INJECTIVE ENCODINGS TO ELLIPTIC CURVES

PIERRE-ALAIN FOUQUE, ANTOINE JOUX, AND MEHDI TIBOUCHI

Abstract. We investigate the problem of constructing efficient, efficiently invertible injective maps with large image to the set of rational points of an elliptic curve over a finite field, and obtain an essentially optimal solution for a large families of curves, including all Edwards curves with a complete addition law.

1. Introduction

Various cryptographic protocols based on the hardness of Diffie-Hellman-like problems in a certain group $\mathbb{G}$, such as El Gamal encryption [6] or Lindell's recent universally-composable commitment scheme [12], assume the existence of an efficient (possibly randomized) algorithm f mapping messages $m \in \{0,1\}^\ell$ to elements of $\mathbb{G}$, in such a way that m can also be recovered efficiently from $f(m)$. For example, El Gamal encryption is *a priori* defined on group elements, so that a message needs to be mapped to an element of $\mathbb{G}$ before encrypting it, and mapped back to a bit string upon decryption. Moreover, the size ℓ of supported bit strings should preferably as close as possible to the bit size of $\mathbb{G}$ to maximize bandwidth. We call such an algorithm f an injective encoding.

For certain groups $\mathbb{G}$, like multiplicative groups of finite fields or certain supersingular elliptic curves, it is not difficult to construct injective encodings achieving the optimal value of ℓ. On the other hand, for a general group $\mathbb{G}$, it is not obvious how to construct a function f with ℓ even super-logarithmic in the size of $\mathbb{G}$. In §2.3, we prove that this is not possible with a deterministic generic group algorithm.

When $\mathbb{G}$ is the group of points of any elliptic curve over a finite field, one can construct a probabilistic injective encoding with ℓ equal to about half of the size of $\mathbb{G}$, as we show in §2.4, but we do not know constructions achieving a better ℓ in general. Recently, however, a solution was proposed by Farashahi [7] in the special case of Hessian elliptic curves over finite fields $\mathbb{F}_q$ with $q \equiv 2 \pmod 3$.

In §4, we propose an essentially optimal construction for a new, large class of ordinary elliptic curves over fields $\mathbb{F}_q$ with $q \equiv 3 \pmod 4$, including all curves with exactly two rational points of exact order 4 (these are birational to the well-known Edwards curves with complete addition law, studied by Edwards and Bernstein–Lange [1]. Our construction is based on the bijective encoding from [9] to certain hyperelliptic curves of genus 2 and 3, and on the observation from [10] that those curves are quadratic covers of elliptic curves.

1991 *Mathematics Subject Classification.* 14G50, 94A60, 14G15.
Key words and phrases. Elliptic Curve Cryptography, Injective Encoding, Algebraic Curves.

2. Injective encodings

2.1. **Definition.** To fix ideas, and although it is not essential for our main purpose, let us first give a formal definition of what we mean by an "injective encoding".

Let us say that a *cyclic group family* $(\mathbb{G}_k)_{k\in\mathbb{N}}$ consists in the data of a sequence of integers $n_k \geq 1$ converging to infinity, a sequence of integers $s_k \geq 0$ that is at most polynomial in $\log n_k$, and for each k, an efficiently computable bijection σ_k between the cyclic group $\mathbb{Z}/n_k\mathbb{Z}$ of order n_k and a set $\mathbb{G}_k \subset \{0,1\}^{s_k}$ of bit strings of length s_k, as well as efficient algorithms:

$$\oplus_k \colon \{0,1\}^{s_k} \times \{0,1\}^{s_k} \to \{0,1\}^{s_k} \cup \{\bot\} \qquad \ominus_k \colon \{0,1\}^{s_k} \to \{0,1\}^{s_k} \cup \{\bot\}$$

which induce on the $\mathbb{G}_k$ the group addition and negation obtained by transport of structure via σ_k. Here, "efficient" means with a time complexity polynomial in $\log n_k$ (or equivalently, in s_k).

For example, if q_k is an increasing sequence of positive prime powers, we can construct a cyclic group family $\mathbb{G}_k = \mathbb{F}_{q^k}^*$ with $n_k = q_k - 1$ and $s_k = O(\log q_k)$ by representing invertible elements in $\mathbb{F}_{q_k}$ as polynomials over the corresponding prime field (themselves the concatenation their coefficients as bit strings). Similarly, if E is an elliptic curve over $\mathbb{Z}[1/N]$ with N coprime with the q_k's such that $E(\mathbb{F}_{q_k})$ is cyclic for all k, we have a cyclic group family $\mathbb{G}_k = E(\mathbb{F}_{q_k})$ with $n_k = q_k + O(\sqrt{q_k})$ and $s_k = O(\log q_k)$ obtained by representing curve points in e.g. affine coordinates (with a special string for the point at infinity).

Given such a cyclic group family $(\mathbb{G}_k)$ and a sequence of non negative integers ℓ_k, we define an *ℓ_k-injective encoding* to $(\mathbb{G}_k)$ be the data consisting of a pair of efficient, possibly randomized algorithms:

$$\mathscr{F}_k \colon \{0,1\}^{\ell_k} \to \mathbb{G}_k \subset \{0,1\}^{s_k} \qquad \mathscr{I}_k \colon \{0,1\}^{s_k} \to \{0,1\}^{\ell_k} \cup \{\bot\}$$

for all k, which satisfy $\mathscr{I}_k(\mathscr{F}_k(m)) = m$ for all $m \in \{0,1\}^{\ell_k}$ with overwhelming probability over the randomness involved. We will typically express ℓ_k in terms of $\nu_k = \lfloor \log_2 n_k \rfloor$, which is the optimal bound, in the sense that we clearly have $\ell_k \leq \nu_k$ for all k by injectivity.

In what follows, the indices k, as well as references to sequences of integers and groups, will be omitted most of the time for simplicity's sake.

2.2. **Some simple, optimal examples.** Let p be an odd prime number. The bijection $[1, p-1] \to \mathbb{F}_p^*$ yields an obvious injective injective encoding to the multiplicative group $\mathbb{G} = \mathbb{F}_p^*$ which is optimal, in the sense that $\ell = \nu$.

