Concentration bounds for almost k-wise independence with applications to non-uniform security [*]

Nick Gravin[†] Siyao Guo[‡] Tsz Chiu Kwok[†] Pinyan Lu[†]

Abstract

We prove a few concentration inequalities for the sum of n binary random variables under weaker conditions than k-wise independence. Namely, we consider two standard conditions that are satisfied in many applications: (a) direct product conditions (b) the XOR condition. Both conditions are weaker than mutual independence and both imply strong concentration bounds (similar to Chernoff-Hoeffding) on the tail probability of the sum of bounded random variables ([Impagliazzo and Kabanets, APPROX-RANDOM 10], [Unger, FOCS 09]). Our inequalities can be stated as the implication of threshold direct product theorems from either k-wise direct product conditions, or the k-wise XOR condition. By proving optimality of our inequalities, we show a clear separation for $k \ll n$ between k-wise product conditions and XOR condition as well as a stark contrast between k-wise and n-wise product theorems.

We use these bounds in the cryptographic application that provides provable security against algorithms with S-bit advice. Namely, we show how the problem reduces to proving S-wise direct product theorems or S-wise XOR lemmas for certain ranges of parameters. Finally, we derive a new S-wise XOR lemma, which yields a tight non-uniform bound for length increasing pseudorandom generators, resolving a 10-year-old open problem from [De, Trevisan, and Tulsiani, CRYPTO 10].

1 Introduction

Concentration inequalities such as Chernoff [Che52] and Hoeffding [Hoe63] bounds on the tail probability of the sums of bounded random variables are important tools in the design and analysis of many algorithms, security protocols, and complexity theory. These bounds require independence or similar weaker conditions on the random variables, e.g., mutual independence in Chernoff-Hoeffding and k-wise independence in [SSS93]. In fact, it might not be wrong to say that the major application of independence condition for a set of random variables is some sort of concentration inequality.

Different relaxations of independence assumption in Chernoff-Hoeffding (CH) inequality to weaker conditions are explicitly studied in a few independent lines of work. One such important relaxation is k-wise independence. Indeed, Schmidt et al. [SSS93] used similar to CH exponential concentration inequality for k-wise independent random variables to improve performance of job scheduling and oblivious packet routing algorithms, while Bellare and Rompel [BR94] used concentration inequalities for k-wise independent random variables to reduce the number of rounds and amount of randomness in Arthur-Merlin protocols. Another type of important relaxations considers approximate versions of the characteristic property of the product distribution, e.g., such as direct product condition $\mathbf{Pr}[\bigwedge_{i\in I} X_i = 1] = \prod_{i\in I} \mathbf{Pr}[X_i = 1]$ that is equivalent to mutual independence for binary random variables. These relaxations are needed in complexity theory for the tasks of hardness amplification (Direct Product Theorems) and are used, for example, in the cryptographic application CAPTCHA. In particular, Impagliazzo and Kabanets [IK10] show that CH type bounds corresponds to threshold direct product theorem and can be obtained from the approximate product condition $\forall I \subseteq [n]\, \mathbf{Pr}[\bigwedge_{i\in I} X_i = 1] \le \delta^{|I|}$ for some $\delta < 1$. Similarly, in an earlier paper Unger [Ung09] showed that CH bound follows from a weaker than independence condition on the co-moments, i.e., a bound on the bias of XORs for any subset of random variables $\forall I \subseteq [n]\, \mathrm{Bias}(\oplus_{i\in I} X_i) \le \beta^{|I|}$. The condition that can be verified in many application domains by using various XOR lemmas.

These relaxation can be roughly divided into two trends: k-wise independence and approximation, both trends having good motivation and practical significance.

[*] This work is supported by Science and Technology Innovation 2030 –"New Generation of Artificial Intelligence" Major Project No.(2018AAA0100903), Program for Innovative Research Team of Shanghai University of Finance and Economics (IRTSHUFE) and the Fundamental Research Funds for the Central Universities.

[†] {nikolai,kwok,lu.pinyan}@mail.shufe.edu.cn, ITCS, Shanghai University of Finance and Economics.

[‡] siyao.guo@nyu.edu, New York University Shanghai. Supported by Shanghai Eastern Young Scholar Program (SMEC No.0920000169).

Motivation for k-wise independence. Exponential concentration bounds show their true power only on the instances with large number n of random variables, i.e., they apply only to high-dimensional joint distributions. The mutual independence condition imposes strong requirements on the joint distribution and often is too difficult to achieve in practice. The k-wise relaxation for k much smaller than n allows to mitigate these requirements. For example, a basic task of testing whether n-dimensional distribution is k-wise independent has computational and sample complexity of $n^{O(k)}$ [AAK+07] for fixed k. I.e., it is effectively impossible to test if many variables are mutually independent, while it is still manageable to test k-wise independence for small k.

Constructing a high-dimensional product distribution is quite costly, as it has high entropy and requires $\Omega(n)$ random bits. Thus k-wise distributions are more common in applications such as hashing and pseudo random generators. Indeed, there are many constructions for k-wise independent distributions that utilize significantly lower amount of randomness.

Motivation for approximation. Mutual independence and k-wise independence of random variables are often idealistic abstractions that only approximately capture reality. It is especially true in complexity theory and cryptographic applications, where many primitives entail certain margin for errors. For example, the direct product theorems (e.g., [Imp95]) and XOR lemmas (Yao's XOR Lemma, Vazirani's Parity Lemma [Vaz87], or Unger's XOR lemma [Ung09]) mentioned earlier use approximate bounds. These statements allow to amplify hardness of many cryptographic constructions and thereby achieve better security guarantees (see, e.g., [MT09]). These approximation results are especially useful in practice when they require only local conditions, i.e., when they hold only for small subsets of variables similar to k-wise independence (e.g., [Imp95]).

Motivation for k-wise approximation. Interestingly, our initial motivation for studying k-wise approximation comes from a fundamental and challenging problem in a seeming unrelated context — proving *tight non-uniform security bounds* for one of the basic cryptographic primitives — pseudorandom generators (PRGs). Along the way, we discover a fruitful connection between concentration bounds for k-wise approximations and analysis techniques for non-uniform security, illustrated below.

Cryptography usually models the attacker as non-uniform, meaning he can obtain an arbitrary (but bounded) advice before attacking the system. Specifically, a non-uniform PRG attacker $\mathcal{A}$ for a length increasing function $\mathcal{O} : [N] \to [2N]$ consists of an offline algorithm $\mathcal{A}_1$ and an online algorithm $\mathcal{A}_2$. In the preprocessing stage, $\mathcal{A}_1$ makes arbitrary number of queries to $\mathcal{O}$ and produces S bits of advice. Then $\mathcal{A}_2$ uses the advice and makes T queries to distinguish whether a given y is a random image of $\mathcal{O}$, or a uniform random string from $[2N]$.

De, Trevisan, and Tulsiani [DTT10] showed that an attacker with an S-bit advice and $T = O(1)$ queries, can achieve advantage $\Omega(\sqrt{S/N})$ for any given $\mathcal{O}$. In addition, they gave an $O(\sqrt{ST/N})$ upper bound on the advantage, even for the attacks in a special case of $\mathcal{O} = (f, P)$ where $f : [N] \to [N]$ is a random permutation, and $P : [N] \to \{0, 1\}$ is a random predicate. Recently, Dodis, Guo and Katz [DGK17], Coretti, Dodis, Guo, and Steignberger [CDGS18] proved the same security bound (i.e., the upper bound on the advantage) for the special case of a random function $\mathcal{O}$.

These security bounds only match the attack by De et al. in the extremal case of $T = O(1)$, i.e., only for a constant number of queries. De et al. left an intriguing open question of what parameters' range can lead to distinguishability despite that inversion of one-way permutations or functions is impossible. Specifically, is the advantage $\Omega(\sqrt{T/N})$ achievable for $S = O(1)$?

Surprisingly, even for the extremal case of $S = 1$, there has been no progress on either attacks, or security bounds in the past decade. One difficulty in obtaining tight non-uniform security is the lack of applicable techniques. Unlike the uniform setting (i.e., no advice is allowed), there are only two major techniques, imcompressibility argument [DTT10, DGK17, CK18] and the presampling technique [Unr07, CDGS18, CDG18] in the non-uniform setting. Both of them fail to obtain better security bounds for the PRG problem.

Do we have other techniques? An elegant and short proof for tight non-uniform security of the one way permutation (OWP) problem by Impagliazzo [Imp11] is one such example[1]. Moreover, this proof has a strong concentration bound flavor. In particular, Impagliazzo showed that, any adversary without advice can achieve advantage ε (i.e., invert at least ε fraction of points in $[N]$), for at most $2^{-S}/N$ fraction of a random permutation $f : [N] \to [N]$ (for some parameters $0 < \varepsilon < 1$). Then, by a simple union bound, any adversary with S-bit advice can achieve advantage ε for at most $1/N$ fraction of a random permutation. Therefore, any adversary with S-bit advice achieves at most $\varepsilon + 1/N \approx \varepsilon$ (ε usually dominates $1/N$) advantage

[1] The proof is included in the appendix [Imp11] and stated for random injective functions.

for a random permutation

It is somewhat surprising that the simple idea of union bound works so nicely. It reduces the task from showing security for attackers with S-bit advice, to showing a roughly 2^{-S} concentration bound on the advantages for attackers with *no advice*. Another useful insight from [Imp11] is that, for any choice of permutation f, the advantage of a OWP attacker can be written as an average of $\{0,1\}$ variables $X_1,\dots,X_N$ where X_i indicates whether the attacker succeeds in inverting i (i.e., outputting $f^{-1}(i)$). Using this language, Impagliazzo [Imp11] showed that $X_1+\cdots+X_N \geq \varepsilon N$ happens with probability at most $2^{-S}/N$ over the randomness of f.

How to prove such concentration bounds for $X_1,\dots,X_N$? No independence condition is guaranteed. Impagliazzo [Imp11] showed that those variables satisfy the following condition, for every subset $I \subseteq [N]$ of size k, $\mathbf{Pr}[\Pi_{i\in I} X_i = 1] \leq (6kT/N)^k$, which does not give N-wise approximation because the inside factor becomes worse when size of I grows and becomes trivial when $|I| \geq N/6T$. By using above condition for sets of size $(S+\log N)$, Impagliazzo [Imp11] showed the desired concentration bounds, which yield a tight non-uniform security bounds for one-way permutations.

The proof strategy of Impagliazzo [Imp11] can be adapted to general cryptographic applications. The success probability of any (deterministic) attacker is typically captured by the fraction of solved challenges among N possible challenges, which can be written as the averaging of $X_1,\dots,X_N$ where $X_1,\dots,X_N$ are $\{0,1\}$ variables indicating whether $\mathcal{A}$ succeeds in the i-th challenge. Hence, proving non-uniform security bounds reduces to proving $2^{-S}/N$ concentration bounds for the event of $X_1+\cdots+X_N \geq \varepsilon N$.

We employ similar to [Imp11] proof strategy to an equivalent variant of PRG problem, called hardcore predicate problem. We obtain corresponding $X_1,\dots,X_N$ variables and observe that they satisfy a variety of approximate conditions including $(1/2 - 6kT/N)^k \leq \mathbf{Pr}[\Pi_{i\in I} X_i = 1] \leq (1/2 + 6kT/N)^k$ and more. To prove the tight bound, we would need a tail bound as strong as for perfect k-wise independent variables $X_1,\dots,X_N$. This motivates us to study k-wise approximation notions with strong concentration bounds.

In this paper we investigate the power of different approximation notions for k-wise independence. Our goal is to find conditions that (i) yield exponential concentration inequalities similar to CH bounds (ii) hold in various cryptographic application. In particular, we compare to the concentration bound for the exact k-wise independence of Schmidt et al. [SSS93] (Theorem 5 case I.(a)): if $X = X_1+\ldots+X_n$ are k-wise independent $\{0,1\}$ random variables with $\mathbf{E}[X_i] = 1/2$, then $\mathbf{Pr}[|X - n/2| > O(\sqrt{kn})] \leq e^{-\lfloor k/2 \rfloor}$. As discussed above, we are explicitly interested in the application of proving tight non-uniform security bounds for PseudoRandom Generators (PRGs).

1.1 Our results The first part of our results is dedicated to finding a good notion of approximation for the k-wise independence condition on a set of n random variables. Our goal is to identify a condition that can be verified in the wide range of applications and would imply a strong enough concentration, i.e., similar to CH a bound which is exponential in n and k. To this end, we do a "case study" on the most basic setting with n binary random variables $X_1,\dots,X_n$ with values in $\{0,1\}$. As our motivation comes from cryptographic applications, we take a special interest in the unbiased case when $\mathbf{Pr}[X_i = 1]$ is close to $1/2$, but we consider other regimes as well.

An obvious candidate for the approximation of k-wise independence is the *direct product condition* (corresponds to direct product theorems) that $a^{|I|} \leq \mathbf{Pr}[\Pi_{i\in I} X_i = 1] \leq b^{|I|}$ for any subset I of size at most k of n variables and some $a < b < 1$. We would like to bound the probability that the average value is larger than $c > b$: $\mathbf{Pr}[\sum_{i=1}^n X_i > cn]$. One of the most common regimes of parameters would be $a = (1-\varepsilon)/2$, $b = (1+\varepsilon)/2$, and $c = (1+\gamma)/2$ where $\gamma > 2\varepsilon$ (ideally, $\varepsilon, \gamma = o(1)$ decrease with n). We also consider the *one-sided* version of the product condition, as in some applications it is easier to care only about one side of the inequality in the direct product condition, namely that $\mathbf{Pr}[\Pi_{i\in I} X_i = 1] \leq b^{|I|}$. Finally, we look at the *k-wise XOR condition* (similar in spirit to [Ung09]) corresponding to XOR lemmas, where the XOR of $\{X_i\}_{i\in I}$ in any subset I of size at most k satisfies $\mathbf{Pr}[\oplus_{i\in I} X_i = 1] - \mathbf{Pr}[\oplus_{i\in I} X_i = 0] \leq \varepsilon^{|I|}$.

