luby-Rackoff Backwards: Increasing Security by Making Block Ciphers Non-invertible. In: Nyberg, K. (ed.) EUROCRYPT 1998. Lecture Notes in Computer Science, vol. 1403, pp. 266–280. Springer (1998). https://doi.org/10.1007/BFb0054132

3. Beyne, T., Chen, Y.L.: Information-Theoretic Security with Asymmetries. In: Reyzin, L., Stebila, D. (eds.) CRYPTO 2024. Lecture Notes in Computer Science, vol. 14923, pp. 463–494. Springer (2024)
4. Bhattacharya, S., Nandi, M.: Revisiting Variable Output Length XOR Pseudorandom Function. IACR Trans. Symmetric Cryptol. **2018**(1), 314–335 (2018). https://doi.org/10.13154/tosc.v2018.i1.314-335
5. Bhattacharya, S., Nandi, M.: Luby-Rackoff Backwards with More Users and More Security. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021. Lecture Notes in Computer Science, vol. 13092, pp. 345–375. Springer (2021). https://doi.org/10.1007/978-3-030-92078-4_12
6. Chen, S., Lampe, R., Lee, J., Seurin, Y., Steinberger, J.P.: Minimizing the Two-Round Even-Mansour Cipher. J. Cryptol. **31**(4), 1064–1119 (2018). https://doi.org/10.1007/s00145-018-9295-y
7. Chen, Y.L., Choi, W., Lee, C.: Improved Multi-user Security Using the Squared-Ratio Method. In: Handschuh, H., Lysyanskaya, A. (eds.) CRYPTO 2023. Lecture Notes in Computer Science, vol. 14082, pp. 694–724. Springer (2023). https://doi.org/10.1007/978-3-031-38545-2_23
8. Choi, W., Kim, H., Lee, J., Lee, Y.: Multi-user Security of the Sum of Truncated Random Permutations. In: Agrawal, S., Lin, D. (eds.) ASIACRYPT 2022. Lecture Notes in Computer Science, vol. 13792, pp. 682–710. Springer (2022). https://doi.org/10.1007/978-3-031-22966-4_23
9. Cogliati, B., Dutta, A., Nandi, M., Patarin, J., Saha, A.: Proof of Mirror Theory for a Wide Range of $\xi_{\max }$. In: Hazay, C., Stam, M. (eds.) EUROCRYPT 2023. Lecture Notes in Computer Science, vol. 14007, pp. 470–501. Springer (2023). https://doi.org/10.1007/978-3-031-30634-1_16
10. Cogliati, B., Lampe, R., Patarin, J.: The indistinguishability of the XOR of k permutations. In: Cid, C., Rechberger, C. (eds.) FSE 2014. Lecture Notes in Computer Science, vol. 8540, pp. 285–302. Springer (2014). https://doi.org/10.1007/978-3-662-46706-0_15
11. Dai, W., Hoang, V.T., Tessaro, S.: Information-Theoretic Indistinguishability via the Chi-Squared Method. In: Katz, J., Shacham, H. (eds.) CRYPTO 2017. Lecture Notes in Computer Science, vol. 10403, pp. 497–523. Springer (2017). https://doi.org/10.1007/978-3-319-63697-9_17
12. Dinur, I.: Tight indistinguishability bounds for the XOR of independent random permutations by fourier analysis. In: Joye, M., Leander, G. (eds.) EUROCRYPT 2024. Lecture Notes in Computer Science, vol. 14651, pp. 33–62. Springer (2024)
13. Dutta, A., Nandi, M., Saha, A.: Proof of Mirror Theory for $\xi_{\max }$ = 2. IEEE Trans. Inf. Theory **68**(9), 6218–6232 (2022). https://doi.org/10.1109/TIT.2022.3171178
14. Eberhard, S.: More on additive triples of bijections (2017), https://arxiv.org/abs/1704.02407
15. Eberhard, S., Manners, F., Mrazović, R.: Additive triples of bijections, or the toroidal semiqueens problem. J. Eur. Math. Soc. **21**(2), 441–463 (2018). https://doi.org/10.4171/JEMS/841
16. Golomb, S.W., Gong, G., Mittenthal, L.: Constructions of Orthomorphisms of $\mathbb{F}_2^n$. In: Gong, G., Mittenthal, L. (eds.) Finite Fields and Applications. pp. 178–195. Springer Berlin Heidelberg (2001). https://doi.org/10.1007/978-3-642-56755-1_15
17. Gueron, S.: Double Nonce Derive Key AES-GCM (DNDK-GCM), Active Internet-Draft (2024), https://datatracker.ietf.org/doc/draft-gueron-cfrg-dndkgcm/

18. Hall, C., Wagner, D.A., Kelsey, J., Schneier, B.: Building prfs from prps. In: Krawczyk, H. (ed.) CRYPTO 1998. vol. 1462, pp. 370–389. Springer (1998). https://doi.org/10.1007/BFb0055742
19. Hoang, V.T., Shen, Y.: Security of Streaming Encryption in Google's Tink Library. In: Ligatti, J., Ou, X., Katz, J., Vigna, G. (eds.) CCS 2020. pp. 243–262. ACM (2020). https://doi.org/10.1145/3372297.3417273, https://doi.org/10.1145/3372297.3417273
20. Iwata, T.: New Blockcipher Modes of Operation with Beyond the Birthday Bound Security. In: Robshaw, M.J.B. (ed.) FSE 2006. vol. 4047, pp. 310–327. Springer (2006)
21. Iwata, T., Mennink, B., Vizár, D.: CENC is Optimally Secure. IACR Cryptol. ePrint Arch. p. 1087 (2016), http://eprint.iacr.org/2016/1087
22. Lucks, S.: The Sum of PRPs Is a Secure PRF. In: Preneel, B. (ed.) EUROCRYPT 2000. Lecture Notes in Computer Science, vol. 1807, pp. 470–484. Springer (2000). https://doi.org/10.1007/3-540-45539-6_34
23. Mennink, B., Preneel, B.: On the XOR of Multiple Random Permutations. In: Malkin, T., Kolesnikov, V., Lewko, A.B., Polychronakis, M. (eds.) ACNS 2015. Lecture Notes in Computer Science, vol. 9092, pp. 619–634. Springer (2015). https://doi.org/10.1007/978-3-319-28166-7_30
24. O'Donnell, R.: Analysis of Boolean Functions. Cambridge University Press (2014)
25. Patarin, J.: A Proof of Security in $O(2^n)$ for the Xor of Two Random Permutations. In: Safavi-Naini, R. (ed.) ICITS 2008. Lecture Notes in Computer Science, vol. 5155, pp. 232–248. Springer (2008). https://doi.org/10.1007/978-3-540-85093-9_22
26. Patarin, J.: Introduction to mirror theory: Analysis of systems of linear equalities and linear non equalities for cryptography. IACR Cryptol. ePrint Arch. p. 287 (2010), http://eprint.iacr.org/2010/287
27. Patarin, J.: Generic Attacks for the Xor of k Random Permutations. In: Jr., M.J.J., Locasto, M.E., Mohassel, P., Safavi-Naini, R. (eds.) ACNS 2013. vol. 7954, pp. 154–169. Springer (2013). https://doi.org/10.1007/978-3-642-38980-1_10

A Missing Proofs and Details from Section 2

Proposition 20. *Let $a, b, c, d, k \in \mathbb{R}^{\geq 0}$. Define the functions $B(k) = (ak + b)^{ck+d}$ and $C(k) = \frac{1}{(b-ak)^{ck+d}}$. Then*

$$\frac{B(k+1)}{B(k)} \leq (a(k+1)+b)^c e^{\frac{a(ck+d)}{ak+b}}, \text{ and } \frac{C(k+1)}{C(k)} \leq \frac{1}{(b-a(k+1))^c} e^{\frac{a(ck+d)}{b-a(k+1)}},$$

where the last inequality assumes $b > a(k+1)$.

Proof (of Proposition 20). We have

$$\begin{aligned}\tfrac{B(k+1)}{B(k)} &= \tfrac{(a(k+1)+b)^{c(k+1)+d}}{(ak+b)^{ck+d}} = \tfrac{(a(k+1)+b)^{c(k+1)+d}}{(a(k+1)+b)^{ck+d}} \tfrac{(a(k+1)+b)^{ck+d}}{(ak+b)^{ck+d}} \\ &= (a(k+1)+b)^c (1+\tfrac{a}{ak+b})^{ck+d} \leq (a(k+1)+b)^c e^{\frac{a(ck+d)}{ak+b}}.\end{aligned}$$

and

$$\begin{aligned}\tfrac{C(k+1)}{C(k)} &= \tfrac{(b-a(k+1))^{-c(k+1)-d}}{(b-ak)^{-ck-d}} = \tfrac{(b-a(k+1))^{-c(k+1)-d}}{(b-a(k+1))^{-ck-d}} \tfrac{(b-a(k+1))^{-ck-d}}{(b-ak)^{-ck-d}} \\ &= \tfrac{1}{(b-a(k+1))^c} (1+\tfrac{a}{b-a(k+1)})^{ck+d} \leq \tfrac{1}{(b-a(k+1))^c} e^{\frac{a(ck+d)}{b-a(k+1)}}.\end{aligned}$$

■

Proof (of Proposition 3). We have

$$\begin{aligned}2\,\mathrm{SD}(\varphi, \mathbf{1}_{qn}) &= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\varphi(x) - 1| = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \alpha\neq 0}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \\ &= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{S}} \widehat{\varphi}(\alpha)\chi_\alpha(x) + \sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \\ &\leq \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{S}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| + \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)|\end{aligned}$$

For the upper bound, it remains to prove that $\mathrm{E}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \leq \sqrt{\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)^2}$. Applying the Cauchy-Schwarz inequality,

$$\begin{aligned}&\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \leq \sqrt{\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} [\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)]^2} \\ &= \sqrt{\sum_{(\alpha,\beta)\in\overline{\mathcal{S}}\times\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\widehat{\varphi}(\beta) \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} [\chi_\alpha(x)\chi_\beta(x)]} = \sqrt{\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)^2},\end{aligned}$$

where the final equality is by orthogonality of the characters.