Similarly, we obtain an optimal injective encoding to the group of squares $\mathbb{G} = (\mathbb{F}_p^*)^2 \subset \mathbb{F}_p^*$ from the bijection $[1, \frac{p-1}{2}] \to (\mathbb{F}_p^*)^2$ given by $x \mapsto x^2$. The inversion algorithm $\mathscr{I}$ then computes the unique square root of an element in $(\mathbb{F}_p^*)^2$ contained in $[1, \frac{p-1}{2}]$. This is sufficient to obtain IND–CPA El Gamal encryption in the group $(\mathbb{F}_p^*)^2$ when p is a safe prime, assuming the Decisional Diffie–Hellman assumption in that group (though one typically wouldn't want to use it for efficiency reasons). On the other hand, it is not clear how to construct a close to optimal injective encoding to the subgroup of prime order q in $\mathbb{F}_p^*$ when p is a Diffie–Hellman prime $p = 2r \cdot q + 1$.

Some elliptic curve groups also have optimal injective encodings. This is for example the case for the supersingular elliptic curves given by an equation of the form:

$$E \colon y^2 = x^3 + b$$

over a field $\mathbb{F}_q$ with $q \equiv 2 \pmod 3$. Then, as observed e.g. by Boneh and Franklin [3], the map $\mathbb{F}_q \to E(\mathbb{F}_p)\setminus\{\infty\}$ given by $u \mapsto \big((u^2-b)^{1/3}; u\big)$ is an efficient bijection, and its inverse is clearly efficient as well. This gives, again, an optimal injective encoding to $\mathbb{G} = E(\mathbb{F}_q)$. Similarly, the genus 1 case of the construction we proposed in [9] provides an optimal injective encoding to supersingular elliptic curves of the form:

$$E\colon y^2 = x^3 + ax$$

over fields $\mathbb{F}_q$ with $q \equiv 3 \pmod 4$. However, we are not aware of any strictly optimal injective encoding to groups of points of ordinary elliptic curves.

2.3. **Generic injective encodings.** It is easy to construct ℓ_k-injective encodings to any cyclic group family $(\mathbb{G}_k)$ provided that $\ell_k = O(\log \nu_k)$ (and of course $\ell_k \leq \nu_k$ for all k). Indeed, in that case, the set $\{0,1\}^{\ell_k}$ of elements to be encoded contains only polynomially many elements: therefore, $\mathscr{F}_k$ and $\mathscr{I}_k$ can be defined as mutually inverse dictionary lookups for each k, and still be efficient. For example, we can define $\mathscr{F}_k$ to be the restriction of σ_k to $\{0,1,\ldots,2^{\ell_k}-1\} \subset \mathbb{Z}/n_k\mathbb{Z}$ (coded as bit strings in the obvious way), and $\mathscr{I}_k$ as a series of 2^{ℓ_k} successive comparisons. Moreover, $\mathscr{F}_k$ and $\mathscr{I}_k$ are clearly *generic group algorithms* in the sense of Shoup [14].

On the other hand, if $\ell_k = \omega(\log \nu_k)$, then it is easy to see that $\mathscr{F}_k$ and $\mathscr{I}_k$ cannot be both generic group algorithms for all k if the $\mathscr{F}_k$'s are deterministic. Indeed, suppose that it were the case. Since it doesn't take any group element as input, $\mathscr{F}_k$ must be of the form:

$$\mathscr{F}_k(m) = \sigma_k\big(f(m)\big)$$

for some efficiently computable function $f_k\colon \{0,1\}^{\ell_k} \to \mathbb{Z}/n_k\mathbb{Z}$, by definition of a generic group algorithm. Then, let $S = \mathscr{F}_k(\{0,1\}^{\ell_k})$ be the image of $\mathscr{F}_k$. The generic group algorithm $f_k \circ \mathscr{I}_k$ computes the discrete logarithm $\sigma_k^{-1}(g)$ of any element $g \in S$ with overwhelming probability in $\mathrm{poly}(\nu_k)$ steps. As a result, by Shoup's argument, we must have $\#S = \mathrm{poly}(\nu_k)$: a contradiction.

This means that deterministic injective encodings from sets of superlogarithmic bit size to must use the particular representation of individual group elements. We conjecture that no *probabilistic* generic $\omega(\log \nu)$-injective encoding exists either, although this seems less easy to establish.

2.4. **Injective encodings to elliptic curves.** For groups of points of arbitrary (even ordinary) elliptic curves over finite prime fields, it is possible to construct ℓ-injective encodings for much larger values ℓ than in the generic case. We propose one such construction here.

Let E be an elliptic curve over $\mathbb{F}_p$ ($p \geq 5$) in short Weierstrass form, and ℓ an integer such that $\ell \leq (1/2-\varepsilon)\log_2 p$ for some fixed constant $\varepsilon \in (0, 1/2)$. We define the encoding algorithm $\mathscr{F}\colon \{0,1\}^{\ell} \to E(\mathbb{F}_p)$ as follows. To compute $\mathscr{F}(m)$, pick a random integer x in $[0, p-1]$ whose least significant ℓ bits coincide with m. If there are points in $E(\mathbb{F}_p)$ of abscissa $x \bmod p$, return one of those (at most two) points; otherwise, start over. The inversion algorithm $\mathscr{I}$ then simply maps a point $(x, y) \in E(\mathbb{F}_p)$ to the bit string m formed by the ℓ least significant bits of x.

To prove that this method works, it suffices to show that $\mathscr{F}$ terminates in expected polynomial time on any input m. We obtain the following, more precise result.

