

# Hard-Core Predicates for a Diffie-Hellman Problem over Finite Fields

Nelly Fazio<sup>1,2</sup>, Rosario Gennaro<sup>1,2</sup>, Irippuge Milinda Perera<sup>2</sup>,  
and William E. Skeith III<sup>1,2</sup>

<sup>1</sup> The City College of CUNY

{fazio,rosario,wes}@cs.ccny.cuny.edu

<sup>2</sup> The Graduate Center of CUNY

iperera@gc.cuny.edu

**Abstract.** A long-standing open problem in cryptography is proving the existence of (deterministic) hard-core predicates for the Diffie-Hellman problem defined over finite fields. In this paper, we make progress on this problem by defining a very natural variation of the Diffie-Hellman problem over  $\mathbb{F}_{p^2}$  and proving the unpredictability of every single bit of one of the coordinates of the secret DH value.

To achieve our result, we modify an idea presented at CRYPTO'01 by Boneh and Shparlinski [4] originally developed to prove that the LSB of the elliptic curve Diffie-Hellman problem is hard. We extend this idea in two novel ways:

1. We generalize it to the case of finite fields  $\mathbb{F}_{p^2}$ ;
2. We prove that any bit, not just the LSB, is hard using the list decoding techniques of Akavia et al. [1] (FOCS'03) as generalized at CRYPTO'12 by Duc and Jetchev [6].

In the process, we prove several other interesting results:

- Our result also hold for a larger class of predicates, called *segment predicates* in [1];
- We extend the result of Boneh and Shparlinski to prove that every bit (and every segment predicate) of the elliptic curve Diffie-Hellman problem is hard-core;
- We define the notion of *partial one-way function* over finite fields  $\mathbb{F}_{p^2}$  and prove that every bit (and every segment predicate) of one of the input coordinates for these functions is hard-core.

**Keywords:** Hard-Core Bits, Diffie-Hellman Problem, Finite Fields, Elliptic Curves.

## 1 Introduction

A long-standing open problem in cryptography is proving the existence of (deterministic) hard-core predicates for the Diffie-Hellman problem defined over finite fields. In this paper we make progress on this problem by defining a very natural extension of the Diffie-Hellman problem over  $\mathbb{F}_{p^2}$  and proving that a

large class of predicates (including every single bit of one of the coordinates) are unpredictable under the assumption that this problem is hard.

In their seminal paper that introduced public-key cryptography [5] Diffie and Hellman defined the following key exchange protocol, which works in arbitrary finite cyclic groups. Let  $G$  be such a group, generated by  $g$  of order  $n$ . Two parties, Alice and Bob, want to establish a secret value. Alice chooses a random value  $a \in \mathbb{Z}_n$  and sends the value  $A = g^a$  to Bob. Similarly Bob chooses a random value  $b \in \mathbb{Z}_n$  and sends the value  $B = g^b$  to Alice. At this point they share the common Diffie-Hellman secret value  $K = g^{ab} = A^b = B^a$ .

The *Computational Diffie-Hellman Assumption* (CDH) over the group  $G$  informally states that no efficient algorithm can compute  $K = g^{ab}$  when given only  $g, A = g^a, B = g^b$ . The hardness of computing the entire value  $K$ , however does not rule out an efficient way to compute some of the bits of  $K$ , or even just predict them with a probability better than a random guess. This property is very important because without it, Alice and Bob do not have any guarantee about the “pseudorandomness” of any bit of the secret value  $K$ , and those are the properties needed by  $K$  in order to be used as a secret key in a subsequent cryptographic scheme. This problem is usually addressed by making a much stronger assumption on the hardness of the Diffie-Hellman problem: the so-called *Decisional Diffie-Hellman Assumption* (DDH) states that the value  $K$  is computationally indistinguishable from a random element of  $G$ . While the DDH guarantees that the entire value of  $K$  is pseudorandom, there are groups  $G$  where the DDH is false, even when the CDH is still conjectured to be hard.

Ideally, however, one would like to prove that certain bits (or more generally, certain predicates) of the value  $K$  are unpredictable, when given  $g^a$  and  $g^b$ , simply under the CDH assumption. Such results were established quite early for other conjectured hard-problems (e.g., Blum and Micali’s result on the hardness of discrete log bits [3] and Alexi et al. work on the hardness of the RSA input bits [2]). However for the case of the Diffie-Hellman problem no such result has been proven (except for the result by Boneh and Shparlinksi [4] in a slightly different model and which we discuss below). The only hard-core predicates known for the Diffie-Hellman function are the generic “randomized” predicates which work over any computationally hard problem (e.g., the Goldreich-Levin and Näslund hard-core bits [8, 10]).

**HARD-CORE PREDICATES.** Let  $\pi : G \rightarrow \{\pm 1\}$  be a predicate<sup>1</sup> defined over  $G$ . To prove that  $\pi$  is hard-core for the CDH problem one has to construct a reduction from guessing  $\pi$  better than at random, to solving the CDH problem. More specifically, assume we have an oracle  $\Omega$  which on input  $g, g^a, g^b$  outputs the correct  $\pi(g^{ab})$  with probability (taken over the choice of  $a, b$ ) substantially better than<sup>2</sup>  $1/2$ , then there is an efficient algorithm  $A$  which invokes  $\Omega$  and solves the CDH problem.

<sup>1</sup> For reasons that will become clearer in the technical section of the paper, we adopt the convention that predicates map a value to  $\pm 1$  instead of  $\{0, 1\}$ .

<sup>2</sup> Let’s assume for now that  $\pi$  is balanced. In the rest of the paper we take into account the possible bias of  $\pi$ .

Note that a crucial step of this reduction is to “correct” the answers of the oracle  $\Omega$  which are guaranteed to be right only slightly more than half of the times. This step requires randomizing the queries to  $\Omega$  while still keeping its answers useful to the solution of the underlying CDH problem. This proves somewhat difficult, due to the limited random self-reducibility of the Diffie-Hellman problem.

**RANDOMIZING THE PROBLEM REPRESENTATION.** Boneh and Shparlinksi in [4] achieved a breakthrough for the elliptic curve Diffie-Hellman problem, i.e., the CDH problem defined over the group  $G$  of points of an elliptic curve. They were able to prove that the least significant bit of each coordinate of the Diffie-Hellman secret value  $K$  is hard-core, when the probability space of the oracle  $\Omega$  also includes a random choice for the representation of the curve.

More specifically: let  $p$  be a prime and let  $E$  be an elliptic curve defined over  $\mathbb{F}_p$ , the finite field with  $p$  elements. To represent  $E$  we use a short Weierstrass equation  $W : y^2 = x^3 + ax + b$ , with  $a, b \in \mathbb{F}_p$  and  $4a^3 + 27b^2 \neq 0$ . Let  $W(E)$  be the set of Weierstrass equations representing  $E$ . It is well known that  $W(E)$  is defined by the equations  $W_\lambda$  of the form  $y^2 = x^3 + \lambda^4 ax + \lambda^6 b$  for  $\lambda \in \mathbb{F}_p^\times$ . If  $Q = (Q_x, Q_y)$  is a point satisfying  $W$  then the point  $Q_\lambda = (Q_{\lambda,x} = \lambda^2 Q_x, Q_{\lambda,y} = \lambda^3 Q_y)$  satisfies  $W_\lambda$ . Furthermore, the points of  $E$  form a group under a certain operation, and the mapping  $\Phi_\lambda : E \rightarrow E$  defined as  $\Phi_\lambda(Q) = Q_\lambda$  is an isomorphism with respect to such group operation over  $E$ .

Let  $G$  be a cyclic subgroup of  $E$  generated by a point  $P$ . Switching to additive notation for the group operation, the elliptic curve CDH (EC-CDH) assumption says that given  $W, P, aP, bP$  it is hard to compute  $abP$ .

In [4] they prove that if there exists an oracle  $\Omega$  that works on a random representation of  $E$ , i.e., such that

$$\Pr_{\lambda,a,b} [\Omega(\lambda, P, aP, bP) = \text{LSB}([\Phi_\lambda(abP)]_x)] > 1/2 + \epsilon$$

for a non-negligible value  $\epsilon$ , then it is possible to solve EC-CDH on any curve (a similar result holds for the  $y$ -coordinate of  $abP$ ).

