Efficient Arithmetic on Subfield Elliptic Curves over Small Odd Characteristics

Keisuke Hakuta[1], Hisayoshi Sato[1], and Tsuyoshi Takagi[2]

[1] Hitachi, Ltd., Systems Development Laboratory,
1099, Ohzenji, Asao-ku, Kawasaki, 215-0013, Japan
{k-hakuta,hisato}@sdl.hitachi.co.jp
[2] Future University - Hakodate,
School of Systems Information Science,
116-2, Kamedanakano-cho, Hakodate, 041-8655, Japan
takagi@fun.ac.jp

Abstract. In elliptic curve cryptosystems, scalar multiplications performed on the curves have much effect on the efficiency of the schemes, and many efficient methods have been proposed. In particular, recoding methods of the scalars play an important role in the performance of the algorithm used. For integer radices, non-adjacent form (NAF) and its generalizations (e.g., generalized non-adjacent form (GNAF) and radix-r non-adjacent form (rNAF) [4, 21]) are proposed for minimizing the non-zero densities in the representations of the scalars. On the other hand, for subfield elliptic curves, Frobenius-adic expansions of the scalars can be used for improving efficiency ([18]). Unfortunately, there are only a few methods apply the techniques of NAF or its analogue to Frobenius-adic expansion, namely τ-adic NAF techniques ([11, 20, 2] and [6]) for Koblitz curves and hyperelliptic Koblitz curves. In this paper, we try to combine these techniques, namely recoding methods for reducing non-zero density and Frobenius-adic expansion, and propose two new efficient recoding methods of scalars for more general family of subfield elliptic curves over odd characteristics. We also prove that the non-zero densities for the new methods are same as those for original GNAF and rNAF. As a result, the speed of the proposed schemes improve between 12.5% and 79% over that for previously known schemes.

Keywords: *generalized non-adjacent form (GNAF), radix-r non-adjacent form (rNAF), Frobenius-adic expansions, τ-adic NAF (τ-NAF), ϕ-adic NAF (ϕ-NAF), Elliptic Curve Cryptosystems (ECC)*

1 Introduction

Elliptic curve cryptosystems (ECC) were proposed in 1985 independently by Victor Miller [13] and by Neal Koblitz [9]. Since ECC provide many advantages, for example, shorter key length and faster computation speed than those of RSA cryptosystems, ECC have been the focus of much attention. In ECC, each protocol such as ECDH, ECElGamal, and ECDSA involves scalar multiplications for given points on an elliptic curve by large integers. These multiplications have much effect on the efficiency of the schemes, and many efficient methods have been proposed.

As one such method, the use of subfield elliptic curves (i.e. elliptic curves over finite fields which are actually defined over some subfield [3]) is especially attractive because by using the Frobenius maps, which is efficiently calculated, scalar multiplication on subfield elliptic curves can be performed much faster than that on curves over prime fields. Indeed, Smart ([18]) shows that every element $d \in \mathbb{Z}[\phi]$ can be written as $d = \sum_{i=0}^{\ell-1} d_i \phi^i$, where $d_i \in \{0, \pm 1, \cdots, \pm(q-1)/2\}$, q is an order of the defining field of an $\mathbb{F}_q$-subfield elliptic curve E, and ϕ is the q-th power Frobenius map on E. Therefore, we can use not point doubling but Frobenius map when we perform scalar multiplication on a subfield elliptic curve. Note that neither of these methods can be applied in the case of curves over prime fields (the case in which the group of prime field rational points is used to cryptosystem). In [7], the authors proposed efficiently computable endomorphisms other than Frobenius endomorphisms can be used for fast scalar multiplication. Moreover, in [14], the authors proposed two kinds of endomorphisms in [7] that can be used together for a certain class of curves, and they also presented a new expansion method.

On the other hand, recoding method of the scalars also plays an important role in the performance. In general, smaller non-zero densities in the representations of scalars improve the efficiency. Non-adjacent form (NAF) and its generalizations such as generalized non-adjacent form (GNAF [4]) and radix-r non-adjacent form (rNAF [21]), are methods used for minimizing the non-zero densities. So as to achieve further improvement, it has been tried to combine the subfield curve method with the recoding methods. In [20], Solinas proposed an efficient method of scalar multiplication for Koblitz curves, namely τ-adic NAF (τ-NAF), and [11] proposed τ-adic NAF for some supersingular elliptic curves defined over the prime field of characteristic three using the Frobenius endomorphism of the curves. In addition, [6] proposed a generalization of τ-adic NAF for hyperelliptic Koblitz curves. Recently, in [2], the authors proposed the radix-τ width-w NAF for every integer in all Euclidean quadratic imaginary fields. However, only a few curves are available for the above methods so far.

1.1 Contribution of this paper

The contribution of this paper is to propose two generalizations of τ-NAF, that is, two classes of ϕ-adic NAF (ϕ-GNAF and ϕ-rNAF) using the techniques of GNAF and rNAF, respectively, which can be applied to a family of subfield elliptic curves defined over finite fields of odd characteristics. The digit set of NAF is $\{0, \pm 1\}$ and the digit set of Frobenius-adic expansion is $\{0, \pm 1, \cdots, \pm(q-1)/2\}$. We can not directly apply the technique of NAF to Frobenius-adic expansions except for τ-NAF for Koblitz curves because of the narrowness of the digit set of NAF. Thus as a natural development, we apply the GNAF and rNAF techniques, which are the generalizations of ordinary NAF, to generalize τ-NAF to elliptic curves over odd characteristics. For the resulting recoding methods, ϕ-GNAF and ϕ-rNAF, if the radix is small (e.g., $3, 5$), the difference between the computational costs for the precomputation tables of ϕ-GNAF and ϕ-rNAF is relatively small (at most, several elliptic additions). But, if the radix is significantly large, the computational cost for the precomputation table of ϕ-rNAF is quite large compared to that for ϕ-GNAF. However the non-zero density for ϕ-rNAF is significantly smaller than that for ϕ-GNAF. Thus, these two generalizations are used for the most appropriate applications. The speed of the proposed schemes improved between 12.5% and 79% over that for previously known schemes. In this paper, as the first step in the generalizations of ϕ-NAF, we concentrate on investigating only ϕ-GNAF and ϕ-rNAF, and we do not deal with the width-w versions of these.

This paper is organized as follows. Section 2 reviews ordinary GNAF, rNAF, and τ-adic NAF for Koblitz curves. Section 3 shows how to generalize τ-NAF for Koblitz curves to two classes of ϕ-adic NAF for a family of subfield elliptic curves and proves some properties of ϕ-GNAF and ϕ-rNAF. Section 4 compares the total computational costs of several previous methods and the proposed methods.

2 Preliminaries

In this paper, in general, for any complex number $\psi(\neq 0)$, we denote $\sum_{i=0}^{\ell-1} c_i\psi^i$ with $c_i \in \mathbb{Z}$ by $(c_{\ell-1}, \cdots, c_0)_\psi$. Hamming weight of $(c_{\ell-1}, \cdots, c_0)_\psi$ is defined by the number of non-zero c_i's. According to convention, we denote $-a$ by $\bar{a}$ for any natural number a.