Theorem 1. *If p is large enough, the expected number of iterations in $\mathscr{F}$ on any input is less than 3.*

Proof. Let $P(m)$ be the success probability of $\mathscr{F}$ on input m after a single iteration; in other words, $P(m)$ is the probability that a random integer x in $[0, p-1]$ whose least significant ℓ bits coincide with m is the abscissa of a point in $E(\mathbb{F}_p)$. Since for each such x there are at most two corresponding points in $E(\mathbb{F}_p)$, we have:

$$P(m) \geq \frac{1}{2} \cdot \frac{\#\{(x,y) \in E(\mathbb{F}_p) \mid \textsc{lsb}_\ell(x) = m\}}{\#\{x \in [0, p-1] \mid \textsc{lsb}_\ell(x) = m\}} \tag{1}$$

where $\textsc{lsb}_\ell(x)$ denotes the bit string formed by the ℓ least significant bits of x. Clearly we have

$$\#\{x \in [0, p-1] \mid \textsc{lsb}_\ell(x) = m\} \leq 2^{-\ell} \cdot p.$$

On the other hand, the value $\#\{(x,y) \in E(\mathbb{F}_p) \mid \textsc{lsb}_\ell(x) = m\}$ can be estimated as in [8, §6]. It is the number of $\mathbb{F}_p$-points (x,y) of E such that x/p is in a certain interval of $\mathbb{R}/\mathbb{Z}$ of length $\geq 2^{-\ell} \cdot (1 - 2/p)$ (because x can be of the form $m + 2^\ell \cdot r$ at least for any $r \in [0, \lfloor p/2^\ell \rfloor - 1]$). But the values x/p in $\mathbb{R}/\mathbb{Z}$ for $(x,y) \in E(\mathbb{F}_p)$ are close to equidistributed. More precisely, we know from Bombieri's bound on character sums [2] that for any nontrivial additive character ψ of $\mathbb{F}_p$, we have:

$$T(\psi) = \left| \sum_{(x,y) \in E(\mathbb{F}_p) \setminus \{\infty\}} \psi(x) \right| \leq 4\sqrt{p}. \tag{2}$$

As a result, the (1-dimensional version of) the Erdős–Turán–Koksma inequality [5, Th. 1.21] gives, for any interval $I \subset \mathbb{R}/\mathbb{Z}$ of length L and any positive integer $H < p$:

$$\left| \frac{\#\{(x,y) \in E(\mathbb{F}_p) \setminus \{\infty\} \mid \frac{x}{p} \in I\}}{\#E(\mathbb{F}_p) \setminus \{\infty\}} - L \right| \leq \frac{3}{H+1} + \frac{3}{\#E(\mathbb{F}_p) \setminus \{\infty\}} \sum_{h=1}^{H} \frac{T(\psi_h)}{h}$$

where ψ_h is the additive character $x \mapsto e^{2i\pi hx/p}$. Setting $H = \sqrt{p} - 1$ and $N = \#E(\mathbb{F}_p) \setminus \{\infty\}$, we get, in view of (2):

$$\begin{aligned} \#\{(x,y) \in E(\mathbb{F}_p) \setminus \{\infty\} \mid \frac{x}{p} \in I\} &\geq L \cdot N - \frac{3N}{\sqrt{p}} - 3 \cdot 4\sqrt{p} \log \sqrt{p} \\ &\geq L \cdot p - 2L\sqrt{p} - 3\sqrt{p} - 6 - 6\sqrt{p} \log p \\ &\geq L \cdot p - 12\sqrt{p} \log p \end{aligned}$$

since $|N - p| \leq 2\sqrt{p}$ by the Hasse bound. Plugging this estimate back into (1), we finally obtain:

$$P(m) \geq \frac{1}{2} \cdot \frac{2^{-\ell}(1 - 2/p)p - 12\sqrt{p} \log p}{2^{-\ell} \cdot p} = \frac{1}{2} - \frac{7 \log p}{p^\varepsilon}$$

since $\ell \leq (1/2 - \varepsilon) \log_2 p$. Hence, the expected number of iteration in $\mathscr{F}$ is $1/P(m) \leq 3$ for large enough p as required. □

Thus, we can construct ℓ-injective encodings to elliptic curves over prime fields for $\ell = (1/2 - \varepsilon)\nu$: this is much better than the logarithmic bound we get in the generic case, but this still falls short of optimality by a constant factor greater than 2. It is conceivable that the same algorithm does in fact work with a larger ℓ still, possibly as large as $(1 - \varepsilon)\nu$ or even $\nu - \log^{O(1)} \nu$; we doubt that current results on the distribution of points on elliptic curves are sufficient to prove that the algorithm

terminates on all inputs on those cases, however (though it should be possible to bound its complexity *on average* over all inputs m).

The only injective encoding to ordinary elliptic curves in the literature achieving a better bound is, to our knowledge, the one proposed by Farashahi in [7]. It applies to Hessian curves (i.e. elliptic curves with a rational point of order 3) over fields $\mathbb{F}_q$ with $q \equiv 2 \pmod 3$, and achieves $\ell = \nu - 1$, a single bit short of optimal. In the next sections, we construct a similar deterministic injective encoding to elliptic curves over fields $\mathbb{F}_q$ with $q \equiv 3 \pmod 4$ which have rational point of order 4 and only one point of order 2 (these are birational to the computationally interesting Edwards curves), also with $\ell = \nu - 1$.

3. On some special families of hyperelliptic curves

3.1. **Mapping to "odd" hyperelliptic curves.** Let f be an odd polynomial over a finite field $\mathbb{F}_q$ with $q \equiv 3 \pmod 4$, which has simple roots in $\overline{\mathbb{F}}_q$. We denote its degree by $2g+1$, and consider the hyperelliptic curve over $\mathbb{F}_q$ defined by:

$$H\colon y^2 = f(x) = a_0 x^{2g+1} + a_1 x^{2g-1} + \cdots + a_g x.$$

In [9], we defined an "almost bijective" map $F\colon \mathbb{F}_q \to H(\mathbb{F}_q)$ as follows.

Denote by $\sqrt{\cdot}$ the usual square root function on the set of quadratic residues in $\mathbb{F}_q$ (exponentiation by $(q+1)/4$), and by $\chi_q(\cdot)$ the nontrivial quadratic character of $\mathbb{F}_q^*$, extended by 0 to $\mathbb{F}_q$ as usual. Let further $\varepsilon(t) = \chi_q(f(t))$ for all $t \in \mathbb{F}_q$. Then, we define F as:

$$\begin{aligned} F\colon \mathbb{F}_q &\longrightarrow H(\mathbb{F}_q) \\ t &\longmapsto \left(\varepsilon(t)\cdot t \;;\; \varepsilon(t)\sqrt{\varepsilon(t)\cdot f(t)}\right). \end{aligned} \tag{3}$$

Let $W \subset H(\mathbb{F}_q)$ be the set of $\mathbb{F}_q$-rational Weierstrass points of H (the points where the rational function y is ramified, i.e. the point at infinity and those of the form $(x,0)$), and $T \subset \mathbb{F}_q$ the set of roots of f in $\mathbb{F}_q$. In [9], we proved:

