The Distinction Between Fixed and Random Generators in Group-Based Assumptions

James Bartusek* Fermi Ma† Mark Zhandry‡

Princeton University

Abstract

There is surprisingly little consensus on the precise role of the generator g in group-based assumptions such as DDH. Some works consider g to be a fixed part of the group description, while others take it to be random. We study this subtle distinction from a number of angles.

- In the generic group model, we demonstrate the plausibility of groups in which random-generator DDH (resp. CDH) is hard but fixed-generator DDH (resp. CDH) is easy. We observe that such groups have interesting cryptographic applications.
- We find that seemingly tight generic lower bounds for the Discrete-Log and CDH problems with preprocessing (Corrigan-Gibbs and Kogan, Eurocrypt 2018) are not tight in the sub-constant success probability regime if the generator is random. We resolve this by proving tight lower bounds for the random generator variants; our results formalize the intuition that using a random generator will reduce the effectiveness of preprocessing attacks.
- We observe that DDH-like assumptions in which exponents are drawn from low-entropy distributions are particularly sensitive to the fixed- vs. random-generator distinction. Most notably, we discover that the Strong Power DDH assumption of Komargodski and Yogev (Komargodski and Yogev, Eurocrypt 2018) used for non-malleable point obfuscation is in fact *false* precisely because it requires a fixed generator. In response, we formulate an alternative fixed-generator assumption that suffices for a new construction of non-malleable point obfuscation, and we prove the assumption holds in the generic group model. We also give a generic group proof for the security of fixed-generator, low-entropy DDH (Canetti, Crypto 1997).

1 Introduction

Starting with the seminal work of Diffie and Hellman [DH76], the *Computational Diffie-Hellman* (CDH) assumption in certain cyclic groups has become a core pillar of modern cryptography. For a finite cyclic group G and generator g, the assumption holds if it is hard to compute g^{ab} given (g, g^a, g^b) for random a, b. The corresponding *Decisional Diffie-Hellman* (DDH) assumption, introduced by Brands [Bra94], is that given (g, g^a, g^b) for random a, b, it is hard to distinguish g^{ab} from g^c for random c.

A somewhat subtle issue is the precise role of g in these assumptions: is it fixed in the group description, or is it randomly chosen along with a and b? For CDH in groups where the totient of the order is known, a folklore equivalence between the fixed and random generator variants exists (e.g. see Chapter 21 of Galbraith's textbook [Gal12]). For DDH, Shoup [Sho99] observed that the fixed generator assumption appears to be a *stronger* assumption than the random generator version, though a formal separation between

*bartusek.james@gmail.com.

†fermima1@gmail.com.

‡mzhandry@princeton.edu.

the two is unknown. Despite this apparent distinction, the cryptographic literature commonly refers to both the fixed and random generator variants simply as "DDH".[1]

A likely explanation for this practice is that in most applications of cryptographic groups, it is straightforward to switch between fixed and random generators. For example, in ElGamal encryption [ElG84], users who want the additional security of random-generator DDH can easily specify a random generator in their public key.

Sadeghi and Steiner [SS01] observed that this justification does not apply in settings where the choice of group generator is left to a potentially untrusted party.[2] They give the example of a bank that offers its customers an anonymous payment system, claiming provable security under group-based assumptions. If the bank is free to choose parameters such as the group generator, then for security it is crucial that any underlying assumptions hold in their (stronger) fixed generator form. While Sadeghi and Steiner did not point to specific assumptions that can be broken simply by fixing the group generator, they stressed that continuing to conflate these distinct assumptions could lead to serious ambiguities and mistakes in the future.

In the nearly two decades since Sadeghi and Steiner [SS01] first called attention to the above issue, dozens of new and increasingly sophisticated group-based assumptions have been introduced. Accordingly, researchers have devoted significant effort to evaluating the plausibility of these assumptions (e.g. [BFF+14, DHZ14]), frequently in idealized models such as the generic group model [Nec94, Sho97, Mau05]. We observe that these generic group justifications generally ignore the question of whether the generator is fixed or random, but that in most cases this distinction does not seem to affect real world security of these assumptions.

In this work, however, we will see that this is not *always* the case.

1.1 Our Results

We first examine how the fixed vs. random generator distinction affects the classical Discrete-Log, CDH, and DDH problems in a variety of different settings, obtaining the following results:

- **Generic Separations for CDH and DDH.** We prove that fixed- and random-generator DDH are *inequivalent* assumptions in the generic group model [Nec94, Sho97, Mau05]. We show that for groups of *unknown order*, fixed- and random-generator CDH are also inequivalent assumptions in the generic group model. In addition, we give evidence (relying on a new assumption about arithmetic circuits) that they are inequivalent even if the group order is known but its factorization is not.[3]

- **Split-CDH and Split-DDH Groups.** We define Split-CDH (resp. Split-DDH) groups for which the fixed-generator variant of CDH (resp. DDH) is easy but the random-generator variant is hard, and we observe that such groups imply interesting cryptographic applications. A split-CDH group can be turned into a *self-bilinear map* [YYHK14, KKS15] where the random-generator variant of the Multilinear CDH assumption holds. This implies powerful primitives such as multiparty non-interactive key agreement (with trusted setup).[4] A split-DDH group can be used to instantiate a variant of the Boneh-Franklin identity-based encryption [BF01] scheme. We stress here that giving candidate constructions of these groups is outside of the scope of this work. On the negative side, we prove that a natural class of non-interactive key exchange protocols (without trusted setup) are *insecure* in certain split-CDH groups.

[1] For example, the Katz-Lindell textbook [KL] defines DDH with a fixed generator, while Cramer-Shoup [CS98] defines DDH with a random generator.

[2] Sadeghi and Steiner [SS01] actually consider the more general possibility of the untrusted party choosing the group itself maliciously. This question is beyond the scope of our work, but in many cases it is an equally important consideration.

[3] This inequivalence was also suggested by Saxena and Soh [SS06].

[4] A similar observation was also made in [SS06].

- **Asymptotic Bounds for Discrete-Log and CDH with Preprocessing.** We revisit the recent work of Corrigan-Gibbs and Kogan [CK18], which seemingly resolves the generic hardness of Discrete-Log and CDH with preprocessing. We observe that while their lower bounds are tight for the fixed-generator variants, they leave a gap in the random-generator setting for algorithms with sub-constant success probability. We close these gaps by proving tight lower bounds for the random-generator variants. Our bounds suggest that using a random generator can reduce the impact of preprocessing attacks, and in turn group parameters can be set more aggressively than previously thought in situations where random-generator Discrete-Log or CDH are sufficient.

Next, we turn our attention to the class of Diffie-Hellman-like assumptions involving *non-uniform random exponents*. An example of such an assumption is Canetti's "DDH-II" assumption [Can97], which states that DDH remains hard even if the exponent a in (g, g^a, g^b, g^{ab}) is drawn from a well-spread distribution (so that a has super-logarithmic min-entropy). While these assumptions are somewhat undesirable due to their non-standard nature [GK16], Wee [Wee05] showed that these assumptions (ones that require hardness given only super-logarithmic entropy) are *necessary* for applications such as point-function obfuscation.

Before we rely on such assumptions, it is important to rule out idealized adversaries that attack the underlying structure of the assumption. The most common technique for achieving this is to prove the assumption holds in the generic group model [Nec94, Sho97, Mau05]. Such proofs certainly do not imply the validity of the assumption; instead, these proofs are generally viewed as a *minimal level of guarantee* we need to gain confidence in an assumption [BFF+14].

Our central focus is on the recently proposed "Strong Power DDH" assumption of Komargodski and Yogev [KY18a]. The assumption states that for x sampled from any arbitrary well-spread distribution $\mathcal{D}$, that $g^x, g^{x^2}, \ldots, g^{x^k}$ is indistinguishable from k uniformly random group elements. Our results are the following:

- **Strong Power DDH is False for a Fixed Generator.** We demonstrate the "Strong Power DDH" assumption underlying Komargodski and Yogev's non-malleable point obfuscator [KY18a] as well as Fenteany and Fuller's non-malleable digital locker [FF18] is *false* in the fixed-generator setting. This results from a subtle issue in the order of quantifiers; if g is fixed, an arbitrary well-spread distribution could depend on g. For example, x can come from the distribution that conditions on the bit-representation of g^x beginning with 0. Unfortunately, these constructions can only be instantiated with a fixed generator, so the original security proofs in [KY18a] and [FF18] must rely on a false assumption.[5,6]

  In response to private communication from the authors of this work, Komargodski and Yogev have offered a simple fix [KY18b] for their original construction through a new "Entropic Power DDH" assumption.[7] This new assumption suffices for non-malleable point obfuscation and is formulated precisely to address the vulnerability described above.

- **Fixing Non-Malleable Point Obfuscation and Justifying Assumptions in the Generic Group Model.** In this work, we offer an alternative resolution. We construct a new non-malleable point obfuscator that is qualitatively different from the one in [KY18a]. Security of our construction relies on a newly formulated fixed-generator entropic assumption that we prove holds in the generic group model. Note that neither the Strong Power DDH Assumption [KY18a] nor the revised Entropic Power DDH Assumption [KY18b] come with generic group proofs of security.

[5] Relying on a random generator would require a common random string, which is not the model considered in [KY18a] or in the version of [FF18] dated Jan 30, 2019 at eprint.iacr.org/2018/957/20190130:190441.

[6] This issue appears in the Eurocrypt 2018 version of [KY18a], in an older ePrint version of [KY18b] dated May 1, 2018 at eprint.iacr.org/2018/149/20180211:142746, and in the ePrint version of [FF18] dated Jan 30, 2019 at eprint.iacr.org/2018/957/20190130:190441.

[7] This refers to the newer ePrint version of [KY18b] dated Feb 21, 2019 available at https://eprint.iacr.org/2018/149/20190221:133556.

Along the way, we develop general techniques (based heavily on [CDG18]) for proving generic security of non-standard, entropic assumptions. As a final contribution, we demonstrate the applicability of these techniques by showing that the fixed- and random-generator versions of Canetti's DDH-II assumption [Can97] hold in the generic group model.[8] This assumption has been used in both its fixed-generator form (e.g. [KLRZ08, CD08, DHZ14]) and random-generator form (e.g. [Can97, BC10]).

1.2 Reader's Guide

Our contributions (and the following technical overview) are divided into 4 parts.

- Part 1 is collection of generic-group-based results that explore the fixed- or random-generator distinction for Discrete-Log, CDH, and DDH. We also describe our new split-CDH and split-DDH groups. These results are contained in Section 3.
- Part 2 is a discussion on the negative implications of the fixed- or random-generator distinction, with a focus on trusted set-up in Diffie-Hellman Key Exchange. The key technical result in this section is a black-box separation between random-generator CDH and a natural class of non-interactive key exchange protocols. This part is in Section 4.
- Part 3 considers the problem of generic algorithms with preprocessing in the random-generator setting. Our lower bound for random-generator Discrete-Log and CDH is in Section 5.
- Part 4 studies the problem of non-malleable point obfuscation. We give our construction and prove security under a new assumption in Section 6. We justify our new assumption with a generic group model proof in Section 7. Our generic group model proof for DDH-II can be found in Section 7.3.

We encourage any readers interested in non-malleable point obfuscation or generic group proof techniques to first read Part 4 in the following technical overview before visiting the proofs in Section 6, Section 7, and Section 7.3.

1.3 Technical Overview

1.3.1 Part 1: Generic Separations and Split Groups.

Formalizing the Distinction. We will assume some process for generating a group description G of order N. This group description is assumed to include a generator g. The *fixed-generator* DDH assumption, or f-DDH, states that the tuples (g^x, g^y, g^{xy}) and (g^x, g^y, g^z) are computationally indistinguishable, given the description of G. Here, x, y, z are chosen randomly in $\mathbb{Z}_N$. On the other hand, the *random-generator* DDH assumption, or r-DDH, states that the tuples (h, h^x, h^y, h^{xy}) and (h, h^x, h^y, h^z) are computationally indistinguishable. Here, x, y, z are chosen randomly in $\mathbb{Z}_N$, and h is a random generator of G (chosen, say, by setting $h = g^r$ for a random r in $\mathbb{Z}_N^*$). We can also define fixed- and random-generator variants of Computational Diffie-Hellman (CDH) and Discrete-Log (DLog). For example, f-CDH states that given (g^x, g^y) for random x, y, it is computationally infeasible to find g^{xy}.

We consider the following three settings of groups: known prime group order, known composite group order of *unknown* factorization, and unknown group order. For each of the three assumptions and three settings (for 9 instances in total) we explore the relationship between the fixed- and random-generator variants. Trivially, the f- variants of the assumptions are at least as strong as the r- variants. In the other direction, some instances have known or folklore reductions showing equivalence [Gal12]. For each of the cases that do not have a proof of equivalence, we provide a separation. This is formalized by augmenting the generic group model [Sho97] with an oracle for the f- variant, and showing (potentially under reasonable computational assumptions) that the r- variant still holds. Table 1 summarizes our findings.

[8] Previously, such proofs had been obtained by Bitanksy and Canetti [BC10] and Damgård, Hazay, and Zottarel [DHZ14], who considered the random- and fixed-generator versions, respectively. We observe that both of these proofs treat the well-spread distribution as independent of the generic group labeling. Our proof handles distributions with arbitrary dependence on the labels; for more discussion refer to Part 4 of Section 1.3.

	DLog	CDH	DDH
Known Order	✓ (FL/Lemma 5)	✓ (FL)	× (Theorem 2)
Unknown Factorization	✓ (FL/Lemma 5)	×? (Theorem 4)	× (Theorem 2)
Unknown order	✓ (FL/Lemma 5)	×? ([YYHK18]/Theorem 3)	× (Theorem 2)

Table 1: Generic equivalences and separations. FL denotes a folklore result. ✓ means that the fixed and random generator versions are equivalent. × means that the random generator version is harder than the fixed generator version (in the generic model). ×? means the result holds under a plausible conjecture.

Applications of Split Groups. Looking at Table 1, we see that in the case of DDH, there is the potential for a group where f-DDH is easy but r-DDH is hard. We will call such groups split-DDH groups. Similarly, if the group order is unknown, potentially f-CDH is easy but r-CDH is hard; we call such groups split-CDH groups. In this section, we will see that such split Diffie-Hellman groups have useful cryptographic applications.

First, we observe that a split-CDH group is very close to a self-bilinear map [YYHK14]. A self-bilinear map is a group G together with a pairing $e : G^2 \to G$ such that $e(g^x, g^y) = e(g, g)^{xy}$. Let $g_1 = g$ and $g_n = e(g, g_{n-1})$. A typical computational assumption on self-bilinear maps would be the multilinear CDH assumption [BS02]: for any $n > 1$, given $g^{x_0}, \ldots, g^{x_n}$, it is hard to compute $g_n^{\prod_{i=0}^{n} x_i}$. Notice that by applying the mapping $e(\cdot, \cdot)$, it is only possible to compute $g_{n+1}^{\prod_{i=0}^{n} x_i}$.

An f-CDH oracle gives such an oracle where $e(g, g) = g$. Therefore, a split-CDH group gives all the functionality of a self-bilinear map. But notice that since $e(g, g) = g$, $g_n = g$ for any n. Therefore, the multilinear CDH assumption is false. However, we observe that if we choose a random element h, then $e(h, h) = h^r$ where $h = g^r$. As such, the f-CDH oracle would also give a self-bilinear map with respect to the random generator h. We then show that multilinear CDH is actually hard relative to h, assuming r-CDH is hard. Thus, we obtain a self-bilinear map from any split-CDH group. As a consequence, following [YYHK14] we would immediately obtain multiparty non-interactive key agreement, broadcast encryption satisfying a distributed setup notion [BZ14], and attribute-based encryption for circuits.

In Section 3.2 we show that Split-DDH groups allow for a simple identity-based encryption (IBE) scheme based on the Boneh-Franklin [BF01] construction.

Our results above demonstrate that finding groups where f- and r- assumptions are separated yields interesting applications. In the next part, we discuss the negative implications of differing hardness between f- and r- assumptions.

1.3.2 Part 2: Trusted Setup Assumptions

The previous sections demonstrated that the f- and r-DDH assumptions are distinct assumptions that may not both be true. But then which DDH assumption should be used? In practice, g is typically part of a standards library chosen by a trusted third party (e.g. NIST). As such, users have essentially three choices:

1. Believe that the trusted third party chose g at random, and use the r-DDH assumption.
2. Do not trust the third party, but instead assume that there are no bad g. In other words, rely on the f-DDH assumption for g.
3. Do not trust the third party, but instead have one of the users generate a random g and distribute it to everyone else. Then rely on r-DDH.

Option 1 means that users need to trust that no one could have subverted g and chosen a bad generator for which DDH is actually easy; history has shown such trust could very well be misplaced. Only Options 2 and 3 remove the need to trust a third party.

Remark 1. *Note that to remove trusted setup assumptions entirely, we would need to ensure that G itself is guaranteed to satisfy f-DDH. One option is to assume that both G and g were generated by a deterministic process, so that all parties can calculate G, g for themselves without any setup. For groups based on finite fields, this requires deterministically generating large primes; while no polynomial-time provable algorithms are known, there are very simple heuristic algorithms. For elliptic curve-based groups, other options are available (e.g. using a field with small characteristic). For one approach to deterministic curve generation, see [BCLN16].*

In most cases, it is straightforward to switch between Options 2 and 3. A scheme designed for f-DDH can often be converted into a scheme that relies only on r-DDH by having one of the parties choose a random generator. On the other hand, a scheme designed for r-DDH can often be converted into an f-DDH scheme by fixing a group element and not including it with the user's messages, saving slightly on transmission costs.

The above means slightly different parameter sizes for the two assumptions. For example, for public key schemes, the extra group element would naturally go in the public key. The result is that schemes secure under r-DDH naturally require one additional group element in the public key relative to the f-DDH analog. As authors often compare parameter sizes in terms of group elements (e.g. [Fuj16]), it is important that they clearly identify which assumption is used.

In some cases, however, switching between f-DDH and r-DDH will have a more profound impact. For example, in a protocol between mutually distrusting parties, which party will be entrusted to come up with the generator? While we are not aware of any instances of protocols in the literature that cannot be made to work with a random generator, it is straightforward to devise protocols where no single party can be trusted to choose the generator. As such, care must be taken when using the r-DDH assumption in these settings.

Diffie-Hellman Key Exchange. For the remainder of this section, we will focus on a concrete setting where it is not possible to trivially switch between f-DDH and r-DDH: Diffie-Hellman key exchange. In the protocol, Alice chooses a random $a \leftarrow \mathbb{Z}_N$ and computes $A = g^a$, and Bob chooses a random $b \leftarrow \mathbb{Z}_N$ and computes $B = g^b$. Then the two parties exchange A, B. In most treatments, Diffie-Hellman is a *non-interactive key exchange* (NIKE), which means that A and B are sent simultaneously. Alice then computes the secret key $K = g^{ab} = B^a$ and Bob computes $K = g^{ab} = A^b$. By the DDH assumption, an eavesdropper who learns A, B can learn nothing about K.

The key issue here is that Alice and Bob need to know g in order to generate their first message. So if we want one of them, say Alice, to come up with the generator, the result is an *interactive* protocol with Alice sending the first message, and only then can Bob send his. Therefore, in addition to requiring slightly more communication, Option 3 actually changes the nature of the protocol. What we see is that Diffie-Hellman can only remain a setupless NIKE under the f-DDH assumption.

Now, it is possible to alter Diffie-Hellman to work with CDH by extracting hardcore bits from the unpredictable key. By the equivalence of f-CDH and r-CDH in known prime-order groups, we can obtain a setupless NIKE protocol from r-CDH (and hence also r-DDH). In groups of unknown order, however, this does not apply. As our main technical result from this section, we give evidence that in groups where the totient of the order is *unknown*, r-CDH alone is *insufficient* for constructing setupless NIKE. This is formalized by assuming that f-CDH is easy and demonstrating an attack on a wide class of key agreement protocols that generalize the classical Diffie-Hellman protocol.

1.3.3 Part 3: Random-Generator Discrete-Log and CDH with Preprocessing.

A recent line of works [Mih, LCH11, BL13, CK18, CDG18] have explored *non-uniform* attacks on various problems in cryptographic groups. Here, a computationally expensive offline pre-processing stage generates an advice string, which in a later online stage can be used to speed up computation in the group. We are interested in the relationship between the length S of the advice string, the running time T of the online stage, the group order N, and the success probability ϵ.

Very recently, Corrigan-Gibbs and Kogan [CK18] seemingly resolve the non-uniform hardness of the *discrete logarithm* problem. Namely, they show in the generic group model that $\epsilon = \widetilde{O}(ST^2/N)$, where the $\widetilde{O}$ hides logarithmic factors. This matches known upper bounds (attacks) up to logarithmic factors.

However, all the works in this line (both lower bounds and attacks) only consider the fixed generator version of discrete log. Corrigan-Gibbs and Kogan briefly mention this, concluding that "using a fixed generator is essentially without loss of generality" since a discrete log with respect to one generator can be solved by solving two discrete logs with respect to a different generator.

When considering just polynomial reductions between problems, the above is certainly true. However, when it comes to precisely quantifying hardness, the problem no longer remains identical for different generators. In particular, suppose we have an algorithm that solves discrete log with respect to generator g with probability ϵ and we want to solve a discrete log instance with respect to generator $h = g^r$. To do so, on input h^x, we apply the algorithm twice to find the discrete logs of h and h^x with respect to g. This gives r and rx, allowing us to solve for x. But since we needed to solve both instances correctly, our overall success probability is only ϵ^2. Of course if ϵ is a constant so is ϵ^2, but in the low success probability regime, squaring the advantage significantly changes the hardness of the problem.

We resolve the question of the hardness of random-generator discrete log in the pre-processing setting, showing that $\epsilon = \widetilde{\Theta}\left(\frac{T^2}{N} + \frac{S^2T^4}{N^2}\right)$. The attack side is simple: there are two natural ways to attack a random-generator discrete log instance h, h^x. One is to ignore the pre-processing, and apply the Baby-step Giant-step algorithm, with success $\Omega(\frac{T^2}{N})$. The other is to use the pre-processing to solve two discrete log instances relative to some fixed generator g, in the manner described above. This gives success $\Omega((\frac{ST^2}{N})^2)$, as shown in [CK18]. By choosing which algorithm to use based on the parameters S, T, N, one obtains $\epsilon = \Omega\left(\frac{T^2}{N} + \frac{S^2T^4}{N^2}\right)$.