## 1.1 Our Results

Our main technical contribution is to show that the Boneh-Shparlinksi idea of randomizing the representation of the underlying group for the CDH problem can be also applied to the case of finite fields  $\mathbb{F}_{p^2}$ .

For a given prime  $p$ , there are many different fields  $\mathbb{F}_{p^2}$ , but they are all isomorphic to each other. Let  $h(x) = x^2 + h_1x + h_0$  be a monic irreducible polynomial of degree 2 in  $\mathbb{F}_p$ . It is well known that  $\mathbb{F}_{p^2}$  is isomorphic to the field  $\mathbb{F}_p[x]/(h)$ , and therefore elements of  $\mathbb{F}_{p^2}$  can be written as linear polynomials: if  $g \in \mathbb{F}_{p^2}$  then  $g = g_0 + g_1x$  and addition and multiplication are performed as polynomial operations modulo  $h$ . In the following, given  $g \in \mathbb{F}_{p^2}$  we denote with  $[g]_i$  the coefficient of the degree- $i$  term.

Let  $I_2(p)$  be the set of monic irreducible polynomials of degree 2 in  $\mathbb{F}_p$ . For  $h, \hat{h} \in I_2(p)$  we know that there exists an (easily computable) isomorphism

$$\phi_{h,\hat{h}} : \mathbb{F}_p[x]/(h) \rightarrow \mathbb{F}_p[x]/(\hat{h}).$$

Finally, denote with  $g$  a generator of the multiplicative group of  $\mathbb{F}_{p^2}$  which is known to be cyclic.

Our first attempt was to use the approach from [4] over  $\mathbb{F}_{p^2}$ . That is, we hoped to prove that given an oracle  $\Omega$  which, on input random values  $g^a, g^b$  and a random description of  $\mathbb{F}_{p^2}$ , outputs  $\text{LSB}([g^{ab}]_i)$ , then we can solve the CDH over  $\mathbb{F}_{p^2}$ . Unfortunately there are several technical complications with directly applying the approach of [4] to the finite field case, one of them being the fact that representations of an elliptic curve are in bijective correspondence with  $\mathbb{F}_p$  allowing them to be represented by a single element of  $\mathbb{F}_p$ . Conversely the representations of  $\mathbb{F}_{p^2}$  are in bijective correspondence with  $I_2(p)$  which has  $\approx p^2/2$  elements.

**A NEW DIFFIE-HELLMAN PROBLEM.** To solve these technical problems we had to define the following variant of the CDH problem over  $\mathbb{F}_{p^2}$ : informally we say that the *Partial-CDH* problem is hard in  $\mathbb{F}_{p^2}$  if no efficient algorithm given  $g, A = g^a, B = g^b \in \mathbb{F}_{p^2}$  can compute  $K = [g^{ab}]_1 \in \mathbb{F}_p$  (i.e., the coefficient of the degree 1 term of  $g^{ab}$ ).

We note that the Partial-CDH problem is obviously weaker than the regular CDH problem over  $\mathbb{F}_{p^2}$ , but that it still allows Alice and Bob to agree on a common secret value in  $\mathbb{F}_p$ , via the traditional Diffie-Hellman protocol.

**OUR MAIN RESULT.** Assuming the hardness of the Partial-CDH problem we prove that for a large class of predicates  $\pi$  (described below – it includes every individual bit of  $K$ ), the bit  $\pi(K)$  is unpredictable given  $g^a, g^b$  and a random representation of  $\mathbb{F}_{p^2}$ . More specifically, we prove that if there exists an oracle  $\Omega$  such that for any  $h \in I_2(p)$  it holds that

$$\Pr_{\hat{h}, a, b} \left[ \Omega \left( h, \hat{h}, g, g^a, g^b \right) = \pi \left( \left[ \phi_{h, \hat{h}}(g^{ab}) \right]_1 \right) \right] > 1/2 + \epsilon$$

for a non-negligible value  $\epsilon$ , then it is possible to solve Partial-CDH on  $\mathbb{F}_p[x]/(h)$ .

We may define an analogous problem for the general case of  $\mathbb{F}_{p^t}$  with any  $t > 1$ . The Partial-CDH problem is defined as outputting the coefficient of the term of degree  $t - 1$ . However our hard-core results hold only for the quadratic ( $\mathbb{F}_{p^2}$ ) case. See the conclusion (Section 6) for a discussion.

**OUR TECHNIQUES.** To achieve our result we divert from the techniques used in [4] in another fundamental way. To prove that the predicate  $\pi$  is hard-core for the Partial-CDH problem in  $\mathbb{F}_{p^2}$  we use the list-decoding approach pioneered by Akavia et al. [1] as extended by Duc and Jetchev in [6] to the case of prediction oracles which also take as input a random representation of the underlying group.

We describe the approach in detail in Section 3. For now we just remind the reader that as defined originally in [1] this approach allows one to prove the security of so-called *segment predicates* which include both the most and least significant bits of the input. In [9] the technique was extended to work for any input bit. So the class of predicates  $P$  described above includes every individual bit of the input and also segment predicates as defined in [1].

ADDITIONAL RESULTS. Since the list-decoding approach works for a larger class of predicates, we obtain two additional results:

1. In the elliptic curve scenario, we are able to extend the [4] result for EC-CDH to any predicate  $\pi$  as above, not just the LSB.
2. For the finite field case we prove that the predicates  $\pi$  are hard-core for a much larger class of conjectured computationally hard problems. Consider a function  $f : \mathbb{F}_{p^2} \rightarrow S$  for an arbitrary set  $S$ . We say that  $f$  is a *finite field-based partial one-way function* (FFB-POWF) if the following conditions hold:
  - $f$  is “independent” of the representation used for  $\mathbb{F}_{p^2}$  (see Section 5.2 for a precise definition);
  - no efficient algorithm, given  $f(x)$  can compute  $[x]_1$ , i.e., the coefficient of the degree 1 term of  $x$ .

Then we can prove that if  $f$  is a FFB-POWF then it is hard to predict  $\pi([x]_1)$  better than at random (over a random representation of  $\mathbb{F}_{p^2}$ ) when given only  $f(x)$ .

INTERPRETATION OF OUR RESULTS. One way to interpret our results is to think of the group representation as part of the input to the computational hard problem (be it a one-way function, or the CDH problem) being used. This means that our results do not apply to the case when the Diffie-Hellman key exchange protocol is performed over a fixed representation of the finite field (or the elliptic curve). Rather it is necessary for Alice and Bob to choose a random representation (an irreducible polynomial for  $\mathbb{F}_{p^2}$  or a Weierstrass equation for the curve  $E$ ) over which to run the protocol.

## 1.2 Paper Organization

Section 2 reviews some relevant background, particularly the notion of Fourier transform for codes. In Section 3, we cover the list-decoding approach to prove hard-core predicates [1] and its generalization to the case of elliptic curves from [6]. Sections 4 and 5 present our original results. First, as a warm-up we prove that every bit of the EC-CDH problem is hard-core. Then we present our main result on the bit security of Partial-CDH over finite fields, and its extension to FFB-POWF. Finally, we conclude in Section 6 with some discussion about our results and a list of interesting problems left open by our work.