Lemma 2. *The function F given by (3) is well-defined, maps all points in T to $(0,0) \in W$, and induces a bijection $\mathbb{F}_q \setminus T \to H(\mathbb{F}_q) \setminus W$.*

3.2. **Hyperelliptic curves with a certain non-hyperelliptic involution.** We now turn our attention to hyperelliptic curves over $\mathbb{F}_q$ (where $\mathbb{F}_q$ is as above, and in particular of odd characteristic) defined by another special type of polynomials. Let $f(x) = b_{2g+1}x^{2g+1} + b_{2g}x^{2g} + \cdots$ be a squarefree polynomial of degree $2g+1$ over $\mathbb{F}_q$, whose coefficients b_j satisfy the relation $b_{2g+2-j} = b_j$. In other words:

$$x^{2g+2} f(1/x) = f(x). \tag{4}$$

Then we obtain immediately that the hyperelliptic curve:

$$H\colon y^2 = f(x)$$

admits the following non-hyperelliptic involution σ defined over $\mathbb{F}_q$:

$$\begin{aligned} \sigma\colon H &\longrightarrow H \\ (x;y) &\longmapsto \left(\frac{1}{x} \;;\; \frac{y}{x^{g+1}}\right). \end{aligned} \tag{5}$$

We claim that the quotient curve $H/\langle\sigma\rangle$ is a hyperelliptic curve of genus $\lfloor g/2 \rfloor$ for which we can give an explicit equation.

Theorem 3. *There is a (unique) squarefree polynomial $h \in \mathbb{F}_q[u]$ of degree $g+1$ satisfying the relation:*

$$h_0(u^2) = (1+u)^{2g+2} f\left(\frac{1-u}{1+u}\right).$$

The quotient curve $H/\langle\sigma\rangle$ is then isomorphic to the hyperelliptic curve of genus $\lfloor g/2 \rfloor$ defined by $H_0 \colon v^2 = h_0(u)$, and the quotient map $H \to H/\langle\sigma\rangle \cong H_0$ is given by:

$$(x;y) \mapsto (u;v) = \left(\left(\frac{1-x}{1+x}\right)^2 \;;\; y\left(\frac{2}{1+x}\right)^{g+1}\right).$$

Proof. To establish the first claim, we introduce the polynomial:

$$h(t) = (1+t)^{2g+2} f\left(\frac{1-t}{1+t}\right) = b_{2g+1}(1-t)^{2g+1}(1+t) + b_{2g}(1-t)^{2g}(1+t)^2 + \cdots \in \mathbb{F}_q[t].$$

Clearly, h is of degree at most $2g+2$. Moreover, the coefficient of t^{2g+2} is $-b_{2g+1} + b_{2g} - b_{2g-1} + \cdots = f(-1)$, which we can show is non zero.

Indeed, suppose $f(-1) = 0$. Taking the derivative of (4), we see that $f'(x) = (2g+2)x^{2g+1} f(1/x) - x^{2g} f'(1/x)$, which for $x = -1$ gives $f'(-1) = 0 - f'(-1)$, hence $f'(-1) = 0$. As a result, we see that -1 is a double root of f, which is a contradiction since f was assumed to be squarefree.

Hence, h is a polynomial of degree $2g+2$. In addition, we have:

$$h(-t) = (1-t)^{2g+2} f\left(\frac{1+t}{1-t}\right) = (1-t)^{2g+2} \left(\frac{1+t}{1-t}\right)^{2g+2} f\left(\frac{1-t}{1+t}\right) = h(t)$$

so that the polynomial h is even, and there is thus a unique polynomial h_0 of degree exactly $g+1$ such that:

$$h_0(t^2) = h(t) = (1+t)^{2g+2} f\left(\frac{1-t}{1+t}\right). \tag{6}$$

It remains to see that h_0 is squarefree. Suppose that h_0 has a double root $u_0 \in \overline{\mathbb{F}}_q$. We can write it as $u_0 = t_0^2$ for some $t_0 \in \overline{\mathbb{F}}_q$, and since we can replace t_0 by $-t_0$, we can assume $t_0 \neq -1$. Then, let $x_0 = (1-t_0)/(1+t_0)$. By (6) we have $f(x_0) = 0$. Moreover, taking the derivative of (6) we obtain the following relation in $\mathbb{F}_q[t]$:

$$2t \cdot h_0'(t^2) = (2g+2)(1+t)^{2g+1} f\left(\frac{1-t}{1+t}\right) - 2(1+t)^{2g} f'\left(\frac{1-t}{1+t}\right).$$

For $t = t_0$, this gives $-2(1+t_0)^{2g} f'(x_0) = 0$, hence x_0 is a double root of f in $\overline{\mathbb{F}}_q$, a contradiction.

Hence, the curve $H_0 \colon v^2 = h_0(u)$ is a hyperelliptic curve of genus $\lfloor g/2 \rfloor$ as required. Evidently, it is the quotient of $H_1 \colon v^2 = h_0(t^2) = h(t)$ by the involution $\sigma_1 \colon (t;v) \mapsto (-t;v)$, and the rational maps:

$$\begin{aligned} H &\longrightarrow H_1 \\ (x;y) &\longmapsto \left(\frac{1-x}{1+x} \;;\; y\left(\frac{2}{1+x}\right)^{g+1}\right) \\ \left(\frac{1-t}{1+t} \;;\; \frac{v}{(1+t)^{g+1}}\right) &\longmapsto\!\shortmid\; (t;v) \end{aligned}$$

are inverse of each other. Since the involutions σ and σ_1 correspond to each other under these isomorphisms, this completes the proof. □

Remark 4. (1) *The previous results also apply to a polynomial f of degree $2g+2$. We focused on degree $2g+1$ as we will be considering odd polynomials f in the next section.*
(2) *The one-line idea of the proof is of course that the rational function $t = (1-x)/(1+x)$ of degree 1 on H satisfies $\sigma^* t = -t$; hence, writing the equation of H in terms of t leads to the form $v^2 = h_0(t^2)$ for some h_0, which makes the quotient is easy to find.*
(3) *The fact that $H/\langle\sigma\rangle$ is a curve of genus $\lfloor g/2 \rfloor$ can be seen directly with the Riemann-Hurwitz formula, and occurs in various forms in the literature. See for example* [13, Th. 4].