On the other hand, to prove the lower bound we need to show, essentially, that the two algorithms above are the only possible algorithms. This does not follow from the analysis of [CK18]. Instead, we use the tools developed in subsequent works [CDGS18, CDG18] (based on the earlier pre-sampling techniques developed by Unruh [Unr07] for the Random Oracle model) to switch to a "bit-fixing" model, where we then show the optimality of the algorithms. In addition, we show that the same relationship holds as well for r-CDH. Generically, auxiliary input r-CDH is as hard as either using the auxiliary information to solve two discrete logarithms, or ignoring the input and solving one discrete logarithm.

1.4 Part 4: Low-Entropy Fixed-Generator Assumptions

Background: Point Obfuscation from Low-Entropy Assumptions. Our discussion thus far has focused on Discrete Log/Diffie-Hellman-type assumptions where g^a, g^b are uniformly random group elements. However, the security of many important cryptographic applications often relies on a stronger version of these assumptions in which a and/or b might not be drawn uniformly at random.

Canetti's construction of point function obfuscation is perhaps the most well-known example.[9] A point function $f_x(\cdot)$ is a boolean function that accepts on x and rejects on all other inputs. Roughly speaking, an obfuscated point function $\mathcal{O}(f_x(\cdot))$ implements the same input/output functionality as $f_x(\cdot)$, but leaks no information about x beyond what can be learned through black-box oracle queries to $f_x(\cdot)$. In other words, the obfuscated program acts as a *virtual black box* for evaluating the function.[10] Canetti's point function obfuscator is simple: to obfuscate $f_x(\cdot)$, draw a random group element g^b and output (g^b, g^{xb}). Evaluation on input y is done by computing $(g^b)^y$ and accepting if it matches g^{xb}.

The security of this construction follows from an assumption Canetti refers to as DHI-II (in subsequent works it has been renamed to "DDH-II"; we will adopt this name), which states that $(g, g^a, g^b, g^{ab}) \approx_C (g, g^a, g^b, g^c)$ where g is a random generator, b, c are chosen uniformly at random, and a has super-logarithmic

[9] Canetti's results [Can97] were originally described in the language of "oracle hashing"; the equivalence to point function obfuscation was later pointed out by Wee [Wee05]. We describe Canetti's results in Wee's terminology.

[10] We defer a more detailed discussion on virtual-black-box obfuscation to [BGI+01] (see [Wee05] for specifics on point function obfuscation).

min-entropy, i.e. it is sampled from a *well-spread* distribution $\mathcal{D}$. We stress that DDH-II is technically an infinite family of assumptions, since it requires indistinguishability if $\mathcal{D}$ is *any* well-spread distribution (even ones that are not efficiently sampleable).

Under DDH-II, the obfuscated program (g^b, g^{xb}) hides all information about the point x as long as x is drawn from a well-spread distribution, since g^{xb} is indistinguishable from g^c. This immediately implies a notion of average-case virtual-black-box (VBB) security. Canetti proves that if a point function obfuscator is average-case VBB for *any* well-spread distribution, this implies full (worst-case) VBB security. It was later shown by Wee ([Wee05], Section 4.2) that Canetti's approach is essentially inherent: VBB-secure point function obfuscation *requires* strong assumptions that are hard for any well-spread distribution.

Background: Non-Malleable Point Obfuscation. Canetti's original motivation for studying point obfuscation was to realize useful properties of random oracles [BR93] in the standard model. If $H(\cdot)$ is a random oracle, observe that $H(x)$ is a secure point obfuscation of $f_x(\cdot)$, where evaluation is a single random oracle call followed by a comparison. Komargodski and Yogev [KY18a] observe that the random oracle obfuscator $H(x)$ satisfies a strong *non-malleability* property, in the sense that given $H(x)$ it is impossible to compute $H(f(x))$ for any (meaningfully) related point $f(x)$, without first recovering x. This property is missing from Canetti's point obfuscator [Can97], e.g. since given (g^b, g^{xb}), one can easily compute $(g^b, g^{(x+1)b})$, which is an obfuscation of the related point $f(x) = x + 1$.

Komargodski and Yogev [KY18a] propose the following modification to Canetti's point obfuscator. To obfuscate the point x, sample a random b and output $(g^b, (g^b)^{g^{x^4+x^3+x^2+x}})$. Note that for this expression to make sense, $g^{x^4+x^3+x^2+x}$ must be mapped back into the exponent space under some fixed public mapping. Evaluation on input y is done by computing $g^{y^4+y^3+y^2+y}$, mapping this element back to the exponent space and raising g^b to that power, and finally comparing to $(g^b)^{g^{x^4+x^3+x^2+x}}$.

Komargodski and Yogev [KY18a] argue their obfuscation resists bounded-degree polynomial *mauling* attacks, in which an adversary given an obfuscation of x attempts to produce an obfuscation of $P(x)$ for some bounded-degree polynomial $P(\cdot)$. Roughly, the intuition is that the adversary cannot replace g^b with any other $g^{b'}$, since generating $(g^{b'})^{g^{P(x)}}$ does not appear possible given only $(g^b)^{g^{x^4+x^3+x^2+x}}$. But if the adversary cannot change g^b, the argument is that the linear constraints imposed by the form of $x^4 + x^3 + x^2 + x$ make it impossible to replace x with $P(x)$.

Formally, security in [KY18a] is proved under the newly introduced "Strong Power DDH" assumption, which states it is hard to distinguish $g^x, g^{x^2}, \ldots, g^{x^\ell}$ from ℓ random group elements, if x is drawn from any well-spread distribution.

Fixed-Generator Strong Power DDH is False. In stating the assumption, Komargodski and Yogev [KY18a] do not specify how g is chosen or the relationship between g and the distribution over x. We observe that if g is a fixed generator, then their assumption is false. For a uniformly random group element, there must be some bit in its description with noticeable entropy. If it is bit i, we let $\mathcal{D}$ be the distribution over all points x such that the ith bit of the description of g^x is 0. Then $\mathcal{D}$ has high min-entropy, and moreover g^x for $x \leftarrow \mathcal{D}$ is distinguishable from a random group element by inspecting the ith bit.

If the assumption is taken in its random-generator formulation, the security proof in [KY18a] breaks down, since an adversary can potentially replace g with a different generator g'. A natural idea to fix the construction would be to generate g using a public source of randomness.[11] However, this would move the construction into the CRS model, where strong non-malleability results were previously known [CV09].

Fixing Non-Malleable Point Obfuscation. We remedy this situation by giving an alternative low-entropy fixed-generator assumption, and proving that this assumption is sufficient to achieve their notion of

[11] As noted in Section 1.1, Komargodski and Yogev have offered a fix through a new Entropic Power DDH Assumption in a revised ePrint posting [KY18b], which does not come with a generic group proof. The goal of this section is to build non-malleable point obfuscation from an assumption that holds against generic adversaries.

non-malleable point obfuscation. We formulate our assumption in a way that allows us to prove it holds in the generic group model. Our assumption is the following:

> Let $p \in [2^{\lambda-1}, 2^{\lambda}]$ and let n be at most $\mathsf{poly}(\lambda)$. Fix a group G of order p along with a generator g, and draw $k_2, \ldots, k_n$ uniformly at random from $\mathbb{Z}_p$. For any well-spread distribution $\mathcal{D}$ over $\mathbb{Z}_p$, no efficient adversary can distinguish $\{g^{k_i x + x^i}\}_{i \in \{2,\ldots,n\}}$ for $x \leftarrow \mathcal{D}$ from $n-1$ uniformly random group elements, even given $k_2, \ldots, k_n$.[12]

The intuition for the design of this assumption is the following. We want to modify the group elements $g^x, g^{x^2}, \ldots$ in Strong Power DDH to block distributions $\mathcal{D}$ which "condition" on the fixed g, as we have already seen how such distributions falsify the assumption. However, we are restricted to modifications that preserve our ability to perform a security reduction for the proof of non-malleability, as in [KY18a].

Without delving into the non-malleability security proof itself, the key requirement is that the reduction must be able to construct specific polynomials (in x) in the exponent. We tweak the construction so that the reduction can construct a polynomial of the form $ax + x^2 + x^3 + x^4 + x^5$, where a is an arbitrary but known scalar.[13] Then by using terms of the form $g^{k_i x + x^i}$, we enable the reduction to construct this polynomial by simply multiplying the $i = 2, \ldots, 5$ terms; it will know a since the k_i's are given in the clear. Intuitively, the k_i scalars contribute enough randomness to prevent distributions $\mathcal{D}$ which make the $g^{k_i x + x^i}$ terms distinguishable from random.

Our resulting construction of non-malleable point obfuscation is (essentially) $a, g^{ax + x^2 + x^3 + x^4 + x^5}$. We note that our construction does not require the "double exponentiation" of [KY18a]. The full construction comes with two additional scalars and group elements that ensure that x is the only accepting input.

Discussion: Low-Entropy Fixed Generator Assumptions in the Generic Group Model. In order to gain confidence in our assumption, we prove it secure in the generic group model. As discussed in Section 1.1, this is usually viewed as a minimum requirement in order to gain confidence in a new group-based assumption. Recall that in the generic group model, group elements g^x are replaced with random "labels" $\sigma(x)$, where σ is a uniformly random injection from the space of exponents to some space of labels. An oracle stores the entire description of σ, and allows the generic adversary oracle access to honest group operations. For example, an adversary with labels $\sigma(x), \sigma(y)$ can request the label for $\sigma(x + y)$.

We find that in the setting of fixed generator lower entropy assumptions, the standard intuition for designing generic group model proofs falls short. Our goal is to prove no generic adversary can distinguish between $\{k_i, \sigma(k_i x + x^i)\}_{i \in \{2,\ldots,n\}}$ and $\{k_i, \sigma(r_i)\}_{i \in \{2,\ldots,n\}}$ for uniformly random k_i, r_i, and $x \leftarrow \mathcal{D}$. Since the group and generator are fixed in this assumption, we *must* consider distributions which depend on the group description itself. So in the generic model, any distribution $\mathcal{D}$ should be viewed as the output distribution of a potentially *inefficient* sampling algorithm $\mathcal{S}$ that is free to scan the entire labeling function σ.[14] The only requirement we enforce is that given σ, the point $x \leftarrow \mathcal{S}(\sigma)$ has super-logarithmic entropy.

To illustrate the difference in this setting, suppose for a moment that the sampler $\mathcal{S}$ had to output x without seeing σ (as is the case when x is drawn uniformly at random from $\mathbb{Z}_p$). The standard generic group argument for indistinguishability would use the following structure:

> Imagine treating x as a formal variable instead of as a randomly drawn value. This replaces the group exponent space $\mathbb{Z}_p$ with formal polynomials $\mathbb{Z}_p[x]$, so the oracle now returns labels by sampling a uniformly random label from the image of σ each time it encounters a distinct formal polynomial. Observe that there are no (non-trivial) linear combinations of the $\{k_i x + x^i\}_i$

[12] We remark that the assumption we actually use is slightly different: instead of stating indistinguishability from uniform, we require indistinguishability from $\{g^{k_i y + y^i}\}_{i \in \{2,\ldots,n\}}$ for the same $\{k_i\}_i$ but uniformly random y. We can prove both forms of this assumption hold in the GGM, but this second form yields a simpler proof of VBB security. For the purposes of this technical overview this distinction can be ignored.

[13] In the full proof of non-malleability, the reduction must be able to construct other higher degree polynomials, which is why we need the assumption to hold for $n = \mathsf{poly}(\lambda)$.

[14] In order to leverage Canetti's proof that average-case VBB implies worst-case VBB for point functions, average-case security must hold even for inefficiently sampleable well-spread distributions.

polynomials (taken as formal polynomials in x) that evaluate to identically zero polynomials over x. This implies that the adversary will never encounter non-trivial collisions in the labels it sees, and we can use the Schwartz-Zippel Lemma to argue that the adversary's view is identical in the world where x is random instead of a formal variable.

This type of argument breaks down if $\mathcal{S}$ can choose x *after* seeing the labeling function σ. Now $\mathcal{S}$ can try to pick x so that $\sigma(k_i x + x^i)$ conveys non-trivial distinguishing information to the adversary. In particular, it is no longer accurate to argue that we can produce an identical view for the adversary by replacing x with a formal variable.

We could intuitively hope that $\mathcal{S}$ is powerless to pick x that can bias the distribution of $\sigma(k_i x + x^i)$ away from uniform, as it does not know the random k_i. However, this intuition proves tricky to formalize, especially since $\mathcal{S}$ is given unlimited computational power and access to the entire function σ.

Connection to Preprocessing Attacks. To solve this problem, we apply the "bit-fixing" technique from Coretti, Dodis, and Guo [CDG18]. They consider generic algorithms which are given an additional advice string, computed beforehand using a computationally unbounded algorithm with access to σ. Conditioned on the advice string, it is no longer accurate to argue σ is a random labeling function. However, they show (roughly) that if we obtain at most P bits of advice about σ, this only leaks useful information about σ on $O(P)$ points. So for generic security proofs, this allows us to switch to a setting in which σ is a random labeling function on all but $O(P)$ inputs.

We apply these techniques to our setting by re-casting the sampler $\mathcal{S}$ outputting x as a computationally unbounded algorithm outputting x as "advice". However in our setting, the challenger is the one receiving the advice instead of the adversary. It turns out that the [CDG18] techniques still apply here, allowing us to argue that σ can be re-sampled on all but polynomially many points. Once we perform this re-sampling, we show that the adversary will not be able to apply group operations to its set of initial group elements and produce a point that was not re-sampled, except with negligible probability. Once this is established, standard generic group techniques suffice to complete the proof.

Generic Hardness of DDH-II. As a final contribution, we also prove the generic hardness of Canetti's DDH-II assumption. We remark that previous proofs of DDH-II [BC10, DHZ14] operate in a highly idealized model that assumes the sampler is independent of the labeling function σ. Preventing the sampler from seeing the labels implicitly relies on the group itself being drawn at random, which in particular leads to counterexamples when dealing with fixed generator assumptions. For example, the Strong Power DDH assumption with fixed generator can be proven in this model even though it is false in the real world.

In the case of DDH-II, one of the elements the adversary receives is $\sigma(a)$ for low entropy a. We must show at a minimum that this does not allow the adversary to recover a (i.e. compute the discrete log), as distinguishing would then be trivial. Such a claim might not be immediately obvious, especially considering that we can *distinguish* $\sigma(a)$ from $\sigma(r)$ for uniform r for certain distributions on a. We observe that any adversary which succeeds in solving discrete log of $\sigma(a)$ with noticeable advantage for a well-spread distribution is also an adversary that solves discrete log (with much smaller advantage) for the uniform distribution. However, the resulting advantage exceeds the known generic bounds for discrete log algorithms [Sho97]. The remainder of our proof makes use of bit-fixing techniques to reduce the problem of distinguishing the DDH-II instance to the problem of recovering a given just $\sigma(a)$.

2 Preliminaries

For $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, \ldots, n\}$. We specify formal variables by bold letters $\mathbf{x}$. For a function f, let $im(f)$ denote the image of f.

Throughout, we let $\lambda \in \mathbb{N}$ be the security parameter. We use the usual Landau notations. A function $f(\lambda)$ is said to be negligible if it is $\lambda^{-\omega(1)}$ and we denote it by $f(\lambda) \coloneqq \mathsf{negl}(\lambda)$. A function $f(\lambda)$ is said to have polynomial growth rate if it is $\lambda^{O(1)}$ and we denote it by $f(\lambda) \coloneqq \mathsf{poly}(\lambda)$. A probability $p(\lambda)$ is said to

be overwhelming if it is $1 - \lambda^{-\omega(1)}$. We refer to $\mathcal{A}$ as PPT if it is a probabilistic polynomial time algorithm. If $\mathcal{A}$ has access to an oracle $\mathcal{O}$, we write $\mathcal{A}^{\mathcal{O}}$.

The statistical distance between two distributions D_1 and D_2 over a countable support S is defined to be $\Delta(D_1, D_2) \coloneqq \frac{1}{2}\sum_{x \in S} |D_1(x) - D_2(x)|$. Let $\gamma > 0$. We say that two distributions D_1 and D_2 are γ-close if $\Delta(D_1, D_2) \leq \gamma$. We let $x \leftarrow \mathcal{D}$ denote drawing x from the distribution $\mathcal{D}$. When X is a set, then $x \leftarrow X$ denotes drawing x *uniformly at random* from the set X. The following definition regarding infinite families of distributions will be used throughout.

Definition 1 (Well-Spread Distribution Ensemble)**.** *An ensemble of distributions $\{\mathcal{D}_\lambda\}_\lambda$ over domains $\{\mathcal{X}_\lambda\}_\lambda$ is well-spread if for all large enough $\lambda \in \mathbb{N}$,*

$$H_\infty(\mathcal{D}_\lambda) = -\min_{x \in \mathcal{X}_\lambda} \log_2 \Pr[x \leftarrow \mathcal{D}_\lambda] = \omega(\log(\lambda)).$$

2.1 Generic Group Model

Definition 2 (Generic Group Model (GGM) [Nec94, Sho97])**.** *An application in the generic group model is defined as an interaction between a T-attacker $\mathcal{A}$ and a challenger $\mathcal{C}$. For a cyclic group of order N with fixed generator g, a random injective function $\sigma : [N] \to [M]$ is sampled, mapping group exponents in $\mathbb{Z}_N$ to a set of labels $\mathcal{L}$. Label $\sigma(x)$ for $x \in \mathbb{Z}_N$ corresponds to the group element g^x.*

$\mathcal{C}$ initializes $\mathcal{A}$ with some set of labels $\{\sigma(x_i)\}_i$. It then implements the group operation oracle $\mathcal{O}_G(\cdot,\cdot)$, which on inputs $\sigma_1, \sigma_2 \in [M]$ does the following:

- *If either of σ_1 or σ_2 is not in $\mathcal{L}$, return $\perp$.*
- *Otherwise, set $x = \sigma^{-1}(\sigma_1)$ and $y = \sigma^{-1}(\sigma_2)$, compute $x + y \in \mathbb{Z}_N$, and return $\sigma(x + y)$.*

$\mathcal{A}$ is allowed at most T queries to the oracle, after which $\mathcal{C}$ outputs a bit indicating whether $\mathcal{A}$ was successful. We refer to the probability that this bit is 1 as $\mathsf{Succ}_{\mathcal{C}}(\mathcal{A})$.

Remark 2. *It will often be convenient to represent each query to $\mathcal{O}_G$ as a linear polynomial over the initial set of elements $\{x_i\}_i$ given to $\mathcal{A}$.*

For an *indistinguishability* application, we define the *advantage* of attacker $\mathcal{A}$ as $\mathsf{Adv}_{\mathcal{C}}(\mathcal{A}) = 2|\mathsf{Succ}_{\mathcal{C}}(\mathcal{A}) - 1/2|$. For an *unpredictability* application, the advantage is defined as $\mathsf{Adv}_{\mathcal{C}}(\mathcal{A}) = \mathsf{Succ}_{\mathcal{C}}(\mathcal{A})$. An application with associated challenger $\mathcal{C}$ is (T, ϵ)-secure in the GGM is for every T-attacker $\mathcal{A}$, $\mathsf{Adv}_{\mathcal{C}}(\mathcal{A}) \leq \epsilon$.

Definition 3 (Auxiliary-Input Generic Group Model (AI-GGM))**.** *We now consider (S, T)-attackers $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$. First $\sigma : [N] \to [M]$ is sampled. $\mathcal{A}_1$ receives σ as input and outputs an S-bit string aux. Then the challenger $\mathcal{C}$ operates as before, modeling interaction between $\mathcal{A}_2$ and $\mathcal{O}_G(\cdot,\cdot)$. Now $\mathcal{A}_2$ receives aux as input and is allowed T queries to the oracle. Success, advantage, and security are defined analogously.*

Definition 4 (Bit-Fixing Generic Group Model (BF-GGM))**.** *We now consider (S, T, P)-attackers $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$. First $\sigma : [N] \to [M]$ is sampled. $\mathcal{A}_1$ receives σ as input and outputs an S-bit string aux along with a set $\mathcal{P} \subseteq \mathbb{Z}_N$ of size P. Then σ is uniformly re-sampled on all but the points $\mathcal{P}$ (conditioned on maintaining the same image), producing the injection σ'. We let $im(\mathcal{P})$ refer to the images under σ and σ' of the points in $\mathcal{P}$. Then the challenger $\mathcal{C}$ operates as before, modeling interaction between $\mathcal{A}_2$ and $\mathcal{O}_G(\cdot,\cdot)$, where $\mathcal{O}_G(\cdot,\cdot)$ uses σ' to answer queries. $\mathcal{A}_2$ receives aux as input and is allowed T queries to the oracle. Success, advantage, and security are defined analogously.*

Theorem 1 ([CDG18])**.** *Let $N, M, P \in \mathbb{N}$, $N \geq 16$, and $\gamma > 0$. If an unpredictability application with challenger $\mathcal{C}$ that initializes $\mathcal{A}$ with T' group elements is $((S, T, P), \epsilon')$-secure in the BF-GGM for*

$$P \geq 18(S + \log(\gamma^{-1}))(T + T'),$$

then it is $((S, T, P), \epsilon)$-secure in the AI-GGM for $\epsilon \leq 2\epsilon' + \gamma$.

3 Generic Separations Between Fixed and Random Generator Assumptions

In this section, we explore the relationship between fixed generator and random generator assumptions, in particular looking at settings and assumptions where the fixed generator and random generator versions are plausibly inequivalent. As we then show, groups where f-DDH is easy but r-DDH is hard (split-DDH groups), or groups where f-CDH is easy but r-CDH is hard (split-CDH groups), can be used to build powerful cryptographic applications. On the other hand, we also show that a large class of setupless NIKE protocols are not secure in split-CDH groups.

We now formalize the six problems that we consider in this section. We assume a group generation algorithm $(\mathbb{G}_\lambda, g, N) \leftarrow \mathsf{GroupGen}(1^\lambda)$ where λ is the security parameter and $\mathbb{G}_\lambda$ is the description of a cyclic group with generator g and order N. All the assumptions below are relative to a *fixed* [15] output of the group generation algorithm on input the security parameter λ.