For the lower bound, observe similarly that

$$\begin{aligned}
2\,\mathrm{SD}(\varphi, \mathbf{1}_{qn}) &= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{S}} \widehat{\varphi}(\alpha)\chi_\alpha(x) + \sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \\
&\geq \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{S}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| - \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| \\
&\geq \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{S}} \widehat{\varphi}(\alpha)\chi_\alpha(x)| - \sqrt{\sum_{\alpha\in\overline{\mathcal{S}}} \widehat{\varphi}(\alpha)^2}.
\end{aligned}$$

■

Proof (of Proposition 5). First, by (1) and Proposition 3,

$$\mathrm{Opt}_H^{\mathrm{prf}}(q) = \mathrm{SD}(\varphi_H^{n,q}, \mathbf{1}_{qn}) \leq \tfrac{1}{2}\sqrt{\mathrm{Var}[\varphi_H^{n,q}]}.$$

Second, by (2), Proposition 3 and Proposition 4,

$$\begin{aligned}
\mathrm{Opt}_{H,u}^{\text{mu-prf}}(q_{max}) &= \mathrm{SD}((\varphi_H^{n,q_{\max}})^{\times u}, \mathbf{1}_{uq_{\max}n}) \\
&\leq \tfrac{1}{2}\sqrt{\mathrm{Var}[\varphi_H^{n,q_{\max}})^{\times u}]} \leq \tfrac{1}{\sqrt{2}}\sqrt{u\,\mathrm{Var}[\varphi_H^{n,q_{\max}}]}.
\end{aligned}$$

■

Proof (of Proposition 7). Let $y \in \mathbb{F}_2^n$ be arbitrary. Observe that for $x \in \mathbb{F}_2^{k\times n}$, $\mu_{n,k}(x) = \mu_{n,k}(x_1 \oplus y, \ldots, x_k \oplus y)$. Therefore,

$$\begin{aligned}
\widehat{\mu}_{n,k}(\alpha) &= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x)\chi_\alpha(x)] = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x_1 \oplus y, \ldots, x_k \oplus y)\chi_\alpha(x_1, \ldots, x_k)] \\
&= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x_1 \oplus y, \ldots, x_k \oplus y)\chi_\alpha(x_1 \oplus y, \ldots, x_k \oplus y)]\chi_{(\oplus_{i\in[k]}\alpha_i)}(y) \\
&= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x)\chi_\alpha(x)]\chi_{(\oplus_{i\in[k]}\alpha_i)}(y) = \widehat{\mu}_{n,k}(\alpha)\chi_{(\oplus_{i\in[k]}\alpha_i)}(y).
\end{aligned}$$

If $\widehat{\mu}_{n,k}(\alpha) \neq 0$, we divide both sides by $\widehat{\mu}_{n,k}(\alpha)$. We deduce that for every $y \in \mathbb{F}_2^n$, $\chi_{(\oplus_{i\in[k]}\alpha_i)}(y) = 1$, implying that $\oplus_{i\in[k]}\alpha_i = 0$. ■

Proof (of Proposition 8). Denote $k_0 = k$. We assume that $\mathcal{NZ}_\alpha = [k_0] = [k]$, which is possible without loss of generality by Proposition 6.

We further assume that $k_1 = k$, as adding or removing zero elements from α does not change $\widehat{\mu}_{n,k}(\alpha)$. Finally, using Proposition 6 we assume without loss of

generality that $j = k$. By Proposition 1,

$$\begin{aligned}
\widehat{\mu}_{n,k}(\alpha) &= \mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)] = \mathop{\mathrm{E}}_{x\sim\mu_{n,k-1}}\Big[\mathop{\mathrm{E}}_{x_k\sim\mathbb{F}_2^n\setminus\{x_1,\ldots,x_{k-1}\}}[\chi_\alpha(x_{1..k-1},x_k)]\Big] \\
&= \tfrac{N}{N-k+1}\mathop{\mathrm{E}}_{x\sim\mu_{n,k-1}}\Big[\mathop{\mathrm{E}}_{x_k\sim\mathbb{F}_2^n}[\chi_\alpha(x_{1..k-1},x_k)]\Big] \\
&\quad - \tfrac{k-1}{N-k+1}\mathop{\mathrm{E}}_{x\sim\mu_{n,k-1}}\Big[\mathop{\mathrm{E}}_{x_k\sim\{x_1,\ldots,x_{k-1}\}}[\chi_\alpha(x_{1..k-1},x_k)]\Big] \\
&= \tfrac{N}{N-k+1}\mathop{\mathrm{E}}_{x\sim\mu_{n,k-1}}[\chi_{\alpha_{1..k-1}}(x_{1..k-1})]\mathop{\mathrm{E}}_{x_k\sim\mathbb{F}_2^n}[\chi_{\alpha_k}(x_k)] \\
&\quad - \tfrac{1}{N-k+1}\sum_{i=1}^{k-1}\mathop{\mathrm{E}}_{x\sim\mu_{n,k-1}}[\chi_\alpha(x_{1..k-1},x_i)] \\
&= 0 - \tfrac{1}{N-k+1}\sum_{i=1}^{k-1}\mathop{\mathrm{E}}_{x\sim\mu_{n,k-1}}[\chi_{(\alpha_{1..,i-1},\alpha_i\oplus\alpha_k,\alpha_{i+1..,k-1})}(x_{1..k-1})] \\
&= -\tfrac{1}{N-k+1}\sum_{i=1}^{k-1}\widehat{\mu}_{n,k-1}(\alpha^{\oplus(k,i)}) = -\tfrac{1}{N-k+1}\sum_{i=1}^{k-1}\widehat{\mu}_{n,k}(\alpha^{\oplus(k,i)}),
\end{aligned}$$

where in the fifth equality we used $\mathrm{E}_{x_k\sim\mathbb{F}_2^n}[\chi_{\alpha_k}(x_k)] = \mathrm{E}[\chi_{\alpha_k}\chi_0] = 0$, which holds by orthogonality of characters since $\alpha_k \neq 0$. ■

Proof (of Lemma 1). We may assume that $k_0 = k_1 = k$, as adding and removing 0 elements from α does not change $\widehat{\mu}_{n,k}(\alpha)$. The proof is by induction on k.

For $k = 0$, we have $\mathrm{M}^{=0}[\mu_{n,k}] = 1 = \frac{1}{\sqrt{\binom{N}{0}}}$.

Next, let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ have $\#\alpha = k$. For $k = 1$, by Proposition 7, $|\widehat{\mu}_{n,k}(\alpha)| = 0 < \frac{1}{\sqrt{\binom{N}{1}}}$. For $k \geq 2$, by Proposition 8 and the triangle inequality,

$$|\widehat{\mu}_{n,k}(\alpha)| = \left|-\tfrac{1}{N-k+1}\sum_{i=1}^{k-1}\widehat{\mu}_{n,k}(\alpha^{\oplus(k,i)})\right| \leq \tfrac{1}{N-k+1}\sum_{i=1}^{k-1}|\widehat{\mu}_{n,k}(\alpha^{\oplus(k,i)})|.$$

We have $\#\alpha^{\oplus(k,i)} \in \{k-1, k-2\}$. Assume that for m values of $i \in [k-1]$, $\#\alpha^{\oplus(k,i)} = k-2$ holds. Then, by the induction hypothesis (assuming $k \leq N/2$),

$$\begin{aligned}
|\widehat{\mu}_{n,k}(\alpha)| &\leq \tfrac{m}{N-k+1}\mathrm{M}^{=k-2}[\mu_{n,k}] + \tfrac{k-1-m}{N-k+1}\mathrm{M}^{=k-1}[\mu_{n,k}] \\
&\leq \tfrac{m}{N-k+1}\frac{1}{\sqrt{\binom{N}{k-2}}} + \tfrac{k-1-m}{N-k+1}\frac{1}{\sqrt{\binom{N}{k-1}}} \leq \tfrac{k-1}{N-k+1}\frac{1}{\sqrt{\binom{N}{k-2}}} \\
&= \tfrac{k-1}{N-k+1}\sqrt{\tfrac{k-2}{N}\tfrac{k-3}{N-1}\cdots\tfrac{1}{N-(k-3)}} \\
&\leq \sqrt{\tfrac{k}{N-k+2}\tfrac{k-1}{N-k+1}}\sqrt{\tfrac{k-2}{N}\tfrac{k-3}{N-1}\cdots\tfrac{1}{N-(k-3)}} = \frac{1}{\sqrt{\binom{N}{k}}}.
\end{aligned}$$

■

Proof (of Proposition 10). We assume without loss of generality that $k_0 = k_1 = k$. The proof is by induction on k. By Proposition 8,

$$\widehat{\mu}_{n,k}(\alpha) = -\tfrac{1}{N-k+1} \sum_{i=1}^{k-1} \widehat{\mu}_{n,k}(\alpha^{\oplus(k,i)}).$$

For $k = 2$, this gives $-\frac{1}{N-1}\widehat{\mu}_{n,2}(\alpha^{\oplus(2,1)}) = -\frac{1}{N-1}$, as $\#(\alpha^{\oplus(2,1)}) = 0$ and hence $\widehat{\mu}_{n,2}(\alpha^{\oplus(2,1)}) = 1$.

For $k > 2$, for all $i \in [k-1]$, $\alpha^{\oplus(k,i)}$ is equal to $\alpha^{\oplus(k,1)}$ (up to a permutation of the elements). Therefore, $\widehat{\mu}_{n,k}(\alpha) = -\frac{k-1}{N-k+1}\widehat{\mu}_{n,k}(\alpha^{\oplus(k,1)})$. Since $\#(\alpha^{\oplus(k,1)}) = k-2$, and $\alpha^{\oplus(k,1)}$ has all non-zero elements equal (as α), we apply the induction hypothesis to $\alpha^{\oplus(k,1)}$ and deduce

$$\widehat{\mu}_{n,k}(\alpha) = -\tfrac{k-1}{N-k+1}(-1)^{k/2-1}\tfrac{k-3}{N-1}\cdots\tfrac{1}{N-(k-3)} = (-1)^{k/2}\tfrac{k-1}{N-1}\tfrac{k-3}{N-3}\cdots\tfrac{1}{N-(k-1)}.$$

Next, note that $|\widehat{\mu}_{n,k}(\alpha)| \le \frac{1}{\sqrt{\binom{N}{k}}}$ holds by Lemma 1. It remains to prove that $\frac{1}{\sqrt{k\binom{N}{k}}} \le |\widehat{\mu}_{n,k}(\alpha)|$. Indeed,

$$\begin{aligned}\sqrt{k}|\widehat{\mu}_{n,k}(\alpha)| &= \sqrt{k}\tfrac{k-1}{N-1}\tfrac{k-3}{N-3}\cdots\tfrac{1}{N-(k-1)} \ge \sqrt{\tfrac{k}{N-(k-1)}}\sqrt{\tfrac{k-1}{N-1}\tfrac{k-2}{N-2}\cdots\tfrac{1}{N-(k-1)}} \\ &\ge \sqrt{\tfrac{k}{N}}\sqrt{\tfrac{k-1}{N-1}\tfrac{k-2}{N-2}\cdots\tfrac{1}{N-(k-1)}} = \frac{1}{\sqrt{\binom{N}{k}}}.\end{aligned}$$