Let $X_1,\dots,X_n$ be $\{0,1\}$ random variables and $X = X_1+\ldots+X_n$. For different approximation conditions on k-wise independence we have

One-sided direct product. A *tight* upper bound on $\mathbf{Pr}[X > c\cdot n]$ of order $\approx \left(\frac{b}{c}\right)^k$ in Theorem 3.1. This is rather weak bound. As an example, to compare it to the bound for exact k-wise independent condition, we set $b = 1/2$, $c = 1/2 + O(\sqrt{k/n})$. Then instead of $e^{-\lfloor k/2 \rfloor}$, we get the bound of $(1 - O(\sqrt{k/n}))^k = 1 - o(1)$ for $k = o(n^{1/3})$.

Two-sided direct product. This condition is a proper approximation to k-wise independence for binary random variables, as it coincides with k-wise independence when $a = b$. If $a = b$, the results of Schmidt et al. [SSS93] imply a $e^{-\lfloor k/2 \rfloor}$ bound. An important question is how the bound degrades with the approximation depending on the gap between a and b.

1. First, we derive a succinctly represented LP and its dual (Lemma 3.1), that describes tight bounds on the tail probability. The LP can be generalized to a broader set of approximate product type conditions.
2. We apply these LPs to two-sided (a, b)-product condition and derive almost tight bounds in Theorem 3.2 for the approximation gap as small as $b - a > k/n$. [2] In particular, when compared to the bound in [SSS93], Theorem 3.2 implies that $\mathbf{Pr}[X > c \cdot n]$ may be as large as $\Omega(1/k) \gg e^{-\Omega(k)}$, for $a = 1/2 - k/n$, $b = 1/2 + k/n$, and $c = 1/2 + O(\sqrt{k/n})$.

XOR condition. We show in Theorem 3.3 that XOR condition gives the following tail bound

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \geq \frac{n(1+\gamma)}{2}\right] \leq \left(\frac{\varepsilon + \sqrt{\frac{k}{n}}}{\gamma}\right)^k .$$

Note that this bound matches the bound of [SSS93] for unbiased variables, when approximation error $\varepsilon = 0$. In a stark contrast to approximate k-wise direct product condition, this bound is of order $e^{-\Omega(k)}$ in the interesting regime $\gamma > 2\varepsilon$ and $\gamma = \Omega(\sqrt{k/n})$, i.e., for the same deviation from the mean as in [SSS93] and approximation error within a constant factor from the deviation's threshold.

For completeness, we consider a generalized version of the XOR condition for the collection of arbitrary bounded random variables. We show a similar to Theorem 3.3 result Theorem B.1 in Appendix B.

It should not be too surprising that the one-sided version of the k-wise direct product condition does not imply good concentration. It is interesting that the two-sided (a, b)-approximation of the k-wise independence suffers from a similar problem for any reasonable approximation error ($b-a > k/n$). This is surprising in the light of Impagliazzo and Kabanets [IK10] results, who showed that the approximate direct product condition for n-wise (mutual) independence implies CH bound. On the other hand, the k-wise XOR condition does yield the desired tail bound for all values of k. Thus our bounds effectively dismiss the approximation of k-wise independence via the direct product condition (unless the approximation error is very small), and show that k-wise XOR condition is most useful for small k.

Related works on almost k-wise independence. There are other notions of almost k-wise independence in the literature. Naor and Naor [NN93] define k-wise ϵ-biased variables, which requires the bias of every subset I of size $\leq k$ to be at most ϵ (instead of $\epsilon^{|I|}$ in the XOR condition), and provide a concentration bound with this assumption. Alon et al. [AGHP92] define almost k-wise independent variables to be for every subset I of size $\leq k$, $|\mathbf{Pr}[\prod_{i \in I} X_i = 1] - a^{|I|}| \leq \epsilon$. Under this assumption, Kabanets [Kab03] proves a bound on $\mathbf{Pr}[X_i = 1\ \forall i]$. Both of the bounds cannot be compared directly to our works.

Assuming the one-sided direct product condition, techniques in multiple works [SSS93, IK10, LL14, PR17] are able to give the concentration bound in Theorem 3.1. Our work provides also a matching example to show that the bound is tight.

Non-uniform security of PRGs. In our second set of results, we focus on the application of our concentration inequalities for k-wise XOR condition to non-uniform security in pseudorandom generators. We obtain the following security guarantee. The notation $\widetilde{O}(\cdot)$ hides factors that are polynomial in $\log N$.

Theorem 1.1. (Pseudorandom Generators) *Let $P : [N] \to \{0, 1\}$ be a random function and $f : [N] \to [N]$ be a random permutation. For any pair of algorithms $\mathcal{A}_1, \mathcal{A}_2$ that $\mathcal{A}_1$ outputs an S-bit advice ($S \geq 1$) with oracle access to $\mathcal{O} := (f, P)$ and $\mathcal{A}_2$ makes T adaptive oracle queries to $\mathcal{O}$, it holds that*

$$\left|\mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, \mathcal{O}(x)) = 1\right] - \mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, u) = 1\right]\right| = \widetilde{O}\left(\frac{ST}{N} + \sqrt{\frac{S}{N}}\right) ,$$

where x is uniformly drawn from $[N]$, and u is uniformly drawn from $[N] \times \{0, 1\}$.

[2] The approximation gap of $b - a < k/n$ is really small. To see this, consider any k-wise independent distribution of $\mathbf{Y} = (Y_1, \ldots, Y_n)$ binary random variables with $\mathbf{E}[Y_i] = a$, where $a < 1/2$. Consider a modified random variable $\mathbf{Y}^+$, in which we independently set $Y_i^+ = 1$ with probability $2k/n$ and $Y_i^+ = Y_i$ otherwise for each $i \in [n]$. Both distributions $\mathbf{Y}^+ = (Y_1^+, \ldots, Y_1^+)$ and $\mathbf{Y}$ are k-wise independent, with $\mathbf{Pr}[\prod_{i \in I} Y_i = 1] = a^{|I|}$ and $\mathbf{Pr}[\prod_{i \in I} Y_i^+] > (a + k/n)^{|I|}$ for any subset I of size at most k. I.e., $\mathbf{Y}^+$ is above the threshold and $\mathbf{Y}$ is exactly at its lower side for any set of I variables in k-wise (a, b)-product condition, if $b - a < k/n$. On the other hand, we have changed in expectation at most $2k$ variables in $\mathbf{Y}$ to get $\mathbf{Y}^+$. In fact, when $k = o(n)$, the probability that we have changed more than $4k$ random variables is less than $\mathbf{Pr}[Poisson(2k) > 4k] = e^{-\Omega(k)}$.

Our results answer the open question posed by De et al. In particular, The above theorem shows that the advantage of order $\Omega(\sqrt{T/N})$ is not achievable for $S = O(1)$. Moreover, it strictly improves the result by De et al., and shows optimality of the best known attack for any $\mathcal{O}$ of the form (f, P), due to combining Hellman's permutation inversion algorithm [Hel80] and the attack by De et al. In particular, Hellman's permutation inversion algorithm inverts any permutation f on ST/N fraction of points. This algorithm yields a natural distinguisher with advantage ST/N (by inverting x and then comparing $P(x)$ with the last bit of the given y). When combined with the $\sqrt{S/N}$ attack by De et al., a non-uniform attacker achieves advantage of $\Omega(ST/N + \sqrt{S/N})$.

Our results trivially imply the same security bound for attacking a worst case $\mathcal{O}$, which is tight up to the additive term ST/N, for all range of parameters. For $ST^2 \leq N$, in which case $\sqrt{S/N}$ term dominates ST/N, our bound suggests that the attack from De et al. is the best possible. Strengthening ST/N term, which corresponds to the security bound for function inversion, is a long standing open problem. Corrigan-Gibbs and Kogan [CK19] showed that any improvement of this bound will give new lower bounds for depth two circuits with arbitrary gates, and strong improvements would imply breakthrough circuit lower bounds on linear-size log-depth circuits.

We prove our main result by establishing the same (and tight) non-uniform security bounds for hardcore predicates.

Theorem 1.2. (Hard-core Predicates) *Let $P : [N] \to \{0,1\}$ be a random function and $f : [N] \to [N]$ be a random permutation. For any oracle algorithms $\mathcal{A}_1, \mathcal{A}_2$, such that $\mathcal{A}_1$ outputs an S-bit advice ($S \geq 1$), and $\mathcal{A}_2$ makes T queries, it holds that*

$$\mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, f(x)) = P(x)\right] = \frac{1}{2} + \widetilde{O}\left(\frac{ST}{N} + \sqrt{\frac{S}{N}}\right) ,$$

where x is uniformly drawn from $[N]$, and $\mathcal{O} = (f, P)$.

Lastly, we present a new attack for Yao's box problem, matching the known $\Omega(\sqrt{ST/N})$ security bound due to De et al. [DTT10].

Theorem 1.3. (Yao's Box Problem) *Let $P : [N] \to \{0,1\}$ be a random function. There exist oracle algorithms $\mathcal{A}_1, \mathcal{A}_2$, that $\mathcal{A}_1$ outputs an S-bit advice, and $\mathcal{A}_2$ makes T queries without querying the given input x, such that*

$$\mathbf{Pr}\left[\mathcal{A}_2^{P}(\mathcal{A}_1^{P}, x) = P(x)\right] = \frac{1}{2} + \Omega\left(\sqrt{\frac{ST}{N}}\right) .$$

where x is uniformly drawn from $[N]$.

Yao's box problem has been used as an intermediate problem, which yields an upper bound on non-uniform bounds of hardcore predicates and pseudorandom generators [DTT10, DGK17]. Similar formulations of such problem are recently studied in the context of circuit complexity [ST18, MW19], which led to a new result on depth-3 circuits.

An attack with perfect advantage was known for the special case $S(T + 1) = N$ [Yao90]. Our new attack works for the general case, and is tight (up to constant factors) for all range of parameters. Combining Theorem 1.3 and Theorem 1.2, we obtain an interesting separation between the non-uniform security of Yao's box problem and hardcore predicates.

A general approach for proving non-uniform security. Our proof in Section 4.1 for establishing non-uniform security of hardcore predicates (see Theorem 1.2) is elementary given our concentration inequalities for the XOR condition (see Theorem 3.3). We find it conceptually simpler to abstract our proof and techniques in following modular way, which may be useful for proving non-uniform security for other cryptographic problems. We abstract our approach and provide a proof of Theorem 1.2 under this abstraction in Appendix 5.

Based on our concentration inequalities, we provide a general approach for proving non-uniform security. Informally, it reduces proving non-uniform security of cryptographic applications against S-bit advice to S-wise direct product theorems or S-wise XOR lemmas of a given problem.

Lemma 1.1. *(Informal) Let G be a problem with instance space $[N]$ (or a cryptographic application with challenge space $[N]$).*

1. *For a search problem G (or an unpredictability cryptographic application), if its k-wise direct product problem is ε^k-secure against attackers with zero-advice and Tk-query for any $k \leq S + \log N$, then G is ε'-secure against attacker with S-bit advice and T-query, where $\varepsilon' = 6\varepsilon + \widetilde{O}\left(\frac{S}{N}\right)$.*

2. *For a decision problem G (or an indistinguishability cryptographic application), if its k-wise XOR problem is ε^k-secure against attackers with zero-advice and Tk-query, for any $k \leq S + \log N$, then G is ε'-secure against attacker with S-bit advice and T-query, where $\varepsilon' = 2\varepsilon + \widetilde{O}\left(\sqrt{\frac{S}{N}}\right)$.*

The formal statement of Lemma 1.1 (together with definition of product and xor problem) is given in Section 5. An example of search problem is inverting one-way permutation. An example of decision problem is to distinguish the output of a pseudorandom generator from a random string.

A major difficulty in analyzing non-uniform security is that most standard techniques are developed for zero-advice attackers, and do not hold when preprocessing is allowed. The above lemma reduces the problem to proving the security against attackers with *no advice*, which is considered to be much simpler task, because standard techniques are applicable again.

A nice feature of Lemma 1.1 is that the blow up from ε to ε' is almost the best we can hope for. The multiplicative factor of ε is a small constant. The additive terms of S/N and $\sqrt{S/N}$ are necessary in such a general statement[3], and are often dominated by ε. The small blow up feature is appealing and crucial for bypassing barriers of previous techniques (incompressibility argument [DTT10, DGK17, CK18], and the presampling technique [Unr07, CDGS18, CDG18]).

We remark that the first item of Lemma 1.1 has been implicitly used in the prior work of Impagliazzo [Imp11] for proving tight non-uniform security of one-way permutations. We don't use it for our main result and only include it for illustration purpose. Very recently, its power has been recognized and used to prove tight bounds for several unpredictable cryptographic applications, including finding short collisions in Merkle-Damgård hash functions [ACDW20], and function inversion against affine non-adaptive decoders [CHM20]. In addition, it recently has been generalized to the quantum advice setting and yields tight quantum time-space tradeoffs for the function inversion problem [CGLQ20]. Presampling and incompressibility arguments are stuck in proving tight bounds for these problems, and are difficult to be generalized to the quantum advice setting.

We consider the second item as the main contribution of Lemma 1.1. Although product condition gives optimal bounds for search problems, attempts of using product condition only yield sub-optimal bounds for decision problems. Our main conceptual contribution is showing inherent limitations of product conditions and putting forward a strictly stronger condition (the XOR condition) for decision problems. Besides the PRG problem, we believe that this approach has a great potential to close the gap for other fundamental problems such as the decisional Diffie Hellman (DDH) [4], and to be generalized to the quantum advice setting using the machinery of [CGLQ20].

We remark that the reliance on XOR, as opposed to AND, is indeed the "natural" thing to do when capturing multi-instance security of indistinguishability based notions. This was argued in the work of Bellare et al. [BRT12], and was recently used in work by Auerbach et al. [AGK20]. Our work shows this in a technical sense.

Based on the second item of Lemma 1.1, Theorem 1.2 immediately follows from an XOR lemma for hardcore predicates (see Lemma 4.1), which says k-wise XOR problem of hardcore predicates is $(6kT/N)^k$-secure against attackers with zero advice and Tk queries.

2 Preliminaries

We use $[n]$ to denote the set $\{1,\ldots,n\}$. Let $\mathbf{X} = (X_1,\ldots,X_n)$ be n binary ($X_i \in \{0,1\}$) random variables. We consider the following notions of approximation for the k-wise independence of $\mathbf{X}$.