## 2 Background

### 2.1 Fourier Transforms

Let  $\mathbb{Z}_n$  denote the additive group of integers modulo  $n$ . For any two functions  $f, g : \mathbb{Z}_n \rightarrow \mathbb{C}$ , their *inner product* is defined as  $\langle f, g \rangle = \frac{1}{n} \sum_{x \in \mathbb{Z}_n} f(x) \overline{g(x)}$ . Let  $\mathbb{C}(\mathbb{Z}_n)$  denote the vector space formed by all functions  $f : \mathbb{Z}_n \rightarrow \mathbb{C}$ . The  $\ell_2$ -norm of  $f$  on  $\mathbb{C}(\mathbb{Z}_n)$  is defined as  $\|f\|_2 = \sqrt{\langle f, f \rangle}$ . A *character* of  $\mathbb{Z}_n$  is a

homomorphism  $\chi : \mathbb{Z}_n \rightarrow \mathbb{C}^\times$ , such that  $\forall_{x,y \in \mathbb{Z}_n} \chi(x+y) = \chi(x)\chi(y)$ . These characters are defined by  $\chi_\alpha(x) = \omega_n^{\alpha x}$ , where  $\alpha \in \mathbb{Z}_n$  and  $\omega_n = e^{2\pi i/n}$ . The set of all characters form a group  $\widehat{\mathbb{Z}}_n$ . Since the members of  $\widehat{\mathbb{Z}}_n$  are orthogonal and  $|\widehat{\mathbb{Z}}_n| = |\mathbb{Z}_n|$ , they form an orthogonal basis, termed the *Fourier basis*, for  $\mathbb{C}(\mathbb{Z}_n)$ . The *Fourier transform*  $\widehat{f} : \widehat{\mathbb{Z}}_n \rightarrow \mathbb{C}$  of  $f$  is defined as  $\widehat{f}(\chi) = \langle f, \chi \rangle$ . The *Fourier expansion* of  $f$  is written as  $\sum_{\chi \in \widehat{\mathbb{Z}}_n} \widehat{f}(\chi)\chi$ . For  $\Gamma \subset \widehat{\mathbb{Z}}_n$  the restriction of  $f$  to  $\Gamma$  is the function  $f|_\Gamma : \mathbb{Z}_n \rightarrow \mathbb{C}$  defined by  $f|_\Gamma = \sum_{\chi \in \Gamma} \widehat{f}(\chi)\chi$ . The *Fourier coefficients* of  $f$  are the coefficients  $\widehat{f}(\chi)$  in the Fourier basis  $\widehat{\mathbb{Z}}_n$ . The *weight* of a Fourier coefficient is denoted by  $|\widehat{f}(\chi)|^2$ . Definition 2.1 formalizes the notion of *heavy characters* with respect to  $f$ .

**Definition 2.1 ( $\tau$ -heavy Characters).** *Let  $\tau \in \mathbb{R}^+$  be a threshold and  $f : \mathbb{Z}_n \rightarrow \mathbb{C}$  be an arbitrary function. We say a character  $\chi \in \widehat{\mathbb{Z}}_n$  is  $\tau$ -heavy if the weight of its corresponding Fourier coefficient is at least  $\tau$ . The set of all such character is denoted by  $\text{Heavy}_\tau(f)$ , i.e.,*

$$\text{Heavy}_\tau(f) = \{\chi \in \widehat{\mathbb{Z}}_n : |\widehat{f}(\chi)|^2 \geq \tau\}.$$

## 2.2 Codes and Their Properties

In what follows, we report a few useful known definitions [6] and lemmata [1] about codes over  $\mathbb{Z}_n$ . As in [6], we will regard  $\mathbb{Z}_n$ -codes as associating an element  $x \in \mathbb{Z}_n$  to a  $\mathbb{Z}_n$ -codeword  $C_x$ , which we will in turn see interchangeably as a function  $C_x : \mathbb{Z}_n \rightarrow \{\pm 1\}$  or as a length- $n$  sequence of  $\{\pm 1\}$ .

**Definition 2.2 ( $\epsilon$ -concentrated Function).** *We say a function  $f : \mathbb{Z}_n \rightarrow \{\pm 1\}$  is Fourier  $\epsilon$ -concentrated if there exist a size  $\text{poly}(n, 1/\epsilon)$ ,  $\epsilon > 0$ , set of characters  $\Gamma \subset \widehat{\mathbb{Z}}_n$  such that  $\|f - f|_\Gamma\|_2 \leq \epsilon$ . We say a function is Fourier concentrated if it is  $\epsilon$ -concentrated for every  $\epsilon > 0$ .*

**Definition 2.3 ( $\epsilon$ -concentrated Code).** *We say a code  $\mathcal{C} = \{C_x : \mathbb{Z}_n \rightarrow \{\pm 1\}\}$  is  $\epsilon$ -concentrated if all its codewords  $C_x$  are Fourier  $\epsilon$ -concentrated. We say a code is Fourier concentrated if it is  $\epsilon$ -concentrated for every  $\epsilon > 0$ .*

**Definition 2.4 (Code Recoverability).** *We say a code  $\mathcal{C} = \{C_x : \mathbb{Z}_n \rightarrow \{\pm 1\}\}$  is recoverable if there exists an algorithm that, given as input a threshold  $\tau$  and a character  $\chi \in \widehat{\mathbb{Z}}_n$ , produces a list of all elements  $x$  associated with codewords  $C_x$  for which  $\chi$  is a  $\tau$ -heavy coefficient, that is,  $\{x \in \mathbb{Z}_n : \chi \in \text{Heavy}_\tau(C_x)\}$ , in time polynomial in  $\log n$  and  $1/\tau$ .*

The following two results appear in [1]. Lemma 2.5 shows that, in a concentrated code  $\mathcal{C}$ , any noisy version  $\tilde{C}_x$  of codeword  $C_x$  share at least one heavy coefficient with  $C_x$ . Theorem 2.6 shows that one can efficiently learn all the heavy characters of any function when given query access to it. Therefore having query access to  $\tilde{C}_x$  (which in our case is obtained by querying the prediction oracle  $\Omega$ ), one can learn at least one heavy coefficient of  $C_x$ , and that if the code is also recoverable, then one can recover  $x$ .

**Lemma 2.5 (Lem. 1 of [1]).** *Let  $f, g : \mathbb{Z}_n \rightarrow \{\pm 1\}$  such that  $f$  is Fourier concentrated and, for some  $\epsilon > 0$ ,*

$$\Pr_{x \in \mathbb{Z}_n} f(x) = g(x) \geq \text{maj}_f + \epsilon,$$

where  $\text{maj}_f$  denotes the bias of the function  $f$ , i.e.,  $\text{maj}_f = \max_{b \in \{\pm 1\}} \Pr_{x \in \mathbb{Z}_n} f(x) = b$ . Then there exist a threshold  $\tau$  such that  $1/\tau$  is polynomial in  $\epsilon$  and  $\log n$ , and there exists a character  $\chi \neq 0$  heavy for  $f$  and  $g$ :  $\chi \in \text{Heavy}_\tau(f) \wedge \text{Heavy}_\tau(g)$ .

**Theorem 2.6 (Thm. 6 of [1]).** *There exists a randomized learning algorithm over  $\mathbb{Z}_n$  that, given query access to a function  $w : \mathbb{Z}_n \rightarrow \{\pm 1\}$ ,  $\tau > 0$  and  $0 < \delta < 1$ , returns a list of  $O(1/\tau)$  characters containing  $\text{Heavy}_\tau(w)$  with probability at least  $1 - \delta$ . The probability is taken over the random coins of the algorithm, whose running time is*

$$\tilde{O}\left(\log(n) \ln^2 \frac{(1/\delta)}{\tau^{5.5}}\right).$$

An overview of the above learning algorithm [1] is provided in Appendix A of the full version [7].

### 3 Hard-Core Predicates by List Decoding

In this section, we review the work of Akavia et al. [1] on how to prove that certain predicates are hard-core for a one-way function  $f$  using list decoding of a particular error-correcting code. We also summarize the extensions by Duc and Jetchev [6] to the case of elliptic-curve based one-way functions.

Let  $f : \mathbb{Z}_n \rightarrow S$  be a one-way function and let  $y = f(x)$  for  $x \in \mathbb{Z}_n$ . Let also  $\pi : \mathbb{Z}_n \rightarrow \{\pm 1\}$  denote a predicate (with the convention that a 0 bit is encoded as +1). Finally we denote with  $\beta_\pi$  the bias of the predicate  $\pi$ , i.e.,  $\beta_\pi = \max_{b \in \{\pm 1\}} \Pr_x[\pi(x) = b]$ .

The goal is to prove that  $\pi$  is a hard-core predicate for the function  $f$ . The proof goes as usual by contradiction by assuming that there exists an oracle  $\Omega$  which, when queried on  $f(x)$ , returns a bit  $b$  which is equal to  $\pi(x)$  with probability  $\beta_\pi + \epsilon$  for a non-negligible  $\epsilon$ , and then using  $\Omega$  to invert  $f$ , i.e., find  $x$  given  $y$ .

