Non-Uniform Bounds in the Random-Permutation, Ideal-Cipher, and Generic-Group Models

Sandro Coretti*
New York University
corettis@nyu.edu

Yevgeniy Dodis†
New York University
dodis@cs.nyu.edu

Siyao Guo‡
Northeastern University
s.guo@neu.edu

June 4, 2018

Abstract

The random-permutation model (RPM) and the ideal-cipher model (ICM) are idealized models that offer a simple and intuitive way to assess the conjectured standard-model security of many important symmetric-key and hash-function constructions. Similarly, the generic-group model (GGM) captures generic algorithms against assumptions in cyclic groups by modeling encodings of group elements as random injections and allows to derive simple bounds on the advantage of such algorithms.

Unfortunately, both well-known attacks, e.g., based on rainbow tables (Hellman, IEEE Transactions on Information Theory '80), and more recent ones, e.g., against the discrete-logarithm problem (Corrigan-Gibbs and Kogan, EUROCRYPT '18), suggest that the concrete security bounds one obtains from such idealized proofs are often *completely inaccurate* if one considers *non-uniform* or *preprocessing* attacks in the standard model. To remedy this situation, this work

- defines the auxiliary-input (AI) RPM/ICM/GGM, which capture both non-uniform and preprocessing attacks by allowing an attacker to leak an arbitrary (bounded-output) function of the oracle's function table;
- derives the *first* non-uniform bounds for a number of important practical applications in the AI-RPM/ICM, including constructions based on the Merkle-Damgård and sponge paradigms, which underly the SHA hashing standards, and for AI-RPM/ICM applications with computational security; and
- using simpler proofs, recovers the AI-GGM security bounds obtained by Corrigan-Gibbs and Kogan against preprocessing attackers, for a number of assumptions related to cyclic groups, such as discrete logarithms and Diffie-Hellman problems, and provides new bounds for two assumptions.

An important step in obtaining these results is to port the tools used in recent work by Coretti et al. (EUROCRYPT '18) from the ROM to the RPM/ICM/GGM, resulting in very powerful and easy-to-use tools for proving security bounds against non-uniform and preprocessing attacks.

*Supported by NSF grants 1314568 and 1319051.

†Partially supported by gifts from VMware Labs and Google, and NSF grants 1619158, 1319051, 1314568.

‡Supported by NSF grants CNS-1314722 and CNS-1413964; Part of this work done while the author was visiting the Simons Institute for the Theory of Computing at UC Berkeley.

1 Introduction

The random-permutation and ideal-cipher models. The random-permutation model (RPM) and the ideal-cipher model (ICM) are idealized models that offer a simple and intuitive way to prove the (conjectured) security of many important applications. This holds especially true in the realms of symmetric cryptography and hash-function design since most constructions of block ciphers and hash functions currently do not have solid theoretical foundations from the perspective of provable security. In fact, the *exact security bounds* obtained in such idealized models are often viewed as guidance for both designers and cryptanalysts in terms of the *best possible* security level that can be achieved by the corresponding construct in the standard model. By and large, this method has been quite successful in practice, as most separations between the standard model and various idealized models [12, 11, 39, 30, 3, 8] are somewhat contrived and artificial and are not believed to affect the security of widely used applications. In fact, the following *RPM/ICM methodology* appears to be a good way for practitioners to assess the best possible security level of a given (natural) application.

> **RPM/ICM methodology.** *For "natural" applications of hash functions and block ciphers, the concrete security proven in the RPM/ICM is the right bound even in the standard model, assuming the "best possible" instantiation for the idealized component (permutation or block cipher) is chosen.*

Both the RPM and the ICM have numerous very important practical applications. In fact, most practical constructions in symmetric-key cryptography and hash-function design are naturally defined in the RPM/ICM. The following are a few representative examples:

- The famous AES cipher is an example of key-alternating cipher, which can be abstractly described and analyzed in the RPM [2, 13], generalizing the Even-Mansour [24, 23] cipher $\mathsf{EM}_{\pi,s}(x) = \pi(x \oplus s) \oplus s$, where π is a public permutation, s is the secret key, and x is the message.
- The compression function of the SHA-1/2 [47, 42] and MD5 [44] hash functions, as well as the popular HMAC scheme [4], is implemented via the Davies-Meyer (DM) hash function $\mathsf{DM}_E(x, y) = E_x(y) \oplus y$, for a block cipher E. But its collision-resistance can only be analyzed in the ICM [52].
- The round permutation of SHA-3 [41]—as part of the *sponge mode of operation* [6]—can be defined in the RPM: given old n-bit state s and new r-bit block message x (where $r < n$), the new state is $s' = \pi(s \oplus (x \| 0^{n-r}))$, where π is a public permutation. The sponge mode is useful for building CRHFs, message authentication codes (MACs), pseudorandom functions (PRFs) [7], and key derivation functions [26], among others.
- The round function of MD6 [45] can be written as $f_Q(x) = \mathsf{trunc}_r(\pi(x \| Q))$, where Q is a constant, trunc_r is the truncation to r bits, and π is a public permutation. This construction was shown indifferentiable from a random oracle in the RPM [21].
- Many other candidate collision-resistant hash functions can be described using either ideal ciphers (e.g., the large PGV family [9]) or random permutations (e.g., [49, 6, 46, 21]).

The generic group model. Another well-known idealized model is the so-called generic-group model (GGM), which serves the purpose of proving lower bounds on the complexity of generic attacks against common computational problems in cyclic groups used in public-key cryptography,

such as the discrete-logarithm problem (DL), the computational and decisional Diffie-Hellman problems (CDH and DDH), and many more. Generic attacks are algorithms that do not exploit the specific representation of the elements of a group. This property is modeled by considering generic encoding captured by a random injection $\sigma : \mathbb{Z}_N \to [M]$ and allowing the algorithm access to a group-operation oracle, which, given a pair of encodings $(\sigma(x), \sigma(y))$, returns $\sigma(x + y)$.

The justification for the GGM is rooted in the fact that there are no unconditional hardness proofs for important group-related problems, and that there are some groups based on elliptic curves for which no better algorithms than the generic ones are known. Hence, results in the GGM provide at least some indication as to how sensible particular assumptions are. There are a plethora of security bounds proven in the GGM, e.g., lower bounds on the complexity of generic algorithms against DL or CDH/DDH by Shoup [48] or the knowledge-of-exponent assumption by Abe and Fehr [1] and Dent [18].

Non-uniformity and preprocessing. Unfortunately, a closer look reveals that the rosy picture above can only be true if one considers *uniform* attacks (as explained below). In contrast, most works (at least in theoretical cryptography) consider attackers in the *non-uniform* setting, where the attacker is allowed to obtain some arbitrary (but bounded) *advice* before attacking the system. The main rationale for this modeling comes from the realization that a determined attacker will know the parameters of a target system in advance and might be able to invest a significant amount of preprocessing to do something to speed up the actual attack, or to break many instances at once (therefore amortizing the one-time preprocessing cost). Perhaps the best known example of such attacks are *rainbow tables* [33, 40] (see also [35, Section 5.4.3]) for inverting arbitrary functions; the idea is to use one-time preprocessing to initialize a clever data structure in order to dramatically speed up brute-force inversion attacks. Thus, restricting to uniform attackers might not accurately model realistic preprocessing attacks one would like to protect against.

There are also other, more technical, reasons why the choice to consider non-uniform attackers is convenient (see [15] for details), the most important of which is security under composition. A well-known example are zero-knowledge proofs [29, 28], which are *not* closed under (even sequential) composition unless one allows non-uniform attackers and simulators. Of course, being a special case of general protocol composition, this means that any work that uses zero-knowledge proofs as a subroutine must consider security against *non-uniform* attackers in order for the composition to work. Hence, it is widely believed by the theoretical community that *non-uniformity is the right cryptographic modeling of attackers*, despite being overly conservative and including potentially unrealistic attackers—due to the potentially unbounded pre-computation allowed to generate the advice.

Idealized models vs. non-uniformity and preprocessing. When considering non-uniform attackers, it turns out that the RPM/ICM methodology above is blatantly false: once non-uniformity or preprocessing is allowed, the separations between the idealized models and the standard model are *no longer* contrived and artificial, but rather lead to *impossibly good* exact security of most *widely deployed* applications. To see this, consider the following examples:

- **One-way permutations:** Hellman [33] showed that there is a preprocessing attack that takes S bits of advice and makes T queries to a permutation $\pi : [N] \to N$ and inverts a random element of $[N]$ with probability roughly ST/N. Hence, a permutation cannot be one-way against attackers of size beyond $T = S = N^{1/2}$. However, in the RPM, a random permutation is easily shown to be invertible with probability at most T/N, therefore suggesting security

against attackers of size up to N.

- **Even-Mansour cipher:** In a more recent publication, Fouque *et al.* [25] showed a non-uniform $N^{1/3}$ attack against the Even-Mansour cipher that succeeds with constant probability. As with OWPs, the analysis in the RPM model suggests an incorrect security leve, namely, up to the birthday bound since one easily derives an upper bound of T^2/N on the distinguishing advantage of any attacker in RPM.

Similar examples also exist in the GGM:

- **Discrete logarithms:** A generic preprocessing attack by Mihalcik [38] and Bernstein and Lange [5] (and a recent variant by Corrigan-Gibbs and Kogan [16]) solves the DL problem with advantage ST^2/N in a group of order N, whereas the security of DL in the GGM is known to be T^2/N [48].
- **Square DDH:** A generic preprocessing attack by Corrigan-Gibbs and Kogan [16] breaks the so-called *square DDH (sqDDH)* problem—distinguishing (g^x, g^{x^2}) from (g^x, g^y) in a cyclic group $G = \langle g \rangle$ of order N—with advantage $\sqrt{ST^2/N}$, whereas the security of sqDDH in the GGM can be shown to be T^2/N.

1.1 Contributions: Non-Uniform Bounds in the RPM/ICM/GGM

Given the above failure of the idealized-models methodology, this paper revisits security bounds derived in the RPM, ICM, and GGM and re-analyzes a number of applications highly relevant in practice w.r.t. their security against non-uniform attackers or preprocessing. To that end, following the seminal work of Unruh [51] as well as follow-up papers by Dodis *et al.* [19] and Coretti *et al.* [15], the idealized models are replaced by weaker counterparts that adequately capture non-uniformity and preprocessing by allowing the attacker to obtain *oracle-dependent* advice. The resulting models, called the *auxiliary-input* RPM, ICM, and GGM, are parameterized by S ("space") and T ("time") and work as follows: The attacker $\mathcal{A}$ in the AI model consists of two entities $\mathcal{A}_1$ and $\mathcal{A}_2$. The first-stage attacker $\mathcal{A}_1$ is computationally unbounded, gets full access to the idealized primitive $\mathcal{O}$, and computes some advice z of size at most S. This advice is then passed to the second-stage attacker $\mathcal{A}_2$, who may make up to T queries to oracle $\mathcal{O}$ (and, unlike $\mathcal{A}_1$, may have additional application-specific restrictions, such as bounded running time, etc.). The oracle-dependent advice naturally maps to non-uniform advice when the random oracle is instantiated, and, indeed, none of the concerns expressed in the above examples remain valid in the AI-RPM/ICM/GGM.

Symmetric primitives. In the AI-RPM and AI-ICM, this work analyzes and derives non-uniform security bounds for (cf. Table 1 and Section 4):

- *basic applications* such as inverting a random permutation (OWP), the Even-Mansour cipher (EM), using the ideal cipher as a block cipher directly (BC-IC), the PRF security of Davies-Meyer (PRF-DM), the collision resistance of a salted version of the Davies-Meyer compression function (CRHF-DM);
- the collision-resistance, the PRF security, and the MAC security of the *sponge construction*, which underlies the SHA-3 hashing standard;
- the collision-resistance of the *Merkle-Damgård construction with Davies-Meyer (MD-DM)*, which underlies the SHA-1/2 hashing standards, and PRF/MAC security of NMAC and HMAC.

	AI Security	SM Security	Best Attack
OWP	$\frac{ST}{N}$	$\frac{T}{N}$	$\frac{ST}{N}$ [33]
EM	$\left(\frac{ST^2}{N}\right)^{1/2} + \frac{T^2}{N}$	$\frac{T^2}{N}$	$\left(\frac{S}{N}\right)^{1/2}$ [17]
BC-IC	$\left(\frac{ST}{K}\right)^{1/2} + \frac{T}{K}$	$\frac{T}{K}$	$\left(\frac{S}{K}\right)^{1/2}$ [17]
PRF-DM	$\left(\frac{ST}{N}\right)^{1/2} + \frac{T}{N}$	$\frac{T}{N}$	$\left(\frac{S}{N}\right)^{1/2}$ [17]
CRHF-DM	$\frac{(ST)^2}{N}$	$\frac{T^2}{N}$	not known
CRHF-S	$\frac{ST^2}{2^c} + \frac{T^2}{2^r}$	$\frac{T^2}{2^c} + \frac{T^2}{2^r}$	$\frac{ST^2}{N}$ [15]
PRF-S	$\left(\frac{ST^2}{2^c}\right)^{1/2}$	$\frac{T^2}{2^c}$	$\left(\frac{S}{N}\right)^{1/2}$ [17]
MAC-S	$\frac{ST^2}{2^c} + \frac{T}{2^r}$	$\frac{T^2}{2^c} + \frac{T}{2^r}$	$\min\left\{\frac{ST}{N}, \left(\frac{S^2T}{N^2}\right)^{1/3}\right\} + \frac{T}{N}$ [33]
CRHF-MD	$\frac{ST^2}{N}$	$\frac{T^2}{N}$	$\frac{ST^2}{N}$ [15]
PRF-MD-N	$\left(\frac{ST^3}{N}\right)^{1/2} + \frac{T^3}{N}$	$\frac{T^3}{N}$	$\left(\frac{S}{N}\right)^{1/2}$ [17]
NMAC/HMAC	$\frac{ST^3}{N}$	$\frac{T^3}{N}$	$\min\left\{\frac{ST}{N}, \left(\frac{S^2T}{N^2}\right)^{1/3}\right\} + \frac{T}{N}$ [33]

Table 1: *Asymptotic upper and lower bounds on the security of applications in the AI-ICM/AI-RPM and in the standard model (SM) against (S, T)-attackers.*

Surprisingly, except for OWPs [17], no non-uniform bounds were known for any of the above applications; not even for applications as fundamental as BC-IC, Even-Mansour, or HMAC.

The bounds derived for OWP and the collision-resistance (CR) of Sponges and MD-DM are tight, i.e., there exist matching attacks by Hellman [33] (for OWPs) and by Coretti *et al.* [15] (for CR). For the remaining primitives significant gaps remain between the derived security bounds and the best known attacks. Closing these gaps is left as an interesting (and important) open problem.

Generic groups. In the AI-GGM, the following applications are analyzed w.r.t. their security against preprocessing (cf. Table 2 and Section 5): the discrete-logarithm problem (DL), the multiple-discrete-logarithms problem (MDL), the computational Diffie-Hellman problem (CDH), the decisional Diffie-Hellman problem (DDH), the square decisional Diffie-Hellman problem (sqDDH), the one-more discrete-logarithm problem (OM-DL), and the knowledge-of-exponent assumption (KEA).

- For DL, MDL, CDH, DDH, and sqDDH, the derived bounds match those obtained in recent work by Corrigan-Gibbs and Kogan [16]. As highlighted below, however, the techniques used in this paper allow for much simpler proofs than the one based on incompressibility arguments in [16]. All of these bounds are tight, except those for DDH, for which closing the gap remains an open problem.
- The bounds for OM-DL and KEA are new and may be non-trivial to derive using compression techniques.

Computational security. Idealized models such as the ROM, RPM, and ICM are also often used in conjunction with computational hardness assumptions such as one-way functions, hardness of factoring, etc. Therefore, this paper also analyzes the security of public-key encryption based on

	AI-GGM Security	GGM Security	Best Attack
DL/CDH	$\frac{ST^2}{N} + \frac{T^2}{N}$	$\frac{T^2}{N}$	$\frac{ST^2}{N}$ [16, 38, 5]
t-fold MDL	$\left(\frac{S(T+t)^2}{tN} + \frac{(T+t)^2}{tN}\right)^t$	$\left(\frac{(T+t)^2}{tN}\right)^t$	see caption [16]
DDH	$\left(\frac{ST^2}{N}\right)^{1/2} + \frac{T^2}{N}$	$\frac{T^2}{N}$	$\frac{ST^2}{N}$ [16, 38, 5]
sqDDH	$\left(\frac{ST^2}{N}\right)^{1/2} + \frac{T^2}{N}$	$\frac{T^2}{N}$	$\left(\frac{ST^2}{N}\right)^{1/2}$ [16]
OM-DL	$\left(\frac{S(T+t)^2}{N}\right) + \frac{(T+t)^2}{N}$	$\frac{T^2}{N}$	$\frac{ST^2}{N}$ [16, 38, 5]
KEA	$\frac{ST^2}{N}$	$\frac{T^2}{N}$	not known

Table 2: *Asymptotic upper and lower bounds on the security of applications in the generic-group model against (S,T)-attackers in the AI-ROM; new bounds are in a bold-face font. The value t for the one-more DL problem stands for the number of challenges requested by the attacker. The attack against MDL succeeds with constant probability and requires that $ST^2/t + T^2 = \Theta(tN)$.*

trapdoor functions (cf. Section 6) in the AI-RPM, specifically, of a scheme put forth by Phan and Pointcheval [43]. Other schemes in the AI-RPM/ICM, e.g., [32, 34], can be analyzed similarly.

1.2 Methodology: Pre-Sampling

Bit-fixing oracles and pre-sampling. Unfortunately, while solving the issue of not capturing non-uniformity and preprocessing, the AI models are considerably more difficult to analyze than the traditional idealized models. From a technical point, the key difficulty is the following: *conditioned on the leaked value z*, which can depend on the entire function table of $\mathcal{O}$, many of the individual values $\mathcal{O}(x)$ are no longer random to the attacker, which ruins many of the key techniques utilized in the traditional idealized models, such as lazy sampling programmability, etc.

One way of solving the above issues is to use incompressibility arguments, as introduced by Gennaro and Trevisan [27] and successfully applied to OWPs by De *et al.* [17], to the random-oracle model by Dodis *et al.* [15], and to the GGM by Corrigan-Gibbs and Kogan [16]. Compression-based proofs generally lead to tight bounds, but are usually quite involved and, moreover, seem inapplicable to computationally secure applications. Hence, this paper, adopts the much simpler and more powerful pre-sampling approach taken recently by Coretti *et al.* [15] and dating back to Unruh [51]. The pre-sampling technique can be viewed as a general reduction from the auxiliary-input model to the so-called *bit-fixing* (BF) model, where the oracle can be arbitrarily fixed on some P coordinates, for some parameter P, but the remaining coordinates are chosen at random and independently of the fixed coordinates. Moreover, the non-uniform S-bit advice of the attacker in this model can only depend on the P fixed points, but *not* on the remaining truly random points. This makes dealing with the BF model much easier than with the AI model, as many of the traditional proof techniques can again be used, provided that one avoids the fixed coordinates.

Bit-fixing vs. auxiliary input. In order for the BF model to be useful, this work shows that any (S,T)-attack in the AI-RPM/ICM/GGM model will have similar advantage in the P-BF-RPM/ICM/GGM model for an appropriately chosen P, up to an *additive* loss of $\delta(S,T,P) \approx ST/P$. Moreover, for the special case of unpredictability applications (e.g., CRHFs, OWFs, etc.), one can set P to be (roughly) ST, and achieve a *multiplicative* loss of 2 in the exact security. This gives a general recipe for dealing with the AI models as follows: (a) prove security $\varepsilon(S,T,P)$ of the given application

in the P-BF model; (b) for unpredictability applications, set $P \approx ST$, and obtain final AI security roughly $2 \cdot \varepsilon(S, T, ST)$; (c) for general applications, choose P to minimize $\varepsilon(S, T, P) + \delta(S, T, P)$.

The proof of the above connection is based on a similar connection between the AI-ROM and BF-ROM shown by [15] (improving a weaker original bound of Unruh [51]). While borrowing a lot of tools from [15], the key difficulty is ensuring that the P-bit-fixing cipher, which "approximates" the ideal cipher conditioned on the auxiliary input z, is actually a valid cipher: the values at fixed points cannot repeat, and the remaining values are chosen at random from the "unused" values (similar issues arise for generic groups). Indeed, the proof in this paper is more involved and the resulting bounds are slightly worse than those in [15].

Using the power of pre-sampling to analyze the applications presented above, the technical bulk consists of showing the security of these applications in the easy-to-handle BF-RPM/ICM/GGM, and then using Theorem 1 to translate the resulting bound to the AI-RPM/ICM/GGM. Most of BF proofs are remarkably straightforward extensions of the traditional proofs (without auxiliary input), which is a great advantage of the pre-sampling methodology over other approaches, such as compression-based proofs.

Computational security. Note that, unlike compression-based techniques [19, 16], pre-sampling can be applied to computational reductions, by "hardwiring" the pre-sampling set of size P into the attacker breaking the computational assumption. However, this means that P cannot be made larger than the maximum allowed running time t of such an attacker. Since standard pre-sampling incurs additive cost $\Omega(ST/P)$, one cannot achieve final security better that ST/t, irrespective of the value of ε in the (t, ε)-security of the corresponding computational assumption.

Fortunately, the multiplicative variant of pre-sampling for unpredictability applications sets the list size to be roughly $P \approx ST$, which is polynomial for polynomial S and T and can be made smaller than the complexity t of the standard-model attacker for the computational assumption used. Furthermore, even though the security of public-key encryption is *not* an unpredictability application, the analysis in Section 6 shows a way to use multiplicative pre-sampling for the part that involves the reduction to a computational assumption.

1.3 Related Work

Tessaro [50] also adapted the presampling technique by Unruh to the random-*permutation* model; the corresponding bound is suboptimal, however. De *et al.* [17] study the effect of salting for inverting a *permutation* as well as for a specific pseudorandom generator based on one-way permutations.

Corrigan-Gibbs and Kogan [16], investigate the power of preprocessing in the GGM. Besides deriving security bounds for a number of important GGM applications, they also provide new attacks for DL (based on [38, 5]), MDL, and sqDDH.

The most relevant papers in the AI-ROM are those by Unruh [51], Dodis *et al.* [19], and Coretti *et al.* [15]. Chung *et al.* [14] study the effects of salting in the design of collision-resistant hash functions, and used Unruh's pre-sampling technique to argue that salting defeats pre-processing in this important case. However, they did not focus on the exact security and obtained suboptimal bounds (compared to the expected "birthday" bound obtained by [19]). Using salting to obtain non-uniform security was also advocated by Mahmoody and Mohammed [37], who used this technique for obtaining non-uniform black-box separation results.

The realization that multiplicative error is enough for unpredictability applications and can lead to non-trivial savings, is related to the work of Dodis et al. [20] in the context of improved entropy loss of key derivation schemes.

2 Capturing the Models

This section explains how the various idealized models considered in this paper—the ideal-cipher-model (ICM), the random-permutation model (RPM), and the generic-group model (GGM)—are captured. Attackers in these models are modeled as two-stage attackers $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, and applications as (single-stage) challengers C. Both $\mathcal{A}$ and C are given access to an oracle $\mathcal{O}$. Oracles $\mathcal{O}$ have two interfaces pre and main, where pre is accessible only to $\mathcal{A}_1$, which may pass auxiliary information to $\mathcal{A}_2$, and both $\mathcal{A}_2$ and C may access main. In certain scenarios it is also useful to consider an additional interface $\mathsf{main\text{-}c}$ that is only available to the challenger C.