Definition 2.1. (One-sided product) $X_1,\ldots,X_n$ *satisfy* b-product condition *for* $b > 0$ *if,*

$$\forall I \subseteq [n],\ s.t.\ |I| \le k \qquad \mathbf{Pr}\left[\Pi_{i\in I} X_i = 1\right] \le b^{|I|}.$$

Definition 2.2. (Two-sided product) $X_1,\ldots,X_n$ *satisfy* (a,b)-product condition *for* $b > a > 0$ *if,*

$$\forall I \subseteq [n],\ s.t.\ |I| \le k \qquad a^{|I|} \le \mathbf{Pr}\left[\Pi_{i\in I} X_i = 1\right] \le b^{|I|}.$$

Definition 2.3. (XOR) $X_1,\ldots,X_n$ *satisfy* ε-XOR condition *for* $\varepsilon > 0$ *if,* $\forall I \subseteq [n],\ s.t.\ |I| \le k$

$$\mathrm{Bias}(\bigoplus_{i\in I} X_i) \overset{def}{=} \mathbf{Pr}\left[\bigoplus_{i\in I} X_i = 1\right] - \mathbf{Pr}\left[\bigoplus_{i\in I} X_i = 0\right] \le \varepsilon^{|I|}.$$

We also consider *average* versions of these conditions: if for a random set $I \subseteq [n]$ of a given size $|I| = t$ and for each $t \le k$

$$\mathbf{E}_{I,\mathbf{X}}\left[\Pi_{i\in I} X_i = 1\right] \le b^t, \quad \text{or}$$
$$a^t \le \mathbf{E}_{I,\mathbf{X}}\left[\Pi_{i\in I} X_i = 1\right] \le b^t, \quad \text{or}$$
$$\mathbf{E}_{I,\mathbf{X}}\left[\mathrm{Bias}(\oplus_{i\in I} X_i)\right] \le \varepsilon^t,$$

then we say that $\mathbf{X}$ satisfies *average*- b-product, or (a,b)-product, or ε-XOR condition[5], respectively.

[3] There are (many) examples of search problems (such as function inversion) for which $\varepsilon' = \tilde{\Omega}(S/N)$ by storing answers for S instances. There are examples of decision problems (such as distinguishing PRGs) for which $\varepsilon' = \Omega(\sqrt{S/N})$.

[4] The known upper bound on the advantage of non-uniform attackers is $\widetilde{O}\left(\sqrt{ST^2/N}\right)$ and the lower bound is $\tilde{\Omega}(ST^2/N)$ (see [CK18]).

[5] We remark that our notion of bias is defined as the "bias" of a $\{0,1\}$ variable towards 1 instead of 0, as considered in [Ung09, IK10]. These two notions are exchangable after switching 0 and 1. We choose this one because it better connects with the advantage of attackers against indistinguishability applications.

3 Concentration bounds for almost k-wise independence

In this section, we consider approximate notion of k-wise independence, and prove various concentration inequality for the sum of approximate k-wise independent variables. In Subsection 3.1, we give the weakest concentration bound under one-sided b-product condition (see Definition 2.1). We give a simple proof and show that this weak tail bound is tight. Next, we consider a stronger two-sided (a, b)-product condition (see Definition 2.2). The previous approach for one-sided product condition does not naturally extend. In Subsection 3.2, we show that *exact tail bounds* on sum of approximate k-wise independent variables are captured by primal and dual linear programs. The approach of analyzing LPs gives a unified way to obtain tight bounds for product conditions, and more generally any conditions that can be "symmetrized". In Subsection 3.3, we obtain tight bounds for two-sided product condition using this approach for a large range of parameters. Finally in Subsection 3.4, we derive our strongest concentration bound under the xor condition, which turns out to be strong enough for our application that we discuss in Section 4.

3.1 One-sided Product Condition

We obtain the following concentration inequality theorem for the one-sided product condition.

Theorem 3.1. *Let $X_1, \ldots, X_n$ be $\{0,1\}$ random variables such that for any set $I \subseteq [n]$ of size at most k, $\mathbf{Pr}[\Pi_{i\in I} X_i = 1] \le b^{|I|}$. Then for $c \ge 0$,*

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge c\cdot n\right] \le \min_{0\le i\le k}\left(\frac{b^i\binom{n}{i}}{\binom{cn}{i}}\right) . \tag{3.1}$$

Moreover, the bound (3.1) *holds for average b-product condition, and is tight.*

Theorem 3.1 captures the exact tail bound for all range of parameters. The upper bound is at most $(b/c)^k \approx \frac{b^k\binom{n}{k}}{\binom{cn}{k}}$ for a large range of parameters. It is interesting to check two regimes of parameters. First, when b is close to 0 and we are interested in multiplicative tail bound, i.e., $c = C \cdot b$ for a large constant C. Theorem 3.1 implies that the tail bound is $\exp(-\Omega(k))$ for any $C > e$. Second, when b is close to $1/2$, and we want to get dependency on the additive error in the tail bound, i.e., $c = b + \gamma$ for a small constant $\gamma > 0$. Theorem 3.1 implies that the tail bound is $\exp(-\Omega(\gamma k))$.

Proof. Our proof resembles in many ways the proof from Impagliazzo and Kabanets [IK10], that gave concentration bounds given product conditions for arbitrary subsets (not only those of size at most k). We denote $p_i \overset{\text{def}}{=} b^i \cdot \binom{n}{i}/\binom{cn}{i}$ and let t be the minimizer for $p_0, \ldots, p_k$. Let I be a random subset of size t. Let $\xi(\mathbf{X}) \overset{\text{def}}{=} \mathbf{E}_I[\Pi_{i\in I} X_i = 1]$. By the b-product condition,

$$\mathbf{E}_{\mathbf{X}}\left[\xi(\mathbf{X})\right] = \mathbf{E}_I\left[\mathbf{E}_{\mathbf{X}}\left[\Pi_{i\in I} X_i = 1\right]\right] \le b^t . \tag{3.2}$$

On the other hand, if we condition on the event that $\mathbf{X}$ has at least $c \cdot n$ coordinates equal to 1, then such an X has at least of $\binom{cn}{t}$ size-t subsets I such that $\Pi_{i\in I} X_i$ is 1. Therefore,

$$\mathbf{E}_{I,\mathbf{X}}\left[\Pi_{i\in I} X_i = 1 \;\middle|\; \sum_{i=1}^{n} X_i \ge cn\right] \ge \frac{\binom{cn}{t}}{\binom{n}{t}} .$$

Because $\Pi_{i\in I} X_i$ is a non-negative random variable, we have that

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge cn\right] \le \frac{\mathbf{E}[\Pi_{i\in I} X_i = 1]}{\mathbf{E}[\Pi_{i\in I} X_i = 1 | \sum_{i=1}^{n} X_i \ge cn]} \le \frac{b^t\binom{n}{t}}{\binom{cn}{t}} = p_t .$$

The desired upper bound follows. Because average-b-product condition also implies (3.2), the same proof extends to variables satisfying average-b-product condition.

To show p_t is attainable, consider distribution $\mathbf{X} = (X_1, \ldots, X_n)$ which outputs 0^n with probability $1 - p_t$, and outputs a uniformly random string from strings with exactly cn ones with probability p_t. Notice that p_t is exactly the probability that $\mathbf{X}$ has more than cn ones, i.e.,

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge cn\right] = p_t .$$

It remains to show that for any I of size $i \le k$, the product condition holds.

$$\mathbf{Pr}\left[\Pi_{i\in I} X_i = 1\right] = \frac{p_t}{\binom{n}{cn}} \cdot \binom{n-i}{cn-i} = p_t \cdot \frac{\binom{cn}{i}}{\binom{n}{i}} = p_t \cdot \frac{b^i}{p_i} \le b^i$$

where the second equality is because of binomial coefficient identity $\binom{n-i}{cn-i}\binom{n}{i} = \binom{n}{cn}\binom{cn}{i}$, the third equality is by definition of p_i, and the last inequality is because $p_t \le p_i$ for $0 \le i \le k$. □

Remark 3.1. *Theorem 3.1 can be extended to* $[0,1]$*-valued random variables $Y_1, \ldots, Y_n$ by reducing to* $\{0,1\}$ *variables $X_1, \ldots, X_n$ where $X_i = 1$ with probability* $\mathbf{E}[Y_i]$.

3.2 An LP Approach to Product Conditions

In the previous subsection, we characterized the tail bounds when only the upper bounds for the products of small subsets are given. A natural question is what can be done if we also have lower bounds on these products. Can we still characterize the exact tail bounds in this case? Note that in the case when the lower and upper bounds coincide, i.e., $a = b$, the distribution of $\mathbf{X}$ must be exactly k-wise independent, which implies strong concentration bounds [SSS93].

One difficulty in analyzing two sided product condition is that the proof of Theorem 3.1 does not extend naturally. In this subsection, we prove the following lemma, which proposes a unified approach for studying product conditions based on linear programs.

Definition 3.1. *For any set of parameter* $1 > c > b > a > 0$*, let*

$$p_{a,b,c} \overset{def}{=} \max_{\mathcal{D}} \Pr_{\mathbf{X}\sim\mathcal{D}}\left[\sum_{i=1}^{n} X_i \ge cn\right]$$
$$s.t.\ a^{|I|} \le \Pr\left[\Pi_{i\in I} X_i = 1\right] \le b^{|I|} \quad \forall I \subseteq [n], |I| \le k\ .$$

Lemma 3.1. *$p_{a,b,c}$ is the objective value of following two programs.*

$$\begin{aligned} \mathbf{obj} = \max \sum_{j\ge cn} \binom{n}{j} x_j & & & \textit{primal}\\ s.t.\ \sum_{j=0}^{n} -\binom{n-i}{j-i} x_j &\le -a^i & i\in[0,k] & \quad (\underline{\lambda}_i)\\ \sum_{j=0}^{n} \binom{n-i}{j-i} x_j &\le \ \ b^i & i\in[0,k] & \quad (\overline{\lambda}_i)\\ x_j &\ge \ \ 0 & \textit{for } j\in[0,n]. & \end{aligned}$$

$$\begin{aligned} \mathbf{obj} = \min \sum_{i=0}^{k} (\overline{\lambda}_i b^i - \underline{\lambda}_i a^i) & & & \textit{dual}\\ s.t.\ \sum_{i=0}^{k} \binom{n-i}{j-i} (\overline{\lambda}_i - \underline{\lambda}_i) &\ge \binom{n}{j} & & j \ge cn\\ \sum_{i=0}^{k} \binom{n-i}{j-i} (\overline{\lambda}_i - \underline{\lambda}_i) &\ge \ \ 0 & & j < cn\\ \overline{\lambda}_i \ge 0,\ \ \underline{\lambda}_i &\ge \ \ 0 & & i\in[0,k]. \end{aligned}$$

By convention we assume $\binom{n}{i} = 0$ for $i < 0$. The objective values of the LPs in Lemma 3.1 capture the best bound achievable on the tail probability. By proposing feasible solutions to the primal and dual LPs, we respectively obtain lower and upper bounds on the tail probabilities.

Proof. We first show that we can focus on symmetric distributions for $\mathbf{X}$. Given a random variable $(X_1, \dots, X_n) = \mathbf{X} \sim \mathcal{D}$ we construct a symmetric random variable $\mathbf{Y}$: $\mathbf{Y} = (Y_1, \dots, Y_n) \overset{\text{def}}{=} (X_{\sigma(1)}, \dots, X_{\sigma(n)})$ where we choose uniformly at random a permutation $\sigma \sim [n!]$ and draw $\mathbf{X} \sim \mathcal{D}$. Let $p_S \overset{\text{def}}{=} \Pr[X_i = 1\ \forall i \in S,\ X_i = 0\ \forall i \notin S]$ for any $S \subseteq [n]$. Then for any $S \subseteq [n]$ we have

$$\begin{aligned} \Pr\left[Y_i = 1\ \forall i \in S,\ Y_i = 0\ \forall i \notin S\right] &= \frac{1}{n!} \sum_{\sigma\in[n!]} p_{\sigma(S)}\\ &= \frac{|S|! \cdot (n-|S|)!}{n!} \sum_{\substack{T\subseteq[n],\\ |T|=|S|}} p_{\sigma(T)} \end{aligned}$$

is the same for all sets S of the same size. Thus $\mathbf{Y}$ has a symmetric distribution. Moreover, for any $S \subseteq [n]$ such that $|S| \le k$, we have

$$\Pr\left[\prod_{i\in S} Y_i = 1\right] = \mathop{\mathbf{E}}_{\sigma}\left[\Pr\left[\prod_{i\in S} X_{\sigma(i)} = 1\right]\right] \in [a^{|S|}, b^{|S|}],$$

because the probability term inside the expectation is in the interval $[a^{|S|}, b^{|S|}]$ by (a,b)-product condition.

Hence, to prove the tail bound for almost k-wise independent random variables, it suffices to only consider symmetrized instances. From now on, we shall assume that distribution of $\mathbf{X}$ is symmetric. To characterize a symmetric distribution, it suffices to specify

$$x_j = p_S = \Pr\left[X_i = 1\ \forall i \in S, X_i = 0\ \forall i \notin S\right],$$

for any $|S| = j$ and $j = 0, 1, \dots, n$.

The (a,b)-product criteria can be written as a set of linear constraints in $(x_j)_{j=0}^n$. Specifically, if we fix set S of size $\ell \le k$ and write $\Pr[\prod_{i\in S} X_i = 1]$ as a combination of x_j, we get

$$a^\ell \le \sum_{j=0}^{n} \binom{n-\ell}{j-\ell} x_j \le b^\ell.$$

Hence we obtain the first linear program for the tail bound. By taking the dual, we obtain the second linear program. By LP duality theorem, both linear programs have the same objective value. Note that these linear programs give the exact tail bound. □

Observe that the case $a = 0$ corresponds to one-sided product condition, Lemma 3.1 provides an alternative approach for proving Theorem 3.1. Most importantly for us, Lemma 3.1 allows to reason about tightness of the concentration bounds for the (a,b)-product condition.

3.3 Two-sided Product Condition

In this subsection, we provide a tight analysis for the product condition when $n \gg k$.