■

Cryptographic Preliminaries

We use the standard notion of PRF security, as defined below. Let $H : \mathcal{K} \times \{0,1\}^{m_1} \mapsto \{0,1\}^{m_2}$ be a family of functions and $\mathrm{Func}(m_1, m_2)$ be the set of all functions $g : \{0,1\}^{m_1} \mapsto \{0,1\}^{m_2}$. Let A be an algorithm with oracle access to a function $f : \{0,1\}^{m_1} \mapsto \{0,1\}^{m_2}$. The PRF advantage of A against H is

$$\mathrm{Adv}^{\mathrm{prf}}_H(A) = \left| \Pr_{K\sim\mathcal{K}}[A^{H_K(\cdot)} \Rightarrow 1] - \Pr_{f\sim\mathrm{Func}(m_1,m_2)}[A^{f(\cdot)} \Rightarrow 1] \right|.$$

We further define the optimal advantage

$$\mathrm{Opt}^{\mathrm{prf}}_H(q) = \max\{\mathrm{Adv}^{\mathrm{prf}}_H(A) \colon A \text{ makes } q \text{ queries}\}.$$

In the multi-user setting we have u users, each with an independent instantiation of the cryptosystem. The adversary can issue (up to) $q_{\max}$ queries to each user with the goal of distinguishing the u instantiations of the cryptosystem from u

instantiations of a random function. We define the PRF advantage of A against H in the multi-user setting as

$$\begin{aligned}\mathrm{Adv}^{\text{mu-prf}}_{H,u}(A) = \big| &\Pr_{K_1,\ldots,K_u\sim\mathcal{K}}[A^{H_{K_1}(\cdot),\ldots,H_{K_u}(\cdot)} \Rightarrow 1] \\ &- \Pr_{f_1,\ldots,f_u\sim\mathrm{Func}(m_1,m_2)}[A^{f_1(\cdot),\ldots,f_u(\cdot)} \Rightarrow 1]\big|\end{aligned}$$

We further define the optimal advantage

$$\mathrm{Opt}^{\text{mu-prf}}_{H,u}(q_{\max}) = \max\{\mathrm{Adv}^{\text{mu-prf}}_{H,u}(A)\colon A \text{ makes } q_{\max} \text{ queries to each user}\}.$$

B Bounds on Advantage for Symmetric Functions

Proposition 21 (Bounds on advantage for symmetric functions). *Assume that the output distribution generated by $H : \mathcal{K} \times \{0,1\}^m \mapsto \{0,1\}^n$ is independent of the queries of the adversary. Denote by $\varphi^{n,q}_H : \mathbb{F}_2^{q\times n} \to \mathbb{R}^{\geq 0}$ the density function generated by H. Moreover, assume that $\varphi^{n,q}_H$ is symmetric in the sense that every element of the sample is marginally distributed as $\varphi^{n,1}_H$. Then,*

$$Opt^{prf}_H(q) \leq q\,\mathrm{SD}(\varphi^{n,1}_H, \mathbf{1}_n) + \frac{1}{2}\sqrt{\sum_{k=2}^{q} \mathrm{W}^{=k}[\varphi^{n,q}_H]}.$$

Proof. Let $\mathcal{S} = \mathcal{M}^n_{=1,q} = \{\alpha \in \widehat{\mathbb{F}}_2^{q\times n}\colon \#\alpha = 1\}$. By (1) and the upper bound of Proposition 3,

$$\begin{aligned}&2\mathrm{Opt}^{\text{prf}}_H(q) = 2\,\mathrm{SD}(\varphi^{n,q}_H, \mathbf{1}_{qn}) \\ \leq &\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\varphi}^{n,q}_H(\alpha)\chi_\alpha(x)| + \sqrt{\sum_{k=2}^{q} \mathrm{W}^{=k}[\varphi^{n,q}_H]}.\end{aligned}$$

It remain to prove that

$$\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\varphi}^{n,q}_H(\alpha)\chi_\alpha(x)| \leq 2q\,\mathrm{SD}(\varphi^{n,1}_H, \mathbf{1}_n).$$

For $\alpha \in \widehat{\mathbb{F}}_2^{q\times n}$ with $\#\alpha = 1$, define $in(\alpha)$ to be the unique index i with $\alpha_i \neq 0$. By symmetry of $\varphi_H^{n,q}$, we have $\widehat{\varphi}_H^{n,q}(\alpha) = \widehat{\varphi}_H^{n,1}(\alpha_{in(\alpha)})$. Therefore,

$$\begin{aligned}
&\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\varphi}_H^{n,q}(\alpha)\chi_\alpha(x)| = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\varphi}_H^{n,1}(\alpha_{in(\alpha)}) \prod_{i\in[q]} \chi_{\alpha_i}(x_i)| \\
&= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\varphi}_H^{n,1}(\alpha_{in(\alpha)})\chi_{\alpha_{in(\alpha)}}(x_{in(\alpha)})| \\
&= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{i=1}^{q} \sum_{\substack{\beta\in\widehat{\mathbb{F}}_2^n \\ \beta\neq 0}} \widehat{\varphi}_H^{n,1}(\beta)\chi_\beta(x_i)| = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{i=1}^{q} (\varphi_H^{n,1}(x_i) - \widehat{\varphi}_H^{n,1}(0)\chi_0(x_i))| \\
&\leq \sum_{i=1}^{q} \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |(\varphi_H^{n,1}(x_i) - 1)| = q \mathop{\mathrm{E}}_{y\sim\mathbb{F}_2^n} |\varphi_H^{n,1}(y) - 1| = 2q\,\mathrm{SD}(\varphi_H^{n,1}, \mathbf{1}_n).
\end{aligned}$$

■

C Missing Proofs and Additional Details from Section 3

Lemma 3 is proved using three additional propositions which we state and prove below.

Proposition 22 (Recursive validity of Setting 1). *Assume that* calcW *is applicable up to depth $d_{\max}$ with parameters $\mathcal{S}$, T and a PIS, pis. Then, for each node v at depth at most $d_{\max}$, **(d1)** T_v is injective on the elements of $\mathcal{S}_v$, **(d2)** for every $\alpha \in \mathcal{S}_v$ and every $\ell \in [k']$, $T_v(\alpha)_\ell \neq 0$ if and only if $\ell \in \mathcal{N}_v$.*

Proof. The proof is by induction on the depth $d \leq d_{\max}$ of v. The two restrictions hold at the root ($d = 0$) by assumption. Assume correctness up to depth $d \leq d_{\max} - 1$ and let v be a node of depth d. Consider a child node $v_{i,b}$ for $i \in \mathcal{N}_v \backslash \{j\}$ and $b \in \{0, 1\}$. Recall that $T_{v_{i,b}}(\alpha) = T_{v_i}(\alpha) = T_v(\alpha)^{\oplus(j,i)}$ only changes entries i, j of $T_v(\alpha)$.

We prove (**d1**). Consider $\alpha, \beta \in \mathcal{S}_{v_{i,b}}$ such that $T_{v_{i,b}}(\alpha) = T_{v_{i,b}}(\beta)$. We show that $\alpha = \beta$. Since calcW is applicable up to depth $d_{\max}$, condition (**b2**) of Definition 6 implies that $T_v(\alpha) = T_v(\beta)$. Indeed, if $T_v(\alpha) \neq T_v(\beta)$ but $T_{v_{i,b}}(\alpha) = T_{v_{i,b}}(\beta)$ then $T_v(\alpha)$ cannot be uniquely recovered from $T_{v_{i,b}}(\alpha) = T_{v_i}(\alpha)$.

Since $T_v(\alpha) = T_v(\beta)$, the induction hypothesis implies that $\alpha = \beta$ (as $\mathcal{S}_{v_{i,b}} \subseteq \mathcal{S}_v$ and T_v is injective of $\mathcal{S}_v$). This proves (**d1**).

We prove (**d2**). Consider $\alpha \in \mathcal{S}_{v_{i,b}}$ and let $\ell \in [k']$. If $\ell = j$, then $T_{v_{i,b}}(\alpha)_\ell = 0$ and $\ell \notin \mathcal{N}_{v_{i,b}}$ by definition of calcW.

Next, consider $\ell = i$. Then $T_v(\alpha)_\ell \neq 0$ and $\ell \in \mathcal{N}_v$ by the hypothesis. Therefore, if $T_{v_{i,b}}(\alpha)_\ell \neq 0$, then $b = 1$ and also $i \in \mathcal{N}_{v_{i,1}}$, while if $T_{v_{i,b}}(\alpha)_\ell = 0$, then $b = 0$ and also $i \notin \mathcal{N}_{v_{i,0}}$ (by definition of calcW).

Otherwise $\ell \notin \{i, j\}$. Then $T_{v_{i,b}}(\alpha)_\ell = T_v(\alpha)_\ell$, so $T_{v_{i,b}}(\alpha)_\ell \neq 0$ if and only if $T_v(\alpha)_\ell \neq 0$. By the induction hypothesis, this holds if and only if $\ell \in \mathcal{N}_v$, which holds if and only if $\ell \in \mathcal{N}_{v_{i,b}}$ (by definition of calcW). This completes the proof. ■

Proposition 23. *Assume that* calcW *is applicable up to depth* $d_{\max} \geq 0$ *with parameters* $(\mathcal{S}, T)$ *and a PIS,* pis*. Then, for every node* v *with depth* $d \leq d_{\max}$ *such that* $|\mathcal{N}_v| \leq \frac{N}{2}$*,*

$$\sum_{\alpha \in \mathcal{S}_v} \widehat{\mu}_{n,k'}(T_v(\alpha))^2 \leq \mathrm{calcW}_{pis,d_{\max}}(d, v, \mathcal{S}_v, T_v, \mathcal{N}_v).$$

Proof (of Proposition 23). We prove the result by induction on $d \leq d_{\max}$ (starting with $d = d_{\max}$, down to $d = 0$). Let $k'_v = |\mathcal{N}_v|$. For $d = d_{\max}$,

$$\begin{aligned}\sum_{\alpha \in \mathcal{S}_v} \widehat{\mu}_{n,k'}(T_v(\alpha))^2 &\leq \sum_{\beta \in \widehat{\mathbb{F}}_2^{k'_v}} \widehat{\mu}_{n,k'_v}(\beta)^2 = \mathrm{W}^{=k'_v}[\mu_{n,k'_v}] \leq (\tfrac{k'_v}{N-k'_v})^{k'_v/2} \\ &= \mathrm{calcW}_{pis,d_{\max}}(d, v, \mathcal{S}_v, T_v, \mathcal{N}_v),\end{aligned}$$

where the first inequality relies on (**d2**) in Proposition 22, as we trim the $k' - |\mathcal{N}_v|$ zero entries that are common to all $T_v(\alpha)$ for $\alpha \in \mathcal{S}_v$. It further relies on (**d1**) in Proposition 22, as each $\alpha \in \mathcal{S}_v$ is mapped to a single $\beta \in \widehat{\mathbb{F}}_2^{k'_v}$ after removing the common zero entries. The second inequality is by Lemma 2. We remark that the assumption that calcW is applicable up to depth $d_{\max}$ implies $2d_{\max} < |\mathcal{N}|$ and hence $k'_v = |\mathcal{N}_v| \geq |\mathcal{N}| - 2d_{\max} > 0$, so Lemma 2 can indeed be applied.