Notation. Throughout this paper, P, K, N, and M are natural numbers and $[x] = \{0, \ldots, x-1\}$ for $x \in \mathbb{N}$. For applications in the generic-group model, $[N]$ is identified with the cyclic group $\mathbb{Z}_N$. Furthermore, denote by $\mathcal{P}_N$ the set of permutations $\pi : [N] \to [N]$ and by $\mathcal{I}_{N,M}$ the set of injections $f : [N] \to [M]$.

Oracles. An oracle $\mathcal{O}$ has two interfaces $\mathcal{O}.\mathsf{pre}$ and $\mathcal{O}.\mathsf{main}$, where $\mathcal{O}.\mathsf{pre}$ is accessible only once before any calls to $\mathcal{O}.\mathsf{main}$ are made. Some oracles may also have an additional interface $\mathcal{O}.\mathsf{main\text{-}c}$. Oracles used in this work are:

- **Auxiliary-input ideal cipher** $\mathsf{AI\text{-}IC}(K, N)$**:** Samples a random permutation $\pi_k \leftarrow \mathcal{P}_N$ for each $k \in [K]$; outputs all π_k at $\mathcal{O}.\mathsf{pre}$; answers both forward and backward queries $(k, x) \in [N]$ at $\mathcal{O}.\mathsf{main}$ by the corresponding value $\pi_k(x) \in [N]$ or $\pi_k^{-1}(x) \in [N]$, respectively.
- **Bit-fixing ideal cipher** $\mathsf{BF\text{-}IC}(P, K, N)$**:** Takes a list at $\mathcal{O}.\mathsf{pre}$ of at most P query/answer pairs (without collisions for each k); samples a random permutation $\pi_k \leftarrow \mathcal{P}_N$ consistent with said list for each k; the other interfaces behave as with $\mathsf{AI\text{-}IC}$.
- **Auxiliary-input random permutation** $\mathsf{AI\text{-}RP}(N)$**:** Special case of an auxiliary-input ideal cipher with $K = 1$.
- **Bit-fixing random permutation** $\mathsf{BF\text{-}RP}(P, N)$**:** Special case of a bit-fixing ideal cipher with $K = 1$.
- **Auxiliary-input generic group** $\mathsf{AI\text{-}GG}(N, M)$**:** Samples a random injection $\sigma \leftarrow \mathcal{I}_{N,M}$; outputs all of σ at $\mathcal{O}.\mathsf{pre}$; answers *forward* queries $x \in [N]$ at $\mathcal{O}.\mathsf{main}$ by the corresponding value $\sigma(x) \in [N]$; answers *group-operation* queries (s, s') at $\mathcal{O}.\mathsf{main}$ as follows: if $s = \sigma(x)$ and $s' = \sigma(y)$ for some x, y, the oracle replies by $\sigma(x + y)$ and by $\perp$ otherwise; answers *inverse* queries s at interface $\mathcal{O}.\mathsf{main\text{-}c}$ by returning $\sigma^{-1}(s)$ if s is in the range of F and by $\perp$ otherwise.
- **Bit-fixing generic group** $\mathsf{BF\text{-}GG}(P, N, M)$**:** Samples a random size-N subset $\mathcal{Y}$ of $[M]$ and outputs $\mathcal{Y}$ at $\mathcal{O}.\mathsf{pre}$; takes a list at $\mathcal{O}.\mathsf{pre}$ of at most P query/answer pairs without collisions and all answers in $\mathcal{Y}$; samples a random injection $\sigma \leftarrow \mathcal{I}_{N,M}$ with range $\mathcal{Y}$ and consistent with said list; the other interfaces behave as with $\mathsf{AI\text{-}GG}$.
- **Standard model:** None of the interfaces offer any functionality.

The parameters P, K, N, and M are occasionally omitted in contexts where they are of no relevance. Similarly, whenever evident from the context, explicitly specifying which interface is queried is omitted. Note that the non-auxiliary-input versions of the above oracles can be defined by not offering any functionality at $\mathcal{O}.\mathsf{pre}$. However, they are not used in this paper.

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

Definition 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* $\mathcal{O}.\mathsf{pre}$*, and outputs an* S*-bit string, and*
- $\mathcal{A}_2$*, which takes an* S*-bit auxiliary input and makes at most* T *queries to* $\mathcal{O}.\mathsf{main}$*.*

In certain contexts, if additional restrictions, captured by some parameters p, are imposed on $\mathcal{A}_2$ (e.g., time and space requirements of $\mathcal{A}_2$ or a limit on the number of queries of a particular type that $\mathcal{A}_2$ makes to a challenger it interacts with), $\mathcal{A}$ is referred to as (S, T, p)-attacker.

Applications. Let $\mathcal{O}$ be an arbitrary oracle. An application G in the $\mathcal{O}$-model is defined by specifying a challenger C, which is an oracle algorithm that has access to $\mathcal{O}.\mathsf{main}$ as well as possibly to $\mathcal{O}.\mathsf{main\text{-}c}$, interacts with 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

$$\mathrm{Succ}_{G,\mathcal{O}}(\mathcal{A}) \;:=\; \mathsf{P}\big[\mathcal{A}_2^{\mathcal{O}.\mathsf{main}}(\mathcal{A}_1^{\mathcal{O}.\mathsf{pre}}) \leftrightarrow \mathsf{C}^{\mathcal{O}.\mathsf{main},\mathcal{O}.\mathsf{main\text{-}c}} = 1\big],$$

where $\mathcal{A}_2^{\mathcal{O}.\mathsf{main}}(\mathcal{A}_1^{\mathcal{O}.\mathsf{pre}}) \leftrightarrow \mathsf{C}^{\mathcal{O}.\mathsf{main},\mathcal{O}.\mathsf{main\text{-}c}}$ denotes the bit output by C after its interaction with the attacker. This work considers two types of applications, captured by the next definition.

Definition 2. *For an* indistinguishability *application* G *in the* $\mathcal{O}$*-model, the* advantage of an attacker $\mathcal{A}$ *is defined as*

$$\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}) \;:=\; 2\left|\mathrm{Succ}_{G,\mathcal{O}}(\mathcal{A}) - \frac{1}{2}\right|.$$

For an unpredictability *application* G*, the advantage is defined as*

$$\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}) \;:=\; \mathrm{Succ}_{G,\mathcal{O}}(\mathcal{A}).$$

An application G *is said to be* $((S, T, p), \varepsilon)$-secure *in the* $\mathcal{O}$*-model if for every* (S, T, p)*-attacker* $\mathcal{A}$*,*

$$\mathrm{Adv}_{G,\mathcal{O}}(\mathcal{A}) \;\leq\; \varepsilon\,.$$

Combined query complexity. In order to state and prove Theorem 1 in Section 3, the interaction of some attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ with a challenger C in the $\mathcal{O}$-model must be "merged" into a single entity $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2)$ that interacts with oracle $\mathcal{O}$. That is, $\mathcal{D}_1^{(\cdot)} := \mathcal{A}_1^{(\cdot)}$ and $\mathcal{D}_2^{(\cdot)}(z) := \mathcal{A}_2^{(\cdot)}(z) \leftrightarrow \mathsf{C}^{(\cdot)}$ for $z \in \{0,1\}^S$. $\mathcal{D}$ is called the *combination of* $\mathcal{A}$ *and* C, and the number of queries it makes to its oracle is referred to as *the combined query complexity of* $\mathcal{A}$ *and* C. For all applications in this work, there exists an upper bound $T_G^{\mathsf{comb}} = T_G^{\mathsf{comb}}(S, T, p)$ on the combined query complexity of any attacker and the challenger.

3 Auxiliary Input vs. Bit Fixing

Since dealing with idealized models with *auxiliary input (AI)* directly is difficult, this section establishes useful connections between AI models and their *bit-fixing (BF)* counterparts, which are much less cumbersome to analyze. Specifically, for ideal ciphers, random permutations (as special cases of ideal ciphers), and generic groups, Theorem 1 below relates the advantage of attackers in a BF model to that in the corresponding AI model, allowing to translate the security of (1) *any* application at an *additive* security loss and of (2) *unpredictability* applications at a *multiplicative* security loss from the BF setting to the AI setting.

Theorem 1. *Let* $P, K, N, M \in \mathbb{N}$*,* $N \geq 16$*, and* $\gamma > 0$*. Moreover, let*

$$(\mathsf{AI}, \mathsf{BF}) \in \{(\mathsf{AI\text{-}IC}(K, N), \mathsf{BF\text{-}IC}(P, K, N)), (\mathsf{AI\text{-}GG}(N, M), \mathsf{BF\text{-}GG}(P, N, M))\} \ .$$

Then,

1. *if an application* G *is* $((S, T, p), \varepsilon')$*-secure in the* BF*-model, it is* $((S, T, p), \varepsilon)$*-secure in the* AI*-model, where*
$$\varepsilon \ \leq \ \varepsilon' + \frac{6(S + \log \gamma^{-1}) \cdot T_G^{\mathsf{comb}}}{P} + \gamma \, ;$$
2. *if an* unpredictability *application* G *is* $((S, T, p), \varepsilon')$*-secure in the* BF*-model for*
$$P \ \geq \ 6(S + \log \gamma^{-1}) \cdot T_G^{\mathsf{comb}},$$
it is $((S, T, p), \varepsilon)$*-secure in the* AI*-model for*
$$\varepsilon \ \leq \ 2\varepsilon' + \gamma \, ,$$

where T_G^{comb} *is the combined query complexity corresponding to* G*.*

Proof Outline

This section contains a brief outline of the proof of Theorem 1. The full proof of Theorem 1 is provided in Appendix A; it follows the high-level structure of the proof in [15], where a similar theorem is shown for the random-oracle model.

1. **Leaky sources vs. dense sources:** A (K, N)-cipher source X is the random variable corresponding to the function table of a cipher $F : [K] \times [N] \to [M]$. It turns out that if X has min-entropy $H_\infty(X) = K \log N! - S$ for some S, it can be replaced by a convex combination of so-called *dense* sources, which are fixed on a subset of the coordinates and have almost full min-entropy everywhere else:

   Definition 3. *A* (K, N)*-cipher source* X *is called* $(\bar{P}, 1 - \delta)$-dense *for* $\bar{P} = (P_1, \ldots, P_K) \in [N]^K$ *if it is fixed on at most* P_k *coordinates* $(k, \cdot)$ *for each* $k \in [K]$ *and if for all families* $I = \{I_k\}_{k \in [K]}$ *of subsets* I_k *of* non-fixed *coordinates* $(k, \cdot)$*,*
   $$H_\infty(X_I) \ \geq \ (1 - \delta) \sum_{k=1}^{K} \log(N - P_k)^{\underline{|I_k|}} \, ,$$
   where $a^{\underline{b}} := a!/(a-b)!$ *and* X_I *is* X *restricted to the coordinates in* I*.* X *is called* $(1-\delta)$-dense *if it is* $(0, 1 - \delta)$*-dense, and* $\bar{P}$-fixed *if it is* $(\bar{P}, 1)$*-dense.*

More concretely, one can prove that a cipher source X as above is close to a convex combination of finitely many $(\bar{P}', 1-\delta)$-dense sources for some $\bar{P} = (P_1, \ldots, P_K)$ satisfying

$$\sum_{k=1}^{K} P_k \;\approx\; \frac{S}{\delta}\,.$$

The proof is an adaptation of the proof of the corresponding lemma for random functions in [15], the difference being that the version here handles cipher sources.

2. **Dense sources vs. bit-fixing sources:** Any dense source has a corresponding bit-fixing source, which is simply a function table chosen uniformly at random from all those that agree with the P fixed positions. It turns out that a T-query distinguisher's

   - *advantage* at telling a dense source and its corresponding bit-fixing source apart can be upper bounded by approximately $T\delta$, and that its
   - *probability of outputting* 1 is at most a factor of approximately $2^{T\delta}$ larger when interacting with the bit-fixing as compared to the dense source.

   Compared to the case of random functions [15], some additional care is needed to properly handle *inverse* queries. Given the above, by setting

   $$\delta \;\approx\; \frac{S}{P}\,,$$

   one obtains additive and multiplicative errors of roughly ST/P and $2^{ST/P}$, respectively.

3. **From bit fixing to auxiliary input:** The above almost immediately implies that an application that is $((S,T),\varepsilon)$-secure in the BF-ICM is $((S,T),\varepsilon')$-secure in the AI-ICM for

   $$\varepsilon' \;\approx\; \varepsilon + \frac{ST}{P}$$

   and even

   $$\varepsilon' \;\approx\; 2\varepsilon$$

   if it is an unpredictability application, by setting $P \approx ST$. Observe that for the additive case, the final security bound in the AI-ICM is obtained by choosing P in a way that minimizes $\varepsilon(P) + ST/P$.

   For the generic-group model, the proof proceeds similarly, with two important observations:

   (a) once the range is fixed, a random injection behaves like a random permutation, which is covered by ideal ciphers as a special case;

   (b) the group-operation oracle can be implemented by three (two inverse and one forward) calls to the injection.

4 Non-Uniform Bounds for Hash Functions and Symmetric Primitives

This section derives non-uniform security bounds for a number of primitives commonly analyzed in either the random-permutation model (RPM) or the ideal-cipher model (ICM). The primitives in question can be grouped into *basic*, *sponge-based*, and *Merkle-Damgård-based* applications.

In the following, for primitives in the RPM, $\pi, \pi^{-1} : [N] \to [N]$ denote the permutation and its inverse to which $\mathsf{AI\text{-}RP}(N)$ and $\mathsf{BF\text{-}RP}(P,N)$ offer access at interface main. Similarly, for primitives in the ICM $E, E^{-1} : [K] \times [N] \to [N]$ denote the ideal cipher and its inverse to which $\mathsf{AI\text{-}IC}(K,N)$ and $\mathsf{BF\text{-}IC}(P,K,N)$ offer access at interface main (cf. Section 2).

Basic applications. The security of the following basic applications in the RPM resp. ICM is considered:

- **One-way permutation inversion (OWP):** Given $\pi(x)$ for an $x \in [N]$ chosen uniformly at random, an attacker has to find x.
- **Even-Mansour cipher (EM):** The PRF security of the Even-Mansour cipher

$$\mathsf{EM}_{\pi,s}(m) \ := \ \pi(m \oplus s_2) \oplus s_1$$

  with key $s = (s_1, s_2)$.
- **Ideal cipher as block cipher (ICM):** The PRF security of the ideal cipher used as a block cipher directly.
- **PRF security of Davies-Meyer (PRF-DM):** The PRF security of the Davies-Meyer (DM) compression function DM_E

$$\mathsf{DM}_E(h, m) \ := \ E(m, h) \oplus h$$

  when h is used as the key.
- **A collision-resistant variant of Davies-Meyer (CRHF-DM):** The collision-resistance of a salted variant

$$\mathsf{DM}_{E,a,b}(h, m) \ := \ E(m, h) + am + bh$$

  of the DM compression function, where the first-stage attacker $\mathcal{A}_1$ is unaware of the public random salt value (a, b).

Sponge-based constructions. The sponge construction is a popular hash-function design paradigm and underlies the SHA-3 hash-function standard. For $N = 2^n$, $r \leq n$, $c = n - r$, it hashes a message $m = m_1 \cdots m_\ell$ consisting of r-bit blocks m_i to $y := \mathsf{Sponge}_{\pi,\mathsf{IV}}(m)$ as follows, where $\mathsf{IV} \in \{0,1\}^c$ is a c-bit initialization vector (IV):[1]

1. Set $s_0 \leftarrow 0^r \| \mathsf{IV}$.
2. For $i = 1, \ldots, \ell$: set $s_i \leftarrow \pi(m_i \oplus s_{i-1}^{(1)} \| s_{i-1}^{(2)})$, where $s_{i-1}^{(1)}$ denotes the first r bits of s_{i-1} and $s_{i-1}^{(2)}$ the remaining c bits.
3. Output $y := s_\ell^{(1)}$.

This work considers the following applications based on the sponge paradigm:

- **Collision-resistance:** The collision resistance of the sponge construction for a randomly chosen public IV unknown to the first-stage attacker $\mathcal{A}_1$.
- **PRF security:** The PRF security of the sponge construction with the IV serving as the key.
- **MAC security:** The MAC security of the sponge construction with the IV serving as the key.

[1] To keep things simple, no padding is considered here.

Merkle-Damgård constructions with Davies-Meyer: Another widely used approach to the design of hash functions is the well-known Merkle-Damgård paradigm. For a compression function $f : [N] \times [K] \to [N]$ and an IV $\mathsf{IV} \in [N]$, a message $\overline{m} = m_1 \cdots m_\ell$ consisting of ℓ blocks $m_i \in [K]$, is hashed to $y := \mathsf{MD}_{f,\mathsf{IV}}(\overline{m})$ as follows:[2]

1. Set $h_0 \leftarrow \mathsf{IV}$.
2. For $i = 1, \ldots, \ell$: set $h_i \leftarrow f(h_{i-1}, m_i)$.
3. Output $y := h_\ell$.

This work considers the Merkle-Damgård construction with f instantiated by the Davies-Meyer compression function

$$\mathsf{DM}_E(h, m) \;:=\; E(m, h) \oplus h\,,$$

resulting in the *Merkle-Damgård-with-Davies-Meyer* function (MD-DM)

$$\mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(\overline{m}) \;:=\; \mathsf{MD}_{\mathsf{DM}_E,\mathsf{IV}}(\overline{m})\,,$$

which underlies the SHA-2 hashing standard. This work considers the following applications based on the MD-DM hash function:

- **Collision-resistance:** The collision resistance of the MD-DM construction for a randomly chosen public IV unknown to the first-stage attacker $\mathcal{A}_1$.
- **PRF security:** The PRF security of the NMAC/HMAC variants

  $$\mathsf{NMAC}_{E,k}(\overline{m}) \;:=\; \mathsf{DM}_E(k_1, \mathsf{MD\text{-}DM}_{E,k_2}(\overline{m}))$$

  of the MD-DM construction with key $k = (k_1, k_2)$.
- **MAC security:** The MAC security of the NMAC/HMAC variant of the MD-DM construction.

Discussion. The asymptotic security bounds derived for the applications listed above are summarized in Table 1. No non-uniform bounds were previously known for any of these primitives, except for OWPs, for which the same bound was derived by De *et al.* [17] using an involved, compression-based proof.

As can be seen from Table 1, a matching attack, derived by Hellman *et al.* [33], is known for OWPs. Moreover, for CRHFs based on sponges and Merkle-Damgård with Davies-Meyer, a variant of a recent attack by Coretti *et al.* [15] closely matches the derived bounds.[3] For the remaining applications, significant gaps remain: For indistinguishability applications such as BI-IC and PRFs, adapting an attack on PRGs by De *et al.* [17] results in an advantage of roughly $\sqrt{S/N}$. For the MAC applications, the best attacks are based on rainbow tables for inverting functions [33].

All security bounds are derived by following the bit-fixing approach: the security of a particular application is assessed in the *bit-fixing (BF)* RPM/ICM, and then Theorem 1 is invoked to obtain a corresponding bound in the *auxiliary-input (AI)* RPM/ICM and similarly for the random-permutation model. Deriving security bounds in the BF-ICM/RPM turns out to be quite straightforward, and all of the proofs closely follow the corresponding proofs in the ICM/RPM without

[2] As with the sponge construction, for simplicity no padding is considered here.

[3] The original attack by [15] was devised for Merkle-Damgård with a random compression function.

auxiliary input; intuitively, the only difference is that one needs to take the list $\mathcal{L}$ of the at most P input/output pairs where $\mathcal{A}_1$ fixes the random permutation or the ideal cipher.

The security proofs for one-way permutations, the ideal cipher as block cipher, the collision-resistant variant of Davies-Meyer, collision-resistance of the sponge construction, and the PRF and MAC security of NMAC/HMAC with Davies-Meyer are provided after the brief overview below. The precise definitions of the remaining applications as well as the corresponding theorems and proofs can be found in Sections B, C, and D of the appendix.

4.1 One-Way Permutations

The one-way-permutation inversion application G^{OWP} is defined via the challenger $\mathsf{C}^{\mathsf{OWP}}$ that randomly and uniformly picks an $x \in N$, passes $y := \pi(x)$ to the attacker, and outputs 1 if and only if the attacker returns x.

Theorem 2 below provides an upper bound on the success probability of any attacker in inverting π in the $\mathsf{AI\text{-}RP'}$-model, which is defined as the $\mathsf{AI\text{-}RP}$-model, except that no queries to π^{-1} are allowed. The bound matches known attacks (up to logarithmic factors) and are also shown by De et al. [17] via a more involved compression argument.

Theorem 2. *The application G^{OWP} is $\left((S,T), \tilde{O}\left(\frac{ST}{N}\right)\right)$-secure in the $\mathsf{AI\text{-}RP'}(N)$-model for $N \geq 16$.*

Proof. It suffices to show that G^{OWP} is $\left((S,T), O\left(\frac{P+T}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}RP'}(P,N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{OWP}}} = T+1$, setting $\gamma := 1/N$ and $P = 2(S+\log N)(T+1) = \tilde{O}(ST)$, and applying Theorem 1, the desired conclusion follows.

Assume $P + T < N/2$ since, otherwise, the bound of $O((P+T)/N)$ holds trivially. Let $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ be an (S,T)-attacker. Without loss of generality, assume $\mathcal{A}$ is deterministic and $\mathcal{A}_2$ makes distinct queries and always queries its output. Let $\mathcal{L} = \{(x'_1, y'_1), \ldots, (x'_P, y'_P)\}$ be the list submitted by $\mathcal{A}_1$. Recall that the challenger uniformly and randomly picks an x from $[N]$ and outputs $y := \pi(x)$. Let $x_1, \ldots, x_T$ denote the queries made by $\mathcal{A}_2$ and let $y_i := \pi(x_i)$ for $i \in [T]$ be the corresponding answers. Let $\mathcal{E}$ be the event that y appears in $\mathcal{L}$ namely $x = x'_i$ for some $i \in [P]$. Note that

$$\begin{aligned}\mathrm{Succ}_{G,\mathsf{BF\text{-}RP}}(\mathcal{A}) &\leq \mathsf{P}[\mathcal{E}] + \mathsf{P}[\exists i \in [T], x_i = x | \neg\mathcal{E}] \\ &\leq \mathsf{P}[\mathcal{E}] + \sum_{i=1}^{T} \mathsf{P}[x_i = x | \neg\mathcal{E}, x_1 \neq x, \ldots, x_{i-1} \neq x] \ .\end{aligned}$$

Observe that $\mathsf{P}[\mathcal{E}] \leq P/N$. Moreover, conditioned on $y \notin \mathcal{L}$ and any fixed choice of $(x_1, y_1), \ldots, (x_{i-1}, y_{i-1})$, x_i is a deterministic value while x is uniformly distributed over $[N] \backslash \{x_1, \ldots, x_{i-1}, x'_1, \ldots, x'_P\}$. Thus,

$$\mathsf{P}[x_i = x | \neg\mathcal{E}, x_1 \neq x, \ldots, x_{i-1} \neq x] \ \leq \ 1/(N-P-T) \ \leq \ 2/N \ ,$$

where the second inequality uses $P+T < N/2$. Therefore, $\mathrm{Succ}_{G,\mathsf{BF\text{-}RP}}(\mathcal{A}) \leq \frac{P}{N} + \frac{2T}{N} = O(\frac{P+T}{N})$. □

4.2 The Ideal Cipher as a Block Cipher

The ideal cipher can be directly used as a block cipher even in the presence of leakage. The corresponding application $G^{\mathsf{BC\text{-}IC}}$ is defined via the following challenger $\mathsf{C}^{\mathsf{BC\text{-}IC}}$: it initially chooses random bit $b \leftarrow \{0,1\}$; if $b = 0$, it picks a key $k^* \leftarrow [K]$ uniformly at random, and answers forward queries $m \in [N]$ made by $\mathcal{A}_2$ by the value $E(k^*, m)$ and inverse queries $c \in [N]$ by $E^{-1}(k^*, c)$; if

$b = 1$, forward queries m are answered by $f(m)$ and inverse queries c by $f^{-1}(c)$, where f is an independently chosen uniform random permutation; the attacker wins if and only if he correctly guesses b.