Theorem 3.2. *We assume that $\mathbf{X}$ are n binary random variables that satisfy (a,b)-product approximate k-wise independent condition, i.e., $a^{|I|} \le \mathbf{Pr}[\Pi_{i\in I} X_i = 1] \le b^{|I|}\ \ \forall I \subseteq [n]$, and $|I| \le k$. Suppose there is a small gap between a and b, $a + k/n \le b \le c$. In the case $ac^{k-1} \ge b^k$, there exists a unique solution $p \in [0,1]$ and $s \in [0,1]$ to the following system of equations:*

$$\begin{cases} p\cdot c + (1-p)\cdot s = a \\ p\cdot c^k + (1-p)\cdot s^k = b^k. \end{cases} \tag{3.3}$$

Then we have respectively the following tail bound, and complementary tightness example

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge cn\right] \le p + \frac{k^2/n}{(k-1)a^k - ka^{k-1}c + c^k}.$$

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge cn\right] \ge p - \frac{(k/n)\cdot(2(a+k/n))}{(c-a-k/n)^2(k-1)}.$$

If $ac^{k-1} < b^k$ (i.e., the gap between c and b is much smaller than between a and b), then we have respectively the following tail bound, and complementary tightness example

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge cn\right] \le \left(\frac{b}{c-k/n}\right)^k = \frac{b^k}{c^k}\left(1-\frac{k}{cn}\right)^{-k}.$$

$$\mathbf{Pr}\left[\sum_{i=1}^{n} X_i \ge cn\right] \ge \frac{b^k}{c^k}\left(1-\left(\frac{k}{n(b-a)}\right)^k\right).$$

Proof. We prove lower bound by a feasible solution for the primal LP and upper bound by a feasible solution for the dual LP. We choose a primal feasible solution x_j with only two non-zeros variables, one at $j = cn$ to maximize the objective, and the other at $j = sn$ where s is chosen to satisfy the constraints. As n approaches infinity, the binomial coefficients in the constraints of the primal LP can be well approximated by powers, and then the power mean inequality suggests only three constraints (when $i = 0$, $i = 1$ and $i = k$) are relevant. These constraints are approximated by the system of equations. The dual solution we choose contains only three non-zeros, corresponding to the three relevant constraints.

The following approximation will be useful to simplify coefficients in the constraints.

Claim 1. *Let $1 \ge \alpha \ge k/n$, then for any $0 \le i \le k$, we have*

$$\alpha^i \ge \frac{\binom{\alpha n}{i}}{\binom{n}{i}} \ge \left(\alpha - \frac{k}{n}\right)^i.$$

Proof. Note that

$$\frac{\binom{\alpha n}{i}}{\binom{n}{i}} = \frac{\alpha n\cdot(\alpha n-1)\cdot\ldots\cdot(\alpha n-i+1)}{n\cdot(n-1)\cdot\ldots\cdot(n-i+1)} = \prod_{j=0}^{i-1}\frac{\alpha n - j}{n-j}.$$

Now for any $j \le k \le \alpha n$, we have

$$\alpha = \frac{\alpha n - \alpha j}{n-j} \ge \frac{\alpha n - j}{n-j} \ge \frac{\alpha n - j}{n} = \alpha - \frac{j}{n} \ge \alpha - \frac{k}{n}.$$

Hence for any $i \le k \le \alpha n$,

$$\alpha^i \ge \prod_{j=0}^{i-1}\frac{\alpha n - j}{n-j} \ge \left(\alpha - \frac{k}{n}\right)^i.$$

□

Lower bound. We first consider the case $ac^{k-1} \ge b^{k-1}$. Let $f_i(p,s) := pc^i + (1-p)s^i$, and so the system of equations is $f_1(p,s) = a$, $f_k(p,s) = b^k$. Note that if we keep the invariant $f_k(p,s) = b^k$, and move p from 0 to $(b/c)^k$, then s strictly decreases from b to 0, and $f_1(p,s)$ strictly decreases from $b \ge a$ to $b^k/c^{k-1} \le a$. Since the move is smooth, we conclude there is a unique solution (p,s) such that both $f_1(p,s) = a$ and $f_k(p,s) = b^k$ are satisfied. By similar argument, when $a + k/n \le b$, there is another solution (p',s') such that $f_1(p',s') = a + k/n$ and $f_k(p',s') = b^k$. We shall first construct a primal feasible solution with (p',s'), and then analyze the difference between p and p'.

There are two cases. Suppose s' is not too small, $s' \ge k/n$. Our construction is $x_{cn} = p'/\binom{n}{cn}$, $x_{s'n} = (1-p')/\binom{n}{s'n}$, and $x_j = 0$ for any other j. Then the objective value is p', and the i-th constraint becomes

$$a^i \le \binom{n-i}{cn-i}x_{cn} + \binom{n-i}{sn-i}x_{s'n} = p'\frac{\binom{n-i}{cn-i}}{\binom{n}{cn}} + (1-p')\frac{\binom{n-i}{s'n-i}}{\binom{n}{s'n}} = p'\frac{\binom{cn}{i}}{\binom{n}{i}} + (1-p')\frac{\binom{s'n}{i}}{\binom{n}{i}} \le b^i.$$

Note that by the power mean inequality, we have for any $s \ge 0$ and $p \in [0,1]$,

$$f_1(p,s) \le f_i(p,s)^{1/i} \le f_k(p,s)^{1/k}.$$

Hence by Claim 1, we have

$$p'\frac{\binom{cn}{i}}{\binom{n}{i}} + (1-p')\frac{\binom{s'n}{i}}{\binom{n}{i}} \le p'c^i + (1-p')s'^i \le (p'c^k + (1-p')s'^k)^{i/k} = b^i,$$

and

$$p'\frac{\binom{cn}{i}}{\binom{n}{i}}+(1-p')\frac{\binom{s'n}{i}}{\binom{n}{i}} \ge p'\left(c-\frac{k}{n}\right)^i+(1-p')\left(s'-\frac{k}{n}\right)^i$$
$$\ge \left(p'(c-\frac{k}{n})+(1-p')\left(s'-\frac{k}{n}\right)\right)^i = a^i.$$

If $s' < k/n$, then our construction is $x_{cn} = p'/\binom{n}{cn}$, $x_0 = 1-p'$, and $x_j = 0$ for any other j. Then the objective value is p', and the 0-th constraint is satisfied, and the i-th constraint ($1 \le i \le k$) becomes

$$a^i \le p'\frac{\binom{cn}{i}}{\binom{n}{i}} \le b^i.$$

By Claim 1, we have

$$p'\frac{\binom{cn}{i}}{\binom{n}{i}} \le p'c^i \le p'c^i + (1-p')s'^i \le b^i,$$

and

$$p'\frac{\binom{cn}{i}}{\binom{n}{i}} \ge p'\left(c-\frac{k}{n}\right)^i \ge \left(p'\left(c-\frac{k}{n}\right)\right)^i$$
$$\ge \left(p'\left(c-\frac{k}{n}\right)+(1-p')s'-(1-p')\frac{k}{n}\right)^i$$
$$= \left(p'c+(1-p')s'-\frac{k}{n}\right)^i \ge a^i.$$

Hence in both cases, p' is a lower bound to the primal optimum.

Now we analyze the change of p when a is changed to $a+k/n$. We keep the invariant $f_k(p,s) = b$, and hence

$$0 = \frac{d}{dp}f_k(p,s) = c^k - s^k + (1-p)ks^{k-1}\frac{ds}{dp},$$

or

$$\frac{ds}{dp} = -\frac{c^k-s^k}{(1-p)ks^{k-1}}.$$

This implies

$$-\frac{d}{dp}f_1(p,s) = -(c-s)-(1-p)\frac{ds}{dp} = \frac{c^k-s^k}{ks^{k-1}}-(c-s)$$
$$= (c-s)\left(\frac{c^{k-1}+c^{k-2}s+\ldots+s^{k-1}}{ks^{k-1}}-1\right)$$
$$= (c-s)\left(\frac{(c^{k-1}-s^{k-1})+(c^{k-2}-s^{k-2})s+\ldots}{ks^{k-1}}\right)$$
$$\ge (c-s)^2\left(\frac{(k-1)s^{k-2}+(k-2)s^{k-2}+\ldots}{ks^{k-1}}\right)$$
$$= (c-s)^2\frac{k-1}{2s} \ge \frac{(c-a-k/n)^2(k-1)}{2(a+k/n)},$$

where the last inequality holds since s is always upper bounded by a. Therefore the change of p is at most

$$p-p' \le \frac{k}{n}/(-\frac{d}{dp}f_1(p,s)) \le \frac{(k/n)\cdot(2(a+k/n))}{(c-a-k/n)^2(k-1)}.$$

Hence the primal feasible solution we constructed has objective

$$p' \ge p - \frac{(k/n)\cdot(2(a+k/n))}{(c-a-k/n)^2(k-1)}.$$

For the second case $ac^{k-1} < b^k$, there is no solution to the system of equations. Our construction is $s = kb/(n(b-a))$, $p = (b^k-s^k)/(c^k-s^k)$, $x_{cn} = p/\binom{n}{cn}$, $x_{sn} = (1-p)/\binom{n}{sn}$, and $x_j = 0$ for any other j. To show that it is a primal feasible solution, it suffices to show that $f_1(p,s) \ge a+k/n$ and $f_k(p,s) \le b^k$. Our choice of p implies $f_k(p,s) = b^k$. On the other hand,

$$f_1(p,s) = s+(c-s)\frac{b^k-s^k}{c^k-s^k} = s+(b-s)\frac{b^{k-1}+\ldots+s^{k-1}}{c^{k-1}+\ldots+s^{k-1}}$$
$$\ge s+(b-s)\left(\frac{b}{c}\right)^{k-1} \ge s+(b-s)\frac{a}{b} = a+s\left(1-\frac{a}{b}\right) = a+\frac{k}{n}.$$

Hence we constructed a primal feasible solution with objective

$$p = \frac{b^k-s^k}{c^k-s^k} \ge \frac{b^k-s^k}{c^k} = \frac{b^k}{c^k}\left(1-\left(\frac{k}{n(b-a)}\right)^k\right).$$

Upper bound. We first consider the case $ac^{k-1} \ge b^{k-1}$. In such case we have solution (p,s) to the system of equations. Our feasible solution for the dual LP will only have non-zero values at $\overline{\lambda}_0$, $\underline{\lambda}_1$, and $\overline{\lambda}_k$. Let $f(x) = \gamma_0+\gamma_1 x+\gamma_k x^k$ satisfies $f(c) = 1$, $f(s) = f'(s) = 0$. Then we can check that

$$f(x) = \frac{(k-1)s^k-ks^{k-1}x+x^k}{(k-1)s^k-ks^{k-1}c+c^k}.$$

Since $f'(x) < 0$ when $0 < x < s$ and $f'(x) > 0$ when $x > s$, we have $f(x) \ge 0$ for all $x \in [0,c]$ and $f(x) \ge 1$ for all $x \in [c,1]$.

Let $\overline{\lambda}_0 = \gamma_0+(k^2/n)\gamma_k$, $\underline{\lambda}_1 = -\gamma_1$, and $\overline{\lambda}_k = \gamma_k$. Then the j-th constraint of the dual LP is

$$\gamma_0+\frac{k^2}{n}\gamma_k+\gamma_1\frac{j}{n}+\gamma_k\frac{j(j-1)\ldots(j-k+1)}{n(n-1)\ldots(n-k+1)} \ge 1_{j\ge cn}.$$

Since $f(x) \ge 1_{x\ge c}$, we have for any $j \ge k$,

$$\gamma_0 + \frac{k^2}{n}\gamma_k + \gamma_1 \frac{j}{n} + \gamma_k \frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)}$$
$$= f\left(\frac{j}{n}\right) + \frac{k^2}{n}\gamma_k + \gamma_k\left(\frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)} - \left(\frac{j}{n}\right)^k\right) \ge 1_{j\ge cn} + \gamma_k\left(\frac{k^2}{n} + \left(\frac{j}{n}-\frac{k}{n}\right)^k - \left(\frac{j}{n}\right)^k\right)$$
$$\ge 1_{j\ge cn} + \gamma_k\left(\frac{k^2}{n} - \frac{k}{n}\cdot k\cdot\left(\frac{j}{n}\right)^{k-1}\right) \ge 1_{j\ge cn}$$

For $j < k$,

$$\gamma_0 + \frac{k^2}{n}\gamma_k + \gamma_1\frac{j}{n} + \gamma_k\frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)} = f(\frac{j}{n}) + \gamma_k\left(\frac{k^2}{n} - (\frac{j}{n})^k\right) \ge f(\frac{j}{n}) + \gamma_k\left(\frac{k^2}{n} - (\frac{k}{n})^k\right) \ge f(\frac{j}{n}).$$

Hence we constructed a dual feasible solution with objective

$$\frac{(k-1)s^k - ks^{k-1}a + b^k}{(k-1)s^k - ks^{k-1}c + c^k} + \frac{k^2}{n}\gamma_k.$$

The first term is exactly p. This is because $p(c^k - s^k) = b^k - s^k$ and $p(c-s) = a - s$, and hence

$$p((k-1)s^k - ks^{k-1}c + c^k) = p(c^k - s^k) - ks^{k-1}p(c-s)$$
$$= b^k - s^k - ks^{k-1}(a-s).$$

Finally since $(k-1)s^k - ks^{k-1}c + c^k$ decreases with s as long as $s \le c$ by considering its derivative, we can upper bound γ_k by $((k-1)a^k - ka^{k-1}c + c^k)^{-1}$. Hence the error term is at most

$$\frac{k^2}{n((k-1)a^k - ka^{k-1}c + c^k)}.$$

For the second case $ac^{k-1} < b^k$, we shall just let $\overline{\lambda}_k = (c - k/n)^{-k}$ and all other dual variables to be 0. Then the j-th constraint of the dual LP is

$$\overline{\lambda}_k \frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)} \ge 1_{j\ge cn}.$$

Clearly the constraints are satisfied when $j < cn$, since $\overline{\lambda}_k$ is non-negative. When $j \ge cn$,

$$\overline{\lambda}_k \frac{j(j-1)\dots(j-k+1)}{n(n-1)\dots(n-k+1)} \ge \left(c - \frac{k}{n}\right)^{-k}\left(\frac{j}{n} - \frac{k}{n}\right)^k$$
$$\ge \left(c - \frac{k}{n}\right)^{-k}\left(\frac{cn}{n} - \frac{k}{n}\right)^k = 1.$$

Hence this is a feasible solution, with objective value $\overline{\lambda}_k b^k = b^k(c-k/n)^{-k}$. □

A regime of particular interest is when a, b, and c are close to 1/2: $a = (1-\varepsilon)/2$, $b = (1+\varepsilon)/2$, $c = (1+\gamma)/2$, where $2\varepsilon < \gamma = o(1)$. In this regime, solution to the system of equations (3.3) has $p \approx 2\varepsilon/(2\varepsilon + k\gamma^2)$ (see Appendix A). To compare our bounds in Theorem 3.2 to CH bound for k-wise independent random variables [SSS93], we should consider $\gamma = \Omega(\sqrt{k/n})$. Then, even for a very small approximation error $b - a = 2k/n$, the tightness example in the case $ac^{k-1} \ge b^k$ makes $p \approx 2\varepsilon/(2\varepsilon + k\gamma^2) \approx 2/k$ which yields a lower bound of $\Omega(1/k)$, a much worse guarantee than $e^{-\Omega(k)}$ for the exact k-wise independence.

