Tight Indistinguishability Bounds for the XOR of Independent Random Permutations by Fourier Analysis

Itai Dinur

Department of Computer Science, Ben-Gurion University, Israel
dinuri@bgu.ac.il

Abstract. The XOR of two independent permutations (XoP) is a well-known construction for achieving security beyond the birthday bound when implementing a pseudorandom function using a block cipher (i.e., a pseudorandom permutation). The idealized construction (where the permutations are uniformly chosen and independent) and its variants have been extensively analyzed over nearly 25 years.

The best-known asymptotic information-theoretic indistinguishability bound for the XoP construction is $O(q/2^{1.5n})$, derived by Eberhard in 2017, where q is the number of queries and n is the block length.

A generalization of the XoP construction outputs the XOR of $r \geq 2$ independent permutations, and has also received significant attention in both the single-user and multi-user settings. In particular, for $r = 3$, the best-known bound (obtained by Choi et al. [ASIACRYPT'22]) is about $q^2/2^{2.5n}$ in the single-user setting and $\sqrt{u}q_{\max}^2/2^{2.5n}$ in the multi-user setting (where u is the number of users and $q_{\max}$ is the number of queries per user).

In this paper, we prove an indistinguishability bound of $q/2^{(r-0.5)n}$ for the (generalized) XoP construction in the single-user setting, and a bound of $\sqrt{u}q_{\max}/2^{(r-0.5)n}$ in the multi-user setting. In particular, for $r = 2$, we obtain the bounds $q/2^{1.5n}$ and $\sqrt{u}q_{\max}/2^{1.5n}$ in single-user and multi-user settings, respectively. For $r = 3$ the corresponding bounds are $q/2^{2.5n}$ and $\sqrt{u}q_{\max}/2^{2.5n}$. All of these bounds hold assuming $q < 2^n/2$ (or $q_{\max} < 2^n/2$).

Compared to previous works, we improve all the best-known bounds for the (generalized) XoP construction in the multi-user setting, and the best-known bounds for the generalized XoP construction for $r \geq 3$ in the single-user setting (assuming $q \geq 2^{n/2}$). For the basic two-permutation XoP construction in the single-user setting, our concrete bound of $q/2^{1.5n}$ stands in contrast to the asymptotic bound of $O(q/2^{1.5n})$ by Eberhard. Since all of our bounds are matched (up to constant factors) for $q > 2^{n/2}$ by attacks published by Patarin in 2008 (and their generalizations to the multi-user setting), they are all tight.

We obtain our results by Fourier analysis of Boolean functions. Most of our technical work involves bounding (sums of) Fourier coefficients of the density function associated with sampling without replacement. While the proof of Eberhard relies on similar bounds, our proof is elementary and simpler.

1 Introduction

Many cryptosystems such as encryption modes, MAC algorithms and authenticated encryption schemes require pseudorandom functions to achieve security. However, in practice, pseudorandom functions are typically implemented by block ciphers, which are pseudorandom permutations that are only secure up to the birthday bound of $q = 2^{n/2}$ queries (where n is the block length). In order to overcome this limitation, achieving security beyond the birthday bound has become a prominent research area, initiated by the seminal papers by Bellare, Krovetz, and Rogaway [2], and by Hall, Wagner, Kelsey, and Schneier [18].

1.1 The XoP Construction

One of the main constructions analyzed in the literature for achieving security beyond the birthday bound is the XOR of permutations (XoP) construction, which has two main variants. One variant uses two permutations $\pi_1, \pi_2 : \{0,1\}^n \mapsto \{0,1\}^n$ to define $f_{\pi_1,\pi_2} : \{0,1\}^n \mapsto \{0,1\}^n$ by $f(x) = \pi_1(x) \oplus \pi_2(x)$. In practice, π_1 and π_2 are implemented using a block cipher, instantiated with independent keys. In the following, we simply refer to this variant as the XoP construction. Another variant uses a single permutation $\pi : \{0,1\}^n \mapsto \{0,1\}^n$ to define $f_\pi : \{0,1\}^{n-1} \mapsto \{0,1\}^n$ by $f(x) = \pi(0\|x) \oplus \pi(1\|x)$ (where $\|$ denotes concatenation). We refer to this construction as a single-permutation XoP construction. Similarly to the two-permutation 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.

We note that there are other variants of the XoP construction defined in the literature that we do not deal with in this paper. For example, the recent result [16] by Gunsing et al. analyzes a variant where the underlying permutations are public and the adversary is allowed to query them. Previous works that analyze additional variants include [3,4,7,17].

Previous results. There have been several works on the security of the (idealized) XoP construction [1,20,25,27], analyzing one or both of its variants. Yet, a simple and verifiable proof that the XoP construction variants achieve security up to $q = O(2^n)$ queries was only published in 2017 in a paper by Dai, Hoang, and Tessaro [11]. Specifically, [11] proved that any adversary that makes q queries to the (two-permutation) XoP construction can distinguish it from a truly random function with advantage of at most (about) $\frac{q^{1.5}}{2^{1.5n}}$.

Independently, in [13, Thm. 1.5] Eberhard proved a substantially better indistinguishability bound of $O(\frac{q}{2^{1.5n}})$, relying and extending results of [14] in additive combinatorics. The bound was given in asymptotic form with an unspecified constant. An additional paper that analyzed the XoP construction is [12].

For the single-permutation XoP variant, the distinguishing advantage was bounded in [11,12] by about $\frac{q}{2^n}$. The works of [8,12], essentially confirm (and

improve) the results obtained earlier by Patarin [25,27] (using the so-called mirror theory technique).

The indistinguishability bound $\frac{q}{2^n}$ for the single permutation XoP construction variant is essentially tight. Indeed, it is matched by a simple attack based on the observation that since π is a permutation, for all $x \in \{0,1\}^{n-1}$, $f(x) = \pi(0\|x) \oplus \pi(1\|x) \neq \vec{0}$, while for a random function, $\vec{0}$ is output with probability 2^{-n} for each query.

The attack above does not work for the variant where the permutations are independent, and indeed the bound $O(\frac{q}{2^{1.5n}})$ of [13] for this variant is much better (particularly when q is large). This bound is matched by an attack published by Patarin [26,28], which obtains distinguishing advantage of about $\frac{q}{2^{1.5n}}$, assuming $q = O(2^n)$. Note that if $q = 2^n$, then the distinguishing advantage is close to 1 since the XOR of the outputs of all inputs to $f(x) = \pi_1(x) \oplus \pi_2(x)$ is $\vec{0}$.

Multi-user setting. The XoP construction also recently received attention in the multi-user setting in [6,7]. A trivial extension of the result in [13] gives a bound of $O(\frac{u \cdot q_{\max}}{2^{1.5n}})$ in the multi-user setting, where u is the number of users and $q_{\max}$ is the allowed number of queries to each user.

In terms of attacks, one can generically extend the attacks by Patarin [26,28] to the multi-user setting by independently applying the single-user attacker to each user, and then taking the majority of answers (which attempt to deduce whether the oracle is the XoP construction or a random function). Applying a standard Chernoff bound, the attack achieves a distinguishing advantage of about $\frac{\sqrt{u} q_{\max}}{2^{1.5n}}$.

Generalized XoP construction. A natural generalization of the XOR construction defines f by XORing together $r \geq 2$ permutations, where r is a (small) parameter. As in the case of $r = 2$, the generalized construction also has two variants, but we focus on the case where all permutations are independent.

Previous results. This construction was first analyzed by Lucks [20] and this analysis was improved by Cogliati, Lampe, and Patarin [9], who proved security up to roughly $2^{rn/(r+1)}$ queries (also see [22]). More recently, this analysis has been improved in [11], which obtained an indistinguishability bound to about $(\frac{q}{2^n})^{1.5\lfloor r/2 \rfloor}$ using the generic amplification technique of Maurer, Pietrzak, and Renner [21]. The specific case of $k = 3$ was analyzed in [7] by Choi et al., who proved an indistinguishability bound of about $\frac{\sqrt{u} q_{\max}^2}{2^{2.5n}}$ in the multi-user setting.

On the other hand, the best known attacks on the generalized XoP construction, published in [26,28], obtained distinguishing advantage of about $\frac{q}{2^{(r-0.5)n}}$. One can also consider attacks on the generalized XoP construction in the multi-user setting. Similarly to the case of $r = 2$, the best known attack is the generic extension of the single-user attack by Patarin [26,28] to the multi-user setting, which achieves advantage of about $\frac{\sqrt{u} \cdot q_{\max}}{2^{(r-0.5)n}}$.

1.2 Our Contribution

Our results. In this paper, we prove an indistinguishability bound of $\frac{q}{2^{(r-0.5)n}}$ for the (generalized) XoP construction in the single-user setting, and a bound of $\frac{\sqrt{u}q_{\max}}{2^{(r-0.5)n}}$ in the multi-user setting. Specifically, for the basic two-permutation XoP construction, we obtain a bound of $\frac{q}{2^{1.5n}}$ in the single-user setting and $\frac{\sqrt{u}q_{\max}}{2^{1.5n}}$ in the multi-user settings. All of these bounds have no hidden constants. They hold as long as $q < 2^n/2$ (or $q_{\max} < 2^n/2$ in the multi-user setting), assuming $2^n \geq 1000$.

Compared to previous results, we improve all the best-known bounds for the (generalized) XoP construction in the multi-user setting, and the best-known bounds for the generalized XoP construction for $r \geq 3$ in the single-user setting (assuming $q \geq 2^{n/2}$). For the basic XoP construction (with $r = 2$), our concrete bound of $q/2^{1.5n}$ in the single-user setting stands in contrast to the asymptotic bound of $O(q/2^{1.5n})$, derived in [13].

All of our bounds are tight assuming $q \leq 2^{n/2}$, as they match (up to constant factors) the single-user attacks published by Patarin in [26,28], as well as their trivial generalization to the multi-user setting.

Our techniques. Similarly to [13,14], the main framework that we use to obtain our results is Fourier analysis (of Boolean functions). This is a standard tool for analyzing probability distributions in mathematics, yet it is not commonly used as a main framework in information-theoretic security proofs in symmetric-key cryptography. For example, [5] used Fourier analysis as an auxiliary tool in order to prove an internal lemma, but not as the main framework. The application of Fourier analysis in the more recent work [19] is somewhat more related to ours. We summarize the main ideas of our proof below.

First, the distinguishing advantage of the adversary is bounded by the statistical distance between the distribution generated by the XoP construction and the uniform distribution. Consider a sample in $\mathbb{F}_2^{q\times n}$ composed of q elements in $\{0,1\}^n$, generated by the XoP construction. We can bound the statistical distance of this distribution from the uniform distribution in the "Fourier domain" by bounding the bias (i.e., Fourier coefficient) of each of the $2^{q\cdot n}$ possible masks (i.e., linear equations over $\mathbb{F}_2$) applied to the bits of the sample. To gain intuition, note that for the uniform distribution over $\mathbb{F}_2^{q\times n}$, all non-empty linear equations have 0 bias (i.e., hold with probability 1/2), and thus a distribution that is close to uniform has biases (Fourier coefficients) that are very close to 0.

Our task is thus to bound the Fourier coefficients for the distribution function generated by the XoP construction.[1] Next, we use standard techniques to reduce this task to the task of bounding the Fourier coefficients for the distribution generated by the underlying primitive, namely, a random permutation. Specifically, we consider k elements (for any $1 \leq k \leq q$) drawn uniformly without replace-

[1] More accurately, the task is to bound the Fourier coefficients for the normalized distribution function (i.e., density function) generated by the XoP construction.

ment. Our goal is reduced to bounding 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 Fourier coefficients of level k.

Intuitively, level-k (Fourier) weight is a measure of dependence between k elements drawn from the distribution. For example, the level-k Fourier weight of a q-wise uniform distribution is 0 for any $1 \leq k \leq q$. We remark that calculating the above two Fourier quantities for various levels has the additional advantage of hinting at the best attack strategy. In particular, we show that for the XoP construction, level-2 Fourier coefficients are dominant. This suggests that the best attack strategy should consider pairwise relations, and indeed, the optimal attacks by Patarin [26,28] count pairwise collisions.

Most of our technical work involves bounding the two quantities above, which is non-trivial due to intricate dependencies among the bits of the sample. This analysis does not directly deal with the XoP construction, but rather derives fundamental Fourier properties of the sampling without replacement distribution.

Bounding the quantities. We briefly summarize the main ideas used to bound each of the above quantities. Fix a mask involving bits from exactly k elements. In order to bound the associated bias of the linear equation (in absolute value), we devise an algorithm that allows to partition a subset of the sample space into sample couples with opposite signs (i.e., one satisfies the linear equation and one does not). Thus, the bias (in absolute value) is bounded by the fraction of samples that are not coupled. This fraction is bounded by probabilistic analysis of the algorithm. We note that our analysis does more than merely bound the maximal level-k Fourier coefficients. It actually classifies them into types (or groups) and obtains a refined bound for each type.