To achieve this goal, Akavia et al. in [1] define a *multiplication code*

$$\mathcal{C} = \{C_x : \mathbb{Z}_n \rightarrow \{\pm 1\}\}_{x \in \mathbb{Z}_n}, \text{ where } C_x(\lambda) = \pi(\lambda x).$$

In order for their proof to work this code needs the following properties:

**Accessibility:** Given  $y = f(x)$ , it must be possible to obtain a “noisy” version  $\tilde{C}_x$  of the codeword  $C_x$ , i.e., one that agrees with the correct one with probability  $\beta_\pi + \epsilon$  for a non-negligible  $\epsilon$ . In [1], this is done by assuming that the one-way function has some homomorphic property, i.e given  $y = f(x)$  and  $\lambda \in \mathbb{Z}_n$  it is possible to compute  $y_\lambda = f(\lambda x)$  (modular exponentiation has this property). Then, by querying  $\Omega$  on  $y_\lambda$  one gets the desired accessibility property;

**Concentration:** Every codeword  $C_x$  must be a Fourier concentrated function. Remember that according to the definition above this means that for every  $\epsilon$  there exists a polynomial (in  $\log n$  and  $\epsilon^{-1}$ ) set  $\Gamma$  of Fourier characters, such that  $\|C_x - C_{x,\Gamma}\| \leq \epsilon$  (where  $C_{x,\Gamma}$  is the restriction of  $C_x$  to the Fourier characters in  $\Gamma$ );

**Recoverability:** There exists an algorithm that on input a Fourier character  $\chi$  and a threshold  $\tau$ , outputs a list  $L_\chi$  containing all the values  $x \in \mathbb{Z}_n$  such that  $\chi$  is  $\tau$ -heavy for  $C_x$ . The algorithm runs in polynomial (in  $\log n$  and  $\tau^{-1}$ ) time, which in particular means that the size of  $L_\chi$  is also “small”.

Concentration and recoverability depends on the choice of the predicate  $\pi$ . In [1], the notion of *segment predicates* is defined and shown to be sufficient for the purpose. Later Morillo and Rafols in [9] prove that any individual input bit yields a concentrated and recoverable code (we review this in Appendix B of the full version [7]). We assume  $\pi$  to be one of such predicates in the following.

If the code  $\mathcal{C}$  has the above properties then it is possible to prove that  $\pi$  is a hard-core predicate. Assume we have an oracle  $\Omega$  which when queried on  $f(x)$  returns a bit  $b$  which is equal to  $\pi(x)$  with probability  $\beta_\pi + \epsilon$  where  $\epsilon = 1/\text{poly}(\ell)$  (where  $\ell = |n|$ ). We need to show how to use  $\Omega$  to invert  $f$ .

The inversion works as follows. On input  $y = f(x)$ , the oracle  $\Omega$  allows us to access a “noisy” version  $\tilde{C}_x$  of  $C_x$ , i.e., such that  $\Pr_\lambda[C_x(\lambda) = \tilde{C}_x(\lambda)] > \beta_\pi + \epsilon$ . By applying Lemma 2.5 we know that there exists a threshold  $\tau$  which is polynomial in  $\epsilon$  and at least one Fourier character  $\chi$  which is  $\tau$ -heavy for both  $C_x$  and  $\tilde{C}_x$ . Using the learning algorithm described in Theorem 2.6, we obtain a list containing all the  $\tau$ -heavy Fourier characters for  $\tilde{C}_x$ ; for each such character we use the recovery property to create a polynomial size list of possible pre-images for  $y$  which because of Lemma 2.5 must necessarily include  $x$ . The correct  $x$  can be identified by evaluating the OWF  $f$  over all the possible candidates and comparing with  $y$ . Details can be found in [1] (in any case, in Sections 4 and 5 we present the details of this algorithm as it applies to our results).

### 3.1 Accessibility via Elliptic Curve Isomorphisms

Taking the result of [1] as a starting point, and using techniques first developed in [4], Duc and Jetchev [6] show how to obtain the accessibility property in a different way, when the one-way function is defined over the group  $G$  of points of an elliptic curve. Their result does not require the one-way function  $f$  to have some homomorphic property; on the other hand it requires the oracle to work over a *random* description of the curve.

Let  $p$  be a prime and let  $E$  be an elliptic curve defined over  $\mathbb{F}_p$ . To represent  $E$  we use a short Weirstrass equation  $W : y^2 = x^3 + ax + b$ , with  $a, b \in (\mathbb{F}_p)$  and  $4a^3 + 27b^2 \neq 0$ . Let  $W(E)$  be the set of Weirstrass equations representing  $E$ : so  $W \in W(E)$ . It is well known that  $W(E)$  is defined by the equations  $W_\lambda$  of the form  $y^2 = x^3 + \lambda^4 ax + \lambda^6 b$  for  $\lambda \in \mathbb{F}_p^\times$ . If  $Q = (Q_x, Q_y)$  is a point satisfying  $W$  then the point  $Q_\lambda = (Q_{\lambda,x} = \lambda^2 Q_x, Q_{\lambda,y} = \lambda^3 Q_y)$  satisfies  $W_\lambda$ . It is not hard to see that the mapping  $\Phi_\lambda : E \rightarrow E$  defined as  $\Phi_\lambda(Q) = Q_\lambda$  is an isomorphism with respect to the group operation over  $E$ .



Boneh and Shparlinski were the first to note that this isomorphism gives rise to a natural extension of the prediction oracle  $\Omega$ , by requiring that the input distribution for  $\Omega$  also include  $\lambda$ . Following this idea, the oracle in [6] takes as input  $f(Q)$  where  $f$  is a one-way function defined over the group  $E$ , and *also* a value  $\lambda$  (i.e., a representation  $W_\lambda$  of  $E$ ). The oracle returns a bit  $b$  such that  $b = \pi(Q_{\lambda,x})$  with probability  $\beta_\pi + \epsilon$  (for a non-negligible  $\epsilon$ ) where the probability is not only over the choice of  $Q$  (and the internal random coins of  $\Omega$ ) but *also* over the choice of  $\lambda \in \mathbb{F}_p^\times$ .

As defined, the prediction oracle  $\Omega$  gives noisy access to the *quadratic* code-word  $C_Q(\lambda) = \pi(\lambda^2 Q_x)$ , which would complicate matters (in particular it makes it hard to prove concentration and recovery, see [6] for a discussion). To apply the techniques of [1], we need noisy access to the multiplication code  $C_Q : \mathbb{F}_p \rightarrow \{\pm 1\}$  defined as  $C_Q(\lambda) = \pi(\lambda Q_x)$ .

Following [4] again, Duc and Jethchev defined a modified oracle  $\Omega'$  which queries  $\Omega$  if  $\lambda$  is a square in  $\mathbb{F}_p^\times$ , otherwise tosses a  $\beta_\pi$ -biased coin. It is not hard to see that if  $\Omega$  had advantage  $\epsilon$ , then  $\Omega'$  has advantage  $\epsilon/2$  (see [4]).

Using  $\Omega'$ , the generic approach on [1] shows that  $\pi$  is a hard-core predicate for *any* one-way function  $f$  defined over  $E$ , provided that the output of  $f$  does not depend on the Weierstrass equation used to describe  $E$  (in other words that the function  $f$  is defined over the group of points, irrespective of its representation). Duc and Jethchev call such a function an *elliptic curve-based one-way function* (ECB-OWF) and discuss the application of their result to bilinear pairings defined over elliptic curves, which are indeed a conjectured example of ECB-OWF.

## 4 Hard-Core Predicates for the Diffie-Hellman Problem over Elliptic Curves

In this section, we show our first original result: if the Diffie-Hellman problem over elliptic curves is hard, then every bit (and every segment predicate) of a secret Diffie-Hellman value is unpredictable. This generalizes the result of Boneh and Shparlinski [4] which holds only for the least significant bit.

For a security parameter  $\ell$ , consider an instance generator  $\mathcal{E}$  which on input  $1^\ell$  outputs  $E_\ell$  an elliptic curve defined over  $\mathbb{F}_{p_\ell}$  where  $p_\ell$  is a  $\ell$ -bit prime, such that  $G_\ell$  is a cyclic subgroup of  $E_\ell$  (under the standard group operation defined over the curve points) generated by a point  $P_\ell$ . In the following, we will drop the suffix  $\ell$  when it is clear from the context. We also use the additive notation for the group operation over  $E$ , therefore every point  $Q \in G$  can be written as  $Q = aP$  for some  $a \in \{1, \dots, |G|\}$ .