Theorem 3. *Application* $G^{\mathsf{BC\text{-}IC}}$ *is* $\left((S,T,q), \tilde{O}\left(\frac{T}{K} + \sqrt{S(T+q)/K}\right)\right)$*-secure in the* $\mathsf{AI\text{-}IC}(K,N)$*-model for* $N \geq 16$.

Proof. It suffices to show that $G^{\mathsf{BC\text{-}IC}}$ is $((S,T,q), O((T+P)/K))$-secure in the $\mathsf{BF\text{-}IC}(P,K,N)$-model since then the theorem follows by observing that $T^{\mathsf{comb}}_{G^{\mathsf{BC\text{-}IC}}} = T + q$, setting $\gamma := 1/N$ and

$$P \;:=\; \sqrt{(S + \log N)(T+q)K} \;=\; \tilde{\Theta}\left(\sqrt{S(T+q)K}\right),$$

and applying Theorem 1.

Clearly, $\mathcal{A}_2$ only has non-zero advantage in guessing bit b if it makes a (forward or inverse) query involving the key k^* chosen by the challenger or if k^* appears in one of the prefixed query/answer pairs. The latter occurs with probability at most P/K, whereas the former occurs with probability at most $T/(K-(T+P)) \leq 2T/K$, using that $T + P \leq K/2$, an assumption one can always make since, otherwise, $G^{\mathsf{BC\text{-}IC}}$ is trivially $O((T+P)/K)$-secure. □

4.3 A Collision-Resistant Variant of Davies-Meyer

The plain Davies-Meyer (DM) compression function cannot be collision-resistant against non-uniform attackers, which begs the question of if and how it can be salted to withstand non-uniform attacks. To that end, let $N = K = 2^\kappa$ for some $\kappa \in \mathbb{N}$ and interpret $[N]$ as a finite field of size N. For two values $a, b \in [N]$, let

$$\mathsf{DM}_{E,a,b}(h,m) \;:=\; E(m,h) + am + bh\ .$$

Note that for $a = 0$ and $b = 1$, $\mathsf{DM}_{E,a,b}$ is the usual DM compression function.

The application $G^{\mathsf{CRHF\text{-}DM}}$ of collision-resistance of the salted DM function is defined via the following challenger $\mathsf{C}^{\mathsf{CRHF\text{-}DM}}$: it picks two random values $a, b \in [N]$ and passes them to the attacker; the attacker wins if and only if it returns two pairs $(h,m) \neq (h',m')$ such that $\mathsf{DM}_{E,a,b}(h,m) = \mathsf{DM}_{E,a,b}(h',m')$.

Theorem 4. $G^{\mathsf{CRHF\text{-}DM}}$ *is* $\left((S,T), \tilde{O}\left(\frac{(ST)^2}{N}\right)\right)$*-secure in the* $\mathsf{AI\text{-}IC}(N,N)$*-model for* $N \geq 16$.

Proof. At the cost of at most 2 additional queries to E, assume that the pairs (h,m) and (h',m') output by $\mathcal{A}_2$ are such that $\mathcal{A}_2$ has queried its oracle E on all points $\mathsf{DM}_{E,a,b}$ would query E when evaluated on (h,m) and (h',m'). It suffices to show that $G^{\mathsf{CRHF\text{-}DM}}$ is $\left((S,T), O\left(\frac{T^2}{N} + \frac{P(P+T)}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}IC}(P,N,N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{CRHF\text{-}DM}}} = T + 2$, setting $\gamma := 1/N$ and $P = 2(S + \log N)(T+2) = \tilde{O}(ST)$, and applying Theorem 1, the desired conclusion follows.

Set $T' := T+2$ and consider an interaction of $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ and $\mathsf{C}^{\mathsf{CRHF\text{-}DM}}$ in the $\mathsf{BF\text{-}IC}(P,N,N)$-model. Denote by $((k'_i, x'_i), y'_i)$ for $i = 1, \ldots, P$ the query/answer pairs prefixed by $\mathcal{A}_1$ and by $((k_i, x_i), y_i)$ for $i = 1, \ldots, T'$ the queries $\mathcal{A}_2$ makes to E. Let $\mathcal{E}$ be the event that there exists *no* collision among the prefixed values, i.e., there exist no $i \neq j$ such that

$$E(k'_i, x_i) + ak'_i + bh'_j \;=\; E(k'_j, x_j) + ak'_j + bh'_j \tag{1}$$

and that $b \neq 0$. For any fixed $i \neq j$, consider two cases:

1. $k_i \neq k_j$: in this case, the two pairs cause a collision if and only if
$$a \;=\; \frac{(y'_j - y'_i) - b(x'_i - x'_j)}{k'_i - k'_j} \;,$$
which happens with probability at most $1/N$.
2. $k_i = k_j$: in this case, $x'_i \neq x'_j$, and the two pairs cause a collision if and only if $b = (y'_j - y'_i)/(x'_i - x'_j)$, which happens with probability at most $1/N$ as well.

Summarizing, $\mathsf{P}[\neg\mathcal{E}] \leq (P^2+1)/N = O\left(P^2/N\right)$.

Moving to queries made by $\mathcal{A}_2$, let $\mathcal{E}'_i$ be the event that after the i^{th} query made by $\mathcal{A}_2$, there exists no collision between any query pair and a prefixed pair or among the query pairs themselves; the corresponding conditions are analogous to (1). Consider the probability $\mathsf{P}[\neg\mathcal{E}'_i|\mathcal{E}'_{i-1},\mathcal{E}]$. If the i^{th} query is a forward query, then a collision occurs only if $y_i = a(k_i - k_j) + b(x_i - x_j) + y_j$ for some $j < i$ or if the analogous condition holds for a collision with a prefixed pair and some $j \in \{1,\ldots,P\}$; if the i^{th} query is a backward query, then a collision occurs only if
$$x_i \;=\; \frac{a(k_i - k_j) - (y_j - y_i)}{b}$$
for some $j < i$ or if the analogous condition holds for a collision with a prefixed pair and some $j \in \{1,\ldots,P\}$ (using that $b \neq 0$). In either case,
$$\mathsf{P}[\neg\mathcal{E}'_i|\mathcal{E}'_{i-1},\mathcal{E}] \;\leq\; \frac{(i-1)+P}{N-(T'+P)} \;\leq\; \frac{2((i-1)+P)}{N} \;,$$
using that $T' + P \leq N/2$, an assumption on may always make since, otherwise, the desired bound holds trivially. Summarizing, setting $\mathcal{E}' := \mathcal{E}'_{T'}$,
$$\begin{aligned}\mathsf{P}[\neg\mathcal{E}'|\mathcal{E}] \;=\; \mathsf{P}[\neg\mathcal{E}'_{T'}|\mathcal{E}] \;&\leq\; \sum_{i=1}^{T'} \mathsf{P}[\neg\mathcal{E}'_i|\mathcal{E}'_{i-1},\mathcal{E}] \\ &\leq\; \sum_{i=1}^{T'} \frac{2((i-1)+P)}{N} \\ &=\; O\left(\frac{T^2}{N} + \frac{TP}{N}\right) \;.\end{aligned}$$

Clearly, $\mathcal{A}_2$ only wins if $\mathcal{E}$ or $\mathcal{E}'$ occurs, and hence the overall security in the $\mathsf{BF\text{-}IC}(P,N,N)$-model is $O\left(\frac{T^2}{N} + \frac{P(P+T)}{N}\right)$. □

4.4 CRHFs from Unkeyed Sponges

The application $G^{\mathsf{CRHF\text{-}S}}$ of collision resistance for the sponge construction is defined via the following challenger $\mathsf{C}^{\mathsf{CRHF\text{-}S}}$: it picks an initialization vector $\mathsf{IV} \leftarrow \{0,1\}^c$ uniformly at random, passes it to the attacker, and outputs 1 if and only if the attacker returns two messages $m \neq m'$ such that $\mathsf{Sponge}_{\pi,\mathsf{IV}}(m) = \mathsf{Sponge}_{\pi,\mathsf{IV}}(m')$.

The following theorem provides an upper bound on the probability that an (S,T,ℓ)-attacker finds a collision of the sponge construction in the AI-RPM, where ℓ is an upper bound on the lengths of the messages m and m' the attacker submits to the challenger. The proof follows the approach by Bertoni *et al.* [6].

Theorem 5. *Application* $G^{\mathsf{CRHF\text{-}S}}$ *is* $\left((S,T,\ell),\tilde{O}\left(\frac{S(T+\ell)^2}{2^c}+\frac{(T+\ell)^2}{2^r}\right)\right)$*-secure in the* $\mathsf{AI\text{-}RP}(N)$*-model, for* $N=2^n=2^{r+c}\geq 16$.

Node graphs. A useful formalism for security proofs of sponge-based constructions is that of node and supernode graphs, as introduced by Bertoni *et al.* [6]. For a permutation $\pi:\{0,1\}^n\to\{0,1\}^n$, consider the following (directed) *node graph* $G_\pi=(V,E)$ with $V=\{0,1\}^r\times\{0,1\}^c=\{0,1\}^n$ and $E=\{(s,t)\mid \pi(s)=t\}$. Moreover, let $G'_\pi=(V',E')$ be the (directed) *supernode graph*, with $V'=\{0,1\}^c$ and $(s^{(2)},t^{(2)})\in E'$ iff $((s^{(1)},s^{(2)}),(t^{(1)},t^{(2)}))\in E$ for some $s^{(1)},t^{(1)}\in\{0,1\}^r$. Observe that the value of $\mathsf{Sponge}_{\pi,\mathsf{IV}}(m)$ for an ℓ-block message $m=m_1\cdots m_\ell$ is obtained by starting at $s_0:=(0^r,\mathsf{IV})\in\{0,1\}^n$ in G_π, moving to $s_i\leftarrow\pi(m_i\oplus s_{i-1}^{(1)}\|s_{i-1}^{(2)})$ for $i=1,\ldots,\ell$, and outputting $s_\ell^{(1)}$. In other words, in the supernode graph, m corresponds to a path of length ℓ starting at node IV and ending at $s_\ell^{(2)}$, and $s_1^{(1)},\ldots,s_\ell^{(1)}\in\{0,1\}^r$ are the values that *appear* on that path.

Proof. At the cost of at most 2ℓ additional queries to π, assume that the messages m and m' output by $\mathcal{A}_2$ are such that $\mathcal{A}_2$ has queried its oracle π on all points $\mathsf{Sponge}_{\pi,\mathsf{IV}}(\cdot)$ would query π when evaluated on m and m'.

It suffices to show that $G^{\mathsf{CRHF\text{-}S}}$ is $\left((S,T,\ell),O\left(\frac{(T+\ell)^2}{2^r}+\frac{(T+\ell)^2+(T+\ell)P)}{2^c}\right)\right)$-secure in the $\mathsf{BF\text{-}RP}(P,N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{CRHF\text{-}S}}}=T+2\ell$, setting $\gamma:=1/N$ and $P:=2(S+\log N)(T+\ell)=\tilde{O}\left(S(T+\ell)\right)$, and applying Theorem 1, the desired conclusion follows.

Consider now an interaction of $\mathcal{A}_2$ with $\mathsf{C}^{\mathsf{CRHF\text{-}S}}$ and incrementally build the node and supernode graphs (as defined above), adding edges when $\mathcal{A}_2$ makes the corresponding (forward or inverse) query to π, and starting with the edges that correspond to the at most P prefixed query/answer pairs.

Let $\mathcal{E}_{\mathsf{coll}}$ be the event that a (valid) collision occurs. Clearly, this happens if and only if there exists a value $s^{(1)}\in\{0,1\}^r$ that appears as the last value on two different paths from IV. Let $\mathcal{E}_{\mathsf{path,i}}$ be the event that after the i^{th} query to π, there is a unique path from IV to any node in the supernode graph and that no prefixed supernode is reachable from IV.

Observe that when $\mathcal{E}_{\mathsf{path}}:=\mathcal{E}_{\mathsf{path,T+2\ell}}$ occurs, the values that appear on these paths are uniformly random and independent since every node inside a supernode has the same probability of being chosen. Hence,

$$\mathsf{P}[\mathcal{E}_{\mathsf{coll}}|\mathcal{E}_{\mathsf{path}}]\ \leq\ \binom{T+2\ell}{2}\cdot 2^{-r}\ =\ O\left(\frac{(T+\ell)^2}{2^r}\right).$$

Moreover,

$$\mathsf{P}[\neg\mathcal{E}_{\mathsf{path,i}}|\mathcal{E}_{\mathsf{path,i-1}}]\ \leq\ \frac{(i+P)\cdot 2^r}{2^{r+c}-(i-1+P)}\ \leq\ \frac{i+P}{2^c-(T+2\ell+P)/2^r}\ \leq\ \frac{i+P}{2^{c-1}}$$

if the i^{th} query is a forward query, and

$$\mathsf{P}[\neg\mathcal{E}_{\mathsf{path,i}}|\mathcal{E}_{\mathsf{path,i-1}}]\ \leq\ \frac{i\cdot 2^r}{2^{r+c}-(i-1+P)}\ \leq\ \frac{i}{2^{c-1}}$$

if the i^{th} query is an inverse query, using that $T+2\ell+P\leq N/2$, an assumption one may always

make since, otherwise, the lemma holds trivially. Letting $T' := T + 2\ell$,

$$\begin{aligned}
\mathsf{P}[\neg\mathcal{E}_{\mathsf{path}}] \;=\; \mathsf{P}[\neg\mathcal{E}_{\mathsf{path},\mathsf{T}'}] \;&\leq\; \mathsf{P}[\neg\mathcal{E}_{\mathsf{path},\mathsf{T}'}|\mathcal{E}_{\mathsf{path},\mathsf{T}'-1}] + \mathsf{P}[\neg\mathcal{E}_{\mathsf{path},\mathsf{T}'-1}] \\
&\leq\; \sum_{i=1}^{T'} \mathsf{P}[\neg\mathcal{E}_{\mathsf{path},\mathsf{i}}|\mathcal{E}_{\mathsf{path},\mathsf{i}-1}] + \mathsf{P}[\mathcal{E}_{\mathsf{path},0}] \\
&\leq\; \sum_{i=0}^{T'} \frac{(i+P)}{2^{c-1}} \;=\; O\left(\frac{(T+\ell)(T+\ell+P)}{2^{-c}}\right),
\end{aligned}$$

observing that $\mathsf{P}[\mathcal{E}_{\mathsf{path},0}] \leq \frac{P}{2^c}$, the probability that a node inside supernode IV is prefixed. □

4.5 PRFs via NMAC with Davies-Meyer

For simplicity, let $K = N$. Recall that the NMAC construction using the Davies-Meyer compression function is defined as

$$\mathsf{NMAC}_{E,k}(\overline{m}) \;:=\; \mathsf{DM}_E(k_1, \mathsf{MD\text{-}DM}_{E,k_2}(\overline{m}))$$

where $k = (k_1, k_2)$.

The application $G^{\mathsf{PRF\text{-}MD\text{-}N}}$ of PRF security for NMAC is defined via the following challenger $\mathsf{C}^{\mathsf{PRF\text{-}MD\text{-}N}}$: it picks a random bit $b \leftarrow \{0,1\}$ and a key $k \leftarrow [N]$; when the attacker queries a message $m = m_1 \cdots m_\ell$ consisting of blocks m_i, if $b = 0$, the challenger answers by $\mathsf{NMAC}_{E,k}(\overline{m})$, and, if $b = 1$, the challenger answers by a value chosen uniformly at random for each $\overline{m}$. The attacker wins, if and only if he correctly guesses the bit b.

The following theorem provides an upper bound on the advantage of an (S, T, q, ℓ)-attacker in distinguishing the sponge construction from a random function in the AI-ICM, where q is an upper bound on the number of messages $\overline{m}$ the attacker submits to the challenger and ℓ is an upper bound on the length of those messages.

Theorem 6. *$G^{\mathsf{PRF\text{-}MD\text{-}N}}$ is $\left((S,T,q,\ell), \tilde{O}\left(\frac{Tq^2\ell}{N} + \sqrt{\frac{S(T+\ell q)q^2\ell}{N}}\right)\right)$-secure in the $\mathsf{AI\text{-}IC}(N,N)$-model, for $N \geq 16$.*

Proof. Follows from Lemma 7 by observing that $T^{\mathsf{comb}}_{G^{\mathsf{PRF\text{-}MD\text{-}N}}} = T + q\ell$, setting $\gamma := 1/N$ and $P := \sqrt{\frac{S(T+q\ell)N}{q^2\ell}}$, and applying Theorem 1. □

Lemma 7. *For any $P, N \in \mathbb{N}$, $G^{\mathsf{PRF\text{-}MD\text{-}N}}$ is $\left((S,T,q,\ell), O\left(q^2\ell\frac{T+P}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}IC}(P,N,N)$-model.*

The proof of Lemma 7 uses the fact the Merkle-Damgård construction with the DM function is *almost-universal* in the BF-ICM; this property is captured by the application $G^{\mathsf{AU\text{-}MD}}$ defined by the following challenger $\mathsf{C}^{\mathsf{AU\text{-}MD}}$: It expects $\mathcal{A}_2$ to submit two messages $\overline{m}$ and $\overline{m}'$. Then, it picks a random key k. The attacker wins if $\mathsf{MD\text{-}DM}_{E,k}(\overline{m}) = \mathsf{MD\text{-}DM}_{E,k}(\overline{m}')$.

The proof of almost-universality uses Lemma 24, which states that the DM function is a PRF when keyed by h .

Lemma 8. *For any $P, N \in \mathbb{N}$, $G^{\mathsf{AU\text{-}MD}}$ is $\left((S,T,q,\ell), O\left(\ell\frac{T+P}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}IC}(P,N,N)$-model.*

Proof (sketch). Consider a sequence of ℓ hybrid experiments, where in the i^{th} hybrid, instead of evaluating $\mathsf{MD\text{-}DM}_{E,k}(\overline{m})$ for $\overline{m} = m_1 \cdots m_\ell$, the challenger computes $\mathsf{MD\text{-}DM}_{E,k'}(m_{i+1} \cdots m_\ell)$, where $k' \leftarrow f(m_1 \cdots m_i)$ for a uniformly random function $f : [N]^i \to N$. By the PRF security of the Davies-Meyer function (cf. Lemma 24), the distance between successive hybrids is at most $8(T+P)/N$. Moreover, in the last hybrid, the success probability of $\mathcal{A}_2$ is at most $1/N$. □

Proof (of Lemma 7, sketch). Using the PRF security of the Davies-Meyer (DM) function (cf. Lemma 24), it suffices to show security in the hybrid experiment in which the outer DM evaluation is replaced by a uniform random function f. In this hybrid experiment, $\mathcal{A}_2$ only has non-zero advantage in guessing bit b if two of its q queries to the challenger cause a collision right before f. Let ε be the probability that this event occurs.

Consider the following attacker $\mathcal{A}' := (\mathcal{A}_1, \mathcal{A}'_2)$ against the $\mathsf{C}^{\mathsf{AU\text{-}MD}}$: $\mathcal{A}'_2$ runs $\mathcal{A}_2$ internally, forwarding its oracle queries to and back from its own oracle, and answering every query $\mathcal{A}_2$ would make to its challenger by a fresh uniformly random value. Once $\mathcal{A}_2$ terminates, $\mathcal{A}'_2$ picks a pair of queries made by $\mathcal{A}_2$ uniformly at random and subits it to its own challenger. It is easily seen that the advantage of $\mathcal{A}'_2$ is at least ε/q^2. Therefore, the final PRF security of NMAC is $q^2\ell(T+P)/N$. □

4.6 MACs via NMAC with Davies-Meyer

The application $G^{\mathsf{MAC\text{-}MD\text{-}N}}$ of MAC security of the NMAC construction is defined via the following challenger $\mathsf{C}^{\mathsf{MAC\text{-}MD\text{-}N}}$: it initially picks a random key $k \leftarrow [N]$; when the attacker queries a message $\overline{m} = m_1 \cdots m_\ell$ consisting of blocks m_i, the challenger answers by $\mathsf{MD\text{-}DM}_{\mathcal{O},k}(\overline{m})$. The attacker wins if he submits a pair $(\overline{m}, y)$ with $\mathsf{MD\text{-}DM}_{\mathcal{O},k}(\overline{m}) = y$ for a previously unqueried $\overline{m}$.

Theorem 9. *$G^{\mathsf{MAC\text{-}MD\text{-}N}}$ is $\left((S,T,q,\ell), \tilde{O}\left(q^2\ell\frac{S(T+q\ell)}{N}\right)\right)$-secure in the $\mathsf{AI\text{-}IC}(N)$-model, for $N \geq 16$.*

Proof. It suffices to show that $G^{\mathsf{MAC\text{-}MD\text{-}N}}$ is $\left((S,T,q,\ell), O\left(q^2\ell\frac{T+P}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}IC}(P,N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{MAC\text{-}MD\text{-}N}}} = T + q\ell$, setting $\gamma := 1/N$ and $P = 2(S + \log N)(T + q\ell) = \tilde{O}\left(S(T+q\ell)\right)$ and applying Theorem 1, the desired conclusion follows.

The bound in the $\mathsf{BF\text{-}IC}(P,N)$-model follows immediately from Lemma 7 and the fact that with a truly random function, the adversary's success probability at breaking the MAC is at most q/N. □

4.7 Extensions to HMAC

Recall that, for simplicity, $K = N$. The HMAC construction using the Davies-Meyer compression function is defined as

$$\mathsf{HMAC}_{E,k}(\overline{m}) \;:=\; \mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(k \oplus \mathsf{opad}, \mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(k \oplus \mathsf{ipad}, \overline{m}))\ ,$$

where $\mathsf{IV} \in [N]$ is some fixed initialization vector. As usual, results for NMAC carry over to HMAC, even in the presence of leakage about the ideal cipher. More precisely, the HMAC construction can be seen as a special case of the NMAC by observing that

$$\mathsf{HMAC}_{E,k}(\overline{m}) \;=\; \mathsf{NMAC}_{E,k_1,k_2}(\overline{m})$$

for $k_1 = E(k \oplus \mathsf{ipad}, \mathsf{IV}) \oplus \mathsf{IV}$ and $k_2 = E(k \oplus \mathsf{opad}, \mathsf{IV}) \oplus \mathsf{IV}$. Hence, in the $\mathsf{BF\text{-}IC}$-model, unless $(k \oplus \mathsf{opad}, \mathsf{IV})$ or $(k \oplus \mathsf{ipad}, \mathsf{IV})$ are prefixed by $\mathcal{A}_1$ or queried by $\mathcal{A}_2$, which happens with probability $O\left((T+P)/N\right)$, the NMAC analysis applies.