Our bound on the maximal level-k Fourier coefficient is tight, yet by itself, it is not sufficient in order to derive tight indistinguishability bounds for the XoP construction. For this purpose, we bound the level-k Fourier weight of the sampling without replacement distribution. While an exact expression for the weight is relatively easy to derive, this expression is a complex sum of terms, and therefore not immediately useful. Hence, we manipulate this expression in two main steps. First, we show how to compute the level-k Fourier weight via a recursive formula, and then we bound this weight by induction. Overall, although the weight is bounded by elementary analysis, it requires insight which is somewhat non-trivial.

Remark 1. Our bounds on the level-k Fourier weight can be formulated in terms of the so-called Efron–Stein orthogonal decomposition [24, Ch. 8] of the density function of sampling without replacement. This decomposition is independent of a specific Fourier basis, and thus these bounds apply more generally to the density function of sampling without replacement from an arbitrary set.

Technical comparison to previous works. Below, we compare our techniques to those of [13,14]. We then compare them to additional proof techniques.

Comparison to [13,14]. The papers [13,14] obtained several results in additive combinatorics. One of them is [13, Thm. 1.5], which gives an asymptotic indistinguishability bound of $O(\frac{q}{2^{1.5n}})$ for the two-permutation XoP construction. We compare our result and techniques to the ones of [13,14], focusing on the aforementioned result.

Both our proof and the one of [13,14] use Fourier analysis and (in our language) bound the (sums of) Fourier coefficients of the density function of sampling without replacement. However, the proof of [13, Thm. 1.5] is significantly more complicated. In particular, it relies on several bounds which are not required to obtain our result. Moreover, it uses complex analysis, whereas our proof is elementary.

We remark, however, that [13, Thm. 1.5] is stronger than our result in the sense that it is applicable all the way up to $q \leq 2^n - 1$, whereas our analogous result is only applicable up to $q \leq 2^{n-1} - 1$. While it is not difficult to extend our result to $q \leq C \cdot 2^n$ for some constant $C > 1/2$, our bounds do not seem sufficient to reach $q = 2^n - 1$. For such a result, the additional bounds of [13, Thm. 1.5] indeed seem necessary. On the other hand, while the difference between $q \leq 2^n - 1$ and $q \leq 2^{n-1} - 1$ is crucial in [13, Thm. 1.5], it is not very important from a cryptographic viewpoint.

The two bounds that we use (mentioned above) have comparable bounds in [13,14]. The analog of our first bound (the maximal level-k Fourier coefficient in absolute value) is [13, Lem. 4.1]. After normalization, our bound is identical for even k and slightly better for odd k. It is proved using a completely different technique. The analog of our second bound (the level-k Fourier weight) is [13, Thm 2.3] ([14, Thm. 5.1]). After normalization, our bound is somewhat inferior for small k (e.g., for $k \leq 2^{n/3}$), and becomes better for large k (e.g., denoting $N = 2^n$, it is better by a factor of $2^{\Omega(N)}$ for $k \geq \Omega(N)$). However, such an improvement seems insignificant to the asymptotic results of [13,14]. Our proof of the second bound begins by deriving an exact expression for the weight, as the proof of [14, Thm. 5.1]. On the other hand, our analysis of this expression is elementary, while the one of [14] is based on complex analysis.

In terms of generality of results, [13, Thm. 1.5] was proved for a (generalized variant of the) XoP construction defined over an arbitrary additive abelian group. While our results only apply to the original XoP construction, it is not difficult to extend them to the variant defined over an arbitrary abelian group. In fact, our second bound is already independent of the actual group (see Remark 1), and it only remains to modify the proof of the first bound. However, we leave this to future work.

Techniques not based on Fourier analysis. Additional techniques for obtaining indistinguishability results for the XoP construction used either the chi-squared method (devised in [11]), or mirror theory (devised in [25,27]), or a combination (or a variant) of them. However, when applied to the XoP construction (e.g.,

with $r = 2$), these methods have shortcomings that result in sub-optimal bounds. We elaborate on these shortcomings below, and explain why they are not present when using Fourier analysis.

The shortcoming of the chi-squared method is that when applying it the XoP construction, the analysis seems to require revealing the intermediate values output by the permutations. This extra information cannot be deduced from the output of the XoP construction, and it gives the adversary extra power, typically leading to a loose bound. While revealing the intermediate permutation values is not a strict prerequisite for the application of the chi-squared method, it seems very difficult to analyze the XoP construction without it. On the other hand, Fourier analysis deals with bounding the bias of the XOR of subsets of bits of the sample, and does not reveal any extra information to the adversary. We remark, however, that both Fourier analysis and the chi-squared method essentially bound the distance in L^2 norm between the output distribution of the cryptosystem and the uniform distribution, treated as vectors over $\mathbb{R}$ (yet, internally this is done using different techniques).

In contrast to the chi-squared method, mirror theory does not reveal extra information to the adversary. Yet, its main shortcoming is that it (essentially) considers a worst-case scenario over the samples received by the adversary. In other words, it bounds the maximal distinguishing advantage of the adversary over the output space of the XoP construction. This bound can potentially be much larger than the average bound over the samples, which, by definition, is the actual distinguishing advantage. On the other hand, Fourier analysis considers the average case over the sample when analysing biases of linear equations on subsets of its bits.

Other related papers. Superficially related papers in the area of linear cryptanalysis [10,23] show that most *fixed* idealized block ciphers (permutations) do not have a good linear approximation. Such results are different than ours as we deal with a distribution, rather than any fixed permutation. We are not aware of previous works (other than [13,14]) that analyze Fourier properties of the sampling without replacement distribution over the domain $\{0,1\}^n$.

Moreover, a linear approximation table of an n-bit block cipher includes a bias per each of the 2^{2n} possible masks (subsets of input and output bits), assuming the block cipher input is selected uniformly. On the other hand, we analyze biases (only) among subsets of output bits of a random permutation, as a function of the number of elements of $\{0,1\}^n$ involved in the XOR relation. The number of possible masks over k elements is about $2^{kn} \gg 2^{2n}$ (when $k > 2$). Such masks can have rather complex structures that the analysis needs to account for.

1.3 Paper Structure

The rest of this paper is organized as follows. We describe preliminaries in Section 2. In Section 3, we summarize our bounds on the two Fourier properties of sampling without replacement, and use them to prove indistinguishability bounds for the XoP construction. Finally, we prove these bounds in Section 4

and Section 5. Specifically, in Section 4 we bound the maximal (absolute value of the) level-k Fourier coefficient of the sampling without replacement density function, while in Section 5, we bound its level-k Fourier weight.

2 Preliminaries

For a natural number m, denote $[m] = \{1, 2, \ldots, m\}$. For natural numbers m_1 and m_2 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 $\mathbb{F}$ 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 vectors $v, u \in \mathbb{F}^k$, we denote by $\langle u, v \rangle_{\mathbb{F}} = \sum_{i \in [k]} u_i v_i$ their inner product. Similarly, for matrices $v, u \in \mathbb{F}^{k_1 \times k_2}$, $\langle u, v \rangle_{\mathbb{F}} = \sum_{(i,j) \in [k_1] \times [k_2]} u_{i,j} v_{i,j}$.

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 $N = 2^n$. We further denote by $(e_1, e_2, \ldots, e_n)$ the standard basis vectors of $\mathbb{F}_2^n$.

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 random string drawn from the associated probability distribution, defined by

$$\Pr_{x \sim \varphi}[x = y] = \varphi(y)/2^{n \cdot q} \text{ for every } y \in \mathbb{F}_2^{q \times n}.$$

In particular, the uniform probability density function over $\mathbb{F}_2^{q \times n}$ is the constant function 1, and we denote it by $\mathbf{1}_{q \cdot n}$.

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 $\mathcal{A}$.

Definition 2 (Collision probability). *The collision probability of a density function $\varphi : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}^{\geq 0}$ is*

$$Col[\varphi] = \Pr_{\substack{x, x' \sim \varphi \\ independently}}[x = x'].$$

Definition 3 (Convolution). *Let $f, g : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}$. Their convolution is the function $f * g : \mathbb{F}_2^{q \times n} \mapsto \mathbb{R}$ defined by*

$$(f * g)(x) = \mathrm{E}_{y \sim \mathbb{F}_2^{q \times n}}[f(y) g(x \oplus y)].$$

For a function $f : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}$ and a natural number $r \geq 2$, we denote the r-fold convolution of f with itself by $f^{(*r)} = f * f * \ldots * f$ (in particular $f^{(*2)} = f * f$).

Proposition 1 ([24], Proposition 1.26). *If $\varphi, \psi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ are density functions, then so is $\varphi * \psi$. It represents the distribution over $\mathbb{F}_2^{q\times n}$ given by choosing $y \sim \varphi$ and $z \sim \psi$ independently and setting $x = y \oplus z$.*

Definition 4 (Statistical distance). *The statistical distance between two probability density functions $\varphi, \psi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ is*

$$\mathrm{SD}(\varphi, \psi) = 1/2 \cdot \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}} |\varphi(x) - \psi(x)|.$$

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 taken from [24].

Definition 5 (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 $\{f \mid f : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}\}$, 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 $\{f \mid f : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}\}$ 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 $\widehat{\mathbb{F}}_{\mathbb{F}_2^{q\times n}}$, and thus

$$f(x) = \sum_{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}} \widehat{f}(\alpha)\chi_\alpha(x).$$

For a mask $\alpha \in \widehat{\mathbb{F}}_2^{q\times n}$, we write

$$\mathrm{supp}(\alpha) = \{i \mid \alpha_i \neq 0\} \text{ and } \#\alpha = |\mathrm{supp}(\alpha)|.$$

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

Definition 6 (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*

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

The Fourier weight of f up to level k is $\mathrm{W}^{\leq k}[f] = \sum_{i=0}^{k} \mathrm{W}^{=i}[f]$.
The maximal magnitude of a level-k Fourier coefficient of f is

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

Finally, let $\mathrm{M}^{\geq 1}[f] = \max_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \alpha\neq 0}} \{|\widehat{f}(\alpha)|\}$ denote the maximal magnitude of a Fourier coefficient on a non-zero mask.

Proposition 2 ([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*

$$\mathop{\mathrm{E}}_{x\sim\varphi}[f(x)] = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{q\times n}}[\varphi(x)f(x)].$$

Proposition 3 ([24], Theorem 1.27 – Fourier coefficients of convolution). *Let $f, g : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}$. Then for all $\alpha \in \widehat{\mathbb{F}}_2^{q\times n}$, $\widehat{f * g}(\alpha) = \widehat{f}(\alpha)\widehat{g}(\alpha)$.*

Proposition 4 ([24], Exercise 1.23 – relation between Fourier weight and collision probability). *For a density function $\varphi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$,*

$$\mathrm{W}^{\leq q}[\varphi] = Col[\varphi] \cdot 2^{q\cdot n}.$$

Proposition 5 ([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 6 ([24], Exercise 1.23 – bound on statistical distance from uniform). *Let $\varphi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ be a density function. Then*

$$\mathrm{SD}(\varphi, \mathbf{1}_{q\cdot n}) \leq \frac{1}{2}\sqrt{\mathrm{Var}[\varphi]}.$$

We prove two additional basic results regarding variance.