**Assumption 4.1.** *We say that the Diffie-Hellman problem over  $\mathcal{E}$  is hard if for every polynomial time machine  $A$ , we have that the probability*

$$\Pr[A(E_\ell, P_\ell, aP_\ell, bP_\ell) = abP_\ell \mid E_\ell \leftarrow \mathcal{E}(1^\ell); a, b \leftarrow \{1, \dots, |G|\}]$$

*is negligible in  $\ell$ .*

For every point  $Q \in E$  we denote with  $Q_x$  the  $x$ -coordinate of  $Q$ . As before we denote with  $W(E)$  the set of short Weirstrass equations describing a curve  $E$ ; recall that each  $W \in W(E)$  can be uniquely associated with a  $\lambda \in \mathbb{F}_p^\times$  which gives rise to the isomorphism  $\Phi_\lambda$  defined in the previous section.

Let  $B_k : \mathbb{F}_p \rightarrow \{\pm 1\}$  denote the  $k$ -th bit predicate and let  $\beta_k$  be the bias of  $B_k$ . We now state our first main theorem. Intuitively it says that under Assumption 4.1, every bit of the binary expansion of the  $x$ -coordinate of  $abP$  is unpredictable (e.g., pseudorandom) for a random representation of the curve  $E$ .

**Theorem 4.2.** *Under Assumption 4.1, for any polynomial time machine  $\Omega$ ,*

$$\left| \Pr[\Omega(\lambda, P, aP, bP) = B_k([\Phi_\lambda(abP)]_x) \mid \lambda \leftarrow \mathbb{F}_p^\times; a, b \leftarrow \{1, \dots, |G|\}] - \beta_k \right|$$

*must be negligible.*

The intuition of the proof is as follows. The crucial observation is that the techniques of Duc and Jetchev [6] apply not just to ECB-OWFs but to any computation which “respects” the isomorphism  $\Phi_\lambda$  defined by a change in the Weirstrass representation of the curve. The Diffie-Hellman problem is one such problem since applying the Diffie-Hellman transform to  $\Phi_\lambda(aP), \Phi_\lambda(bP)$  yields the value  $\Phi_\lambda(abP)$  – indeed this is at the basis of the result of [4]. Therefore, an oracle  $\Omega$  contradicting Theorem 4.2 on input  $aP, bP$  and a curve  $W_\lambda$  defined by a parameter  $\lambda \in \mathbb{F}_p^\times$  would output a bit equal to  $B_k(\lambda^2[abP]_x)$  with non-negligible advantage. This allows us to construct a multiplication code with the required properties and apply the framework of [1] to prove that the predicate is hard-core.

*Remark 4.3.* The extension to segment predicates follow from using the concentration and recoverability arguments for those predicates as presented in [1].

*Proof.* Assume that there exists an oracle  $\Omega$  such that the quantity

$$\left| \Pr[\Omega(\lambda, P, aP, bP) = B_k([\Phi_\lambda(abP)]_x) \mid \lambda \leftarrow \mathbb{F}_p^\times; a, b \leftarrow \{1, \dots, |G|\}] - \beta_k \right|$$

is larger than a non-negligible quantity  $\epsilon$ .

From this oracle we build a modified oracle  $\Omega'$  which queries  $\Omega$  if  $\lambda$  is a square in  $\mathbb{F}_p^\times$ , otherwise tosses a  $\beta_k$ -biased coin. It is not hard to see [4] that if  $\Omega$  had advantage  $\epsilon$ , then  $\Omega'$  has advantage  $\epsilon/2$ . We now show how to use  $\Omega'$  to break Assumption 4.1.

Let  $E$  be an elliptic curve defined by an equation  $W \in W(E)$  over  $\mathbb{F}_p$  and let  $G$  be a cyclic subgroup of  $|E|$  generated by the point  $P$ . Given  $P, aP, bP$  we want to compute  $Q = abP$  with non-negligible probability.

Consider the codeword:

$$C_Q : \mathbb{F}_p \rightarrow \{\pm 1\} \text{ defined as } C_Q(\lambda) = B_k(\lambda Q_x).$$

The following properties hold for  $C_Q$ .

**Accessibility:** The oracle  $\Omega'$  gives us access to a noisy version  $\tilde{C}_Q$  of this codeword defined as  $\tilde{C}_Q = \Omega'(\lambda, P, aP, bP)$ . Because  $\Omega'$  has advantage  $\epsilon/2$  we know that  $\Pr_\lambda[C_Q(\lambda) = \tilde{C}_Q(\lambda)] > \beta_k + \epsilon/2$ .

**Concentration:** The codeword  $C_Q$  is a Fourier concentrated function. Indeed for a threshold  $\tau$  the  $\tau$ -heavy characters of  $C_Q$  must belong to the set

$$\Gamma_{Q,\tau} = \{\chi_\beta : \beta = \alpha Q_x \text{ mod } p \text{ for } \alpha \in \Gamma_\tau\},$$

where  $\Gamma_\tau$  is a set of size  $O(\tau^{-2})$  containing the  $\tau$ -heavy coefficients of the function  $B_k$ . We refer the reader to [6, 9] for a proof of this statement and also the definition of  $\Gamma_\tau$  which shows that the elements of  $\Gamma_\tau$  can be easily enumerated. See also Appendix B of the full version [7].

**Recoverability.** Given a Fourier character  $\chi_\beta$  we want to find a set  $L_\beta$  containing all the points  $Q$  such that  $\chi_\beta$  is  $\tau$ -heavy for  $C_Q$ . If  $\chi_\beta$  is  $\tau$ -heavy for  $C_Q$  then  $\chi_\beta \in \Gamma_{Q,\tau}$  and therefore  $Q_x = \beta\alpha^{-1} \text{ mod } p$  for  $\alpha \in \Gamma_\tau$ , therefore

$$L_\beta = \{Q : Q_x = \beta\alpha^{-1} \text{ mod } p \text{ for } \alpha \in \Gamma_\tau\}.$$

By applying Lemma 2.5 we know that there exists a threshold  $\tau$  which is polynomial in  $\epsilon$  and at least one Fourier character  $\chi$  which is  $\tau$ -heavy for both  $C_Q$  and  $\tilde{C}_Q$ .

We then invoke Theorem 2.6 and use the learning algorithm of [1] to learn a polynomial-size list  $L_Q$  of all the  $\tau$ -heavy Fourier characters for  $\tilde{C}_Q$ . For each such character  $\chi_\beta \in L_Q$  we use the recovery property to create a polynomial size list  $L_\beta$  of possible values for  $Q$ . Let  $L = \cup_{\chi_\beta \in L_Q} L_\beta$ ; this is a polynomial-size set and because of Lemma 2.5 it must necessarily include  $Q$ .

More specifically, on input  $E, P, aP, bP$  and with access to  $\Omega$ , the following algorithm produces a polynomial size list of points in  $E$  which is guaranteed to contain  $Q$  with probability  $1 - \delta$ :

1. Let  $\tau$  be the threshold determined by Lemma 2.5 ; note that  $\tau^{-1}$  is polynomial in  $\ell = |p|$ , since  $\epsilon^{-1}$  is.
2. Learn the polynomial-size set  $L_Q$  containing all  $\tau$ -heavy Fourier characters of  $\tilde{C}_Q$ , using the learning algorithm in [1], which is correct with probability  $1 - \delta$ . This algorithms uses oracle  $\Omega'$  to obtain the required query access to  $\tilde{C}_x$ . By applying Lemma 2.5, we know that there exists at least one Fourier character  $\chi$  which is  $\tau$ -heavy for  $C_Q$  and  $\chi \in L_Q$ .
3. Use the recovery algorithm to construct a polynomial-size list of candidates values for  $Q$ . For each  $\chi_\beta \in L_Q$  let

$$\begin{aligned} L_\beta &= \{R \in E : \chi_\beta \text{ is } \tau\text{-heavy for } C_R\} \\ &= \{R \in E : R_x = \beta\alpha^{-1} \text{ mod } p \text{ for } \alpha \in \Gamma\}. \end{aligned}$$

Let  $L = \cup_{\chi_\beta \in L_Q} L_\beta$ . Note that  $L$ 's size is polynomial in  $\ell$  and that  $Q \in L$  with probability  $1 - \delta$ .