For $d < d_{\max}$, by reordering elements, we assume without loss of generality that $\mathcal{N}_v = [k'_v]$ and $pis(v) = k'_v$. Then, by Proposition 9,

$$\begin{aligned}\sum_{\alpha \in \mathcal{S}_v} \widehat{\mu}_{n,k'}(T_v(\alpha))^2 &\leq \tfrac{k'_v - 1}{(N-k'_v+1)^2} \sum_{\alpha \in \mathcal{S}_v} \sum_{i=1}^{k'_v - 1} \widehat{\mu}_{n,k'}(T_v(\alpha)^{\oplus(k'_v,i)})^2 \\ = \tfrac{k'_v - 1}{(N-k'_v+1)^2} \sum_{i=1}^{k'_v - 1} \Big(&\sum_{\alpha \in \mathcal{S}_{v_{i,0}}} \widehat{\mu}_{n,k'}(T_v(\alpha)^{\oplus(k'_v,i)})^2 + \sum_{\alpha \in \mathcal{S}_{v_{i,1}}} \widehat{\mu}_{n,k'}(T_v(\alpha)^{\oplus(k'_v,i)})^2 \Big),\end{aligned} \tag{16}$$

where we use the fact that $\mathcal{S}_{v_{i,0}} \cup \mathcal{S}_{v_{i,1}} = \mathcal{S}_v$ for every $i \in [k'_v - 1]$. Also, $k'_v = |\mathcal{N}_v| \geq |\mathcal{N}| - 2(d_{\max} - 1) > 2$, so Proposition 9 can indeed be applied.

We have

$$\begin{aligned}\sum_{\alpha \in \mathcal{S}_{v_{i,0}}} \widehat{\mu}_{n,k'}(T_v(\alpha)^{\oplus(k'_v,i)})^2 &= \sum_{\alpha \in \mathcal{S}_{v_{i,0}}} \widehat{\mu}_{n,k'}(T_{v_{i,0}}(\alpha))^2 \\ &\leq \mathrm{calcW}_{pis,d_{\max}}(d+1, v_{i,0}, \mathcal{S}_{v_{i,0}}, T_{v_{i,0}}, \mathcal{N}_{v_{i,0}}),\end{aligned}$$

where the inequality is by the induction hypothesis (relying on applicability up to depth $d_{\max}$).

Moreover, a similar inequality holds for the sum over $\mathcal{S}_{v_{i,1}}$. Plugging these inequalities into (16), and comparing with the return value of calcW, we deduce $\sum_{\alpha \in \mathcal{S}_v} \widehat{\mu}_{n,k'}(T_v(\alpha))^2 \leq \mathrm{calcW}_{pis,d_{\max}}(d, v, \mathcal{S}_v, T_v, \mathcal{N}_v)$, concluding the proof. ■

Proposition 24. *Let v be a node of depth d such that $d \le d_{\max}$ and $k'_v = |\mathcal{N}_v| \le N/8$. Denote $d' = d_{\max} - d$. Then,*

$$\mathrm{calcW}_{pis,d_{\max}}(d, v, \mathcal{S}_v, T_v, \mathcal{N}_v) \le 2^{d'} \frac{(k'_v)^{2d'}(k'_v - 2d')^{k'_v/2-d'}}{(N-k'_v)^{k'_v/2+d'}}.$$

Proof. To simplify notation, denote $k = k'_v$. The recursion tree starting from v is of depth $d' = d_{\max} - d$. Each leaf u contributes to the output at most

$$(\tfrac{k}{(N-k)^2})^{d'} (\tfrac{k'_u}{N-k'_u})^{k'_u/2}, \tag{17}$$

where we used the fact that for each internal node w, $k'_w \le k$ and thus $\frac{k'_w-1}{(N-k'_w+1)^2} \le \frac{k-1}{(N-k+1)^2} \le \frac{k}{(N-k)^2}$.

Initially, $|\mathcal{N}_v| = k'_v = k$. For each internal node w, for each $i \in \mathcal{N}_w \backslash \{j\}$, $|\mathcal{N}_{w_{i,0}}| = |\mathcal{N}_w| - 2$ (there are $k'_w < k$ such children $w_{i,0}$) and $|\mathcal{N}_{w_{i,1}}| = |\mathcal{N}_w| - 1$ (there are $k'_w < k$ such children $w_{i,1}$).

Therefore, for every leaf u, $k'_u \in [k-d', k-2d']$. More specifically for $c \in \{0, 1, \ldots, d'\}$, the number of leaf nodes u with $k'_u = k - 2d' + c$ is at most $k^{d'}\binom{d'}{c}$. Hence, using (17), we bound

$$\begin{aligned}
&\mathrm{calcW}_{pis,d_{\max}}(d, v, \mathcal{S}_v, T_v, \mathcal{N}_v) \le (\tfrac{k}{(N-k)^2})^{d'} \sum_{u \text{ leaf}} (\tfrac{k'_u}{N-k'_u})^{k'_u/2} \\
&\le (\tfrac{k}{(N-k)^2})^{d'} k^{d'} \sum_{c=0}^{d'} \tbinom{d'}{c} (\tfrac{k-2d'+c}{N-k+2d'-c})^{(k-2d'+c)/2} \qquad (18) \\
&\le (\tfrac{k}{N-k})^{2d'} 2^{d'} \max_{c \in \{0,\ldots,d'\}} \{(\tfrac{k-2d'+c}{N-k+2d'-c})^{(k-2d'+c)/2}\}.
\end{aligned}$$

Denote $B(c) = (\frac{k-2d'+c}{N-k+2d'-c})^{(k-2d'+c)/2}$. For $c+1 \le d'$, by Proposition 20,

$$\begin{aligned}
\tfrac{B(c+1)}{B(c)} &\le e^{\frac{(k-2d'+c)/2}{k-2d'+c} + \frac{(k-2d'+c)/2}{N-k+2d'-c-1}} (\tfrac{k-2d'+c+1}{N-k+2d'-c-1})^{1/2} \\
&\le e^{\frac{1}{2} + \frac{k/2}{N-k}} (\tfrac{k}{N-k})^{1/2} \le e^{4/7} (\tfrac{1}{7})^{1/2} \le 1,
\end{aligned}$$

where we have used the assumption that $k \le N/8$. Thus,

$$\begin{aligned}
\max_{c \in \{0,\ldots,d'\}} \{(\tfrac{k-2d'+c}{N-k+2d'-c})^{(k-2d'+c)/2}\} &= \max_{c \in \{0,\ldots,d'\}} \{B(c)\} = B(0) \\
&= (\tfrac{k-2d'}{N-k+2d'})^{(k-2d')/2} \le (\tfrac{k-2d'}{N-k})^{(k-2d')/2}.
\end{aligned}$$

Finally, plugging this back into (18) we deduce

$$\begin{aligned}
&\mathrm{calcW}_{pis,d_{\max}}(d, v, \mathcal{S}_v, T_v, \mathcal{N}_v) \le (\tfrac{k}{N-k})^{2d'} 2^{d'} \max_{c \in \{0,\ldots,d'\}} \{B(c)\} \\
&\le 2^{d'} (\tfrac{k}{N-k})^{2d'} (\tfrac{k-2d'}{N-k})^{(k-2d')/2} = 2^{d'} \frac{k^{2d'}(k-2d')^{(k-2d')/2}}{(N-k)^{k/2+d'}}.
\end{aligned}$$

■

Proof (of Lemma 3). Let $\mathcal{N}$ be the set defined in Setting 1. By Proposition 23 (with $d = 0, v = (\text{NULL}), k'_v = k_0$) and Proposition 24,

$$\sum_{\alpha \in \mathcal{S}} \widehat{\mu}_{n,k'}(T(\alpha))^2 \leq \text{calcW}_{pis,d}(0, (\text{NULL}), \mathcal{S}, T, \mathcal{N}) \leq 2^d \frac{(k_0)^{2d}(k_0 - 2d)^{k_0/2-d}}{(N-k_0)^{k_0/2+d}}.$$

■

Possible Variants of calcW

There are many possible variants of calcW that may give better bounds in different settings, but are not used in this paper. We summarize a few below.

1. Instead of fixing the maximal depth $d_{\max}$ in advance, we can continue recursive calls from a node v as long as condition (**c2**) of Proposition 11 holds.
2. The purpose of condition (**b2**) of Definition 6 (or condition (**c2**) of Proposition 11) is to assure that T_v remains injective on the elements of $\mathcal{S}_v$ at all nodes v. This can be assured without this condition if we partition $\mathcal{S}_v$ into more subsets that result in more recursive calls (with additional information about the masks added to the recursion stack v to assure injectivity).
3. Instead of using the bound derived from Lemma 2, $(\frac{k'_v}{N-k'_v})^{k'_v/2}$, at the leaves with $d = d_{\max}$, we can use a bound derived from Lemma 1 (or a minimum of these bounds).

D Missing Proofs from Section 4

D.1 Proof of Theorem 1

Proof (of Theorem 1). We prove the inequalities of the theorem.

Proof of (6). For r odd, by Lemma 6, $\text{Var}[\nu^{(r)}_{n,q}] \leq 2^{2r} r^{2r} \frac{q^2}{N^{2r-1}}$. Both inequalities then follow by Proposition 5.

Proof of inequalities for even r. For even $r \geq 4$, we have $\text{Var}[\nu^{(r)}_{n,q}] \leq 2r^r \frac{q}{N^{r-1}}$ by Lemma 6. Combined with Proposition 5, this proves the first multi-user inequality of (8).