Proposition 7 (Variance reduction by convolution). *Let $\varphi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ be a density function. Let r_1, r_2 be integers such that $0 < r_2 < r_1$. Then,*

$$\mathrm{Var}[\varphi^{(*r_1)}] \leq (\mathrm{M}^{\geq 1}[\varphi])^{2(r_1-r_2)} \mathrm{Var}[\varphi^{(*r_2)}].$$

Proof. By Proposition 5 and Proposition 3,

$$\mathrm{Var}[\varphi^{(*r_1)}] = \sum_{\substack{\alpha \in \widehat{\mathbb{F}_2^{q\times n}} \\ \alpha \neq 0}} \widehat{\varphi^{(*r_1)}}(\alpha)^2 = \sum_{\substack{\alpha \in \widehat{\mathbb{F}_2^{q\times n}} \\ \alpha \neq 0}} \widehat{\varphi}(\alpha)^{2r_1}$$
$$\leq (\mathrm{M}^{\geq 1}[\varphi])^{2(r_1-r_2)} \sum_{\substack{\alpha \in \widehat{\mathbb{F}_2^{q\times n}} \\ \alpha \neq 0}} \widehat{\varphi}(\alpha)^{2r_2} = (\mathrm{M}^{\geq 1}[\varphi])^{2(r_1-r_2)}\, \mathrm{Var}[\varphi^{(*r_2)}].$$

■

Proposition 8 (Variance of independent samples). *Let $\varphi : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ be a density function. Let u be a natural number and let $\varphi^{\times u} : \mathbb{F}_2^{(q\cdot u)\times n} \mapsto \mathbb{R}^{\geq 0}$ be the density function obtained by concatenating u independent samples drawn from φ. Then,*

$$\mathrm{Var}[\varphi^{\times u}] \leq 2u \cdot \mathrm{Var}[\varphi], \;\; \textit{assuming } u \cdot \mathrm{Var}[\varphi] \leq 1/2.$$

Proof. By independence of the u samples, we have $\mathrm{Col}[\varphi^{\times u}] = \mathrm{Col}[\varphi]^u$. Applying Proposition 4 and Proposition 5,

$$\mathrm{W}^{\leq q\cdot u}[\varphi^{\times u}] = \mathrm{Col}[\varphi^{\times u}] \cdot 2^{q\cdot n\cdot u} = (\mathrm{Col}[\varphi] \cdot 2^{q\cdot n})^u = (\mathrm{W}^{\leq q}[\varphi])^u = \left(\widehat{\varphi}(0)^2 + \mathrm{Var}[\varphi]\right)^u.$$

Writing $z = \mathrm{Var}[\varphi]$ and noting that $\widehat{\varphi}(0)^2 = 1$ since φ is a density function, we have $\mathrm{W}^{\leq q\cdot u}[\varphi^{\times u}] = (1+z)^u = 1+\sum_{i=1}^{u} \binom{u}{i} z^i$. The ratio between two consecutive terms in the sum $\sum_{i=1}^{u} \binom{u}{i} z^i$ is upper bounded by $u\cdot z \leq 1/2$ (by the assumption). Thus, the sum is upper bounded by a geometric series with ratio $1/2$ (i.e., twice the first term). We conclude that

$$\mathrm{W}^{\leq q\cdot u}[\varphi^{\times u}] \leq 1 + 2u \cdot z = \widehat{\varphi^{\times u}}(0)^2 + 2u \cdot z.$$

Hence, by Proposition 5, $\mathrm{Var}[\varphi^{\times u}] = \sum_{k=1}^{q\cdot u} \mathrm{W}^{=k}[\varphi^{\times u}] \leq 2u \cdot z$. ■

2.3 Cryptographic Preliminaries and Sampling Without Replacement

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 also define the optimal advantage

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

In this paper we also consider the multi-user setting, where 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. Extending the single-user definitions, 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

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

The XoP$[r,n]$ construction and sampling without replacement. 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 natural numbers r, n such that $r \geq 2$, define the family of functions $\mathrm{XoP}[r,n] : (\mathrm{Perm}(n))^r \times \{0,1\}^n \mapsto \{0,1\}^n$ by

$$\mathrm{XoP}[r,n](\pi_1,\ldots,\pi_r,i) = \pi_1(i) \oplus \pi_2(i) \oplus \ldots \oplus \pi_r(i).$$

The main goal of this paper is to bound $\mathrm{Opt}^{\text{prf}}_{\mathrm{XoP}[r,n]}(q)$ as a function of the parameters r, n, q. By symmetry of the randomly chosen permutations $\pi_1,\ldots,\pi_r$, an adversary against $\mathrm{XoP}[r,n]$ obtains the XOR of r independent samples, each containing q elements of $\{0,1\}^n$, chosen uniformly without replacement (regardless of the actual queries). Below, we formalize this statement.

Definition 7 (Density function of sampling without replacement). *For natural numbers n, q such that $1 \leq q \leq 2^n$, let $\mu_{n,q} : \mathbb{F}_2^{q\times n} \mapsto \mathbb{R}^{\geq 0}$ be the density function associated with the process of uniformly sampling q elements from $\mathbb{F}_2^n$ without replacement. Specifically, for $x \in \mathbb{F}_2^{q\times n}$,*

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

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

Then, by Proposition 1 an adversary against $\mathrm{XoP}[r,n]$ that makes q distinct queries obtains a sample from $\mu^{(*r)}_{n,q}$. By well-known properties of statistical distance,

$$\mathrm{Opt}^{\text{prf}}_{\mathrm{XoP}[r,n]}(q) \leq \mathrm{SD}(\mu^{(*r)}_{n,q}, \mathbf{1}_{q\cdot n}). \tag{1}$$

Therefore, our task reduces to upper bounding $\mathrm{SD}(\mu^{(*r)}_{n,q}, \mathbf{1}_{q\cdot n})$.

We further consider the multi-user setting. Observe that in this setting, an adversary against $\mathrm{XoP}[r,n]$ obtains a sample of $(\mu^{(*r)}_{n,q_{\max}})^{\times u} : \mathbb{F}_2^{(q_{\max}\cdot u)\times n} \mapsto$

$\mathbb{R}^{\geq 0}$, where $(\mu^{(*r)}_{n,q_{\max}})^{\times u}$ is the density function obtained by concatenating u independent samples drawn from $\mu^{(*r)}_{n,q_{\max}}$. Similarly to the single-user setting,

$$\mathrm{Opt}^{\text{mu-prf}}_{\mathrm{XoP}[r,n],u}(q_{max}) \leq \mathrm{SD}((\mu^{(*r)}_{n,q_{\max}})^{\times u}, \mathbf{1}_{u\cdot q_{\max}\cdot n}). \tag{2}$$

Therefore, in this setting our task reduces to upper bounding $\mathrm{SD}((\mu^{(*r)}_{n,q_{\max}})^{\times u}, \mathbf{1}_{u\cdot q_{\max}\cdot n})$.

3 Indistinguishability Bounds for XoP$[r,n]$ Using Fourier Properties of Sampling Without Replacement

In this section we derive tight indistinguishability bounds for $\mathrm{XoP}[r,n]$ and then extend them to the multi-user setting. For this purpose, we start by stating the fundamental Fourier properties of $\mu_{n,k}$ that we prove in this paper.

3.1 Basic Properties of $\mu_{n,k}$

We will obtain bounds for the maximal magnitude of Fourier coefficients by level, namely $\mathrm{M}^{=k}[\mu_{n,q}]$, and Fourier weight by level, namely $\mathrm{W}^{=k}[\mu_{n,q}]$. First, note that if $x \sim \mu_{n,q}$, 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 without replacement from $\mathbb{F}_2^{k\times n}$, namely, $(x_{i_1}, \ldots, x_{i_k}) \sim \mu_{n,k}$. Therefore, for $1 \leq k \leq q$, we have $\mathrm{M}^{=k}[\mu_{n,q}] = \mathrm{M}^{=k}[\mu_{n,k}]$ and

$$\begin{aligned}\mathrm{W}^{=k}[\mu_{n,q}] = \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \#\alpha=k}} \widehat{\mu_{n,q}}(\alpha)^2 &= \sum_{\{i_1,\ldots,i_k\}\subseteq[q] \text{ distinct}} \sum_{\substack{\beta\in\widehat{\mathbb{F}}_2^{k\times n}\\ \mathrm{supp}(\beta)=\{i_1,\ldots,i_k\}}} \widehat{\mu_{n,k}}(\beta)^2 \\ &= \sum_{\{i_1,\ldots,i_k\}\subseteq[q] \text{ distinct}} \mathrm{W}^{=k}[\mu_{n,k}] = \binom{q}{k}\mathrm{W}^{=k}[\mu_{n,k}].\end{aligned}$$

Consequently, our main results bound $\mathrm{M}^{=k}[\mu_{n,k}]$ and $\mathrm{W}^{=k}[\mu_{n,k}]$. Lemma 1 below is proved in Section 4, while Lemma 2 is proved in Section 5.

Lemma 1 (Bounds on magnitude of level-k Fourier coefficients). *We have* $\mathrm{M}^{=2}[\mu_{n,2}] \leq \frac{1}{N-1}$. *Generally,*

$$\mathrm{M}^{=k}[\mu_{n,k}]^2 \leq \begin{cases} \frac{1}{\binom{N}{k}} & \text{if } k < N/2 \text{ is even,} \\ \frac{1}{\binom{N}{k}}\cdot\frac{k+1}{N-k} < \frac{1}{\binom{N}{k}} & \text{if } k < N/2 \text{ is odd.}\end{cases}$$

Note that the bound $\mathrm{M}^{=2}[\mu_{n,2}] \leq \frac{1}{N-1}$ is slightly better (by a factor of about $\sqrt{2}$) than the generic bound for $k=2$. The quantity $\mathrm{M}^{=2}[\mu_{n,2}]$ plays a significant role in our analysis, as it is the maximal magnitude of a Fourier coefficient with a non-zero mask ($\mathrm{M}^{=1}[\mu_{n,1}] = 0$ can be deduced from Lemma 2 below).

Lemma 2 (Bounds on weight of level-k Fourier coefficients). *We have*

$$\mathrm{W}^{=1}[\mu_{n,1}] = 0, \mathrm{W}^{=2}[\mu_{n,2}] = \frac{1}{N-1}, \text{ and } \mathrm{W}^{=3}[\mu_{n,3}] = \frac{4}{(N-1)(N-2)}.$$

Generally,

$$\mathrm{W}^{=k}[\mu_{n,k}] \leq \begin{cases} \frac{(N(k-1))^{k/2}}{(N)_k} \leq \Psi_N(k) & \text{if } k \geq 2 \text{ is even,} \\ \frac{(N(k-1))^{(k+1)/2}}{(N)_{k+1}} \leq \Psi_N(k+1) & \text{if } k \geq 3 \text{ is odd,} \end{cases}$$

where

$$\Psi_N(k) = \left(\frac{k}{N-k}\right)^{k/2} \exp\left(-\frac{k(k-2)}{8N(N-k)+2\cdot k^2}\right)^{k/2}.$$

Remark 2. The fact that $\mathrm{W}^{=1}[\mu_{n,1}] = 0$ is obvious since $\mu_{n,1}$ is the uniform distribution over $\{0,1\}^n$, and thus all non-empty linear equations on these bits are unbiased.

Remark 3. For $k < N/2$, $\frac{k}{N-k} < 1$. Therefore, the lemma shows that the Fourier weight of $\mu_{n,k}$ at level k is exponentially small in k up to $k < N/2$. In particular, in the extreme case of $k \approx N/2$, we have

$$\exp\left(-\frac{k(k-2)}{8N(N-k)+2\cdot k^2}\right)^{k/2} \approx \exp\left(-\frac{N^2/4}{4N^2+N^2/2}\right)^{N/4} = e^{-N/72} \approx e^{-k/36}.$$

Nevertheless, we will only use a simpler bound of the form $\mathrm{W}^{=k}[\mu_{n,k}] \leq \left(\frac{k}{N-k}\right)^{k/2}$ in our application. Furthermore, since $\mathrm{W}^{=k}[\mu_{n,q}] = \binom{q}{k}\mathrm{W}^{=k}[\mu_{n,k}]$, the number of queries q obviously also plays a significant role in the analysis.

3.2 Application to Indistinguishability Bounds for XoP$[r,n]$

We now use the results about $\mu_{n,k}$ in our main application to derive indistinguishability bounds for XoP$[r,n]$, starting with $r = 2$.