- r-DLog: Given (g^r, g^{rx}) compute x, where $r \leftarrow \mathbb{Z}_N^*, x \leftarrow \mathbb{Z}_N$
- f-DLog: Given (g^x) compute x, where $x \leftarrow \mathbb{Z}_N$
- r-CDH: Given (g^r, g^{rx}, g^{ry}) compute g^{rxy}, where $r \leftarrow \mathbb{Z}_N^*, x, y \leftarrow \mathbb{Z}_N$
- f-CDH: Given (g^x, g^y) compute g^{xy}, where $x, y \leftarrow \mathbb{Z}_N$
- r-DDH: Given $(g^r, g^{rx}, g^{ry}, g^{brxy+(1-b)z})$ determine b, where $r \leftarrow \mathbb{Z}_N^*, x, y, z \leftarrow \mathbb{Z}_N, b \leftarrow \{0,1\}$
- f-DDH: Given $(g^x, g^y, g^{bxy+(1-b)z})$ determine b, where $x, y, z \leftarrow \mathbb{Z}_N, b \leftarrow \{0,1\}$

We use the generic group model to explore the relationships among the above problems, and operate in three different settings. First is the usual setting, described in the preliminaries, where we implicitly assume the order of the group and its factorization is known. Then we model the setting where the order of the group is known but the factorization is not explicitly given. We call this the Unknown Factorization Generic Group Model (UF-GGM) (previously considered in [AJR08, AM09]). Note that in this setting, an advantage bound $\mathsf{Adv}_{\mathcal{C}}(\mathcal{A})$ is only meaningful for PPT adversaries $\mathcal{A}$ since otherwise we can assume $\mathcal{A}$ factors the group order and we are back in the usual setting. Contrast this with the usual generic group setting where the computational complexity of $\mathcal{A}$ is arbitrary and we only have a bound on the number of group operation queries to the oracle.

Finally, we model a third setting where the order of the group itself is unknown, called the Unknown Order Generic Group Model (UO-GGM) (first described in [DK02]). We adopt the convention considered in [YYHK18] for generic rings of unknown order, where the order is drawn at random from all primes with bit length at most λ. We augment this model with an explicit group inverse oracle that on input $\sigma(x)$ returns $\sigma(-x)$. Note that when the group order is known, this operation can be done with logarithmic queries to the group operation oracle, so this is usually not explicitly given in that setting.

We will also consider augmenting these generic group models with additional oracles beyond the group operation/inverse oracles. More specifically, we consider an f-DDH oracle $\mathcal{O}_{fDDH}(\cdot,\cdot,\cdot)$ that on input $\sigma(x), \sigma(y), \sigma(z)$ returns 1 if $xy = z$ and 0 otherwise, and an f-CDH oracle $\mathcal{O}_{fCDH}(\cdot,\cdot)$ that on input $\sigma(x), \sigma(y)$ returns $\sigma(xy)$. Now when we give an adversary T oracle queries, we mean combined queries between all oracles it has access to.

In the following proofs, we always assume that for group order $N \in [2^{\lambda-1}, 2^\lambda]$, the smallest prime dividing N has bit length $\omega(\log(\lambda))$. This is required to prevent trivial attacks, and is useful when arguing that polynomials modulo N will vanish with low probability over random inputs. We cannot apply Schwartz-Zippel right away since $\mathbb{Z}_N$ is not necessarily a field, and instead apply the following lemma.

[15] Random generator assumptions can be formulated with respect to a randomized generator h output by $\mathsf{GroupGen}$. Here we assume a fixed generator g and randomize it as part of the game itself, letting $h = g^r$.

Lemma 1. *Let $N \in [2^{\lambda-1}, 2^\lambda]$ be such that the bit length of its largest prime divisor is $\omega(\log(\lambda))$. Let $Q(x_1, \ldots, x_m)$ be a polynomial of total degree $n = \mathsf{poly}(\lambda)$. Then $\Pr[Q(r_1, \ldots, r_m) = 0 \bmod N] = \mathsf{negl}(\lambda)$, where each r_i is drawn uniformly and independently from either $\mathbb{Z}_N$ or $\mathbb{Z}_N^*$.*

Proof. Let p be a prime and k an integer such that $p^k \mid N$ but $p^{k+1} \nmid N$. By assumption, we know that $p = \omega(\mathsf{poly}(\lambda))$ and $k = \mathsf{poly}(\lambda)$. Any r_i drawn uniformly from $\mathbb{Z}_N$ when taken mod p^k is uniform from $\mathbb{Z}_{p^k}$, and by the Chinese Remainder Theorem, any r_i drawn uniformly from $\mathbb{Z}_N^*$ when taken mod p^k is uniform from $\mathbb{Z}_{p^k} \setminus p\mathbb{Z}$. Both of these domains have size $\omega(\mathsf{poly}(\lambda))$, thus we can apply Schwartz-Zippel to show that $\Pr[Q(r_1, \ldots, r_n) = 0 \bmod p^k] = \mathsf{negl}(\lambda)$. This immediately gives the result. □

3.1 DDH Separations

The following theorem establishes the inequivalence of the f-DDH and r-DDH problems in all three settings we consider. In particular, it gives heuristic evidence that a group might exist where f-DDH is easy (in fact, can be solved with probability 1) but r-DDH is hard.

Theorem 2. *r-DDH is (T, ϵ)-secure in the GGM (or UF-GGM or UO-GGM) augmented with an f-DDH oracle, for group order $N \in [2^{\lambda-1}, 2^\lambda], T = \mathsf{poly}(\lambda)$, and $\epsilon = \mathsf{negl}(\lambda)$.*

Proof. It suffices to give the proof in the regular GGM, since applications can only become harder in the UF-GGM and UO-GGM. In the r-DDH game, the challenger $\mathcal{C}$ is initialized with a labeling σ where $\mathcal{L} := im(\sigma)$ and chooses $r \leftarrow \mathbb{Z}_N^*, x, y, z \leftarrow \mathbb{Z}_N$, and $b \leftarrow \{0, 1\}$. The adversary $\mathcal{A}$ is initialized with $(\sigma(1), \sigma(r), \sigma(rx), \sigma(ry), \sigma((1-b)rxy + bz))$. It interacts with the generic group oracle $\mathcal{O}_G$ and the f-DDH oracle $\mathcal{O}_{fDDH}$, outputs a bit b', and wins if $b' = b$.

Let T' be the number of $\mathcal{O}_G$ queries made by $\mathcal{A}$ and T'' the number of $\mathcal{O}_{fDDH}$ queries. We represent the set of queries made to $\mathcal{O}_G$ as the set of linear polynomials

$$\{a_1^{(t)} r + a_2^{(t)} rx + a_3^{(t)} ry + a_4^{(t)}((1-b)rxy + bz) + a_5^{(t)}\}_{t \in [T']},$$

and the set of queries made to $\mathcal{O}_{fDDH}$ as

$$\begin{aligned}\{&b_1^{(t)} r + b_2^{(t)} rx + b_3^{(t)} ry + b_4^{(t)}((1-b)rxy + bz) + b_5^{(t)},\\ &c_1^{(t)} r + c_2^{(t)} rx + c_3^{(t)} ry + c_4^{(t)}((1-b)rxy + bz) + c_5^{(t)},\\ &d_1^{(t)} r + d_2^{(t)} rx + d_3^{(t)} ry + d_4^{(t)}((1-b)rxy + bz) + d_5^{(t)}\}_{t \in [T'']}.\end{aligned}$$

First we switch to a hybrid game where r, x, y, z are left as formal variables by the challenger. The challenger maintains a table mapping polynomials in $\mathbb{Z}_N[\mathbf{r}, \mathbf{x}, \mathbf{y}, \mathbf{z}]$ to labels in $\mathcal{L}$, picking a new uniformly random label among those unused so far each time $\mathcal{A}$ queries $\mathcal{O}_G$ with a new polynomial. When $\mathcal{A}$ queries $\mathcal{O}_{fDDH}$ with three linear polynomials ℓ_b, ℓ_c, ℓ_d, it responds with 0 if and only if $\ell_b \ell_c - \ell_d$ is identically zero over $\mathbf{r}, \mathbf{x}, \mathbf{y}, \mathbf{z}$. At the end of the game, the challenger draws the values of r, x, y and z from the appropriate distributions.

Now there are two ways in which $\mathcal{A}$ can distinguish this from the original game. First, it could query $\mathcal{O}_G$ on two separate polynomials that are not the same in $\mathbb{Z}_N[\mathbf{r}, \mathbf{x}, \mathbf{y}, \mathbf{z}]$, but evaluate to the same element once r, x, y, z are plugged in. Second, it could query $\mathcal{O}_{fDDH}$ on ℓ_b, ℓ_c, ℓ_d where $\ell_b \ell_c - \ell_d$ is not identically zero, but evaluates to zero when r, x, y, z are plugged in. All of these events can be represented by some non-constant polynomial of total degree at most 6 evaluating to zero over the randomness of r, x, y, z. By Lemma 1, the probability that any one of them evaluates to zero is $\mathsf{negl}(\lambda)$. Then a union bound over the $O(T^2)$ polynomials shows that $\mathcal{A}$ can notice this change with $\mathsf{negl}(\lambda)$ probability.

Now we argue that any adversary $\mathcal{A}$ has advantage 0 in this hybrid game. Regardless of the value of b, the four initial handles received by $\mathcal{A}$ are distinct monomials of the formal variables $\mathbf{r}, \mathbf{x}, \mathbf{y}, \mathbf{z}$. Thus any set of distinct linear combinations of them will be distinct polynomials over $\mathbb{Z}_N[\mathbf{r}, \mathbf{x}, \mathbf{y}, \mathbf{z}]$, and the answers from $\mathcal{O}_G$ will be identically distributed. Thus if $\mathcal{A}$ distinguishes, it must do so via a $\mathcal{O}_{fDDH}$ query. If $b = 0$, then

any query to $\mathcal{O}_{fDDH}$ returns 0 if and only if

$$(b_1\mathbf{r} + b_2\mathbf{rx} + b_3\mathbf{ry} + b_4\mathbf{rxy} + b_5)(c_1\mathbf{r} + c_2\mathbf{rx} + c_3\mathbf{ry} + c_4\mathbf{rxy} + c_5) \\ -(d_1\mathbf{r} + d_2\mathbf{rx} + d_3\mathbf{ry} + d_4\mathbf{rxy} + d_5) \equiv 0 \bmod N.$$

If $b = 1$, then any query to $\mathcal{O}_{fDDH}$ returns 0 if and only if

$$(b_1\mathbf{r} + b_2\mathbf{rx} + b_3\mathbf{ry} + b_4\mathbf{z} + b_5)(c_1\mathbf{r} + c_2\mathbf{rx} + c_3\mathbf{ry} + c_4\mathbf{z} + c_5) \\ -(d_1\mathbf{r} + d_2\mathbf{rx} + d_3\mathbf{ry} + d_4\mathbf{z} + d_5) \equiv 0 \bmod N.$$

It is clear that if the $b = 1$ equation is identically zero, then the $b = 0$ equation is also. Therefore, $\mathcal{A}$ can only distinguish if it finds a query such that the $b = 0$ equation is identically zero but the $b = 1$ equation is not. By expanding the left hand side of each, we see that the only difference lies in the following monomials (all others being distinct monomials over $\mathbf{r}$,$\mathbf{x}$,$\mathbf{y}$,$\mathbf{z}$ with the same coefficient regardless of b):

$$b = 0 : (b_1c_4 + b_4c_1 + b_2c_3 + b_3c_2)\mathbf{r}^2\mathbf{xy}$$
$$b = 1 : (b_1c_4 + b_4c_1)\mathbf{rz} + (b_2c_3 + b_3c_2)\mathbf{r}^2\mathbf{xy}.$$

Thus we need a setting of coefficients such that the $b = 0$ equation is identically zero but

$$b_1c_4 + b_4c_1 \not\equiv 0 \bmod N \text{ or } b_2c_3 + b_3c_2 \not\equiv 0 \bmod N.$$

First, since the $b = 0$ equation is identically zero, we know that

$$b_1c_4 + b_4c_1 \equiv -(b_2c_3 + b_3c_2) \bmod N. \tag{1}$$

Next, since $b_1c_4 + b_4c_1 \not\equiv 0 \bmod N$, let p be a prime such that $p \mid N$ but $p \nmid b_1c_4 + b_4c_1$. Then either $p \nmid b_1c_4$ or $p \nmid b_4c_1$ so assume the former (the other case is symmetric). Thus $p \nmid b_1$. Also, since the $b = 0$ equation is identically zero, we must have that the coefficients on $\mathbf{r}^2$, $\mathbf{r}^2\mathbf{x}$, and $\mathbf{r}^2\mathbf{y}$ are 0 mod N, so

$$b_1c_1 \equiv 0 \bmod N,$$
$$b_1c_2 + b_2c_1 \equiv 0 \bmod N,$$
$$b_1c_3 + b_3c_1 \equiv 0 \bmod N.$$

Since $p \mid N$ and p is prime, the first equation above shows that $p \mid c_1$. Combining with the second equation shows that $p \mid c_2$ and combining with the third equation shows that $p \mid c_3$. Thus, p divides the RHS of Equation 1 which is a contradiction since we started with the assumption that p does not divide the LHS. □

3.2 Split-DDH Groups

The previous section demonstrates the feasibility of split-DDH groups. Now we show that such groups can be used to build identity based encryption. Let $\mathbb{G}$ be a split-DDH group with fixed generator g and order N. Let $\mathcal{ID}$ be a space of IDs and let H be a random oracle mapping IDs to group exponents in $\mathbb{Z}_N$. We assume the existence of an f-DDH oracle $\mathcal{O}(\cdot,\cdot,\cdot)$ that is correct on any fixed instance with overwhelming probability. This follows from the the fact the f-DDH is easy in split-DDH groups and folklore random self-reduction and self-correction algorithms for DDH. See [BF01] for definitions of IBE and related security notions.

- KeyGen : Choose $s \leftarrow \mathbb{Z}_N^*$ and let $\mathsf{msk} = s$ and $\mathsf{mpk} = g^s$.
- $\mathsf{Extract}$: On input $\mathsf{id} \in \{0,1\}^*$, let $t_{\mathsf{id}} = H(\mathsf{id})$, $\mathsf{sk}_{\mathsf{id}} = g^{t_{\mathsf{id}}s}$, and $\mathsf{pk}_{\mathsf{id}} = g^{t_{\mathsf{id}}}$.
- $\mathsf{Encrypt}$: On input $m, \mathsf{id} \in \{0,1\}^*$, choose $r_i, u_i \leftarrow \mathbb{Z}_N$ for $i \in [|m|]$ and output $\{\mathsf{mpk}^{r_i}, \mathsf{pk}_{\mathsf{id}}^{r_i^{-1}m_i + u_i(1-m_i)}\}_{i\in[|m|]}$.

- Decrypt : On input an encryption $\{c_i, d_i\}_i$ and id, output $\{\mathcal{O}(c_i, d_i, \mathsf{sk}_{\mathsf{id}})\}_i$.

For correctness, note that if $m_i = 0$, we can write the ith input to $\mathcal{O}$ as $(g^{sr_i}, g^{t_{\mathsf{id}} u}, g^{t_{\mathsf{id}} s})$ where s, u, r_i are uniformly random and independent. This is a valid f-DDH instance with negligibly probability. However, when $b = 1$, we can write the inputs as $(g^{sr_i}, g^{t_{\mathsf{id}} r_i^{-1}}, g^{t_{\mathsf{id}} s})$ which is a valid f-DDH instance. Now we argue security.

Lemma 2. *If H is a random oracle, the above scheme is semantically secure under adaptive identity attacks.*

Proof. We prove security in the case where $\mathcal{A}$'s challenge messages are one bit long. In other words, $\mathcal{A}$ doesn't pick two messages, rather, the challenger encrypts a uniformly random bit b. This easily extends to longer messages via a standard hybrid argument.

Assume the existence of a PPT adversary $\mathcal{A}$ that makes $n = \mathsf{poly}(\lambda)$ random oracle queries and wins the adaptive identity attack game with $\epsilon = 1/\mathsf{poly}(\lambda)$ probability. Following the proof of Boneh-Franklin security, we show a reduction $\mathcal{B}$ that solves r-DDH in $\mathbb{G}$ with $1/\mathsf{poly}(\lambda)$ probability. We reformulate the r-DDH game to be, given $g^x, g^y, g^z, g^{bx^{-1}yz+(1-b)u}$, determine b. First, $\mathcal{B}$ chooses an $i \leftarrow [n]$. We condition on $\mathcal{A}$'s challenge identity id^* corresponding to its i'th random oracle query, which occurs with probability $1/n$. We implicitly set the msk s equal to z and $H(\mathsf{id}^*) = t_{\mathsf{id}^*} = y$. $\mathcal{B}$ initializes $\mathcal{A}$ with g^z. On $\mathcal{A}$'s jth identity query id_j for $j \neq i$, $\mathcal{B}$ will choose a uniformly random $t_{\mathsf{id}_j} \leftarrow \mathbb{Z}_N$, return $g^{t_{\mathsf{id}_j}}$, and store the mapping $(\mathsf{id}_j, t_{\mathsf{id}_j})$. On $\mathcal{A}$'s ith identity query id^*, $\mathcal{B}$ returns g^y.

Now $\mathcal{A}$'s challenge can be answered as follows. Implicitly setting $sr = x$, $\mathcal{B}$ returns $(g^x, g^{bx^{-1}yz+(1-b)u})$. Note that $g^{x^{-1}yz} = g^{s^{-1}r^{-1}t_{\mathsf{id}^*}s} = \mathsf{pk}_{\mathsf{id}^*}^{r^{-1}}$, so this is a valid encryption of the bit b chosen by the r-DDH challenger. Thus $\mathcal{B}$'s advantage in determining this bit is exactly $\mathcal{A}$'s advantage in determining the bit encrypted. So overall, we have that $\mathcal{B}$ succeeds in the r-DDH game with probability $\epsilon/n = 1/\mathsf{poly}(\lambda)$, a contradiction. □

3.3 CDH Separations

Next, we study the relationship between f-CDH and r-CDH. In the usual setting where the totient of the group order is known, we refer to [Gal12] Chapter 21, which contains a proof of the polynomial equivalence of these two assumptions. However, as observed previously (for example in [SS06]), the reduction from r-CDH to f-CDH seems to require the ability to compute multiplicative inverses of unknown exponents. In a group where the totient of the group order N is known, this is easy to do by raising to the power of $\phi(N) - 1$. Otherwise, it seems that the reduction cannot go through. We give heuristic evidence that this is indeed the case, by observing that results about the generic *ring* model from [YYHK18] can be re-cast to give a separation between the two assumptions in the UO-GGM (under a variant of the factoring assumption) and giving our own separation in the UF-GGM (under a new but plausible computational assumption).

We assume throughout this section that adversaries in the UF-GGM and UO-GGM can generate uniformly random elements mod the group order. This is not exactly true, but since the group order N is known to be bounded above by 2^λ, the adversary can pick an integer k such that $k \bmod N$ is $1/2^\lambda$-close to uniform by picking k uniformly from $\mathbb{Z}_{2^{3\lambda}}$.

We first confirm that r-CDH is random self-reducible in a group of unknown order. This requires showing a reduction that does not require taking multiplicative inverses mod the group order.

Lemma 3. *Consider a group $\mathbb{G}$ of order $N \in [2^{\lambda-1}, 2^\lambda]$. An adversary $\mathcal{A}$ that does not know N and solves a uniformly random r-CDH instance in $\mathbb{G}$ with probability $\epsilon = 1/\mathsf{poly}(\lambda)$ implies the existence of an adversary $\mathcal{B}$ that does not know N and solves any fixed r-CDH instance in $\mathbb{G}$ with probability ϵ.*

Proof. First observe that the r-CDH problem described above is equivalent to the following: given (g^x, g^y, g^z), compute $g^{x^{-1}yz}$ where $x \leftarrow \mathbb{Z}_N^*, y, z \leftarrow \mathbb{Z}_N$, and inverses in the exponent are taken in $\mathbb{Z}_N^*$.

Now let (g^x, g^y, g^z) be the fixed r-CDH instance for which we want to compute $g^{x^{-1}yz}$. Choose uniform t, s, u mod the group order and compute

$$(g^{xu}, g^{(y+xt)u}, g^{z+xs}).$$

Note that g^{xu}, g^{xt} and g^{xs} are uniformly random and independent group elements, so this is a uniformly random r-CDH triple. Solving, we get

$$g^{(xu)^{-1}(y+xt)u(z+xs)} = g^{x^{-1}(xy+xys+xzt+x^2st)} = g^{x^{-1}yz+ys+zt+xst}.$$

Now, g^{ys}, g^{zt}, and g^{xst} can all be calculated from the the original three elements. Then we can take the group inverse of each and divide to recover $g^{x^{-1}yz}$. □

Now we give a formal statement of the unbalanced modulus factoring assumption and then use the results of [YYHK18] to give a separation between f-CDH and r-CDH in the UO-GGM under this assumption.

Assumption 1 (Unbalanced Modulus Factoring Assumption [Lip94, YYHK18])**.** *Let $\mathcal{P}_\lambda$ be the set of primes with bit length at most λ. Then for any PPT $\mathcal{A}$ and polynomial $\mathsf{poly}(\cdot)$,*

$$\Pr[p \leftarrow \mathcal{A}(1^\lambda, pq) \mid p \leftarrow \mathcal{P}_\lambda, q \leftarrow \mathcal{P}_{\mathsf{poly}(\lambda)}] = \mathsf{negl}(\lambda)$$

Theorem 3. *r-CDH is (T, ϵ)-secure in the UO-GGM augmented with an f-CDH oracle, for group order $p \leftarrow \mathcal{P}_\lambda, T = \mathsf{poly}(\lambda)$, and $\epsilon = \mathsf{negl}(\lambda)$.*[16]

Proof. As seen above, if an adversary $\mathcal{A}$ can solve a random instance r-CDH with an f-CDH oracle $\mathcal{O}_{fCDH}$, then it can solve r-CDH on any fixed instance. In particular, given (g^x, g^1, g^1) for some fixed prime x, $\mathcal{A}^{\mathcal{O}_{fCDH}}$ will output g^{-x} with $1/\mathsf{poly}(\lambda)$ probability. Now, we can simulate the behavior of $\mathcal{A}^{\mathcal{O}_{fCDH}}$ in the UO-GGM via access to a generic ring model of unknown (prime) order. We simply let the ring be the exponent space of the group, and implement generic group queries with additions and $\mathcal{O}_{fCDH}$ queries with multiplications. This generic ring setting of unknown order is exactly the setting considered by Yamakawa et al. [YYHK18], who show that under Assumption 1, it is hard to compute the inverse of any prime x. □

We now turn to separating f-CDH and r-CDH in the setting where the factorization of the group order is unknown (i.e. in the UF-GGM model).[17] We make the following knowledge assumption:

Assumption 2. *Let $\mathcal{A}$ be a PPT that on input an integer N of $\Theta(\lambda)$ bits, outputs an arithmetic circuit C. Say that with probability $\epsilon(\lambda)$ over the randomness of $\mathcal{A}$ and N, C implements a univariate polynomial $f(\cdot)$ whose coefficients are not all zero modulo N, but satisfies $f(x) \equiv 0 \mod N$ for all $x \in \mathbb{Z}_N$. Then there exists an "extractor" $\overline{\mathcal{A}}$ which, given the same inputs as $\mathcal{A}$, can factor N with the same probability $\epsilon(\lambda)$.*

Remarks on Assumption 2. Let $N = pq$. By elementary number theory and the Chinese Remainder Theorem, we can show that all polynomials $f(x)$ satisfying the conditions in Assumption 2 take the form $f(x) \equiv (x^p - x)r(x) + (x^q - x)s(x) \mod N$ for arbitrary (but not both zero) polynomials $r(x)$ and $s(x)$. Intuitively, it seems that the only way $\mathcal{A}$ can output an arithmetic circuit C is to "know" at least one of p and q, but we were unable to prove this.