This variance bound gives a bound of $O_r\left(\frac{\sqrt{q}}{N^{(r-1)/2}}\right)$ on the statistical distance from uniform for $r \geq 4$, and a similar bound for $r = 2$ is obtained by Lemma 5. However, these bounds are not tight. For example, for $r = 2$, we obtain $O\left(\sqrt{\frac{q}{N}}\right)$, where the tight bound is known to be $O\left(\frac{q}{N}\right)$.

In order to improve the bound we use Proposition 3 with

$$\mathcal{S} = \mathcal{M}^n_{=1,q} = \{\alpha \in \widehat{\mathbb{F}}_2^{q \times n} \colon \#\alpha = 1\}.$$

Thus, combining (1) in Section 2 and Proposition 3 we obtain

$$\begin{aligned}2\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) &\le 2\,\mathrm{SD}(\nu^{(r)}_{n,q}, \mathbf{1}_{qn}) \\ &\le \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)| + \sqrt{\sum_{\alpha\in\mathcal{M}^n_{\ge 2,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)^2}. \end{aligned} \tag{19}$$

By Lemma 4, the first term in (19) is bounded by

$$\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)| \le \frac{2q}{\sqrt{\binom{N}{r}}} \le 2r^{r/2}\frac{q}{N^{r/2}}. \tag{20}$$

Proof of (7). For $r = 2$, by Lemma 5, the second term in (19) is bounded by

$$\sqrt{\sum_{\alpha\in\mathcal{M}^n_{\ge 2,q}} \widehat{\nu}^{(2)}_{n,q}(\alpha)^2} \le \sqrt{\frac{18q^2}{N^2}} = \frac{\sqrt{18}q}{N}.$$

Therefore, using (19) with (20) and the bound on the second term above,

$$\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[2,n]}(q) \le \frac{2q}{N} + \frac{\sqrt{4.5}q}{N} \le \frac{5q}{N}.$$

Proof of (8). The first multi-user inequality of (8) was proved above. It remains to prove the single-user and second multi-user inequalities.

For $r \ge 4$, we apply Lemma 6 to bound the second term in (19) by

$$\begin{aligned}&\sqrt{\sum_{\alpha\in\mathcal{M}^n_{\ge 2,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)^2} \le \sqrt{2^{2r+1}r^{2r}\frac{q^2}{N^{2r-2}}} = 2^{r+1/2}r^r\frac{q}{N^{r-1}} = 2^{r+1/2}r^r\frac{1}{N^{r/2-1}}\frac{q}{N^{r/2}} \\ &\le 2^{r+1/2}r^r\frac{1}{(2^{13}r)^{r/2-1}}\frac{q}{N^{r/2}} = 2^{-5.5r+13.5}r^{r/2+1}\frac{q}{N^{r/2}}.\end{aligned} \tag{21}$$

where we have used the fact that $N \ge 2^{13}r$.

Therefore, using (19) with (20) and the bound on the second term above,

$$\begin{aligned}\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) &\le r^{r/2}\frac{q}{N^{r/2}} + 2^{-5.5r+12.5}r^{r/2+1}\frac{q}{N^{r/2}} \\ &= r^{r/2}\frac{q}{N^{r/2}}(1 + 2^{-5.5r+12.5}r) \le 2r^{r/2}\frac{q}{N^{r/2}},\end{aligned}$$

where we have used the fact that for $r \ge 4$, $2^{-5.5r+12.5}r \le 1$. The second part of the multi-user bound of (8) (the second term inside min) follows from the single-user bound above by a straightforward triangle inequality.

Proof of (9). For the other direction, by Proposition 3,

$$2\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) \ge |\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)| - \sqrt{\sum_{\alpha\in\mathcal{M}^n_{\ge 2,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)^2}.$$

By Lemma 4, the first term is lower bounded as

$$|\sum_{\alpha\in\mathcal{M}^n_{=1,q}} \widehat{\nu}^{(r)}_{n,q}(\alpha)\chi_\alpha(x)| \geq \frac{3q}{2\sqrt{r\binom{N}{r}}} \geq \tfrac{3}{2}e^{-r/2}r^{(r-1)/2}\tfrac{q}{N^{r/2}}.$$

Combining with the upper bound on the second term (21) we obtain

$$\begin{aligned}
2\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{SXoP}[r,n]}(q) &\geq \tfrac{3}{2}e^{-r/2}r^{(r-1)/2}\tfrac{q}{N^{r/2}} - 2^{-5.5r+13.5}r^{r/2+1}\tfrac{q}{N^{r/2}} \\
&\geq \tfrac{3}{2}e^{-r/2}r^{(r-1)/2}\tfrac{q}{N^{r/2}}(1-2^{-5.5r+13.5}e^{r/2}r^{3/2}) \\
&\geq \tfrac{3}{2}e^{-r/2}r^{(r-1)/2}\tfrac{q}{N^{r/2}}(1-\tfrac{1}{3}) = e^{-r/2}r^{(r-1)/2}\tfrac{q}{N^{r/2}},
\end{aligned}$$

where the second inequality is based on the assumption $r \geq 4$.

■

D.2 Proof of Lemma 5

We prove simple bounds that are similar to (3) and (4), proved in the motivating example of Section 3. We then use these results to prove Lemma 5.

Proposition 25 (Bound 1 on level-k Fourier weight of $\nu^{(r)}_{n,q}$). *Assume that $rq \leq N/2$. Then, for even r*

$$\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \leq \binom{q}{k}N^k\frac{1}{\binom{N}{rk}} \leq \binom{q}{k}(rk)^k\left(\frac{rk}{N}\right)^{(r-1)k}.$$

For odd r, $\mathrm{W}^{=1}[\nu^{(r)}_{n,q}] = 0$ and

$$\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \leq \binom{q}{k}N^{k-1}\frac{1}{\binom{N}{rk}} \leq \binom{q}{k}(rk)^{k-1}\left(\frac{rk}{N}\right)^{(r-1)k+1}.$$

Proposition 26 (Bound 2 on level-k Fourier weight of $\nu^{(r)}_{n,q}$). *Assume that $rq \leq N/2$. Then,*

$$\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \leq \binom{q}{k}\left(\frac{rk}{N-rk}\right)^{rk/2} \leq \binom{q}{k}\left(\frac{2rk}{N}\right)^{rk/2}.$$

We remark that Proposition 25 gives a better bound than Proposition 26 for small values of k, while Proposition 26 is better for large values of k. However, both are very far from being tight in general.

Proof (of Proposition 25). Applying Proposition 12 and then Lemma 1,

$$\begin{aligned}
\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] &= \tbinom{q}{k}\mathrm{W}^{=k}[\nu^{(r)}_{n,k}] = \tbinom{q}{k}\sum_{\alpha\in\mathcal{M}^n_{=k,k}}\widehat{\nu}^{(r)}_{n,k}(\alpha)^2 \\
&= \tbinom{q}{k}\sum_{\alpha\in\mathcal{M}^n_{=k,k}}\widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 \leq \tbinom{q}{k}N^k(\mathrm{M}^{=rk}[\mu_{n,rk}])^2 \leq \tbinom{q}{k}N^k\tfrac{1}{\binom{N}{rk}}.
\end{aligned} \tag{22}$$

When r is odd, then by Proposition 7, $\widehat{\mu}_{n,rk}(\alpha^{\odot r}) \neq 0$ only if $0 = \oplus_{i\in[rk]}(\alpha^{\odot r})_i = \oplus_{i\in[k]}\alpha_i$, which holds only for at most N^{k-1} of the masks in $\mathcal{M}^n_{=k,k} \subset \widehat{\mathbb{F}}_2^{k\times n}$. Hence for odd r the bound is improved by a factor of N.

For the particular case of $k = 1$, we have $\oplus_{i\in[k]}\alpha_i \neq 0$ when $\#\alpha = 1$, and hence $\mathrm{W}^{=1}[\nu^{(r)}_{n,q}] = 0$. ■

Proof (of Proposition 26). Applying Proposition 12 (similarly to (22) above) and then Lemma 2,

$$\begin{aligned}\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] &= \tbinom{q}{k}\sum_{\alpha\in\mathcal{M}^n_{=k,k}}\widehat{\mu}_{n,rk}(\alpha^{\odot r})^2 \le \tbinom{q}{k}\sum_{\beta\in\mathcal{M}^n_{=rk,rk}}\widehat{\mu}_{n,rk}(\beta)^2 \\ &= \tbinom{q}{k}\mathrm{W}^{=rk}[\mu_{n,rk}] \le \tbinom{q}{k}(\tfrac{rk}{N-rk})^{rk/2}.\end{aligned}$$

■

Proof (of Lemma 5). By Proposition 26, for $r = 2$,

$$\sum_{k=2}^{q}\mathrm{W}^{=k}[\nu^{(2)}_{n,q}] \le \sum_{k=2}^{q}\tbinom{q}{k}(\tfrac{2k}{N-2k})^k = \sum_{k=2}^{q}\tbinom{q}{k}2^k(\tfrac{k}{N-2k})^k.$$

Denote $B(k) = \binom{q}{k}2^k(\frac{k}{N-2k})^k$. Assuming $k+1 \le q$, by Proposition 20,

$$\tfrac{B(k+1)}{B(k)} \le 2\tfrac{q}{k+1}\tfrac{k+1}{N-2k-2}e^{\frac{k}{k}+\frac{2k}{N-2k-2}} \le \tfrac{2q}{N-2q}e^{1+\frac{2q}{N-2q}} \le \tfrac{1}{7}e^{8/7} \le \tfrac{1}{2},$$

as $q \le N/16$. Therefore,

$$\sum_{k=2}^{q}\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \le \sum_{k=2}^{q}B(k) \le 2B(2) = 2\tbinom{q}{2}2^2\tfrac{2^2}{(N-4)^2} \le 18(\tfrac{q}{N})^2,$$

as $N \ge 100$. Combining with Proposition 25 that asserts $\mathrm{W}^{=1}[\nu^{(2)}_{n,q}] \le qN\frac{1}{\binom{N}{2}} = 2\frac{q}{N-1}$, we deduce

$$\mathrm{Var}[\nu^{(2)}_{n,q}] = \sum_{k=1}^{q}\mathrm{W}^{=k}[\nu^{(2)}_{n,q}] \le 2\tfrac{q}{N-1} + 18\tfrac{q^2}{N^2} \le 2\tfrac{q}{N-1} + \tfrac{18}{16}\tfrac{q}{N} \le \tfrac{4q}{N},$$

as $\frac{q}{N} \le \frac{1}{16}$ and $N \ge 100$. ■

D.3 Proof of Lemma 6

Proof (of Lemma 6). Consider any $r \ge 3$. By Proposition 13,

$$\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \le \tbinom{q}{k}2^{(r-2)k/2+c_{rk}}(\tfrac{rk}{N-rk})^{(r-1)k+c_{rk}}.$$