Theorem 1. *For $N \geq 1000$ and $q < N/2$,*

$$\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{XoP}[2,n]}(q) \leq \frac{q}{2\cdot(N-1)^{3/2}} < \frac{q}{N^{3/2}}.$$

Proof. Using (1), and applying Proposition 6,

$$\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{XoP}[2,n]}(q) \leq \mathrm{SD}(\mu_{n,q} * \mu_{n,q}, \mathbf{1}_{q\cdot n}) \leq \frac{1}{2}\sqrt{\mathrm{Var}[\mu_{n,q} * \mu_{n,q}]}.$$

Thus, it remains to prove that

$$\mathrm{Var}[\mu_{n,q} * \mu_{n,q}] \leq \frac{q^2}{(N-1)^3}. \tag{3}$$

Applying Proposition 5, and then Proposition 3, we have

$$\mathrm{Var}[\mu_{n,q}^{(*2)}] = \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \alpha\neq 0}} \widehat{\mu_{n,q} * \mu_{n,q}}(\alpha)^2 = \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \alpha\neq 0}} \widehat{\mu_{n,q}}(\alpha)^4 = \sum_{k=1}^{q} \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \#\alpha=k}} \widehat{\mu_{n,q}}(\alpha)^4$$

$$\leq \sum_{k=1}^{q} \mathrm{M}^{=k}[\mu_{n,q}]^2 \sum_{\substack{\alpha\in\widehat{\mathbb{F}}_2^{q\times n}\\ \#\alpha=k}} \widehat{\mu_{n,q}}(\alpha)^2 = \sum_{k=1}^{q} \mathrm{M}^{=k}[\mu_{n,q}]^2 \cdot \mathrm{W}^{=k}[\mu_{n,q}]$$

$$= \sum_{k=1}^{q} \mathrm{M}^{=k}[\mu_{n,k}]^2 \cdot \binom{q}{k} \mathrm{W}^{=k}[\mu_{n,k}],$$

where the final equality exploits the symmetry of $\mu_{n,q}$. Next, applying Lemma 1, and using the fact that $\mathrm{W}^{=1}[\mu_{n,1}] = 0$ (by Lemma 2),

$$\mathrm{Var}[\mu_{n,q}^{(*2)}] \leq \frac{1}{(N-1)^2} \cdot \binom{q}{2} \cdot \mathrm{W}^{=2}[\mu_{n,2}] + \sum_{k=3}^{q} \frac{\binom{q}{k}}{\binom{N}{k}} \mathrm{W}^{=k}[\mu_{n,k}]$$

$$\leq \frac{q^2}{(N-1)^2} \cdot (1/2) \cdot \mathrm{W}^{=2}[\mu_{n,2}] + \sum_{k=3}^{q} \frac{(q)(q-1)\dots(q-(k-1))}{(N)(N-1)\dots(N-(k-1))} \mathrm{W}^{=k}[\mu_{n,k}]$$

$$\leq \frac{q^2}{(N-1)^2} \cdot (1/2) \cdot \mathrm{W}^{=2}[\mu_{n,2}] + \sum_{k=3}^{q} (q/N)^k \cdot \mathrm{W}^{=k}[\mu_{n,k}]$$

$$\leq \frac{q^2}{(N-1)^2} \left((1/2) \cdot \mathrm{W}^{=2}[\mu_{n,2}] + \sum_{k=3}^{q} (q/N)^{k-2} \cdot \mathrm{W}^{=k}[\mu_{n,k}] \right)$$

We now apply Lemma 2. We will also separate the term $\mathrm{W}^{=3}[\mu_{n,3}] = \frac{4}{(N-1)(N-2)}$ from the sum of terms for $k \geq 4$. For these we use a simple bound

$$\mathrm{W}^{=k}[\mu_{n,k}] \leq \left(\frac{k+1}{N-k-1} \right)^{k/2} \leq \left(\frac{2(k+1)}{N} \right)^{k/2},$$

which holds both for even and odd k, and uses the fact that $k \leq q < N/2$. We will further split the remaining sum at $k = 4n$ and use once again the fact that $q/N < 1/2$. Thus, $\mathrm{Var}[\mu_{n,q}^{(*2)}]$ is upper bounded by

$$\frac{q^2}{(N-1)^2} \cdot \left((1/2) \cdot \mathrm{W}^{=2}[\mu_{n,2}] + (q/N) \cdot \mathrm{W}^{=3}[\mu_{n,3}] + \sum_{k=4}^{4n} (q/N)^{k-2} \cdot \mathrm{W}^{=k}[\mu_{n,k}] \right)$$

$$+ \sum_{k=4n+1}^{q} (q/N)^k \cdot \mathrm{W}^{=k}[\mu_{n,k}]$$

$$\leq \frac{q^2}{(N-1)^2} \cdot \left(\frac{1}{2(N-1)} + \frac{2}{(N-1)(N-2)} + 4 \sum_{k=4}^{4n} 2^{-k} \cdot \left(\frac{2(k+1)}{N} \right)^{k/2} \right) + \sum_{k=4n+1}^{q} 2^{-k}$$

$$\leq \frac{q^2}{(N-1)^2} \cdot \left(\frac{1}{2(N-1)} + \frac{2}{(N-1)(N-2)} + 4 \sum_{k=4}^{4n} \left(\frac{k+1}{2N} \right)^{k/2} \right) + N^{-4}.$$

We now upper bound $\sum_{k=4}^{4n}\left(\frac{k+1}{2N}\right)^{k/2}$. The (inverse) squared ratio between two consecutive terms is

$$\frac{((k+1)/2N)^k}{((k+2)/2N)^{k+1}} = \left(\frac{k+1}{k+2}\right)^k \cdot \frac{2N}{k+2} = \left(1-\frac{1}{k+2}\right)^k \cdot \frac{2N}{k+2}$$
$$\geq e^{-2k/(k+2)}\frac{2N}{k+2} \geq e^{-2}\frac{2N}{k+2} \geq \frac{2N}{(4n+2)e^2}.$$

where we have used the inequality $1-(x/2) > e^{-x}$, which holds for $0 < x \leq 1$, as well as the fact that $k \leq 4n$ in the analyzed sum. Since $\frac{2N}{(4n+2)e^2} \geq 4$ holds for $N \geq 1000$, the sum is upper bounded by the sum of a geometric series with ratio at most $1/2$. Hence, $\sum_{k=4}^{4n}\left(\frac{k+1}{2N}\right)^{k/2} \leq 2\left(\frac{5}{2N}\right)^2 = \frac{25}{2N^2}$. Also, noting that $N^{-4} \leq (q^2/(N-1)^2)\cdot 1/N^2$, we plug these into the above bound and obtain

$$\mathrm{Var}[\mu_{n,q}^{(*2)}] \leq \frac{q^2}{(N-1)^2}\cdot\left(\frac{1}{2(N-1)}+\frac{2}{(N-1)(N-2)}+\frac{50}{N^2}+\frac{1}{N^2}\right).$$

As each one of the last three summands is bounded by $\frac{1}{8(N-1)}$ assuming $N \geq 1000$, we conclude that $\mathrm{Var}[\mu_{n,q}^{(*2)}] \leq \frac{q^2}{(N-1)^3}$ as in (3). ■

Next, we generalize Theorem 1 to derive indistinguishability bounds for $\mathrm{XoP}[r,n]$ for arbitrary $r \geq 2$.

Theorem 2. *For $N \geq 1000$, $q < N/2$ and $r \geq 2$,*

$$\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{XoP}[r,n]}(q) \leq \frac{q}{2\cdot(N-1)^{r-(1/2)}} < \frac{q}{N^{r-(1/2)}},$$

where the last inequality assumes $r \leq N/2$.

Proof. By (1) and Proposition 6, $\mathrm{Opt}^{\mathrm{prf}}_{\mathrm{XoP}[r,n]}(q) \leq \mathrm{SD}(\mu_{n,q}^{(*r)}, \mathbf{1}_{q\cdot n}) \leq \frac{1}{2}\sqrt{\mathrm{Var}[\mu_{n,q}^{(*r)}]}$, and thus it remains to prove that

$$\mathrm{Var}[\mu_{n,q}^{(*r)}] \leq \frac{q^2}{(N-1)^{2r-1}}. \tag{4}$$

Applying Proposition 6 and then Proposition 7 (with $r_2 = 2$),

$$\mathrm{Var}[\mu_{n,q}^{(*r)}] \leq (\mathrm{M}^{\geq 1}[\mu_{n,q}])^{2r-4}\cdot\mathrm{Var}[\mu_{n,q}^{(*2)}] = (\max_{0<k\leq q}\{\mathrm{M}^{=k}[\mu_{n,k}]\})^{2r-4}\cdot\mathrm{Var}[\mu_{n,q}^{(*2)}],$$

where the final equality is by symmetry of $\mu_{n,q}$. Next, note from Lemma 1 that (the bound on) $\mathrm{M}^{=k}[\mu_{n,k}]$ is maximized for $k=2$ assuming $q < N/2$, and $\mathrm{M}^{=2}[\mu_{n,2}] \leq \frac{1}{N-1}$. Moreover $\mathrm{Var}[\mu_{n,q}^{(*2)}] \leq \frac{q^2}{(N-1)^3}$ by (3). Hence,

$$\mathrm{Var}[\mu_{n,q}^{(*r)}] \leq \frac{1}{(N-1)^{2r-4}}\frac{q^2}{(N-1)^3} = \frac{q^2}{(N-1)^{2r-1}}.$$

■

The multi-user setting. We extend Theorem 2 to derive indistinguishability bounds for $\mathrm{XoP}[r,n]$ in the multi-user setting.

Theorem 3. *For $N \geq 1000$, $q < N/2$ and $r \geq 2$,*

$$\mathrm{Opt}^{\text{mu-prf}}_{\mathrm{XoP}[r,n],u}(q_{max}) \leq \frac{\sqrt{u/2}\cdot q_{\max}}{(N-1)^{r-(1/2)}} \leq \frac{\sqrt{u}\cdot q_{\max}}{N^{r-(1/2)}},$$

assuming $\frac{\sqrt{u/2}\cdot q_{\max}}{(N-1)^{r-(1/2)}} \leq 1/2$ (and $r \leq N/3$ for the last inequality).

Proof. By (2) and Proposition 6,

$$\mathrm{Opt}^{\text{mu-prf}}_{\mathrm{XoP}[r,n],u}(q_{max}) \leq \mathrm{SD}((\mu^{(*r)}_{n,q_{\max}})^{\times u}, \mathbf{1}_{u\cdot q_{\max}\cdot n}) \leq \frac{1}{2}\sqrt{\mathrm{Var}[(\mu^{(*r)}_{n,q_{\max}})^{\times u}]},$$

and thus is remains to prove that $\mathrm{Var}[(\mu^{(*r)}_{n,q_{\max}})^{\times u}] \leq \frac{2u\cdot q^2_{\max}}{(N-1)^{2r-1}}$.

Applying Proposition 8 (assuming $u\cdot\mathrm{Var}[\mu^{(*r)}_{n,q_{\max}}] \leq 1/2$), we have

$$\mathrm{Var}[(\mu^{(*r)}_{n,q_{\max}})^{\times u}] \leq 2u\cdot\mathrm{Var}[\mu^{(*r)}_{n,q_{\max}}] \leq \frac{2u\cdot q^2_{\max}}{(N-1)^{2r-1}},$$

where the final inequality is by (4). Finally, note that by (4), $u\cdot\mathrm{Var}[\mu^{(*r)}_{n,q_{\max}}] \leq \frac{u\cdot q^2_{\max}}{(N-1)^{2r-1}}$, so the condition for applying Proposition 8 is assured if $\frac{u\cdot q^2_{\max}}{(N-1)^{2r-1}} \leq 1/2$, namely $\frac{\sqrt{u/2}\cdot q_{\max}}{(N-1)^{r-(1/2)}} \leq 1/2$. ∎

4 Bounding $\mathbf{M}^{=k}[\boldsymbol{\mu}_{n,k}]$ (Proof of Lemma 1)

The goal of this section is to prove Lemma 1. We first bound the Fourier coefficients on a specific subset of masks (called masks of type $K=(k)$). We will later generalize these results to all mask.

4.1 Bounding $|\widehat{\mu_{n,k}}(\alpha)|$ for α of Type $K=(k)$

Definition 8 (Mask of type $K=(k)$). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ be a non-zero mask such that $\#\alpha = k$ (i.e., $\alpha_i \neq 0$ for all $i \in [k]$). We define the type of α to be $K=(k)$, if for every $i \in [k]$, $\alpha_{i,1}=1$.*

In other words, α is of type $K=(k)$ if the first bit of all of its k elements is 1. The bounds on the Fourier coefficients are formulated using the following function.

Definition 9. *For natural numbers a, b such that b is even and $a \geq b$ let*

$$\Gamma(a,b) = \prod_{i=1,3,\ldots,b-1} \frac{b-i}{a-i}.$$

The main result of this section is as follows.

Proposition 9. *Let $\alpha \in \widehat{\mathbb{F}_2^{k\times n}}$ be of type $K = (k)$. Then,*

$$|\widehat{\mu_{n,k}}(\alpha)| \leq \Gamma(N,k) = \prod_{i=1,3,\ldots,k-1} \frac{k-i}{N-i}$$

if k is even and 0 otherwise.

In particular,

$$|\widehat{\mu_{n,1}}(\alpha)| = 0, |\widehat{\mu_{n,2}}(\alpha)| \leq \frac{1}{N-1}, |\widehat{\mu_{n,3}}(\alpha)| = 0, |\widehat{\mu_{n,4}}(\alpha)| \leq \frac{3}{(N-1)(N-3)},$$

etc. We need the following definitions.

Definition 10 (Pairing of two elements). *Two elements $a, b \in \mathbb{F}_2^n$ are paired on bit $j \in [n]$ if $a \oplus b = e_j$ (where $e_j \in \mathbb{F}_2^n$ is the j'th vector of the standard basis).*

Definition 11 (Pairing of a sequence of elements). *Let $x = (x_1, \ldots, x_k) \in \mathbb{F}_2^{k\times n}$. Then, x is self-paired on bit $j \in [n]$ if $(x_1, \ldots, x_k)$ are distinct (i.e., $x_{i_1} \neq x_{i_2}$ for $i_1 \neq i_2$), and for every $i_1 \in [k]$, there exists $i_2 \in [k]$ such that (x_{i_1}, x_{i_2}) are paired on bit j.*

Note that since $(x_1, \ldots, x_k)$ are distinct, each element x_i cannot be paired to more than one other element on bit j, and thus if x is self-paired (on any $j \in [n]$), then k is even.