3.3. **Behavior of F with respect to the involution.** Putting the results of §3.1 and §3.2 together, consider a polynomial f over $\mathbb{F}_q$ of degree $2g+1$ which is odd and satisfies $f(x) = x^{2g+2} f(1/x)$. Then the hyperelliptic curve

$$H\colon y^2 = f(x)$$

admits both the non-hyperelliptic involution σ from §3.2 and the encoding function $F\colon \mathbb{F}_q \to H(\mathbb{F}_q)$ from §3.1. In this paragraph, we look into the relation between the two.

For any $t \in \mathbb{F}_q^*$, we have:

$$\varepsilon\left(\frac{1}{t}\right) = \chi_q\Big(f\Big(\frac{1}{t}\Big)\Big) = \chi_q\big(t^{-2g-2} \cdot f(t)\big) = \chi_q(f(t)) = \varepsilon(t).$$

Moreover, for any $u \in \mathbb{F}_q$, $\sqrt{u^2} = u^{(q+1)/2} = \chi_q(u) \cdot u$. As a result, if we let $F(t) = (x; y)$, we can compute:

$$\begin{aligned} F\left(\frac{1}{t}\right) &= \left(\varepsilon\Big(\frac{1}{t}\Big)\frac{1}{t}\ ;\ \varepsilon\Big(\frac{1}{t}\Big)\sqrt{\varepsilon\Big(\frac{1}{t}\Big) f\Big(\frac{1}{t}\Big)}\right) \\ &= \left(\frac{\varepsilon(t)}{t}\ ;\ \varepsilon(t)\sqrt{\varepsilon(t)\frac{1}{t^{2g+2}} f(t)}\right) \\ &= \left(\frac{\varepsilon(t)}{t}\ ;\ \varepsilon(t)\frac{\chi_q(t)^{g+1}}{t^{g+1}}\sqrt{\varepsilon(t) f(t)}\right) \\ &= \left(\frac{1}{x}\ ;\ \big(\varepsilon(t)\chi_q(t)\big)^{g+1} \cdot \frac{y}{x^{g+1}}\right). \end{aligned}$$

In particular, if g is odd, the following simple relation always holds:

$$F(1/t) = \left(\frac{1}{x}\ ;\ \frac{y}{x^{g+1}}\right) = \sigma\big(F(t)\big).$$

This makes it easy to construct an injective encoding to $H_0(\mathbb{F}_q)$ where $H_0 = H/\langle\sigma\rangle$, whose image contains about $q/2$ points. However, this case is actually never interesting, because H_0 is then isomorphic to an "odd" hyperelliptic curve of genus $(g-1)/2$ (in the sense of §3.1). To see this, note that with the same notation as in the proof of Theorem 3, the polynomial h is self-reciprocal (because f is odd),

and hence so is h_0. As a result, we have $h_0(u) = u^{g+1}h_0(1/u)$. It follows that the polynomial:

$$r(x) = (1+x)^{g+1} \cdot h_0\left(\frac{1-x}{1+x}\right)$$

is odd, and since $g+1$ is even, H_0 is isomorphic to the odd hyperelliptic curve $y^2 = r(x)$. We can thus use the results of [9] to obtain an "almost bijective" map $\mathbb{F}_q \to H_0(\mathbb{F}_q)$, as well as a $\big(\nu - O(1)\big)$-injective encoding to the Jacobian of H_0.

Therefore, we concentrate on the case when g is even, which we study in more details from now.

3.4. **An injective function when g is even.** In the situation of the previous paragraph, we see that when g is even, $F(1/t)$ can be either $\sigma\big(F(t)\big)$ or its image under the hyperelliptic involution, depending on t, so the map does not directly pass to the quotient. To overcome this problem, we restrict ourselves to an even more special case.

We now choose f squarefree of the form $f(x) = x^{g+1}f_1(x^2)f_1(1/x^2)$, where f_1 is an even, squarefree polynomial of degree $g/2$ whose leading and constant coefficients $c_{g/2}$ and c_0 are such that $c_{g/2}c_0 = \pm 1$, and that satisfies $f_1(1) \neq 0$. This shape directly implies that f is of degree $2g+1$, odd, and satisfies $f(x) = x^{2g+2}f(1/x)$ so we are indeed in a particular case of the above discussion. Then, for all $t \in \mathbb{F}_q$, let:

$$\alpha(t) = \chi_q\big(f_1(t^2)\big)\,.$$

We claim that for all $t \in \mathbb{F}_q$ with $f(t) \neq 0$, the following holds:

$$\alpha(t) = \varepsilon(t)\chi_q(t) \cdot \alpha(1/t). \tag{7}$$

Indeed, for any $t \in \mathbb{F}_q^*$, we have:

$$\varepsilon(t)\chi_q(t) = \chi_q\left(t^{g+2} \cdot f_1(t^2) \cdot f_1\Big(\frac{1}{t^2}\Big)\right) = 1 \cdot \alpha(t) \cdot \alpha(1/t)$$

which, if t is not a root of f, yields (7) by multiplying both sides by $\alpha(1/t)$, which is then necessarily non zero.

Now, introduce a "twisted" variant F_α of F defined for all $t \in \mathbb{F}_q$ by $F_\alpha(t) = \big(x; \alpha(t)y\big)$, where $(x;y) = F(t)$. Since $F(-t) = (x;-y)$, we have for all $t \in \mathbb{F}_q$:

$$F_\alpha(-t) = \big(x; \alpha(-t)\cdot(-y)\big) = \big(x; -\alpha(t)y\big)$$

so that $F_\alpha(-t)$ is the image of $F_\alpha(t)$ under the hyperelliptic involution of H. Moreover, we obtain, for all $t \in \mathbb{F}_q$ other than roots of f:

$$F_\alpha\left(\frac{1}{t}\right) = \left(\frac{1}{x}\ ;\ \alpha\Big(\frac{1}{t}\Big)\cdot\varepsilon(t)\chi_q(t)\cdot\frac{y}{x^{g+1}}\right) = \left(\frac{1}{x}\ ;\ \alpha(t)\cdot\frac{y}{x^{g+1}}\right) = \sigma\big(F_\alpha(t)\big),$$

so that F_α does pass to the quotient and gives a map to the points of $H_0 = H/\langle\sigma\rangle$ which in turn yields the injective map we are looking for. To see that, denote by G the map $H(\mathbb{F}_q) \to H_0(\mathbb{F}_q)$ induced by the covering. We first prove the following result.