Write $N = Mr$ and define

$$B_r(k) = \tbinom{q}{k}2^{(r-2)k/2+c_{rk}}(\tfrac{rk}{N-rk})^{(r-1)k+c_{rk}} = \tbinom{q}{k}2^{(r-2)k/2+c_{rk}}(\tfrac{k}{M-k})^{(r-1)k+c_{rk}}.$$

Then, for $2 \le k \le q-2$ (noting that $c_{rk} = c_{r(k+2)}$), by Proposition 20,

$$\begin{aligned}
&\tfrac{B_r(k+2)}{B_r(k)} \le \tfrac{q^2}{(k+1)(k+2)} 2^{r-2} e^{\frac{2((r-1)k+c_{rk})}{k} + \frac{2((r-1)k+c_{rk})}{M-k-2}} (\tfrac{k+2}{M-k-2})^{2(r-1)} \\
&\le 2^{r-2} e^{2(r-1)+\frac{1}{k}+\frac{2(r-1)q}{M-q}} \tfrac{k+2}{k+1} \tfrac{q^2}{(k+2)^2} (\tfrac{k+2}{M-q})^{2(r-1)} \\
&\le 2^{r-2} \tfrac{k+2}{k+1} e^{\frac{16(r-1)}{7}+\frac{1}{k}} (\tfrac{q}{7q})^{2(r-1)} \le \tfrac{4}{3} (\tfrac{2e^{16/7}}{49})^{r-1} \tfrac{1}{2} e^{\frac{1}{2}} \le (\tfrac{1}{2})^2 \tfrac{2}{3} e^{\frac{1}{2}} \le \tfrac{1}{2},
\end{aligned}$$

where we have used the facts $k \ge 2$, $M = \frac{N}{r} \ge 8q$ and $r \ge 3$.

Therefore, using the fact that $N \ge 2^{13} r$,

$$\begin{aligned}
&\sum_{k=3}^{q} \mathrm{W}^{=k}[\nu_{n,q}^{(r)}] \le \sum_{k=3}^{q} B_r(k) \le 2B_r(3) + 2B_r(4) \\
&= 2\tbinom{q}{3} 2^{3(r-2)/2+c_{3r}} (\tfrac{3r}{N-3r})^{3(r-1)+c_{3r}} + 2\tbinom{q}{4} 2^{2(r-2)} (\tfrac{4r}{N-4r})^{4(r-1)} \\
&\le 2^{3r/2-4+c_{3r}} q^3 (\tfrac{4r}{N})^{3(r-1)+c_{3r}} + 2^{2r-7} q^4 (\tfrac{8r}{N})^{4(r-1)} \\
&= 2^{7.5r-10+3c_{3r}} r^{3r-3+c_{3r}} \tfrac{q^3}{N^{3r-3+c_{3r}}} + 2^{14r-19} r^{4r-4} \tfrac{q^4}{N^{4r-4}} \\
&= 2^{7.5r-10+3c_{3r}} r^{3r-3+c_{3r}} \tfrac{q^3}{N^{3r-3+c_{3r}}} (1 + 2^{6.5r-9-3c_{3r}} r^{r-1-c_{3r}} \tfrac{q}{N^{r-1-c_{3r}}}).
\end{aligned}$$

We have

$$\begin{aligned}
&2^{6.5r-9-3c_{3r}} r^{r-1-c_{3r}} \tfrac{q}{N^{r-1-c_{3r}}} = 2^{6.5r-9-3c_{3r}} (\tfrac{r}{N})^{r-2-c_{3r}} \tfrac{rq}{N} \\
&\le 2^{6.5r-9-3c_{3r}} 2^{-13(r-2-c_{3r})} \tfrac{1}{8} = 2^{-6.5r-9+26-3+10c_r} \\
&\le 2^{-6.5r+14+10c_r} \le 2^{-6.5r+19} \le 1,
\end{aligned}$$

where we have used the assumptions $r \ge 3$, $rq \le N/8$ and $N \ge 2^{13} r$. Plugging this into the previous inequality, we deduce

$$\sum_{k=3}^{q} \mathrm{W}^{=k}[\nu_{n,q}^{(r)}] \le 2^{7.5r-9+3c_{3r}} r^{3r-3+c_{3r}} \tfrac{q^3}{N^{3r-3+c_{3r}}}. \tag{23}$$

Assume that r is odd. Then, by Proposition 25, $\mathrm{W}^{=1}[\nu_{n,q}^{(r)}] = 0$. Moreover, by Proposition 25 (which gives a better bound on $\mathrm{W}^{=2}[\nu_{n,q}^{(r)}]$ than Proposition 13)

$$\mathrm{W}^{=2}[\nu_{n,q}^{(r)}] \le \tbinom{q}{2} (2r) \left(\tfrac{2r}{N}\right)^{2r-1} \le 2^{2r-1} r^{2r} \tfrac{q^2}{N^{2r-1}}.$$

Hence by the above results and (23) (noting that $c_{3r} = \frac{1}{2}$ and recalling that $rq \le N/8$),

$$\begin{aligned}
\mathrm{Var}[\nu_{n,q}^{(r)}] &= \sum_{k=2}^{q} \mathrm{W}^{=k}[\nu_{n,q}^{(r)}] \le 2^{2r-1} r^{2r} \tfrac{q^2}{N^{2r-1}} + 2^{7.5r-9+1.5} r^{3r-3+0.5} \tfrac{q^3}{N^{3r-3+0.5}} \\
&= 2^{2r-1} r^{2r} \tfrac{q^2}{N^{2r-1}} (1 + 2^{5.5r-6.5} r^{r-2.5} \tfrac{q}{N^{r-1.5}}).
\end{aligned}$$

We have

$$2^{5.5r-6.5}r^{r-2.5}\tfrac{q}{N^{r-1.5}} = 2^{5.5r-6.5}(\tfrac{r}{N})^{r-2.5}\tfrac{q}{N} \le 2^{5.5r-6.5}2^{-13(r-2.5)}\tfrac{1}{8\cdot 3}$$
$$\le 2^{-7.5r-6.5+32.5-4} = 2^{-7.5r+22} \le 1,$$

where we have used the assumptions $r \ge 3$, $rq \le N/8$ and $N \ge 2^{13}r$. Plugging this into the previous inequality, we deduce the claimed inequality

$$\mathrm{Var}[\nu^{(r)}_{n,q}] \le 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-1}}.$$

For even $r \ge 4$, by Proposition 25, $\mathrm{W}^{=2}[\nu^{(r)}_{n,q}] \le \binom{q}{2}(2r)^2\left(\frac{2r}{N}\right)^{2r-2} = 2^{2r}r^{2r}\frac{q^2}{N^{2r-2}}$. Therefore, by the above inequality and (23) (with $c_{rk} = 0$),

$$\sum_{k=2}^{q}\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \le 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-2}} + 2^{7.5r-9}r^{3r-3}\tfrac{q^3}{N^{3r-3}}$$
$$= 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-2}}(1 + 2^{5.5r-9}r^{r-3}\tfrac{q}{N^{r-1}}) = 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-2}}(1 + 2^{5.5r-9}(\tfrac{r}{N})^{r-3}\tfrac{q}{N}\tfrac{1}{N})$$
$$\le 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-2}}(1 + 2^{5.5r-9}2^{-13(r-3)}\tfrac{1}{8\cdot 4}2^{-13}) \le 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-2}}(1 + 2^{-7.5r-9+39-5-13})$$
$$= 2^{2r}r^{2r}\tfrac{q^2}{N^{2r-2}}(1 + 2^{-7.5r+12}) \le 2^{2r+1}r^{2r}\tfrac{q^2}{N^{2r-2}},$$

where we have used the assumptions $r \ge 4$, $rq \le N/8$ and $N \ge 2^{13}r$.

Finally, by Proposition 25 and the above inequality (again using the assumptions $r \ge 4$, $rq \le N/8$ and $N \ge 2^{13}r$),

$$\mathrm{Var}[\nu^{(r)}_{n,q}] = \sum_{k=1}^{q}\mathrm{W}^{=k}[\nu^{(r)}_{n,q}] \le qr^r\tfrac{1}{N^{r-1}} + 2^{2r+1}r^{2r}\tfrac{q^2}{N^{2r-2}}$$
$$= r^r\tfrac{q}{N^{r-1}}(1 + 2^{2r+1}r^r\tfrac{q}{N^{r-1}}) \le r^r\tfrac{q}{N^{r-1}}(1 + 2^{2r+1}r\tfrac{rq}{N}(\tfrac{r}{N})^{r-2})$$
$$\le r^r\tfrac{q}{N^{r-1}}(1 + 2^{2r+1}r\tfrac{1}{8}2^{-13(r-2)}) = r^r\tfrac{q}{N^{r-1}}(1 + 2^{-11r+1-3+26}r)$$
$$= r^r\tfrac{q}{N^{r-1}}(1 + 2^{-11r+24}r) \le 2r^r\tfrac{q}{N^{r-1}}.$$

■

E Missing Proofs from Section 5

Proof (of Lemma 7). By Proposition 15, $\mathrm{W}^{=1}[\xi^{(L)}_{n,q}] = 0$. Hence, by Proposition 16,

$$\mathrm{Var}[\xi^{(L)}_{n,q}] = \sum_{k=1}^{q}\mathrm{W}^{=k}[\xi^{(L)}_{n,q}] = \sum_{k=2}^{q}\mathrm{W}^{=k}[\xi^{(L)}_{n,q}] \le \sum_{k=2}^{q}\tbinom{q}{k}2^{3k/2+3c_k}\tfrac{k^{k+2c_k}(k-2c_k)^{k/2-c_k}}{(N-2k)^{3k/2+c_k}}.$$