In order to prove Proposition 9, we define the following algorithm.

1. Sample $x \sim \mu_{n,k}$.
2. If x is self-paired on bit 1, return 1. Else, return 0.

Define the random variable $T(x)$ for the output of the algorithm.

We will prove the following two claims, whose combination immediately implies Proposition 9.

Proposition 10 (Magnitude of Fourier coefficient bounded by success probability). $|\widehat{\mu_{n,k}}(\alpha)| \leq \Pr_{x\sim\mu_{n,k}}[T(x) = 1]$.

Proposition 11 (Bound on success probability).

$$\Pr_{x\sim\mu_{n,k}}[T(x) = 1] = \begin{cases} \Gamma(N,k) & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Proof (of Proposition 10). By Proposition 2,

$$\begin{aligned}
&|\widehat{\mu_{n,k}}(\alpha)| = |\mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^n}[\mu_{n,k}(x)\chi_\alpha(x)]| = |\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)]| \\
=&|\Pr_{x\sim\mu_{n,k}}[T(x)=1]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=1] \\
+&\Pr_{x\sim\mu_{n,k}}[T(x)=0]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=0]| \\
\le&|\Pr_{x\sim\mu_{n,k}}[T(x)=1]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=1]| \\
+&|\Pr_{x\sim\mu_{n,k}}[T(x)=0]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=0]| \\
\le&|\Pr_{x\sim\mu_{n,k}}[T(x)=1]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[|\chi_\alpha(x)|\mid T(x)=1]| \\
+&|\Pr_{x\sim\mu_{n,k}}[T(x)=0]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=0]| \\
=&\Pr_{x\sim\mu_{n,k}}[T(x)=1]+|\Pr_{x\sim\mu_{n,k}}[T(x)=0]\mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=0]|.
\end{aligned} \tag{5}$$

Next, we prove that $\mathrm{E}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=0]=0$, which concludes the proof. This is proved by partitioning the sample space of the algorithm conditioned on $T(x)=0$ into couples of the form (x,x') such that $\chi_\alpha(x)=-\chi_\alpha(x')$. Since all samples in the space (conditioned on $T(x)=0$) have identical probability, the total contribution of each couple to the expectation is $\chi_\alpha(x)+\chi_\alpha(x')=0$, which proves that $\mathrm{E}_{x\sim\mu_{n,k}}[\chi_\alpha(x)\mid T(x)=0]=0$.

We now define how to couple the samples. Assume that $T(x)=0$. Then, there exists an element of x that is not paired. Define $in(x)\in[k]$ to be the index of the first unpaired element in $[k]$. Then, $x'=(x_1,\ldots,x_{in(x)-1},x_{in(x)}\oplus e_1,x_{in(x)+1},\ldots,x_k)$ is a valid sample from the space (conditioned on $T(x)=0$). We couple together (x,x'). Note that we need to prove that this is a valid coupling, i.e., if x is coupled to x', then x' is coupled to x. This indeed holds since $in(x')=in(x)$, as x and x' only differ on the element with index $in(x)$.

Finally, we prove that $\chi_\alpha(x)=-\chi_\alpha(x')$ or $\chi_\alpha(x)\chi_\alpha(x')=-1$. As $\alpha\in\widehat{\mathbb{F}}_2^{k\times n}$ is of type $K=(k)$, then $\alpha_{i,1}=1$ for any $i\in[k]$. Therefore,

$$\begin{aligned}
\chi_\alpha(x)\chi_\alpha(x') &= (-1)^{\langle\alpha,x\rangle_{\mathbb{F}_2}}(-1)^{\langle\alpha,x'\rangle_{\mathbb{F}_2}} = (-1)^{\langle\alpha,x\oplus x'\rangle_{\mathbb{F}_2}} \\
&= (-1)^{\langle\alpha_{in(x)},e_1\rangle_{\mathbb{F}_2}} = (-1)^{1\cdot1} = -1.
\end{aligned}$$

■

Proof (of Proposition 11). First, if k is odd, then x cannot be self-paired. Hence, $\Pr_{x\sim\mu_{n,k}}[T(x)=0]=1$ and $\Pr_{x\sim\mu_{n,k}}[T(x)=1]=0$.

Next, assume that k is even and consider x_1. There is a single element it can be paired to on bit 1, which is $x_1\oplus e_1$. The probability that $x_1\oplus e_1$ appears among $x_2,\ldots,x_k$ is $\frac{k-1}{N-1}$. Next, assuming x_1 is paired, continue by induction after removing the pair from the set of available elements. We obtain

$$\Pr_{x\sim\mu_{n,k}}[T(x)=1]=\frac{k-1}{N-1}\frac{k-3}{N-3}\cdots\frac{1}{N-k+1}=\Gamma(N,k),$$

as claimed. ■

4.2 Classification of Masks

Towards proving bounds on the magnitude of Fourier coefficients on general masks, we define two basic operations on masks and prove that they preserve Fourier coefficients. These operations will allow us to focus on a subset of masks whose associated Fourier coefficient is easier to bound. Bounds on the magnitude of Fourier coefficients on the remaining masks will follow by preservation of Fourier coefficients.

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 the mask $\alpha^\pi \in \widehat{\mathbb{F}}_2^{k\times n}$ by $\alpha_i^\pi = \alpha_{\pi(i)}$ for $i \in [k]$. Then, $\widehat{\mu_{n,k}}(\alpha^\pi) = \widehat{\mu_{n,k}}(\alpha)$.*

Proof. Similarly to the definition of α^π, for $x \in \mathbb{F}_2^{k\times n}$, define $x^\pi \in \mathbb{F}_2^{k\times n}$ by $x_i^\pi = x_{\pi(i)}$ for $i \in [k]$. Observe that since π merely permutes the elements of x, it preserves equality and inequality among elements, and thus $\mu_{n,k}(x) = \mu_{n,k}(x^\pi)$. Furthermore $\chi_\alpha(x) = \chi_{\alpha^\pi}(x^\pi)$ as inner product in invariant under permutation of elements of α and x. Combining these observations,

$$\begin{aligned}\widehat{\mu_{n,k}}(\alpha) &= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x)\chi_\alpha(x)] = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x^\pi)\chi_{\alpha^\pi}(x^\pi)] \\ &= \mathop{\mathrm{E}}_{y\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(y)\chi_{\alpha^\pi}(y)] = \widehat{\mu_{n,k}}(\alpha^\pi).\end{aligned}$$

■

Proposition 13 (Invertible element-wise linear operations preserve Fourier coefficients). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$. Let $L \in \mathbb{F}_2^{n\times n}$ be an invertible matrix and define the mask $\alpha^L \in \widehat{\mathbb{F}}_2^{k\times n}$ by $\alpha_i^L = \alpha_i \cdot L$ for $i \in [k]$ (where we view α_i as a row vector in $\mathbb{F}_2^n$, multiplied with L). Then, $\widehat{\mu_{n,k}}(\alpha^L) = \widehat{\mu_{n,k}}(\alpha)$.*

Proof. For $x \in \mathbb{F}_2^{k\times n}$, define $x^L \in \mathbb{F}_2^{k\times n}$ similarly to the definition of α^L. By the properties of the inner product, for any $a, b \in \mathbb{F}_2^n$,

$$\langle a, b\rangle_{\mathbb{F}_2} = \langle a \cdot L \cdot L^{-1}, b\rangle_{\mathbb{F}_2} = \langle a \cdot L, b \cdot L^{-T}\rangle_{\mathbb{F}_2},$$

where L^T is the transpose of L and L^{-T} is the inverse of L^T. Hence, $\chi_\alpha(x) = \chi_{\alpha^L}(x^{L^{-T}})$. Furthermore, since L^{-T} is an invertible transformation on the elements of x, it preserves equality and inequality among elements, and thus $\mu_{n,k}(x) = \mu_{n,k}(x^{L^{-T}})$. Therefore,

$$\begin{aligned}\widehat{\mu_{n,k}}(\alpha) &= \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x)\chi_\alpha(x)] = \mathop{\mathrm{E}}_{x\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(x^{L^{-T}})\chi_{\alpha^L}(x^{L^{-T}})] \\ &= \mathop{\mathrm{E}}_{y\sim\mathbb{F}_2^{k\times n}}[\mu_{n,k}(y)\chi_{\alpha^L}(y)] = \widehat{\mu_{n,k}}(\alpha^L).\end{aligned}$$

■

These two propositions motivate the following definition.

Definition 12 (Equivalence of masks). *Masks $\alpha, \beta \in \widehat{\mathbb{F}}_2^{k\times n}$ are called equivalent (with respect to $\mu_{n,k}$) if β can be obtained from α by permuting its elements and performing invertible element-wise linear operations.*

By invertibility of the basic operations, equivalence of masks is a well-defined equivalence relation. By the above propositions, if α and β are equivalent, then $\widehat{\mu_{n,k}}(\alpha) = \widehat{\mu_{n,k}}(\beta)$ (and obviously $\#\alpha = \#\beta$).

We now define a classification of masks that will later be used to bound their associated Fourier coefficients.

Definition 13 (Rank of mask). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ be a non-zero mask. We define the rank of α as its rank when viewed as a $k \times n$ matrix over $\mathbb{F}_2$.*

The following definition generalizes Definition 8.

Definition 14 (Type of mask). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ be a mask such that $\#\alpha = k > 0$. Let $K = (k_1, k_2, \ldots, k_t)$ be a t-tuple of natural positive indices such that $k_j < k_{j+1}$ for all $j \in [t-1]$ and $k_t = k$. Define $k_0 = 0$. We define the type of α to be K, if for every $j \in [t]$, the following two conditions hold:*

1. *For every $i \in [k_{j-1}+1, k_j]$, $\alpha_{i,j} = 1$.*
2. *For every $i \in [k_j+1, k]$, $\alpha_{i,j} = 0$.*

If α is not of type K for any tuple K, then we define its type to be NULL.

In other words, α is of type $K = (k_1, k_2, \ldots, k_t)$ if the first bit of its first k_1 elements is 1, and the first bits of elements $x_{k_1+1}, \ldots, x_k$ is 0. Next, bit 2 of elements $x_{k_1+1}, \ldots, x_{k_2}$ is 1, while bit 2 of elements $x_{k_2+1}, \ldots, x_k$ is 0, and so forth.

Example 1. Let $n = 4$ and $k = 3$ and assume the leftmost bit is the first bit. Then, the mask $(1011, 1101, 1001)$ is of type (3), $(1011, 0110, 0101)$ is of type $(1, 3)$, $(1011, 0110, 0011)$ is of type $(1, 2, 3)$, while $(1011, 0101, 1001)$, $(1011, 0010, 0101)$ and $(1011, 0110, 0001)$ are all of type NULL.

While many non-zero masks have type NULL, they can be easily transformed to a non-NULL type by basic operations. More specifically, the following holds.

Proposition 14 (Every non-zero mask is equivalent to a mask of non-NULL type). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k\times n}$ have $\#\alpha = k > 0$ and rank r. Then, α is equivalent to some $\beta \in \widehat{\mathbb{F}}_2^{k\times n}$ of type $K = (k_1, \ldots, k_t)$, such that $k_t = k$ and $t = r$.*

Proposition 14 thus allows us to focus on bounding the Fourier coefficients on masks of non-NULL type.

Proof. We transform α to β by basic operations as follows. First, since the rank of α is r, it contains r linearly independent elements. Define and apply to α an invertible linear transformation that maps the first r linearly independent elements (in lexicographical order) to the first r vectors of the standard basis of $\mathbb{F}_2^n$, $e_1, \ldots, e_r$. Denote the outcome by α'.

Next, permute the elements of α' by moving all elements α'_i such that $\alpha'_{i,1} = 1$ to be first, and elements with $\alpha'_{i,1} = 0$ to be last. Let k_1 be the index such that $\alpha'_{i,1} = 1$ if $i \le k_1$ and $\alpha'_{i,1} = 0$ if $i > k_1$. Note that $k_1 \ge 1$ since the first bit of e_1 is 1 and $k_1 \le k - r + 1$, as the first bit of all the elements $e_2, \ldots, e_r$ is 0. If $r = 1$, then since $\#\alpha = k$ we must have $k_1 = k$ (otherwise, α has two linearly independent elements). Thus, define $\beta = \alpha'$, which is of type (k), and we are done after 1 step. If $r > 1$, define k_2 after permuting the elements $\alpha'_{k_1+1}, \ldots, \alpha'_k$ according to their second bit and continue inductively. After the process terminates, define $\beta = \alpha'$.