We stress that this is a new and ad-hoc assumption, but that its statement is completely independent of cryptographic groups. In other words, we can show that the problem of separating f-CDH from r-CDH in the UO-GGM reduces to validating a knowledge assumption concerning arithmetic circuits and factoring.

Unfortunately we were unable to meaningfully relate the strength of Assumption 2 to other cryptographic assumptions or even to complexity-theoretic statements. In particular, it might be possible to give an unconditional proof that Assumption 2 holds.

Thus our purpose in stating this assumption is to mark partial progress towards proving this separation in the hopes of affirmatively resolving this problem. In the words of Goldwasser and Kalai, this assumption should be taken as an open invitation to refute or simplify [GK16].

Theorem 4. *Under Assumption 2, r-CDH is (T, ϵ)-secure in the UF-GGM augmented with an f-CDH oracle, for group order $N \in [2^{\lambda-1}, 2^\lambda], T = \mathsf{poly}(\lambda)$, and $\epsilon = \mathsf{negl}(\lambda)$.*

[16] We thank Takashi Yamakawa for pointing out an error in an earlier version of this proof.

[17] We note that the existence of such a separation was essentially conjectured by Saxena and Soh [SS06], but that they do not give any evidence to support the conjecture.

Proof. We refer to [Gal12][21.3] for a proof that r-CDH is random self-correctable, meaning an algorithm that has $1/\mathsf{poly}(\lambda)$ success probability on a random r-CDH instance can be boosted to one with $1 - \mathsf{negl}(\lambda)$ success probability on a random instance. We note this self-correction algorithm works even if the group order is unknown, as it does not require inverses in the exponent.

Then Lemma 3 and the result from [Gal12] can be combined to show that a generic adversary $\mathcal{A}$ with non-negligible advantage in the r-CDH game, given an f-CDH oracle, can be boosted to one that has overwhelming success probability on any fixed r-CDH instance. Thus we can assume that any $\mathcal{A}$ contradicting the theorem statement actually succeeds on an overwhelming fraction of inputs $(\sigma(x), \sigma(1), \sigma(1))$. Now we describe a reduction $\mathcal{B}$ that contradicts Assumption 2. On input N, $\mathcal{B}$ will let $\mathbf{x}$ be a formal variable, and simulate the generic group and f-CDH oracles over $\mathbb{Z}_N[\mathbf{x}]$, drawing the labeling function σ on the fly. It initializes $\mathcal{A}$ with $(\sigma(x), \sigma(1), \sigma(1))$ and maintains a table mapping polynomials in $\mathbb{Z}_N[\mathbf{x}]$ to labels. When $\mathcal{A}$ returns a label, $\mathcal{B}$ finds which polynomial P over $\mathbf{x}$ it corresponds to (represented as a polynomial size circuit). We are guaranteed that $P(x) = x^{-1} \bmod N$ with overwhelming probability over the choice of x. So $Q(\mathbf{x}) := \mathbf{x}P(\mathbf{x}) - 1$ is zero on an overwhelming fraction $1 - \epsilon(\lambda)$ of inputs $x \in \mathbb{Z}_N$. $\mathcal{B}$ will choose λ elements r_i uniformly at random from $\mathbb{Z}_N$ and form the polynomial $Q'(\mathbf{x}) = Q(\mathbf{x}+r_1)\dots Q(\mathbf{x}+r_\lambda)$. For any fixed x, the probability that $Q'(x) \neq 0$ is at most $\epsilon(\lambda)^\lambda$. So by a union bound, the probability that Q' is not identically zero is at most $N\epsilon(\lambda)^\lambda \leq (2\epsilon(\lambda))^\lambda = \mathsf{negl}(\lambda)$. □

3.4 Split-CDH Groups

The previous section demonstrates the feasibility of split-CDH groups. Observe that split-CDH groups are very similar to self-bilinear maps [YYHK14] (i.e. symmetric bilinear maps where the target group is the same as the source group) which in turn are sufficient for instantiating *multilinear maps* [BS02], since they allow for taking repeated products in the exponent. Consider the following random generator version of the Multilinear Computational Diffie Hellman problem:

Definition 5 (r-MCDH)**.** *Fix a group description $\mathbb{G}$ with generator g and order N. Given $(g^r, g^{rx_0}, \dots, g^{rx_n})$ compute $g^{r^n \prod_{i=0}^n x_i}$, where $r, x_i \leftarrow \mathbb{Z}_N$*

This assumption is quite powerful; for example it has been shown to imply multiparty non-interactive key exchange with trusted setup [BS02] and distributed broadcast encryption [YYHK14]. We now show that the r-MCDH problem is hard in any split-CDH group.

Lemma 4. *No PPT adversary $\mathcal{A}$ can solve r-MCDH with $1/\mathsf{poly}(\lambda)$ probability in a split-CDH group with order N of bit length $\Theta(\lambda)$.*

Proof. Say there exists an adversary $\mathcal{A}$ such that given $(g^r, g^{rx_0}, \dots, g^{rx_0})$ and an f-CDH oracle, computes $g^{r^n \prod_{i=0}^n x_i}$ with $1/\mathsf{poly}(\lambda)$ probability. Define the reduction $\mathcal{B}$ attacking r-CDH, which receives g^a, g^b, g^c as input. It picks $z_2, \dots, z_n$ uniformly at random, letting $s := \prod_{i=2}^n z_i$. Give $(g^{as}, g^b, g^c, g^{z_2}, \dots, g^{z_n})$ as input to $\mathcal{A}$. We are implicitly setting $r = as, x_0 = a^{-1}s^{-1}b, x_1 = a^{-1}s^{-1}c, x_2 = a^{-1}s^{-1}z_2, \dots, x_n = a^{-1}s^{-1}z_n$. Then $r^n \prod_{i=0}^n x_i = a^n s^n a^{-(n+1)} s^{-(n+1)} bcs = a^{-1}bc$, so if $\mathcal{A}$ is successful in breaking r-MCDH, it returns the solution to $\mathcal{B}$'s r-CDH instance. □

3.5 Discrete Log

To conclude, we observe that unlike DDH and CDH, the fixed and random generator versions of discrete log are polynomially equivalent in the three settings we consider. The random generator variant is clearly at least as hard as the fixed generator variant, so we just give the following reduction, which we believe is likely folklore.

Lemma 5. *If r-Dlog is $(T = \mathsf{poly}(\lambda), \epsilon = \mathsf{negl}(\lambda))$-secure in the GGM (or UF-GGM or UO-GGM) with group order $N \in [2^{\lambda-1}, 2^\lambda]$, then f-Dlog is $(\mathsf{poly}(\lambda), \mathsf{negl}(\lambda))$-secure.*

Proof. Assume the existence of an adversary $\mathcal{A}$ with $T = \mathsf{poly}(\lambda)$ queries that solves f-Dlog with $1/\mathsf{poly}(\lambda)$ probability. We first show that in the UO-GGM, there exists a reduction $\mathcal{B}$ with $\mathsf{poly}(\lambda)$ queries which uses $\mathcal{A}$ to determine the group order. $\mathcal{B}$ will repeatedly pick random integers t such that $t \bmod N$ is negligibly close to uniformly random, and query $\mathcal{A}$ on each t. If $\mathcal{A}$ is successful, it returns $t' \equiv t \bmod N$. In this case $N \mid t - t'$, and this can be tested by querying the generic group oracle on two integers $t - t'$ apart and seeing if the same handle is returned. By repeating a $\mathsf{poly}(\lambda)$ number of times, eventually $\mathcal{A}$ will be successful and $\mathcal{B}$ will recover a random multiple of N. This can be repeated to obtain many multiples of N and $\mathcal{B}$ finishes by taking the GCD.

Thus we can assume to be in the UF-GGM or regular GGM. Now we describe a reduction $\mathcal{C}$ which uses $\mathcal{A}$ to solve r-Dlog. On input $\sigma(x), \sigma(xy)$, $\mathcal{C}$ will query $\mathcal{A}$ on both. $\mathcal{A}$ is successful on both with $1/\mathsf{poly}(\lambda)$ and in this case $\mathcal{C}$ recovers x and xy in the clear. Now since N is known, $\mathcal{C}$ can calculate $x^{-1} \bmod N$ via the Extended Euclidean algorithm and multiply the result with xy to recover y. □

4 A Black Box Separation Between Random-Generator CDH and NIKE

We show a limit to the power of split-CDH groups. As discussed in the introduction, the r-CDH assumption can suffice for setupless Diffie-Hellman key exchange in groups where the factorization of the order is known (when r-CDH and f-CDH are equivalent).

4.1 Our Techniques

In the classic Diffie-Hellman scheme, Alice draws a random a and sends g^a, Bob draws a random b and sends g^b, and the two agree on g^{ab}. We show that in certain split-CDH groups, a large class of protocols that generalize this idea cannot be secure. In particular, we will allow Alice to draw a vector of $n = \mathsf{poly}(\lambda)$ random values $\vec{x} = (x_1, \ldots, x_n)$, compute m different multivariate polynomials $\vec{R}(\vec{x}) = (R_1(x_1, \ldots, x_n), \ldots, R_m(x_1, \ldots, x_n))$ over these values, and send $g^{\vec{R}(\vec{x})}$ as her message (interpreted as the m group elements obtained by raising g to each component of $\vec{R}(\vec{x})$). Bob does the same, sending the message $g^{\vec{R}(\vec{y})}$.

To agree on a group element, Alice and Bob compute $g^{P(\vec{R}(\vec{x}),\vec{y})} = g^{P(\vec{R}(\vec{y}),\vec{x})}$ for some polynomial P. Ideally, security is argued the same way as in standard Diffie-Hellman. We want to show that there exists a polynomial P so that Alice and Bob can agree on a group element, and that this group element looks random to an adversary who only sees $g^{\vec{R}(\vec{x})}$ and $g^{\vec{R}(\vec{y})}$.

To prove the insufficiency of r-CDH alone, we first prove Lemma 6. This is a purely mathematical claim that whenever P and $\vec{R}$ satisfy $P(\vec{R}(\vec{x}), \vec{y}) = P(\vec{R}(\vec{y}), \vec{x})$, there in fact exists another polynomial Q such that $Q(\vec{R}(\vec{x}), \vec{R}(\vec{y})) = P(\vec{R}(\vec{x}), \vec{y}) = P(\vec{R}(\vec{y}), \vec{x})$. If an adversary can break f-CDH with overwhelming advantage[18], an adversary can break security by simply computing Q in the exponent.

This of course requires Q to be efficiently computable, which places a number of restrictions on the parameters this separation applies to. We also remark that the following separation only holds in split-CDH groups in which the adversary can calculate multiplicative inverses of *constants* mod the group order. This is not a generic feature of split-CDH groups, but holds for example in split-CDH groups of known order but unknown factorization. The reason for this restriction is that we must represent the coefficients of the R_i, P, and Q polynomials below as rational numbers.

4.2 Separating r-CDH and setupless NIKE

We first define the generalized Diffie-Hellman protocols we consider, which we call "polynomial-based" non-interactive key exchange.

[18] This can be achieved by applying self-correction to an algorithm breaking f-CDH with any non-negligible advantage, see e.g. [Gal12] Section 21.3

Definition 6. *Polynomial-based non-interactive key exchange (NIKE) with parameters (n, m, d) is a non-interactive key exchange protocol taking the following form:*

- *Fix a group $\mathbb{G}$ with order p and associated generator g, a vector of n-variate polynomials $\vec{R} = (R_1, \ldots, R_m)$ and an $m+n$-variate polynomial P such that*

$$P(\vec{R}(\vec{\boldsymbol{x}}), \vec{\boldsymbol{y}}) \equiv P(\vec{R}(\vec{\boldsymbol{y}}), \vec{\boldsymbol{x}}) \bmod N$$

  where $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{y}}$ are each vectors of n formal variables and the total degree of P over $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{y}}$ is d.

- *Alice chooses $x_1, \ldots, x_n \leftarrow \mathbb{Z}_p$ and publishes $g^{\vec{R}(x_1, \ldots, x_n)}$.*

- *Bob chooses $y_1, \ldots, y_n \leftarrow \mathbb{Z}_p$ and publishes $g^{\vec{R}(y_1, \ldots, y_n)}$.*

- *Alice uses the f-CDH oracle to compute $k = P(\vec{R}(y_1, \ldots, y_n), x_1, \ldots, x_n)$ in the exponent and Bob uses the f-CDH oracle to compute $k = P(\vec{R}(x_1, \ldots, x_n), y_1, \ldots, y_n)$ in the exponent, and g^k is the shared secret.*

This class of NIKE schemes includes basic Diffie-Hellman, as well as natural extensions that would take advantage of the extra functionality given by an f-CDH oracle.

Theorem 5. *In a split-CDH group of known order, there does not exist secure polynomial-based NIKE with $(n = \mathsf{poly}(\lambda), m = \mathsf{poly}(\lambda), d = O(1))$ or $(n = \mathsf{poly}(\lambda), m = O(1), d = \mathsf{poly}(\lambda))$.*

Notation. We will use the notation $\deg_{\vec{\mathbf{x}}}(P(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ to denote the total degree of a polynomial P over only the $\mathbf{x}_1, \ldots, \mathbf{x}_n$ formal variables (i.e. treating the $\mathbf{y}_1, \ldots, \mathbf{y}_n$ formal variables as part of the "coefficients"), and define $\deg_{\vec{\mathbf{y}}}(P(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ analogously. For polynomials over only the $\vec{\mathbf{x}}$ variables, we drop the subscript and let $\deg(P(\vec{\mathbf{x}}))$ denote the total degree. Similarly, we let $\mathrm{coeffs}_{\vec{\mathbf{x}}}(P(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ denote the vector of all coefficients of $P(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ when only the $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are treated as formal variables (i.e. the coefficients are themselves polynomials over $\mathbf{y}_1, \ldots, \mathbf{y}_n$), and we define $\mathrm{coeffs}_{\vec{\mathbf{y}}}(P(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ analogously; as before, we drop the subscript when only one set $\vec{\mathbf{x}}$ or $\vec{\mathbf{y}}$ of formal variables are used. We define the maximum number of monomials in an n-variate polynomial of total degree d to be $k_{n,d} := \sum_{i=0}^{d} \binom{i+n-1}{n-1}$.

Proof. Lemma 6 below, applied with $P_1 = P_2$, shows that for any such scheme, there exists a fixed polynomial Q over the exponents of the public group elements that evaluates to the exponent of the shared secret key. Furthermore, the number of monomials in Q will be at most $k_{2m,d} = \mathsf{poly}(\lambda)$, so can be efficiently computed with an f-CDH oracle. □

Lemma 6. *Let $\vec{\boldsymbol{x}} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ and $\vec{\boldsymbol{y}} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ each denote length n vectors of formal variables. Let $\vec{R}$ denote a length m vector of n-variate polynomials $(R_1(\cdot), \ldots, R_m(\cdot))$, let $\vec{R}(\vec{\boldsymbol{x}})$ denote the result of evaluating $(R_1(\vec{\boldsymbol{x}}), \ldots, R_m(\vec{\boldsymbol{x}}))$.*

Suppose there exists two $(m+n)$-variate polynomials P_1, P_2 such that

$$P_1(\vec{R}(\vec{\boldsymbol{x}}), \vec{\boldsymbol{y}}) \equiv P_2(\vec{R}(\vec{\boldsymbol{y}}), \vec{\boldsymbol{x}})$$

as polynomials over formal variables $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$. Let

$$\begin{aligned} d = \max\{ & deg_{\vec{\boldsymbol{x}}}(P_1(\vec{R}(\vec{\boldsymbol{x}}), \vec{\boldsymbol{y}})), deg_{\vec{\boldsymbol{x}}}(P_2(\vec{R}(\vec{\boldsymbol{y}}), \vec{\boldsymbol{x}})) \\ & deg_{\vec{\boldsymbol{y}}}(P_1(\vec{R}(\vec{\boldsymbol{x}}), \vec{\boldsymbol{y}})), deg_{\vec{\boldsymbol{y}}}(P_2(\vec{R}(\vec{\boldsymbol{x}}), \vec{\boldsymbol{y}}))\}. \end{aligned}$$

Then there exists a $2m$-variate polynomial Q of total degree at most $2d$ such that

$$Q(\vec{R}(\vec{\boldsymbol{x}}), \vec{R}(\vec{\boldsymbol{y}})) \equiv P_1(\vec{R}(\vec{\boldsymbol{x}}), \vec{\boldsymbol{y}})$$

as polynomials over the same formal variables.

Proof. We consider the set of polynomials $J(x)$ that can be represented as $A(\vec{R}(\vec{\mathbf{x}}))$ for some m-variate polynomial where $\deg(A(\vec{R}(\vec{\mathbf{x}}))) \leq d$. Observe that the set of such $J(x)$ form a linear subspace of the set of all n-variate polynomials over $\vec{\mathbf{x}}$ with degree at most d. Let $\vec{v} = \text{coeffs}(A(\vec{R}(\vec{\mathbf{x}})))$ denote the $k_{n,d}$-dimensional vector of coefficients of $A(\vec{R}(\vec{\mathbf{x}}))$ (writing the coefficients according to some fixed ordering on the monomials), where $k_{n,d}$ defined above is the number of monomials over n variables of total degree at most d. Then there exists some $k_{n,d} \times k_{n,d}$ dimensional matrix $C^{(\vec{R})}$ such that $C^{(\vec{R})} \cdot \vec{v} = 0$ if and only if the polynomial satisfying $\text{coeffs}(p(\vec{\mathbf{x}})) = \vec{v}$ is of the form $p(\vec{x}) = A(\vec{R}(\vec{\mathbf{x}}))$ for some m-variate polynomial $A(\cdot)$. In other words, the set of polynomials $A(\vec{R}(\vec{\mathbf{x}}))$ form a linear subspace; in particular the subspace is spanned by the coefficient vectors corresponding to $\prod_{i=1}^{m} R_i(\vec{\mathbf{x}})^{a_i}$ for each choice of $\{a_i\}$ where $\sum_i a_i r_i \leq d, a_i \geq 0$ where r_i denotes the degree of R_i.

We can extend this characterization to describe the set of polynomials $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ which can be written as $A(\vec{R}(\vec{\mathbf{x}}), \vec{\mathbf{y}})$ for some $(m+n)$-variate polynomial $A(\cdot)$ where $\max\{\deg_{\vec{\mathbf{x}}}(A(\vec{R}(\vec{\mathbf{x}}), \vec{\mathbf{y}})), \deg_{\vec{\mathbf{y}}}(A(\vec{R}(\vec{\mathbf{x}}), \vec{\mathbf{y}}))\} \leq d$. This follows from viewing J as a polynomial over only the $\mathbf{x}_1, \ldots, \mathbf{x}_n$ formal variables, so that the coefficients are now themselves polynomials over the $\mathbf{y}_1, \ldots, \mathbf{y}_n$ formal variables. In other words, if $\text{coeffs}_{\vec{\mathbf{x}}}(J(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ denotes the coefficients of $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ when viewed as a polynomial over the $\mathbf{x}_1, \ldots, \mathbf{x}_n$ formal variables, then $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ can be written in the form $A(\vec{R}(\vec{\mathbf{x}}), \vec{\mathbf{y}})$ if and only if $C^{(\vec{R})} \cdot \vec{v}(\vec{\mathbf{y}}) = 0$.

Observe that $\vec{v}(\vec{\mathbf{y}})$ can be written in the form $V \cdot \text{powers}(\vec{\mathbf{y}}, d)$ where V is a $k_{n,d} \times k_{n,d}$ dimensional matrix consisting of the coefficients of $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ (taking the coefficients to be field elements, not polynomials) whose rows are indexed by monomials on the $\vec{\mathbf{x}}$ variables and columns are indexed by monomials on the $\vec{\mathbf{y}}$ variables, and $\text{powers}(\vec{\mathbf{y}}, d)$ denotes the $k_{n,d}$-dimensional vector of all monomials of degree at most d over the $\mathbf{y}_1, \ldots, \mathbf{y}_n$ formal variables.

Then the following three statements are equivalent:

- $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ can be written as $A(\vec{R}(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ for some $(m+n)$-variate polynomial A.
- $C^{(\vec{R})} \cdot \vec{v}(\vec{\mathbf{y}}) \equiv 0^{k_{n,d}}$.
- $C^{(\vec{R})} \cdot V = 0^{k_{n,d} \times k_{n,d}}$.

The equivalence between the last two statements follows immediately from the fact that $C^{(\vec{R})} \cdot \vec{v}(\vec{\mathbf{y}}) \equiv 0^{k_{n,d}}$ holds if and only if this vector of $k_{n,d}$ formal polynomials is a vector of $k_{n,d}$ identically zero polynomials, which holds if and only if all the coefficients of these formal polynomial equal 0; the coefficients of the polynomial in the ith entry of $C^{(\vec{R})} \cdot \vec{v}(\vec{\mathbf{y}})$ are precisely the ith row of V.