Lemma 5. *Let $S \subset \mathbb{F}_q$ be any subset of $\mathbb{F}_q$ containing no root of f, and such that $S \cap S^{-1} = \varnothing$ (i.e. for all $x \in S$, $1/x \notin S$). Then, the restriction of $G \circ F_\alpha$ to S is injective.*

Proof. Consider $t, t' \in S$ such that $G\big(F_\alpha(t)\big) = G\big(F_\alpha(t')\big)$. We must have either $F_\alpha(t) = F_\alpha(t')$ or $F_\alpha(t) = \sigma\big(F_\alpha(t')\big) = F_\alpha(1/t')$.

In the latter case, we see in particular that the first coordinates of $F_\alpha(t)$ and $F_\alpha(1/t')$ coincide, so that $t = \pm 1/t'$. By definition of S, $t = 1/t'$ is excluded, so we must have $t = -1/t'$. Now observe that G commutes with the hyperelliptic involution, i.e. we have $G \circ \tau_H = \tau_{H_0} \circ G$, where τ_H, τ_{H_0} are the maps induced on $H(\mathbb{F}_q)$ and $H_0(\mathbb{F}_q)$ by the corresponding hyperelliptic involutions: this is obvious from the explicit expression of G given in the proof of Theorem 3. We can thus write:

$$G\big(F_\alpha(t')\big) = G\big(F_\alpha(-1/t')\big) = G\big(\tau_H F_\alpha(1/t')\big) = \tau_{H_0} G\big(\sigma F_\alpha(t')\big) = \tau_{H_0} G\big(F_\alpha(t')\big).$$

Therefore, $G\big(F_\alpha(t')\big)$ is a Weierstrass point on H_0. Given the expression of G, this implies that $F_\alpha(t')$ is a Weierstrass point on H, and hence that t' is a root of f, which is a contradiction.

If on the other hand $F_\alpha(t) = F_\alpha(t')$, we see in particular by comparing the first coordinates of $F_\alpha(t)$ and $F_\alpha(t')$ that $t' = \pm t$. But since S contains no root of f, $F_\alpha(t)$ is not a Weierstrass point, so it is not equal to its image $F_\alpha(-t)$ under the hyperelliptic involution τ_H. Hence $t' = -t$ is impossible, and we must have $t = t'$ as required. □

Corollary 6. *Let S be as in Lemma 5. The restriction of $G \circ F_\alpha$ to $S \cup \{0, 1\}$ is injective.*

Proof. The images of 0 and 1 under $G \circ F_\alpha$ in $H_0(\mathbb{F}_q)$ are:

$$\begin{aligned} G\big(F_\alpha(0)\big) &= G\big((0;0)\big) = (1;0) \\ G\big(F_\alpha(1)\big) &= G\big((\varepsilon(1); \alpha(1)\varepsilon(1)\sqrt{\varepsilon(1)f(1)})\big) = G\big((1; f_1(1))\big) = \big(0; f_1(1)\big) \end{aligned}$$

since $f(1) = f_1(1)^2$ and hence $\varepsilon(1) = 1$ and $\sqrt{f(1)} = \alpha(1) f_1(1)$. Thus, we see that the images of 0 and 1 under $G \circ F_\alpha$ are distinct. Moreover, it follows from the previous proof that for all $t \in S$, $G\big(F_\alpha(t)\big)$ is never a Weierstrass point on H_0, and hence is always distinct from $G\big(F_\alpha(0)\big)$. And finally, if there was some $t \in S$ such that $G\big(F_\alpha(t)\big) = G\big(F_\alpha(1)\big)$, then, again following the previous proof, we would have $t = \pm 1$, which is impossible since $S \cap S^{-1} = \varnothing$. □

Now fix $I \subset \mathbb{F}_q$ a subset of $\mathbb{F}_q$ of cardinal $(q-1)/2$ such that $I \cap (-I) = \varnothing$ and $-1 \notin I$. If q is prime, we can simply take $I = [1, (q-1)/2]$. If q is a power of the prime p, we can choose a basis of $\mathbb{F}_q$ as an $\mathbb{F}_p$-vector space, and set I as the set of elements in $\mathbb{F}_q^*$ whose first non zero component on that basis is in $[1, (p-1)/2]$.

Then, remove from I all the elements of the form $\frac{1-t}{1+t}$ where t is a root of f, and add 0 and 1 to the resulting set to get I_0. In other words:

$$I_0 = \left(I \setminus \left\{ \frac{1-t}{1+t} \mid f(t) = 0 \right\} \right) \cup \{0, 1\}.$$

Note that for any non zero root t of f, the four values t, $-t$, $1/t$ and $-1/t$ are pairwise distinct roots of f (because $f(\pm 1) \neq 0$ and -1 is a non quadratic residue), and exactly two of them satisfy $\frac{1-t}{1+t} \in I$. As a result, the number of roots of f is of the form $4\rho + 1$, and we have $\#I_0 = (q-1)/2 - (2\rho+1) + 2 = (q+1)/2 - 2\rho$.

We are then in a position to define our injective function:

$$\begin{aligned} F_{\text{inj}} \colon I_0 &\longrightarrow H_0(\mathbb{F}_q) \\ u &\longmapsto (G \circ F_\alpha)\left(\frac{1-u}{1+u}\right). \end{aligned} \tag{8}$$

The following is an immediate consequence of Corollary 6.

Theorem 7. *The function F_{inj} is well-defined and injective.*

In the next section, we make it explicit how this function F_{inj} yields $(\nu-1)$-injective encodings to the group of points of a large family of elliptic curves.

4. A new encoding to elliptic curves

4.1. **Explicit encoding.** In the situation of §3.4, the quotient curve H_0 is an elliptic curve when $g = 2$. In that case, the polynomial f_1 is of degree 1, of the form $f_1(x) = cx + \delta/c$ for some $c \in \mathbb{F}_q^*$ and $\delta = \pm 1$. Moreover, since $f(x) = x^3 f_1(x) f_1(1/x)$ is squarefree, we must have $c \neq \pm 1$, and all the other values of c give a valid polynomial f. The hyperelliptic curve H is then:

$$H \colon y^2 = f(x) = \delta x^5 + \left(c^2 + \frac{1}{c^2}\right) x^3 + \delta x.$$

Its quotient $H_0 = H/\langle\sigma\rangle$ is the elliptic curve of equation $v^2 = h_0(u)$ where h_0 is the polynomial of degree 3 defined in Theorem 3, and the function F_{inj} readily provides an injective encoding to $H_0(\mathbb{F}_q)$.