5 The Generic-Group Model with Preprocessing

This section analyzes the hardness of various problems in the generic-group model (GGM) with preprocessing. Specifically, the following applications are considered, where $N \in \mathbb{N}$ is an arbitrary prime and σ the random injection used in the GGM:

- **Discrete-logarithm problem (DL):** Given $\sigma(x)$ for a uniformly random $x \in [N]$, find x.
- **Multiple-discrete-logarithms problem (MDL):** Given $(\sigma(x_1), \ldots, \sigma(x_t))$ for uniformly random and independent $x_i \in [N]$, find $(x_1, \ldots, x_t)$.
- **Computational Diffie-Hellman problem (CDH):** Given $(\sigma(x), \sigma(y))$ for uniformly random and independent $x, y \in [N]$, find xy.
- **Decisional Diffie-Hellman problem (DDH):** Distinguish $(\sigma(x), \sigma(y), \sigma(xy))$ from $(\sigma(x), \sigma(y), \sigma(z))$ for uniformly random and independent $x, y, z \in [N]$.
- **Square decisional Diffie-Hellman problem (sqDDH):** Distinguish $(\sigma(x), \sigma(x^2))$ from $(\sigma(x), \sigma(y))$ for uniformly random and independent $x, y \in [N]$.
- **One-more-discrete-logarithm problem (OM-DL):** Given access to an oracle creating DL challenges $\sigma(x_i)$, for uniformly random and independent $x_i \in [N]$, as well as a DL oracle, make t queries to the challenge oracle and at most $t-1$ queries to the DL oracle, and solve *all* t challenges, i.e., find $(x_1, \ldots, x_t)$.
- **Knowledge-of-exponent assumption (KEA):** The KEA assumption states that if an attacker $\mathcal{A}$ is given $\sigma(x)$, for $x \in [N]$ chosen uniformly at random, and outputs A and $\hat{A}$ with $A = \sigma(a)$ and $\hat{A} = \sigma(ax)$, then it must know discrete logarithm a of A. This is formalized by requiring that for every $\mathcal{A}$ there exist an *extractor* $\mathcal{X}_{\mathcal{A}}$ that is run on the same random coins as $\mathcal{A}$ and must output the value a.

The asymptotic security bounds derived for the above applications are summarized in Table 2. The bounds for DL, MDL, CDH, DDH, and sqDDH match previously known bounds from [16, 38, 5]; they are tight in that there is a matching attack, except for the DDH problem, for which, remarkably, closing the gap remains an open problem. The bounds for OM-DL and KEA are new.

Note that all bounds with preprocessing are considerably worse than those without. For example, in the classical GGM, DL is secure up to roughly $N^{1/2}$ queries, whereas it becomes insecure for $S = T = N^{1/3}$ in the AI-GGM.

All security bounds are derived by following the bit-fixing approach: the security of a particular application is assessed in the *bit-fixing (BF)* GGM, and then Theorem 1 is invoked to obtain a corresponding bound in the *auxiliary-input (AI)* GGM. This approach features great simplicity since deriving security bounds in the BF-GGM turns out to be remarkably straightforward, and all of the proofs closely follow the original proofs in the classical GGM without preprocessing; the only difference is that one needs to take the list $\mathcal{L}$ of the at most P input/output pairs where $\mathcal{A}_1$ fixes σ into account.

Besides simplicity, another advantage of the bit-fixing methodology is applicability: using bit-fixing, in addition to recovering all of the bounds obtained in [16] via much more involved compression-based proofs, one also easily derives bounds for applications that may be challenging to derive using compression-based proofs, such as, e.g., the knowledge-of-exponent assumption.

As representative examples, the proofs for the DL problem and the KEA are provided below. Readers familiar with the original proofs by Shoup [48] for DL and by Abe and Fehr [1] and Dent [18] for the KEA may immediately observe the similarity. The precise definitions of the remaining applications as well as the corresponding theorems and proofs can be found in Section E of the appendix .

5.1 Discrete Logarithms

The discrete-logarithm application G^{DL} is defined via the challenger C^{DL} that randomly and uniformly picks an $x \in [N]$, passes $\sigma(x)$ to the attacker, and outputs 1 if and only if the attacker returns x.

Theorem 10 below provides an upper bound on the success probability of any attacker at computing discrete logarithms in the AI-GGM. The bound is matched by the attack of Mihalcik [38] and Bernstein and Lange [5]; a variation of said attack has recently also been presented by Corrigan-Gibbs and Kogan [16].

Theorem 10. *G^{DL} is $((S,T),\varepsilon)$-secure in the $\mathsf{AI\text{-}GG}(N,M)$-model for any prime $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\frac{ST^2}{N} + \frac{T^2}{N}\right).$$

Proof. It suffices to show that the application G^{DL} is $\left((S,T), O\left(\frac{TP+T^2}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}GG}(P, N, M)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{DL}}} = T+1$, setting $\gamma := 1/N$ and $P = 6(S+\log N)(T+1) = \tilde{O}(ST)$, and applying the second part of Theorem 1, the desired conclusion follows.

Consider now the interaction of $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ with C^{DL} in the $\mathsf{BF\text{-}GG}$-model. Recall that the $\mathsf{BF\text{-}GG}$-oracle outputs the range $\mathcal{Y}$ of the underlying random injection σ to $\mathcal{A}_1$ via interface pre. Condition on a particular realization of this set for the remainder of the proof.

Define the following hybrid experiment involving $\mathcal{A}_1$ and $\mathcal{A}_2$:

- For each of the at most P query/answer pairs (a', s') where $\mathcal{A}_1$ fixes σ, define a (constant) polynomial $v(X) := a'$ and store the pair (v, s').
- To create the challenge, choose a value s^* uniformly at random from all unused values in $\mathcal{Y}$, define the polynomial $u^*(X) := X$, and store (u^*, s^*).
- A forward query a by $\mathcal{A}_2$ to $\mathsf{BF\text{-}GG}$ is answered as follows: define the (constant) polynomial $u(X) := a$, choose a value s uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u, s), and return s.
- A group-operation query (s_1, s_2) by $\mathcal{A}_2$ is answered as follows:
  - If s_1 or s_2 is not in $\mathcal{Y}$, return $\perp$.
  - If s_1 has not been recorded, choose a random unused $a \in [N]$, define the (constant) polynomial $u(X) := a$, and store the pair (u, a). Proceed similarly if s_2 has not been recorded. Go to the next item.
  - Let u_1 and u_2 be the polynomials recorded with s_1 and s_2, respectively. If, for $u' := u_1 + u_2$, a pair (u', s') has been recorded, return s'. Otherwise, choose a value s' uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u', s'), and return s'.

- When $\mathcal{A}_2$ outputs a value x', pick a value $x \in [N]$ uniformly at random and output 1 if and only if $x' = x$.

Observe that the hybrid experiment only differs from the original one if for a group-operation query (s_1, s_2), $u'(x) = v(x)$ for some recorded v or $u'(x) = u(x)$ for some recorded u—and similarly for the polynomial u^* corresponding to the challenge. Since in the hybrid experiment, x is chosen uniformly at random *at the end* of the execution, the probability of this event is at most $((T+1)P+(T+1)^2)/N$ by the Schwartz-Zippel Lemma (cf. Lemma 37 in Section G of the appendix) and a union bound. Moreover, in the hybrid experiment, the probability that $x' = x$ is $1/N$. The theorem follows. □

5.2 Knowledge-of-Exponent Assumption

Informally, the knowledge-of-exponent assumption (KEA) states that if an attacker $\mathcal{A}$ is given (h, h^x), for a generator h of a cyclic group of order N and $x \in [N]$ chosen uniformly at random, and outputs group elements A and $\hat{A}$ with $\hat{A} = A^x$, then it must know discrete logarithm a of A. This is formalized by requiring that for every $\mathcal{A}$ there exist an *extractor* $\mathcal{X}_\mathcal{A}$ that is run on the same random coins as $\mathcal{A}$ and must output the value a.

The above is captured in the GGM by considering the following experiment $\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_\mathcal{A}}$ parameterized by an attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, an extractor $\mathcal{X}_\mathcal{A}$, and an oracle $\mathcal{O} \in \{\mathsf{AI\text{-}GG}(N, M), \mathsf{BF\text{-}GG}(N, M)\}$:

1. Run $\mathcal{A}_1$ to obtain $z \leftarrow \mathcal{A}_1^{\mathcal{O}}$.
2. Choose $x \in [N]$ uniformly at random, let $y \leftarrow \sigma(x)$, pick random coins ρ, and run
   (a) $\mathcal{A}_2$ to get $(A, \hat{A}) \leftarrow \mathcal{A}_2(z, y; \rho)$, and
   (b) $\mathcal{X}_\mathcal{A}$ to get $a \leftarrow \mathcal{X}_\mathcal{A}(z, y; \rho)$.
3. Output 1 if and only if $A = \sigma(a')$ and $\hat{A} = \sigma(a'x)$ for some a', but $a \neq a'$.

The KEA says that for every attacker $\mathcal{A}$ there exists an extractor $\mathcal{X}_\mathcal{A}$ such that the probability of the above experiment outputting 1 is negligible. The following theorem is equivalent to saying that the KEA holds in the AI-GGM.

Theorem 11. *For every attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, there exists an extractor $\mathcal{X}_\mathcal{A}$ such that*

$$\mathsf{P}[\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_\mathcal{A}} = 1] \;\leq\; \tilde{O}\left(\frac{ST^2}{N}\right).$$

Sketch. The extractor $\mathcal{X}_\mathcal{A}$ internally runs $\mathcal{A}_2$ on the inputs received and keeps track of $\mathcal{A}_2$'s oracle queries using polynomials as in the proof of Theorem 10. If at the end the polynomials u_A and $u_{\hat{A}}$ corresponding to $\mathcal{A}_2$'s outputs $(A, \hat{A})$ have the form $u_A(X) = a$ and $u_{\hat{A}}(X) = aX$, then $\mathcal{X}_\mathcal{A}$ outputs a and otherwise $\perp$.

Observe that if the experiment outputs 1, then

- $u_{\hat{A}} \neq X \cdot u_A$ since $\mathcal{A}_2$ only creates polynomials of degree at most 1, but
- $u_{\hat{A}}(x) = x \cdot u_A(x)$ for the challenge x.

Hence, the extractor only fails if at least two of the polynomials involved (including $u_{\hat{A}}$ and $X \cdot u_A$) collide on x, which is already analyzed in the proof of Theorem 10.

The experiment $\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_\mathcal{A}}$ defining KEA does not exactly match the syntax of challenger and attacker to which Theorem 1 caters, but it is easily checked that the corresponding proof can be adapted to fit $\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_\mathcal{A}}$. □

6 Computationally Secure Applications

A main advantage of the pre-sampling methodology over other approaches (such as compression) to dealing with auxiliary-input in idealized models is that it also applies to applications that rely on computational hardness assumptions. To illustrate this fact, this section considers a public-key encryption scheme based on trapdoor functions (cf. Section F for a definition)by Phan and Pointcheval [43] in the auxiliary-input random-permutation model (AI-RPM). Other schemes in the AI-RPM/ICM, e.g., [32, 34], can be analyzed similarly.

FDP encryption. Let F be a trapdoor family (TDF) generator. Full-domain permutation (FDP) encryption in the random-permutation model with oracle $\mathcal{O}$ is defined as follows:

- **Key generation:** Run the TDF generator to obtain $(f, f^{-1}) \leftarrow F$, where $f, f^{-1} : [N] \to [N]$. Set the public key $\mathsf{pk} := f$ and the secret key $\mathsf{sk} := f^{-1}$.
- **Encryption:** To encrypt a message m with randomness r and public key $\mathsf{pk} = f$, compute $\tilde{y} \leftarrow f(y)$ for $y \leftarrow \mathcal{O}(m\|r)$ and output $c = \tilde{y}$.
- **Decryption:** To decrypt a ciphertext $c = y$ with secret key $\mathsf{sk} = f^{-1}$, compute $m\|r \leftarrow \mathcal{O}^{-1}(f^{-1}(y))$ and output m.

The following theorem relates to the CPA security (cf. Section F for a definition) of FDP encryption in the AI-RPM.

Theorem 12. *Let Π be FDP encryption with F. If $G^{\mathsf{TDF},F}$ is $((S', *, t', s'), \varepsilon')$-secure, then, for any $T \in \mathbb{N}$, $G^{\mathsf{PKE},\Pi}$ is $((S, T, t, s), \varepsilon)$-secure in the $\mathsf{AI\text{-}RP}(N, N)$-model, where*

$$\varepsilon \;=\; \tilde{O}\left(\varepsilon' + \sqrt{\frac{ST}{2^{\rho}}}\right)$$

and $S = S' - \tilde{O}(ST)$, $t = t' - \tilde{O}(t_{\mathsf{tdf}} \cdot T)$, and $s = s' - \tilde{O}(ST)$, where t_{tdf} is the time required to evaluate the TDF.

The straightforward approach to proving the security of FDP encryption in the AI-RPM would be to analyze the scheme in the BF-RPM with list size P and then use the general part of Theorem 1 to obtain a bound in the AI-RPM. However, such an approach, due to the additive error in the order of ST/P would require a very large list and therefore make the reduction to TDF security extremely loose.

Instead, the actual proof, which is sketched in Appendix F, follows the same high-level structure as that of TDF encryption in the AI-ROM, analyzed in [15]:

1. It first considers a hybrid experiment that is only distinguishable from the original CPA experiment if the attacker queries a particular value to the random permutation. To bound the probability of this event occurring, the proof moves to the BF-RPM and the analysis there—which involves the reduction to TDF security—is carried back to the AI-RPM via the *unpredictability* part of Theorem 1. This allows the list size to remain a moderate $P' \approx ST$ and hence for a tight reduction.
2. To analyze the advantage of the attacker in the hybrid experiment, the BF-RPM is used again, but using the general part of Theorem 1, which requires a larger list size P. However, since this second step involves no reduction to TDF security and is purely information-theoretic, this does not pose a problem.

Acknowledgments

The authors thank Dan Boneh, Henry Corrigan-Gibbs, and Dmitry Kogan for valuable discussions on pre-processing in generic-group models.

References

[1] Masayuki Abe and Serge Fehr. Perfect NIZK with adaptive soundness. In *Theory of Cryptography, 4th Theory of Cryptography Conference, TCC 2007, Amsterdam, The Netherlands, February 21-24, 2007, Proceedings*, pages 118–136, 2007.

[2] Elena Andreeva, Andrey Bogdanov, Yevgeniy Dodis, Bart Mennink, and John P. Steinberger. On the indifferentiability of key-alternating ciphers. In *Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, pages 531–550, 2013.

[3] Mihir Bellare, Alexandra Boldyreva, and Adriana Palacio. An uninstantiable random-oracle-model scheme for a hybrid-encryption problem. In *Advances in Cryptology - EUROCRYPT 2004, International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004, Proceedings*, pages 171–188, 2004.

[4] Mihir Bellare, Ran Canetti, and Hugo Krawczyk. Keying hash functions for message authentication. In *Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings*, pages 1–15, 1996.

[5] Daniel J. Bernstein and Tanja Lange. Non-uniform cracks in the concrete: The power of free precomputation. In *Advances in Cryptology - ASIACRYPT 2013 - 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part II*, pages 321–340, 2013.

[6] Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. On the indifferentiability of the sponge construction. In *Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, pages 181–197, 2008.

[7] Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. On the security of the keyed sponge construction. In *Symmetric Key Encryption Workshop (SKEW)*, 2011.

[8] John Black. The ideal-cipher model, revisited: An uninstantiable blockcipher-based hash function. In *Fast Software Encryption, 13th International Workshop, FSE 2006, Graz, Austria, March 15-17, 2006, Revised Selected Papers*, pages 328–340, 2006.

[9] John Black, Phillip Rogaway, and Thomas Shrimpton. Black-box analysis of the block-cipher-based hash-function constructions from PGV. In *Advances in Cryptology - CRYPTO 2002, 22nd Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 2002, Proceedings*, pages 320–335, 2002.

[10] John Black, Phillip Rogaway, Thomas Shrimpton, and Martijn Stam. An analysis of the blockcipher-based hash functions from PGV. *J. Cryptology*, 23(4):519–545, 2010.

[11] Ran Canetti, Oded Goldreich, and Shai Halevi. On the random-oracle methodology as applied to length-restricted signature schemes. In *Theory of Cryptography, First Theory of Cryptography Conference, TCC 2004, Cambridge, MA, USA, February 19-21, 2004, Proceedings*, pages 40–57, 2004.

[12] Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. *J. ACM*, 51(4):557–594, 2004.

[13] Shan Chen and John P. Steinberger. Tight security bounds for key-alternating ciphers. In *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, pages 327–350, 2014.

[14] Kai-Min Chung, Huijia Lin, Mohammad Mahmoody, and Rafael Pass. On the power of nonuniformity in proofs of security. In *Innovations in Theoretical Computer Science, ITCS '13, Berkeley, CA, USA, January 9-12, 2013*, pages 389–400, 2013.

[15] Sandro Coretti, Yevgeniy Dodis, Siyao Guo, and John Steinberger. Random oracles and nonuniformity. In *Advances in Cryptology - EUROCRYPT 2018 - To appear*, 2018.

[16] Henry Corrigan-Gibbs and Dmitry Kogan. The discrete-logarithm problem with preprocessing. In *Advances in Cryptology - EUROCRYPT 2018 - To appear*, 2018.

[17] Anindya De, Luca Trevisan, and Madhur Tulsiani. Time space tradeoffs for attacks against one-way functions and prgs. In *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings*, pages 649–665, 2010.

[18] Alexander W. Dent. The hardness of the dhk problem in the generic group model. Cryptology ePrint Archive, Report 2006/156, 2006. `https://eprint.iacr.org/2006/156`.

[19] 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*, 2017.

[20] Yevgeniy Dodis, Krzysztof Pietrzak, and Daniel Wichs. Key derivation without entropy waste. In *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, pages 93–110, 2014.

[21] Yevgeniy Dodis, Leonid Reyzin, Ronald L. Rivest, and Emily Shen. Indifferentiability of permutation-based compression functions and tree-based modes of operation, with applications to MD6. In *Fast Software Encryption, 16th International Workshop, FSE 2009, Leuven, Belgium, February 22-25, 2009, Revised Selected Papers*, pages 104–121, 2009.

[22] Orr Dunkelman, Nathan Keller, and Adi Shamir. Minimalism in cryptography: The even-mansour scheme revisited. In *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, April 15-19, 2012. Proceedings*, pages 336–354, 2012.

[23] Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. In *Advances in Cryptology - ASIACRYPT '91, International Conference on the Theory and Applications of Cryptology, Fujiyoshida, Japan, November 11-14, 1991, Proceedings*, pages 210–224, 1991.

[24] Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. *J. Cryptology*, 10(3):151–162, 1997.

[25] Pierre-Alain Fouque, Antoine Joux, and Chrysanthi Mavromati. Multi-user collisions: Applications to discrete logarithm, even-mansour and PRINCE. In *Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014. Proceedings, Part I*, pages 420–438, 2014.

[26] Peter Gazi and Stefano Tessaro. Provably robust sponge-based prngs and kdfs. In *Advances in Cryptology - EUROCRYPT 2016 - 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part I*, pages 87–116, 2016.

[27] Rosario Gennaro and Luca Trevisan. Lower bounds on the efficiency of generic cryptographic constructions. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 305–313, 2000.

[28] Oded Goldreich and Hugo Krawczyk. On the composition of zero-knowledge proof systems. *SIAM J. Comput.*, 25(1):169–192, 1996.

[29] Oded Goldreich and Yair Oren. Definitions and properties of zero-knowledge proof systems. *J. Cryptology*, 7(1):1–32, 1994.

[30] Shafi Goldwasser and Yael Tauman Kalai. On the (in)security of the fiat-shamir paradigm. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 102–113, 2003.

[31] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM J. Comput.*, 45(5):1835–1869, 2016.

[32] Louis Granboulan. Short signatures in the random oracle model. In *Advances in Cryptology - ASIACRYPT 2002, 8th International Conference on the Theory and Application of Cryptology and Information Security, Queenstown, New Zealand, December 1-5, 2002, Proceedings*, pages 364–378, 2002.

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

[34] Jakob Jonsson. An OAEP variant with a tight security proof. *IACR Cryptology ePrint Archive*, 2002:34, 2002.

[35] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography.* Chapman and Hall/CRC Press, 2007.

[36] Pravesh Kothari, Raghu Meka, and Prasad Raghavendra. Approximating rectangles by juntas and weakly-exponential lower bounds for LP relaxations of csps. *STOC*, 2017.

[37] Mohammad Mahmoody and Ameer Mohammed. On the power of hierarchical identity-based encryption. In *Advances in Cryptology - EUROCRYPT 2016 - 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part II*, pages 243–272, 2016.

[38] Joseph P. Mihalcik. An analysis of algorithms for solving discrete logarithms in fixed groups. Master's thesis, Naval Postgraduate School, Monterey, California, 2010.

[39] Jesper Buus Nielsen. Separating random oracle proofs from complexity theoretic proofs: The non-committing encryption case. In *Advances in Cryptology - CRYPTO 2002, 22nd Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 2002, Proceedings*, pages 111–126, 2002.

[40] Philippe Oechslin. Making a faster cryptanalytic time-memory trade-off. In *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, pages 617–630, 2003.

[41] National Institute of Standards and Technology (NIST). FIPS 202. sha-3 standard: Permutation-based hash and extendable-output functions. Technical report, US Department of Commerce, Apr 2014.

[42] National Institute of Standards and Technology (NIST). FIPS 180-4. secure hash standard. Technical report, US Department of Commerce, Aug 2015.

[43] Duong Hieu Phan and David Pointcheval. Chosen-ciphertext security without redundancy. In *Advances in Cryptology - ASIACRYPT 2003, 9th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, November 30 - December 4, 2003, Proceedings*, pages 1–18, 2003.

[44] Ronald L. Rivest. The MD5 Message-Digest algorithm (RFC 1321). `http://www.ietf.org/rfc/rfc1321.txt?number=1321`.

[45] Ronald L. Rivest, Benjamin Agre, Daniel V. Bailey, Christopher Crutchfield, Yevgeniy Dodis, Kermin Elliott, Fleming Asif Khan, Jayant Krishnamurthy, Yuncheng Lin, Leo Reyzin, Emily Shen, Jim Sukha, Drew Sutherland, Eran Tromer, and Yiqun Lisa Yin. The md6 hash function: A proposal to nist for sha-3, 2008.

[46] Phillip Rogaway and John Steinberger. *Constructing Cryptographic Hash Functions from Fixed-Key Blockciphers*, pages 433–450. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[47] National Technical Information Service. FIPS 180-1. secure hash standard. Technical report, US Department of Commerce, Apr 1995.