3.4 XOR Condition Here we derive our strongest tail bound using k-wise XOR condition on $\{0,1\}$-random variables $\mathbf{X}$.

Theorem 3.3. *Let $\gamma > 0$. Suppose for any set $I \subseteq [n]$ of size at most k, $\mathrm{Bias}(\oplus_{i\in I} X_i) \le \varepsilon^{|I|}$, then*

$$\mathbf{Pr}\left[\sum_{i=1}^n X_i \ge \frac{n(1+\gamma)}{2}\right] \le \left(\frac{\varepsilon + \sqrt{\frac{k}{n}}}{\gamma}\right)^k.$$

The same conclusion holds even for the average-ε-XOR condition.

This result is most useful when expectations of each variable $\mathbf{E}[X_i]$ is around 1/2. With a more careful analysis, it is possible to improve the parameters slightly and extend Theorem 3.3 to bounded range variables[6] (see Theorem B.1 in Appendix B).

The concentration bound in Theorem 3.3 bypasses the limitation of (a,b)-product conditions when $a = (1-\varepsilon)/2$ and $b = (1+\varepsilon)/2$ are around 1/2, $c = (1+\gamma)/2$, and $k \ll n$. For $\gamma \ge 2(\varepsilon + \sqrt{k/n})$, product conditions yield $\exp(-\Omega(\gamma k))$, while XOR condition yields $\exp(-\Omega(k))$.

Proof. Let $Y_i = 2X_i - 1$, so that $Y_i \in \{-1, 1\}$. Consider

[6] For the clarity of exposition and because we need this specific form in our cryptographic application, we decided to state the theorem in this form. Theorem 3.3 considers one-sided XOR condition, i.e., the condition is satisfied when $\mathbf{Pr}[\oplus_{i\in I} X_i = 0]$ is close to 1. One can similarly define two-sided XOR condition, and define corresponding LP approach to analyze the exact tail bounds. However, the tail bound for the one-sided XOR condition is already nearly optimal compared to k-wise independence case. It also suffices for our main application. Thus we leave the question of improving concentration bounds for two-sided XOR condition for future work.

the k-th moment method.

$$\mathbf{Pr}\left[\sum_{i=1}^{n} Y_i \geq \gamma n\right] \leq \mathbf{Pr}\left[\left(\sum_{i=1}^{n} Y_i\right)^k \geq (\gamma n)^k\right]$$
$$\leq \mathbf{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^k / (\gamma n)^k\right].$$

We shall expand $(\sum_i Y_i)^k$ and regroup the terms. Each term is in the form $\prod_{l=1}^{k} Y_{i_l} = \prod_{i=1}^{n} Y_i^{n_i}$, where n_i is the number of times Y_i appears in the product. Let j be the total number of n_is that are odd. Then $\mathbf{E}[\prod_{l=1}^{k} Y_{i_l}] \leq \varepsilon^j$ by replacing every Y_i with 1 except j of them (note that $Y_i^2 = 1$).

Now we count the number of terms that have exactly j different variables Y_i with an odd degree n_i. We can uniquely describe such a term by the following procedure. Let $S \subseteq [k]$ be the set of indices ℓ in the product $\prod_{\ell=1}^{k} Y_{i_\ell}$, such that $\ell \in S$ if and only if the degree n_{i_ℓ} of Y_{i_ℓ} is odd and ℓ is the smallest index for i_ℓ (if there is more than one Y_{i_ℓ} in the product). Thus $|S| = j$. Then we choose different $i_\ell \in [n]$ for each of the $\ell \in S$. So far we have at most $\binom{k}{j}$ possibilities for choosing $S \subseteq [k]$ and at most n^j possibilities for choosing each of the i_ℓ. The remaining indices in $[k] \setminus S$ can be grouped into pairs. We have at most n choices for i_ℓ and at most $k-j-1$ choices for the position of its pair; the second unmatched index and its pair have at most $n \cdot (k-j-3)$ possibilities, and so on. Therefore, there are at most $n^{(k-j)/2} \cdot (k-j-1)(k-j-3) \cdot \ldots \cdot 1 \leq n^{(k-j)/2} k^{(k-j)/2} = (\sqrt{nk})^{k-j}$ possibilities for the indices outside of S. Hence we get

$$\mathbf{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^k\right] \leq \sum_{j=0}^{k} \varepsilon^j \binom{k}{j} n^j (\sqrt{nk})^{k-j}$$
$$= \sum_{j=0}^{k} \binom{k}{j} (\varepsilon n)^j (\sqrt{nk})^{k-j} = \left(\varepsilon n + \sqrt{nk}\right)^k.$$

This completes the proof. The same proof holds for average-ε-XOR condition, because the coefficient is the same for all terms that have exactly j different variables with odd degree, so that average-ε-XOR condition gives the same upper bound on their grouping sum as ε-XOR condition does. □

4 Non-uniform security of pseudorandom generators and related problems

In this section, we prove Theorem 1.1 (Pseudorandom Generators), Theorem 1.2 (Hard-core Predicates) and Theorem 1.3 (Yao's Box Problem).

By Yao's reduction of distinguishers to predictors [Yao82], Theorem 1.2 immediately gives Theorem 1.1. So we omit the proof of Theorem 1.1 and prove Theorem 1.2 and Theorem 1.3. We recall the statement of Theorem 1.2.

THEOREM 4.1. *Let $P : [N] \to \{0,1\}$ be a random function and $f : [N] \to [N]$ be a random permutation. For any oracle algorithms $\mathcal{A}_1, \mathcal{A}_2$, such that $\mathcal{A}_1$ outputs an S-bit advice ($S \geq 1$), and $\mathcal{A}_2$ makes T queries, it holds that*

$$\mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, f(x)) = P(x)\right] = \frac{1}{2} + \widetilde{O}\left(\frac{ST}{N} + \sqrt{\frac{S}{N}}\right).$$

where x is uniformly drawn from $[N]$, and $\mathcal{O} = (f, P)$.

4.1 A tight security bound for hardcore predicates: Proof of Theorem 1.2

Let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be an attacker for the hardcore predicate problem. Without loss of generality, we assume that $\mathcal{A}$ is deterministic (by fixing the best choice of random coins). We will show that for any fixed advice string $w \in \{0,1\}^S$, it holds that

$$(4.4) \quad \Pr_{\mathcal{O}}\left[\mathbf{Pr}_x\left[\mathcal{A}_2^{\mathcal{O}}(w, f(x)) = P(x)\right] \geq \frac{1}{2} + \varepsilon + \sqrt{\frac{S + \log(1/\gamma)}{N}}\right] \leq 2^{-S} \cdot \gamma.$$

where $\gamma := 1/N$, $\varepsilon := 6(S + \log(1/\gamma))T/N$. By the union bound over 2^S advice strings,

$$\Pr_{\mathcal{O}}\left[\mathbf{Pr}_x\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, f(x)) = P(x)\right] \geq \frac{1}{2} + \varepsilon + \sqrt{\frac{S + \log(1/\gamma)}{N}}\right] \leq \gamma.$$

Then by an averaging argument over $\mathcal{O}$, we obtain

$$\mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, f(x)) = P(x)\right] \leq \frac{1}{2} + \varepsilon + \sqrt{\frac{S + \log(1/\gamma)}{N}} + \gamma = \frac{1}{2} + \widetilde{O}\left(\frac{ST}{N} + \sqrt{\frac{S}{N}}\right),$$

which gives Theorem 1.2.

Fix an arbitrary $w \in \{0,1\}^S$, and consider boolean variables $X_1, \ldots, X_N$, where X_y is the indicator for the event $\mathcal{A}_2^{\mathcal{O}}(w, y) = P(f^{-1}(y))$. For any permutation f, because $(f^{-1}(y), y)$ distributes the same as $(x, f(x))$ for uniformly random x, y from $[N]$, (4.4) can be rewritten as

$$\mathbf{Pr}\left[\frac{1}{N}\sum_{y\in[N]} X_y \geq \frac{1}{2}+\varepsilon+\sqrt{\frac{S+\log(1/\gamma)}{N}}\right] \leq 2^{-S}\cdot\gamma\ .$$

By Theorem 3.3, it suffices to prove that for any $I \subseteq [N]$ of size at most $S+\log(1/\gamma)$, $\mathrm{Bias}(\oplus_{i\in I} X_i) \leq \varepsilon^{|I|}$. This is implied by the following lemma.

Lemma 4.1. *For any* $I \subseteq [N]$, $\mathrm{Bias}(\oplus_{i\in I} X_i) \leq (6|I|T/N)^{|I|}$.

Proof. For $|I| \geq N/6T$, the statement holds trivially because $\mathrm{Bias}(\oplus_{i\in I} X_i) \leq 1$ and $6|I|T/N \geq 1$. Fix a set I of size k where $k \leq N/6T$. For any $y \in [N]$, observe that,

$$X_y = 1_{\mathcal{A}_2^{\mathcal{O}}(w,y)=P(f^{-1}(y))} = P(f^{-1}(y)) \oplus (\mathcal{A}_2^{\mathcal{O}}(w,y)) \oplus 1.$$

By the property of XOR function,

$$\mathop{\oplus}_{y\in I} X_y = \left(\mathop{\oplus}_{y\in I} P(f^{-1}(y))\right) \oplus \left(\mathop{\oplus}_{y\in I} \mathcal{A}_2^{\mathcal{O}}(w,y)\right) \oplus \left(\mathop{\oplus}_{y\in I} 1\right)$$

Hence to obtain the claimed upper bound on $\mathrm{Bias}(\oplus_{y\in I} X_y)$, it suffices to prove that

$$2\Pr[\mathop{\oplus}_{y\in I} \mathcal{A}_2^{\mathcal{O}}(w,y) = \left(\mathop{\oplus}_{y\in I} P(f^{-1}(y))\right) \oplus b'] \leq 1 + \left(\frac{6kT}{N}\right)^k \tag{4.5}$$

where $b' \in \{0,1\}$ is fixed bit only depending on k, and the randomness is only over $\mathcal{O}$ (recall that I is fixed and $\mathcal{A}_2$ is deterministic).

Let Γ be the random variable corresponding to the transcript containing query/answer pairs $(x_1, \mathcal{O}(x_1), \ldots, x_q, \mathcal{O}(x_q))$ resulting from simulating $\mathcal{A}_2$ with advice w for every $y \in I$ where $q := kT$. Let E denote the event that every preimage of $y \in I$ appears as a query in Γ, in other words, $f^{-1}(y) = x_i$ for some $i \in [q]$.

Claim 2. $2\,\mathbf{Pr}[\mathop{\oplus}_{y\in I} \mathcal{A}_2^{\mathcal{O}}(w,y) = (\mathop{\oplus}_{y\in I} P(f^{-1}(y))) \oplus b'] \leq 1 + \mathbf{Pr}[E]$.

Proof. Let τ be an arbitrary possible transcript. Conditioned on $\Gamma = \tau$, because $\mathcal{A}_2$ is deterministic, $\oplus_{y\in I}\mathcal{A}_2^{\mathcal{O}}(w,y)$ is a fixed bit. Let $I_\tau \subseteq I$ denote the set of elements in I whose preimage appears as a query in τ. Then

$$\mathop{\oplus}_{y\in I} P(f^{-1}(y)) = \left(\mathop{\oplus}_{y\in I_\tau} P(f^{-1}(y))\right) \oplus \left(\mathop{\oplus}_{y\in I\setminus I_\tau} P(f^{-1}(y))\right)$$

is fixed if $I_\tau = I$, otherwise it is a random bit because P is a random function on $f^{-1}(I \setminus I_\tau)$. Therefore,

$$\begin{aligned}\mathbf{Pr}\left[\mathcal{Z}=1\right] &= \mathbf{Pr}\left[\mathcal{Z}=1|I_\Gamma = I\right]\cdot\mathbf{Pr}\left[I_\Gamma = I\right] \\ &\quad + \mathbf{Pr}\left[\mathcal{Z}=1|I_\Gamma \neq I\right]\cdot\mathbf{Pr}\left[I_\Gamma \neq I\right] \\ &\leq 1\cdot\mathbf{Pr}\left[I_\Gamma = I\right] + \frac{1}{2}\cdot\mathbf{Pr}\left[I_\Gamma \neq I\right] = \frac{1+\mathbf{Pr}[I_\Gamma = I]}{2}\ .\end{aligned}$$

Observe that the event $I_\Gamma = I$ is exactly E. The desired conclusion follows. □

To prove (4.5) and complete the proof, it remains to show that $\mathbf{Pr}[E] \leq (6Tk/N)^k$. Let E_i denote the event that the i-th distinct query gets mapped to some y in I under f. Conditioning on all previous query/answer pairs, then each E_i has conditional probability at most $k/(N-q) \leq 2k/N$ (recall $k \leq N/6T$). There are at most q such events. If E happens, there must be at least k of these events that are true. For each of the $\binom{q}{k}$ sets of these events, the probability that all of these events occur simultaneously is at most $(2k/N)^k$. So E happens with probability at most $\binom{q}{k}(2k/N)^{|I|} \leq (2eq/N)^k \leq (6Tk/N)^k$. □

4.2 A tight attack for Yao's box problem: Proof of Theorem 1.3

We recall the statement of Theorem 1.3.