The algorithm runs in polynomial time, since the learning algorithm of [1] is efficient and all the enumerations in the algorithm are over polynomial-size lists.

To contradict Assumption 4.1 at this point, it would be sufficient to choose a random point in  $L$ . The probability to select the correct point  $Q$  is  $1/|L|$  and

therefore the algorithm outputs the correct  $Q$  with probability  $(1 - \delta)/|L|$  which is non-negligible since  $|L|$  is of polynomial-size.

Another option is to use the above algorithm as a subroutine in Shoup’s “self-corrector” for the Diffie-Hellman problem (Theorem 7 in [11]). Shoup shows how an algorithm  $A$  that runs in time  $T_A$  and produces a list of  $m$  points, which contains the correct Diffie-Hellman value with probability  $> 7/8$  can be easily converted into an algorithm  $B$  that output only the correct Diffie-Hellman value with overwhelming probability and runs in time  $T_A \ell + \text{poly}(m, \ell)$ .

## 5 Hard-Core Predicates for the Diffie-Hellman Problem over Finite Fields

In this section, we state and prove our main result: after defining a natural (though weaker) variation of the Diffie-Hellman problem over finite fields  $\mathbb{F}_{p^t}$  for  $t > 1$ , we prove that in the case of quadratic extensions ( $t = 2$ ), this problem admits a large class of hard-core predicates, including every single bit of one of the coordinates of the secret value.

For a given prime  $p$ , there are many different fields  $\mathbb{F}_{p^2}$ , but they are all isomorphic to each other. Let  $h(x) = x^2 + h_1x + h_0$  be a monic irreducible polynomial of degree 2 in  $\mathbb{F}_p$ . It is well known that  $\mathbb{F}_{p^2}$  is isomorphic to the field  $\mathbb{F}_p[x]/(h)$ , and therefore elements of  $\mathbb{F}_{p^2}$  can be written as linear polynomials: if  $g \in \mathbb{F}_{p^2}$  then  $g = g_0 + g_1x$  and addition and multiplication are performed as polynomial operations modulo  $h$ . In the following, given  $g \in \mathbb{F}_{p^2}$  we denote with  $[g]_i$  the coefficient of the degree- $i$  term.

Let  $I_2(p)$  be the set of monic irreducible polynomials of degree 2 in  $\mathbb{F}_p$ . For  $h, \hat{h} \in I_2(p)$  we know that there exists an (easily computable) isomorphism

$$\phi_{h, \hat{h}} : \mathbb{F}_p[x]/(h) \rightarrow \mathbb{F}_p[x]/(\hat{h}).$$

Finally, denote with  $g$  a generator of the multiplicative group of  $\mathbb{F}_{p^2}$  which is known to be cyclic.

**A NEW DIFFIE-HELLMAN PROBLEM.** Denote with  $g$  the generator of the multiplicative group of  $\mathbb{F}_{p^2}$  which is known to be cyclic. We define the following variant of the CDH problem over  $\mathbb{F}_{p^2}$ : informally we say that the *Partial-CDH* problem is hard in  $\mathbb{F}_{p^2}$  if no efficient algorithm given  $g, A = g^a, B = g^b \in \mathbb{F}_{p^2}$  can compute  $K = [g^{ab}]_1 \in \mathbb{F}_p$ , for any representation of  $\mathbb{F}_{p^2}$ .

More formally, for a security parameter  $\ell$ , consider an instance generator  $\mathcal{F}$  which on input  $1^\ell$  outputs  $p_\ell$  an  $\ell$ -bit prime. Let  $g_\ell$  be a generator of the multiplicative group of the finite field  $\mathbb{F}_{p_\ell^2}$ . In the following, we will drop the suffix  $\ell$  when it is clear from the context.

**Assumption 5.1.** *We say that the Partial Diffie-Hellman problem over  $\mathcal{F}$  is hard if for every polynomial time machine  $A$ , we have that for all  $h_\ell \in I_2(p_\ell)$  the following probability is negligible in  $\ell$ :*

$$\Pr[A(p_\ell, h_\ell, g_\ell, g_\ell^a, g_\ell^b) = [g_\ell^{ab}]_1 \mid p_\ell \leftarrow \mathcal{F}(1^\ell); a, b \leftarrow \{1, \dots, p_\ell^2 - 1\}].$$

Note that  $A$  gets as input a representation  $h_\ell$  of the field, and that  $A$ 's advantage must be negligible for all representations.

We now state our second main theorem. We show that, when given an oracle  $\Omega$  which predicts the  $k$ th bit of the degree-1 coefficient of the Diffie-Hellman secret with non-negligible advantage (where the probability is taken over the input pair), *as well as the representation of the field*, then one can efficiently solve the Partial Diffie-Hellman problem with non-negligible probability.

**Theorem 5.2.** *Under Assumption 5.1, for any polynomial time machine  $\Omega$  we have that the following quantity must be negligible for all  $h \in I_2(p)$ :*

$$\left| \Pr \left[ \Omega \left( h, \hat{h}, g, g^a, g^b \right) = B_k \left( \left[ \phi_{h, \hat{h}} \left( g^{ab} \right) \right]_1 \right) \right. \right. \\ \left. \left. \mid \hat{h} \leftarrow I_2(p); a, b \leftarrow \{1, \dots, p^2 - 1\} \right] - \beta_k \right|.$$

The proof of Theorem 5.2 appears in Section 5.1. Here we give an informal intuition of the proof.

Our goal is to construct a code similar to that of [6], which must be accessible by querying  $\Omega$  over many different representation of the field. For an element  $\alpha \in \mathbb{F}_{p^2}$ , and a fixed  $h \in I_2(p)$ , a natural definition for a codeword is as follows:

$$C_\alpha(\hat{h}) = B_k \left( \left[ \phi_{h, \hat{h}}(\alpha) \right]_1 \right). \quad (1)$$

This code is accessible using  $\Omega$ , however it is defined over  $I_2(p)$ , and it is not immediately seen to be a multiplication code like the ones used in [1, 6]. Note, however, that the predicate  $B_k$  is evaluated *only on the first coordinate of  $\phi_{h, \hat{h}}(\alpha)$* . In this case, it holds that  $\left[ \phi_{h, \hat{h}}(\alpha) \right]_1 = \lambda[\alpha]_1$  for some  $\lambda \in \mathbb{F}_p^\times$  (see Lemma 5.5 below).

Consider then the following multiplication code over  $\mathbb{F}_p$ : for  $\alpha \in \mathbb{F}_{p^2}$  and for  $\lambda \in \mathbb{F}_p^\times$ , set

$$C_\alpha(\lambda) = B_k(\lambda[\alpha]_1)$$

extended with  $C_\alpha(0) = -1$ . We stress that in light of Lemma 5.5, the above code is conceptually the same as equation (1) in that codewords are obtained by evaluating a predicate over all possible representations of elements. We've simply restricted attention to the degree-1 coordinate. Therefore the multiplication is accessible via  $\Omega$  and then the proof follows similarly to the one in [1, 6].

*Remark 5.3 (List of Candidate Solutions).* The list-decoding algorithm of [1] applied to the code above returns a polynomial size list of possible candidates for  $[\alpha]_1$ . In our reduction  $\alpha = g^{ab}$  and therefore it will be sufficient to output a random element of the list to contradict Assumption 5.1. In contrast to Theorem 4.2, we will not be able to apply Shoup's "self-corrector" in this case to identify the correct solution with high probability, as we have only a single coordinate for  $g^{ab}$ .

*Remark 5.4 (Segment Predicates).* While Theorem 5.2 is stated only for the predicate  $B_k$ , it holds for any predicate  $\pi$  such that the corresponding code  $C_\alpha$  can be proven to be concentrated and recoverable; in particular, it holds for the segment predicates defined in [1].

### 5.1 Proof of Theorem 5.2

We start with a lemma that gives a simple characterization of the isomorphisms between two different representations of the field  $\mathbb{F}_{p^2}$ . When describing such maps, it will be convenient for us to view them as matrices in  $GL_2(\mathbb{F}_p)$ .