We can apply a symmetric argument to say that the same polynomial $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ can also be written in the form $A'(\vec{R}(\vec{\mathbf{y}}), \vec{\mathbf{x}})$ for some m-variate $A'(\cdot)$ if and only if $C^{(\vec{R})} \cdot V^\top = 0^{k_{n,d} \times k_{n,d}}$. This means that if a polynomial $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ can be written as both $A(\vec{R}(\vec{\mathbf{x}}), \vec{\mathbf{y}}))$ for some $A(\cdot)$ as well as $A'(\vec{R}(\vec{\mathbf{y}}), \vec{\mathbf{x}})$, for some $A'(\cdot)$, then its coefficient matrix V satisfies

$$C^{(\vec{R})} \cdot V = C^{(\vec{R})} \cdot V^\top = 0^{k_{n,d} \times k_{n,d}}.$$

The possible matrices V satisfying the above are spanned by rank 1 matrices $\vec{w}_i \cdot \vec{w}_i^\top$ where $C^{(\vec{R})} \cdot \vec{w}_i = 0^{k_{n,d}}$. For any $\vec{w}_i \cdot \vec{w}_i^\top$ satisfying

$$C^{(\vec{R})} \cdot \vec{w}_i \cdot \vec{w}_i^\top = C^{(\vec{R})} \cdot \vec{w}_i \cdot \vec{w}_i^\top = 0^{k_{n,d} \times k_{n,d}},$$

we note that the corresponding polynomial $W_i(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ satisfying $\vec{w}_i \cdot \vec{w}_i^\top =$ $\text{coeffs}(W_i(\vec{\mathbf{x}}, \vec{\mathbf{y}}))$ factors as $W_i(\vec{\mathbf{x}}, \vec{\mathbf{y}}) = W_i'(\vec{\mathbf{x}}) \cdot W_i'(\vec{\mathbf{y}})$ where $\vec{w}_i = \text{coeffs}(W_i'(\vec{\mathbf{x}})) = \text{coeffs}(W_i'(\vec{\mathbf{y}}))$, and furthermore that $W_i'(\vec{\mathbf{x}})$ has the form $A_i(\vec{R}(\vec{\mathbf{x}}))$ for some m-variate polynomial A_i. We can therefore write

$$W_i(\vec{\mathbf{x}}, \vec{\mathbf{y}}) = A_i(\vec{R}(\vec{\mathbf{x}})) \cdot A_i(\vec{R}(\vec{\mathbf{y}})).$$

Since any valid $J(\vec{\mathbf{x}}, \vec{\mathbf{y}})$ is a linear combination of such $W_i(\vec{\mathbf{x}}, \vec{\mathbf{y}})$, it follows that $J(\vec{\mathbf{x}}, \vec{\mathbf{y}}) = Q(\vec{R}(\vec{\mathbf{x}}), \vec{R}(\vec{\mathbf{y}}))$ for some polynomial Q. Note that the total degree of each W_i is at most d, so the total degree of the resulting Q is at most $2d$. □

5 Lower Bounds for Random Generator Discrete Log and CDH

We proceed to give tight lower bounds (up to logarithmic factors) for r-DLog and r-CDH in the AI-GGM, making use of the following special case of a lemma due to Yun [Yun15].

Lemma 7 (Search-by-Hyperplane-Queries [Yun15] (SHQ))**.** *Consider drawing z_1, z_2 uniformly at random from $\mathbb{Z}_N$, and allowing an adversary $\mathcal{A}$* hyperplane queries *of the form (a_1, a_2, b) where 1 is returned if $a_1 z_1 + a_2 z_2 = b$ and 0 otherwise. Then the probability that $\mathcal{A}$ outputs (z_1, z_2) after q hyperplane queries is at most q^2/N^2.*

Theorem 6. *The r-Dlog problem is $((S,T),\epsilon)$-secure in the AI-GGM for any prime $N \geq 16$ and*

$$\epsilon = \widetilde{O}\left(\frac{T^2}{N} + \left(\frac{ST^2}{N}\right)^2\right).$$

Proof. In the r-Dlog game, the challenger $\mathcal{C}$ draws $x \leftarrow \mathbb{Z}_N^*, y \leftarrow \mathbb{Z}_N$ and initializes $\mathcal{A}$ with $(\sigma(1), \sigma(x), \sigma(xy))$. $\mathcal{A}$ is successful if it outputs y after at most T generic group queries. We show that r-Dlog is $\left((S,T), O\left(\frac{T^2}{N} + \frac{T^2P^2+T^3P}{N^2}\right)\right)$-secure in the BF-GGM. Then we can apply Theorem 1 with $\gamma = 1/N$ to get the result, noting that $T' = 3$ and $\log(1/\gamma) = \log(N)$, so $P = \widetilde{O}(ST)$.

$\mathcal{A} := \mathcal{A}_2$ takes as input the advice string aux generated by $\mathcal{A}_1$, makes T adaptive queries $\{c_1^{(t)}\sigma(x) + c_2^{(t)}\sigma(xy) + c_3^{(t)}\sigma(1)\}_{t\in[T]}$ to the generic group oracle and receives $\{\sigma(c_1^{(t)}x + c_2^{(t)}xy + c_3^{(t)})\}_{t\in[T]}$ in return. Let E be the event that there exists an $a \in \mathcal{P}$ and $t \in [T]$ such that $c_1^{(t)}x + c_2^{(t)}xy + c_3^{(t)} = a$ and $c_3^{(j)} \neq a$. Then

$$\Pr_{\sigma,x,y}[y \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux})] \leq \Pr_{\sigma,x,y}[y \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid E] + \Pr_{\sigma,x,y}[y \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid \neg E].$$

We begin by analyzing the first probability in the sum. Condition on a particular image $\mathcal{L}$ of σ and a particular set of fixed points $\mathcal{P}$. The following holds for any such choice. We set up a reduction $\mathcal{B}$ which plays the SHQ game defined above and perfectly simulates the generic group game for $\mathcal{A}$. $\mathcal{B}$ has access to $\mathcal{L}, \mathcal{P}, im(\mathcal{P})$, and hyperplane query access to uniform values z_1, z_2 in $\mathbb{Z}_N$ which we implicitly set to be x, xy. We assume that $z_1 \neq 0$, which happens except with probability $1/N$. $\mathcal{B}$ operates as follows.

- Maintain a table mapping linear polynomials in $\mathbb{Z}_N[\mathbf{z_1}, \mathbf{z_2}]$ to $\mathcal{L}$. For each $a \in \mathcal{P}$, record the pair $(a, \sigma(a))$.
- Query the SHQ oracle on hyperplane $(1, 0, a)$ for each $a \in \mathcal{P}$. If any query returns 1, record the pair $(\mathbf{z_1}, \sigma(a))$, otherwise choose a uniform value r from all unused values in $\mathcal{L} \setminus im(\mathcal{P})$ and record $(\mathbf{z_1}, r)$. Do the same for $\mathbf{z_2}$. Next, store 1 along with its image. If $1 \in \mathcal{P}$ this is already done. If not, query $(1, 0, 1)$ to determine if $z_1 = 1$ and if so store 1 along with the image of $\mathbf{z_1}$. Do the same for $\mathbf{z}_2$. Otherwise, draw a uniform value r from all unused values in $\mathcal{L} \setminus im(\mathcal{P})$ and record $(1, r)$. Initialize $\mathcal{A}$ with the images of 1, $\mathbf{z_1}$, and $\mathbf{z_2}$.
- When $\mathcal{A}$ submits a query $c_1\mathbf{z_1} + c_2\mathbf{z_2} + c_3$, subtract each previously stored polynomial $Q(\mathbf{z_1}, \mathbf{z_2})$, resulting in some polynomial $k_1\mathbf{z_1} + k_2\mathbf{z_2} + k_3$. Query the SHQ oracle on $(k_1, k_2, -k_3)$. If 1 is returned, let s be the element stored along with $Q(\mathbf{z_1}, \mathbf{z_2})$, record $(c_1\mathbf{z_1} + c_2\mathbf{z_2} + c_3, s)$, and return s to $\mathcal{A}$. Otherwise, choose a uniform value r from all unused values in $\mathcal{L} \setminus im(\mathcal{P})$, record $(c_1\mathbf{z_1} + c_2\mathbf{z_2} + c_3, r)$ and return r.
- If E occurs, $\mathcal{B}$ will see a 1 returned by the SHQ oracle on a hyperplane query (k_1, k_2, k_3) for $k_3 \neq 0$, meaning at least one of $k_1, k_2 \neq 0$. Record this tuple. At the end of the interaction, $\mathcal{A}$ will return a $y \in \mathbb{Z}_N$. Now $\mathcal{B}$ outputs $(k_3(k_1 + k_2 y)^{-1}, y)$.

Setting $z_1 = x$ and $z_2 = xy$, it is clear that $\mathcal{B}$ perfectly simulates the r-Dlog game for $\mathcal{A}$. If E occurs, we know that $k_3 = k_1 x + k_2 xy = x(k_1 + k_2 y)$, and $k_3 \neq 0$, so $k_1 + k_2 y \neq 0$. Thus if $\mathcal{A}$ is successful and

returns y, $\mathcal{B}$ successfully computes $x = k_3(k_1 + k_2 y)^{-1}$. Applying Lemma 7, and noting that $\mathcal{B}$ makes less than $2(P+1) + T(P+T) = O(TP + T^2)$ queries, we get that

$$\Pr_{\sigma,x,y}[y \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid E] = O\left(\frac{T^2P^2 + T^3P + T^4}{N^2}\right).$$

To analyze the second probability, we move to a hybrid game in the BF-GGM where x and y are set to be formal variables $\mathbf{x}$ and $\mathbf{y}$ at the beginning of the game. The challenger implements group operations over $\mathbb{Z}_N[\mathbf{x}, \mathbf{y}]$, initializing its table with the points in $(a, \sigma(a))$ for all $a \in \mathcal{P}$. Every time $\mathcal{A}$ queries for a new polynomial, $\mathcal{C}$ chooses a uniform element in $\mathcal{L} \setminus im(\mathcal{P})$ among those unused so far. When $\mathcal{A}$ outputs a guess for y at the end of the game, the true value is chosen uniformly at random, so $\mathcal{A}$ wins with probability $1/N$. Given that E does not occur, $\mathcal{A}$'s probability of distinguishing these two games is bounded by the probability that in the original game, two of its T queries are different polynomials over $\mathbf{x}$ and $\mathbf{y}$ but evaluate to the same element, or there exists some query $c_1^{(j)}\mathbf{x} + c_2^{(j)}\mathbf{xy} + c_3^{(j)}$ such that $c_1^{(j)}x + c_2^{(j)}xy = 0$ and at least one of $c_1^{(j)}, c_2^{(j)} \neq 0$. So there are $O(T^2)$ possible equations that could be satisfied and by Schwartz-Zippel, each occurs with probability $O(1/N)$ over the random choice of x and y. Thus by a union bound, $\mathcal{A}$'s probability of distinguishing is $O(T^2/N)$.

Combining, we have that $\mathcal{A}$'s probability of success is

$$O\left(\frac{T^2P^2 + T^3P + T^4}{N^2}\right) + O\left(\frac{T^2}{N}\right) + O\left(\frac{1}{N}\right) = O\left(\frac{T^2}{N} + \frac{T^2P^2 + T^3P}{N^2}\right).$$

□

.

Theorem 7. *The r-CDH problem is $((S,T),\epsilon)$-secure in the AI-GGM for any prime $N \geq 16$ and*

$$\epsilon = \widetilde{O}\left(\frac{T^2}{N} + \left(\frac{ST^2}{N}\right)^2\right).$$

Proof. In the r-CDH game, the challenger $\mathcal{C}$ draws uniformly random $a \leftarrow \mathbb{Z}_N^*, x, y \leftarrow \mathbb{Z}_N$ and initializes $\mathcal{A}$ with $\sigma(1), \sigma(a), \sigma(ax), \sigma(ay)$. $\mathcal{A}$ is successful if it outputs $\sigma(axy)$ after at most T generic group queries. We show that r-CDH is $\left((S,T), O\left(\frac{T^2}{N} + \frac{T^2P^2+T^3P}{N^2}\right)\right)$-secure in the BF-GGM and apply Theorem 1 with $\gamma = 1/N$ to get the result, noting that $T' = 4$ and $\log(1/\gamma) = \log(N)$, so $P = \widetilde{O}(ST)$.

$\mathcal{A} := \mathcal{A}_2$ takes as input the advice string aux generated by $\mathcal{A}_1$, makes T adaptive queries $\{c_1^{(t)}\sigma(ax) + c_2^{(t)}\sigma(ay) + c_3^{(t)}\sigma(a) + c_4^{(t)}\sigma(1)\}_{t\in[T]}$ to the generic group oracle, and receives $\{\sigma(c_1^{(t)}ax + c_2^{(t)}ay + c_3^{(t)}a + c_4^{(t)})\}_{t\in[T]}$ in return. We define a number of events:

- E_1: there exists a $t \in [T]$ and $i \in \{1,2,3,4\}$ such that $c_i^{(t)} \in \{x, y\}$.
- E_2: there exists a $t \in [T], p \in \mathcal{P}$ such that $c_1^{(t)}ax + c_2^{(t)}ay + c_3^{(t)}a + c_4^{(t)} = p$ with $c_1^{(t)} \neq 0$ or $c_2^{(t)} \neq 0$.
- E_3: there exists a $t \in [T], p \in \mathcal{P}$ such that $c_1^{(t)}ax + c_2^{(t)}ay + c_3^{(t)}a + c_4^{(t)} = p$ with $c_1^{(t)}, c_2^{(t)} = 0, c_3^{(t)} \neq 0$.

Now we can write

$$\begin{aligned}\Pr_{\sigma,a,x,y}[\sigma(axy) \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux})] \leq{}& \Pr_{\sigma,a,x,y}[E_1] \\ &+ \Pr_{\sigma,a,x,y}[\sigma(axy) \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid \neg E_1 \wedge E_2] \\ &+ \Pr_{\sigma,a,x,y}[\sigma(axy) \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid \neg(E_1 \vee E_2) \wedge E_3] \\ &+ \Pr_{\sigma,a,x,y}[\sigma(axy) \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid \neg(E_1 \vee E_2 \vee E_3)].\end{aligned}$$

To analyze the first probability in the sum, we set up a simple reduction to r-Dlog in the BF-GGM. Assume that $c_i^{(t)} = x$ in E_1, the other case being symmetric. Then $\mathcal{B}$ will take as input $\sigma(a), \sigma(ax)$, draw y uniformly at random, and simulate $\mathcal{A}$'s view of the generic group oracle with input $\sigma(1), \sigma(a), \sigma(ax), \sigma(ay)$. This is possible since all of $\mathcal{A}$'s queries will be linear in a and ax. Whenever $\mathcal{A}$ makes a query $c_1 ax + c_2 ay + c_3 a + c_4$, $\mathcal{B}$ will query its own oracle on $c_i \sigma(a)$ for each $i \in \{1, 2, 3, 4\}$. If any queries result in the same handle as $\sigma(ax)$, $\mathcal{B}$ has determined x and will return it. Since E_1 occurs, this will eventually happen. The number of queries $\mathcal{B}$ makes to its oracle is 5 times that of $\mathcal{A}$, meaning we can appeal to the proof of Theorem 6 to say that

$$\Pr_{\sigma,a,x,y}[E_1] = O\left(\frac{T^2}{N} + \frac{T^2P^2 + T^3P}{N^2}\right).$$

To analyze the second probability in the sum, we follow the proof of Theorem 6, setting up a reduction $\mathcal{B}$ which plays the SHQ game. We only describe the differences. $\mathcal{B}$'s goal will still be to output (z_1, z_2), except now we let $\mathcal{B}$ draw uniform $a \leftarrow \mathbb{Z}_N^*$ at the beginning of the game and implicitly set $z_1 = ax$ and $z_2 = ay$. We assume that $z_1, z_2 \neq 0$ (so also that $x, y \neq 0$), with a $O(1/N)$ loss in success probability. If $\mathcal{A}$ is successful, $\mathcal{B}$ has the following equations at the end of the game (where the second equation comes from mapping $\sigma(axy)$ back to the polynomial associated with it):

- $c_1^{(t)} ax + c_2^{(t)} ay + c_3^{(t)} a + c_4^{(t)} = p$, with $c_1^{(t)} \neq 0$ or $c_2^{(t)} \neq 0$.
- $k_1 ax + k_2 ay + k_3 a + k_4 = axy$.

There are 3 cases. First let both $c_1^{(t)}, c_2^{(t)} \neq 0$. Write $x = (p - c_4^{(t)} - c_3^{(t)} a - c_2^{(t)} ay)(c_1^{(t)})^{-1} a^{-1}$ from the first equation and plug into the second equation. This results in a quadratic equation over y with non-zero coefficient $-c_2^{(t)} a (c_1^{(t)})^{-1}$ on y^2. Solve for y and plug in to the first equation to recover x and thus $z_1 = ax$ and $z_2 = ay$.

Now say that $c_1^{(t)} \neq 0$ but $c_2^{(t)} = 0$. Then solve for $x = (p - c_4^{(t)} - c_3^{(t)} a)(c_1^{(t)})^{-1} a^{-1}$ which can be recovered in the clear since there is no term involving y and $\mathcal{B}$ knows a. Then plug in a and x to the second equation, resulting in a linear polynomial over y with coefficient on y equal to $a(k_2 - x)$. Since E_1 does not occur (so $k_2 \neq x$), we know this coefficient is non-zero, so $\mathcal{B}$ can successfully solve for y. The case where $c_1^{(t)} = 0$ but $c_2^{(t)} \neq 0$ is symmetric. Thus by Lemma 7,

$$\Pr_{\sigma,a,x,y}[\sigma(axy) \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid \neg E_1 \wedge E_2] = O\left(\frac{T^2P^2 + T^3P + T^4}{N^2}\right).$$

To analyze the third probability in the sum, we again set up a reduction $\mathcal{B}$ which plays the SHQ game. This time we implicitly set $z_1 = a$ and $z_2 = ax$. Since we have a query such that $c_3^{(t)} a + c_4^{(t)} = p$, $\mathcal{B}$ determines the value of a. Then $\mathcal{B}$ learns $k_1 as + k_2 ay + k_3 a + k_4 = axy$. The coefficient of ax is $k_1 - y$, and since we know $k_1 \neq y$, this allows $\mathcal{B}$ to solve for ax. So again we have

$$\Pr_{\sigma,a,x,y}[\sigma(axy) \leftarrow \mathcal{A}^{\mathcal{O}_G}(\mathsf{aux}) \mid \neg(E_1 \vee E_2) \wedge E_3] = O\left(\frac{T^2P^2 + T^3P + T^4}{N^2}\right).$$

Finally, to determine the fourth probability, we can repeat exactly the same analysis from Theorem 6, moving to a hybrid where a, x, and y are formal variables. We again obtain a distinguishing advantage of $O(T^2/N)$. Combining all probabilities, we get that $\mathcal{A}$'s probability of success is

$$O\left(\frac{T^2}{N} + \frac{T^2P^2 + T^3P}{N^2}\right).$$

□

6 Non-Malleable Point Obfuscation

In this section, we construct a non-malleable point obfuscator secure against *polynomial mauling attacks*, which were first considered by Komargodski and Yogev [KY18a]. We first briefly review relevant definitions.

6.1 Definitions

Denote by $\mathcal{I}_x$ the function that returns 1 on input x and 0 otherwise.

Definition 7. *(Point Obfuscation) A point obfuscator for a domain $\{\mathcal{X}_\lambda\}_\lambda$ of inputs is a PPT* Obf *that takes as input a point $x \in \mathcal{X}_\lambda$ and outputs a circuit such that the following hold.*

- ***Functionality Preservation:*** *For all $\lambda \in \mathbb{N}$, there exists a negligible function μ such that for all $x \in \mathcal{X}_\lambda$,*
$$\Pr[\mathsf{Obf}(x) \equiv \mathcal{I}_x] = 1 - \mu(\lambda).$$
- ***Virtual Black Box (VBB) Security:*** *For all PPT $\mathcal{A}$ and any polynomial function p, there exists a PPT $\mathcal{S}$ such that for all $x \in \mathcal{X}_\lambda$ and any predicate $P : \mathcal{X}_\lambda \to \{0,1\}$, and all large enough λ,*
$$\left|\Pr[\mathcal{A}(\mathsf{Obf}(x)) = P(x)] - \Pr[\mathcal{S}^{\mathcal{I}_x}(\mathsf{Obf}(x)) = P(x))]\right| \leq \frac{1}{p(\lambda)}.$$

We give another property of point obfuscators first considered in [Can97] and re-defined in [BC14].

Definition 8 (Distributional Indistinguishability)**.** *Let $\{\mathcal{X}_\lambda\}_\lambda$ be a family of domains. Then a point obfuscator* Obf *for $\{\mathcal{X}_\lambda\}_\lambda$ satisfies Distributional Indistinguishability if for all PPT $\mathcal{A}$ and well-spread ensembles of distributions $\{\mathcal{D}_\lambda\}_\lambda$ over $\{\mathcal{X}_\lambda\}_\lambda$, there exists a negligble function $\mu(\lambda)$ such that*

$$|\Pr[\mathcal{A}(\mathsf{Obf}(x)) = 1] - \Pr[\mathcal{A}(\mathsf{Obf}(u)) = 1]| = \mu(\lambda),$$

where $x \leftarrow \mathcal{D}_\lambda$ and u is drawn from the uniform distribution over $\mathcal{X}_\lambda$.

[Can97, BC14] show that Distributional Indistiguishability is equivalent to VBB security for point obfuscators. Now we give the [KY18a] definition of non-malleability. This definition involves the notion of a Verifier algorithm, which simply checks that the potentially mauled obfuscation is valid.

Definition 9. *(Verifier) A PPT $\mathcal{V}$ for a point obfuscator* Obf *for an ensemble of domains $\{\mathcal{X}_\lambda\}_\lambda$ is called a Verifier if for all $\lambda \in \mathbb{N}$ and $x \in \mathcal{X}_\lambda$, it holds that $\Pr[\mathcal{V}(\mathsf{Obf}(x)) = 1] = 1$, where the probability is taken over the randomness of $\mathcal{V}$ and* Obf*.*

Definition 10. *(Non-malleable Point Function Obfuscation) Let* Obf *be a point function obfuscator for an ensemble of domains $\{\mathcal{X}_\lambda\}_\lambda$ with an associated verifier $\mathcal{V}$. Let $\{\mathcal{F}_\lambda\}_\lambda = \{f : \mathcal{X}_\lambda \to \mathcal{X}_\lambda\}_\lambda$ be an ensemble of families of functions, and let $\{\mathcal{D}_\lambda\}_\lambda$ be an ensemble of distributions over $\mathcal{X}_\lambda$. Then* Obf *is a non-malleable point obfuscator for $\mathcal{F}$ and $\mathcal{D}$ if for any PPT $\mathcal{A}$, there exists a negligible function μ such that for any $\lambda \in \mathbb{N}$,*

$$\Pr[\mathcal{V}(C) = 1, f \in \mathcal{F}_\lambda, C \equiv \mathcal{I}_{f(x)} \mid x \leftarrow \mathcal{D}_\lambda, (C, f) \leftarrow \mathcal{A}(\mathsf{Obf}(x))] \leq \mu(\lambda).$$

In the following, we rely on the existence of a *pseudo-deterministic* $\mathsf{GroupGen}$ algorithm that may use randomness, but on input the security parameter 1^λ outputs a *unique* description of a group $\mathbb{G}_\lambda$ with a unique generator g and prime order $p(\lambda) \in [2^{\lambda-1}, 2^\lambda]$. As discussed in the introduction, this would involve psuedo-deterministic generation of large primes. This is not provably efficient, but we can rely for example on Cramer's conjecture to argue efficiency. See [GG11] for further discussion on psuedo-deterministic algorithms, including group generator generation.