2.1 GNAF, rNAF

In this section, we review ordinary GNAF and rNAF. Let r, α be relatively prime positive integers. We denote $D_{r,\alpha}$ a set defined as follows.

$$D_{r,\alpha} := \begin{cases} \{0, \pm 1, \cdots, \pm\alpha\} & \text{if } \alpha < r, \\ \{0, \pm 1, \cdots, \pm\alpha\} \setminus \{\pm r, \pm 2r, \cdots, \pm\lfloor \alpha/r \rfloor r\} & \text{otherwise.} \end{cases}$$

For an integer radix $r \geq 2$, GNAF and rNAF are proposed for minimizing the numbers of non-zero densities in the representations of integer scalars. In [4] and [21], the authors calculate the non-zero densities using Markov chains. In this paper, we regard non-zero densities of some representations as average densities of non-zero digits of the representations (See Section 3 for precise definitions).

Definition 1 [GNAF [4]] *A radix-r generalized non-adjacent form (GNAF) of a positive integer d is a representation $d = \sum_{i=0}^{\ell-1} e_i r^i$ where $e_i \in D_{r,r-1}, e_{\ell-1} \neq 0$ and for each i, one of the following holds : (1) $e_{i+1}e_i = 0$, (2) if $e_{i+1}e_i > 0$, then $|e_{i+1} + e_i| < r$, (3) if $e_{i+1}e_i < 0$, then $|e_{i+1}| > |e_i|$. The length of the GNAF is ℓ. For $a, b \in D_{r,r-1}$, if a, b satisfy one of the followings : (1) $ab = 0$, (2) if $ab > 0$, then $|a + b| < r$, (3) if $ab < 0$, then $|a| > |b|$, then we call a pair (a, b) radix-r admissible pair, and otherwise, we call (a, b) radix-r non-admissible pair.*

Definition 2 [r**NAF [21]**] *A radix-r non-adjacent form (rNAF) of a positive integer d is a representation $d = \sum_{i=0}^{\ell-1} e_i r^i$ where $e_i \in D_{r,(r^2-1)/2}$, $e_{\ell-1} \neq 0$ and for each i, it satisfies $e_{i+1}e_i = 0$ where we define $e_\ell = 0$. The length of the rNAF is ℓ. For $a, b \in D_{r,(r^2-1)/2}$, if $ab = 0$, then we call a pair (a, b) radix-r non-adjacent pair, and otherwise, we call (a, b) radix-r adjacent pair.*

In the above definitions, note that for the radix $r = 2$, GNAF and rNAF coincide, and in this case, we call these recoding method "NAF"([21], pp.104). We can see that GNAF and rNAF have some desired properties. For details, consult [4] for GNAF, and [21] for rNAF.

Proposition 1 [Properties of GNAF (*resp.* r**NAF)]**
(1) *Every positive integer d has a unique GNAF (resp. rNAF) representation.*
(2) *GNAF (resp. rNAF) representation of d has the smallest Hamming weight among all signed representations of d with digit set $D_{r,r-1}$ (resp. $D_{r,(r^2-1)/2}$).*
(3) *The average non-zero density of GNAF (resp. rNAF) is asymptotically $(r-1)/(r+1)$ (resp. $(r-1)/(2r-1)$).*

2.2 Subfield elliptic curves and Frobenius-adic expansion

We briefly review subfield elliptic curves and Frobenius-adic expansion on the curves which we focus on in this paper. For detail, refer [18], [3] and [17].

Definition 3 [$\mathbb{F}_q$**-subfield elliptic curves**] *Let p be an odd prime, $q = p^r$ a power of p, and $\mathbb{F}_q$ the finite field with q-elements. An elliptic curve defined over $\mathbb{F}_q$ is called an "$\mathbb{F}_q$-subfield elliptic curve" if for some cryptographic usage, we focus on the group of $\mathbb{F}_{q^n}$-rational points $E(\mathbb{F}_{q^n})$ for some $n \geq 2$. An $\mathbb{F}_q$-subfield elliptic curve E is given by a Weierstrass equation $y^2 = x^3 + ax + b$, where $a, b \in \mathbb{F}_q$. Let us denote ϕ the q^{th}-power Frobenius map on E.*

$$\phi : E \to E, \quad (x, y) \mapsto (x^q, y^q),$$

and put $t_n := q^n + 1 - \#E(\mathbb{F}_{q^n})$, where $E(\mathbb{F}_{q^n})$ means the set of $\mathbb{F}_{q^n}$-rational points on E. We can regard ϕ as a complex number which satisfies the following characteristic equation : $\phi^2 - t_1\phi + q = 0$.

In [18], the authors show that every element $d \in \mathbb{Z}[\phi]$ has a ϕ-adic representation with some digit set. More precisely, they show the followings.

Theorem 1 [Frobenius-adic expansion for subfield elliptic curves] *Let E be an elliptic curve over $\mathbb{F}_q$ and ϕ be its q^{th}-power Frobenius map of E. Let $d \in \mathbb{Z}[\phi]$, then we can write $d = \sum_{i=0}^{\ell-1} c_i\phi^i$, where $c_i \in \{0, \pm1, \cdots, \pm(q-1)/2\}$ and $\ell \leqslant \lceil 2\log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)}\rceil + 4$.*

2.3 τ-NAF for Koblitz curves and supersingular Koblitz curves

Next, we review τ-NAF for Koblitz curves and supersingular Koblitz curves.

Definition 4 [τ-NAF for ordinary Koblitz curves] *Let E_a be an elliptic curve defined over $\mathbb{F}_2$ with defining equation as follows.*

$$E_a : y^2 + xy = x^3 + ax^2 + 1, \quad a = 0 \text{ or } 1.$$

It can be proven that these curves are ordinary, and we call these "ordinary Koblitz curves". Let τ be the 2nd power Frobenius map of E_a. A τ-adic NAF representation (τ-NAF) for ordinary Koblitz curves of an element $d \in \mathbb{Z}[\tau]$ is a representation $d = \sum_{i=0}^{\ell-1} e_i \tau^i$ where $e_i \in D_{2,1}, e_{\ell-1} \neq 0$ and no two consecutive digits e_i are nonzero. The length of the τ-NAF is ℓ.

Theorem 2 [Properties of τ-NAF [20], [1]]
(1) *Every $d \in \mathbb{Z}[\tau]$ has a unique τ-NAF.*
(2) *τ-NAF of d has the smallest Hamming weight with digit set $D_{2,1}$.*
(3) *The average non-zero density of τ-NAF is asymptotically $1/3$ and it has the same non-zero density of ordinary NAF.*

Koblitz [11] also proposed another possibility of NAF-like recoding method for Frobenius-adic representations on supersingular elliptic curves $E_a : y^2 = x^3 - x - (-1)^a/\mathbb{F}_3$, where $a = 0$ or 1. Similarly to the case of ordinary Koblitz curves, the uniqueness of τ-NAF for these curves can be proven. Moreover, it is also proven that the non-zero density for this method is 2/5. See [11] for details. It is unknown any analogues of τ-NAF representation for another subfield elliptic curves except for Koblitz curves and curves in [2]. But in [2], they proposed only non-adjacent radix-τ expansions for integers in all Euclidean quadratic imaginary fields. In the next section, we will propose two classes of ϕ-NAF representation for a family of subfield elliptic curves. In the following, we call two classes of ϕ-NAF representation ϕ-GNAF, ϕ-rNAF, respectively.