[48] Victor Shoup. Lower bounds for discrete logarithms and related problems. In *Advances in Cryptology - EUROCRYPT '97, International Conference on the Theory and Application of Cryptographic Techniques, Konstanz, Germany, May 11-15, 1997, Proceeding*, pages 256–266, 1997.

[49] Thomas Shrimpton and Martijn Stam. Building a collision-resistant compression function from non-compressing primitives. In *Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part II - Track B: Logic, Semantics, and Theory of Programming & Track C: Security and Cryptography Foundations*, pages 643–654, 2008.

[50] Stefano Tessaro. Security amplification for the cascade of arbitrarily weak prps: Tight bounds via the interactive hardcore lemma. In *Theory of Cryptography - 8th Theory of Cryptography Conference, TCC 2011, Providence, RI, USA, March 28-30, 2011. Proceedings*, pages 37–54, 2011.

[51] Dominique Unruh. Random oracles and auxiliary input. In *Advances in Cryptology - CRYPTO 2007, 27th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2007, Proceedings*, pages 205–223, 2007.

[52] Robert S. Winternitz. A secure one-way hash function built from DES. In *Proceedings of the 1984 IEEE Symposium on Security and Privacy, Oakland, California, USA, April 29 - May 2, 1984*, pages 88–90, 1984.

[53] Aaram Yun. Generic hardness of the multiple discrete logarithm problem. In *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part II*, pages 817–836, 2015.

A Auxiliary Input vs. Bit Fixing: Proof

This section contains the proof of Theorem 1. To that end one first shows that any ideal cipher (IC) about which an attacker may have a certain amount of auxiliary information can be replaced by a suitably chosen convex combination of so-called *dense* ICs (cf. Definition 3 and Lemma 13 in Section A.1). In turn, dense ICs can be replaced by *bit-fixing* ICs; this substitution comes at the price of either an additive term to the distinguishing advantage or a multiplicative one to the probability that an attacker outputs 1 (cf. Lemma 16 in Section A.2).

With the above one can immediately prove Theorem 1 for ICs—and random permutations (RPs) as a special case. To establish the theorem for the generic-group model (GGM), one observes that a random injection essentially becomes a permutation once the range is fixed. Therefore, the bit-fixing result for RPs can be reused for the GGM since a group-operation query corresponds to three queries to the underlying injection: two inverse queries and one foreward query (cf. Section A.3).

A.1 Replacing Leaky Sources by Dense Sources

A (K, N)-*cipher source* X is a random variable that takes on as value function tables corresponding to functions $F : [K] \times [N] \to [N]$, where for each $k \in [K]$, $F[k, \cdot]$ is a permutation. Consider the following definition of dense cipher sources.

Definition 3. *A (K, N)-cipher source X is called* $(\bar{P}, 1-\delta)$-dense *for $\bar{P} = (P_1, \ldots, P_K) \in [N]^K$ if it is fixed on at most P_k coordinates $(k, \cdot)$ for each $k \in [K]$ and if for all families $I = \{I_k\}_{k\in[K]}$ of subsets I_k of* non-fixed *coordinates $(k, \cdot)$,*

$$H_\infty(X_I) \;\geq\; (1-\delta)\sum_{k=1}^{K} \log(N - P_k)^{\underline{|I_k|}},$$

where $a^{\underline{b}} := a!/(a-b)!$ and X_I is X restricted to the coordinates in I. X is called $(1-\delta)$-dense *if it is $(0, 1-\delta)$-dense, and* $\bar{P}$-fixed *if it is $(\bar{P}, 1)$-dense.*

It turns out that cipher sources with high min-entropy are close to dense cipher sources. The following lemma is an adaptation of the corresponding lemma for random functions due to Göös *et al.* [31], who used it in the area of communication complexity. The technique was also adopted in a paper by Kothari *et al.* [36], who gave a simplified argument for decomposing high-entropy sources into bit-fixing sources with constant density (cf. Definition 3). For self-containment, a proof is provided here.

Lemma 13. *For every $\delta > 0$, a cipher source X is γ-close to a convex combination of finitely many $(\bar{P}', 1-\delta)$-dense sources for some $\bar{P} = (P_1, \ldots, P_K)$ satisfying*

$$\sum_{k=1}^{K} P_k \;\leq\; \frac{S + \log 1/\gamma}{\delta \cdot \log(N/e)}\;,$$

where $S = K \log N! - H_\infty(X)$ is the min-entropy deficiency of X.

Proof. For ease of notation, let $S := S_z$ and $X := X_z$. Suppose X is not $(1-\delta)$-dense, as otherwise there is nothing to show. Let $Y := X$ and $I = \{I_k\}_{k\in[K]}$ be a maximal family of subsets such that there exists a y_I with

$$\mathsf{P}[Y_I = y_I] \;>\; 2^{-(1-\delta)\sum_{k=1}^{K} \log N^{\underline{|I_k|}}}, \tag{2}$$

where "maximal" means that none of the sets I_k can be extended without violating property (2). Let Y' be the distribution of Y conditioned on $Y_I = y_I$.

Claim 14. *Y' is a $((|I_1|, \ldots, |I_k|), 1-\delta)$-dense IC source.*

Proof. Suppose Y' is not $((|I_1|, \ldots, |I_k|), 1-\delta)$-dense. Then, there exists a family $J = \{J_k\}_{k\in[K]}$ with at least one non-empty set $J_k \subseteq \overline{I_k}$ and y_J such that

$$\mathsf{P}[Y'_J = y_J] \;=\; \mathsf{P}[Y_J = y_J | Y_I = y_I] \;>\; 2^{-(1-\delta)\sum_{k=1}^{K} \log(N-|I_k|)^{\underline{|J_k|}}}.$$

The set family $I \cup J$ defined by the component-wise union now satisfies

$$\begin{aligned}
\mathsf{P}[Y_{I\cup J} = y_{I\cup J}] &= \mathsf{P}[Y_I = y_I \wedge Y_J = y_J] \\
&= \mathsf{P}[Y_I = y_I] \cdot \mathsf{P}[Y_J = y_J | Y_I = y_I] \\
&> 2^{-(1-\delta)\sum_{k=1}^{K} \log N^{\underline{|I_k|}}} \cdot 2^{-(1-\delta)\sum_{k=1}^{K} \log(N-|I_k|)^{\underline{|J_k|}}} \\
&= 2^{-(1-\delta)\sum_{k=1}^{K} \log N^{\underline{|I_k \cup J_k|}}}.
\end{aligned}$$

since I_k and J_k are disjoint for every k. This, however, contradicts the maximality of I. □

Claim 15. $\sum_{k=1}^{K} |I_k| \le S/(\delta \log(N/e))$.

Proof. For a vector y_I, let $\mathcal{B}_{y_I}$ be the set of vectors $y_{\bar{I}}$ such that y_I and $y_{\bar{I}}$ are a valid IC. Then, on the one hand, $H_\infty(Y) \ge K \log N! - S$ implies that for any y_I,

$$\begin{aligned}
\mathsf{P}[Y_I = y_I] &= \sum_{y_{\bar{I}} \in \mathcal{B}_{y_I}} \mathsf{P}[Y_I = y_I \wedge Y_{\bar{I}} = y_{\bar{I}}] \\
&\le 2^{\sum_{k=1}^{K} \log(N-|I_k|)!} \cdot 2^{-(K \log N! - S)} \\
&= 2^{-\left(\sum_{k=1}^{K} \log N^{\underline{|I_k|}} - S\right)},
\end{aligned}$$

and, hence, $H_\infty(Y_I) \ge \sum_{k=1}^{K} \log N^{\underline{|I_k|}} - S$. On the other hand, because Y_I is not $(1-\delta)$-dense, $H_\infty(Y_I) < (1-\delta)\sum_{k=1}^{K} \log N^{\underline{|I_k|}}$. Combining the above two inequalities, one obtains

$$\sum_{k=1}^{K} \log N^{\underline{|I_k|}} \;\le\; S/\delta \;.$$

Using Proposition 38, the desired conclusion follows. □

Set Y now to be Y conditioned on $Y_I \neq y_I$ and recursively decompose Y as long as

$$\mathsf{P}[X \in \mathrm{supp}(Y)] \;>\; \gamma. \tag{3}$$

Observe that $H_\infty(Y) \ge K \log N! - (S + \log 1/\gamma)$ at any point in this decomposition process since

$$\begin{aligned}
\mathsf{P}[Y = y] &= \mathsf{P}[X = y | X \in \mathrm{supp}(Y)] \\
&\le \frac{\mathsf{P}[X = y]}{\mathsf{P}[X \in \mathrm{supp}(Y)]} \\
&\le \frac{2^{-(K \log N! - S)}}{\gamma} \;=\; 2^{-(K \log N! - (S + \log 1/\gamma))}.
\end{aligned}$$

Note that $|\mathrm{supp}(Y)|$ decreases in every step, and since $\mathrm{supp}(X)$ is finite, after finitely many steps, this process ends with a Y_{final} with $\mathsf{P}[X \in \mathrm{supp}(Y_{\mathsf{final}})] \leq \gamma$. Hence, X is a convex combination of finitely many $(\bar{P}, 1-\delta)$-dense sources for $\bar{P} = (P_1, \ldots, P_k)$ with

$$\sum_{k=1}^{K} P_k \;\leq\; (S + \log 1/\gamma)/(\delta \log(N/e))$$

and Y_{final}. This implies that X is γ-*close* to a convex combination of $(\bar{P}, 1-\delta)$-dense sources (e.g., the convex combination obtained by replacing Y_{final} by the uniform distribution). □

A.2 Replacing Dense Sources by Bit-Fixing Sources

Lemma 16. *Let X be an ideal (K, N)-cipher and $f : (\mathcal{P}_N)^K \to \{0,1\}^S$ an arbitrary function. For any $\gamma > 0$ and $P \in \mathbb{N}$, there exists $\bar{P} = (P_1, \ldots, P_K) \in [N]^K$ with $\sum_{k=1}^{K} P_k = P$ and a family $\{Y_z\}_{z \in \{0,1\}^S}$ of convex combinations Y_z of P-fixed (K, N)-cipher sources such that for any distinguisher D taking an S-bit input and querying at most T coordinates of its oracle*

$$\left|\mathsf{P}\left[\mathcal{D}^{X}(f(X)) = 1\right] - \mathsf{P}\left[\mathcal{D}^{Y_{f(X)}}(f(X)) = 1\right]\right| \;\leq\; \frac{2(S + \log 1/\gamma) \cdot T}{P} + \gamma,$$

and

$$\mathsf{P}\left[\mathcal{D}^{X}(f(X)) = 1\right] \;\leq\; 2^{2(S_z + \log 1/\gamma)T/P} \cdot \mathsf{P}\left[\mathcal{D}^{Y'_{f(X)}}(f(X)) = 1\right] + \gamma,$$

provided $N \geq 16$.

Proof. Fix $\gamma > 0$ and an arbitrary $z \in \{0,1\}^S$ and let X_z be the distribution of X conditioned on $f(X) = z$. Let $S_z = K \log N! - H_\infty(X_z)$ be the min-entropy deficiency of X_z. The following claim follows immediately from Lemma 13.

Claim 17. *For every $\delta > 0$, X_z is γ-close to a convex combination of finitely many $(\bar{P}', 1-\delta)$-dense sources for some $\bar{P}' = (P'_1, \ldots, P'_K)$ satisfying*

$$\sum_{k=1}^{K} P'_k \;\leq\; \frac{S_z + \log 1/\gamma}{\delta \cdot \log(N/e)} \;.$$

Let X'_z be the convex combination of $(\bar{P}', 1-\delta)$-dense sources that is γ-close to X_z for a $\delta = \delta_z$ to be determined later. For every $(\bar{P}', 1-\delta)$ source X' in said convex combination, let Y' be the corresponding P'-fixed source Y', i.e., X' and Y' are fixed on the same coordinates to the same values. The following claim bounds the distinguishing advantage between X' and Y' for any T-query distinguisher.

Claim 18. *For any $(\bar{P}', 1-\delta)$-dense (K, N)-cipher source X' and its corresponding P'-fixed source Y', it holds that for any (adaptive) distinguisher $\mathcal{D}$ that makes at most, for each k, T_k* forward or inverse *queries to its oracle,*

$$\left|\mathsf{P}\left[\mathcal{D}^{X'} = 1\right] - \mathsf{P}\left[\mathcal{D}^{Y'} = 1\right]\right| \;\leq\; T\delta \cdot \log N \;,$$

and

$$\mathsf{P}\left[\mathcal{D}^{X'} = 1\right] \;\leq\; N^{T\delta} \cdot \mathsf{P}\left[\mathcal{D}^{Y'} = 1\right] \;,$$

where $T = \sum_{k=1}^{K} T_k$.

Proof. Assume without loss of generality that $\mathcal{D}$ is deterministic and does not query any of the fixed positions, make the same query twice, or make an inverse query after making the corresponding forward query or vice-versa. Let $T_{X'}$ and $T_{Y'}$ be the random variables corresponding to the transcripts containing the query/answer pairs resulting from $\mathcal{D}$'s interaction with X' and Y', respectively. More specifically, τ contains a pair $((k,x),y)$ whenever either (k,x) was queried and y was the answer (forward query) or the other way around (inverse query); it will not matter which was the case, however, for the probabilities computed below.

For a fixed transcript τ, denote by $\mathsf{p}_{X'}(\tau)$ and $\mathsf{p}_{Y'}(\tau)$ the probabilities that X' and Y', respectively, produce the answers in τ if asked the queries in τ. Since $\mathcal{D}$ is deterministic, $\mathsf{P}[T_{X'}=\tau] \in \{0, \mathsf{p}_{X'}(\tau)\}$, and similarly, $\mathsf{P}[T_{Y'}=\tau] \in \{0, \mathsf{p}_{Y'}(\tau)\}$. Denote by $\mathcal{T}_X$ the set of all transcripts τ for which $\mathsf{P}[T_{X'}=\tau] > 0$. For such τ, $\mathsf{P}[T_{X'}=\tau] = \mathsf{p}_{X'}(\tau)$ and also $\mathsf{P}[T_{Y'}=\tau] = \mathsf{p}_{Y'}(\tau)$.

For every of the K possible keys, in order to bound the probabilities $\mathsf{p}_{X'}(\tau)$ and $\mathsf{p}_{Y'}(\tau)$, consider the complete bipartite graph $K_{N,N}$. The fixed coordinates form a matching of size (at most) P' in $K_{N,N}$, and choosing a permutation according to either X' or Y' amounts to extending said matching to a perfect matching. Similarly, the query/answer pairs in τ form a matching compatible with that defined by the fixed coordinates. Therefore, $\mathsf{p}_{X'}(\tau)$ and $\mathsf{p}_{Y'}(\tau)$ are equal to the probabilities that an extension according to X' and Y', respectively, contains the matching defined by τ. That is,

$$\mathsf{p}_{X'}(\tau) \;\leq\; 2^{-(1-\delta)\sum_{k=1}^{K} \log(N-P'_k)^{\underline{T_k}}} \qquad \text{and} \qquad \mathsf{p}_{Y'}(\tau) \;=\; 2^{-\sum_{k=1}^{K} \log(N-P'_k)^{\underline{T_k}}} \tag{4}$$

as X' is $(\bar{P}', 1-\delta)$-dense and Y' is $\bar{P}'$-fixed.

Towards proving the first part of the lemma, observe that $\mathcal{D}$'s output can be computed from the transcript (including whether a query was a forward or an inverse query) by just running $\mathcal{D}$ and providing the answers to its queries from the transcript. Hence,

$$\begin{aligned}
\left|\mathsf{P}\big[\mathcal{D}^{X'}=1\big] - \mathsf{P}\big[\mathcal{D}^{Y'}=1\big]\right| \;&\leq\; \mathsf{SD}(T_{X'}, T_{Y'}) \\
&=\; \sum_{\tau} \max\{0, \mathsf{P}[T_{X'}=\tau] - \mathsf{P}[T_{Y'}=\tau]\} \\
&=\; \sum_{\tau\in\mathcal{T}_X} \max\{0, \mathsf{p}_{X'}(\tau) - \mathsf{p}_{Y'}(\tau)\} \\
&=\; \sum_{\tau\in\mathcal{T}_X} \mathsf{p}_{X'}(\tau) \cdot \max\left\{0, 1 - \frac{\mathsf{p}_{Y'}(\tau)}{\mathsf{p}_{X'}(\tau)}\right\} \\
&\leq\; 1 - 2^{-\delta\sum_{k=1}^{K} \log(N-P'_k)^{\underline{T_k}}} \\
&\leq\; 1 - 2^{-\delta T \log N} \;\leq\; T\delta \cdot \log N \;,
\end{aligned}$$

where the first sum is over all possible transcripts and where the last inequality uses $2^{-x} \geq 1-x$ for $x \geq 0$ and $a^{\underline{b}} \leq a^b$ for $a, b \in \mathbb{N}$.

As for the second part of the lemma, observe that due to (4) and the support of $T_{X'}$ being a subset of $T_{Y'}$,

$$\mathsf{P}[T_{X'}=\tau] \;\leq\; 2^{\delta\sum_{k=1}^{K} \log(N-P'_k)^{\underline{T}}} \cdot \mathsf{P}[T_{Y'}=\tau] \;=\; 2^{T\delta\log N} \cdot \mathsf{P}[T_{Y'}=\tau]$$

for any transcript τ. Let $\mathcal{T}_{\mathcal{D}}$ be the set of transcripts where $\mathcal{D}$ outputs 1. Then,

$$\begin{aligned}\mathsf{P}[\mathcal{D}^{X'}=1] &= \sum_{\tau\in\mathcal{T}_{\mathcal{D}}} \mathsf{P}[T_{X'}=\tau] \\ &\le 2^{T\delta\log N}\cdot\sum_{\tau\in\mathcal{T}_{\mathcal{D}}}\mathsf{P}[T_{Y'}=\tau] \\ &= 2^{T\delta\log N}\cdot\mathsf{P}[\mathcal{D}^{Y'}=1]\,.\end{aligned}$$

□

Let Y'_z be obtained by replacing X' by Y' in X'_z. Setting

$$\delta_z \;=\; \frac{S_z+\log 1/\gamma}{P\cdot\log(N/e)}\,,$$

The above claims imply

$$\Big|\mathsf{P}\big[\mathcal{D}^{X_z}(z)=1\big]-\mathsf{P}\big[\mathcal{D}^{Y'_z}(z)=1\big]\Big| \;\le\; 2\frac{(S_z+\log 1/\gamma)\cdot T}{P}+\gamma\,, \tag{5}$$

as well as

$$\mathsf{P}\big[\mathcal{D}^{X_z}(z)=1\big] \;\le\; 2^{2(S_z+\log 1/\gamma)T/P}\cdot\mathsf{P}\big[\mathcal{D}^{Y'_z}(z)=1\big]+\gamma\,, \tag{6}$$

using that $\log(N)/\log(N/e)\le 2$ for $N\ge 16$. Moreover, note that for the above choice of δ_z, Lemma 13 implies $\sum_{k=1}^{K}P'_k\le P$, i.e., the sources Y' are fixed on at most P coordinates in total.

Claim 19. *$\mathbf{E}_z[S_z]\le S$ and $\mathbf{E}_z[2^{S_zT/P}]\le 2^{ST/P}$.*

Proof. Observe that $H_\infty(X_z)=H_\infty(X|Z=z)=H(X|Z=z)$ since, conditioned on $Z=z$, X is distributed uniformly over all values x with $f(x)=z$. Therefore,

$$\begin{aligned}\mathbf{E}_z[S_z] \;=\; K\log N!-\mathbf{E}_z[H_\infty(X|Z=z)] &= K\log N!-\mathbf{E}_z[H(X|Z=z)] \\ &= K\log N!-H(X|Z) \;\le\; S.\end{aligned}$$

Again due to the uniformity of X, $\mathsf{P}[f(X)=z]=2^{-S_z}$. Hence,

$$\mathbf{E}_z[2^{S_z/S}] \;=\; \sum_z 2^{-S_z}\cdot 2^{S_zT/P} \;=\; \sum_z 2^{-S_z(1-T/P)}.$$

Because $T<P$, one can use Hölder's inequality and obtain

$$\begin{aligned}\sum_z 2^{-S_z(1-T/P)} &\le \left(\sum_z\left(2^{-S_z(1-T/P)}\right)^{1/(1-T/P)}\right)^{1-T/P}\cdot\left(\sum_z 1^{P/T}\right)^{T/P} \\ &= 1^{1-T/P}\cdot 2^{ST/P} \;=\; 2^{ST/P}.\end{aligned}$$

□

The lemma now follows (using $Y_z:=Y'_z$) by taking expectations over z of (5) and (6) and applying the above claim. □

For injections, consider the following definition and an analog of Lemma 20.

Definition 4. *An* (N, M)-injection source X *is a random variable that takes on as value function tables corresponding to injections* $F : [N] \to [M]$*. An* (N, M)*-injection source* X *is called* $(P, \mathcal{Y})$-fixed *for* $\mathcal{Y} \subseteq [M]$ *if its range is* $\mathcal{Y}$ *and it is fixed on at most* P *coordinates.*

Lemma 20. *Let* X *be a uniform random* (N, M)*-injection and* $f : \mathcal{F}_{N,M} \to \{0,1\}^S$ *an arbitrary function. For any* $\gamma > 0$ *and* $P \in \mathbb{N}$*, there exists a family* $\{Y_{z,\mathcal{Y}}\}_{z,\mathcal{Y}}$*, indexed by values* $z \in \{0,1\}^S$ *and size-*N *subsets* $\mathcal{Y}$ *of* $[M]$*, of convex combinations* $Y_{z,\mathcal{Y}}$ *of* $(P, \mathcal{Y})$*-fixed injection sources such that for any distinguisher* D *taking an* S*-bit input and making at most* T *forward or backward queries to its oracle*

$$\left|\mathsf{P}\left[\mathcal{D}^{X}(f(X)) = 1\right] - \mathsf{P}\left[\mathcal{D}^{Y_{f(X),\mathsf{im}(X)}}(f(X)) = 1\right]\right| \;\leq\; \frac{2(S + \log 1/\gamma) \cdot T}{P} + \gamma \;,$$

and

$$\mathsf{P}\left[\mathcal{D}^{X}(f(X)) = 1\right] \;\leq\; 2^{2(S_z + \log 1/\gamma)T/P} \cdot \mathsf{P}\left[\mathcal{D}^{Y'_{f(X)}}(f(X)) = 1\right] + \gamma \;,$$

provided $N \geq 16$*.*

Proof. The lemma is proved by conditioning on the range $\mathsf{im}(X)$ of X, which turns X into a bijection, and then proceeding as in the proof of Lemma 16. □

A.3 From Bit-Fixing to Auxiliary-Input

For convenience, Theorem 1 is restated here.