6.2 Assumptions

Assumption 3. *Let* $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$*, where* $2^{\lambda-1} < p(\lambda) < 2^\lambda$*. Let* $\{\mathcal{D}_\lambda\}$ *be a family of well-spread distributions where the domain of* $\mathcal{D}_\lambda$ *is* $\mathbb{Z}_{p(\lambda)}$*. Then for any* $n = poly(\lambda)$*, for any PPT* $\mathcal{A}$*,*

$$\left|\Pr[\mathcal{A}(\{k_i, g^{k_i x + x^i}\}_{i\in[2,\ldots,n]}) = 1] - \Pr[\mathcal{A}(\{k_i, g^{k_i r + r^i}\}_{i\in[2,\ldots,n]}) = 1]\right| = \mathsf{negl}(\lambda),$$

where $x \leftarrow \mathcal{D}_\lambda$*,* $r \leftarrow \mathbb{Z}_{p(\lambda)}$*, and* $k_i \leftarrow \mathbb{Z}_{p(\lambda)}$*.*

Assumption 4. *Let* $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$*, where* $2^{\lambda-1} < p(\lambda) < 2^\lambda$*. Let* $\{\mathcal{D}_\lambda\}$ *be a family of well-spread distributions where the domain of* $\mathcal{D}_\lambda$ *is* $\mathbb{Z}_{p(\lambda)}$*. Then for any* $n = poly(\lambda)$*, for any PPT* $\mathcal{A}$*,*

$$\Pr[g^x \leftarrow \mathcal{A}(\{k_i, g^{k_i x + x^i}\}_{i\in[2,\ldots,n]})] = \mathsf{negl}(\lambda),$$

where $x \leftarrow \mathcal{D}_\lambda$ *and* $k_i \leftarrow \mathbb{Z}_{p(\lambda)}$*.*

.

Lemma 8. *Assumption 3 implies Assumption 4.*

Proof. We first give the following intermediate assumption. For binary strings s_1, s_2, let $\langle s_1, s_2\rangle$ denote their inner product mod 2.

Assumption 5. *Let* $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$*, where* $2^{\lambda-1} < p(\lambda) < 2^\lambda$*. Let* $\sigma : \mathbb{G}_\lambda \to \{0,1\}^{\ell(\lambda)}$ *be an arbitrary embedding of group elements into binary strings. Let* $\{\mathcal{D}_\lambda\}$ *be a family of well-spread distributions where the domain of* $\mathcal{D}_\lambda$ *is* $\mathbb{Z}_{p(\lambda)}$*. Then for any* $n = poly(\lambda)$*, for any PPT* $\mathcal{A}$*,*

$$\begin{aligned}&\Big|\Pr[\mathcal{A}(s, \langle s, \sigma(g^x)\rangle, \{k_i, g^{k_i x + x^i}\}_{i\in[2,\ldots,n]}) = 1]\\ &- \Pr[\mathcal{A}(s, \langle s, \sigma(g^x)\rangle, \{k_i, g^{k_i r + r^i}\}_{i\in[2,\ldots,n]}) = 1]\Big| = \mathsf{negl}(\lambda),\end{aligned}$$

where $x \leftarrow \mathcal{D}_\lambda$*,* $r \leftarrow \mathbb{Z}_{p(\lambda)}, k_i \leftarrow \mathbb{Z}_{p(\lambda)}$*, and* $s \leftarrow \{0,1\}^{\ell(\lambda)}$*.*

First we argue that Assumption 3 implies Assumption 5. Say there exists an adversary $\mathcal{A}$ and well-spread distribution ensemble $\{\mathcal{D}_\lambda\}$ such that $\mathcal{A}$ breaks Assumption 5. Then for infinitely many λ, there exists a fixed string s^*_λ for which $\mathcal{A}$ distinguishes with $1/\mathsf{poly}(\lambda)$ advantage. Let $\mathcal{D}^{(0)}_\lambda$ be the distribution $\mathcal{D}_\lambda$ restricted to x such that $\langle s^*_\lambda, \sigma(g^x)\rangle = 0$ and define $\mathcal{D}^{(1)}_\lambda$ analogously. Now there are two cases. If $\Pr[\langle s^*_\lambda, \sigma(g^x)\rangle = b \mid x \leftarrow \mathcal{D}_\lambda] = \mathsf{negl}(\lambda)$ for some bit b, then $\mathcal{A}$ must break Assumption 3 when $x \leftarrow \mathcal{D}^{(1-b)}_\lambda$, since $\mathcal{D}^{(1-b)}_\lambda$ is negligibly close to $\mathcal{D}_\lambda$, and since $(s, \langle s, \sigma(g^x)\rangle)$ can be fixed to be constant parameters $(s^*_\lambda, 1-b)$. Otherwise, both $\mathcal{D}^{(0)}_\lambda$ and $\mathcal{D}^{(1)}_\lambda$ have $\omega(\log(\lambda))$ min-entropy. Obverse that there must be a fixed bit b for which $\mathcal{A}$ distinguishes with $1/\mathsf{poly}(\lambda)$ advantage when $s = s^*_\lambda$ and $\langle s^*_\lambda, \sigma(g^x)\rangle = b$. But this is exactly the Assumption 3 distinguishing game with well-spread distribution $\mathcal{D}^{(b)}_\lambda$. Thus there exists a distribution for which $\mathcal{A}$ breaks Assumption 3 for infinitely many settings of λ.

Next we argue that Assumption 5 implies Assumption 4. Assume the existence of an adversary $\mathcal{A}$ that succeeds in the Assumption 4 game with non-negligible probability $\epsilon(\lambda)$. Consider the following reduction $\mathcal{B}$ that takes as input $s, b := \langle s, \sigma(g^x)\rangle$ and either $\{k_i, g^{k_i x + x^i}\}_{i\in[2,\ldots,n]}$ or $\{k_i, g^{k_i r + r^i}\}_{i\in[2,\ldots,n]}$. It forwards its set of $n-1$ group elements to $\mathcal{A}$, which returns some group element h. $\mathcal{B}$ outputs 1 if $\langle s, \sigma(h)\rangle = b$. If $\mathcal{B}$ received $\{k_i, g^{k_i r + r_i}\}_{i\in[2,\ldots,n]}$, then the view of $\mathcal{A}$ is independent of s and g^x. Thus we can imagine drawing s and x after $\mathcal{A}$ has returned the group element h. We condition on $h \neq g^x$, which occurs with overwhelming probability due the min-entropy requirement on x. Then the probability $\mathcal{B}$ outputs 1 is the probability that $\langle s, (\sigma(g^x) - \sigma(h))\rangle = 0$, which is exactly $1/2$ since $\sigma(g^x) - \sigma(h) \neq 0^{\ell(\lambda)}$. So overall, the probability that $\mathcal{B}$ outputs 1 is at most $1/2 + \mathsf{negl}(\lambda)$. Now if $\mathcal{B}$ received $\{k_i, g^{k_i x + x^i}\}_{i\in[2,\ldots,n]}$, then with probability $\epsilon(\lambda)$, $\mathcal{A}$ returns $h = g^x$ and with probability $1 - \epsilon(\lambda)$, $\mathcal{A}$ returns $h = g^y$ for some $y \neq x$. Since s is independent

of $\mathcal{A}$'s view, we can use the same argument to show that in this latter case, $\mathcal{B}$ outputs 1 with probability exactly 1/2. In the former case, $\mathcal{B}$ outputs 1 with probability 1. Then overall, $\mathcal{B}$ outputs 1 with probability $\epsilon(\lambda) + \frac{1}{2}(1-\epsilon(\lambda)) = \frac{1}{2} + \frac{\epsilon(\lambda)}{2}$, which is non-negligibly greater $\frac{1}{2} + \mathsf{negl}(\lambda)$, thus breaking Assumption 5.

□

6.3 The Obfuscator

Our obfuscation consists of three scalars and three group elements. We remark that the first group element is sufficient for our proof on non-malleability, but that we include the next two to obtain functionality preservation.

- $\mathsf{Obf}(1^\lambda, x)$: Compute $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$. Draw $a, b, c \leftarrow \mathbb{Z}_{p(\lambda)}$ and output

$$a, b, c, g^{ax+x^2+x^3+x^4+x^5}, g^{bx+x^6}, g^{cx+x^7}.$$

- $\mathsf{Eval}(1^\lambda, (a, b, c, h_a, h_b, h_c), x)$: Compute $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$. Accept if and only if

$$h_a = g^{ax+x^2+x^3+x^4+x^5}, h_b = g^{bx+x^6}, h_c = g^{cx+x^7}.$$

Theorem 8. *The above point obfuscator satisfies functionality preservation.*

Proof. Fix a point $x \in \mathbb{Z}_{p(\lambda)}$. We show the probability that there exists a $y \neq x$ such that $\mathsf{Eval}(1^\lambda, \mathsf{Obf}(1^\lambda, x), y)$ accepts is at most $4/p(\lambda)^2$. Union bounding over all x completes the proof.

The randomness in Obf consists of the elements a, b, c. Fix just a for now and let $t = ax+x^2+x^3+x^4+x^5$. Then any y which causes Eval to accept satisfies $ay+y^2+y^3+y^4+y^5 = t$. This leaves four possible $y \neq x$. For each such y, we write $P(\mathbf{b}) = (x^6-y^6)+(x-y)\mathbf{b}$ and $Q(\mathbf{c}) = (x^7-y^7)+(x-y)\mathbf{c}$ which are linear polynomials over $\mathbf{b}$ and $\mathbf{c}$ respectively with non-zero linear coefficient. Then y only causes Eval to accept if $P(b) = 0$ and $P(c) = 0$. But these occur simultaneously with probability $1/p(\lambda)^2$ over the uniform randomness of b, c. So by a union bound, there exists a $y \neq x$ such that $\mathsf{Eval}(1^\lambda, \mathsf{Obf}(1^\lambda, x), y)$ accepts with probability at most $4/p(\lambda)^2$. □

Theorem 9. *Under Assumption 3, the above point obfuscator satisfies Virtual Black Box Security.*

Proof. The obfuscator satisfies distributional indistinguishability, which follows directly from Assumption 3 with $n = 7$. A reduction simply receives $\{k_i, h_i\}_{i\in[2,\ldots,7]}$ and forms the obfuscation $(\sum_{i=2}^5 k_i, k_6, k_7, \prod_{i=2}^5 h_i, h_6, h_7)$. As mentioned earlier, this is equivalent to VBB security.

□

Theorem 10. *Let $\{\mathcal{D}_\lambda\}$ be a well-spread distribution ensemble with domain $\{\mathbb{Z}_{p(\lambda)}\}_\lambda$. Let $\mathcal{F}_{poly} = \{f_\lambda : \mathbb{Z}_{p(\lambda)} \to \mathbb{Z}_{p(\lambda)}\}_\lambda$ be the ensemble of functions where f_λ is the set of non-constant, non-identity polynomials [19] in $\mathbb{Z}_{p(\lambda)}[x]$ with $\mathsf{poly}(\lambda)$ degree. Then under Assumption 3, the above obfuscator is non-malleable for $\mathcal{F}_{poly}$ and distribution ensemble $\{\mathcal{D}_\lambda\}$.*

Proof. First, we fix the verifier to check that the Eval circuit is using the g output by $\mathsf{GroupGen}(1^\lambda)$. Now we show that any mauling adversary $\mathcal{A}$ can be used to break Assumption 4, which as seen above follows from Assumption 3.

We first handle the case where $\mathcal{A}$ outputs an f of degree at least 2. Let $m \geq 2$ be the degree of $\mathcal{A}$'s polynomial. We define the following reduction $\mathcal{B}$.

- Receive $\{k_i, h_i\}_{i\in[2,\ldots,7m]} := \{k_i, g^{k_i x + x^i}\}_{i\in[2,\ldots,7m]}$ from the Assumption 4 challenger, where $x \leftarrow \mathcal{D}_\lambda$.

[19] Note that constant and identity polynomials correspond to 'trivial' mauling attacks that cannot be prevented. A constant polynomial corresponds to picking an unrelated y and obfuscating y, while the identity polynomial corresponds to doing nothing.

- Send $(\sum_{i=2}^{5} k_i, k_6, k_7, \prod_{i=2}^{5} h_i, h_6, h_7)$ to $\mathcal{A}$, which returns $(f, a, b, c, j_a, j_b, j_c)$ where $a, b, c \in \mathbb{Z}_{p(\lambda)}$ and j_a, j_b, j_c are group elements.
- Compute $cf(x) + f(x)^7 = \ell_0 + \ell_1 x + \cdots + \ell_{7m} x^{7m}$.
- Return $(j_c/(g^{\ell_0} \prod_{i=2}^{7m} (h_i^{\ell_i})))^{1/(\ell_1 - \sum_{i=2}^{7m} k_i \ell_i)}$.

$\mathcal{B}$ perfectly simulates the obfuscation for $x \leftarrow \mathcal{D}_\lambda$ for $\mathcal{A}$, which is guaranteed to return a valid obfuscation of $f(x)$ with $1/\mathsf{poly}(\lambda)$ probability. In this case, $f_c = g^{\ell_0 + \ell_1 x + \cdots + \ell_{7m} x^{7m}}$. Then $\mathcal{B}$ successfully computes g^x unless $\ell_1 - \sum_{i=2}^{7m} k_i \ell_i = 0$. We know that $\ell_{7m} \neq 0$ and that k_{7m} is uniformly random and independent of $\mathcal{A}$'s view, so this occurs with probability at most $1/p(\lambda) = \mathsf{negl}(\lambda)$. Thus, $\mathcal{B}$ breaks Assumption 4 with $1/\mathsf{poly}(\lambda)$ probability.

In the case that f is linear, we set up the same reduction $\mathcal{B}$, except for the last two steps.

- Compute $af(x) + f(x)^2 + f(x)^3 + f(x)^4 + f(x)^5 = \ell_0 + \ell_1 x + \cdots + \ell_5 x^5$.
- Return $(j_a/(g^{\ell_0} \prod_{i=2}^{5} (h_i^{\ell_i})))^{1/(\ell_1 - \sum_{i=2}^{5} k_i \ell_i)}$.

Like before, it suffices to argue that $\ell_1 - \sum_{i=2}^{5} k_i \ell_i \neq 0$ except with negligible probability. In this case, the adversary receives $z := k_2 + k_3 + k_4 + k_5$. Thus letting $k_5 = z - k_2 - k_3 - k_4$, there are 3 free variables k_2, k_3, k_4 in $\mathcal{A}$'s view. We can then re-write $\ell_1 - \sum_{i=2}^{5} k_i \ell_i \neq 0$ as

$$\ell_1 - \ell_5 z + (\ell_5 - \ell_2)k_2 + (\ell_5 - \ell_3)k_3 + (\ell_5 - \ell_4)k_4.$$

So in order for this to evaluate to 0 with non-negligible probability, each of the coefficients on k_2, k_3, k_4 must be 0. Let $f(x) = rx + s$. Then writing out what the ℓ_i are, we see that the following must hold.

$$r^5 = 5r^4 s + r^4 = 10r^3 s^2 + 4r^3 s + r^3 = 10r^2 s^3 + 6r^2 s^2 + 3r^2 s + r^2$$

It is easily verified that the only solutions to the above system are when $r = 0$ or $(r = 1, s = 0)$. These correspond to when f is constant or the identity, so we can conclude that if $\mathcal{A}$ succeeds in breaking non-malleability, $\mathcal{B}$ breaks Assumption 4 with $1/\mathsf{poly}(\lambda)$ probability.

□

7 Justifying Assumptions in the Generic Group Model

We will need some additional background from [CDG18], plus a couple of new simple lemmas. Note that while we make use of techniques from [CDG18] that establish theorems relating the AI-GGM and BF-GGM, we never technically operate in the BF-GGM. We need a more fine-grained approach, starting in the plain GGM and modifying the labeling function and challenger's game incrementally.

7.1 Background

Definition 11 ([CDG18])**.** *An (N, M)-injection source Σ is a random variable that takes on as value function tables corresponding to injections $\sigma : [N] \to [M]$. An (N, M)-injection source Σ is called $(P, \mathcal{L}, 1 - \delta)$-dense for $\mathcal{L} \subseteq [M]$ if it is fixed on at most P coordinates and if for every subset I of non-fixed coordinates,*

$$H_\infty(\Sigma_I) \geq (1 - \delta) \log \left(\frac{(N - P)!}{(N - P - |I|)!} \right),$$

where Σ_I is the random variable Σ restricted to the coordinates in I. When $\delta = 0$, the source is called $(P, \mathcal{L})$-fixed.

Remark 3. *We denote by $\mathcal{A}^{\Sigma}$ an algorithm that has oracle access to an injection σ drawn from the source Σ. This means that $\mathcal{A}$ can perform forward queries where on input x the oracle returns $\sigma(x)$ or backward queries where on input x the oracle returns $\sigma^{-1}(x)$.*

Lemma 9 ([CDG18])**.** *Let Σ be a uniform (N, M)-injection source and $f : [M]^{[N]} \to \{0,1\}^S$ a potentially randomized function. Let $\Sigma_{f,x,\mathcal{L}}$ be the random variable corresponding to the distribution of Σ conditioned on $f(\Sigma) = x$ and $im(\Sigma) = \mathcal{L}$. Then for any $\gamma > 0, P \in \mathbb{N}$, there exists a family $\{Y_{x,\mathcal{L}}\}_{x,\mathcal{L}}$, indexed by values $x \in \{0,1\}^S$ and size-N subsets $\mathcal{L}$ of $[M]$, of convex combinations $Y_{x,\mathcal{L}}$ of $(P, \mathcal{L}, 1 - \frac{S+\log(1/\gamma)}{P \log(N/e)})$-dense sources, such that $\Sigma_{f,x,\mathcal{L}}$ is γ-close to $Y_{x,\mathcal{L}}$. Furthermore, replacing each $Y_{x,\mathcal{L}}$ with its corresponding convex combination $Z_{x,\mathcal{L}}$ of $(P, \mathcal{L})$-fixed sources, we have that for any distinguisher $\mathcal{D}$ taking an S-bit input and making at most T queries to its injection oracle,*

$$|\Pr[\mathcal{D}^{\Sigma}(f(\Sigma)) = 1] - \Pr[\mathcal{D}^{Z_{f(\Sigma),im(\Sigma)}}(f(\Sigma)) = 1]| \leq \frac{2(S + \log 1/\gamma) \cdot T}{P} + \gamma.$$

The above is actually slightly modified from the statement in [CDG18], with the only difference being that we allow f to be randomized. The only place in their proof that makes use of f being deterministic is Claim 19, essentially that (where everything is conditioned on some range $\mathcal{L}$), $E_x[H_\infty(\Sigma|f(\Sigma) = x)] \geq \log(N!) - S$. Their proof of this claim can easily be adapted to allow randomized f. Say that f uses k uniformly random bits. Then define the deterministic function $f' : \{0,1\}^k \times [N]^{[N]} \to \{0,1\}^S$ that runs f using its first input as the randomness. Let K be the random variable corresponding to drawing a uniformly random string in $\{0,1\}^k$. Now by averaging, we have that for any x, $H_\infty(\Sigma|X = x) \geq H_\infty((K, \Sigma)|X = x) - k$. Then, following the proof in [CDG18],

$$\begin{aligned} &E_x[H_\infty(\Sigma|f(\Sigma) = x)] \geq E_x[H_\infty((K, \Sigma)|f'(K, \Sigma) = x)] - k \\ &= E_x[H((K, \Sigma)|f'(K, \Sigma) = x)] - k \geq \log(N!) + k - S - k = \log(N!) - S, \end{aligned}$$

where H is Shannon entropy, and the equality is due to the fact that conditioned on x, (K, Σ) is uniform over all values (r, σ) such that $f'(r, \sigma) = x$.

Lemma 10 ([CDG18])**.** *For any $(P, N, 1-\delta)$-dense (N, N)-injection (bijection) source Y and its corresponding (P, N)-fixed source Z, it holds that for any (adaptive) distinguisher $\mathcal{D}$ that makes at most T queries to its oracle,*

$$|\Pr[\mathcal{D}^Y = 1] - \Pr[\mathcal{D}^Z = 1]| \leq T\delta \log N.$$

Now we give two additional lemmas, useful for proving Theorem 11.

Lemma 11. *Let Σ be a uniform (N, N)-injection (bijection) source with $\log(N) = \Theta(\lambda)$ and $f : [N]^{[N]} \to \{0,1\}^S$ a potentially randomized function. Let Σ' be the random variable on σ' that results from drawing $\sigma \leftarrow \Sigma$, $x \leftarrow f(\sigma)$, and then $\sigma' \leftarrow \Sigma_{f,x,[N]}$ defined in Lemma 9. Say that for all σ, $H_\infty(X|\Sigma = \sigma) = \omega(\log(\lambda))$. Then*

$$E_{\Sigma'}[\max_x \{\Pr[X = x|\Sigma' = \sigma]\}] = \mathsf{negl}(\lambda).$$

Proof. With two applications of Bayes' Theorem, we see that for any $x \in \{0,1\}^S$

$$\begin{aligned} \Pr[X = x|\Sigma' = \sigma] &= \frac{\Pr[\Sigma' = \sigma|X = x] \Pr[X = x]}{\Pr[\Sigma' = \sigma]} = \frac{\Pr[\Sigma = \sigma|X = x] \Pr[X = x]}{Pr[\Sigma' = \sigma]} \\ &= \frac{\left(\frac{\Pr[X=x|\Sigma=\sigma] \Pr[\Sigma=\sigma]}{\Pr[X=x]}\right) \Pr[X = x]}{\Pr[\Sigma' = \sigma]} = \Pr[X = x|\Sigma = \sigma] \left(\frac{\Pr[\Sigma = \sigma]}{\Pr[\Sigma' = \sigma]}\right). \end{aligned}$$

So plugging in,

$$E_{\Sigma'}[\max_x\{\Pr[X=x|\Sigma'=\sigma]\}] = \sum_\sigma \max_x\{\Pr[X=x|\Sigma'=\sigma]\}\Pr[\Sigma'=\sigma]$$
$$= \sum_\sigma \max_x\{\Pr[X=x|\Sigma=\sigma]\Pr[\Sigma=\sigma]\} \leq \max_{x,\sigma}\{\Pr[X=x|\Sigma=\sigma]\} = \mathsf{negl}(\lambda).$$

□

Lemma 12. *Consider n events $X_1,\ldots,X_n$ such that each event occurs with probability at least α, where $\alpha > 2/n$. Then for a uniformly random $i,j \leftarrow [n]$, $\Pr[X_i \wedge X_j] \geq \frac{\alpha^2}{4}$.*

Proof. Let $m = 2/\alpha < n$. Picking a uniform pair i,j is equivalent to picking a uniform subset of size m and then picking i,j from that set. Let $Y_1,...,Y_m$ be these m events, where each still occurs with probability at least α. Then

$$1 \geq \Pr[Y_1 \vee \cdots \vee Y_m] \geq \sum_{1\leq i\leq m}\Pr[Y_i] - \sum_{1\leq i<j\leq m}\Pr[Y_i \wedge Y_j]$$
$$= \sum_{1\leq i\leq m}\Pr[Y_i] - m^2\Pr[Y_i \wedge Y_j : i,j \leftarrow [m]]$$
$$\geq m\alpha - m^2\Pr[Y_i \wedge Y_j : i,j \leftarrow [m]].$$

Solving for $\Pr[Y_i \wedge Y_j : i,j \leftarrow [m]]$ gives $\frac{1}{m}(\alpha - \frac{1}{m}) = \frac{\alpha^2}{4}$. □

.