3 Proposed Methods (Two classes of ϕ-NAF)

In this section, we investigate how to expand two classes of ϕ-NAF for a family of subfield elliptic curves. Let $SEC_{t_1}[\mathbb{F}_q]$ be the set of $\mathbb{F}_q$-subfield elliptic curves with the q^{th} power Frobenius trace t_1. In the following, **we focus on the case** $t_1 = 1$ and consider a scalar multiplication for a given integer d and for a given point $P \in E(\in SEC_1[\mathbb{F}_q])$. Note that although curves of trace one are anomalous, this does not mean the attack on anomalous curves in [19, 15] and [16] applies in this context. For all $P \in E(\mathbb{F}_{q^n})$, it satisfies that $(\phi^n - 1)P = \mathcal{O}$. Hence $dP = (d \bmod (\phi^n - 1))\, P$ for any integer scalar d. From [18], there exist $Q, d' \in \mathbb{Z}[\phi]$ such that $d = Q(\phi^n - 1) + d'$ with $d' = 0$ or $\Psi(d') < \lambda \Psi(\phi^n - 1)$, where Ψ is a multiplicative function. Note that this provides a 50% improvement in the performance about the length of Frobenius-adic expansion of d. In this paper, "Frobenius-adic expansion" means the expansion in [18].

3.1 The first ϕ-NAF (ϕ-GNAF)

At first, we show how to expand the multiplication by d map on $E(\mathbb{F}_{q^n})$ in terms of ϕ-GNAF and prove some properties of this. We begin with the definition of ϕ-GNAF for every subfield elliptic curves, and give an algorithm which calculates the ϕ-GNAF representation for a given $d \in \mathbb{Z}[\phi]$, where $q \geqslant 7$.

Definition 5 **[ϕ-GNAF]** *Let $E \in SEC_t[\mathbb{F}_q]$ and $d \in \mathbb{Z}[\phi]$. A ϕ-adic GNAF representation (ϕ-GNAF) of d on E is a representation $d = \sum_{i=0}^{\ell-1} e_i \phi^i$ where $e_i \in D_{q,q-1}$ for each i, $e_{\ell-1} \neq 0$, and one of the followings holds*
(1) $e_{i+1}e_i = 0$,
(2) *if* $e_{i+1}e_i > 0$, *then* $|e_{i+1} + e_i| < q$,
(3) *if* $e_{i+1}e_i < 0$, *then* $|e_{i+1}| > |e_i|$.

Let $a, b \in D_{q,q-1}$. If a, b satisfy one of the followings : (1) $ab = 0$, (2) if $ab > 0$, then $|a + b| < r$, (3) if $ab < 0$, then $|a| > |b|$, then we call a pair $(a, b)_\phi$ ϕ-admissible pair follow the lead of [22]. Otherwise, we call $(a, b)_\phi$ ϕ-non-admissible pair.

Algorithm 1: ϕ-GNAF for subfield elliptic curves with Frobenius trace 1 ($q \geqslant 7$)

Input: $d \in \mathbb{Z}[\phi]$
Output: ϕ-GNAF representation of d

1. Compute $d' = d \bmod (\phi^n - 1)$.
2. Compute $\ell = \lceil 2 \log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)} \rceil + 4$.
3. Compute Frobenius-adic expansion $(c_{\ell-1}, c_{\ell-2}, \cdots, c_1, c_0)$ of d'.
4. $b_0 \leftarrow c_0$, $b_1 \leftarrow c_1$, $b_\ell \leftarrow 0$, $b_{\ell+1} \leftarrow 0$.
5. $i \leftarrow 0$.
6. While $i \leqslant \ell$ do
   6.1. If (b_{i+1}, b_i) : ϕ-admissible pair then $b_{i+2} \leftarrow c_{i+2}$, $e_i \leftarrow b_i$.
   6.2. else if $b_i > 0$ then $b_{i+2} \leftarrow c_{i+2} - 1$, $b_{i+1} \leftarrow b_{i+1} + 1$, $e_i \leftarrow b_i - q$.
   6.3. else ($b_i < 0$) then $b_{i+2} \leftarrow c_{i+2} + 1$, $b_{i+1} \leftarrow b_{i+1} - 1$, $e_i \leftarrow b_i + q$.
   6.4. $i \leftarrow i + 1$.
7. Return $(e_{\ell+1}, e_\ell, \cdots, e_1, e_0)$.

The following lemma and theorem show the correctness of Algorithm 1, thus the existence of ϕ-GNAF. From the lemma, for any given Frobenius-adic expansion, we can have a sequence with digits in $D_{q,q-1}$ such that any adjacent digits are ϕ-admissible. Moreover, the theorem below gives the finiteness of the sequence (hence ϕ-GNAF) and evaluates the upper bound of the length of ϕ-GNAF. For the proof of the lemma, it is easily seen that the proof of Theorem 12.2.3 in [22] can be applied. For details, refer [22].

Lemma 1 *Let $b \in D_{q,(q+1)/2}$, $b', e \in D_{q,(q-1)/2}$. We assume that $(b', e)_\phi$ is a ϕ-admissible pair and $(b, b')_\phi$ is a ϕ-non-admissible pair. If we convert*

$$(b, b', e)_\phi \mapsto \begin{cases} (\bar{1}, c, c', e) := (\bar{1}, b+1, b'-q, e)_\phi & \text{if } b' > 0, \\ (1, c, c', e) := (1, b-1, b'+q, e)_\phi & \text{otherwise,} \end{cases}$$

then $(c, c')_\phi$, $(c', e)_\phi$ are ϕ-admissible pairs.

Theorem 3 [Finiteness of the length of ϕ-GNAF] *Let $d \in \mathbb{Z}[\phi]$ and ℓ be the length of Frobenius-adic expansion of d. Then d has a ϕ-GNAF representation with digit set $D_{q,q-1}$ such that the length is at most $\ell + 2$.*

Proof. Let ℓ' be the length of ϕ-GNAF representation of d. In process of Algorithm 1, we will have a representation $d = (c, b, b', e_{\ell-4}, \cdots, e_1, e_0)_\phi$, where $(e_{\ell-4}, \cdots, e_1, e_0)_\phi$ is a ϕ-GNAF. Now we scan the two digits $(b, b')_\phi$. If the pairs $(c, b)_\phi$ and $(b, b')_\phi$ are ϕ-admissible, there is nothing to do and $\ell' = \ell$. So we can assume that $(b, b')_\phi$ is ϕ-non-admissible. Then we convert

$$(b, b')_\phi \mapsto (c, c', e_{\ell-3})_\phi := \begin{cases} (c-1, b+1, b'-q)_\phi & \text{if } b' > 0, \\ (c+1, b-1, b'+q)_\phi & \text{otherwise.} \end{cases}$$