Theorem 4.2. *Let* $P : [N] \to \{0,1\}$ *be a random function. There exist oracle algorithms* $\mathcal{A}_1, \mathcal{A}_2$*, that* $\mathcal{A}_1$ *outputs an* S*-bit advice, and* $\mathcal{A}_2$ *makes* T *queries without querying the given input* x*, such that*

$$\mathbf{Pr}\left[\mathcal{A}_2^P(\mathcal{A}_1^P, x) = P(x)\right] = \frac{1}{2} + \Omega\left(\sqrt{\frac{ST}{N}}\right)\ .$$

where x *is uniformly drawn from* $[N]$.

We present the proof for the case when N is a multiple of ST and $\ell := N/S$ is an odd number. The general case can be analyzed in a similar way by allowing some sets resulted from the partition to have smaller size than others. Let $X_1, \ldots, X_S$ denote the natural partition of $[N]$ into consecutive blocks of size ℓ: $X_i := \{(i-1)\ell + 1, \ldots, i\ell\}$. Similarly, let $X_{i,1}, \ldots, X_{i,\ell/T}$ denote the natural partition of X_i such that each set has size T.

The offline algorithm $\mathcal{A}_1^P$ computes $z_{i,j} = \oplus_{x\in X_{i,j}} P(x)$ for every i, j, then outputs an S-bit state where the i-th bit is equal to $\mathrm{maj}(z_{i,1}, \ldots, z_{i,\ell/T})$ where maj is the majority function. The online algorithm $\mathcal{A}_2^P$, on given S-bit advice $b_1, \ldots, b_S$ and $x \in [N]$, computes $z'_{i,j} := \oplus_{y\in X_{i,j}\setminus\{x\}} P(y)$ (by querying P on all points in $X_{i,j}$ except x), and outputs $z'_{i,j} \oplus b_i$ as the prediction for $P(x)$, where i, j is the unique pair such that $x \in X_{i,j}$.

Fix an arbitrary $x \in [N]$ together with i, j such that $x \in X_{i,j}$. Since $z'_{i,j} \oplus b_i = z_{i,j} \oplus b_i \oplus P(x)$, it holds that $z'_{i,j} \oplus b_i = P(x)$ if and only if $b_i = z_{i,j}$. Hence, for any fixed $x \in [N]$,

$$\begin{aligned}\mathbf{Pr}\left[\mathcal{A}_2^P(\mathcal{A}_1^P, x) = P(x)\right] &= \mathbf{Pr}\left[b_i = z_{i,j}\right] \\ &= \mathbf{Pr}\left[\mathrm{maj}(z_{i,1}, \ldots, z_{i,\ell/T}) = z_{i,j}\right].\end{aligned}$$

Because P is a random function, and $X_{i,1}, \ldots, X_{i,\ell/T}$ are disjoint, $z_{i,1}, \ldots, z_{i,\ell/T}$ are random and mutually independent.

$$\begin{aligned}&\mathbf{Pr}\left[\mathrm{maj}(z_{i,1}, \ldots, z_{i,\ell/T}) = z_{i,j}\right] = \\ &\mathbf{Pr}\left[\sum_{k \neq j} z_{i,k} = \frac{\ell/T - 1}{2}\right] + \frac{1}{2}\mathbf{Pr}\left[\sum_{k \neq j} z_{i,k} \neq \frac{\ell/T - 1}{2}\right] \\ &\quad = \frac{1}{2} + \binom{\ell/T - 1}{\frac{\ell/T-1}{2}} \cdot 2^{-\ell/T} = \frac{1}{2} + \Omega\left(\sqrt{\frac{T}{\ell}}\right).\end{aligned}$$

The first equality holds because if $\sum_{k \neq j} z_{i,k} = \frac{\ell/T-1}{2}$, $z_{i,j}$ determines the majority, otherwise $z_{i,j}$ is an independent bit of the majority. The last equation holds, because $\sqrt{2\pi n}\,(n/e)^n \leq n! \leq e\sqrt{n}\,(n/e)^n$ so $\binom{n}{n/2} = \Omega(2^n/\sqrt{n})$. Finally, by averaging over x, we conclude that $\mathcal{A}$ predicts correctly with probability $1/2 + \Omega(\sqrt{T/\ell}) = 1/2 + \Omega(\sqrt{ST/N})$. This completes the proof.

5 A general approach for proving non-uniform security

In this section, we provide a general approach for proving non-uniform security bounds of cryptographic applications, by "translating" our concentration bounds into the statements about security. To illustrate usage of this approach, we reprove Theorem 1.2 as an example.

5.1 Setting up the language To illustrate our approach, we formally define attackers with oracle-dependent advice, and other notions from relevant cryptographic applications. Our definitions are adopted from Coretti et al. [CDGS18] (with slight simplifications). We stress that to derive results in our application, it suffices to focus on concrete settings, which we will mention along the way.

Oracle model. An $\mathcal{O}$-model is defined by specifying a domain $[N]$, a range $[M]$, and a distribution over oracle functions from $[N]$ to $[M]$. $\mathcal{O}$-models capture a variety of idealized model in cryptography, including the random oracle model (i.e., $\mathcal{O} : [N] \to [M]$ is a random function), the random permutation model (i.e, $\mathcal{O} : [N] \to [N]$ is a random permutation), generic group models (i.e., $\mathcal{O} : [N] \to [M]$ is a random injection), etc. For our main result, we only need to consider $\mathcal{O} := (f, P)$-model where $f : [N] \to [N]$ is a random permutation, and $P : [N] \to \{0, 1\}$ is a random function.

Attackers with oracle-dependent advice. Attackers $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ consist of a preprocessing procedure $\mathcal{A}_1$ and an online algorithm $\mathcal{A}_2$, which carries out the actual attack using the output of $\mathcal{A}_1$. In the presence of an oracle $\mathcal{O}$, both $\mathcal{A}_1$ and $\mathcal{A}_2$ interacts with $\mathcal{O}$.

Definition 5.1. *An* (S, T)-attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ in the $\mathcal{O}$-model *consists of two procedures*

- $\mathcal{A}_1$*, which is computationally unbounded, interacts with* $\mathcal{O}$*, and outputs an* S*-bit string, and*
- $\mathcal{A}_2$*, which takes an* S*-bit auxiliary input and makes at most* T *queries to* $\mathcal{O}$*.*

Cryptographic applications. An application G in the $\mathcal{O}$-model is defined by specifying a challenger C, which is an oracle algorithm that has access to $\mathcal{O}$, interacts with the main stage $\mathcal{A}_2$ of an attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, and outputs a bit at the end of the interaction. The *success* of $\mathcal{A}$ on G in the $\mathcal{O}$-model is defined as

$$\mathrm{Succ}_{G,\mathcal{O}}(\mathcal{A}) \;:=\; \mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}) \leftrightarrow \mathsf{C}^{\mathcal{O}} = 1\right],$$

where $\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}) \leftrightarrow \mathsf{C}^{\mathcal{O}}$ denotes the bit output by C after its interaction with the attacker.

We consider two types of applications, which are captured by the following definition.

Definition 5.2. *For an application* G *in the* $\mathcal{O}$*-model, we define*

Advantage of attacker $\mathcal{A}$ for indistinguishability G:
$$\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}) \overset{def}{=} 2 \cdot \mathrm{Succ}_{G,\mathcal{O}}(\mathcal{A}) - 1\,.$$

Advantage of attacker $\mathcal{A}$ for unpredicatability G:
$$\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}) \overset{def}{=} \mathrm{Succ}_{G,\mathcal{O}}(\mathcal{A})\,.$$

$((S, T), \varepsilon)$**-secure:** G *is* $((S, T), \varepsilon)$-secure *if for every* (S, T)*-attacker* $\mathcal{A}$ $\quad \mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}) \leq \varepsilon$.

A typical example of unpredictability application is one-way permutation denoted as OWP. The challenger C sends a y to $\mathcal{A}$ for a random y from $[N]$, and outputs 1 if $\mathcal{A}$ returns of $f^{-1}(y)$ where $f : [N] \to [N]$ is a random permutation. The advantage of $\mathcal{A}$ is the same as the success probability of $\mathcal{A}$, captured by the following probability,

$$\begin{aligned}\mathrm{Adv}_{\mathrm{OWP},f}(\mathcal{A}) &= \mathrm{Succ}_{G,f}(\mathcal{A}) \\ &= \mathbf{Pr}\left[\mathcal{A}_2^f(\mathcal{A}_1^f, y) = f^{-1}(y)\right],\end{aligned}$$

where the probability is taken over f,y and the randomness of $\mathcal{A}$.

Unlike unpredictability applications, indistinguishability applications admit a random guess attack, which easily achieves 1/2 success probability. Therefore, the advantage for indistinguishability application is measured by how much it is better than 1/2.

To prove our main result, we only need to consider the indistinguishability application of hardcore predicate HCP in the $\mathcal{O} = (f, P)$-model. The application is specified by the C, which samples a random y from $[N]$, sends y to $\mathcal{A}$, and outputs 1 if and only if $\mathcal{A}$ answers $P(f^{-1}(y))$[7]. The advantage of $\mathcal{A}$ is captured by the following probability.

$$\begin{aligned}\mathrm{Adv}_{\mathrm{HCP},\mathcal{O}}(\mathcal{A}) &= 2 \cdot \mathrm{Succ}_{\mathrm{HCP},\mathcal{O}}(\mathcal{A}) - 1 \\ &= 2\,\mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, y) = P(f^{-1}(y))\right] - 1,\end{aligned}$$

where the probability is taken over $\mathcal{O}$,y and the randomness of $\mathcal{A}$.

k-wise product and XOR applications. The translation of our product and XOR conditions, will be captured by the security of the following applications. Let $\mathcal{O}$ be an arbitrary oracle and G be an application in the $\mathcal{O}$-model.

The k-wise product application of G, denoted as $G^{\otimes k}$, is specified by the following challenger $\mathsf{C}^{\otimes k}$. $\mathsf{C}^{\otimes k}$ samples a random subset $I := \{r_1, \ldots, r_k\} \subseteq [N]$ of size k, where $[N]$ is the randomness space of C. $\mathsf{C}^{\otimes k}$ simulates C on $r_1, \ldots, r_k \in [N]$ (one by one), and outputs 1 if only if C outputs 1 on every $r_1, \ldots, r_k$ after its interaction with the attacker. The advantage of $\mathcal{A}$ for OWP$^{\oplus k}$ is captured by the following probability.

$$\mathrm{Adv}_{\mathrm{OWP}^{\otimes k},\mathcal{O}}(\mathcal{A}) = \mathrm{Succ}_{\mathrm{OWP}^{\otimes k},\mathcal{O}}(\mathcal{A}) = \mathbf{Pr}\left[Z = 1\right] \ .$$

where probability is over $\mathcal{O}$,I and the randomness of $\mathcal{A}$.

As an example, we consider the k-wise product application of OWP. The challenger $\mathsf{C}^{\times k}$ sends a random subset of k elements $I = \{y_1, \ldots, y_k\}$ to $\mathcal{A}$, and outputs $\mathcal{Z} := \cap_{i=1}^k (1_{P^{-1}(y_i)=x_i})$ after $\mathcal{A}$ returns $x_1, \ldots, x_k$.

The k-wise XOR application of G, denoted as $G^{\otimes k}$, is specified by the challenger $\mathsf{C}^{\oplus k}$, which $\mathsf{C}^{\oplus k}$ simulates C on a random size-k subset $I := \{r_1, \ldots, r_k\} \subseteq [N]$ (one by one), and outputs the parity of the output of C on $r_1, \ldots, r_k$ after its interaction with the attacker.

To prove our main result, we only need to consider the k-wise XOR application of HCP, which is also an indistinguishability application. The $\mathsf{C}^{\oplus k}$ sends a random subset of k elements $I = \{y_1, \ldots, y_k\}$ to $\mathcal{A}$, and outputs $\mathcal{Z} := \oplus_{i=1}^k (1_{P(f^{-1}(y_i))=b_i})$ after $\mathcal{A}$ returns $b_1, \ldots, b_k$. Notice that because $1_{P(f^{-1}(y_i))=b_i} = P(f^{-1}(y_i)) \oplus b_i \oplus 1$, $\mathcal{Z}$ can be written as

$$\mathcal{Z} = \left(\oplus_{i=1}^k P(f^{-1}(y_i))\right) \oplus \left(\oplus_{i=1}^k b_i\right) \oplus \left(\oplus_{i=1}^k 1\right) \ .$$

Hence the advantage of $\mathcal{A}$ for HCP$^{\oplus k}$ is captured by the following probability.

$$\begin{aligned}\mathrm{Adv}_{\mathrm{HCP}^{\oplus k},\mathcal{O}}(\mathcal{A}) &= 2 \cdot \mathrm{Succ}_{\mathrm{HCP}^{\oplus k},\mathcal{O}}(\mathcal{A}) - 1 \\ &= 2\,\mathbf{Pr}\left[Z = 1\right] - 1 \ .\end{aligned}$$

where probability is over $\mathcal{O}$,I and the randomness of $\mathcal{A}$. Equivalently, $\mathcal{A}$ aims to predict $\oplus_{i=1}^k P(f^{-1}(y_i))$ (or its negation) in the application of HCP$^{\oplus k}$.

5.2 Reducing non-uniform security to product or XOR security: Proof of Lemma 1.1

LEMMA 5.1. *Let $\mathcal{O}$ be an arbitrary model, and G be an arbitrary application in $\mathcal{O}$-model.*

1. *If G is an unpredictability application, and $G^{\otimes k}$ is $((0, Tk), \varepsilon^k)$-secure for any positive integer $k \leq S + \log(1/\gamma)$ in the $\mathcal{O}$-model, then G is $((S, T), \varepsilon')$-secure in the $\mathcal{O}$-model, where*
$$\varepsilon' \leq 6\varepsilon + \frac{S + \log(1/\gamma)}{N} + \gamma \ .$$
2. *If G is an indistinguishability application, and $G^{\oplus k}$ is $((0, Tk), \varepsilon^k)$-secure for any positive integer $k \leq S + \log(1/\gamma)$ in the $\mathcal{O}$-model, then G is $((S, T), \varepsilon')$-secure in the $\mathcal{O}$-model, where*
$$\varepsilon' \leq 2\varepsilon + 2\sqrt{\frac{S + \log(1/\gamma)}{N}} + \gamma \ .$$

where $[N]$ is the randomness space used by the challenger associated with G.