Clearly, when $\delta = +1$, the polynomial $f_1(x^2)$ has no root in $\mathbb{F}_q$, and as a result, 0 is the only root of f. Therefore, the range I_0 of F_{inj} is of cardinality $(q+1)/2$; when q is prime, it is the interval $[0, (q-1)/2]$.

On the other hand, when $\delta = -1$, the roots of f are $0, \pm c, \pm 1/c$, and I_0 is then of cardinality $(q+1)/2 - 2 = (q-3)/2$; when q is prime, it is the interval $[0, (q-1)/2]$ from which one has removed $\pm t, \pm 1/t$ where $t = \frac{1-c}{1+c}$.

In both cases, we see that the size of the set from which we encode is a single bit less than the cardinality $\#H_0(\mathbb{F}_q) = q + O(\sqrt{q})$ of the target group. Hence, we do get a deterministic $(\nu-1)$-injective encoding as stated.

4.2. **Target elliptic curves.** Let us also look more closely at the elliptic curve H_0. We can compute h_0 and find that:

$$h_0(1-u) = \left(c + \frac{\delta}{c}\right)^2 u^3 - 16\delta u^2 + 16\delta u.$$

By applying the change of coordinates $(u; v) \mapsto (1-u; v)$ and then scaling the coordinates as $(u; v) \mapsto \big(4(c+\delta/c)^2 u; 8(c+\delta/c)^2 v\big)$, we finally find that H_0 is isomorphic to the elliptic curve:

$$E_c^\delta \colon y^2 = x^3 - 4\delta x^2 + \delta\left(c + \delta/c\right)^2 x.$$

This curve obviously has a rational point of exact order 2, namely $(0; 0)$. When $\delta = +1$, it is the only one; indeed, the trinomial $x^2 - 4x + (c+1/c)^2$ has discriminant $16 - 4(c+1/c)^2 = -4(c-1/c)^2$ which is a non quadratic residue. On the other hand, if $\delta = -1$, all three points of exact order 2 are rational, since $x^2 + 4x - (c-1/c)^2$ has discriminant $16 + 4(c-1/c)^2 = 4(c+1/c)^2$ which is a square.

Furthermore, there is a rational point P such that $[2]P = (0; 0)$ if and only if $\delta = +1$. To see that, it suffices to show that there is a line through $(0; 0)$ which is

tangent to the curve, since the intersection point will clearly satisfy the requirement. Now if $y = tx$ is a line through $(0;0)$, the other intersection points with the curve have their abscissa given by $t^2x = x^2 - 4\delta x + \delta(c + 1/c)^2$, and the line is tangent when the discriminant of this quadratic equation vanishes, i.e. when t satisfies:

$$(4\delta + t^2)^2 = 4\delta \left(c + \frac{1}{c}\right)^2.$$

There is no solution when $\delta = -1$ since the right-hand side is not square. On the other hand, when $\delta = +1$, this is equivalent to:

$$t^2 = -4 \pm 2\left(c + \frac{1}{c}\right)$$

and this equation has a solution for one of the two possible signs, because $\big(-4 + 2(c+1/c)\big) \cdot \big(-4-2(c+1/c)\big) = -4(c-1/c)^2$ is a non quadratic residue, and hence exactly one of the factors must be square.

Thus, in all cases, we see that the curve admits a rational subgroup of order 4. In fact, the rational 4-torsion is of order 4 or 8: namely $\mathbb{Z}/4\mathbb{Z}$ when $\delta = +1$, and $(\mathbb{Z}/2\mathbb{Z})^2$ or $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ when $\delta = -1$ depending on whether one of the points of order 2 other than $(0;0)$ is divisible by two.

4.3. **Curves isomorphic to $E_c^{\pm}$.** Conversely, we claim that, *up to a quadratic twist*, any elliptic curve over $\mathbb{F}_q$ with a point of order 4 and only one point of order 2 is isomorphic to E_c^+ for some c. Indeed, let E be any such elliptic curve. We can put E in Weierstrass form, translate so that the point of order 2 is $(0;0)$, and scale the coordinates to get an equation of the form:

$$E\colon y^2 = x^3 \pm 4x^2 + ax$$

for some $a \in \mathbb{F}_q$, with $a \neq 0, 4$ since the right-hand side must have no double root. Note that the nontrivial quadratic twist of E has the same equation, only with the sign of the coefficient of x^2 reversed.

Since there is a single point of order 2, the discriminant $16 - 4a$ of the trinomial $x^2 \pm 4x + a$ must be a non quadratic residue. Hence, $a-4$ is a square. Moreover, $(0;0)$ is divisible by two: therefore, there exists a t such that the line of slope t through $(0;0)$ is tangent to the curve. This t is such that the equation $t^2x = x^2 \pm 4x + a$ has a double root, so we must have $(-t^2 \pm 4)^2 - 4a = 0$, hence a is a square. And the discriminant of the trinomial $c^2 - \sqrt{a} \cdot c + 1$ is $a - 4$, so there is a $c \in \mathbb{F}_q \setminus \{0, \pm 1\}$ such that $a = (c + 1/c)^2$. This shows that E is either E_c^+ or its quadratic twist as required.

Note that curves with a point of order 4 and only one point of order 2 are birational to Edwards curves $x^2 + y^2 = 1 + dx^2y^2$ with non square d [1]. Bernstein and Lange showed that these curves are quite interesting for computation and cryptography, as they admit a complete addition law, and admit the fastest arithmetic known to date.

It seems a bit more difficult to find a nice characterization of curves isomorphic to E_c^- or to its twist. Consider any elliptic curve E over $\mathbb{F}_q$ with full rational 2-torsion (these curves are also isomorphic to computationally interesting curves, namely twisted Huff curves [11]). As above, we can put E in the form:

$$E\colon y^2 = x^3 \pm 4x^2 + ax$$

for some $a \in \mathbb{F}_q$ with $a \neq 0, 4$, and since the right-hand side splits in linear factors, $4 - a$ is a square. *If in addition* a happens to be a non quadratic residue, then E is isomorphic to either E_c^- or its twist. Indeed, $-a$ is then a square, and the discriminant of the trinomial $c^2 - \sqrt{-a} \cdot c - 1$ is $-a + 4$ which is a square as well; hence, we can find $c \in \mathbb{F}_q \setminus \{0, \pm 1\}$ such that $a = -(c - 1/c)^2$. However, a nice characterization of the cases when a is a non quadratic residue escapes us at the time of this writing (it is not the case, in particular, that a must be a non quadratic residue whenever $(0; 0)$ is not divisible by two).