If $(c, c')_\phi$ is ϕ-admissible, then $\ell' = \ell$. Otherwise, we convert

$$(c, c')_\phi \mapsto (a, a', e_{\ell-2})_\phi := \begin{cases} (\bar{1}, c+1, c'-q)_\phi & \text{if } c' > 0, \\ (1, c-1, c'+q)_\phi & \text{otherwise,} \end{cases}$$

then $a = 1$ or $\bar{1}$. If $(a, a')_\phi$ is ϕ-admissible, then $\ell' = \ell + 1$. If $(a, a')_\phi$ is ϕ-non-admissible, then by the definition of GNAF, it is obvious that $a = 1, a' < 0$ or $a = \bar{1}, a' > 0$. In this case, we convert

$$(a, a')_\phi \mapsto \begin{cases} (1, 0, a'+q)_\phi & a' < 0, \\ (\bar{1}, 0, a'-q)_\phi & a' > 0. \end{cases}$$

Then $\ell' = \ell + 2$. Therefore the length of ϕ-GNAF representation of $d \in \mathbb{Z}[\phi]$ is at most 2 more than that of Frobenius-adic expansion of d. □

We can also extend Algorithm 1 in the case of $q = 3$ or 5. However, in this case, there is possibility that b_i is a multiple of q and $b_i \neq 0$. If b_i is a non-zero multiple of q, we convert $(b_{i+1}, b_i)_\phi \mapsto (-b_i/q, b_{i+1} + b_i/q, 0)_\phi$. It is easy to show that if b_i is a non-zero multiple of q, then it satisfies that $b_i = \pm q$. Thus for all i, we always have $|b_{i+1}| \leqslant (q+1)/2$, $|b_i| \leqslant (q+3)/2$, as Algorithm 1 (or Theorem 3). Remark that it does not occur $b_{i+1} \bmod q = 0$ except for $b_{i+1} = 0$. This shows the correctness of Algorithm 2.

Let ϕ-GNAF$_\ell$ be the set of ϕ-GNAF representations of length ℓ. We put $A_\ell = \#\phi\text{-GNAF}_\ell$, $S_\ell = \sum_{d \in \phi\text{-GNAF}_\ell} (\ell - w(d))$, and $C_\ell = \#\{d \in \phi\text{-GNAF}_\ell \mid w(d) = \ell\}$, where $w(d)$ means the Hamming weight of d. In other words, c_ℓ is the number of ϕ-GNAF with length ℓ such that all digits are non-zero. Then the non-zero density of ϕ-GNAF is defined by

$$1 - \lim_{\ell \to \infty} S_\ell / (\ell A_\ell).$$

ϕ-GNAF has properties same as GNAF. For details, refer the appendix.

Algorithm 2: ϕ-GNAF for subfield elliptic curves with Frobenius trace 1 ($q = 3$ or 5)

Input: $d \in \mathbb{Z}[\phi]$
Output: ϕ-GNAF representation of d

1. Compute $d' = d \bmod (\phi^n - 1)$.
2. Compute $\ell = \lceil 2\log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)}\rceil + 4$.
3. Compute Frobenius-adic expansion $(c_{\ell-1}, c_{\ell-2}, \cdots, c_1, c_0)$ of d'.
4. $b_0 \leftarrow c_0$, $b_1 \leftarrow c_1$, $b_\ell \leftarrow 0$, $b_{\ell+1} \leftarrow 0$.
5. $i \leftarrow 0$.
6. While $i \leqslant \ell$ do
   6.1. If $b_i \bmod q = 0$ then $b_{i+1} \leftarrow b_{i+1} - b_i/q$, $b_{i+2} \leftarrow c_{i+2} + b_i/q$, $e_i \leftarrow 0$.
   6.2. If (b_{i+1}, b_i) : ϕ-admissible pair then $b_{i+2} \leftarrow c_{i+2}$, $e_i \leftarrow b_i$.
   6.3. else if $b_i > 0$ then $b_{i+2} \leftarrow c_{i+2} - 1$, $b_{i+1} \leftarrow b_{i+1} + 1$, $e_i \leftarrow b_i - q$.
   6.4. else $(b_i < 0)$ then $b_{i+2} \leftarrow c_{i+2} + 1$, $b_{i+1} \leftarrow b_{i+1} - 1$, $e_i \leftarrow b_i + q$.
   6.5. $i \leftarrow i + 1$.
7. Return $(e_{\ell+1}, e_\ell, \cdots, e_1, e_0)$.

Proposition 2 [Properties of ϕ-GNAF]

(1) *Every $d \in \mathbb{Z}[\phi]$ has a unique ϕ-GNAF representation.*

(2) *The average number of non-zero digits for ℓ digits numbers in $\mathbb{Z}[\phi]$ is equal to $((q-1)/(q+1))\ell + 2/(q+1) + O((-1/q)^\ell)$. In particular, the average non-zero density among ϕ-GNAF representations of length ℓ is asymptotically $(q-1)/(q+1)$.*

The following algorithm calculates the ϕ-GNAF representation for a given $d \in \mathbb{Z}[\phi]$ without the calculation of Frobenius-adic expansion of d and reduces the memory consumption to calculates the ϕ-GNAF representation compared to Algorithm 1 and 2. From Theorem 3 (especially the finiteness of the length of ϕ-GNAF), it is easy to show the correctness of Algorithm 3.

Algorithm 3: ϕ-GNAF for subfield elliptic curves with Frobenius trace 1

Input: $d \in \mathbb{Z}[\phi]$
Output: ϕ-GNAF representation of d

1. Compute $d' := d_0 + d_1\phi = d \bmod (\phi^n - 1) (d_0, d_1 \in \mathbb{Z})$.
2. Compute $\ell = \lceil 2\log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)}\rceil + 4$.
3. Compute $Q, b \in \mathbb{Z}$ such that $d_0 = Qq + b$ $(b \in D_{q,(q-1)/2})$ (see [18]).
4. $d_0 \leftarrow Q + d_1$, $d_1 \leftarrow -Q$.
5. $i \leftarrow 1$.
6. While $i \leqslant \ell + 1$ do
   6.1. Compute $Q, a \in \mathbb{Z}$ such that $d_0 = Qq + a$ $(a \in D_{q,(q-1)/2})$.
   6.2. $d_0 \leftarrow Q + d_1$, $d_1 \leftarrow -Q$.
   6.3. If $(a, b)_\phi$: ϕ-admissible pair then $e_{i-1} \leftarrow b$, $b \leftarrow a$.
   6.4. else if $b > 0$ then $e_{i-1} \leftarrow b - q$, $b \leftarrow a + 1$, $d_0 \leftarrow d_0 - 1$.
   6.5. else $e_{i-1} \leftarrow b + q$, $b \leftarrow a - 1$, $d_0 \leftarrow d_0 + 1$.
7. Return $(e_{\ell+1}, e_\ell, \cdots, e_1, e_0)$.

3.2 The second ϕ-NAF (ϕ-rNAF)

Next, we show how to expand the multiplication by d map on $E(\mathbb{F}_{q^n})$ in terms of ϕ-adic rNAF and prove some properties of this. As with the previous section, we begin with the definition of ϕ-rNAF for every subfield elliptic curves, and give an algorithm which calculates the ϕ-rNAF for a given $d \in \mathbb{Z}[\phi]$, where $q \geqslant 7$.