Denote $B(k) = \binom{q}{k}2^{3k/2+3c_k}\frac{k^{k+2c_k}(k-2c_k)^{k/2-c_k}}{(N-2k)^{3k/2+c_k}}$. Noting that $c_k = c_{k+2}$ and assuming $2 \le k \le q-2$, by Proposition 20,

$$\tfrac{B(k+2)}{B(k)} \le \tfrac{q^2}{(k+1)(k+2)}2^3 e^{\frac{2(k+2c_k)}{k}+\frac{2(k/2-c_k)}{k-2c_k}+\frac{2^2(3k/2+c_k)}{N-2k-4}\frac{(k+2)^2(k+2-2c_k)}{(N-2k-4)^3}}$$
$$\le 8q^2e^{2+\frac{4c_k}{k}+1+\frac{6k+2}{N-2k-4}\frac{k+2}{k+1}\frac{k+2-2c_k}{(N-2k-4)^3}} \le 8\tfrac{4}{3}e^{3+\frac{2}{3}+\frac{6q}{14q}}q^2\tfrac{q}{(14q)^3} \le 8\tfrac{4}{3}e^{4.1}14^{-3} \le \tfrac{1}{3},$$

where we have used the facts $k \geq 2$, $\frac{4c_k}{k} \leq \frac{2}{3}$ and $q \leq N/16$. Therefore,

$$\sum_{k=2}^{q} \mathrm{W}^{=k}[\xi^{(L)}_{n,q}] \leq \tfrac{1}{1-\frac{1}{3}}(B(2)+B(3)) \leq \tfrac{3}{2}\tbinom{q}{2}2^3\tfrac{2^3}{(N-4)^3} + \tfrac{3}{2}\tbinom{q}{3}2^6\tfrac{3^4\cdot 2}{(N-6)^5}$$
$$\leq \tfrac{48q^2}{(N-4)^3} + \tfrac{2^5 3^4 q^3}{(N-6)^5} \leq 60\tfrac{q^2}{N^3} + 2^6 3^4 \tfrac{q}{N}\tfrac{1}{N}\tfrac{q^2}{N^3} \leq 60\tfrac{q^2}{N^3} + \tfrac{q^2}{N^3} \leq 2^6\tfrac{q^2}{N^3},$$

where we have used the assumptions that $q \leq N/16$ and $N \geq 2^{10}$. ■

F Missing Proofs from Section F

Proof (of Lemma 8). By Proposition 17,

$$\mathrm{W}^{=1}[\xi^{(L)}_{n,2,q}] = \tbinom{q}{1}\mathrm{W}^{=1}[\xi^{(L)}_{n,2,1}] = q \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{2n}\\ \#\alpha=1}} \widehat{\xi}^{(L)}_{n,2,1}(\alpha)^2 = q \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{2n}\\ \#\alpha=1}} \widehat{\mu}_{n,3}(t(\alpha))^2.$$

Let $\alpha = (\alpha^{(1)}, \alpha^{(2)}) \in \widehat{\mathbb{F}}_2^{2n}$, hence $t(\alpha) = (\alpha^{(1)}, L^{\mathrm{T}}(\alpha^{(2)}), \alpha^{(2)} \oplus L^{\mathrm{T}}(\alpha^{(1)}))$. By Proposition 7, we have $\widehat{\mu}_{n,3}(t(\alpha)) \neq 0$ only if $\alpha^{(1)} \oplus L^{\mathrm{T}}(\alpha^{(2)}) \oplus \alpha^{(2)} \oplus L^{\mathrm{T}}(\alpha^{(1)}) = 0$. In this case, $\alpha^1 \oplus L^{\mathrm{T}}(\alpha^{(1)}) = \alpha^{(2)} \oplus L^{\mathrm{T}}(\alpha^{(2)})$ and thus $(L')^{\mathrm{T}}(\alpha^{(1)}) = (L')^{\mathrm{T}}(\alpha^{(2)})$. Hence, by invertibility of $(L')^{\mathrm{T}}$, $\alpha^{(1)} = \alpha^{(2)}$, which implies that $t(\alpha) = (\alpha^{(1)}, L^{\mathrm{T}}(\alpha^{(1)}), \alpha^{(1)} \oplus L^{\mathrm{T}}(\alpha^{(1)}))$. In particular, since L^{T} and $(L')^{\mathrm{T}}$ are invertible then $\#t(\alpha) = 3$.

By Proposition 8, for every $\alpha^{(1)} \neq 0$, exact computation gives $|\widehat{\mu}_{n,3}(\alpha^{(1)}, L^{\mathrm{T}}(\alpha^{(1)}), \alpha^{(1)} \oplus L^{\mathrm{T}}(\alpha^{(1)}))| = \frac{2}{N-2}\frac{1}{N-1}$. Since $\alpha^{(1)} \in \widehat{\mathbb{F}}_2^n$ can attain $N-1$ non-zero values, we conclude that

$$\mathrm{W}^{=1}[\xi^{(L)}_{n,2,q}] = q \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{2n}\\ \#\alpha=1}} \widehat{\mu}_{n,3}(t(\alpha))^2 = q(N-1)\tfrac{4}{(N-1)^2(N-2)^2} = \tfrac{4q}{(N-1)(N-2)^2}.$$

■

Proof (of Proposition 19). By Proposition 17 and Proposition 18,

$$\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] = \tbinom{q}{k}\mathrm{W}^{=k}[\xi^{(L)}_{n,2,k}] = \tbinom{q}{k} \sum_{\alpha\in\mathcal{M}^{2n}_{=k,k}} \widehat{\xi}^{(L)}_{n,2,k}(\alpha)^2 = \tbinom{q}{k} \sum_{\alpha\in\mathcal{M}^{2n}_{=k,k}} \widehat{\mu}_{n,3k}(t(\alpha))^2$$
$$= \tbinom{q}{k} \sum_{m=0}^{k} \sum_{\substack{\alpha\in\mathcal{M}^{2n}_{=k,k}\\ \#_3\alpha=m}} \widehat{\mu}_{n,3k}(t(\alpha))^2.$$

For $m \in \{0,\ldots,k\}$, denote $\mathcal{S}^{(k,m)} = \{\alpha \in \widehat{\mathbb{F}}_2^{k\times 2n} \colon \#\alpha = k \wedge \#_3\alpha = m\}$. We have shown that

$$\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] = \tbinom{q}{k} \sum_{m=0}^{k} \sum_{\alpha\in\mathcal{S}^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2. \tag{24}$$

Fix a pair (k,m). We would like to upper bound $\sum_{\alpha\in\mathcal{S}^{(k,m)}}\widehat{\mu}_{n,3k}(t(\alpha))^2$ using Lemma 3. For this purpose, according to restriction (**a2**) of Setting 1, we first need to partition the set $\mathcal{S}^{(k,m)}$ into subsets such that the (transformed) masks in each subset, $t(\alpha)$, share the same non-zero entries (over $\widehat{\mathbb{F}}_2^n$).

We now analyze this partition of $\mathcal{S}^{(k,m)}$. By proposition 18, every $i\in[k]$ with $\#t(\alpha_i)=2$ has 3 possible structures that determine which 2 of its 3 entries are non-zero over $\widehat{\mathbb{F}}_2^n$. For every $\alpha\in\mathcal{S}^{(k,m)}$, there are $k-m$ such indices i with $\#t(\alpha_i)=2$. Therefore, there are $\binom{k}{m}3^{k-m}$ possible non-zero index sets (with non-zero values over $\widehat{\mathbb{F}}_2^n$). Note that every such index set has size equal to $\#t(\alpha)=2k+m$ for $\alpha\in\mathcal{S}^{(k,m)}$.

Denote by $\mathit{\Lambda}_{k,m}$ the collection of these $\binom{k}{m}3^{k-m}$ non-zero index sets, where every $\lambda\subseteq[k]\times[3]$ is of size $2k+m$. We thus partition $\mathcal{S}^{(k,m)}$ into $\binom{k}{m}3^{k-m}$ subsets, denoted $\{\mathcal{S}_\lambda^{(k,m)}\}_{\lambda\in\mathit{\Lambda}_{k,m}}$, each with common non-zero entries of $t(\alpha)$ over $\widehat{\mathbb{F}}_2^n$. Concretely, $\alpha\in\mathcal{S}^{(k,m)}$ satisfies $\alpha\in\mathcal{S}_\lambda^{(k,m)}$ if for every $(i,j)\in[k]\times[3]$, $t(\alpha_i)_j\neq 0$ if and only if $(i,j)\in\lambda$. We have

$$\sum_{\alpha\in\mathcal{S}^{(k,m)}}\widehat{\mu}_{n,3k}(t(\alpha))^2=\sum_{\lambda\in\mathit{\Lambda}_{k,m}}\sum_{\alpha\in\mathcal{S}_\lambda^{(k,m)}}\widehat{\mu}_{n,3k}(t(\alpha))^2. \tag{25}$$

Applying Lemma 3. Fix any $\lambda\in\mathit{\Lambda}_{k,m}$. We now use Lemma 3 to bound $\sum_{\alpha\in\mathcal{S}_\lambda^{(k,m)}}\widehat{\mu}_{n,3k}(t(\alpha))^2$. For this purpose, let $\mathcal{S}=\mathcal{S}_\lambda^{(k,m)}$ and define $T(\alpha)=T_k(\alpha)=t(\alpha)$ for every $\alpha\in\mathcal{S}_\lambda^{(k,m)}$. In this case, $k'=3k$ and $k_0=2k+m$.

The PIS *pis* resembles the one defined in the proof of Proposition 16. Given a node v, we say that an index $\ell\in[k]$ is redundant if all 3 indices $3\ell-2,3\ell-1$ and 3ℓ are unaltered by Definition 7.

At a given node v, let $\ell\in[k]$ be the largest redundant index. The PIS *pis* will select as primary index the smallest index in the triplet $\{3\ell-2,3\ell-1,3\ell\}$ that is in $\mathcal{N}_v$.

The recursion is executed up to depth $d=\lceil k/2\rceil$. As in the proof of Proposition 16, a redundant index is guaranteed to exist up to depth $d-1=\lceil k/2\rceil-1$ and *pis* is well-defined.

In order to invoke Lemma 3, we prove that the two conditions of Proposition 11 hold. First, by our definition of $\mathcal{S}=\mathcal{S}_\lambda^{(k,m)}$, the pair $(\mathcal{S},T)$ defined above satisfies the restrictions of Setting 1, and condition (**c1**) holds (note that $2d=2\lceil k/2\rceil\leq k+1<2k$, as $k\geq 2$).

It remains to prove condition (**c2**). Specifically, for a node v such that $j=pis(v)$ and $\alpha\in\mathcal{S}_v$, we prove that $T_v(\alpha)_j$ can be computed from $T_{v_i}(\alpha)=T_v(\alpha)^{\oplus(j,i)}$ (where i is a secondary index).