Theorem 1. *Let* $P, K, N, M \in \mathbb{N}$*,* $N \geq 16$*, and* $\gamma > 0$*. Moreover, let*

$$(\mathsf{AI}, \mathsf{BF}) \in \{(\mathsf{AI\text{-}IC}(K, N), \mathsf{BF\text{-}IC}(P, K, N)), (\mathsf{AI\text{-}GG}(N, M), \mathsf{BF\text{-}GG}(P, N, M))\} \;.$$

Then,

1. *if an application* G *is* $((S, T, p), \varepsilon')$*-secure in the* BF*-model, it is* $((S, T, p), \varepsilon)$*-secure in the* AI*-model, where*
$$\varepsilon \;\leq\; \varepsilon' + \frac{6(S + \log \gamma^{-1}) \cdot T_G^{\mathsf{comb}}}{P} + \gamma \,;$$
2. *if an* unpredictability *application* G *is* $((S, T, p), \varepsilon')$*-secure in the* BF*-model for*
$$P \;\geq\; 6(S + \log \gamma^{-1}) \cdot T_G^{\mathsf{comb}},$$
it is $((S, T, p), \varepsilon)$*-secure in the* AI*-model for*
$$\varepsilon \;\leq\; 2\varepsilon' + \gamma \,,$$

where T_G^{comb} *is the combined query complexity corresponding to* G*.*

Proof. The proof is split into two parts, where the first part deals with the ideal-cipher (and, hence, the random-permutation model) and the second part with the generic group model.

- **Ideal cipher model:** Fix P as well as γ. Set $\mathsf{BF\text{-}IC} := \mathsf{BF\text{-}IC}(P)$ and let G be an arbitrary application and C the corresponding challenger. Moreover, fix an (S, T)-attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, and let $\{Y_z\}_{z \in \{0,1\}^S}$ be the family of distributions guaranteed to exist by Lemma 16, where the function f is defined by $\mathcal{A}_1$. Consider the following (S, T)-attacker $\mathcal{A}' = (\mathcal{A}'_1, \mathcal{A}'_2)$ (expecting to interact with $\mathsf{BF\text{-}IC}$):

– $\mathcal{A}'_1$ internally simulates $\mathcal{A}_1$ to compute $z \leftarrow \mathcal{A}_1^{\mathsf{AI\text{-}IC.pre}}$. Then, it samples one of the P-bit-fixing sources Y' making up Y_z and presets $\mathsf{BF\text{-}IC}$ to match Y' on the at most P points where Y' is fixed. The output of $\mathcal{A}'_1$ is z.

– $\mathcal{A}'_2$ works exactly as $\mathcal{A}_2$.

Let $\mathcal{D}$ be the combination of $\mathcal{A}_2 = \mathcal{A}'_2$ and C. Hence, $\mathcal{D}$ is a distinguisher taking an S-bit input and making at most T_G^{comb} queries to its oracle. Therefore, by the first part of Lemma 16,

$$\mathrm{Succ}_{G,\mathsf{AI\text{-}IC}}(\mathcal{A}) \;\le\; \mathrm{Succ}_{G,\mathsf{BF\text{-}IC}}(\mathcal{A}') + \frac{2(S+\log\gamma^{-1})\cdot T_G^{\mathsf{comb}}}{P} + \gamma \;.$$

Since there is only an additive term between the two success probabilities, the above inequality implies

$$\mathrm{Adv}_{G,\mathsf{AI\text{-}IC}}(\mathcal{A}) \;\le\; \mathrm{Adv}_{G,\mathsf{BF\text{-}IC}}(\mathcal{A}') + \frac{2(S+\log\gamma^{-1})\cdot T_G^{\mathsf{comb}}}{P} + \gamma$$

for both indistinguishability and unpredictability applications.

Using the same attacker $\mathcal{A}'$ as above and applying the second part of Lemma 16, one obtains, for any $P \ge 2(S+\log\gamma^{-1})\cdot T_G^{\mathsf{comb}}$,

$$\begin{aligned}\mathrm{Succ}_{G,\mathsf{AI\text{-}IC}}(\mathcal{A}) &\le 2^{2(S+\log 1/\gamma)T_G^{\mathsf{comb}}/P}\cdot \mathrm{Succ}_{G,\mathsf{BF\text{-}IC}}(\mathcal{A}') + \gamma \\ &\le 2\cdot \mathrm{Succ}_{G,\mathsf{BF\text{-}IC}}(\mathcal{A}') + \gamma \;,\end{aligned}$$

which translates into

$$\mathrm{Adv}_{G,\mathsf{AI\text{-}IC}}(\mathcal{A}) \;\le\; 2\cdot \mathrm{Adv}_{G,\mathsf{BF\text{-}IC}}(\mathcal{A}') + \gamma$$

for unpredictability applications.

- Fix P as well as γ. Set $\mathsf{BF\text{-}GG} := \mathsf{BF\text{-}GG}(P)$ and let G be an arbitrary application and C the corresponding challenger. Moreover, fix an (S,T)-attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, and let $\{Y_{z,\mathcal{Y}}\}_{z,\mathcal{Y}}$ be the family of distributions guaranteed to exist by Lemma 20, where the function f is defined by $\mathcal{A}_1$. Consider the following (S,T)-attacker $\mathcal{A}' = (\mathcal{A}'_1, \mathcal{A}'_2)$ (expecting to interact with $\mathsf{BF\text{-}GG}$):

  – $\mathcal{A}'_1$ obtains the set $\mathcal{Y}$ from $\mathsf{BF\text{-}GG.pre}$ and internally simulates $\mathcal{A}_1$ on a uniform random injection X with range $\mathcal{Y}$ to obtain $z \leftarrow \mathcal{A}_1^{\mathsf{AI\text{-}GG.pre}}$. Then, it samples one of the P-bit-fixing sources Y' making up $Y_{z,\mathcal{Y}}$ and presets $\mathsf{BF\text{-}GG}$ to match Y' on the at most P points where Y' is fixed. The output of $\mathcal{A}'_1$ is z.

  – $\mathcal{A}'_2$ works exactly as $\mathcal{A}_2$.

  Let $\mathcal{D}$ be a distinguisher—making forward and backward queries to an injection source—that internally runs the combination of $\mathcal{A}_2 = \mathcal{A}'_2$ and C and answers their queries as follows: (1) It passes forward queries to and back from its own oracle. (2) It answers group-operation queries (s, s') by making two backward queries to its own oracle for s and s', obtaining i and j, respectively, making a forward query $i + j$, and passing the answer to back to $\mathcal{A}_2$ or C (unless one of the answers to the backward queries was $\perp$, in which case $\perp$ is returned).

  Note that $\mathcal{D}^X(f(X))$ is identical to $\mathcal{A}_2^{\mathsf{AI\text{-}GG.main}}(\mathcal{A}_1^{\mathsf{AI\text{-}GG.pre}}) \leftrightarrow \mathsf{C}^{\mathsf{AI\text{-}GG.main}}$, and $\mathcal{D}^{Y_{f(X),\mathrm{im}(X)}}(f(X))$ is identical to $\mathcal{A}_2'^{\mathsf{BF\text{-}GG.main}}(\mathcal{A}_1'^{\mathsf{BF\text{-}GG.pre}}) \leftrightarrow \mathsf{C}^{\mathsf{BF\text{-}GG.main}}$. Furthermore, $\mathcal{D}$ is a distinguisher taking an S-bit input and making at most $3T_G^{\mathsf{comb}}$ queries to its oracle. Therefore, by the first part of Lemma 20,

$$\mathrm{Succ}_{G,\mathsf{AI\text{-}GG}}(\mathcal{A}) \;\le\; \mathrm{Succ}_{G,\mathsf{BF\text{-}GG}}(\mathcal{A}') + \frac{6(S+\log\gamma^{-1})\cdot T_G^{\mathsf{comb}}}{P} + \gamma \;.$$

Since there is only an additive term between the two success probabilities, the above inequality implies

$$\mathrm{Adv}_{G,\mathsf{AI\text{-}GG}}(\mathcal{A}) \;\leq\; \mathrm{Adv}_{G,\mathsf{BF\text{-}GG}}(\mathcal{A}') + \frac{6(S+\log\gamma^{-1})\cdot T_G^{\mathsf{comb}}}{P} + \gamma$$

for both indistinguishability and unpredictability applications.

Using the same attacker $\mathcal{A}'$ and distinguisher $\mathcal{D}$ as above and applying the second part of Lemma 16, one obtains, for any $P \geq 6(S+\log\gamma^{-1})\cdot T_G^{\mathsf{comb}}$,

$$\begin{aligned}\mathrm{Succ}_{G,\mathsf{AI\text{-}GG}}(\mathcal{A}) &\leq 2^{6(S+\log 1/\gamma)T_G^{\mathsf{comb}}/P}\cdot \mathrm{Succ}_{G,\mathsf{BF\text{-}GG}}(\mathcal{A}') + \gamma \\ &\leq 2\cdot \mathrm{Succ}_{G,\mathsf{BF\text{-}GG}}(\mathcal{A}') + \gamma\ ,\end{aligned}$$

which translates into

$$\mathrm{Adv}_{G,\mathsf{AI\text{-}GG}}(\mathcal{A}) \;\leq\; 2\cdot \mathrm{Adv}_{G,\mathsf{BF\text{-}GG}}(\mathcal{A}') + \gamma$$

for unpredictability applications.

□

B Basic Applications: Security Proofs

This section contains the deferred definitions as well security bounds and the corresponding proofs for the basic applications presented in Section 4.

In the following, $E, E^{-1} : [K]\times[N]\to[N]$ denote the ideal cipher and its inverse to which $\mathsf{AI\text{-}IC}(K,N)$ and $\mathsf{BF\text{-}IC}(P,K,N)$ offer access at interface main (cf. Section 2).

B.1 Even-Mansour Cipher

Even and Mansour [23, 24] considered the following simple block-cipher construction:

$$\mathsf{EM}_{\pi,s_1,s_2}(m) \;:=\; \pi(m\oplus s_2)\oplus s_1\ .$$

Dunkelman *et al.* [22] observed that the construction can be simplified to use a single key $s = s_1 = s_2$.

The application $G^{\mathsf{BC\text{-}EM}}$ is defined via the challenger $\mathsf{C}^{\mathsf{BC\text{-}EM}}$: it picks a random bit $b \leftarrow \{0,1\}$ and a key $s \leftarrow [N]$; when the attacker $\mathcal{A}_2$ makes a sequence of q queries, if $b=0$, the challenger answers forward queries m by $\mathsf{EM}_{\pi,s}(m)$ and inverse queries $c\in[N]$ by $\mathsf{EM}^{-1}_{\pi,s}(c)$; if $b=1$, forward queries m are answered by $f(m)$ and inverse queries c by $f^{-1}(c)$, where f is an independently chosen uniform random permutation; the attacker wins if and only if he correctly guesses the bit b.

Theorem 21. *Application* $G^{\mathsf{BC\text{-}EM}}$ *is* $\left((S,T,q),\tilde{O}\left(\frac{Tq}{N}+\sqrt{\frac{S(T+q)q}{N}}\right)\right)$*-secure in the* $\mathsf{AI\text{-}RP}(N)$*-model for* $N\geq 16$.

Proof. It suffices to show that $G^{\mathsf{BC\text{-}EM}}$ is $\left((S,T,q),O\left(\frac{q(T+P)}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}RP}(P,N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{BC\text{-}EM}}} = T+q$, setting $\gamma := 1/N$ and $P=\sqrt{S(T+q)N/q}$, and applying Theorem 1, the desired conclusion follows.

Consider a slightly modified challenger: let $\pi_1=\pi$ if $b=0$ and an independently chosen uniform random permutation if $b=1$; the modified challenger answers forward queries m by $f(m) := s\oplus\pi_1(s\oplus m)$ and inverse queries by f^{-1}. It is easily seen that the two challengers are indistinguishable.

Let $L_0 = \{(x_1, \pi(x_1)), \ldots, (x_P, \pi(x_P)), (x_{P+1}, \pi(x_{P+1})), \ldots, (x_{P+T}, \pi(x_{P+T}))\}$ be the set of the presampled pairs by $\mathcal{A}_1$ and T pairs queried by $\mathcal{A}_2$ to π (or π^{-1}). Let $Q = \{(m_1, c_1), \ldots, (m_q, c_q)\}$ denote the set of pre-image and image pairs queried by $\mathcal{A}_2$ when interacts with the challenger and $b = 1$. Let $L_1 = \{(m_1 \oplus s, c_1 \oplus s), \ldots, (m_q \oplus s, c_q \oplus s)\}$ be the set of corresponding pre-image and image pairs under π_1. We say π and π_1 are not separated if there exists $i \in [P+T]$ and $j \in [q]$ such that $x_i = m_j + s$ or $\pi(x_i) = c_j + s$.

The following lemma bounds the successful probability of $\mathcal{A}$ by the separation faliure probability of π and π_1.

Lemma 22 (Separation Lemma)**.** $\mathrm{Succ}_{G,q}(\mathcal{A}) \leq \mathsf{P}[Z]$ *where* $\mathsf{P}[Z]$ *denotes the separation faliure probability of* π *and* π_1*.*

For every $i \in [P+T], j \in [q]$, because s is completely indepdent of the adversary's view, $x_i = m_j + s$ or $\pi(x_i) = c_j + s$ happens with probability at most $\frac{2}{N}$. By a union bound, $\mathsf{P}[Z] \leq \frac{2(P+T)q}{N}$. Therefore, given this lemma, $\mathrm{Succ}_{G,q}(\mathcal{A}) \leq \frac{2(P+T)q}{N}$ and the desired conclusion follows.

The proof of separation lemma is standard via considering two games. Here we sketch the proof. In both games, the challenger maintains two sets Π_0, Π_1 of pairs $(x, y) \in [N]^2$. Each of them presents a function which can be extended to a permuation on $[N]$. Initially, $\Pi_0 = L$ is the set of presampled pairs by $\mathcal{A}_1$ and $\Pi_1 = \emptyset$. For $i \in \{0, 1\}$, let $\mathrm{Domain}(\Pi_i) := \{x : (x, y) \in \Pi_i\}$, $\mathrm{Range}(\Pi_i) := \{y : (x, y) \in \Pi_i\}$ and $\Pi_i^{-1} = \{(y, x) : (x, y) \in \Pi_i\}$. To simplify the notation, we denote π as π_0.

The first game Game_0 proceeds as application $G^{\mathsf{BC\text{-}EM}}$ except that queries to π_i is proceeded as follows: on forward query x,

1. if $x \in \mathrm{Domain}(\Pi_i)$, return $\Pi_b(x)$.
2. (*) if $x \in \mathrm{Domain}(\Pi_{1-i})$, return $\Pi_{1-i}(x)$.
3. Sample $y \leftarrow [N] \setminus \mathrm{Range}(\Pi_i)$,
4. (*) if $y \in \mathrm{Range}(\Pi_b) := \{y : (x, y) \in \Pi_i\}$, sample $y \leftarrow [N] \setminus (\mathrm{Range}(\Pi_i) \cup \mathrm{Range}(\Pi_{1-i}))$,
5. return y and add (x, y) into π_i.

The inverse query is proceeded in the same way by replacing Π_i by Π_i^{-1}. In the second game Game_1, we simply remove the lines marked with $(*)$ in Game_0.

Clearly, in Game_1, π_0 distributes the same as a random permutation which agrees with L and π_1 distributes as a random permutation. Therefore Game_1 simulates $G^{\mathsf{BC\text{-}EM}}$ perfectly when $b = 1$. In Game_0, π_1 agrees with π_0 which is a random permutation conditioning on L. So Game_0 simulates $G^{\mathsf{BC\text{-}EM}}$ perfectly when $b = 0$. Game_1 proceeds identically as Game_0 unless Z holds so that $\mathrm{Succ}_{G,q}(\mathcal{A}) \leq \mathsf{P}[Z]$. □

B.2 PRF Security of Davies-Meyer

The Davies-Meyer (DM) compression function DM_E is defined as

$$\mathsf{DM}_E(h, m) \;:=\; E(m, h) \oplus h\,.$$

The DM function is a PRF when keyed by h. The corresponding application $G^{\mathsf{PRF\text{-}DM}}$ is defined via the following challenger $\mathsf{C}^{\mathsf{PRF\text{-}DM}}$: it picks a random bit $b \leftarrow \{0, 1\}$ and a key $h \leftarrow [N]$; when

the attacker queries a message m if $b = 0$, the challenger answers by $\mathsf{DM}_E(h, m)$, and, if $b = 1$, the challenger answers by a value chosen uniformly at random for each m. The attacker wins, if and only if he correctly guesses the bit b. Let q denote the number of queries $\mathcal{A}_2$ makes to $\mathsf{C}^{\mathsf{PRF\text{-}DM}}$.

Theorem 23. *The application* $G^{\mathsf{PRF\text{-}DM}}$ *is* $\left((S, T, q), O\left(T/N + \sqrt{S(T+q)/N}\right)\right)$*-secure in the* $\mathsf{AI\text{-}IC}(K, N)$*-model for* $N \geq 16$*.*

Proof. Follows immediately from Lemma 24, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{PRF\text{-}DM}}} = T + q$, setting $\gamma := 1/N$ and $P := \sqrt{S(T+q)N}$, and applying Theorem 1. □

Lemma 24. *For any* $P, N \in \mathbb{N}$, $G^{\mathsf{PRF\text{-}DM}}$ *is* $\left((S, T, q, \ell), O\left(\frac{T+P}{N}\right)\right)$*-secure in the* $\mathsf{BF\text{-}IC}(P, K, N)$*-model.*

Proof. Consider an interaction between $\mathcal{A}_2$ and the challenger $G^{\mathsf{PRF\text{-}MD\text{-}N}}$ and define the hybrid experiment where—after the fixing of up to P positions by $\mathcal{A}_1$—the ideal cipher E is modified to $\hat{E}$ as follows:

1. For each $m = 1, \ldots, K$, a random value y^*_m is chosen uniformly at random and independently.
2. Let h^* be the PRF key chosen by the challenger. Forward and inverse queries to $\hat{E}$ are answered by

$$\hat{E}(m, h) \;:=\; \begin{cases} y^*_m & \text{if } h = h^*, \\ E(m, h) & \text{otherwise,} \end{cases} \qquad \text{and} \qquad \hat{E}^{-1}(m, y) \;:=\; \begin{cases} h^* & \text{if } y = y^*_m, \\ E^{-1}(m, y) & \text{otherwise.} \end{cases}$$

For each m, let y'_m be the value h^* originally mapped to under $E(m, \cdot)$, and let h'_m be the value that maps to y^*_m. Observe that this hybrid experiment behaves differently from the original experiment only if

- the pairs $((m, h^*), y'_m)$ or $((m, h'_m), y^*_m)$ for some m are prefixed by $\mathcal{A}_1$, or
- $\mathcal{A}_2$ queries h'_m to $E(m, \cdot)$ or y'_m to $E^{-1}(m, \cdot)$; call such queries *bad* queries.

The probability of bad prefixing is at most $2P/N$. Moreover, denote by $\mathcal{E}_i$ the event that the first i queries are *not* bad. Note that

$$\mathsf{P}[\neg\mathcal{E}_i | \mathcal{E}_{i-1}] \;\leq\; \frac{2}{N - (P + i - 1)} \;\leq\; \frac{2}{N - (P + T)} \;\leq\; \frac{4}{N}\,,$$

using that $P + T \leq N/2$, an assumption one may always make since, otherwise, the lemma trivially holds. Moreover, note that the above argument holds even if h^* and y^*_m for all m are known. Summarizing, $\mathsf{P}[\mathcal{E}_T] \leq 4T/N$.

In the hybrid experiment, let $\mathcal{E}'_i$ be the event that in the first i queries, $\mathcal{A}_2$ queries neither h^* to $\hat{E}$ nor y^*_m to $\hat{E}^{-1}$ for any m. Since the uniformly random values y^*_m perfectly hide h^* in the answers

$$\mathsf{DM}_{\hat{E}}(h^*, m) \;=\; \hat{E}(m, h^*) \oplus h^* \;=\; y^*_m \oplus h^*\,,$$

one obtains that

$$\mathsf{P}[\neg\mathcal{E}'_i | \mathcal{E}'_{i-1}] \;\leq\; \frac{2}{N - (P + i - 1)} \;\leq\; \frac{2}{N - (P + T)} \;\leq\; \frac{4}{N}\,,$$

and, hence, $\mathsf{P}[\mathcal{E}'_T] \leq 4T/N$. However, unless $\mathcal{E}'_T$ is violated, the advantage of $\mathcal{A}_2$ in guessing b is zero. □

C Sponge-Based Applications: Security Proofs

This section contains the deferred definitions as well security bounds and the corresponding proofs for the sponge-based applications presented in Section 4.

In the following, $\pi, \pi^{-1} : [N] \to [N]$ denote the permutation and its inverse to which $\mathsf{AI\text{-}RP}(N)$ and $\mathsf{BF\text{-}RP}(P, N)$ offer access at interface main (cf. Section 2).

C.1 Definitions

The sponge construction. Assume N is a power of two and let $n = \log N$, $r \leq n$, $c = n - r$. Let $m = m_1 \cdots m_\ell$ be a message consisting of r-bit blocks m_i,[4] and $\mathsf{IV} \in \{0,1\}^c$ an initialization vector. The sponge construction $\mathsf{Sponge}_{\pi,\mathsf{IV}}(m)$ works as follows:

1. Set $s_0 \leftarrow 0^r \| \mathsf{IV}$.
2. For $i = 1, \ldots, \ell$: set $s_i \leftarrow \pi(m_i \oplus s_{i-1}^{(1)} \| s_{i-1}^{(2)})$, where $s_{i-1}^{(1)}$ denotes the first r bits of s_{i-1} and $s_{i-1}^{(2)}$ the remaining c bits.
3. Output $s_\ell^{(1)}$.

Node graphs. A useful formalism for security proofs of sponge-based constructions is that of node and supernode graphs, as introduced by Bertoni *et al.* [6]. For a permutation $\pi : \{0,1\}^n \to \{0,1\}^n$, consider the following (directed) *node graph* $G_\pi = (V, E)$ with $V = \{0,1\}^r \times \{0,1\}^c = \{0,1\}^n$ and $E = \{(s,t) \mid \pi(s) = t\}$. Moreover, let $G'_\pi = (V', E')$ be the (directed) *supernode graph*, with $V' = \{0,1\}^c$ and $(s^{(2)}, t^{(2)}) \in E'$ iff $((s^{(1)}, s^{(2)}), (t^{(1)}, t^{(2)})) \in E$ for some $s^{(1)}, t^{(1)} \in \{0,1\}^r$. Observe that the value of $\mathsf{Sponge}_{\pi,\mathsf{IV}}(m)$ for an ℓ-block message $m = m_1 \cdots m_\ell$ is obtained by starting at $s_0 := (0^r, \mathsf{IV}) \in \{0,1\}^n$ in G_π, moving to $s_i \leftarrow \pi(m_i \oplus s_{i-1}^{(1)} \| s_{i-1}^{(2)})$ for $i = 1, \ldots, \ell$, and outputting $s_\ell^{(1)}$. In other words, in the supernode graph, m corresponds to a path of length ℓ starting at node IV and ending at $s_\ell^{(2)}$, and $s_1^{(1)}, \ldots, s_\ell^{(1)} \in \{0,1\}^r$ are the values that *appear* on that path.

C.2 PRFs from Keyed Sponges

The application $G^{\mathsf{PRF\text{-}S}}$ of PRF security of the sponge construction is defined via the following challenger $\mathsf{C}^{\mathsf{PRF\text{-}S}}$: it picks a random bit $b \leftarrow \{0,1\}$ and a key $k \leftarrow \{0,1\}^c$; when the attacker queries a message $m = m_1 \cdots m_\ell$ consisting of r-bit blocks m_i, if $b = 0$, the challenger answers by $\mathsf{Sponge}_{\pi,k}(m)$, and, if $b = 1$, the challenger answers by a value chosen uniformly at random for each m. The attacker wins, if and only if he correctly guesses the bit b.