Definition 6 [**ϕ-rNAF**] *Let $E \in SEC_t[\mathbb{F}_q]$ and $d \in \mathbb{Z}[\phi]$. A ϕ-adic rNAF representation (ϕ-rNAF) of d on E is a representation $d = \sum_{i=0}^{\ell-1} e_i\phi^i$ such that $e_i \in D_{q,(q^2-1)/2}$, $e_{\ell-1} \neq 0$ and $e_{i+1}e_i = 0$ for each i. Let $a, b \in D_{q,(q^2-1)/2}$. If $ab = 0$, we call $(a,b)_\phi$ ϕ-non-adjacent pair. Otherwise, we call $(a,b)_\phi$ ϕ-adjacent pair.*

Algorithm 4: ϕ-rNAF for subfield elliptic curves with Frobenius trace 1 ($q \geqslant 7$)

Input: $d \in \mathbb{Z}[\phi]$
Output: ϕ-rNAF representation of d

1. Compute $d' = d \bmod (\phi^n - 1)$.
2. Compute $\ell = \lceil 2\log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)}\rceil + 4$.
3. Compute Frobenius-adic expansion $(c_{\ell-1}, c_{\ell-2}, \cdots, c_1, c_0)$ of d'.
4. $b_0 \leftarrow c_0$, $b_1 \leftarrow c_1$, $b_\ell \leftarrow 0$, $b_{\ell+1} \leftarrow 0$, $b_{\ell+2} \leftarrow 0$, $b_{\ell+3} \leftarrow 0$.
5. $i \leftarrow 0$.
6. While $i \leqslant \ell + 2$ do
   6.1. If $b_i = 0$ then $b_{i+2} \leftarrow c_{i+2}$, $e_i \leftarrow 0$, $i \leftarrow i+1$.
   6.2. else if $|b_{i+1}q + b_i| \leqslant (q^2-1)/2$ then $b_{i+3} \leftarrow c_{i+3}$, $b_{i+2} \leftarrow c_{i+2} + b_{i+1}$, $e_{i+1} \leftarrow 0$, $e_i \leftarrow b_{i+1}q + b_i$, $i \leftarrow i+2$.
   6.3. else if $b_{i+1}q + b_i < -(q^2-1)/2$ then $b_{i+3} \leftarrow c_{i+3} + 1$, $b_{i+2} \leftarrow c_{i+2} + b_{i+1} + (q-1)$, $e_{i+1} \leftarrow 0$, $e_i \leftarrow b_{i+1}q + b_i + q^2$, $i \leftarrow i+2$.
   6.4. else $(b_{i+1}q + b_i > (q^2-1)/2)$ then $b_{i+3} \leftarrow c_{i+3} - 1$, $b_{i+2} \leftarrow c_{i+2} + b_{i+1} - (q-1)$, $e_{i+1} \leftarrow 0$, $e_i \leftarrow b_{i+1}q + b_i - q^2$, $i \leftarrow i+2$.
7. Return $(e_{\ell+3}, e_{\ell+2}, \cdots, e_1, e_0)$.

In the above algorithm, it does not occur the conversion :

$$(b_{i+1}, b_i)_\phi \mapsto \begin{cases} (\bar{b}'_i, b_{i+1} + b'_i, 0)_\phi & \text{if } b_i = b'_i q \text{ for some } b'_i \in \mathbb{Z}, \\ (\bar{b}'_{i+1}, b_{i+1}, 0, b_i)_\phi & \text{if } b_i + 1 = b'_{i+1}q \text{ for some } b'_{i+1} \in \mathbb{Z}. \end{cases}$$

In fact, there is no possibility such that $b_i \bmod q \equiv 0, b_i \neq 0$ or $b_{i+1} \bmod q \equiv 0, b_{i+1} \neq 0$ when we scan $(b_{i+1}, b_i)_\phi$.

The lemma and theorem below show the correctness of Algorithm 4, thus the existence of ϕ-rNAF. For the proof of the lemma, consult [21].

Lemma 2 *Let $c, c', c'' \in D_{q,(q-1)/2}$, $b \in D_{q,(q+1)/2}$, $b' \in D_{q,2q-1}$. We convert $(c, c', c'', b, b')_\phi$ from right-to-left according to the following rule and we denote the result of the conversion $(a, a', e, e', e'', e''')_\phi$.*
The rule: We assume that we scan consecutive two digits $(a, b)_\phi$, then

(Rule 1) If $a \neq 0, b \neq 0$, then convert $(a,b)_\phi$

$$(a,b)_\phi \mapsto \begin{cases} (a,0,aq+b)_\phi & \textit{if } |aq+b| \leqslant (q^2-1)/2, \\ (1, a+(q-1), 0, (aq+b)+q^2)_\phi & \textit{else if } aq+b < -(q^2-1)/2, \\ (\bar{1}, a-(q-1), 0, (aq+b)-q^2)_\phi & \textit{otherwise.} \end{cases}$$

(Rule 2) If $a \neq 0, b = 0$, then skip the 1-digit b. We scan the next consecutive two digits which include a.
(Rule 3) if $a = 0, b = 0$, then skip the 2-digits a and b. We scan the next consecutive two digits which do not include a.

Then, it always satisfy that $a \in D_{q,(q+1)/2}$, $a' \in D_{q,2q-1}$ and $a' \not\equiv 0 \bmod q$ except for $a' = 0$.

Theorem 4 [Finiteness of the length of ϕ-rNAF] *Let $d \in \mathbb{Z}[\phi]$ and ℓ be the length of Frobenius-adic expansion of d. Then d has a ϕ-rNAF representation with digit set $D_{q,(q^2-1)/2}$ such that the lenght is at most $\ell+4$.*

Proof. Let ℓ' be the length of ϕ-rNAF representation of d. As in the proof of Theorem 3, in process of Algorithm 4, it is easily seen that we will have two cases as follows.

(Case 1) $d = (b, b', e_{\ell-3}, e_{\ell-4}, \cdots, e_1, e_0)_\phi$, where $(e_{\ell-3}, \cdots, e_0)_\phi$ is a ϕ-rNAF. Now we scan the two digits $(b,b')_\phi$. We can assume that $e_{\ell-3} = 0$ because if $e_{\ell-3} \neq 0$, It will be able to reduce this case to the case 2 (See below the details). If $(b,b')_\phi$ is the ϕ-non-adjacent pair, then $\ell' = \ell$. Otherwise, we convert the pair $(b,b')_\phi$ in the following :

$$(b,b')_\phi \mapsto \begin{cases} (b,0,bq+b')_\phi & |bq+b'| \leqslant (q^2-1)/2, \\ (1, b+(q-1), 0, bq+b'+q^2)_\phi & bq+b' < -(q^2-1)/2, \\ (\bar{1}, b-(q-1), 0, bq+b'-q^2)_\phi & bq+b' > (q^2-1)/2, \end{cases}$$

and we put the result $(c, c', e_{\ell-1}, e_{\ell-2})_\phi$. If $|bq+b'| \leqslant (q^2-1)/2$, then $\ell' \leqslant \ell+1$. Otherwise, since $|b| \leqslant (q+1)/2$,

$$|cq+c'| = |\pm q \pm (q-1) + b| = |\pm(2q-1) + b| \leqslant (q^2-1)/2,$$

when it satisfies that $q \geqslant 5$. We can convert $(1, b+(q-1))_\phi \mapsto (1, 0, b+(2q-1))_\phi$ or $(\bar{1}, b-(q-1))_\phi \mapsto (\bar{1}, 0, b-(2q-1))_\phi$, thus we have $\ell' = \ell+2$.