For most of applications we can set $\gamma = 1/N$, which will be dominated by other terms. Thus we will ignore γ and poly $\log(N)$ terms in the following discussion for the purpose of clarity.

The above lemma works for an arbitrary $\mathcal{O}$-model, thus it applies to a variety of well studied cryptographic idealized model, such as the random oracle model, random permutation and generic group model. A common difficulty in analyzing non-uniform security in those models is that most of the standard techniques do not work when preprocessing and oracle-dependent advice are allowed. In particular, for the random oracle model, many techniques are based on the independence

[7] We remark for any permutation f, C is equivalent to C' which samples a random x from $[N]$, sends $f(x)$ to $\mathcal{A}$, and outputs 1 if and only if $\mathcal{A}$ answers $P(x)$.

of values in different inputs, which does not hold for attackers with even a single bit of advice.

For an arbitrary idealized model, the above lemma reduces proving security of an application G against attackers with S-bit of advice to proving the security of its S-wise product and XOR variants against attackers with *no advice*. The latter statement is considered to be much simpler task, because standard techniques are applicable again.

Moreover, the blow up from ε to ε' is very small, which is appealing and crucial for obtaining tight non-uniform security bounds for our application. In particular, for unpredictability applications, $\varepsilon' = O(\varepsilon + S/N)$, and for indistinguishability applications, $\varepsilon' = O(\varepsilon + \sqrt{S/N})$.

Proof. Let $\mathcal{A}$ be an (S,T)-attacker for G in the $\mathcal{O}$-model. Without loss of generality, we assume that $\mathcal{A}$ is deterministic. Our proof for both cases follows the same template. We will show for some κ and γ, and any fixed S-bit string w,

$$\mathbf{Pr}_{\mathcal{O}}\left[\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}_w,\mathcal{A}_2) \geq \kappa\right] \leq 2^{-S}\cdot\gamma$$

where $\mathcal{A}_w$ outputs the fixed advice string w, and we abuse the notion of $\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}_w,\mathcal{A}_2)$ to denote the advantage of $(\mathcal{A}_w,\mathcal{A}_2)$ respect to a fixed choice of $\mathcal{O}$. Then by an averaging argument and a union bound over 2^S possible advice, we obtain

$$\begin{aligned}\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}_1,\mathcal{A}_2) \leq 1\cdot\mathbf{Pr}_{\mathcal{O}}\left[\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}_1,\mathcal{A}_2)\geq\kappa\right]+\kappa\cdot 1\\ \leq \gamma+\kappa\ .\end{aligned}$$

Fix an arbitrary $w \in \{0,1\}^S$ and let $\mathcal{A} = (\mathcal{A}_w,\mathcal{A}_2)$. For $r\in[N]$, let X_r be the output bit of $(\mathcal{A}^{\mathcal{O}} \leftrightarrow \mathsf{C}^{\mathcal{O}}(;r))$ meaning the output of C using randomness r.

If G is an unpredictability application, we consider the $\mathcal{A}^{\otimes k}$ for $G^{\otimes k}$ which simply runs $\mathcal{A}$ separately on given k challenges from $C^{\otimes k}$ with at most Tk queries and a fixed advice. Suppose $G^{\otimes k}$ is $(0,Tk,\varepsilon^k)$-secure, then,

$$\begin{aligned}\mathbf{Pr}_I\left[\Pi_{i\in I}X_i = 1\right] &= \Pr[(\mathcal{A}^{\otimes k})^{\mathcal{O}} \leftrightarrow (\mathsf{C}^{\otimes k})^{\mathcal{O}} = 1]\\ &= \mathrm{Adv}_{G^{\otimes k},\mathcal{O}}(\mathcal{A}^{\otimes k}) \leq \varepsilon^k.\end{aligned}$$

Hence, $X_1,\ldots,X_N$ satisfies average ε-product condition for $k\leq S+\log 1/\gamma$. By Theorem 3.1,

$$\begin{aligned}\mathbf{Pr}_{\mathcal{O}}\left[X_1+\cdots+X_N \geq 6\varepsilon N+k\right] &\leq \frac{\binom{N}{k}\varepsilon^k}{\binom{6\varepsilon N+k}{k}}\\ &\leq \left(\frac{Ne\varepsilon/k}{6\varepsilon N/k}\right)^k \leq 2^{-k}.\end{aligned}$$

where the second inequality is because $\binom{N}{k}\leq(Ne/k)^k$ and $\binom{6\varepsilon N+k}{k}\geq(6\varepsilon N/k)^k$. Observe that $\mathrm{Adv}_{G,\mathcal{A}}(\mathcal{O})$ is $\sum_{i\in[N]}X_i/N$. By setting $k = S+\log 1/\gamma$ and $\kappa = 6\varepsilon + k/N$, we obtain the desired conclusion.

If G is an indistinguishable application, we consider $\mathcal{A}^{\oplus k}$ for $G^{\oplus k}$ which simply runs $\mathcal{A}$ on given challenges from $C^{\oplus k}$ with at most Tk queries and a fixed advice. Suppose $G^{\oplus k}$ is $(0,Tk,\varepsilon^k)$-secure, then,

$$\begin{aligned}\mathrm{Bias}(\oplus_{i\in I}X_i) &= 2\Pr[(\mathcal{A}^{\otimes k})^{\mathcal{O}} \leftrightarrow (\mathsf{C}^{\otimes k})^{\mathcal{O}} = 1]-1\\ &= \mathrm{Adv}_{G^{\oplus k},\mathcal{O}}(\mathcal{A}^{\oplus k})\leq\varepsilon^k.\end{aligned}$$

Hence, $X_1,\ldots,X_N$ satisfies (average) ε-XOR condition for $k\leq S+\log 1/\gamma$. By Theorem 3.3,

$$\begin{aligned}\mathbf{Pr}_{\mathcal{O}}\left[X_1+\cdots+X_N\geq\frac{N(1+2\varepsilon+2\sqrt{k/N})}{2}\right]\\ \leq\left(\frac{\varepsilon+\sqrt{k/N}}{2\varepsilon+2\sqrt{k/N}}\right)^k = 2^{-k}.\end{aligned}$$

Observe that $\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A})$ is $2\sum_{i\in[N]}X_i/N-1$. By setting $k=S+\log 1/\gamma$, $\kappa = 2\varepsilon+2\sqrt{k/N}$, we obtain the desired conclusion.

□

5.3 An XOR lemma for hardcore predicates

In this section, we prove the following lemma which essentially translates Lemma 4.1 using the language of XOR-security.

LEMMA 5.2. *$\mathrm{HCP}^{\oplus k}$ is $((0,Tk),(6kT/N)^k)$-secure in the $\mathcal{O}=(f,P)$-model, where $\mathcal{O}=(f,P)$, $f:[N]\to[N]$ is a random permutation, and $P:[N]\to\{0,1\}$ is a random function.*

Proof. For $k\geq N/6T$, the statement holds trivial. Let $k\leq N/6T$. Let $\mathcal{A}$ be the best $(0,Tk)$-attacker for HCP^k. Without loss of generality, we assume that $\mathcal{A}$ is deterministic.

Recall that the challenger $\mathsf{C}^{\oplus k}$ for $\mathrm{HCP}^{\oplus k}$ sends a random subset of k elements $I=\{y_1,\ldots,y_k\}\subseteq[N]$ to $\mathcal{A}$, and outputs $\mathcal{Z} := \oplus_{i=1}^k(1_{P(f^{-1}(y_i))=b_i})$ after $\mathcal{A}$ returns $b_1,\ldots,b_k$. And because $1_{P(f^{-1}(y_i))=b_i} = P(f^{-1}(y_i))\oplus b_i\oplus 1$, $\mathcal{Z}$ can be written as

$$\mathcal{Z} = \left(\oplus_{i=1}^k P(f^{-1}(y_i))\right)\oplus\left(\oplus_{i=1}^k b_i\right)\oplus\left(\oplus_{i=1}^k 1\right)\ .$$

Hence, the advantage of $\mathcal{A}$ for $\mathrm{HCP}^{\oplus k}$ is captured by the following probability.

$$\begin{aligned}\mathrm{Adv}_{\mathrm{HCP}^{\oplus k},\mathcal{O}}(\mathcal{A}) &= 2\cdot\mathrm{Succ}_{\mathrm{HCP}^{\oplus k},\mathcal{O}}(\mathcal{A})-1\\ &= 2\,\mathbf{Pr}\left[Z=1\right]-1,\end{aligned}$$

where probability is taken over $\mathcal{O}$,I. Equivalently, $\mathcal{A}$ aims to predict $(\oplus_{i=1}^{k} P(f^{-1}(y_i))) \oplus (\oplus_{i=1}^{k} b_i)$ (or its negation) in the application of $\mathrm{HCP}^{\oplus k}$. It suffices to prove that, conditioning on any fixed choice of $I = \{y_1, \ldots, y_k\}$,

$$2\,\mathbf{Pr}\left[\mathcal{Z} = 1\right] \leq 1 + \left(\frac{6kT}{N}\right)^{k},$$

where the randomness is over $\mathcal{O}$. By averaging over the random I, it implies the desired conclusion.

Let Γ be the random variable corresponding to the transcripts containing query/answer pairs $(x_1, \mathcal{O}(x_1), \ldots, x_q, \mathcal{O}(x_q))$ resulting from $\mathcal{A}$'s interaction with $\mathcal{O}$, where $q = kT$. Let I_Γ be the random variable corresponding to the set of $y \in I$, whose preimage under f has been queried by $\mathcal{A}$, in other words, $f^{-1}(y) = x_i$ for some $i \in [q]$.

Claim 3. $2\,\mathbf{Pr}[\mathcal{Z} = 1] \leq 1 + \mathbf{Pr}[I_\Gamma = I]$.

Proof. Let τ be an arbitrary possible transcript. Consider $\mathcal{Z}$ conditioned on $\Gamma = \tau$. Because $\mathcal{A}$ is deterministic, $\oplus_i^k b_i$ is a fixed bit. However,

$$\oplus_{i=1}^{k} P(f^{-1}(y_i)) = \left(\underset{y \in I_\tau}{\oplus} P(f^{-1}(y))\right) \oplus \left(\underset{y \in I \setminus I_\tau}{\oplus} P(f^{-1}(y))\right)$$

is fixed if $I_\tau = I$, otherwise it is a random bit because P is a random function on $f^{-1}(I \setminus I_\tau)$. Therefore,

$$\begin{aligned} \mathbf{Pr}\left[\mathcal{Z} = 1\right] &= \mathbf{Pr}\left[\mathcal{Z} = 1 | I_\Gamma = I\right] \cdot \mathbf{Pr}\left[I_\Gamma = I\right] \\ &\quad + \mathbf{Pr}\left[\mathcal{Z} = 1 | I_\Gamma \neq I\right] \cdot \mathbf{Pr}\left[I_\Gamma \neq I\right] \\ &\leq 1 \cdot \mathbf{Pr}\left[I_\Gamma = I\right] + \frac{1}{2} \cdot \mathbf{Pr}\left[I_\Gamma \neq I\right] = \frac{1 + \mathbf{Pr}[I_\Gamma = I]}{2}\,. \end{aligned}$$

The desired conclusion follows. $\square$

It remains to upper bound $\mathbf{Pr}[I_\Gamma = I]$ by $(6Tk/N)^k$. $\mathcal{A}_2$ makes at most q (adaptive) queries. Let E_i denote the event that the i-th distinct query gets mapped to some y in I under f. Conditioning on all previous query/answer pairs, then each E_i has conditional probability at most $k/(N-q) \leq 2k/N$ (recall $k \leq N/6T$). There are at most q such events. If $I_\Gamma = I$ happens, there must be at least k of these events that are true. For each of the $\binom{q}{k}$ sets of these events, the probability that all of these events occur simultaneously is at most $(2k/N)^k$. So the probability that $I_\Gamma = I$ happens is at most $\binom{q}{k}(2k/N)^{|I|} \leq (2eq/N)^k \leq (6Tk/N)^k$. $\square$

5.4 Putting things together: Proof of Theorem 1.2 By Lemma 5.2, for any k, $\mathrm{HCP}^{\oplus k}$ is $((0, Tk), (6kT/N)^k)$-secure in the $\mathcal{O}$-model where $\mathcal{O} = (f, P)$, $f : [N] \to [N]$ is a random permutation, and $P : [N] \to \{0, 1\}$ is a random predicate. We choose

$$\gamma := \frac{1}{N} \text{ and } \varepsilon := \frac{6(S + \log(1/\gamma))T}{N}\,.$$

So that, for any $k \leq S + \log(1/\gamma)$, $G^{\oplus k}$ is $(0, Tk, \varepsilon^k)$-secure in the $\mathcal{O}$-model. By Lemma 1.1, the indistinguishability application HCP is (S, T, ε')-secure in the $\mathcal{O}$-model, where

$$\varepsilon' \leq 2\varepsilon + 2\sqrt{\frac{S + \log(1/\gamma)}{N}} + \gamma = \tilde{O}\left(\frac{ST}{N} + \sqrt{\frac{S}{N}}\right)\,.$$

In other words, for any (S, T)-attacker $\mathcal{A}$,

$$\begin{aligned} \mathbf{Pr}\left[\mathcal{A}_2^{\mathcal{O}}(\mathcal{A}_1^{\mathcal{O}}, y) = P(f^{-1}(y))\right] &\leq \frac{1 + \varepsilon'}{2} \\ &= \frac{1}{2} + \tilde{O}\left(\frac{ST}{N} + \sqrt{\frac{S}{N}}\right)\,, \end{aligned}$$

which completes the proof.

Acknowledgments

Siyao Guo is deeply indebted to Yevgeniy Dodis, Daniel Wichs and Jiapeng Zhang for long discussions at the early stage of this work, which led to the question of proving strong concentration bounds for almost k-wise independent variables. We thank anonymous reviewers for their constructive comments.