**Lemma 5.5.** *For any  $h \in I_2(p)$  there exists a unique function  $L_h : \mathbb{F}_p \times \mathbb{F}_p^\times \rightarrow I_2(p)$  which takes a pair  $(a, b)$  to the polynomial  $\hat{h} = L_h(a, b)$  such that the matrix  $\begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}$  defines an isomorphism  $\mathbb{F}_p[x]/(h) \rightarrow \mathbb{F}_p[x]/(\hat{h})$ . Moreover, for any  $\hat{h} \in I_2(p)$ ,  $L_h^{-1}(\hat{h})$  represents the complete set of isomorphisms from  $\mathbb{F}_p[x]/(h) \rightarrow \mathbb{F}_p[x]/(\hat{h})$  using the above matrix identification.*

*Proof.* First note that any isomorphism of fields must send the unit element to itself (and thus fix the entire base field  $\mathbb{F}_p$ ). Thus, when viewing such an isomorphism as a linear transformation, the first basis element  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  must be fixed, which determines the first column of the matrix as  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . Since clearly we must have  $b \neq 0$  if the map is to represent an isomorphism, the completeness would follow immediately, once we establish the existence and uniqueness of the map  $L_h$ . We define  $L_h$  as follows. For  $a, b \in \mathbb{F}_p$  with  $b \neq 0$ , let  $L_h(a, b)(x) = \frac{h(a+bx)}{b^2}$ . To make the notation less cumbersome, we'll fix  $a, b$  in what follows, and refer to this polynomial more simply as  $L_h(x)$ . To see that this definition is as desired, note that to specify a homomorphism  $\phi$  from  $\mathbb{F}_p[x]/(h)$  to another field  $K$  of characteristic  $p$  it is both necessary and sufficient to choose  $\phi(x) = \bar{x} \in K$  such that  $h(\bar{x}) = 0$  in  $K$ . The matrix corresponding to  $(a, b)$  sends  $x \mapsto a + bx$ , and indeed,  $a + bx$  is a root of  $h$  in the ring  $\mathbb{F}_p[x]/(L_h)$  by construction. However, it remains to show that  $L_h \in I_2(p)$ , as well as the uniqueness of  $L_h$ . Towards the first goal: it is an elementary fact that since  $h$  was irreducible over  $\mathbb{F}_p$ , so is  $h(a + bx)$ , and hence  $L_h$ . It is easy to verify additionally that  $L_h$  is monic, and has degree 2, so that  $L_h \in I_2(p)$ . Thus, by the above remarks, the mapping defined by  $x \mapsto a + bx$  is an isomorphism  $\mathbb{F}_p[x]/(h) \rightarrow \mathbb{F}_p[x]/(L_h)$  as desired. The fact that  $L_h$  so constructed is unique (within  $I_2(p)$ ) follows easily as well, since if  $h(a + bx)$ , and hence  $L_h(x)$ , are elements of an ideal  $(h')$  for some other  $h' \in I_2(p)$ , then  $L_h, h'$  are associates, and thus  $L_h = h'$  since both are monic.

*Remark 5.6.* We actually know a little more about the distribution; in particular, we have  $|L_h^{-1}(\hat{h})| = 2$  for any  $\hat{h} \in I_2(\mathbb{F}_p)$ . This follows at once from the fact that every isomorphism has a (unique) matrix representation as above, and that  $\text{Gal}(\mathbb{F}_{p^2}/\mathbb{F}_p) \cong \mathbb{Z}_2$  (so that there are precisely two isomorphisms between any two representations  $\mathbb{F}_p[x]/(h), \mathbb{F}_p[x]/(\hat{h})$ ).

*Proof Sketch (Theorem 5.2).* Suppose that the theorem were false, and that an oracle  $\Omega$  with an advantage that is not negligible exists. Now consider another oracle  $\Omega'$  that takes as input a base representation  $h \in I_2(p)$ , a Diffie-Hellman triple  $g, g^a, g^b$  as well as an element of  $\lambda \in \mathbb{F}_p$  (instead of  $\hat{h} \in I_2(p)$ ), which works as follows. The oracle selects  $a \leftarrow \mathbb{F}_p$  uniformly at random, and constructs

an isomorphism  $\hat{h}$  from the matrix  $\begin{pmatrix} 1 & a \\ 0 & \lambda \end{pmatrix}$  as described in Lemma 5.5.  $\Omega'$  then returns the output of  $\Omega(h, \hat{h}, g, g^a, g^b)$ . One can then show that

$$\left| \Pr_{\lambda, a, b} [\Omega'(h, \lambda, g, g^a, g^b) = B_k(\lambda[g^{ab}]_1)] - \beta_k \right|$$

is also not a negligible function. At this point, the proof follows closely to that of Theorem 4.2. To begin, observe that we can, for any element  $\alpha \in \mathbb{F}_{p^2}$ , construct the following encoding of  $[\alpha]_1$  in its base polynomial representation as an element of  $\mathbb{F}_p[x]/(h)$ :

$$C_\alpha : \mathbb{F}_p \rightarrow \{\pm 1\} \text{ defined as } C_\alpha(\lambda) = B_k(\lambda[\alpha]_1),$$

where  $[\alpha]_1$  is taken under the representation determined by  $h$ . The fact that this code is concentrated and recoverable follows immediately from the proof of Theorem 4.2. The argument for accessibility is the same, but with the added simplification that we no longer need to restrict to squares in  $\mathbb{F}_p$ .

As in Theorem 4.2, we will be able to efficiently construct a list of candidates for  $[g^{ab}]_1$ . As mentioned, we unfortunately will not be able to apply Shoup’s “self-corrector” in this case as we have only a single coordinate. Nevertheless, we still obtain a contradiction by guessing a random element of the list as the value of  $[g^{ab}]_1$ , since the list is of polynomial size.

## 5.2 Finite Field-Based One-Way Functions

The work of [6] introduces “elliptic curve-based one-way functions”, and goes on to prove interesting hardness results for this *entire class* of functions. Loosely speaking, elliptic curve-based OWF’s are one-way functions which are well defined on isomorphism classes of curves, and do not depend on any specific representation. Similarly, we consider *finite field-based OWF’s*, which are those that do not depend on the isomorphism class. When considering only prime-order fields  $\mathbb{F}_p$ , this concept is somewhat trivial, since once one fix a bit representation for integers, there are no non-trivial isomorphisms. However, the situation becomes far more interesting when one considers field extensions. Even with a fixed representation for integers, there are many different representations of even a quadratic extension (see Lemma 5.5). As demonstrated in [6] for the case of elliptic curves, having a one-way function which is well defined on many different representations may give rise to a number of hardness results that apply to the entire class of functions. We demonstrate similar results, showing that for quadratic extensions, an efficient oracle that predicts the  $k$ -th bit of the input *over a random representation* of the field will imply an efficient procedure that can “partially” invert the function (i.e., if  $f$  is the one-way function, given  $f(\alpha)$ , it computes  $[\alpha]_1$ ).

In order to define a function  $f$  on a finite field, we first define the function on a particular “base” representation  $F$ . Then, to define  $f$  on any other isomorphic copy  $F'$ , we wish to simply compute  $f \circ \psi$ , where  $\psi : F' \rightarrow F$  is an isomorphism. The following definition guarantees that  $f$  is well defined on isomorphism classes of finite fields.

**Definition 5.7.** Let  $F \cong \mathbb{F}_{p^t}$  be a concrete representation of a finite field. A function  $f : F \rightarrow Y$  is said to be finite field-based if for any  $F' \cong F$  and any two isomorphisms  $\psi, \psi' : F' \rightarrow F$ , we have  $f \circ \psi = f \circ \psi'$ .

*Remark 5.8.* Note that any function  $f$  satisfying Definition 5.7 is actually defined on a quotient space,  $F / \sim$ , where  $\alpha \sim \alpha'$  if and only if  $\alpha, \alpha'$  have the same minimal polynomial over  $\mathbb{F}_p$ . Furthermore, any function which is well defined on  $F / \sim$  will satisfy the definition. Thus, an equivalent definition would be to require that  $f(\alpha)$  depends only on the minimal polynomial of  $\alpha$ . (This follows from the fact that the Galois group acts transitively on the roots of irreducible polynomials.)