(Case 2) $d = (c, b, b', e_{\ell-4}, \cdots, e_1, e_0)_\phi$, where $(e_{\ell-4}, \cdots, e_1, e_0)_\phi$ is a ϕ-rNAF. Now we scan the two digits $(b,b')_\phi$. We can assume that $b' \neq 0$ because if $b' = 0$, we can reduce this case to the case 1. Note that $e_{\ell-4} = 0$. If $b = 0$, there is nothing to do and $\ell' = \ell$. So we can also assume $b \neq 0$. Then we convert

$$(b,b')_\phi \mapsto \begin{cases} (c+b,0,bq+b')_\phi & |bq+b'| \leqslant (q^2-1)/2, \\ (1, c+b+(q-1), 0, bq+b'+q^2)_\phi & bq+b' < -(q^2-1)/2, \\ (\bar{1}, c+b-(q-1), 0, bq+b'-q^2)_\phi & bq+b' > (q^2-1)/2, \end{cases}$$

and we put the result $(a, a', e_{\ell-2}, e_{\ell-3})_\phi$. If $|bq+b'| \leqslant (q^2-1)/2$, there is nothing to do and $\ell' \leqslant \ell$. Otherwise, the most significant two digits are $(1, (q-1)+b+c)_\phi$ or $(\bar{1}, -(q-1)+b+c)_\phi$, we convert

$$|aq + a'| = |\pm q \pm (q-1) + b + c| = |\pm (2q-1) + b + c| \leqslant (q^2-1)/2,$$

when it satisfies $q \geqslant 9$. We can convert

$$\begin{aligned}(1, (q-1)+b+c)_\phi &\mapsto (1, 0, (2q-1)+b+c)_\phi, \\ (\bar{1}, -(q-1)+b+c)_\phi &\mapsto (\bar{1}, 0, -(2q-1)+b+c)_\phi.\end{aligned}$$

Then we have $\ell' = \ell + 2$ when it satisfies $q \geqslant 9$. If it satisfy that $q = 7$ and $|bq+b'| > (q^2-1)/2$, then we convert

$$\begin{aligned}(1, b+c+(q-1))_\phi &\mapsto (\bar{1}, -(q-2), 0, b+c-(q-1)^2)_\phi, \\ (\bar{1}, b+c-(q-1))_\phi &\mapsto (1, q-2, 0, b+c+(q-1)^2)_\phi.\end{aligned}$$

Note that because of Lemma 2, the following situations do not occur :

$$\begin{aligned}(1, b+c+(q-1))_\phi &\mapsto (1, q, 0, b+c+(q-1)^2)_\phi, \\ (\bar{1}, b+c-(q-1))_\phi &\mapsto (\bar{1}, \bar{q}, 0, b+c-(q-1)^2)_\phi.\end{aligned}$$

Here, $\pm q \pm q - 2 = \pm 2(q-1)$, so $|\pm 2(q-1)| \leqslant (q^2-1)/2$. Then $\ell' = \ell + 4$. Therefore the length of ϕ-rNAF representation of $d \in \mathbb{Z}[\phi]$ is at most 4 more than that of Frobenius-adic expansion of d. □

We can also extend Algorithm 4 to the case of $q = 3$ or 5. However, it does not satisfy Lemma 2 in this case, namely there are possibility that b_i is a multiple of q and $b_i \neq 0$. If b_i is a non-zero multiple of q, we convert

$$(b_{i+1}, b_i)_\phi \mapsto (-b_i/q, b_{i+1} + b_i/q, 0)_\phi.$$

It is easy to show that if b_i is a non-zero multiple of q, then it satisfies that $b_i = \pm q$. Thus for all i, we always have

$$|b_{i+1}| \leqslant (q-1)/2, |b_i| \leqslant q-1 \quad \text{or} \quad |b_{i+1}| \leqslant (q+1)/2, |b_i| \leqslant 2q-1 \ (|b_i| \neq q),$$

as Algorithm 4 (or Theorem 4). Note that it does not occur $b_{i+1} \bmod q = 0$ except for $b_{i+1} = 0$. This shows the correctness of Algorithm 5.

Algorithm 5: ϕ-rNAF for subfield elliptic curves with Frobenius trace 1 ($q = 3$ or 5)

Input: $d \in \mathbb{Z}[\phi]$
Output: ϕ-rNAF representation of d

1. Compute $d' = d \bmod (\phi^n - 1)$.
2. Compute $\ell = \lceil 2 \log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)} \rceil + 4$.
3. Compute Frobenius-adic expansion $(c_{\ell-1}, c_{\ell-2}, \cdots, c_1, c_0)$ of d'.
4. $b_0 \leftarrow c_0$, $b_1 \leftarrow c_1$, $b_\ell \leftarrow 0$, $b_{\ell+1} \leftarrow 0$, $b_{\ell+2} \leftarrow 0$, $b_{\ell+3} \leftarrow 0$.
5. $i \leftarrow 0$.
6. While $i \leqslant \ell + 2$ do
   6.1. If $b_i \bmod q = 0$ then $b_{i+1} \leftarrow b_{i+1} - b_i/q$, $b_{i+2} \leftarrow c_{i+2} + b_i/q$, $i \leftarrow i + 1$.
   6.2. If $b_i = 0$ then $e_i \leftarrow 0$, $b_{i+2} \leftarrow c_{i+2}$, $i \leftarrow i + 1$.
   6.3. else if $|b_{i+1}q + b_i| \leqslant (q^2 - 1)/2$ then $b_{i+3} \leftarrow c_{i+3}$, $b_{i+2} \leftarrow c_{i+2} + b_{i+1}$, $e_{i+1} \leftarrow 0$, $e_i \leftarrow b_{i+1}q + b_i$, $i \leftarrow i + 2$.
   6.4. else $b_{i+1}q + b_i < -(q^2 - 1)/2$ then $b_{i+3} \leftarrow c_{i+3} + 1$, $b_{i+2} \leftarrow c_{i+2} + b_{i+1} + (q - 1)$, $e_{i+1} \leftarrow 0$, $e_i \leftarrow b_{i+1}q + b_i + q^2$, $i \leftarrow i + 2$.
   6.5. else $b_{i+1}q + b_i > (q^2 - 1)/2$ then $b_{i+3} \leftarrow c_{i+3} - 1$, $b_{i+2} \leftarrow c_{i+2} + b_{i+1} - (q - 1)$, $e_{i+1} \leftarrow 0$, $e_i \leftarrow b_{i+1}q + b_i - q^2$, $i \leftarrow i + 2$.
7. Return $(e_{\ell+3}, e_{\ell+2}, \cdots, e_1, e_0)$.