4.4. **Mapping to the twist.** The previous paragraph suggests that if E is e.g. an Edwards curve $x^2 + y^2 = 1 + dx^2y^2$ with non square d, then we know an injective encoding to *either* $E(\mathbb{F}_q)$ itself or to $E'(\mathbb{F}_q)$, where E' is the nontrivial quadratic twist of E. But we can in fact do better and map to $E(\mathbb{F}_q)$ itself!

Indeed, if H is as in §4.1, then it is classical (see e.g. [13] or [4, Ch. 14]) that is does not only cover the elliptic curve $H_0 \colon v^2 = h_0(u)$ given by the quotient by σ, but also $H_0' \colon v^2 = u^3 h_0(1/u)$ given by the quotient by $\sigma\tau$. Moreover, we have $u^3 h_0(1/u) = -h_0(u)$, so that H_0' is the nontrivial quadratic twist of H_0.

It is easy to adapt the discussion of §3.4 to obtain a similar injective function F'_{inj} to $H_0'(\mathbb{F}_q)$, and hence a $(\nu - 1)$-injective encoding to the twists of $E_c^{\pm}$. We conclude:

Theorem 8. *Let E be an elliptic curve over a finite field $\mathbb{F}_q$ with $q \equiv 3 \pmod 4$.*

1. *If E has an $\mathbb{F}_q$-point of exact order 4 and a single point of exact order 2, then there is a $(\nu - 1)$-injective encoding to $E(\mathbb{F}_q)$.*
2. *If $E(\mathbb{F}_q)$ has a subgroup isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$, it admits an equation of the form $y^2 = x^3 \pm 4x^2 + ax$. If a is a non quadratic residue, then there is a $(\nu - 1)$-injective encoding to $E(\mathbb{F}_q)$.*

5. Conclusion

In this paper, we propose an efficient injective encoding with almost optimally large image for a new class of elliptic curves including important examples like Edwards curves. The only previous construction in that direction was for Hessian curves.

Note that, from a cryptographic perspective, this does not completely solve the problem of construction an encoding for El Gamal encryption, as the curves we encode to have a small subgroup which can reveal information about the message (i.e. El Gamal is one-way but not semantically secure in this setting). This is similar to the situation of El Gamal in multiplicative groups $\mathbb{F}_p^*$ when p is not a safe prime.

References

[1] D. J. Bernstein and T. Lange. Faster addition and doubling on elliptic curves. In K. Kurosawa, editor, *ASIACRYPT*, volume 4833 of *Lecture Notes in Computer Science*, pages 29–50. Springer, 2007.

[2] E. Bombieri. On exponential sums in finite fields. In *Les Tendances Géom. en Algèbre et Théorie des Nombres*, pages 37–41. Éditions du CNRS, 1966.

[3] D. Boneh and M. K. Franklin. Identity-based encryption from the weil pairing. In J. Kilian, editor, *CRYPTO*, volume 2139 of *Lecture Notes in Computer Science*, pages 213–229. Springer, 2001.

[4] J. Cassels and E. Flynn. *Prolegomena to a middlebrow arithmetic of curves of genus 2*. Number 230 in London Mathematical Society Lecture Note Series. Cambridge University Press, 1996.

[5] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*. Springer, 1997.
[6] T. El Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472, 1985.
[7] R. R. Farashahi. Hashing into Hessian curves. In A. Nitaj and D. Pointcheval, editors, *AFRICACRYPT*, volume 6737 of *Lecture Notes in Computer Science*, pages 278–289. Springer, 2011.
[8] R. R. Farashahi, P.-A. Fouque, I. E. Shparlinski, M. Tibouchi, and J. F. Voloch. Indifferentiable deterministic hashing to elliptic and hyperelliptic curves. *Math. Comp.*, 2012. To appear.
[9] P.-A. Fouque and M. Tibouchi. Deterministic encoding and hashing to odd hyperelliptic curves. In M. Joye, A. Miyaji, and A. Otsuka, editors, *Pairing*, volume 6487 of *Lecture Notes in Computer Science*, pages 265–277. Springer, 2010.
[10] A. Joux and V. Vitse. Cover and decomposition index calculus on elliptic curves made practical. Application to a previously unreachable curve over $\mathbb{F}_{p^6}$. In D. Pointcheval and T. Johansson, editors, *EUROCRYPT*, volume 7237 of *Lecture Notes in Computer Science*. Springer, 2012.
[11] M. Joye, M. Tibouchi, and D. Vergnaud. Huff's model for elliptic curves. In G. Hanrot, F. Morain, and E. Thomé, editors, *ANTS*, volume 6197 of *Lecture Notes in Computer Science*, pages 234–250. Springer, 2010.
[12] Y. Lindell. Highly-efficient universally-composable commitments based on the DDH assumption. In K. G. Paterson, editor, *EUROCRYPT*, volume 6632 of *Lecture Notes in Computer Science*, pages 446–466. Springer, 2011.
[13] J. Paulhus. Decomposing Jacobians of curves with extra automorphisms. *Acta. Arith.*, 132(3):231–244, 2008.
[14] V. Shoup. Lower bounds for discrete logarithms and related problems. In W. Fumy, editor, *EUROCRYPT*, volume 1233 of *Lecture Notes in Computer Science*, pages 256–266. Springer, 1997.

École normale supérieure, 45 rue d'Ulm, f-75230 Paris Cedex 05, France
E-mail address: pierre-alain.fouque@ens.fr

DGA and Université de Versailles–Saint-Quentin, 45 rue des États-Unis, f-78035 Versailles Cedex, France
E-mail address: antoine.joux@m4x.org

NTT Information Sharing Platform Laboratories, 3–9–11 Midori-cho, Musashino-shi, Tokyo 180–8585, Japan
E-mail address: tibouchi.mehdi@lab.ntt.co.jp