Note that if ℓ is redundant, then for every $\alpha\in\mathcal{S}_v$, the elements $T_v(\alpha)_{3\ell-2},T_v(\alpha)_{3\ell-1},T_v(\alpha)_{3\ell}$ are equal to those of $T(\alpha_\ell)=t(\alpha_\ell)$. Thus, depending on v, according to Proposition 18 we have 4 possibilities for the zero entries of $T_v(\alpha)_{3\ell-2},T_v(\alpha)_{3\ell-1},T_v(\alpha)_{3\ell}$.

First, assume that $\#t(\alpha_\ell)=3$, namely $t(\alpha_\ell)=(\alpha_\ell^{(1)},L^{\mathrm{T}}(\alpha_\ell^{(2)}),\alpha_\ell^{(2)}\oplus L^{\mathrm{T}}(\alpha_\ell^{(1)}))$ with all 3 entries non-zero. Then the first index with value $\alpha_\ell^{(1)}$ is selected as

primary index ($j = 3\ell-2$). We need to verify that $T_v(\alpha)_j = \alpha_\ell^{(1)}$ can be uniquely recovered from $T_v(\alpha)^{\oplus(j,i)}$ regardless of the secondary index i, namely that it can be recovered from the values of either

$$\begin{aligned}
&(1)\ L^{\mathrm{T}}(\alpha_\ell^{(2)}), \alpha_\ell^{(2)} \oplus L^{\mathrm{T}}(\alpha_\ell^{(1)}), && \text{in case } i \notin \{3\ell-1, 3\ell\},\\
&(2)\ L^{\mathrm{T}}(\alpha_\ell^{(2)}) \oplus \alpha_\ell^{(1)}, \alpha_\ell^{(2)} \oplus L^{\mathrm{T}}(\alpha_\ell^{(1)}), && \text{in case } i = 3\ell-1, \text{ or}\\
&(3)\ L^{\mathrm{T}}(\alpha_\ell^{(2)}), \alpha_\ell^{(2)} \oplus L^{\mathrm{T}}(\alpha_\ell^{(1)}) \oplus \alpha_\ell^{(1)}, && \text{in case } i = 3\ell.
\end{aligned}$$

In case (1), $\alpha_\ell^{(1)}$ can be recovered after computing $\alpha_\ell^{(2)}$ due to the invertibility of L^{T}. In case (2), we apply L^{T} to the second value and XOR to the first to obtain the value of

$$\begin{aligned}
(L^2)^{\mathrm{T}}(\alpha_\ell^{(1)}) \oplus \alpha_\ell^{(1)} &= (L^2)^{\mathrm{T}}(\alpha_\ell^{(1)}) \oplus L^{\mathrm{T}}(\alpha_\ell^{(1)}) \oplus L^{\mathrm{T}}(\alpha_\ell^{(1)}) \oplus \alpha_\ell^{(1)}\\
&= L^{\mathrm{T}}((L')^{\mathrm{T}}(\alpha_\ell^{(1)})) \oplus (L')^{\mathrm{T}}(\alpha_\ell^{(1)}) = ((L')^2)^{\mathrm{T}}(\alpha_\ell^{(1)}).
\end{aligned}$$

Since $((L')^2)^{\mathrm{T}}$ is invertible, $\alpha_\ell^{(1)}$ can be uniquely recovered. In case (3), we deduce $\alpha_\ell^{(2)}$ and then $L^{\mathrm{T}}(\alpha_\ell^{(1)}) \oplus \alpha_\ell^{(1)} = (L')^{\mathrm{T}}(\alpha_\ell^{(1)})$, from which we recover $\alpha_\ell^{(1)}$ since $(L')^{\mathrm{T}}$ is invertible.

Second, if $\#t(\alpha_\ell) = 2$, then according to Proposition 18,

$$t(\alpha_\ell) \in \{(0, L^{\mathrm{T}}(\alpha_\ell^{(2)}), \alpha_\ell^{(2)}), (\alpha_\ell^{(1)}, 0, L^{\mathrm{T}}(\alpha_\ell^{(1)})), (\alpha_\ell^{(1)}, (L^2)^{\mathrm{T}}(\alpha_\ell^{(1)}), 0)\}.$$

By similar calculation to the case $\#t(\alpha_\ell) = 3$, one can verify that in each of the 3 cases above the first non-zero entry (the value of the primary index) can be recovered from $T_v(\alpha)^{\oplus(j,i)}$ regardless of the secondary index.

We conclude that the two conditions of Proposition 11 hold. Applying our framework of Lemma 3 (with $d = \lceil k/2 \rceil = k/2 + c_k$, $k_0 = 2k + m$), we obtain

$$\sum_{\alpha \in \mathcal{S}_\lambda^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2 \le 2^d \frac{(k_0)^{2d}(k_0-2d)^{k_0/2-d}}{(N-k_0)^{k_0/2+d}} = 2^{k/2+c_k} \frac{(2k+m)^{k+2c_k}(k+m-2c_k)^{k/2+m/2-c_k}}{(N-2k-m)^{3k/2+m/2+c_k}}.$$

We recall that $|\Lambda_{k,m}| = \binom{k}{m}3^{k-m}$ and $\sum_{m=0}^k \binom{k}{m}3^{k-m} = 4^k$. Using (24), (25) and the inequality above we deduce

$$\mathrm{W}^{=k}[\xi_{n,2,q}^{(L)}] = \tbinom{q}{k} \sum_{m=0}^{k} \sum_{\alpha \in \mathcal{S}^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2 = \tbinom{q}{k} \sum_{m=0}^{k} \sum_{\lambda \in \Lambda_{k,m}} \sum_{\alpha \in \mathcal{S}_\lambda^{(k,m)}} \widehat{\mu}_{n,3k}(t(\alpha))^2$$

$$\le \tbinom{q}{k} \sum_{m=0}^{k} |\Lambda_{k,m}| 2^{k/2+c_k} \frac{(2k+m)^{k+2c_k}(k+m-2c_k)^{k/2+m/2-c_k}}{(N-2k-m)^{3k/2+m/2+c_k}}$$

$$\le \tbinom{q}{k} 4^k \max_{m \in \{0,1,\ldots,k\}} \{2^{k/2+c_k} \frac{(2k+m)^{k+2c_k}(k+m)^{k/2+m/2-c_k}}{(N-2k-m)^{3k/2+m/2+c_k}}\}$$

$$= \tbinom{q}{k} 2^{5k/2+c_k} \max_{m \in \{0,1,\ldots,k\}} \{\frac{(2k+m)^{k+2c_k}(k+m)^{k/2+m/2-c_k}}{(N-2k-m)^{3k/2+m/2+c_k}}\}.$$

Denote $B(m) = \frac{(2k+m)^{k+2c_k}(k+m)^{k/2+m/2-c_k}}{(N-2k-m)^{3k/2+m/2+c_k}}$. Then, assuming $m+1 \le k \le q \le N/32$, by Proposition 20,

$$\frac{B(m+1)}{B(m)} \le e^{\frac{k+2c_k}{2k+m}+\frac{k/2+m/2-c_k}{k+m}+\frac{3k/2+m/2+c_k}{N-2k-m-1}}\left(\frac{k+m+1}{N-2k-m-1}\right)^{1/2}$$
$$\le e^{1+c_k\left(\frac{2}{2k+m}-\frac{1}{k+m}\right)+\frac{2q}{N-3q}}\left(\frac{2q}{N-3q}\right)^{1/2} \le e^{1+\frac{c_k}{k}+\frac{2}{29}}\left(\frac{2}{29}\right)^{1/2} \le e^{\frac{7}{6}+\frac{2}{29}}\left(\frac{2}{29}\right)^{1/2} \le 1,$$

where we also used the facts that $k \ge 2$ and $c_2 = 0$, hence $\frac{c_k}{k} \le \frac{1}{6}$. Therefore,

$$\mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \le \binom{q}{k} 2^{5k/2+c_k} \max_{m\in\{0,1,\ldots,k\}} B(m) \le \binom{q}{k} 2^{5k/2+c_k} B(0)$$
$$= \binom{q}{k} 2^{5k/2+c_k} \frac{(2k)^{k+2c_k}(k)^{k/2-c_k}}{(N-2k)^{3k/2+c_k}} = \binom{q}{k} 2^{7k/2+3c_k}\left(\frac{k}{N-2k}\right)^{3k/2+c_k}.$$

■

Proof (of Lemma 9). Applying Proposition 19,

$$\sum_{k=2}^{q} \mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \le \sum_{k=2}^{q} \binom{q}{k} 2^{7k/2+3+3c_k}\left(\frac{k}{N-2k}\right)^{3k/2+c_k}.$$

Denote $B(k) = \binom{q}{k} 2^{7k/2+3c_k}\left(\frac{k}{N-2k}\right)^{3k/2+c_k}$ and note that $c_{k+2} = c_k$. Then, assuming $k+2 \le q \le N/32$, and recalling that $k \ge 2$ (and $c_2 = 0$), by Proposition 20,

$$\frac{B(k+2)}{B(k)} \le \frac{q^2}{(k+1)(k+2)} 2^7 e^{\frac{3k+2c_k}{k}+\frac{2(3k+2c_k)}{N-2k-4}}\left(\frac{k+2}{N-2k-4}\right)^3 \le 2^7 e^{3+\frac{1}{3}+\frac{6q}{N-2q}} \frac{k+2}{k+1} \frac{q^2(k+2)}{(N-2q)^3}$$
$$\le 2^7 \tfrac{4}{3} e^{3+\frac{1}{3}+\frac{6}{30}} \frac{q^3}{(30q)^3} \le 2^7 \tfrac{4}{3} e^4 (30)^{-3} \le \tfrac{1}{2}.$$

Therefore, using the facts that $N \ge 2^{10}$ and $q \le N/32$,

$$\sum_{k=2}^{q} \mathrm{W}^{=k}[\xi^{(L)}_{n,2,q}] \le \sum_{k=2}^{q} B(k) \le 2B(2) + 2B(3) \le 2\binom{q}{2} 2^7 \left(\frac{2}{N-4}\right)^3 + 2\binom{q}{3} 2^{12}\left(\frac{3}{N-6}\right)^5$$
$$\le \frac{2^{10}q^2}{(N-4)^3} + \frac{2^{12}3^4q^3}{(N-6)^5} \le \frac{2^{10.3}q^2}{N^3} + \frac{2^{20}q^2}{N^3}\frac{q}{N}\frac{1}{N} \le \frac{2^{10.3}q^2}{N^3} + \frac{2^{20}q^2}{N^3}\frac{1}{32}\frac{1}{2^{10}} \le \frac{2^{10.5}q^2}{N^3}.$$

■