We now define a natural relaxation of the notion of one-way functions over finite fields, where it is assumed to be hard to output the maximal degree coordinate of the input. While this definition makes sense for the general case  $p^t$  for  $t > 1$ , we only consider the case of quadratic extensions.

Consider the instance generator  $\mathcal{F}$  which on input a security parameter  $1^\ell$ , outputs  $p_\ell$  (an  $\ell$ -bit prime), and a function  $f_\ell : \mathbb{F}_{p_\ell^2} \rightarrow S_\ell$ , where  $S_\ell$  is an arbitrary set. We drop the suffix  $\ell$  when it is clear from the context.

**Definition 5.9.** We say that  $\mathcal{F}$  is partial one-way if for any efficient algorithm  $A$  the following probability is negligible in  $\ell$  for all  $h_\ell \in I_2(p_\ell)$ :

$$\Pr[A(h_\ell, f_\ell(\alpha)) = [\alpha]_1 \mid p_\ell, f_\ell \leftarrow \mathcal{F}(1^\ell); \alpha \leftarrow \mathbb{F}_{p_\ell}[x]/(h_\ell)].$$

Again, note that  $A$  takes as input a representation of the field, but the probability must be negligible for all representations.

In the case of quadratic extensions, we can obtain results similar to what was shown in [6] for elliptic-curve based OWF. In particular, the existence of a noisy oracle which works with non-negligible probability over the point, as well as the representation of the field, will give rise to an efficient procedure which “partially” inverts  $f$  contradicting Definition 5.9. More formally, we have the following.

**Theorem 5.10.** Suppose that  $f$  is a finite field-based partial one-way function, and fix a base representation  $\mathbb{F}_{p^2} = \mathbb{F}_p[x]/(h)$  for some  $h \in I_2(p)$ . Then, for any probabilistic polynomial time machine  $\Omega$ , it must be that the following quantity is negligible:

$$\left| \Pr \left[ \Omega \left( h, \hat{h}, f(\alpha) \right) = B_k \left( \left[ \phi_{\hat{h}, h}(\alpha) \right]_1 \right) \mid \hat{h} \leftarrow I_2(p); \alpha \leftarrow \mathbb{F}_p[x]/(h) \right] - \beta_k \right|.$$

The proof is a combination of the proofs of Theorems 4.2 and 5.2 and will be presented in the full version [7].

*Remark 5.11.* We note that the Diffie-Hellman problem does *not* satisfy the above definition: apart from the fact that the domain is actually two (or three) field elements, the value  $g^{ab}$  is not independent of the representation. However, if one modifies the usual Diffie-Hellman problem to report the minimal polynomial



of  $g^{ab}$  instead, then the definition is satisfied (with the caveat regarding the input coming from a product space). We also remark that the minimal polynomial is efficiently computable; see for example the work of [12]. Finally, we note that for  $\mathbb{F}_{p^t}$ , each of the equivalence classes under  $\sim$  has size  $t$ . Since  $t$  is usually a small constant (in our case, it is 2), the aforementioned conversion in which one “throws away” some information by only considering the minimal polynomial will not affect the problem’s computational character.

## 6 Conclusion and Future Work

We presented a relaxed variant of the Diffie-Hellman problem over finite fields of the form  $\mathbb{F}_{p^t}$  for  $t > 1$  and proved that for the case of quadratic extensions  $\mathbb{F}_{p^2}$ , this problem admits several hard-core predicates (including every single bit of one coordinate of the secret Diffie-Hellman value) over a random representation of the field. These are the first results known for hard-core predicates for the CDH problem over finite fields. We extended this result to a larger class of computationally hard problems (which we called finite field-based partial one-way functions) over such finite fields.

We also proved that the same class of predicates is hard-core for the elliptic curve Diffie-Hellman, over a random representation of the underlying elliptic curve, thereby extending the Boneh-Shparlinski result [4] which worked only for the least significant bit.

Our results can be interpreted as “augmenting” the input to the computational hard problem (being it a one-way function, or the CDH problem) with a random description of the underlying group being used.

Our work leaves several open questions. Perhaps the most natural is to extend the results to  $\mathbb{F}_{p^t}$  for  $t > 2$ . In the case of  $t = 2$ , the isomorphisms from one representation to another amounted, in some sense, to a linear change of variables:  $x \mapsto a + bx$ . This made the set of isomorphisms between representations easy to analyze, and enabled us to show that when restricting attention to the coefficient of  $x$ , each of these maps acts by translation for some  $\lambda \in \mathbb{F}_p^\times$ . For  $t > 2$ , this is *not* the case, and thus our original techniques must be augmented somehow. Perhaps one can find a large (enough) number of representations for which the isomorphisms have the required properties as a linear map.

Other natural questions include the study of the hardness of the Partial-CDH problem in  $\mathbb{F}_{p^t}$  for  $t > 1$ . While it seems quite a reasonable assumption to make, the ultimate goal would be to reduce it to the “full” CDH over another platform. In particular, is it possible to reduce the Partial-CDH over  $\mathbb{F}_{p^t}$  to the regular CDH problem over  $\mathbb{F}_p$ ? A related question is if we can use the hardness of Partial-CDH over, say  $\mathbb{F}_{p^2}$ , to prove the unpredictability of a predicate for the traditional CDH problem over  $\mathbb{F}_p$ .

Finally it is our hope that the techniques presented in this paper could eventually lead to the proof that CDH over  $\mathbb{F}_p$  does have a (deterministic) hard-core predicate.

**Acknowledgments.** The authors would like to thank Adi Akavia and Dimitar Jetchev for several useful discussions and clarifications.

Nelly Fazio's research is sponsored in part by NSF CAREER award #1253927, and by PSC-CUNY award 64578-00 42, funded by the Professional Staff Congress and CUNY. Nelly Fazio and William E. Skeith III are sponsored in part by NSF award #1117675. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Nelly Fazio and Rosario Genaro are supported in part by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

## References

1. Akavia, A., Goldwasser, S., Safra, S.: Proving hard-core predicates using list decoding. In: IEEE Symposium on Foundations of Computer Science—FOCS, pp. 146–157 (2003)
2. Alexi, W., Chor, B., Goldreich, O., Schnorr, C.: Rsa and rabin functions: Certain parts are as hard as the whole. *SIAM Journal on Computing* 17(2), 194–209 (1988)
3. Blum, M., Micali, S.: How to generate cryptographically strong sequences of pseudorandom bits. *SIAM Journal on Computing* 13(4), 850–864 (1984)
4. Boneh, D., Shparlinski, I.E.: On the unpredictability of bits of the elliptic curve diffie–hellman scheme. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 201–212. Springer, Heidelberg (2001)
5. Diffie, W., Hellman, M.: New directions in cryptography. *IEEE Transactions on Information Theory* IT-22(6), 644–654 (1976)
6. Duc, A., Jetchev, D.: Hardness of computing individual bits for one-way functions on elliptic curves. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 832–849. Springer, Heidelberg (2012)
7. Fazio, N., Gennaro, R., Perera, I.M., Skeith III, W.E.: Hard-core predicates for a diffie-hellman problem over finite fields. *Cryptology ePrint Archive*, Report 2013/134 (2013)
8. Goldreich, O., Levin, L.A.: A hard-core predicate for all one-way functions. In: ACM Symposium on Theory of Computing—STOC, pp. 25–32 (1989)
9. Morillo, P., Ràfols, C.: The security of all bits using list decoding. In: Jarecki, S., Tsudik, G. (eds.) PKC 2009. LNCS, vol. 5443, pp. 15–33. Springer, Heidelberg (2009)
10. Näslund, M.: All bits in  $ax + b \bmod p$  are hard. In: Koblitz, N. (ed.) CRYPTO 1996. LNCS, vol. 1109, pp. 114–128. Springer, Heidelberg (1996)
11. Shoup, V.: Lower bounds for discrete logarithms and related problems. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 256–266. Springer, Heidelberg (1997)
12. Shoup, V.: Efficient computation of minimal polynomials in algebraic extensions of finite fields. In: Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation, pp. 53–58. ACM (1999)