References

[AAK$^+$07] Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, and Ning Xie. Testing k-wise and almost k-wise independence. In David S. Johnson and Uriel Feige, editors, *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 11-13, 2007*, pages 496–505. ACM, 2007.

[ACDW20] Akshima, David Cash, Andrew Drucker, and Hoeteck Wee. Time-space tradeoffs and short collisions in merkle-damgård hash functions. *Advances in Cryptology - CRYPTO 2020 - 40th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2020, Proceedings*, 2020.

[AGHP92] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple construction of almost k-wise independent random variables. *Random Struct. Algorithms*, 3(3):289–304, 1992.

[AGK20] Benedikt Auerbach, Federico Giacon, and Eike Kiltz. Everybody's a target: Scalability in public-key encryption. In *Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part III*, pages 475–506, 2020.

[BR94] Mihir Bellare and John Rompel. Randomness-efficient oblivious sampling. In *35th Annual Symposium on Foundations of Computer Science, Santa Fe, New Mexico, USA, 20-22 November 1994*, pages 276–287. IEEE Computer Society, 1994.

[BRT12] Mihir Bellare, Thomas Ristenpart, and Stefano Tessaro. Multi-instance security and its application to password-based cryptography. In *Advances in Cryptology - CRYPTO 2012 - 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2012. Proceedings*, pages 312–329, 2012.

[CDG18] Sandro Coretti, Yevgeniy Dodis, and Siyao Guo. Non-uniform bounds in the random-permutation, ideal-cipher, and generic-group models. In *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part I*, pages 693–721, 2018.

[CDGS18] Sandro Coretti, Yevgeniy Dodis, Siyao Guo, and John P. Steinberger. Random oracles and non-uniformity. In *EUROCRYPT 2018*, pages 227–258. Springer, 2018.

[CGLQ20] Kai-Min Chung, Siyao Guo, Qipeng Liu, and Luowen Qian. Tight quantum time-space tradeoffs for function inversion. *FOCS 2020*, 2020.

[Che52] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, 23(4):493–507, 1952.

[CHM20] Dror Chawin, Iftach Haitner, and Noam Mazor. Lower bounds on the time/memory tradeoff of function inversion. *Electronic Colloquium on Computational Complexity (ECCC)*, 27:89, 2020.

[CK18] Henry Corrigan-Gibbs and Dmitry Kogan. The discrete-logarithm problem with preprocessing. In *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, pages 415–447, 2018.

[CK19] Henry Corrigan-Gibbs and Dmitry Kogan. The function-inversion problem: Barriers and opportunities. In *Theory of Cryptography - 17th International Conference, TCC 2019, Nuremberg, Germany, December 1-5, 2019, Proceedings, Part I*, pages 393–421, 2019.

[DGK17] Yevgeniy Dodis, Siyao Guo, and Jonathan Katz. Fixing cracks in the concrete: Random oracles with auxiliary input, revisited. In *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part II*, pages 473–495, 2017.

[DTT10] Anindya De, Luca Trevisan, and Madhur Tulsiani. Time space tradeoffs for attacks against one-way functions and PRGs. In *CRYPTO 2010*, pages 649–665. Springer, 2010.

[Hel80] Martin E. Hellman. A cryptanalytic time-memory trade-off. *IEEE Trans. Inf. Theory*, 26(4):401–406, 1980.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[IK10] Russell Impagliazzo and Valentine Kabanets. Constructive proofs of concentration bounds. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 13th International Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010, Barcelona, Spain, September 1-3, 2010. Proceedings*, pages 617–631, 2010.

[Imp95] Russell Impagliazzo. Hard-core distributions for somewhat hard problems. In *36th Annual Symposium on Foundations of Computer Science, Milwaukee, Wisconsin, USA, 23-25 October 1995*, pages 538–545. IEEE Computer Society, 1995.

[Imp11] Russell Impagliazzo. Relativized separations of worst-case and average-case complexities for NP. In *Proceedings of the 26th Annual IEEE Conference on Computational Complexity, CCC 2011, San Jose, California, USA, June 8-10, 2011*, pages 104–114, 2011.

[Kab03] Valentine Kabanets. Almost k-wise independence and hard boolean functions. *Theoretical computer science*, 297(1-3):281–295, 2003.

[LL14] Nathan Linial and Zur Luria. Chernoff's inequality-a very elementary proof. *arXiv preprint arXiv:1403.7739*, 2014.

[MT09] Ueli M. Maurer and Stefano Tessaro. Computational indistinguishability amplification: Tight product theorems for system composition. In Shai Halevi, editor, *Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, volume 5677 of *Lecture Notes in Computer Science*, pages 355–373. Springer, 2009.

[MW19] Or Meir and Avi Wigderson. Prediction from partial information and hindsight, with application to circuit lower bounds. *Comput. Complex.*, 28(2):145–183, 2019.

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM journal on computing*, 22(4):838–856, 1993.

[PR17] Christos Pelekis and Jan Ramon. Hoeffding's inequality for sums of dependent random variables. *Mediterranean Journal of Mathematics*, 14(6):1–16, 2017.

[SSS93] Jeanette P. Schmidt, Alan Siegel, and Aravind Srinivasan. Chernoff-hoeffding bounds for applications with limited independence. In Vijaya Ramachandran, editor, *Proceedings of the Fourth Annual ACM/SIGACT-SIAM Symposium on Discrete Al-*

gorithms, 25-27 January 1993, Austin, Texas, USA, pages 331–340. ACM/SIAM, 1993.

[ST18] Alexander V. Smal and Navid Talebanfard. Prediction from partial information and hindsight, an alternative proof. *Inf. Process. Lett.*, 136:102–104, 2018.

[Ung09] Falk Unger. A probabilistic inequality with applications to threshold direct-product theorems. In *50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA*, pages 221–229. IEEE Computer Society, 2009.

[Unr07] Dominique Unruh. Random oracles and auxiliary input. In *CRYPTO 2007*, pages 205–223. Springer, 2007.

[Vaz87] Umesh V. Vazirani. Efficiency considerations in using semi-random sources (extended abstract). In Alfred V. Aho, editor, *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 160–168. ACM, 1987.

[Yao82] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions. In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 80–91, 1982.

[Yao90] Andrew Chi-Chih Yao. Coherent functions and program checkers (extended abstract). In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*, pages 84–94, 1990.

A Calculation of bounds in Theorem 3.2

CLAIM 4. *Suppose $a = (1-\varepsilon)b$ and $c = (1+\gamma)b$, where $\varepsilon < \gamma < 1/(2k)$. Then the solution p to the system of equations (3.3) ($f_1(p,s) = a$ and $f_k(p,s) = b^k$) satisfies*

$$\frac{2\varepsilon}{(k-1)\gamma^2} \geq p \geq \frac{\varepsilon - k\varepsilon^2}{k\gamma^2 + 2\varepsilon\gamma k + \varepsilon} = \Omega\left(\frac{\varepsilon}{\varepsilon + k\gamma^2}\right).$$

Proof. Observe that $f_1(p,s) = pc + (1-p)s = a$ implies $s = (1 - (\varepsilon + p\gamma)/(1-p))b$. Hence,

$$p(1+\gamma)^k + (1-p)\left(1 - \frac{\varepsilon + p\gamma}{1-p}\right)^k = 1.$$

Since $1 + kx \leq (1+x)^k$ for all x, and $1 + kx + k(k-1)x^2/2 \leq (1+x)^k$ for all $x \geq 0$, we have

$$\begin{aligned} 1 &\geq p\left(1 + k\gamma + \frac{k(k-1)\gamma^2}{2}\right) + (1-p)\left(1 - \frac{k(\varepsilon + p\gamma)}{1-p}\right) \\ &= 1 + pk\gamma + p\frac{k(k-1)\gamma^2}{2} - k(\varepsilon + p\gamma). \end{aligned}$$

This means

$$p \leq \frac{2\varepsilon}{(k-1)\gamma^2}.$$

On the other hand, since $(1+x)^k \leq e^{kx} \leq 1 + kx + k^2x^2$ assuming $kx \leq 1$, we have

$$\begin{aligned} 1 \leq\ & p(1 + k\gamma + k^2\gamma^2) \\ &+ (1-p)\left(1 - k\frac{\varepsilon + p\gamma}{1-p} + k^2\left(\frac{\varepsilon + p\gamma}{1-p}\right)^2\right). \end{aligned}$$

This implies

$$p \geq \frac{\varepsilon - k\varepsilon^2}{k\gamma^2 + 2\varepsilon\gamma k + \varepsilon}.$$

□

When $a = (1-\varepsilon)/2$, $b = (1+\varepsilon)/2$, $c = (1+\gamma)/2$ with $\varepsilon = k/n$ and $\gamma = \Theta(\sqrt{k/n})$, Claim 4 gives us $p \geq \frac{2\varepsilon - k(2\varepsilon)^2}{2\varepsilon + 4k\gamma\varepsilon + k\gamma^2} = \frac{2k/n}{k\gamma^2}(1 + o(1)) = \frac{2}{n\gamma^2}(1 + o(1))$. Theorem 3.2 gives us the an example, i.e., provides a lower bound on the tail probability

$$\begin{aligned} \mathbf{Pr}\left[\sum_{i=1}^{n} X_i \geq cn\right] &\geq p - \frac{(k/n)\cdot(2(a + k/n))}{(c - a - k/n)^2(k-1)} \\ &= p - \frac{k/n}{k(\gamma + 2\varepsilon - k/n)^2}(1 + o(1)) = \frac{2(1 + o(1))}{n\gamma^2} \\ &\quad - \frac{(1 + o(1))}{n\gamma^2} = \frac{1}{n\Theta(k/n)}(1 + o(1)) = \Omega(1/k). \end{aligned}$$

B General XOR Condition

In this subsection, we consider general XOR condition and prove the following theorem.

THEOREM B.1. *Suppose $X_1, X_2, \ldots, X_n$ are random variables, and $a_1, a_2, \ldots, a_n$ are approximate expectations of these variables, with the following constraints satisfied.*

1. *(Bounded range) $|X_i - a_i| \le R$ with probability 1 for all i.*
2. *(General XOR condition) For any subset $I \subseteq [n]$ with $|I| \le k$, we have*

$$\mathbf{E}\left[\prod_{i\in I}(X_i - a_i)\right] \le \varepsilon^{|I|}.$$

Then we have the following concentration bound.

$$\mathbf{Pr}\left[\sum_{i=1}^n X_i \ge \sum_{i=1}^n a_i + \gamma n\right] \le \sqrt{2}\left((\epsilon + R\sqrt{\frac{k}{en}})\frac{1}{\gamma}\right)^k.$$

In particular when X_i are ± 1 random variables with the ϵ-XOR condition, then take $a_i = 0$ and $R = 1$ and we have

$$\mathbf{Pr}\left[\sum_{i=1}^n X_i \ge \gamma n\right] \le \sqrt{2}\left(\left(\epsilon + \sqrt{\frac{k}{en}}\right)\frac{1}{\gamma}\right)^k.$$

Proof. Let $Y_i = X_i - a_i$. Consider the k-th moment method.

$$\mathbf{Pr}\left[\sum_{i=1}^n Y_i \ge \eta n\right] = \mathbf{Pr}\left[\left(\sum_{i=1}^n Y_i\right)^k \ge (\eta n)^k\right] \le \mathbf{E}\left[\left(\sum_{i=1}^n Y_i\right)^k\right]/(\eta n)^k.$$

We shall expand $(\sum_i Y_i)^k$ and regroup the terms. Each term is in the form $\prod_{l=1}^k Y_{i_l} = \prod_{i=1}^n Y_i^{p_i}$, where p_i is the number of times Y_i appears in the term. Let j be the number of odd p_is. Then $\mathbf{E}[\prod_{l=1}^k Y_{i_l}] \le R^{k-j}\epsilon^j$ by replacing every Y_i with R except j of them.

Now we count the number of terms with exactly j odd p_is. We can uniquely determine such a term by the following procedure. First we fix a subset $S \subseteq [k]$ which represents the indices of odd items. So $l \in S$ means Y_{i_l} appears odd times. If i_l appear more than once, we will just choose any of them. So $|S| = j$. Then we choose i_l for each of the $l \in S$. So far we have at most $\binom{k}{j}$ possibilities for choosing S and at most n^j possibilities for choosing each of the i_l.

The remaining items outside S can be grouped into pairs. We consider the first index $i_l \notin S$. We have n choices for i_l and $k - j - 1$ choices for the position of its pair. Similarly the second unmatched index have at most $n \cdot (k - j - 3)$ possibilities, and so on. So the remaining items has at most $n^{(k-j)/2} \cdot (k-j-1)(k-j-3) \cdot \ldots \cdot 1 \le \sqrt{2}n^{(k-j)/2}(\frac{k-j}{e})^{(k-j)/2}$ possibilities. The last equality holds since

$$(m-1)(m-3)\ldots 1 = \frac{m!}{(m/2)!2^{m/2}} \le \sqrt{2}(m/e)^m,$$

as

$$\frac{1}{2}\log\frac{m}{2} + \log(\frac{m}{2}+1) + \log(\frac{m}{2}+2) + \ldots + \log(m-1) + \frac{1}{2}\log m \le \int_{m/2}^m \log x dx = \frac{m}{2}\log m - \frac{m}{2} + \frac{m}{2}\log 2$$

Summing up, we have

$$\mathbf{E}\left[\left(\sum_{i=1}^n Y_i\right)^k\right] \le \sum_{j=0}^n R^{k-j}\epsilon^j\binom{k}{j}n^j\sqrt{2}n^{\frac{k-j}{2}} \cdot \left(\frac{k-j}{e}\right)^{\frac{k-j}{2}} \le \sqrt{2}\sum_{j=0}^n\binom{k}{j}R^{k-j}\epsilon^j n^j n^{\frac{k-j}{2}}\left(\frac{k}{e}\right)^{\frac{k-j}{2}}$$

$$= \sqrt{2}\sum_{j=0}^n\binom{k}{j}(\epsilon n)^j\left(R\sqrt{\frac{nk}{e}}\right)^{k-j}$$

$$= \sqrt{2}\left(\epsilon n + R\sqrt{\frac{nk}{e}}\right)^k.$$

This completes the proof. □