Let ϕ-rNAF$_\ell$ be the set of ϕ-rNAF representation of the length ℓ. We put $B_\ell = \#\phi\text{-}r\mathrm{NAF}_\ell$, $T_\ell = \sum_{d \in \phi\text{-}r\mathrm{NAF}_\ell} (\ell - w(d))$, where $w(d)$ means the Hamming weight of d. Then as is the case with ϕ-GNAF, the non-zero density of ϕ-rNAF is defined by $1 - \lim_{\ell \to \infty} T_\ell/(\ell B_\ell)$.

For ϕ-rNAF, we also have similar properties. For the proof of the following proposition, refer the appendix.

Proposition 3 [Properties of ϕ-rNAF]
(1) *Every $d \in \mathbb{Z}[\phi]$ has a unique ϕ-rNAF representation.*
(2) *The average number of non-zero digits for ℓ digits numbers in $\mathbb{Z}[\phi]$ is equal to $((q-1)/(2q-1))\ell + q/(2q-1) + O(((1-q)/q)^\ell)$. In particular, the average non-zero density among ϕ-rNAF representations of length ℓ is asymptotically $(q-1)/(2q-1)$.*

The following algorithm calculates the ϕ-rNAF representation for a given $d \in \mathbb{Z}[\phi]$ without the calculation of Frobenius-adic expansion of d and reduces the memory consumption to calculates the ϕ-rNAF representation compared to Algorithm 4 and 5. From Theorem 4 (especially the finiteness of the length of ϕ-rNAF), it is easy to show the correctness of Algorithm 6.

Algorithm 6: ϕ-rNAF for subfield elliptic curves with Frobenius trace 1

Input: $d \in \mathbb{Z}[\phi]$
Output: ϕ-rNAF representation of d

1. Compute $d' := d_0 + d_1\phi = d \bmod (\phi^n - 1)(d_0, d_1 \in \mathbb{Z})$.
2. Compute $\ell = \lceil 2\log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d)}\rceil + 4$.
3. $i \leftarrow 0$.
4. While $i \leqslant \ell + 2$ do
   4.1. Compute $Q, b \in \mathbb{Z}$ such that $d_0 = Qq + b$ ($b \in D_{q,(q-1)/2}$) (see [18]).
   4.2. $d_0 \leftarrow Q + d_1$, $d_1 \leftarrow -Q$.
   4.3. If $b = 0$ then $e_i \leftarrow 0$, $i \leftarrow i + 1$.
   4.4. else compute $Q, a \in \mathbb{Z}$ such that $d_0 = Qq + a$ ($a \in D_{q,(q-1)/2}$).
   $d_0 \leftarrow Q + d_1$, $d_1 \leftarrow -Q$.
   if $|aq + b| \leqslant (q^2 - 1)/2$ then $e_i \leftarrow aq + b$, $e_{i+1} \leftarrow 0, d_0 \leftarrow a, d_1 \leftarrow -Q$.
   else if $aq + b < -(q^2 - 1)/2$ then $e_i \leftarrow aq + b + q^2$,
   $e_{i+1} \leftarrow 0, d_0 \leftarrow a + (q - 1), d_1 \leftarrow -Q + 1$.
   else $e_i \leftarrow aq + b - q^2$, $e_{i+1} \leftarrow 0, d_0 \leftarrow a - (q - 1), d_1 \leftarrow -Q - 1$.
   $i \leftarrow i + 2$.
5. Return $(e_{\ell+3}, e_{\ell+2}, \cdots, e_1, e_0)$.

Remark. In this paper, we do not discuss about the minimality of Hamming weight of ϕ-GNAF and ϕ-rNAF among various recoding methods with appropriate digit sets. Although the property of minimality is desired, we can easily seen that for ϕ-GNAF, ϕ-rNAF, conventional proofs (for e.g., GNAF, rNAF, etc.) are not available. It can be considerd that these flaws are caused by the difference between the rational integer ring and quadratic imaginary integer rings or quadratic order, and we will need some deep observation on number theoretical properties of quadratic integer rings. These issues remain to be discussed.

4 Comparisons

We compare several recoding methods for computing scalar multiplication for a point on a subfield elliptic curve with the Frobenius trace 1 using standard Left-to-right method (for details, refer [8]). Let d be a large positive integer and we focus on the group of $\mathbb{F}_{q^n}$-rational points $E(\mathbb{F}_{q^n})$ for sufficient large n which satisfy $d \approx q^n$. Let m_q, ℓ be the length of q-adic expansion of d, the length of ϕ-adic expansion of $d_0 = d \bmod (\phi^n - 1)$, respectively.

As $d \approx q^n$, the norm of d will be equal to $d^2 \approx q^{2n}$ and $d_0 \approx q^{n+1}$ (for detail, refer [18]). So $m = \lfloor \log_q d \rfloor + 1 \approx \lfloor \log_q q^n \rfloor + 1 = n + 1$ and $\ell \leqslant \lceil 2\log_q 2\sqrt{N_{\mathbb{Z}[\phi]/\mathbb{Z}}(d_0)}\rceil + 4 \approx \lceil 2\log_q 2q^{(n+1)/2}\rceil + 4 \leqslant n + 6$. For simplify of evaluation of computational cost, we assume that the length of q-adic expansion of d, the length of ϕ-adic expansion of d_0, the length of ϕ-GNAF of d_0 and the length of ϕ-rNAF of d_0 are equal to each other (Of course, we should analyze each average of the length of ϕ-adic expansion, ϕ-GNAF and ϕ-rNAF among positive integers in the range $[1, \#E - 1]$ to evaluate the exact computational costs. However, we do not deal with this analysis). In practical meaning, the shift operations are essentially free, thus if we use normal basis for $\mathbb{F}_{q^n}$, then the cost of Frobenius map on subfield elliptic curves are free.

In the below tables, ECADD, ECDBL, and ECFRB stand for the computational cost of the point addition, point doubling, and Frobenius map, respectively. Note that when we compute point multiplication by q for ordinary GNAF or rNAF, we use $-(\phi^2 - \phi)$, i.e., the computational cost of point multiplication by q just one time is ECADD+2ECFRB.