The following theorem provides an upper bound on the advantage of an (S, T, q, ℓ)-attacker in distinguishing the sponge construction from a random function in the BF-RPM, where q is an upper bound on the number of messages m the attacker submits to the challenger and ℓ is an upper bound on the length of those messages. The proof follows mostly the approach by [7].

Theorem 25. *$G^{\mathsf{PRF\text{-}S}}$ is $\left((S, T, q, \ell), \tilde{O}\left(\frac{(T+q\ell)^2}{2^c} + \sqrt{\frac{S(T+q\ell)^2}{2^c}}\right)\right)$-secure in the $\mathsf{AI\text{-}RP}(N)$-model, where $N = 2^n = 2^{r+c} \geq 16$.*

[4] To keep things simple, no padding is considered here.

Proof. Follows from Lemma 26 by observing that $T^{\mathsf{comb}}_{G^{\mathsf{PRF\text{-}S}}} = T + q\ell$, setting $\gamma := 1/N$ and $P := \sqrt{\frac{(S+\log N)(T+q\ell)\cdot 2^c}{T+q\ell}} = \tilde{\Theta}\left(\sqrt{\frac{S(T+q\ell)\cdot 2^c}{T+q\ell}}\right)$, and applying Theorem 1. □

Lemma 26. *For any $P \in \mathbb{N}$, $G^{\mathsf{PRF\text{-}S}}$ is $\left((S,T,q,\ell), O\left(\frac{(T+q\ell)^2+(T+q\ell)P)}{2^c}\right)\right)$-secure in the $\mathsf{BF\text{-}RP}(P,N)$-model, where $N = 2^n = 2^{r+c}$.*

Proof. In the following, queries by $\mathcal{A}_2$ to the sponge construction (i.e., to the challenger) are referred to as *construction queries* and (forward or inverse) queries to π as *primitive queries*. Moreover, queries the challenger makes to π as a result of construction queries are called *indirect queries*.

Consider an interaction of $\mathcal{A}_2$ with $\mathsf{C}^{\mathsf{PRF\text{-}S}}$ and incrementally build the node and supernode graphs (as defined above), adding edges corresponding to primitive and indirect queries and starting with the edges that correspond to the at most P prefixed query/answer pairs. Denote by $\mathcal{F}$ (resp. $\mathcal{R}$) the set of supernodes to which edges were added as a result of a primitive or a prefixed (resp. an indirect) query (defining the node corresponding to the key k to be contained in $\mathcal{R}$ from the beginning).

Let $\mathcal{E}_{\mathsf{path,i}}$ be the event that there is a unique path from supernode k to any other supernode reachable from it after the i^{th} edge is added. Similarly, let $\mathcal{E}_{\mathsf{clash,i}}$ be the event that $\mathcal{R} \cap \mathcal{F} = \emptyset$ after the i^{th} edge is added. Moreover, let $\mathcal{E}_{\mathsf{i}} := \mathcal{E}_{\mathsf{path,i}} \cap \mathcal{E}_{\mathsf{clash,i}}$.

Observe that when $\mathcal{E} := \mathcal{E}_{\mathsf{T}+\mathsf{q}\ell}$ occurs, the values that appear on the paths starting at k are uniformly random and independent since every node inside a supernode has the same probability of being chosen, and the distinguishing advantage is zero in this case.

Via a calculation similar to that in the proof of Theorem 5,

$$\mathsf{P}[\neg\mathcal{E}_{\mathsf{path,i}}|\mathcal{E}_{\mathsf{i}-1}] \;\le\; \frac{i+P}{2^{c-1}} \qquad \text{and} \qquad \mathsf{P}[\neg\mathcal{E}_{\mathsf{clash,i}}|\mathcal{E}_{\mathsf{i}-1}] \;\le\; \frac{i+P}{2^{c-1}},$$

and, therefore, once again,

$$\mathsf{P}[\neg\mathcal{E}] \;=\; \mathsf{P}[\neg\mathcal{E}_{\mathsf{T}+\mathsf{q}\ell}] \;=\; O\left(\frac{(T+q\ell)^2 + (T+q\ell)P}{2^c}\right).$$

□

C.3 MACs from Keyed Sponges

The application $G^{\mathsf{MAC\text{-}S}}$ of MAC security of the sponge construction is defined via the following challenger $\mathsf{C}^{\mathsf{MAC\text{-}S}}$: it initially picks a random key $k \leftarrow \{0,1\}^c$; when the attacker queries a message $m = m_1 \cdots m_\ell$ consisting of r-bit blocks m_i, the challenger answers by $\mathsf{Sponge}_{\pi,k}(m)$. The attacker wins if he submits a pair (m, y) with $\mathsf{Sponge}_{\pi,k}(m) = y$ for a previously unqueried m.

Theorem 27. *Application $G^{\mathsf{MAC\text{-}S}}$ is $\left((S,T,q,\ell), \tilde{O}\left(\frac{q}{2^r} + \frac{S(T+q\ell)^2}{2^c}\right)\right)$-secure in the $\mathsf{AI\text{-}RP}(N)$-model, where $N = 2^n = 2^{r+c} \ge 16$.*

Proof. One shows that $G^{\mathsf{MAC\text{-}S}}$ is $\left((S,T,q,\ell), O\left(\frac{q}{2^r} + \frac{(T+q\ell)^2+(T+q\ell)P}{2^c}\right)\right)$-secure in the $\mathsf{BF\text{-}RP}(P,N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{PRF\text{-}S}}} = T + q\ell$, setting $\gamma := 1/N$ and $P = 2(S+\log N)(T+q\ell) = \tilde{O}\left(S(T+q\ell)\right)$ and applying Theorem 1, the desired conclusion follows.

The bound in the $\mathsf{BF\text{-}RP}(P,N)$-model follows immediately from Lemma 26 and the fact that with a truly random function, the adversary's success probability at breaking the MAC is at most $q/2^r$. □

D Applications Based on Merkle-Damgård with Davies-Meyer: Security Proofs

This section contains the deferred definitions as well security bounds and the corresponding proofs for the applications based on the Merkle-Damgård paradigm with the Davies-Meyers compression function presented in Section 4.

In the following, $E, E^{-1} : [K] \times [N] \to [N]$ denote the ideal cipher and its inverse to which $\mathsf{AI\text{-}IC}(K, N)$ and $\mathsf{BF\text{-}IC}(P, K, N)$ offer access at interface main (cf. Section 2).

D.1 Definitions

Let $f : [N] \times [K] \to [N]$ be a compression function, $\overline{m} = m_1 \cdots m_\ell$ be a message consisting of ℓ blocks $m_i \in [K]$,[5] and $\mathsf{IV} \in [N]$. The Merkle-Damgård (MD) transform $\mathsf{MD}_{f,\mathsf{IV}}(\overline{m})$ works as follows:

1. Set $h_0 \leftarrow \mathsf{IV}$.
2. For $i = 1, \ldots, \ell$: set $h_i \leftarrow f(h_{i-1}, m_i)$.
3. Output h_ℓ.

Recall that the Davies-Meyer compression function DM_E is defined as

$$\mathsf{DM}_E(h, m) \;:=\; E(m, h) \oplus h\,.$$

Finally, define the MD transform with the Davies-Meyer compression function (MD-DM) as

$$\mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(\overline{m}) \;:=\; \mathsf{MD}_{\mathsf{DM}_E,\mathsf{IV}}(\overline{m})\,.$$

D.2 CRHFs from Merkle-Damgård with Davies-Meyer

The application $G^{\mathsf{CRHF\text{-}MD}}$ of collision-resistance for the MD-DM construction is defined via the following challenger $\mathsf{C}^{\mathsf{CRHF\text{-}MD}}$: it picks an initialization vector $\mathsf{IV} \leftarrow [N]$ uniformly at random, passes it to the attacker, and outputs 1 if and only if the attacker returns two messages $m \neq m'$ such that $\mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(\overline{m}) = \mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(\overline{m}')$.

The following theorem provides an upper bound on the probability that an (S, T, ℓ)-attacker finds a collision of the Merkle-Damgård construction in the AI-ICM, where ℓ is an upper bound on the lengths of the messages m and m' the attacker submits to the challenger. The proof follows the approach by [10].

Theorem 28. *Application $G^{\mathsf{CRHF\text{-}MD}}$ is $((S, T, \ell), \varepsilon)$-secure in the $\mathsf{AI\text{-}IC}(K, N)$-model for $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\frac{S(T+\ell)^2}{N} + \frac{(T+\ell)^2}{N}\right).$$

Proof. At the cost of at most 2ℓ additional queries to E, assume that the messages $\overline{m}$ and $\overline{m}'$ output by $\mathcal{A}_2$ are such that $\mathcal{A}_2$ has queried its oracle $\mathcal{O}$ on all points $\mathsf{MD\text{-}DM}_{E,\mathsf{IV}}(\cdot)$ would query E when evaluated on $\overline{m}$ and $\overline{m}'$. In the following, let $T' := T + 2\ell$. It suffices to show that $G^{\mathsf{CRHF\text{-}MD}}$ is $\left((S, T, \ell), O\left(\frac{(T+\ell)^2+(T+\ell)P}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}IC}(P, K, N)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{CRHF\text{-}MD}}} = T+2\ell$, setting $\gamma := 1/N$ and $P = 2(S+\log N)(T+2\ell) = \tilde{O}\left(S(T+\ell)\right)$ and applying Theorem 1, the desired conclusion follows.

[5] To keep things simple, no padding is considered here.

Call h *reachable* if $\mathcal{A}_2$'s query/answer list contains values

$$((m_0, h_0), y_0), \ldots, ((m_{j-1}, h_{j-1}), y_{j-1})$$

such that

- $h_0 = \mathsf{IV}$,
- $h_j = y_{j-1} \oplus h_{j-1}$ for $i = 1, \ldots, j-1$, and
- $y_{j-1} \oplus h_{j-1} = h$.

Furthermore, h is said to be *fresh* if there exists no prefixed query/answer pair $((\cdot, h), \cdot)$. Let $\mathcal{E}_{\mathsf{fresh,i}}$ be the event that after the i^{th} query by $\mathcal{A}_2$, all reachable h are fresh, and set $\mathcal{E}_{\mathsf{fresh}} := \mathcal{E}_{\mathsf{fresh,T'}}$. Observe that

$$\mathsf{P}[\neg\mathcal{E}_{\mathsf{fresh,i}} | \mathcal{E}_{\mathsf{fresh,i-1}}] \;\leq\; \frac{\max(P,i)}{N - (P + (i-1))} \;\leq\; \frac{\max(P,i)}{N - (P + T')} \;\leq\; \frac{2\max(P,i)}{N} \;,$$

where the numerator is P for forward queries and i for inverse queries, and using $P + T' \leq N/2$, an assumption one may always make since, otherwise, the lemma holds trivially. Therefore,

$$\begin{aligned}
\mathsf{P}[\neg\mathcal{E}_{\mathsf{fresh}}] \;&=\; \sum_{i=1}^{T'} \mathsf{P}[\neg\mathcal{E}_{\mathsf{fresh,i}} | \mathcal{E}_{\mathsf{fresh,i-1}}] \\
&\leq\; \frac{2(T'^2 + T'P)}{N} \\
&\leq\; O\left(\frac{(T+\ell)^2 + (T+\ell)P}{N}\right) \;.
\end{aligned}$$

Let $\mathcal{E}_{\mathsf{DM\text{-}coll,i}}$ be the event that the first i queries by $\mathcal{A}_2$ include

1. (m, h) and (m', h') such that

   $$E(m,h) \oplus h \;=\; E(m',h') \oplus h'$$

   and h and h' are fresh, *or*

2. (m, h) such that $E(m,h) \oplus h = \mathsf{IV}$ and h is fresh.

Let $\mathcal{E}_{\mathsf{DM\text{-}coll}} := \mathcal{E}_{\mathsf{DM\text{-}coll,q}}$. Observe that, *conditioned on* $\mathcal{E}_{\mathsf{fresh}}$, if $\mathcal{A}_2$ finds a collision, then $\mathcal{E}_{\mathsf{DM\text{-}coll}}$ happens. Hence, it remains to bound

$$\begin{aligned}
\mathsf{P}[\mathcal{E}_{\mathsf{DM\text{-}coll}} | \mathcal{E}_{\mathsf{fresh}}] \;=\; \mathsf{P}[\mathcal{E}_{\mathsf{DM\text{-}coll,T'}} | \mathcal{E}_{\mathsf{fresh}}] \;&\leq\; \sum_{i=1}^{T'} \mathsf{P}[\mathcal{E}_{\mathsf{DM\text{-}coll,i}} | \neg\mathcal{E}_{\mathsf{DM\text{-}coll,i-1}}, \mathcal{E}_{\mathsf{fresh}}] \\
&\leq\; \sum_{i=1}^{T'} \frac{i-1}{N - (P + i - 1)} \\
&\leq\; \sum_{i=1}^{T'} \frac{i}{N - (P + T')} \\
&\leq\; O\left(\frac{(T+\ell)^2}{N}\right).
\end{aligned}$$

□

E GGM Applications: Security Proofs

While Section 5 discusses the DL problem in the AI-GGM in detail, it defers the precise definitions of the remaining applications presented there as well as the theorems with the corresponding security bounds and their proofs. These are provided here. The proofs make use of the well-known Schwartz-Zippel Lemma (cf. Lemma 37 in Section G of the appendix).

E.1 The Computational Diffie-Hellman Problem

The application G^{CDH} corresponding to the computational Diffie-Hellman (CDH) problem is defined via the challenger $\mathsf{C}^{\mathsf{CDH}}$ that picks $x, y \in [N]$ uniformly at random and independently, passes $\sigma(x), \sigma(y)$ to the attacker, and outputs 1 if and only if the attacker returns $s = \sigma(xy)$.

Theorem 29 below provides an upper bound on the success probability of any attacker at solving the CDH problem in the AI-GG-model.

Theorem 29. *G^{CDH} is $((S,T),\varepsilon)$-secure in the AI-GG(N,M)-model for any prime $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\frac{ST^2}{N} + \frac{T^2}{N}\right).$$

Proof. It suffices to show that the application G^{CDH} is $\left((S,T), O\left(\frac{TP+T^2}{N}\right)\right)$-secure in the BF-GG$(P, N, M)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{CDH}}} = T+3$, setting $\gamma := 1/N$ and $P = 6(S+\log N)(T+3) = \tilde{O}(ST)$, and applying the second part of Theorem 1, the desired conclusion follows. One may assume that $P + T \leq N/2$ since otherwise the statement in the BF-GG model holds trivially.

Consider now the interaction of $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ with $\mathsf{C}^{\mathsf{CDH}}$ in the BF-GG-model. Recall that the BF-GG-oracle outputs the range $\mathcal{Y}$ of the underlying random injection σ to $\mathcal{A}_1$ via interface pre. Condition on a particular realization of this set for the remainder of the proof.

Define the following hybrid experiment involving $\mathcal{A}_1$ and $\mathcal{A}_2$:

- For each of the at most P query/answer pairs (a', s') where $\mathcal{A}_1$ fixes σ, define a (constant) polynomial $v(X,Y) := a'$ and store the pair (v, s').
- To create the challenge, choose values s_1^*, s_2^*, and s_3^* uniformly at random and independently from all unused values in $\mathcal{Y}$, define the polynomials $u_1^*(X,Y) := X$, $u_2^*(X,Y) := Y$, and $u_3^*(X,Y) := XY$, and store the pairs (u_1^*, s_1^*), (u_2^*, s_2^*), and (u_3^*, s_3^*).
- A forward query a by $\mathcal{A}_2$ to BF-GG is answered as follows: define the (constant) polynomial $u(X,Y) := a$, choose a value s uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u, s), and return s.
- A group-operation query (s_1, s_2) by $\mathcal{A}_2$ is answered as follows:
  - If s_1 or s_2 is not in $\mathcal{Y}$, return $\perp$.
  - If s_1 has not been recorded, choose a random unused $a \in [N]$, define the (constant) polynomial $u(X,Y) := a$, and store the pair (u, a). Proceed similarly if s_2 has not been recorded. Go to the next item.
  - Let u_1 and u_2 be the polynomials recorded with s_1 and s_2, respectively. If, for $u' := u_1 + u_2$, a pair (u', s') has been recorded, return s'. Otherwise, choose a value s' uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u', s'), and return s'.

- When $\mathcal{A}_2$ outputs a value s', pick values $x, y \in [N]$ uniformly at random and independently and output 1 if and only if $s' = s_3^*$.

Observe that the hybrid experiment only differs from the original one if for a group-operation query (s_1, s_2), $u'(x, y) = v(x, y)$ for some recorded v or $u'(x, y) = u(x, y)$ for some recorded u—and similarly for the polynomials u_1^*, u_2^*, and u_3^* corresponding to the challenge. Since in the hybrid experiment, x and y are chosen uniformly at random *at the end* of the execution, the probability of this event is at most $2((T + 3)P + (T + 3)^2)/N$ by the Schwartz-Zippel lemma and a union bound. Since in the hybrid experiment, in the view of $\mathcal{A}$, s_3^* is distributed uniformly over the value not in the presampled encodings or seen as a reply by the oracle, the probability that $s' = s_3^*$ is $1/(N - (P + T))$. The theorem follows. □

E.2 The Decisional Diffie-Hellman Problem

The application G^{DDH} corresponding to the decisional Diffie-Hellman (DDH) problem is defined via the challenger $\mathsf{C}^{\mathsf{DDH}}$ that picks $x, y, z \in [N]$ and a bit $b \in \{0, 1\}$ uniformly at random and independently, passes $\sigma(x), \sigma(y), \sigma(xy)$ or $\sigma(x), \sigma(y), \sigma(z)$ to the attacker if $b = 0$ or if $b = 1$, respectively, and outputs 1 if and only if the attacker returns $b' = b$.

Theorem 30 below provides an upper bound on the advantage of any attacker for the DDH problem in the AI-GG-model.

Theorem 30. *G^{DDH} is $((S, T), \varepsilon)$-secure in the AI-GG(N, M)-model for any prime $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\sqrt{\frac{ST^2}{N}} + \frac{T^2}{N}\right).$$

Proof. It suffices to show that the application G^{DDH} is $\left((S, T), O\left(\frac{TP+T^2}{N}\right)\right)$-secure in the BF-GG$(P, N, M)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{DDH}}} = T + 4$, setting $\gamma := 1/N$ and $P = \sqrt{SN}$, and applying the first part of Theorem 1, the desired conclusion follows.

In the following, one may assume that $P + T \leq N/2$ since otherwise the statement in the BF-GG model holds trivially. Consider now the interaction of $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ with $\mathsf{C}^{\mathsf{DDH}}$ in the BF-GG-model. Recall that the BF-GG-oracle outputs the range $\mathcal{Y}$ of the underlying random injection σ to $\mathcal{A}_1$ via interface pre. Condition on a particular realization of this set for the remainder of the proof.

Define the following hybrid experiment involving $\mathcal{A}_1$ and $\mathcal{A}_2$:

- For each of the at most P query/answer pairs (a', s') where $\mathcal{A}_1$ fixes σ, define a (constant) polynomial $v(X, Y, Z) := a'$ and store the pair (v, s').
- To create the challenge, choose $b \in \{0, 1\}$ and values s_1^*, s_2^*, s_3^*, and s_4^* uniformly at random and independently from all unused values in $\mathcal{Y}$, define the polynomials $u_1^*(X, Y, Z) := X$, $u_2^*(X, Y, Z) := Y$, $u_3^*(X, Y, Z) := XY$, and $u_4^*(X, Y, Z) := Z$, and store the pairs (u_1^*, s_1^*), (u_2^*, s_2^*), (u_3^*, s_3^*), and (u_4^*, s_4^*).
- A forward query a by $\mathcal{A}_2$ to BF-GG is answered as follows: define the (constant) polynomial $u(X, Y, Z) := a$, choose a value s uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u, s), and return s.
- A group-operation query (s_1, s_2) by $\mathcal{A}_2$ is answered as follows:

- If s_1 or s_2 is not in $\mathcal{Y}$, return $\perp$.
- If s_1 has not been recorded, choose a random unused $a \in [N]$, define the (constant) polynomial $u(X, Y, Z) := a$, and store the pair (u, a). Proceed similarly if s_2 has not been recorded. Go to the next item.
- Let u_1 and u_2 be the polynomials recorded with s_1 and s_2, respectively. If, for $u' := u_1 + u_2$, a pair (u', s') has been recorded, return s'. Otherwise, choose a value s' uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u', s'), and return s'.

- When $\mathcal{A}_2$ outputs a bit b', pick values $x, y, z \in [N]$ uniformly at random and independently and output 1 if and only if $b' = b$.[6]

Observe that the hybrid experiment only differs from the original one if for a group-operation query (s_1, s_2), $u'(x, y, z) = v(x, y, z)$ for some recorded v or $u'(x, y, z) = u(x, y, z)$ for some recorded u—and similarly for the polynomials u_1^*, u_2^*, u_3^*, and u_4^* corresponding to the challenge. Since in the hybrid experiment, x and y are chosen uniformly at random *at the end* of the execution, the probability of this event is at most $2((T+4)P + (T+4)^2)/N$ by the Schwartz-Zippel lemma and a union bound.

All polynomials created by $\mathcal{A}$ during the attack are *linear* in x, y, and xy respectively z, depending on the value of b. Since, therefore, xy can be substituted by a variable z' without affecting the view of $\mathcal{A}$, $\mathcal{A}$ obtains no information about b, and its advantage is zero. □

E.3 The Square Decisional Diffie-Hellman Problem

The application $G^{\mathsf{sq\text{-}DDH}}$ corresponding to the square decisional Diffie-Hellman (sqDDH) problem is defined via the challenger $\mathsf{C}^{\mathsf{sq\text{-}DDH}}$ that picks $x, y, z \in [N]$ and a bit $b \in \{0, 1\}$ uniformly at random and independently, passes $\sigma(x), \sigma(y), \sigma(xy)$ or $\sigma(x), \sigma(y), \sigma(z)$ to the attacker if $b = 0$ or if $b = 1$, respectively, and outputs 1 if and only if the attacker returns $b' = b$.

Theorem 31 below provides an upper bound on the advantage of any attacker for the sqDDH problem in the $\mathsf{AI\text{-}GG}$-model.

Theorem 31. *$G^{\mathsf{sq\text{-}DDH}}$ is $((S, T), \varepsilon)$-secure in the $\mathsf{AI\text{-}GG}(N, M)$-model for any prime $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\sqrt{\frac{ST^2}{N}} + \frac{T^2}{N}\right).$$

The proof is very similar to that of the DDH problem and is therefore omitted.

E.4 The One-More-Discrete-Logarithm Problem

The application $G^{\mathsf{OM\text{-}DL}}$ corresponding to the one-more-discrete-logarithm (OM-DL) problem is defined via the following challenger $\mathsf{C}^{\mathsf{OM\text{-}DL}}$: The attacker $\mathcal{A}_2$ is allowed to make queries to a challenge oracle; $\mathsf{C}^{\mathsf{OM\text{-}DL}}$ answers the i^{th} query to this oracle by choosing a value $x_i \in [N]$ uniformly at random and outputting $\sigma(x_i)$ to $\mathcal{A}_2$. Moreover, $\mathcal{A}_2$ is given access to a discrete-logarithm oracle. At the end of the interaction, $\mathsf{C}^{\mathsf{OM\text{-}DL}}$ outputs 1 if and only if the attacker makes t queries to the challenge oracle, at most $t - 1$ queries to the discrete-logarithm oracle, and outputs $(x'_1, \ldots, x'_t) = (x_1, \ldots, x_t)$.