7.2 Proofs

Theorem 11. *Assumption 3 (Section 6) holds in the Generic Group Model.*

Proof. We define the following hybrid games.

- **Hybrid 0.** The Assumption 3 distinguishing game for generic adversary $\mathcal{A}$.

  Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

  The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma, x)$, chooses $b \leftarrow \{0,1\}, r, k_i \leftarrow \mathbb{Z}_p$ for $i \in [2,...,n]$, and initializes the adversary $\mathcal{A}$ with $\{k_i, \sigma(b(k_i x + x^i) + (1-b)(k_i r + r^i))\}_{i\in[2,\ldots n]}$. The challenger $\mathcal{C}$ proceeds to implement the generic group oracle for $\mathcal{A}$, after which $\mathcal{A}$ outputs a guess $b' \in \{0,1\}$. $\mathcal{A}$ wins if $b' = b$.

- **Hybrid 1.** In this hybrid, we switch to a "bit-fixing" labeling σ'.

  Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

  Let $Z_{x,im(\sigma)}$ be the family defined as in Lemma 9 (parameterized by some $P \in \mathbb{N}$ and $\gamma := 1/2^\lambda$). Sample $\sigma' \leftarrow Z_{x,im(\sigma)}$

  The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma', x)$, chooses $b \leftarrow \{0,1\}, r, k_i \leftarrow \mathbb{Z}_p$ for $i \in [2,...,n]$, and initializes the adversary $\mathcal{A}$ with $\{k_i, \sigma'(b(k_i x + x^i) + (1-b)(k_i r + r^i))\}_{i\in[2,\ldots n]}$. The challenger $\mathcal{C}$ proceeds to implement the generic group oracle for $\mathcal{A}$, after which $\mathcal{A}$ outputs a guess $b' \in \{0,1\}$. $\mathcal{A}$ wins if $b' = b$.

Now we assume the existence of an adversary $\mathcal{A}$ that makes $T(\lambda) = \mathsf{poly}(\lambda)$ queries and attains non-negligible advantage $\epsilon(\lambda)$ in **Hybrid 0**. Let $q(\lambda) = \mathsf{poly}(\lambda)$ be such that $q(\lambda) > 1/\epsilon(\lambda)$ for infinitely many λ. Let $T := T(\lambda)$ and $q := q(\lambda)$. Set $P = 30\lambda T^4 q = \mathsf{poly}(\lambda)$.

Claim 1. *$\mathcal{A}$ attains advantage at least $1/2q$ in **Hybrid 1**.*

Consider the following distinguisher $\mathcal{D}(x)$, which interacts with an oracle injection source mapping $[p] \to [p']$, and receives as input $x \leftarrow S(\sigma)$. $\mathcal{D}$ simulates the interaction between $\mathcal{C}$ and $\mathcal{A}$ described in **Hybrid 0** and outputs a bit indicating whether $\mathcal{A}$ was successful or not. If the injection source that $\mathcal{D}$ is interacting with is σ, then the simulation is exactly **Hybrid 0**. If it is $Z_{x,im(\sigma)}$, then the simulation is exactly **Hybrid 1**.

Applying Lemma 9 with the sampler $x \leftarrow S(\sigma)$ as the function f, we have that the success probability of $\mathcal{A}$ in **Hybrid 1** must be at least

$$\epsilon(\lambda) - \frac{2T(\log p + \log(1/\gamma))}{P} - \gamma \geq \frac{1}{q} - \frac{4\lambda T}{30\lambda T^4 q} - \frac{1}{2^\lambda} \geq \frac{1}{2q}.$$

We show that $\mathcal{A}$ obtaining this advantage leads to a contradiction. Condition on $im(\sigma) = \mathcal{L}$ for some $\mathcal{L}$ where $\mathcal{A}$ obtains at least advantage $1/2q$. Here Σ is defined as in Lemma 9, except $[M]$ is fixed to be $\mathcal{L}$, resulting in a bijection source. We drop subscripts from the associated distributions, so $Y_x := Y_{x,\mathcal{L}}$, $Z_x := Z_{x,\mathcal{L}}$, and $\Sigma_x := \Sigma_{S,x,\mathcal{L}}$. The distribution Z_x is a convex combination of bit-fixing distributions $\mathcal{B}_x^{(j)}$ with associated fixed points $\mathcal{P}_x^{(j)}$. Let this convex combination be $\mathcal{J}_x$. So to draw σ' from Z_x, we draw $j \leftarrow \mathcal{J}_x$, then $\sigma' \leftarrow \mathcal{B}_x^{(j)}$.

Now we analyze the adversary's generic group oracle queries. Any query $\mathcal{A}$ makes can be viewed as a linear polynomial over its challenge elements

$$\ell_1 + \sum_{i=2}^{n} \ell_i(b(k_i x + x^i) + (1-b)(k_i r + r^i)),$$

specified by coefficients $[\ell_1, ..., \ell_n]$. We split these queries into two parts based on whether the linear polynomial is constant or non-constant over the challenge elements (whether there is some $i \in [2, ..., n]$ such that $\ell_i \neq 0$). We will consider each initial handle that $\mathcal{A}$ receives as a non-constant query where $\ell_i = 1$ for some i and $\ell_j = 0$ for $j \neq i$. Assume without loss of generality that all of $\mathcal{A}$'s queries are distinct linear combinations.

Note that constant queries are identically distributed in the $b = 0$ and $b = 1$ cases. Let $\mathcal{T}_c$ denote the set of constants that are queried by $\mathcal{A}$ throughout its interaction. Then observe that if, for both settings of b, all of $\mathcal{A}$'s non-constant queries result in distinct group elements that each lie outside of the set $\mathcal{P}_x^{(j)} \cup \mathcal{T}_c$, the oracle responses are identically distributed in both cases. Now, for any T-query adversary that at some point queries two distinct non-constant linear polynomials that evaluate to the same point, we can define a T^2-query adversary that at some point queries a non-constant linear polynomial that evaluates to zero. Redefine $\mathcal{A}$ to be this latter adversary. Thus if $\mathcal{A}$ distinguishes, it must at some point form a non-constant query that evaluates to a value in $\mathcal{P}_x^{(j)} \cup \mathcal{T}_c \cup \{0\}$.

For a given query t, let $\mathcal{T}_c^{(t)}$ denote the set of constants among the first t queries made by $\mathcal{A}$. There must exist some query t such that both of the following hold with probability $1/(2qT^2)$.

- t is non-constant and evaluates to an element in $\mathcal{P}_x^{(j)} \cup \mathcal{T}_c^{(t)} \cup \{0\}$ OR t is a constant c and there exists an earlier non-constant query t' such that query t' evaluates to c
- all previous non-constant queries (except perhaps t') evaluate to an element outside of $\mathcal{P}_x^{(j)} \cup \mathcal{T}_c^{(t)} \cup \{0\}$

Otherwise, by a union bound, $\mathcal{A}$ could not obtain distinguishing success $1/(2q)$. Note that every non-constant query prior to t except perhaps t' is answered with a uniformly random value in $\mathcal{L} \setminus im(\mathcal{P}_x^{(j)} \cup \mathcal{T}_c^{(t)} \cup$

$\{0\}$). Since $|\mathcal{P}_x^{(j)} \cup \mathcal{T}_c^{(t)} \cup \{0\}| = \mathsf{poly}(\lambda)$, we can imagine instead drawing each response uniformly from $\mathcal{L}$, which by a union bound will change $\mathcal{A}$'s view with negligible probability. Then $\mathcal{A}$ can simulate these answers itself with uniform randomness, with a negligible difference in success probability.

Now we are left with an adversary $\mathcal{A}$ that takes as input $\{k_i\} := \{k_i\}_{i\in[2,\ldots,n]}$, makes at most T^2 queries to σ', and outputs a set of coefficients $[\ell_1, ..., \ell_n]$ (representing the non-constant query t or t'). Define $\mathcal{P}'^{(j)}_x := \mathcal{P}_x^{(j)} \cup \mathcal{T}_c^{(t)} \cup \{0\}$.

Now we break up the analysis into whether $b = 0$ or $b = 1$. If $b = 0$, we are guaranteed that with probability $1/(2qT^2) - \mathsf{negl}(\lambda) = 1/\mathsf{poly}(\lambda)$ over all randomness in the game setup, $\{k_i\}$, and $\mathcal{A}$, the following holds.

$$\ell_1 + \sum_{i=2}^{n} \ell_i(k_i r + r^i) \in \mathcal{P}'^{(j)}_x$$

But note that r is drawn uniformly at random from a set of size p, *independently* of $\mathcal{A}$'s view. Thus by Schwartz-Zippel and a union bound, the above holds with probability at most $(T^2 + P + 1)n/p = \mathsf{negl}(\lambda)$.

Now let $b = 1$. We are guaranteed that with probability $1/(2qT^2) - \mathsf{negl}(\lambda)$ over all randomness in the game setup, $\{k_i\}$, and $\mathcal{A}$, the following holds:

$$\ell_1 + \sum_{i=2}^{n} \ell_i(k_i x + x^i) \in \mathcal{P}'^{(j)}_x.$$

Redefine $\mathcal{A}$ to output the above polynomial $Q(x) \in \mathbb{Z}_p[x]$ on input $\{k_i\}$. Now accounting for all randomness during the course of the game, we have that

$$\Pr_{\substack{\sigma\leftarrow\Sigma, x\leftarrow S(\sigma), j\leftarrow\mathcal{J}_x,\\ \sigma'\leftarrow\mathcal{B}_x^{(j)}, \{k_i\}\leftarrow\mathbb{Z}_p^{n-1}, \mathcal{A}}} [Q(x) \in \mathcal{P}'^{(j)}_x : Q \leftarrow \mathcal{A}^{\sigma'}(\{k_i\})] = \frac{1}{2qT^2} - \mathsf{negl}(\lambda).$$

Now we switch the distribution on σ' from $Z = \{Z_x\}_x$ to $Y = \{Y_x\}_x$. We can still represent Y_x in the same way as Z_x except the $\mathcal{B}_x^{(j)}$'s are replaced by $(P, 1-\delta)$-dense sources $\mathcal{D}_x^{(j)}$. Referring to the Lemma 9 statement, we have that

$$\delta \leq \frac{2\lambda + \log 1/\gamma}{P \log(p/e)} \leq \frac{1}{10T^4 q \log(p/e)}.$$

Now assume towards contradiction that this switch in distribution causes the adversary's success to become at most $1/(4qT^2)$. Then there must exist some fixed choice of σ, x and j such that $\mathcal{A}$'s difference in success over σ' and its input is at least $1/(4qT^2) - \mathsf{negl}(\lambda)$. So we have

$$\Pr_{\sigma'\leftarrow\mathcal{B}_x^{(j)}, \{k_i\}\leftarrow\mathbb{Z}_p^{n-1}, \mathcal{A}} [Q(x) \in \mathcal{P}'^{(j)}_x : Q \leftarrow \mathcal{A}^{\sigma'}(\{k_i\})] -$$
$$\Pr_{\sigma'\leftarrow\mathcal{D}_x^{(j)}, \{k_i\}\leftarrow\mathbb{Z}_p^{n-1}, \mathcal{A}} [Q(x) \in \mathcal{P}'^{(j)}_x : Q \leftarrow \mathcal{A}^{\sigma'}(\{k_i\})] \geq \frac{1}{4qT^2} - \mathsf{negl}(\lambda).$$

But now we can define a distinguisher that contradicts Lemma 10. The distinguisher knows the fixed x and the set of fixed points $\mathcal{P}_x^{(j)}$, and interacts with either $\mathcal{B}_x^{(j)}$ or $\mathcal{D}_x^{(j)}$, simulating $\mathcal{A}$ making T^2 queries. It can tell whether $\mathcal{A}$ succeeds by plugging x into the polynomial produced and comparing the result to the set of fixed points and the set of queries made by $\mathcal{A}$. Yet it can only distinguish with probability at most $T^2\delta \log p \leq 1/(5qT^2)$ which is a contradiction.

Now we imagine picking c uniformly at random from $\mathcal{P}'^{(j)}_x$. Since $1/(4qT^2) = 1/\mathsf{poly}(\lambda)$ and $|\mathcal{P}'^{(j)}_x| \leq T^2 + P + 1 = \mathsf{poly}(\lambda)$, we can say that

$$\Pr_{\substack{\sigma\leftarrow\Sigma, x\leftarrow S(\sigma), j\leftarrow\mathcal{J}_x, \sigma'\leftarrow\mathcal{D}_x^{(j)},\\ c\leftarrow\mathcal{P}'^{(j)}_x, \{k_i\}\leftarrow\mathbb{Z}_p^{n-1}, \mathcal{A}}} [Q(x) = c : Q \leftarrow \mathcal{A}^{\sigma'}(\{k_i\})] = \frac{1}{\mathsf{poly}(\lambda)}.$$

Now there must exist a $1/\mathsf{poly}(\lambda)$ fraction of $\{k_i\}$ such that the above holds with probability $1/\mathsf{poly}(\lambda)$ on each of those inputs. Denote this set $\mathcal{K}$, where $\mathcal{K}_i$ denotes the ith element of the set. We also now give σ' as an input to $\mathcal{A}$ rather than just giving it oracle access. So we have

$$\Pr_{\substack{\sigma\leftarrow\Sigma,x\leftarrow S(\sigma),j\leftarrow\mathcal{J}_x,\\ \sigma'\leftarrow\mathcal{D}_x^{(j)},c\leftarrow\mathcal{P}'^{(j)}_x,\mathcal{A}}}[Q(x)=c:Q\leftarrow\mathcal{A}(\sigma',\mathcal{K}_i)]=\frac{1}{\mathsf{poly}(\lambda)}\ \forall i\in[|\mathcal{K}|].$$

Then by Lemma 12, noting that $|\mathcal{K}|=\omega(\mathsf{poly}(\lambda))$,

$$\Pr_{\substack{\sigma\leftarrow\Sigma,x\leftarrow S(\sigma),j\leftarrow\mathcal{J}_x,\sigma'\leftarrow\mathcal{D}_x^{(j)},\\ c\leftarrow\mathcal{P}'^{(j)}_x,\mathcal{A},i_1,i_2\leftarrow[|\mathcal{K}|]}}\left[Q_1(x)=c=Q_2(x):\begin{array}{l}Q_1\leftarrow\mathcal{A}(\sigma',\mathcal{K}_{i_1})\\ Q_2\leftarrow\mathcal{A}(\sigma',\mathcal{K}_{i_2})\end{array}\right]=\frac{1}{\mathsf{poly}(\lambda)}.$$

Thus we can get rid of c, and are guaranteed that

$$\Pr_{\substack{\sigma\leftarrow\Sigma,x\leftarrow S(\sigma),j\leftarrow\mathcal{J}_x\\ \sigma'\leftarrow\mathcal{D}_x^{(j)},\mathcal{A},i_1,i_2\leftarrow[|\mathcal{K}|]}}\left[Q_1(x)-Q_2(x)=0:\begin{array}{l}Q_1\leftarrow\mathcal{A}(\sigma',\mathcal{K}_{i_1})\\ Q_2\leftarrow\mathcal{A}(\sigma',\mathcal{K}_{i_2})\end{array}\right]=\frac{1}{\mathsf{poly}(\lambda)}.$$

Now since $\mathcal{K}$ is a $1/\mathsf{poly}(\lambda)$ fraction of the entire domain of $\{k_i\}$, we can instead pick these sets from the entire domain, and with $1/\mathsf{poly}(\lambda)$ probability they will both lie in $\mathcal{K}$. This gives

$$\Pr_{\substack{\sigma\leftarrow\Sigma,x\leftarrow S(\sigma),j\leftarrow\mathcal{J}_x,\sigma'\leftarrow\mathcal{D}_x^{(j)},\\ \mathcal{A},\{k_i^{(1)}\},\{k_i^{(2)}\}\leftarrow\mathcal{Z}_p^{n-1}}}\left[Q_1(x)-Q_2(x)=0:\begin{array}{l}Q_1\leftarrow\mathcal{A}(\sigma',\{k_i^{(1)}\})\\ Q_2\leftarrow\mathcal{A}(\sigma',\{k_i^{(2)}\})\end{array}\right]=\frac{1}{\mathsf{poly}(\lambda)}.$$

Now we look at the probability that Q_1 and Q_2 are distinct polynomials. For any fixed Q, there are at most a $1/p$ fraction of sets $\{k_i\}$ such that $\mathcal{A}(\{k_i\})$ could possibly output Q. This follows since given some $\{k_i\}$, the coefficients on $x^2,...,x^n$ in Q determine the $\ell_2,...,\ell_n$ in $\mathcal{A}$'s linear combination. Then there remains a $1/p$ chance that the $\{\ell_i\}$ and $\{k_i\}$ dot product to the correct linear coefficient in Q. So for uniformly random choice of the $\{k_i^{(1)}\}$ and $\{k_i^{(2)}\}$ sets, there is a $\mathsf{negl}(\lambda)$ chance that the resulting Q_1 and Q_2 output by $\mathcal{A}$ could possibly be equal.

Let E_1 be the event that $Q_1(x)-Q_2(x)=0$ and E_2 be the event that $Q_1\neq Q_2$. We want to say that $\Pr[E_1\wedge E_2]=1/\mathsf{poly}(\lambda)$. This follows from a simple union bound: $\Pr[E_1\wedge E_2]=1-\Pr[\neg E_1\vee\neg E_2]\geq 1-\Pr[\neg E_1]-\Pr[\neg E_2]=1-(1-1/\mathsf{poly}(\lambda))-\mathsf{negl}(\lambda)=1/\mathsf{poly}(\lambda)$.

So we redefine $\mathcal{A}$ to generate two random sets $\{k_i^{(1)}\}$ and $\{k_i^{(2)}\}$ for itself, determine the polynomials Q_1 and Q_2, solve for the roots of Q_1-Q_2, and output a uniformly random root. Note that the degree of Q_1-Q_2 will be at most $n=\mathsf{poly}(\lambda)$. Thus the following holds:

$$\Pr_{\sigma\leftarrow\Sigma,x\leftarrow S(\sigma),\sigma'\leftarrow Y_x,\mathcal{A}}[x\leftarrow\mathcal{A}(\sigma')]=\frac{1}{\mathsf{poly}(\lambda)}.$$

Now we can switch Y_x to Σ_x, and claim that

$$\Pr_{\sigma\leftarrow\Sigma,x\leftarrow S(\sigma),\sigma'\leftarrow\Sigma_x,\mathcal{A}}[x\leftarrow\mathcal{A}(\sigma')]=\frac{1}{\mathsf{poly}(\lambda)}.$$

If instead $\mathcal{A}$'s success was negligible after this switch, then there exists a fixed x for which the difference in success is $1/\mathsf{poly}(\lambda)$. But Y_x and Σ_x are γ-close with $\gamma=1/2^\lambda=\mathsf{negl}(\lambda)$ so this is impossible. Then, we can write

$$\Pr_{\sigma'\leftarrow\Sigma',\mathcal{A}}[x\leftarrow\mathcal{A}(\sigma')]=\frac{1}{\mathsf{poly}(\lambda)},$$

where Σ' is defined as in Lemma 11. This contradicts Lemma 11. □

7.3 Generic Hardness of DDH-II

Assumption 6. *(f-DDH-II) Let* $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$*, where* $2^{\lambda-1} < p(\lambda) < 2^\lambda$*. Let* $\{\mathcal{D}_\lambda\}_\lambda$ *be a family of well-spread distributions where the domain of* $\mathcal{D}_\lambda$ *is* $\mathbb{Z}_{p(\lambda)}$*. Then for any PPT* $\mathcal{A}$*,*

$$|\Pr[\mathcal{A}(g^x, g^r, g^{xr}) = 1] - \Pr[\mathcal{A}(g^x, g^r, g^s) = 1]| = \mathsf{negl}(\lambda),$$

where $x \leftarrow \mathcal{D}_\lambda$*, and* $r, s \leftarrow \mathbb{Z}_{p(\lambda)}$*.*

Theorem 12. *Assumption 6 holds in the Generic Group Model.*

Note that this trivially implies generic security of r-DDH-II.

Proof. We define the following sequence of hybrid games.

- **Hybrid 0.** The Assumption 6 distinguishing game for generic adversary $\mathcal{A}$.

  Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

  The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma, x)$, picks a random challenge bit $b \leftarrow \{0,1\}$, and samples uniformly random $r, s \leftarrow \mathbb{Z}_p$. It initializes the adversary $\mathcal{A}$ with $(\sigma(1), \sigma(x), \sigma(r), \sigma(bxr + (1-b)s))$. The challenger $\mathcal{C}$ proceeds to implement the generic group oracle for $\mathcal{A}$, after which $\mathcal{A}$ outputs a guess $b' \in \{0,1\}$. $\mathcal{A}$ wins if $b' = b$.

- **Hybrid 1.** In this hybrid, we switch to a bit-fixed labeling σ'

  Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

  Let $Z_{x,im(\sigma)}$ be the family defined as in Lemma 9 (parameterized by some $P \in \mathbb{N}$ and $\gamma := 1/2^\lambda$). Sample $\sigma' \leftarrow Z_{x,im(\sigma)}$, letting $\mathcal{P}$ denote the fixed points between σ and σ'.

  The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma', x)$, picks a random challenge bit $b \leftarrow \{0,1\}$, and samples uniformly random $r, s \leftarrow \mathbb{Z}_p$. It initializes the adversary $\mathcal{A}$ with $(\sigma'(1), \sigma'(x), \sigma'(r), \sigma'(bxr + (1-b)s))$. The challenger $\mathcal{C}$ proceeds to implement the generic group oracle for $\mathcal{A}$, after which $\mathcal{A}$ outputs a guess $b' \in \{0,1\}$. $\mathcal{A}$ wins if $b' = b$.