Denote by t the total number of steps in the process. The process cannot end with $t < r$ as the first bit set to j in e_j has index j, and thus e_j will be among the elements $\alpha'_{k_{j-1}+1}, \ldots, \alpha'_{k_j}$. On the other hand, the process cannot end with $t > r$ steps, since vectors $\alpha'_{k_1}, \ldots, \alpha'_{k_t}$ are linearly independent. Therefore, $t = r$. Furthermore, $k_t = k$ since $\#\alpha = k$. We conclude that α is equivalent to $\beta = \alpha'$ of type $K = (k_1, \ldots, k_t)$ such that $k_t = k$ and $t = r$. ■

4.3 Bounding $|\widehat{\mu_{n,k}}(\alpha)|$ for general α.

In this section we prove bounds on the magnitude of Fourier coefficients on general masks. The main result of this section is the following.

Proposition 15 (Bounds on Fourier magnitude for general masks). *We have*

$$\mathrm{M}^{=k}[\mu_{n,k}] \le \begin{cases} \Gamma(N,k) & \text{if } k < N/2 \text{ is even,} \\ \Gamma(N,k-1) \cdot \frac{k}{N-k} & \text{if } k < N/2 \text{ is odd.} \end{cases}$$

Equivalently, let $\alpha \in \widehat{\mathbb{F}}_2^{k \times n}$ have $\#\alpha = k$. Then,

$$|\widehat{\mu_{n,k}}(\alpha)| \le \begin{cases} \Gamma(N,k) & \text{if } k < N/2 \text{ is even,} \\ \Gamma(N,k-1) \cdot \frac{k}{N-k} & \text{if } k < N/2 \text{ is odd.} \end{cases}$$

Lemma 1 (stated in Section 3) is proved in Appendix A based on this proposition by a straightforward bound on $\Gamma(N,k)$.

Proposition 15 is a consequence of the following proposition.

Proposition 16 (Bounds on Fourier magnitude for masks of non-NULL type). *Let $\alpha \in \widehat{\mathbb{F}}_2^{k \times n}$ be of type $K = (k_1, \ldots, k_t)$ where $k_t = k$. Then,*

$$|\widehat{\mu_{n,k}}(\alpha)| \le \begin{cases} \Gamma(N,k) & \text{if } k < N/2 \text{ is even,} \\ \Gamma(N,k-1) \cdot \frac{k}{N-k} & \text{if } k < N/2 \text{ is odd.} \end{cases}$$

Proof (of Proposition 15). Let $\alpha \in \widehat{\mathbb{F}}_2^{k \times n}$ have $\#\alpha = k$. Then, by Proposition 14, it is equivalent to some $\beta \in \widehat{\mathbb{F}}_2^{k \times n}$ of type $K = (k_1, \ldots, k_t)$ where $k_t = k$ (with the same rank as α). This proposition follows by applying Proposition 16 to β. ■

It remains to prove Proposition 16. We need the following additional definition.

Definition 15 (Pairing of a subsequence of elements). *Let $x = (x_1, \ldots, x_k) \in \mathbb{F}_2^{k \times n}$. Let $k' \in [k]$. Define $(x_{k'}, \ldots, x_k)$ as paired within $x = (x_1, \ldots, x_k)$ on bit $j \in [n]$ if $(x_1, \ldots, x_k)$ are distinct (i.e, $x_{i_1} \neq x_{i_2}$ for $i_1 \neq i_2$), and for every $i_1 \in [k', k]$, there exists $i_2 \in [k]$ such that (x_{i_1}, x_{i_2}) are paired on bit j.*

We define the following algorithm that generalizes the algorithm of Section 4.1 to handle a mask with arbitrary non-NULL type. It takes as input the tuple $K = (k_1, \ldots, k_t)$ (recall the $k_0 = 0$ by definition).

1. Sample $x \sim \mu_{n,k}$.
2. For all $j \in [t]$:
   (a) If $(x_{k_{j-1}+1}, \ldots, x_{k_j})$ are paired within $(x_1, \ldots, x_{k_j})$ on bit j, continue by incrementing j.
   (b) Otherwise, return 0.
3. Return 1.

For $j \in [t]$, define the random variable $T_j(x)$ to be equal to 1 if the algorithm has not returned 0 in iterations $1, \ldots, j$, and let $T_j(x) = 0$ otherwise. Furthermore, define $T(x) = T_t(x)$ to be the output of the algorithm.

We need the following definition.

Definition 16. *For integers $a, b \geq 0, c \geq 1$ such that $a \geq b + c$ ($a > b + c$ if c is odd), define*

$$\Lambda(a,b,c) = \begin{cases} \prod_{i=1,3\ldots,c-1} \frac{b+c-i}{a-b-i} = \frac{b+c-1}{a-b-1}\frac{b+c-3}{a-b-3}\cdots\frac{b+1}{a-b-c+1} & \text{if } c \text{ is even,} \\ \prod_{i=1,3\ldots,c} \frac{b+c-i}{a-b-i} = \frac{b+c-1}{a-b-1}\frac{b+c-3}{a-b-3}\cdots\frac{b}{a-b-c} & \text{if } c \text{ is odd.} \end{cases}$$

Note that for even k, $\Gamma(N,k) = \Lambda(N,0,k)$.

Proposition 16 immediately follows from the three propositions below (that refer to the type of α, namely $K = (k_1, \ldots, k_t)$).

Proposition 17 (Magnitude of Fourier coefficient bounded by success probability). $|\widehat{\mu_{n,k}}(\alpha)| \leq \Pr_{x \sim \mu_{n,k}}[T(x) = 1]$.

Proposition 18 (Bound on success probability). *If k_1 is even, then*

$$\Pr_{x \sim \mu_{n,k}}[T(x) = 1] \leq \Gamma(N, k_1) \cdot \prod_{j=2}^{t} \Lambda(N, k_{j-1}, k_j - k_{j-1}),$$

while if k_1 is odd then, $\Pr_{x \sim \mu_{n,k}}[T(x) = 1] = 0$.

Proposition 19. *For even k_1, we have*

$$\Gamma(N, k_1) \cdot \prod_{j=2}^{t} \Lambda(N, k_{j-1}, k_j - k_{j-1}) \leq \begin{cases} \Gamma(N,k) & \text{if } k = k_t < N/2 \text{ is even,} \\ \Gamma(N, k-1) \cdot \frac{k}{N-k} & \text{if } k = k_t < N/2 \text{ is odd.} \end{cases}$$

In the rest of this section we will prove Proposition 17 and Proposition 18. Proposition 19 is proved in Appendix A by elementary analysis.

Proof (of Proposition 17). The proof is a generalization of the proof of Proposition 10, and we focus on the differences. As in (5),

$$|\widehat{\mu_{n,k}}(\alpha)| \leq \Pr_{x\sim\mu_{n,k}}[T(x)=1] + |\Pr_{x\sim\mu_{n,k}}[T(x)=0] \mathop{\mathrm{E}}_{x\sim\mu_{n,k}}[\chi_\alpha(x) \mid T(x)=0]|,$$

and it remains to prove that $\mathrm{E}_{x\sim\mu_{n,k}}[\chi_\alpha(x) \mid T(x)=0]=0$. Once again this is proved by partitioning the sample space conditioned on $T(x)=0$ into couples (x,x') that satisfy $\chi_\alpha(x)=-\chi_\alpha(x')$. However, this time the coupling depends on the iteration $j\in[t]$ which the algorithm executed and returned 0, namely, $T_\ell(x)=1$ for $\ell\in[j-1]$ and $T_j(x)=0$. Fix this iteration $j\in[t]$, let $in(x)\in[k_{j-1}+1,k_j]$ be the index of the first unpaired element among $(x_{k_{j-1}+1},\ldots,x_{k_j})$.

We now consider two cases depending on whether $x_{in(x)}\oplus e_j$ appears among $x_{k_j+1},\ldots,x_k$ (note that it does not appear among $(x_1,\ldots,x_{k_j})$ since $x_{in(x)}$ is not paired to any of these elements).

If $x_{in(x)}\oplus e_j$ does not appear among $(x_{k_j+1},\ldots,x_k)$, then it does not appear among $(x_1,\ldots,x_k)$, and thus we couple x and $x'=(x_1,\ldots,x_{in(x)-1},x_{in(x)}\oplus e_j,x_{in(x)+1},\ldots,x_k)$, as in the proof of Proposition 10. Specifically, in this case we have $in(x)=in(x')$. Moreover, since α is of type K, then $\alpha_{i,j}=1$ for all $i\in[k_{j-1}+1,k_j]$, and in particular, $\alpha_{in(x),j}=1$. Since $x_{in(x),j}\neq x'_{in(x),j}$ and they are they equal otherwise, $\chi_\alpha(x)=-\chi_\alpha(x')$. The proof of this case is thus essentially the same as the one of Proposition 10.

We remain with the case that there exists $i\in[k_j+1,k]$ such that $x_i=x_{in(x)}\oplus e_j$. In this case, we couple (x,x'), where x' is defined by exchanging the positions of elements $x_{in(x)}$ and x_i in x, namely, $x'_{in(x)}=x_i$, $x'_i=x_{in(x)}$ and $x'_\ell=x_\ell$ for all $\ell\notin\{in(x),i\}$.

This is indeed a valid coupling since the execution of the algorithm on x' returns 0 for the same iteration j and $in(x)=in(x')$. Moreover, since α is of type K, then $\alpha_{in(x),j}=1$, but $\alpha_{i,j}=0$ (as $i\in[k_j+1,k]$). Thus,

$$\chi_\alpha(x)\chi_\alpha(x')=(-1)^{\langle\alpha,x\oplus x'\rangle_{\mathbb{F}_2}}=(-1)^{\langle\alpha_{in(x)},e_j\rangle_{\mathbb{F}_2}}(-1)^{\langle\alpha_i,e_j\rangle_{\mathbb{F}_2}}=-1\cdot 1=-1,$$

i.e., $\chi_\alpha(x)=-\chi_\alpha(x')$. This concludes the proof. ∎

Proof (of Proposition 18). First, if k_1 is odd then already $T_1(x)=0$ and $\Pr_{x\sim\mu_{n,k}}[T(x)=1]=0$.

Next, assume that k_1 is even. We prove by induction on $j\in[t]$ that

$$\Pr_{x\sim\mu_{n,k}}[T_j(x)=1]\leq \mathit{\Gamma}(N,k_1)\cdot\prod_{\ell=2}^{j}\mathit{\Lambda}(N,k_{\ell-1},k_\ell-k_{\ell-1}).$$

The result then follows since $T(x)=T_t(x)$.

For the base case of $j=1$, we have $\Pr_{x\sim\mu_{n,k}}[T_1(x)=1]\leq\mathit{\Gamma}(N,k_1)$ as in the proof of Proposition 11. For the induction step, we have

$$\Pr_{x\sim\mu_{n,k}}[T_j(x)=1]=\Pr_{x\sim\mu_{n,k}}[T_{j-1}(x)=1]\cdot\Pr_{x\sim\mu_{n,k}}[T_j(x)=1\mid T_{j-1}(x)=1].$$

Thus, we need to prove that

$$\Pr_{x\sim\mu_{n,k}}[T_j(x)=1 \mid T_{j-1}(x)=1] \le \Lambda(N,k_{j-1},k_j-k_{j-1}).$$

Fix any values for $x_1,\ldots,x_{k_{j-1}}$ which have positive probability. We prove the above inequality by taking the probability only over the selection of $x_{k_{j-1}+1},\ldots,x_{k_j}$ (which we may assume are only selected in iteration j of the algorithm).

We show that $\Lambda(N,k_{j-1},k_j-k_{j-1})$ is an upper bound on the probability to pair $(x_{k_{j-1}+1},\ldots,x_{k_j})$ within $(x_1,\ldots,x_{k_j})$ on bit j. For this purpose, we assume that all $x_1,\ldots,x_{k_{j-1}}$ are available for pairing on bit j, namely, they are not paired among themselves on bit j (this assumption can only increase the success probability of the algorithm, i.e., its pairing probability).

We upper bound $\Pr_{x\sim\mu_{n,k}}[T_j(x)=1 \mid T_{j-1}(x)=1]$ as follows: the probability that the first element in $(x_{k_{j-1}+1},\ldots,x_{k_j})$ is paired with one of the k_j-1 other elements in $x_1,\ldots,x_{k_j}$ is (at most) $\frac{k_j-1}{N-k_{j-1}-1}$. Assuming this occurs, we remove both of these elements and then the probability that the next element in $(x_{k_{j-1}+1},\ldots,x_{k_j})$ is paired is either $\frac{k_j-3}{N-k_{j-1}-3}$ (if the first element was paired among $(x_{k_{j-1}+1},\ldots,x_{k_j})$ or $\frac{k_j-3}{N-k_{j-1}-2}$ (if the first element was paired among $(x_1,\ldots,x_{k_{j-1}})$. In any case, this probability is at most $\frac{k_j-3}{N-k_{j-1}-3}$. Continue this way until all elements in $(x_{k_{j-1}+1},\ldots,x_{k_j})$ are paired. Clearly, if k_j-k_{j-1} is even, then at least $(k_j-k_{j-1})/2$ pairings are required (which occurs if $(x_{k_{j-1}+1},\ldots,x_{k_j})$ are only paired among themselves).