[6] The values x, y, and z chosen at the end of the game are only used for the analysis.

Theorem 32 below provides an upper bound on the success probability of any attacker for the OM-DL problem in the AI-GG-model. The number t queries to the discrete-logarithm oracle is captured as an additional parameter of $\mathcal{A}_1$.

Theorem 32. *$G^{\mathsf{OM\text{-}DL}}$ is $((S,T,t),\varepsilon)$-secure in the $\mathsf{AI\text{-}GG}(N,M)$-model for any prime $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\frac{S(T+t)^2}{N} + \frac{(T+t)^2}{N}\right).$$

Proof. It suffices to show that $G^{\mathsf{OM\text{-}DL}}$ is $\left((S,T), O\left(\frac{(T+t)P+(T+t)^2}{N}\right)\right)$-secure in the $\mathsf{BF\text{-}GG}(P,N,M)$-model. Then, by observing that $T^{\mathsf{comb}}_{G^{\mathsf{OM\text{-}DL}}} = T+t$, setting $\gamma := 1/N$ and $P = 6(S+\log N)(T+t) = \tilde{O}\left(S(T+t)\right)$, and applying the second part of Theorem 1, the desired conclusion follows. One may assume that $P+T \leq N/2$ since otherwise the statement in the BF-GG model holds trivially.

Consider now the interaction of $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ with $\mathsf{C}^{\mathsf{OM\text{-}DL}}$ in the BF-GG-model. Recall that the BF-GG-oracle outputs the range $\mathcal{Y}$ of the underlying random injection σ to $\mathcal{A}_1$ via interface pre. Condition on a particular realization of this set for the remainder of the proof.

Let $\overline{X} = (X_1, \ldots, X_t)$ be variables and define the following hybrid experiment involving $\mathcal{A}_1$ and $\mathcal{A}_2$:

- For each of the at most P query/answer pairs (a', s') where $\mathcal{A}_1$ fixes σ, define a (constant) polynomial $v(\overline{X}) := a'$ and store the pair (v, s').
- The i^{th} request for a challenge is answered as follows: choose s_i^* uniformly at random and independently from all unused values in $\mathcal{Y}$, define the polynomial $u_i^*(\overline{X}) := X_i$, and store the pair (u_i^*, s_i^*).
- A forward query a by $\mathcal{A}_2$ to BF-GG is answered as follows: define the (constant) polynomial $u(\overline{X}) := a$, choose a value s uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u, s), and return s.
- A discrete-logarithm query s by $\mathcal{A}_2$ is answered as follows:
  - If s is not in $\mathcal{Y}$, return $\perp$.
  - If s has not been recorded, choose a random unused $a \in [N]$, define the (constant) polynomial $u(\overline{X}) := a$, store the pair (u, a), and return a.
  - Otherwise, let u be the polynomial recorded with s. Choose a random unused $a \in [N]$, record (u, a), and return a.
- A group-operation query (s_1, s_2) by $\mathcal{A}_2$ is answered as follows:
  - If s_1 or s_2 is not in $\mathcal{Y}$, return $\perp$.
  - If s_1 has not been recorded, choose a random unused $a \in [N]$, define the (constant) polynomial $u(\overline{X}) := a$, and store the pair (u, a). Proceed similarly if s_2 has not been recorded. Go to the next item.
  - Let u_1 and u_2 be the polynomials recorded with s_1 and s_2, respectively. If, for $u' := u_1 + u_2$, a pair (u', s') has been recorded, return s'. Otherwise, choose a value s' uniformly at random from all unused values in $\mathcal{Y}$, store the pair (u', s'), and return s'.

- When $\mathcal{A}_2$ outputs a vector $\overline{x}' = (x'_1, \ldots, x'_t)$, pick a vector uniformly at random from the set of $(x_1, \ldots, x_t) \in [N]^t$ satisfying the at most $t-1$ linear equations defined by the pairs (u, a) recorded when answering discrete-logarithm queries, and output 1 if and only if $(x'_1, \ldots, x'_t) = (x_1, \ldots, x_t)$.

Observe that the hybrid experiment only differs from the original one if for a group-operation query (s_1, s_2), $u'(\overline{x}) = v(\overline{x})$ for some recorded v or $u'(\overline{x}) = u(\overline{x})$ for some recorded u—and similarly for the polynomials u_i^* corresponding to the challenges. Since in the hybrid experiment, the vector $\overline{x}$ is chosen uniformly at random from a set of size at least N *at the end* of the execution, the probability of this event is at most $2((T+t)P + (T+t)^2)/N$ by the Schwartz-Zippel lemma and a union bound. Moreover, in the hybrid experiment, the probability of $\overline{x}' = \overline{x}$ is at most $1/N$. □

E.5 Multiple Discrete Logarithms

The multiple-discrete-logarithms application $G^{\mathsf{MDL}}_{N,t}$ is defined via the challenger $\mathsf{C}^{\mathsf{MDL}}_{N,t}$ that randomly and uniformly picks a vector $\overline{x} = (x_1, \ldots, x_t) \in [N]$, passes $\sigma(x_1), \ldots, \sigma(x_t)$ to the attacker, and outputs 1 if and only if the attacker returns x.

Theorem 33 below provides an upper bound on the success probability of any attacker solving the multiple-discrete-logarithms problem in the $\mathsf{AI\text{-}GG}$-model.

Theorem 33. *$G^{\mathsf{MDL}}_{N,t}$ is $((S,T),\varepsilon)$-secure in the $\mathsf{AI\text{-}GG}(N,M)$-model for any prime $N \geq 16$ and*

$$\varepsilon \;=\; \tilde{O}\left(\frac{S(T+t)^2}{tN} + \frac{(T+t)^2}{tN}\right)^t.$$

The proof of Theorem 33 follows the corresponding proof by Yun [53] in the GGM without preprocessing. The generic hardness of computing many discrete logarithms is derived from the hardness of the so-called *search-by-hyperplane-queries (SHQ)* application $G^{\mathsf{SHQ}}_{N,t}$, which is defined via the following challenger $\mathsf{C}^{\mathsf{SHQ}}_{N,t}$ *in the standard model*: $\mathsf{C}^{\mathsf{SHQ}}_{N,t}$ initially chooses $\overline{\alpha} = (\alpha_1, \ldots, \alpha_t) \in [N]^t$ uniformly at random. Thereafter, the attacker $\mathcal{B}$ is allowed to make up to q *hyperplane queries* by submitting the description of a hyperplane H and receiving the answer 1 if $\overline{\alpha}$ is on H and 0 otherwise. $\mathsf{C}^{\mathsf{SHQ}}_{N,t}$ outputs 1 in the end if and only if $\mathcal{B}$ submits $\overline{\alpha}'$ with $\overline{\alpha}' = \overline{\alpha}$. Yun established the following theorem:[7]

Theorem 34 ([53])**.** *Application $G^{\mathsf{SHQ}}_{N,t}$ is (q,ε)-secure in the standard model for*

$$\varepsilon \;=\; \frac{1}{N^t} + \left(\frac{eq}{tN}\right)^t.$$

At the heart of the proof of Theorem 33 is a generalization of Yun's reduction from SHQ to MDL. It accommodates attackers in the BF-ROM and is captured by the following lemma:

Lemma 35. *For any (S,T)-attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ in the $\mathsf{BF\text{-}GG} := \mathsf{BF\text{-}GG}(P,N,M)$-model, there exists a q-attacker $\mathcal{B}$ such that*

$$\mathrm{Adv}_{G^{\mathsf{SHQ}}_{N,t}}(\mathcal{B}) \;\geq\; \mathrm{Adv}_{G^{\mathsf{MDL}}_{N,t},\mathsf{BF\text{-}GG}}(\mathcal{A})$$

and $q \leq (T+t)P + (T+t)^2$.

[7] Since SHQ is a standard-model application, S and T are omitted, and the only parameter considered for SHQ attackers is q. For the same reason, SHQ attacker only have a single phase $\mathcal{B}$.

Proof. The proof is similar in spirit to that of the DL application and uses polynomials $u(\overline{X})$ in variables $\overline{X} = (X_1, \ldots, X_t)$ to track the expressions computed by $\mathcal{A}_2$ using the group-operation oracle. However, instead of relying on the collision probability of each pair of polynomials on the actual values $x_1, \ldots, x_t$, the idea here is to query the SHQ oracle to determine collisions, where the values $\overline{\alpha} = (\alpha_1, \ldots, \alpha_t)$ play the role of $\overline{x} = (x_1, \ldots, x_t)$.

Based on $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, consider the following attacker $\mathcal{B}$:

1. Choose a size-N subset $\mathcal{Y}$ of $[M]$ uniformly at random and pass it to $\mathcal{A}_1$. Obtain a list $\mathcal{L}$ from $\mathcal{A}_1$ of P prefixed input/output pairs compatible with $\mathcal{Y}$. For each pair $(a', s') \in \mathcal{L}$, define the polynomial $v(\overline{X}) := a'$ and store the pair (v, s').

2. For $i = 1, \ldots, t$, create each element s_i^* of the challenge $(s_1^*, \ldots, s_t^*)$ for $\mathcal{A}_2$ as follows: Define the polynomial $u_i^*(\overline{X}) := X_i$ and query the SHQ oracle on the hyperplane $u_i^* = u$ for *every* polynomial u previously recored as (u, s) for some s. If the answer is 1, set $s_i^* := s$; otherwise, choose a random unused value s_i^* from $\mathcal{Y}$. Record the pair (u_i^*, s_i^*).

3. When $\mathcal{A}_2$ makes a forward query a, proceed as follows: Define the polynomial $u(\overline{X}) := a$. For each (non-constant) polynomial previously recorded as (u', s') for some s', query the SHQ oracle on the hyperplane $u = u'$. If the answer is 1, record the pair (u, s'), return s'; otherwise, choose a random unused value s from $\mathcal{Y}$, record the pair (u, s), and return s.

4. When $\mathcal{A}_2$ makes a group operation query (s_1, s_2), proceed as follows:

   - If s_1 or s_2 is not in $\mathcal{Y}$, return $\perp$.
   - If s_1 has not been recorded, choose a random unused $a \in [N]$ such that, for $u(\overline{X}) := a$, $u = u'$ returns 0 for all previously recorded (non-constant) polynomials. Record (u, a). Proceed similarly if s_2 has not been recorded. Go to the next item.
   - Let u_1 and u_2 be the polynomials recorded with s_1 and s_2, respectively. If, for $u := u_1 + u_2$ query the SHQ oracle on the hyperplane $u = u'$ for all polynomials previously recorded as (u', s') for some s'. If the answer is 1, record (u, s'), and return s'; otherwise, choose a random unused value s from $\mathcal{Y}$, record the pair (u, s), and return s.

5. When $\mathcal{A}_2$ outputs a value $\overline{\alpha}'$, return $\overline{\alpha}'$ to $\mathsf{C}_{N,t}^{\mathsf{SHQ}}$.

It is readily verified that $\mathcal{B}$ perfectly simulates the experiment with $\mathsf{C}_{N,t}^{\mathsf{MDL}}$ for $\mathcal{A}_2$. □

Proof of Theorem 10. Lemma 35 implies that application $G_{N,t}^{\mathsf{MDL}}$ is $((S, T), \varepsilon')$-secure in the $\mathsf{BF\text{-}GG}(P, N, M)$-model for

$$\varepsilon' \;=\; \frac{1}{N^t} + \left(\frac{e(T+t)P + (T+t)^2}{tN}\right)^t.$$

Then, by observing that $T_{G_{N,t}^{\mathsf{MDL}}}^{\mathsf{comb}} = T + t$, setting $\gamma := 1/N$ and $P = 6(S + \log N)(T + t) = \tilde{O}\left(S(T+t)\right)$, and applying the second part of Theorem 1, the desired conclusion follows. □

E.6 Knowledge-of-Exponent Assumption

Informally, the knowledge-of-exponent assumption (KEA) states that if an attacker $\mathcal{A}$ is given (h, h^x), for a generator h of a cyclic group of order N and $x \in [N]$ chosen uniformly at random, and outputs group elements A and $\hat{A}$ with $\hat{A} = A^x$, then it must know discrete logarithm a of A.

This is formalized by requiring that for every $\mathcal{A}$ there exist an *extractor* $\mathcal{X}_{\mathcal{A}}$ that is run on the same random coins as $\mathcal{A}$ and must output the value a.

The above is captured in the GGM by considering the following experiment $\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_{\mathcal{A}}}$ parameterized by an attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, an extractor $\mathcal{X}_{\mathcal{A}}$, and an oracle $\mathcal{O} \in \{\mathsf{AI\text{-}GG}(N, M), \mathsf{BF\text{-}GG}(N, M)\}$:

1. Run $\mathcal{A}_1$ to obtain $z \leftarrow \mathcal{A}_1^{\mathcal{O}}$.
2. Choose $x \in [N]$ uniformly at random, let $y \leftarrow \sigma(x)$, pick random coins ρ, and run
   (a) $\mathcal{A}_2$ to get $(A, \hat{A}) \leftarrow \mathcal{A}_2(z, y; \rho)$, and
   (b) $\mathcal{X}_{\mathcal{A}}$ to get $a \leftarrow \mathcal{X}_{\mathcal{A}}(z, y; \rho)$.
3. Output 1 if and only if $A = \sigma(a')$ and $\hat{A} = \sigma(a'x)$ for some a', but $a \neq a'$.

The KEA says that for every attacker $\mathcal{A}$ there exists an extractor $\mathcal{X}_{\mathcal{A}}$ such that the probability of the above experiment outputting 1 is negligible. The following theorem is equivalent to saying that the KEA holds in the AI-GGM.

Theorem 36. *For every attacker $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, there exists an extractor $\mathcal{X}_{\mathcal{A}}$ such that*

$$\mathsf{P}[\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_{\mathcal{A}}} = 1] \;\leq\; \tilde{O}\left(\frac{ST^2}{N}\right).$$

Sketch. The extractor $\mathcal{X}_{\mathcal{A}}$ internally runs $\mathcal{A}_2$ on the inputs received and keeps track of $\mathcal{A}_2$'s oracle queries using polynomials as in the proof of Theorem 10. If at the end the polynomials u_A and $u_{\hat{A}}$ corresponding to $\mathcal{A}_2$'s outputs $(A, \hat{A})$ have the form $u_A(X) = a$ and $u_{\hat{A}}(X) = aX$, then $\mathcal{X}_{\mathcal{A}}$ outputs a and otherwise $\perp$.

Observe that if the experiment outputs 1, then

- $u_{\hat{A}} \neq X \cdot u_A$ since $\mathcal{A}_2$ only creates polynomials of degree at most 1, but
- $u_{\hat{A}}(x) = x \cdot u_A(x)$ for the challenge x.

Hence, the extractor only fails if at least two of the polynomials involved (including $u_{\hat{A}}$ and $X \cdot u_A$) collide on x, which is already analyzed in the proof of Theorem 10.

The experiment $\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_{\mathcal{A}}}$ defining KEA does not exactly match the syntax of challenger and attacker to which Theorem 1 caters, but it is easily checked that the corresponding proof can be adapted to fit $\mathrm{Exp}^{\mathcal{O}}_{\mathcal{A},\mathcal{X}_{\mathcal{A}}}$. □

F Computational Applications: Security Proofs

Trapdoor functions. The inversion problem for a trapdoor function generator F can be phrased as an application $G^{\mathsf{TDF},F}$, defined via the challenger $\mathsf{C}^{\mathsf{TDF},F}$ that generates $(f, f^{-1}) \leftarrow F$, picks a random x, passes $y := f(x)$ to the attacker, and outputs 1 if and only if the attacker finds x. Observe that $G^{\mathsf{TDF},F}$ is a *standard-model* application.

Public-key encryption schemes. A *public-key encryption (PKE) scheme* is a triple of algorithms $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$, where Gen generates a public key pk and a secret key sk, Enc takes a public key pk and a message m and outputs a ciphertext c, and Dec takes a secret key sk and a ciphertext c and outputs a message m. In the $\mathcal{O}$-oracle model, all three algorithms may make calls to $\mathcal{O}.\mathsf{main}$.

The application of public-key encryption $G^{\mathsf{PKE},\Pi}$ is defined via the following challenger $\mathsf{C}^{\mathsf{PKE},\Pi}$, which captures the (standard) CPA security of a public-key encryption scheme: Initially, $\mathsf{C}^{\mathsf{PKE},\Pi}$ generates a key pair $(\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{Gen}$ and passes pk to the attacker. Then, the attacker submits two messages m_0 and m_1 tho the challenger, who answers by choosing a random bit b and returning the challenge $c^* \leftarrow \mathsf{Enc}_{\mathsf{pk}}(m_b)$. In the end, the challenger outputs 1 if and only if the attacker submits a bit b' with $b' = b$.

FDP encryption. Let F be a trapdoor family (TDF) generator. Full-domain permutation (FDP) encryption in the random-permutation model with oracle $\mathcal{O}$ is defined as follows:

- **Key generation:** Run the TDF generator to obtain $(f, f^{-1}) \leftarrow F$, where $f, f^{-1} : [N] \to [N]$. Set the public key $\mathsf{pk} := f$ and the secret key $\mathsf{sk} := f^{-1}$.
- **Encryption:** To encrypt a message m with randomness r and public key $\mathsf{pk} = f$, compute $\tilde{y} \leftarrow f(y)$ for $y \leftarrow \mathcal{O}(m\|r)$ and output $c = \tilde{y}$.
- **Decryption:** To decrypt a ciphertext $c = y$ with secret key $\mathsf{sk} = f^{-1}$, compute $m\|r \leftarrow \mathcal{O}^{-1}(f^{-1}(y))$ and output m.

Theorem 12. *Let Π be FDP encryption with F. If $G^{\mathsf{TDF},F}$ is $((S', *, t', s'), \varepsilon')$-secure, then, for any $T \in \mathbb{N}$, $G^{\mathsf{PKE},\Pi}$ is $((S, T, t, s), \varepsilon)$-secure in the $\mathsf{AI\text{-}RP}(N, N)$-model, where*

$$\varepsilon \;=\; \tilde{O}\left(\varepsilon' + \sqrt{\frac{ST}{2^\rho}}\right)$$

and $S = S' - \tilde{O}(ST)$, $t = t' - \tilde{O}(t_{\mathsf{tdf}} \cdot T)$, and $s = s' - \tilde{O}(ST)$, where t_{tdf} is the time required to evaluate the TDF.

Sketch. Let π, π^{-1} be the permutation and its inverse $\mathsf{AI\text{-}RP}(N)$ offers access to. Moreover, let ρ be the length of r. Consider the interaction of $\mathcal{A}$ and $\mathsf{C}^{\mathsf{PKE},\Pi}$ in the $\mathsf{AI\text{-}RP}(N)$-model and consider a hybrid experiment in which an inverse query y to π^{-1} is answered by $\perp$ instead of $m\|r$.

Clearly, the original experiment and the hybrid behave identically unless $\mathcal{A}_2$ queries y to π^{-1}. Using the unpredictability pre-sampling, an upper bound ν on the probability of this event in the $\mathsf{BF\text{-}RP}(P', N)$-model for $P' \approx ST$ yields a bound of roughly 2ν in the $\mathsf{AI\text{-}RP}(N)$-model.

To obtain ν, there is a straight-forward reduction to the security of F that fails with probability $O(P'(2^{-\rho} + N^{-1}))$, namely if $m\|r$ or y as generated by the TDF challenger fall into the prefixed list $\mathcal{L}$ or if $m\|r$ is queried by the attacker. The reduction must handle $\mathcal{L}$, and hence requires additional space in the order of $O(ST)$. Hence, $\nu = O(P' \cdot 2^{-\rho} + \varepsilon')$.

The advantage of $\mathcal{A}$ in the hybrid experiment can again be analyzed in the $\mathsf{BF\text{-}RP}(P, N)$-model, where it is easily seen to be $O(P \cdot 2^{-\rho})$. Translating this back into the $\mathsf{AI\text{-}RP}(N)$-model, incurs an error of roughly ST/P. Choosing $P = \Theta(\sqrt{ST \cdot 2^\rho})$ results in final security $O(\sqrt{ST/2^\rho})$. □

G Miscellaneous

Lemma 37 (Schwartz-Zippel)**.** *Let $u \in \mathbb{F}[X_1, \ldots, X_t]$ be a polynomial of total degree $d \geq 0$ over a field $\mathbb{F}$. Moreover, let S be a finite subset of $\mathbb{F}$, and let $x_1, \ldots, x_t$ be chosen from S uniformly at random and independently. Then,*

$$\mathsf{P}[F(x_1, \ldots, x_t) = 0] \;\leq\; \frac{d}{|S|} \;.$$

Proposition 38. *$N^{\underline{j}} \geq (N/e)^j$, where $a^{\underline{b}} := a!/(a-b)!$.*

Proof. By Stirlings approximation, for any positive integer n,

$$\sqrt{2\pi} \cdot n^{n+1/2} \cdot e^{-n} \;<\; n! \;<\; \sqrt{2\pi} \cdot n^{n+1/2} \cdot e^{-n} \cdot e^{\frac{1}{12n}} \;.$$

The cases $N = 1$ or $j = N$ or $j = 0$ are trivial. Suppose $N \geq 2$ and $1 \leq j \leq N-1$. Then,

$$N^{\underline{j}} \;>\; \frac{\sqrt{2\pi}}{\sqrt{2\pi}} \cdot \frac{N^{N+1/2}}{(N-j)^{N-j+1/2}} \cdot \frac{e^{-N}}{e^{-(N-j)}} \cdot e^{-\frac{1}{12(N-j)}} \;. \tag{7}$$

The right-hand side of (7) can be simplied to

$$\left(\frac{N}{N-j}\right)^{N-|I|+1/2} \cdot \left(\frac{N}{e}\right)^j \cdot e^{-\frac{1}{12(N-j)}} \;.$$

Note that, for $j \leq N-1$,

$$\ln\left(\frac{N}{N-j}\right) \;=\; -\ln\left(1 - \frac{j}{N}\right) \;=\; \sum_{i=1}^{\infty} \left(\frac{j}{N}\right)^i \cdot \frac{1}{i!} \;\geq\; \frac{j}{N} \;,$$

so that

$$\begin{aligned}
\left(N - j + \frac{1}{2}\right) \ln\left(\frac{N}{N-j}\right) \;&\geq\; \left(N - |I| + \frac{1}{2}\right) \cdot \frac{j}{N} \\
&\geq\; \frac{(N-j) \cdot |I|}{N} \\
&\geq\; 1 - \frac{1}{N} \;\geq\; \frac{1}{12} \;.
\end{aligned}$$

Hence, the RHS of (7) is lower bounded by

$$\left(\frac{N}{N-j}\right)^{N-|I|+1/2} \cdot \left(\frac{N}{e}\right)^j \cdot e^{-\frac{1}{12(N-j)}} \;\geq\; e^{\frac{1}{12} - \frac{1}{12(N-j)}} \cdot \left(\frac{N}{e}\right)^j \;\geq\; \left(\frac{N}{e}\right)^j .$$

The desired conclusion follows. □