- **Hybrid 2.** In this hybrid, we switch s, r from uniformly random values to formal variables $\mathbf{s}, \mathbf{r}$.

  Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

  Let $Z_{x,im(\sigma)}$ be the family defined as in Lemma 9 (parameterized by some $P \in \mathbb{N}$ and $\gamma := 1/2^\lambda$). Sample $\sigma' \leftarrow Z_{x,im(\sigma)}$, letting $\mathcal{P}$ denote the fixed points between σ and σ'.

  The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma', \mathcal{P}, x)$, and picks a random challenge bit $b \leftarrow \{0,1\}$. Fix formal variables $\mathbf{r}, \mathbf{s}$. The challenger $\mathcal{C}$ interprets the intial elements given to $\mathcal{A}$ and $\mathcal{A}$'s oracle queries as formal polynomials over $\mathbb{Z}_p[\mathbf{r}, \mathbf{s}]$, and responds differently based on the resulting formal polynomial. If the polynomial is a constant value k with no dependence on $\mathbf{r}$ or $\mathbf{s}$, $\mathcal{C}$ answers with $\sigma'(k)$ as in Hybrid 1. Otherwise, $\mathcal{C}$ samples a uniformly random value from $im(\sigma') \setminus im(\mathcal{P})$ and responds with this (unless the formal polynomial has been queried before, in which case $\mathcal{C}$ responds with the label it returned on the earlier query). $\mathcal{A}$ outputs a guess $b' \in \{0,1\}$, and $\mathcal{A}$ wins if $b' = b$.

- **Hybrid 3.** In this hybrid, we change the game. We don't consider formal variables $\mathbf{r}$, $\mathbf{s}$, or have the challenger pick b. Instead, we only give $\mathcal{A}$ the handle $\sigma'(x)$ and require that it return the value x to be considered successful.

Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

Let $Z_{x,im(\sigma)}$ be the family defined as in Lemma 9 (parameterized by some $P \in \mathbb{N}$ and $\gamma := 1/2^\lambda$). Sample $\sigma' \leftarrow Z_{x,im(\sigma)}$, letting $\mathcal{P}$ denote the fixed points between σ and σ'.

The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma', x)$. It initializes the adversary $\mathcal{A}$ with $(\sigma'(1), \sigma'(x))$ and proceeds to implement the generic group oracle with σ'. $\mathcal{A}$ then outputs a guess x' for x and wins if $x' = x$.

- **Hybrid 4.** In this hybrid, we move back to the original σ.

  Let $\mathsf{GroupGen}(1^\lambda) = (\mathbb{G}_\lambda, g, p(\lambda))$, where $2^{\lambda-1} < p(\lambda) < 2^\lambda$. Let $p := p(\lambda)$. Sample a uniformly random injection $\sigma : [p] \to [p']$ for an arbitrary $p' > p$. Let $S : [p']^{[p]} \to \mathbb{Z}_p$ be a possibly inefficient randomized algorithm such that $H_\infty(S(\sigma)|\sigma) = \omega(\log(\lambda))$. Sample $x \leftarrow S(\sigma)$.

  The challenger $\mathcal{C}$ receives as input $(\mathbb{G}_\lambda, g, p, \sigma, x)$. It initializes the adversary $\mathcal{A}$ with $(\sigma(1), \sigma(x))$ and proceeds to implement the generic group oracle with σ. $\mathcal{A}$ then outputs a guess x' for x and wins if $x' = x$.

Assume towards contradiction the existence of an adversary $\mathcal{A}$ making $T(\lambda) = \mathsf{poly}(\lambda)$ queries which attains non-negligible advantage $\epsilon(\lambda)$ in **Hybrid 0**. Let $q(\lambda) = \mathsf{poly}(\lambda)$ be such that $q(\lambda) > 1/\epsilon(\lambda)$ for infinitely many λ. Let $T := T(\lambda)$ and $q := q(\lambda)$. Set $P = 24\lambda Tq = \mathsf{poly}(\lambda)$.

Claim 2. *$\mathcal{A}$ attains advantage at least $1/2q$ in **Hybrid 1**.*

Consider the following distinguisher $\mathcal{D}(x)$, which interacts with an oracle injection source mapping $[p] \to [p']$, and receives as input $x \leftarrow S(\sigma)$. $\mathcal{D}$ simulates the interaction between $\mathcal{C}$ and $\mathcal{A}$ described in **Hybrid 0** and outputs a bit indicating whether $\mathcal{A}$ was successful or not. If the injection source that $\mathcal{D}$ is interacting with is σ, then the simulation is exactly **Hybrid 0**. If it is $Z_{x,im(\sigma)}$, then the simulation is exactly **Hybrid 1**.

Applying Lemma 9 with the sampler $x \leftarrow S(\sigma)$ as the function f, we have that the success probability of $\mathcal{A}$ in **Hybrid 1** must be at least

$$\epsilon(\lambda) - \frac{2T(\log p + \lambda)}{P} - \frac{1}{2^\lambda} \geq \frac{1}{q} - \frac{4\lambda T}{24\lambda Tq} - \frac{1}{2^\lambda} \geq \frac{1}{2q}.$$

Claim 3. *$\mathcal{A}$ attains advantage at least $1/2q - \mathsf{negl}(\lambda)$ in **Hybrid 2**.*

Let $\{c_1^{(t)}\sigma'(x) + c_2^{(t)}\sigma'(r) + c_3^{(t)}\sigma'(bxr + (1-b)s) + c_4^{(t)}\sigma'(1)\}_{t\in[T]}$ be the set of $\mathcal{A}$'s queries. We consider separately whether $b = 0$ or 1.

If $b = 0$, then in **Hybrid 2** we have that each query is of the form $c_3^{(t)}\mathbf{s} + c_2^{(t)}\mathbf{r} + (c_1^{(t)}x + c_4^{(t)})$. Let $\mathcal{T}'$ be the set of queries for which $c_3^{(t)}, c_2^{(t)} = 0$. We define the set $\mathcal{S}$ to be the set of points that result from queries in $\mathcal{T}'$, combined with the set of fixed points $\mathcal{P}$. Then if all queries outside of $\mathcal{T}'$ evaluate to distinct points not in $\mathcal{S}$, then $\mathcal{A}$ sees exactly the same distribution in **Hybrid 1** as in **Hybrid 2**. This event can be seen as a collection of linear polynomials over $\mathbf{s}$ and $\mathbf{r}$, where at least one evaluates to 0. The number of such polynomials is at most $(T - |\mathcal{T}'|)(P + |\mathcal{T}'|) + (T - |\mathcal{T}'|)^2 \leq TP + 2T^2 = \mathsf{poly}(\lambda)$. But over the uniform randomness of r, s, each individually goes to zero with probability $1/p = \mathsf{negl}(\lambda)$. Thus the probability that $\mathcal{A}$ can distinguish is $\mathsf{negl}(\lambda)$.

Now if $b = 1$, each query in **Hybrid 2** is of the form $(c_3^{(t)}x + c_2^{(t)})\mathbf{r} + (c_1^{(t)}x + c_4^{(t)})$. We have the same argument as in the $b = 0$ case, defining the set $\mathcal{T}'$ to consist of queries for which $c_3^{(t)}x + c_2^{(t)} = 0$. The rest of the argument is identical, and we get that $\mathcal{A}$ can distinguish with $\mathsf{negl}(\lambda)$ probability.

Claim 4. *There exists an $\mathcal{A}'$ that makes at most $2T$ queries and attains advantage at least $1/2q - \mathsf{negl}(\lambda)$ in **Hybrid 3**.*

First we slightly alter **Hybrid 2** to **Hybrid 2a**, where the challenger responds to non-constant polynomials over $\mathbf{r}$ and $\mathbf{s}$ with random values from $im(\sigma')$ rather than from $im(\sigma') \setminus im(\mathcal{P})$. By a union bound, the probability that this changes $\mathcal{A}$'s view is at most $TP/p = \mathsf{negl}(\lambda)$.

Now we determine what events could cause $\mathcal{A}$ to distinguish between $b = 0$ and $b = 1$ in **Hybrid 2a**. First, it could form a query that is constant in one case but non-constant in the other. It is clear that if a query is constant when $b = 0$, then it is constant when $b = 1$. On the other hand, if a query (c_1, c_2, c_3, c_4) is constant when $b = 1$ and non-constant when $b = 0$, we must have that $c_3 x + c_2 = 0$ where either $c_2, c_3 \neq 0$. This implies $c_3 \neq 0$. Second, it could form two non-constant queries that evaluate to the same element in one case but not the other. Again, it is clear that if two queries evaluate to the same element when $b = 0$, then they do so when $b = 1$. On the other hand, if two queries (c_1, c_2, c_3, c_4) and (c'_1, c'_2, c'_3, c'_4) evaluate to the same element when $b = 1$ but not when $b = 0$, we must have that $c_3 x + c_2 = c'_3 x + c'_2$ and $c_1 x + c_4 = c'_1 x + c'_4$ but either $c_3 \neq c'_3$, $c_2 \neq c'_2$, or $c_1 x + c_4 \neq c'_1 x + c'_4$. This implies $c_3 x + c_2 = c'_3 x + c'_2$ where $c_3 \neq c'_3$.

So we set up a reduction $\mathcal{B}$ that interacts with the **Hybrid 3** game. $\mathcal{B}$ will simulate $\mathcal{A}$'s view of **Hybrid 2a** by considering formal variables $\mathbf{s}$ and $\mathbf{r}$, choosing $b \in \{0, 1\}$ and responding to queries that are constant over $\mathbf{s}$ and $\mathbf{r}$ via the **Hybrid 3** challenger, and picking uniformly among $im(\sigma')$ otherwise. Additionally, on each of $\mathcal{A}$'s queries such that $c_3 \neq 0$, it queries its oracle on $c_3\sigma'(x) + c_2$ and stores the responses. If $\mathcal{A}$ succeeds in **Hybrid 2a**, then eventually $\mathcal{B}$ will see the response $\sigma'(0)$ or see same response twice on different queries. Either way, this gives $\mathcal{B}$ a non-zero linear equation over x, which it can solve. We let $\mathcal{A}'$ be the combination of $\mathcal{B}$ and $\mathcal{A}$, noting that $\mathcal{A}'$ makes at most 2 times as many generic group queries as $\mathcal{A}$.

Claim 5. *$\mathcal{A}'$ attains $1/\mathsf{poly}(\lambda)$ advantage in **Hybrid 4**.*

Consider the following distinguisher $\mathcal{D}(x)$, which interacts with an oracle injection source mapping $[p] \to [p']$, and receives as input $x \leftarrow S(\sigma)$. $\mathcal{D}$ simulates the interaction between $\mathcal{C}$ and $\mathcal{A}'$ described in **Hybrid 3** and outputs a bit indicating whether $\mathcal{A}'$ was successful or not (whether it correctly predicted x). If the injection source that D is interacting with is σ', then the simulation is exactly **Hybrid 3**. If the injection source that D is interacting with is $Z_{x,im(\sigma)}$, then the simulation is exactly **Hybrid 4**.

Applying Lemma 9 with the sampler $x \leftarrow S(\sigma)$ as the function f, we have that the success probability of $\mathcal{A}'$ in **Hybrid 4** must be at least

$$\frac{1}{2q} - \mathsf{negl}(\lambda) - \frac{4T(\log p + \lambda)}{P} - \frac{1}{2^\lambda} = \frac{1}{2q} - \frac{8\lambda T}{24\lambda T q} - \mathsf{negl}(\lambda) = \frac{1}{\mathsf{poly}(\lambda)}.$$

Claim 6. *Every adversary $\mathcal{A}'$ with $T = \mathsf{poly}(\lambda)$ queries has $\mathsf{negl}(\lambda)$ advantage in **Hybrid 4**.*

If $\mathcal{A}'(\sigma(1), \sigma(x))$ outputs x with probability $1/\mathsf{poly}(\lambda)$ when x is sampled from any $\mathcal{D}_\lambda$ with $\omega(\log(\lambda))$ min-entropy, there must exist some $\omega(\mathsf{poly}(\lambda))$-size set $S \subset \mathbb{Z}_p$ such that for all $x \in S$, $\mathcal{A}'(\sigma(x))$ outputs x with probability $1/\mathsf{poly}(\lambda)$. Then for a uniformly random $x \leftarrow \mathbb{Z}_p$, it must be the case that $\mathcal{A}'(\sigma(1), \sigma(x))$ outputs x with probability at least $(\omega(\mathsf{poly}(\lambda))/p) \cdot (1/\mathsf{poly}(\lambda)) = \omega(\mathsf{poly}(\lambda))/p$, contradicting the generic discrete log lower bound of $O(T^2/p)$ [Sho97].

□

8 Acknowledgements

We thank Justin Holmgren for collaboration in the early stages of this work and for contributing a number of extremely valuable insights. We also thank Alon Rosen for helpful feedback regarding exposition and presentation.

This material is based upon work supported by the ARO and DARPA under Contract No. W911NF-15-C-0227. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the ARO and DARPA.

References

[AJR08] Kristina Altmann, Tibor Jager, and Andy Rupp. On black-box ring extraction and integer factorization. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP 2008, Part II*, volume 5126 of *LNCS*, pages 437–448. Springer, Heidelberg, July 2008.

[AM09] Divesh Aggarwal and Ueli Maurer. Breaking RSA generically is equivalent to factoring. In Antoine Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 36–53. Springer, Heidelberg, April 2009.

[BC10] Nir Bitansky and Ran Canetti. On strong simulation and composable point obfuscation. In Tal Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 520–537. Springer, Heidelberg, August 2010.

[BC14] Nir Bitansky and Ran Canetti. On strong simulation and composable point obfuscation. *Journal of Cryptology*, 27(2):317–357, April 2014.

[BCLN16] Joppe W. Bos, Craig Costello, Patrick Longa, and Michael Naehrig. Selecting elliptic curves for cryptography: an efficiency and security analysis. *Journal of Cryptographic Engineering*, 6(4):259–286, Nov 2016.

[BF01] Dan Boneh and Matthew K. Franklin. Identity-based encryption from the Weil pairing. In Joe Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 213–229. Springer, Heidelberg, August 2001.

[BFF+14] Gilles Barthe, Edvard Fagerholm, Dario Fiore, John C. Mitchell, Andre Scedrov, and Benedikt Schmidt. Automated analysis of cryptographic assumptions in generic group models. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 95–112. Springer, Heidelberg, August 2014.

[BGI+01] Boaz Barak, Oded Goldreich, Russell Impagliazzo, Steven Rudich, Amit Sahai, Salil P. Vadhan, and Ke Yang. On the (im)possibility of obfuscating programs. In Joe Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 1–18. Springer, Heidelberg, August 2001.

[BL13] Daniel J. Bernstein and Tanja Lange. Non-uniform cracks in the concrete: The power of free precomputation. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part II*, volume 8270 of *LNCS*, pages 321–340. Springer, Heidelberg, December 2013.

[BR93] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In V. Ashby, editor, *ACM CCS 93*, pages 62–73. ACM Press, November 1993.

[Bra94] Stefan Brands. Untraceable off-line cash in wallets with observers (extended abstract). In Douglas R. Stinson, editor, *CRYPTO'93*, volume 773 of *LNCS*, pages 302–318. Springer, Heidelberg, August 1994.

[BS02] Dan Boneh and Alice Silverberg. Applications of multilinear forms to cryptography. *Contemporary Mathematics*, 324:71–90, 2002.

[BZ14] Dan Boneh and Mark Zhandry. Multiparty key exchange, efficient traitor tracing, and more from indistinguishability obfuscation. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 480–499. Springer, Heidelberg, August 2014.

[Can97] Ran Canetti. Towards realizing random oracles: Hash functions that hide all partial information. In Burton S. Kaliski Jr., editor, *CRYPTO'97*, volume 1294 of *LNCS*, pages 455–469. Springer, Heidelberg, August 1997.

[CD08] Ran Canetti and Ronny Ramzi Dakdouk. Extractable perfectly one-way functions. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP 2008, Part II*, volume 5126 of *LNCS*, pages 449–460. Springer, Heidelberg, July 2008.

[CDG18] Sandro Coretti, Yevgeniy Dodis, and Siyao Guo. Non-uniform bounds in the random-permutation, ideal-cipher, and generic-group models. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO 2018, Part I*, volume 10991 of *LNCS*, pages 693–721. Springer, Heidelberg, August 2018.

[CDGS18] Sandro Coretti, Yevgeniy Dodis, Siyao Guo, and John P. Steinberger. Random oracles and non-uniformity. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part I*, volume 10820 of *LNCS*, pages 227–258. Springer, Heidelberg, April / May 2018.

[CK18] Henry Corrigan-Gibbs and Dmitry Kogan. The discrete-logarithm problem with preprocessing. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part II*, volume 10821 of *LNCS*, pages 415–447. Springer, Heidelberg, April / May 2018.

[CS98] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In Hugo Krawczyk, editor, *CRYPTO'98*, volume 1462 of *LNCS*, pages 13–25. Springer, Heidelberg, August 1998.

[CV09] Ran Canetti and Mayank Varia. Non-malleable obfuscation. In Omer Reingold, editor, *TCC 2009*, volume 5444 of *LNCS*, pages 73–90. Springer, Heidelberg, March 2009.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.

[DHZ14] Ivan Damgård, Carmit Hazay, and Angela Zottarel. Short paper on the generic hardness of ddh-ii. 2014.

[DK02] Ivan Damgård and Maciej Koprowski. Generic lower bounds for root extraction and signature schemes in general groups. In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 256–271. Springer, Heidelberg, April / May 2002.

[ElG84] Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In G. R. Blakley and David Chaum, editors, *CRYPTO'84*, volume 196 of *LNCS*, pages 10–18. Springer, Heidelberg, August 1984.

[FF18] Peter Fenteany and Benjamin Fuller. Non-malleable digital lockers for efficiently sampleable distributions. Cryptology ePrint Archive, Report 2018/957, 2018. https://eprint.iacr.org/2018/957.

[Fuj16] Eiichiro Fujisaki. Improving practical uc-secure commitments based on the ddh assumption. In *Proceedings of the 10th International Conference on Security and Cryptography for Networks - Volume 9841*, pages 257–272, Berlin, Heidelberg, 2016. Springer-Verlag.

[Gal12] Steven D Galbraith. *Mathematics of public key cryptography*. Cambridge University Press, 2012.

[GG11] Eran Gat and Shafi Goldwasser. Probabilistic search algorithms with unique answers and their cryptographic applications. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:136, 2011.

[GK16] Shafi Goldwasser and Yael Tauman Kalai. Cryptographic assumptions: A position paper. In Eyal Kushilevitz and Tal Malkin, editors, *TCC 2016-A, Part I*, volume 9562 of *LNCS*, pages 505–522. Springer, Heidelberg, January 2016.

[KKS15] Jinsu Kim, Sungwook Kim, and Jae Hong Seo. Multilinear map via scale-invariant FHE: Enhancing security and efficiency. Cryptology ePrint Archive, Report 2015/992, 2015. https://ia.cr/2015/992.

[KL] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography, Second Edition.*

[KLRZ08] Yael Tauman Kalai, Xin Li, Anup Rao, and David Zuckerman. Network extractor protocols. In *49th FOCS*, pages 654–663. IEEE Computer Society Press, October 2008.

[KY18a] Ilan Komargodski and Eylon Yogev. Another step towards realizing random oracles: Non-malleable point obfuscation. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part I*, volume 10820 of *LNCS*, pages 259–279. Springer, Heidelberg, April / May 2018.

[KY18b] Ilan Komargodski and Eylon Yogev. Another step towards realizing random oracles: Non-malleable point obfuscation. Cryptology ePrint Archive, Report 2018/149, 2018. https://ia.cr/2018/149.

[LCH11] Hyung Tae Lee, Jung Hee Cheon, and Jin Hong. Accelerating ID-based encryption based on trapdoor DL using pre-computation. Cryptology ePrint Archive, Report 2011/187, 2011. https://ia.cr/2011/187.

[Lip94] Richard J Lipton. Straight-line complexity and integer factorization. In *International Algorithmic Number Theory Symposium*, pages 71–79. Springer, 1994.

[Mau05] Ueli M. Maurer. Abstract models of computation in cryptography (invited paper). In Nigel P. Smart, editor, *10th IMA International Conference on Cryptography and Coding*, volume 3796 of *LNCS*, pages 1–12. Springer, Heidelberg, December 2005.

[Mih] J.P. Mihalcik. An analysis of algorithms for solving discrete logarithms in fixed groups.

[Nec94] V. I. Nechaev. Complexity of a determinate algorithm for the discrete logarithm. *Mathematical Notes*, 55(2):165–172, 1994.

[Sho97] Victor Shoup. Lower bounds for discrete logarithms and related problems. In Walter Fumy, editor, *EUROCRYPT'97*, volume 1233 of *LNCS*, pages 256–266. Springer, Heidelberg, May 1997.

[Sho99] Victor Shoup. On formal models for secure key exchange. Technical Report RZ 3120, IBM, 1999.

[SS01] Ahmad-Reza Sadeghi and Michael Steiner. Assumptions related to discrete logarithms: Why subtleties make a real difference. In Birgit Pfitzmann, editor, *EUROCRYPT 2001*, volume 2045 of *LNCS*, pages 244–261. Springer, Heidelberg, May 2001.

[SS06] Amitabh Saxena and Ben Soh. A new cryptosystem based on hidden order groups. Cryptology ePrint Archive, Report 2006/178, 2006. https://ia.cr/2006/178.

[Unr07] Dominique Unruh. Random oracles and auxiliary input. In Alfred Menezes, editor, *CRYPTO 2007*, volume 4622 of *LNCS*, pages 205–223. Springer, Heidelberg, August 2007.

[Wee05] Hoeteck Wee. On obfuscating point functions. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 523–532. ACM Press, May 2005.

[Yun15] Aaram Yun. Generic hardness of the multiple discrete logarithm problem. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part II*, volume 9057 of *LNCS*, pages 817–836. Springer, Heidelberg, April 2015.

[YYHK14] Takashi Yamakawa, Shota Yamada, Goichiro Hanaoka, and Noboru Kunihiro. Self-bilinear map on unknown order groups from indistinguishability obfuscation and its applications. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part II*, volume 8617 of *LNCS*, pages 90–107. Springer, Heidelberg, August 2014.

[YYHK18] Takashi Yamakawa, Shota Yamada, Goichiro Hanaoka, and Noboru Kunihiro. Generic hardness of inversion on ring and its relation to self-bilinear map. Cryptology ePrint Archive, Report 2018/463, 2018. https://ia.cr/2018/463.