Taking the product of the corresponding $(k_j-k_{j-1})/2$ terms,

$$\begin{aligned}&\Pr_{x\sim\mu_{n,k}}[T_j(x)=1 \mid T_{j-1}(x)=1]\\ &\le \frac{k_j-1}{N-k_{j-1}-1}\frac{k_j-3}{N-k_{j-1}-3}\cdots\frac{k_{j-1}+1}{N-k_j+1} = \Lambda(N,k_{j-1},k_j-k_{j-1}),\end{aligned}$$

as claimed. If k_j-k_{j-1} is odd, then at least $(k_j-k_{j-1}+1)/2$ pairing are required. Similarly,

$$\begin{aligned}&\Pr_{x\sim\mu_{n,k}}[T_j(x)=1 \mid T_{j-1}(x)=1]\\ &\le \frac{k_j-1}{N-k_{j-1}-1}\frac{k_j-3}{N-k_{j-1}-3}\cdots\frac{k_{j-1}}{N-k_j} = \Lambda(N,k_{j-1},k_j-k_{j-1}).\end{aligned}$$

■

5 Bounding $\mathrm{W}^{=k}[\mu_{n,k}]$ (Proof of Lemma 2)

The goal of this section is to prove Lemma 2. We start by deriving an exact (but unwieldy) expression for $\mathrm{W}^{=k}[\mu_{n,k}]$.

Proposition 20.

$$\text{For } 0\le k\le 2^n,\ \mathrm{W}^{=k}[\mu_{n,k}] = \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}\frac{N^i}{(N)_i}.$$

Proof. For any integer $0 \le i \le k$, $\mathrm{Col}[\mu_{n,i}] = \Pr_{x,x' \sim \mu_{n,i}}[x = x'] = \frac{(N-i)!}{N!} = \frac{1}{(N)_i}$. Hence, by Proposition 4,

$$\mathrm{W}^{\le i}[\mu_{n,i}] = \mathrm{Col}[\mu_{n,i}] \cdot N^i = \frac{N^i}{(N)_i}. \tag{6}$$

For a subset $\mathcal{S} \subseteq [k]$ of size $|\mathcal{S}|$, define the functions $h(\mathcal{S}) = \mathrm{W}^{=|\mathcal{S}|}[\mu_{n,|\mathcal{S}|}]$ and $g(\mathcal{S}) = \mathrm{W}^{\le|\mathcal{S}|}[\mu_{n,|\mathcal{S}|}]$. Then, $g(\mathcal{S}) = \sum_{\mathcal{R}\subseteq\mathcal{S}} h(\mathcal{R})$, and by the inclusion-exclusion principle [15, Pg. 1049], $h(\mathcal{S}) = \sum_{\mathcal{R}\subseteq\mathcal{S}} (-1)^{|\mathcal{S}|-|\mathcal{R}|} g(\mathcal{R}) = \sum_{\mathcal{R}\subseteq\mathcal{S}} (-1)^{|\mathcal{S}|-|\mathcal{R}|} \mathrm{W}^{\le|\mathcal{R}|}[\mu_{n,|\mathcal{R}|}]$. Therefore,

$$\begin{aligned}\mathrm{W}^{=k}[\mu_{n,k}] &= h([k]) = \sum_{\mathcal{S}\subseteq[k]} (-1)^{k-|\mathcal{S}|} \mathrm{W}^{\le|\mathcal{S}|}[\mu_{n,|\mathcal{S}|}] = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \mathrm{W}^{\le i}[\mu_{n,i}] \\ &= \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \frac{N^i}{(N)_i},\end{aligned}$$

where the third equality is by the symmetry of $\mu_{n,k}$, and the final equality is by (6). ∎

The following definition will be useful in deriving a useful bound on $\mathrm{W}^{=k}[\mu_{n,k}]$ for all k.

Definition 17. *For a positive integer N and non-negative integers k, a such that $N \ge k + a$, let*

$$F_N(k, a) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \frac{N^i}{(N-a)_i}.$$

Note that by Proposition 20, $\mathrm{W}^{=k}[\mu_{n,k}] = F_N(k, 0)$. We now derive a recursive formula which will allow to analyze $\mathrm{W}^{=k}[\mu_{n,k}]$.

Proposition 21 (Recursive formula for level-k weight). *For $k \ge 2$, $F_N(k, a)$ satisfies the recurrence relation*

$$F_N(k, a) = \frac{a}{N-a} \cdot F_N(k-1, a+1) + \frac{(k-1)N}{(N-a)(N-a-1)} \cdot F_N(k-2, a+2),$$

with the starting conditions $F_N(0, a) = 1$ and $F_N(1, a) = \frac{N}{N-a} - 1 = \frac{a}{N-a}$.

Proof. The starting conditions are easily checked by plugging in the parameters into the explicit formula for $F_N(k, a)$. We now prove the recurrence relation holds assuming $k \ge 2$.

To simplify notation, denote $G_i = \frac{N^i}{(N-a)_i}$ and write $F_N(k, a) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} G_i$. For $1 \le i \le k-1$, substitute $\binom{k}{i} = \binom{k-1}{i} + \binom{k-1}{i-1}$ and $\binom{k}{0} = \binom{k-1}{0}, \binom{k}{k} = \binom{k-1}{k-1}$

into the expression, which divides each term into a pair of terms. We obtain

$$\begin{aligned}
F_N(k,a) &= \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}G_i \\
&= \left((-1)^k\binom{k-1}{0}\cdot G_0 + (-1)^{k-1}\binom{k-1}{0}G_1\right) \\
&+ \left((-1)^{k-1}\binom{k-1}{1}G_1 + (-1)^{k-2}\binom{k-1}{1}G_2\right) \\
+\ldots+ &\left((-1)^{k-(k-1)}\binom{k-1}{k-1}G_{k-1} + (-1)^{k-k}\binom{k-1}{k-1}G_k\right) \\
&= \sum_{i=1}^{k}(-1)^{k-i}\binom{k-1}{i-1}\left(G_i - G_{i-1}\right).
\end{aligned}$$

We have $G_i = G_{i-1}\cdot\frac{N}{N-a-(i-1)}$, so $G_i - G_{i-1} = G_{i-1}\cdot(\frac{N}{N-a-(i-1)} - 1) = G_{i-1}\cdot\frac{a+(i-1)}{N-a-(i-1)}$. Therefore, the above expression is equal to

$$\begin{aligned}
&\sum_{i=1}^{k}(-1)^{k-i}\binom{k-1}{i-1}G_{i-1}\cdot\frac{a+(i-1)}{N-a-(i-1)} \\
&= \sum_{i=1}^{k}(-1)^{k-i}\binom{k-1}{i-1}\frac{(a+(i-1))N^{i-1}}{(N-a)(N-a-1)\ldots(N-a-(i-1))} \\
&= \frac{1}{N-a}\cdot\sum_{i=1}^{k}(-1)^{k-i}\binom{k-1}{i-1}\frac{(a+(i-1))N^{i-1}}{(N-a-1)_{i-1}} \\
&= \frac{1}{N-a}\cdot\sum_{i=0}^{k-1}(-1)^{k-1-i}\binom{k-1}{i}\frac{(a+i)N^{i}}{(N-a-1)_{i}} \\
&= \frac{a}{N-a}\cdot\sum_{i=0}^{k-1}(-1)^{k-1-i}\binom{k-1}{i}\frac{N^{i}}{(N-a-1)_{i}} \\
&+ \frac{1}{N-a}\cdot\sum_{i=1}^{k-1}(-1)^{k-1-i}\binom{k-1}{i}\frac{i\cdot N^{i}}{(N-a-1)_{i}} \\
&= \frac{a}{N-a}\cdot F_N(k-1,a+1) \\
&+ \frac{N}{(N-a-1)(N-a)}\cdot\sum_{i=1}^{k-1}(-1)^{k-1-i}\binom{k-1}{i}\frac{i\cdot N^{i-1}}{(N-a-1)_{i-1}}.
\end{aligned}$$

To complete the proof, it remains to show that

$$\sum_{i=1}^{k-1}(-1)^{k-1-i}\binom{k-1}{i}\frac{i\cdot N^{i-1}}{(N-a-1)_{i-1}} = (k-1)\cdot F_N(k-2,a+2).$$

Observe that $i \cdot \binom{k-1}{i} = (k-1) \cdot \binom{k-2}{i-1}$. Therefore,

$$\begin{aligned}
&\sum_{i=1}^{k-1} (-1)^{k-1-i} \binom{k-1}{i} \frac{i \cdot N^{i-1}}{(N-a-1)_{i-1}} \\
&= (k-1) \cdot \sum_{i=1}^{k-1} (-1)^{k-1-i} \binom{k-2}{i-1} \frac{N^{i-1}}{(N-a-1)_{i-1}} \\
&= (k-1) \cdot \sum_{i=0}^{k-2} (-1)^{k-i} \binom{k-2}{i} \frac{N^{i}}{(N-a-2)_{i}} \\
&= (k-1) \cdot F_N(k-2, a+2).
\end{aligned}$$

This completes the proof. ■

Next, we use the recurrence relation to bound $F_N(k, a)$.

Proposition 22.

$$F_N(k,a) \leq \begin{cases} \frac{(N(a+k-1))^{k/2}}{(N-a)_k} & \text{if } k \text{ is even,} \\ \frac{(N(a+k-1))^{(k-1)/2} \cdot (a+k-1)}{(N-a)_k} & \text{if } k \text{ is odd.} \end{cases}$$

Proof. We prove the result using Proposition 21 by induction on k. It is easy to verify that it holds for $k = 0$ and $k = 1$ by the starting conditions. We prove the induction step.

If k is odd, then by the assumption

$$\begin{aligned}
&F_N(k,a) = \frac{a}{N-a} \cdot F_N(k-1, a+1) + \frac{(k-1)N}{(N-a)(N-a-1)} \cdot F_N(k-2, a+2) \\
&\leq \frac{a}{N-a} \cdot \frac{(N(a+k-1))^{(k-1)/2}}{(N-a-1)_{k-1}} \\
&+ \frac{(k-1)N}{(N-a)(N-a-1)} \cdot \frac{(N(a+k-1))^{(k-3)/2}(a+k-1)}{(N-a-2)_{k-2}} \\
&= a \cdot \frac{(N(a+k-1))^{(k-1)/2}}{(N-a)_k} + (k-1) \cdot \frac{(N(a+k-1))^{(k-1)/2}}{(N-a)_k} \\
&= \frac{(N(a+k-1))^{(k-1)/2} \cdot (a+k-1)}{(N-a)_k},
\end{aligned}$$

as desired. If k is even, then

$$\begin{aligned}
&F_N(k,a) \le \frac{a}{N-a} \cdot \frac{(N(a+k-1))^{(k-2)/2} \cdot (a+k-1)}{(N-a-1)_{k-1}} \\
&+ \frac{(k-1)N}{(N-a)(N-a-1)} \cdot \frac{(N(a+k-1))^{(k-2)/2}}{(N-a-2)_{k-2}} \\
&= a \cdot \frac{(N(a+k-1))^{(k-2)/2} \cdot (a+k-1)}{(N-a)_k} + (k-1)N \cdot \frac{(N(a+k-1))^{(k-2)/2}}{(N-a)_k} \\
&= \frac{(N(a+k-1))^{(k-2)/2}}{(N-a)_k} \cdot (a(a+k-1) + (k-1)N).
\end{aligned}$$

It remains to prove that $a(a+k-1)+(k-1)N \le N(a+k-1)$ or $a+k-1 \le N$, which indeed holds (as the quantity $a+k$ is preserved throughout the recursive calls). ■

Finally, Lemma 2 is proved in Appendix B by straightforward manipulation of the bound on $F_N(k,a)$ of Proposition 22, and based on the fact that by Proposition 20, $\mathrm{W}^{=k}[\mu_{n,k}] = F_N(k,0)$.

Acknowledgements. The author was supported by the Israel Science Foundation through grant no. 1903/20. The author would like to thank Samuel Neves for pointing him to the prior works [13,14].

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. p. 24 (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. 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
4. Bhattacharya, S., Nandi, M.: Luby-Rackoff Backwards with More Users and More Security. In: Tibouchi, M., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2021 - 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6-10, 2021, Proceedings, Part III. Lecture Notes in Computer Science, vol. 13092, pp. 345–375. Springer (2021). https://doi.org/10.1007/978-3-030-92078-4_12
5. 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

6. 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
7. 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
8. 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
9. 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
10. Daemen, J., Rijmen, V.: Probability distributions of correlation and differentials in block ciphers. J. Math. Cryptol. **1**(3), 221–242 (2007). https://doi.org/10.1515/JMC.2007.011
11. Dai, W., Hoang, V.T., Tessaro, S.: Information-Theoretic Indistinguishability via the Chi-Squared Method. In: Katz, J., Shacham, H. (eds.) Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part III. Lecture Notes in Computer Science, vol. 10403, pp. 497–523. Springer (2017). https://doi.org/10.1007/978-3-319-63697-9_17
12. 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
13. Eberhard, S.: More on additive triples of bijections (2017), https://arxiv.org/abs/1704.02407
14. 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
15. Gessel, I.M., Stanley, R.P.: Handbook of Combinatorics, vol. 2, chap. Algebraic enumeration, p. 1021–1061. MIT Press, Cambridge, MA, USA (1996)
16. Gunsing, A., Bhaumik, R., Jha, A., Mennink, B., Shen, Y.: Revisiting the indifferentiability of the sum of permutations. In: Handschuh, H., Lysyanskaya, A. (eds.) CRYPTO 2023. Lecture Notes in Computer Science, vol. 14083, pp. 628–660. Springer (2023). https://doi.org/10.1007/978-3-031-38548-3_21
17. Gunsing, A., Mennink, B.: The Summation-Truncation Hybrid: Reusing Discarded Bits for Free. In: Micciancio, D., Ristenpart, T. (eds.) CRYPTO 2020. Lecture Notes in Computer Science, vol. 12170, pp. 187–217. Springer (2020). https://doi.org/10.1007/978-3-030-56784-2_7
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. Liu, T., Tessaro, S., Vaikuntanathan, V.: The t-wise Independence of Substitution-Permutation Networks. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021. Lecture Notes in Computer Science, vol. 12828, pp. 454–483. Springer (2021). https://doi.org/10.1007/978-3-030-84259-8_16

20. 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
21. Maurer, U.M., Pietrzak, K., Renner, R.: Indistinguishability Amplification. In: Menezes, A. (ed.) CRYPTO 2007. Lecture Notes in Computer Science, vol. 4622, pp. 130–149. Springer (2007). https://doi.org/10.1007/978-3-540-74143-5_8
22. 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
23. O'Connor, L.: Properties of Linear Approximation Tables. In: Preneel, B. (ed.) FSE 1994. Lecture Notes in Computer Science, vol. 1008, pp. 131–136. Springer (1994). https://doi.org/10.1007/3-540-60590-8_10
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.) Information Theoretic Security, Third International Conference, ICITS 2008, Calgary, Canada, August 10-13, 2008, Proceedings. 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.: Generic Attacks for the Xor of k random permutations. IACR Cryptol. ePrint Arch. p. 9 (2008), http://eprint.iacr.org/2008/009
27. 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
28. 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 from Section 4

Proof (of Lemma 1). We use Proposition 15 to bound $\mathrm{M}^{=k}[\mu_{n,k}]$. First we have $\mathrm{M}^{=2}[\mu_{n,2}] \le \Gamma(N,2) = \frac{1}{N-1}$.

Next, for even k, by Proposition 15,

$$\mathrm{M}^{=k}[\mu_{n,k}]^2 \le \Gamma(N,k)^2 = \left(\frac{k-1}{N-1}\right)^2 \left(\frac{k-3}{N-3}\right)^2 \cdots \left(\frac{1}{N-(k-1)}\right)^2$$
$$\le \frac{k}{N}\frac{k-1}{N-1}\frac{k-2}{N-2}\frac{k-3}{N-3}\cdots\frac{2}{N-(k-2)}\frac{1}{N-(k-1)} = \frac{1}{\binom{N}{k}}.$$

Similarly, for odd k,

$$\mathrm{M}^{=k}[\mu_{n,k}]^2 \le \Gamma(N,k-1)^2 \cdot \left(\frac{k}{N-k}\right)^2 \le \frac{1}{\binom{N}{k-1}} \cdot \left(\frac{k}{N-k}\right)^2$$
$$= \frac{1}{\binom{N}{k}} \cdot \frac{N-(k-1)}{k}\left(\frac{k}{N-k}\right)^2 = \frac{1}{\binom{N}{k}} \cdot \frac{N-(k-1)}{N-k}\frac{k}{N-k}$$
$$< \frac{1}{\binom{N}{k}} \cdot \frac{N-(k-1)}{N-k}\frac{k+1}{N-(k-1)} = \frac{1}{\binom{N}{k}} \cdot \frac{k+1}{N-k}.$$

■

Proof (of Proposition 19). First assume that k is even. If $t = 1$, then the claim holds with equality. We thus assume that $t > 1$. For the purpose of analyzing the expressions $\Lambda(N, k_{j-1}, k_j - k_{j-1})$ it is convenient to order the numerators of their terms (of the form a/b) from smallest to largest. Specifically, for even $k_j - k_{j-1}$ write

$$\Lambda(N, k_{j-1}, k_j - k_{j-1}) = \frac{k_j - 1}{N - k_{j-1} - 1} \frac{k_j - 3}{N - k_{j-1} - 3} \cdots \frac{k_{j-1} + 1}{N - k_j + 1}$$
$$= \frac{k_{j-1} + 1}{N - k_{j-1} - 1} \frac{k_{j-1} + 3}{N - k_{j-1} - 3} \cdots \frac{k_j - 1}{N - k_j + 1}.$$

and reorder similarly for odd $k_j - k_{j-1}$. Moreover, write $\Gamma(N, k_1) = \frac{1}{N-1} \frac{3}{N-3} \cdots \frac{k_1 - 1}{N - k_1 + 1}$.

We have

$$\Gamma(N, k) = \frac{1}{N - 1} \frac{3}{N - 3} \cdots \frac{k - 1}{N - k + 1}.$$

Hence, $\Gamma(N, k)$ is a product of $k/2$ terms of the form a/b.

Note that every $\Lambda(N, k_{j-1}, k_j - k_{j-1})$ is a product of at least $(k_j - k_{j-1})/2$ terms, hence the expression $\Gamma(N, k_1) \cdot \prod_{j=2}^{t} \Lambda(N, k_{j-1}, k_j - k_{j-1})$ is a product of at least $k/2$ terms.

We claim that the product of the first (i.e., largest) $k/2$ denominators in the expression $\Gamma(N, k_1) \cdot \prod_{j=2}^{t} \Lambda(N, k_{j-1}, k_j - k_{j-1})$ is at least $(N-1)(N-3)\ldots(N-k+1)$, which is the product of denominators in $\Gamma(N, k)$. Indeed, taking the first $k/2$ denominators in the order they appear in $\Gamma(N, k_1)\Lambda(N, k_1, k_2 - k_1)\ldots\Lambda(N, k_{t-1}, k_t - k_{t-1})$, the first denominator in $\Gamma(N, k_1)$ is $N - 1$, and we claim that consecutive denominators do not drop by more than 2.

This is clear for consecutive denominators inside $\Gamma(N, k_1)$ and $\Lambda(N, k_{j-1}, k_j - k_{j-1})$, which drop by exactly 2. Moreover, the smallest denominator in $\Lambda(N, k_{j-1}, k_j - k_{j-1})$ is either $N - k_j + 1$ or $N - k_j$ (depending on the parity of $k_j - k_{j-1}$), while the largest denominator in $\Lambda(N, k_j, k_{j+1} - k_j)$ is $N - k_j - 1 = (N - k_j + 1) - 2$. This is also true when considering the smallest denominator of $\Gamma(N, k_1)$ and the largest denominator of $\Lambda(N, k_1, k_2 - k_1)$.

Next, we claim that the product of the first (i.e., smallest) $k/2$ numerators is at most $1 \cdot 3 \ldots (k-1)$, which is the product of numerators in $\Gamma(N, k)$. Indeed, the smallest numerator in $\Gamma(N, k_1)$ is 1 and we claim that consecutive numerators do not increase by more than 2.

This is clear for consecutive numerators inside $\Gamma(N, k_1)$ and $\Lambda(N, k_{j-1}, k_j - k_{j-1})$, which increase by exactly 2. Moreover, the largest numerator in $\Lambda(N, k_{j-1}, k_j - k_{j-1})$ is $k_j - 1$, while the smallest numerator in $\Lambda(N, k_j, k_{j+1} - k_j)$ is either k_j or $k_j + 1 = (k_j - 1) + 2$ (depending on the parity of $k_{j+1} - k_j$). This also holds for the largest numerator of $\Gamma(N, k_1)$ and the smallest numerator of $\Lambda(N, k_1, k_2 - k_1)$.

Finally, after factoring out the $k/2$ largest denominators and $k/2$ smaller numerators in $\Gamma(N, k_1) \cdot \prod_{j=2}^{t} \Lambda(N, k_{j-1}, k_j - k_{j-1})$, we remain with an expression that is smaller than 1, as the smallest remaining denominator is at least $N - k$

and the largest remaining numerator is at most $k < N - k$, as $k < N/2$. This proves the claim for even k.

The proof for odd k is similar with the additional observation that $\Gamma(N, k_1) \cdot \prod_{j=2}^{t} \Lambda(N, k_{j-1}, k_j - k_{j-1})$ is a product of at least $(k+1)/2$ terms as k_1 is even. Once again, comparing the largest denominators and smallest numerators to the $(k+1)/2$ terms of $\Gamma(N, k-1) \cdot \frac{k}{N-k}$ proves the result. ■

B Missing Proofs from Section 5

Proof (of Lemma 2). By Proposition 20, $\mathrm{W}^{=k}[\mu_{n,k}] = F_N(k, 0)$. Therefore, the equalities for $\mathrm{W}^{=k}[\mu_{n,k}]$ where $k \in \{1, 2, 3\}$ and the first part of the inequalities for $\mathrm{W}^{=k}[\mu_{n,k}]$ directly follow from Proposition 21 and Proposition 22, respectively. Below we prove that for even k, $\frac{(N(k-1))^{k/2}}{(N)_k} \leq \Psi_N(k)$. Then, for odd k, we have

$$\mathrm{W}^{=k}[\mu_{n,k}] \leq \frac{(N(k-1))^{(k-1)/2} \cdot (k-1)}{(N)_k} \leq \frac{(Nk)^{(k+1)/2}}{(N)_{k+1}} \leq \Psi_N(k+1),$$

as claimed.

By rearranging terms we have

$$\begin{aligned} &\frac{(N(k-1))^{k/2}}{(N)_k} \leq \frac{(Nk)^{k/2}}{(N)_k} \\ &= \frac{(Nk)^{k/2}}{(N(N-k)) \cdot ((N-1)(N-k+1)) \cdot \ldots \cdot (N-k/2+1)(N-k/2)} \qquad (7) \\ &= \prod_{i=0}^{(k/2)-1} \frac{Nk}{(N-i)(N-k+i)}. \end{aligned}$$

We now analyze (the inverse of) each term of the product (recalling that $i < k/2$).

$$\begin{aligned} \frac{(N-i)(N-k+i)}{Nk} &= \frac{N^2 - Nk + ik - i^2}{Nk} = \frac{N-k}{k} + \frac{i(k-i)}{Nk} \\ &\geq \frac{N-k}{k} + \frac{ik}{2Nk} = \frac{N-k}{k} + \frac{i}{2N} = \frac{N-k}{k}\left(1 + \frac{ik}{2N(N-k)}\right). \end{aligned}$$

Next, using the fact that for every $x > -1$, $1 + x \geq \exp(x/(x+1))$, and that $i < k/2$, we upper bound the expression above as

$$\geq \frac{N-k}{k} \exp\left(\frac{ik}{2N(N-k)+ik}\right) \geq \frac{N-k}{k} \exp\left(\frac{ik}{2N(N-k)+k^2/2}\right)$$

Plugging this into (7), we obtain

$$\frac{(N(k-1))^{k/2}}{(N)_k} \leq \left(\frac{k}{N-k}\right)^{k/2} \exp\left(-\sum_{i=0}^{(k/2)-1} \frac{ik}{2N(N-k)+k^2/2}\right)$$

$$= \left(\frac{k}{N-k}\right)^{k/2} \exp\left(-\frac{k^2((k/2)-1)/4}{2N(N-k)+k^2/2}\right)$$

$$= \left(\frac{k}{N-k}\right)^{k/2} \exp\left(-\frac{k(k-2)}{8N(N-k)+2\cdot k^2}\right)^{k/2} = \Psi_N(k).$$

■