Method	#Table	ECADD	ECDBL	ECFRB
w-NAF [20]	2^{w-2}	$(2^{w-2}-1)+m_2/(w+1)$	m_2+1	0
ϕ-adic expansion [18]	1	0.67ℓ	0	ℓ
ord. GNAF [4]	2	1.5ℓ	1	2ℓ
$\boldsymbol{\phi}$**−GNAF**	**2**	$\mathbf{0.5}\boldsymbol{\ell}$	**1**	$\boldsymbol{\ell}$
ord. rNAF [21]	3	1.4ℓ	2	2ℓ
$\boldsymbol{\phi - r}$**NAF**	**3**	$\mathbf{0.4}\boldsymbol{\ell}$	**2**	$\boldsymbol{\ell}$

Table 1. Computational cost for each recoding method ($q = 3$)

Method	#Table	ECADD	ECDBL	ECFRB
w-NAF [20]	2^{w-2}	$(2^{w-2}-1)+m_2/(w+1)$	m_2+1	0
ϕ-adic expansion [18]	$(q-1)/2$	$((q-1)/q)\ell+(q-5)/2$	1	ℓ
ord. GNAF [4]	$q-1$	$(2q/(q+1))\ell+(q-3)$	1	2ℓ
$\boldsymbol{\phi}$**−GNAF**	$\boldsymbol{q-1}$	$\boldsymbol{((q-1)/(q+1))\ell+(q-3)}$	**1**	$\boldsymbol{\ell}$
ord. rNAF [21]	$q(q-1)/2$	$(3q/(2q-1))\ell+(q^2-q-4)/2$	1	2ℓ
$\boldsymbol{\phi - r}$**NAF**	$\boldsymbol{q(q-1)/2}$	$\boldsymbol{((q-1)/(2q-1))\ell+(q^2-q-4)/2}$	**1**	$\boldsymbol{\ell}$

Table 2. Computational cost for each recoding method ($q \geqslant 5$)

In the second column, the value #Table equals the number of elements, that have to be precomputed and stored.

We assume that $q = 3, 5, 7$. The above table shows that the throughputs of [18], [4] are improved by 12.5∼25%, 57∼60%, respectively, using ϕ-GNAF and the throughputs of [18], [21] are improved by 40∼46%, 31.5∼71%, respectively, using ϕ-rNAF on the above assumption.

Moreover we assume that the bit length of $d \approx 200$ and the width $w = 6$. Then $n = 126$ $(q = 3), 86$ $(q = 5), 71$ $(q = 7)$, respectively. In this case, computational cost for each recoding method is the following.

The above table shows that the throughputs of w-NAF are improved by 75% using ϕ-GNAF and 79% using ϕ-rNAF.

5 Conclusion

It has been an unsolved problem to generalize τ-NAF techniques for Koblitz curves to more general family of subfield elliptic curves whose endomorphism rings are not necessarily subrings of Euclidean quadratic imaginary number fields. In this paper, we have described two generalized methods on a family

Method	#Table	ECADD	ECDBL	ECFRB
w-NAF [20]	16	44	201	0
ϕ-GNAF ($q = 3$)	2	63	1	126
ϕ-GNAF ($q = 5$)	4	59	1	86
ϕ-GNAF ($q = 7$)	6	57	1	71
ϕ-rNAF ($q = 3$)	3	50	2	126
ϕ-rNAF ($q = 5$)	10	46	1	86
ϕ-rNAF ($q = 7$)	21	52	1	71

Table 3. Computational cost for each recoding method ($w = 6$, $m_2 = 200$)

of subfield elliptic curves. Those methods are two classes of ϕ-NAF (ϕ-GNAF and ϕ-rNAF). Our proposed methods can be applied to every subfield elliptic curves with Frobenius trace 1 regardless of whether or not the endomorphism rings are Euclidean. We also prove that these representations have the same non-zero densities as the corresponding original GNAF and rNAF. Because of the high efficinecy in computing Frobenius maps, our proposed methods improve the efficiency of scalar multiplication significantly compared to previous methods. The speed of the proposed schemes improve between 12.5% and 79% over that for previously known schemes.

References

1. R.M. Avanzi, C. Heuberger, and H. Prodinger, "*Minimality of the Hamming Weight of the τ-NAF for Koblitz Curves and Improved Combination with Point Halving*," Cryptology ePrint Archive : Report 2005/225, 2005. available from : `http://eprint.iacr.org/2005/225.pdf`
2. I.F. Blake, V.K. Murty and G. Xu, "*Nonadjacent radix-τ Expansions of Integers in Euclidean Imaginary Quadratic Number Fields*," GANITA LABORATORY, November, 2004. available from : `http://www.erin.utoronto.ca/~w3ganita/radix_t.pdf`
3. I. Blake, G. Seroussi, and N.P. Smart, "*Elliptic Curves in Cryptography*," Cambridge University Press, 1999.
4. W.E. Clark and J.J. Liang, "On arithmetic weight for a general radix representation of integers," *IEEE Transactions on Information Theory*, IT-19, pp.823-826, 1973.
5. H. Cohen, "Analysis of the Sliding Window Powering Algorithm," *Journal of Cryptology 18*, pp.63-76, 2005.
6. C. Günther, T. Lange and A. Stein, "Speeding up the Arithmetic on Koblitz Curves of Genus Two," *Selected Areas in Cryptography - SAC 2001*, *LNCS* 2012, pp.106-117, 2001.
7. R. Gallant, R. Lambert, and S. Vanstone, "Faster Point Multiplication on Elliptic Curves with Efficient Endomorphisms," *Advances in Cryptology - CRYPTO 2001*, *LNCS* 2139, pp.190-200, 2001.
8. D. Hankerson, A. Menezes, and S. Vanstone, "*Guide to Elliptic Curve Cryptography*," Springer-Verlag, 2004.
9. N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation 48*, pp.203-209, 1987.

10. N. Koblitz, "CM-curves with good cryptographic properties," *Advances in Cryptology - CRYPTO 1991*, *LNCS* 576, pp.279-287, 1992.
11. N. Koblitz, "An elliptic curve implementation of the finite field digital signature algorithm," *Advances in Cryptology - CRYPTO 1998*, *LNCS* 1462, pp.327-337, 1998.
12. T. Lange, "*Efficient Arithmetic on Hyperelliptic Koblitz Curves*," Ph.D. thesis, University of Essen, 2001.
13. V. Miller, "Uses of elliptic curves in cryptography," *Advances in Cryptology - CRYPTO 1985*, *LNCS* 218, pp.417-426, 1986.
14. T.J. Park, M.K. Lee and K. Park, "New Frobenius Expansions for Elliptic Curves with Efficient Endomorphisms," *International Conference on Information Security and Cryptology - ICISC 2002*, *LNCS* 2587, pp.264-282, 2003.
15. T. Satoh and K. Araki, "Fermat quotients and the polynomial time discrete log algorithm for anomalous elliptic curves," *Commentarii Mathematici Universitatis Sancti Pauli, 47*, pp.81-92, 1998.
16. I.A. Semaev, "Evaluation of discrete logarithms on some elliptic curves," *Mathematics of Computation 67*, pp.353-356, 1998.
17. J.H. Silverman, "*The Arithmetic of Elliptic Curves*," GTM 106, Springer-Verlag, 1986.
18. N.P. Smart, "Elliptic curve cryptosystems over small fields of odd characteristic," *Journal of Cryptology 12*, pp.141-151, 1999.
19. N.P. Smart, "The discrete logarithm problem on elliptic curves of trace one," *Journal of Cryptology 12*, pp.193-196, 1999.
20. J.A. Solinas, "Efficient Arithmetic on Koblitz curves," *Designs, Codes and Cryptography 19*, pp.195-249, 2000.
21. T. Takagi, S.M. Yen and B.C. Wu, "Radix-r Non-adjacent Form," *Information Security Conference - ISC 2004*, *LNCS* 3225, pp.99-110, 2004.
22. J.H. van Lint, "*Introduction to coding theory*," GTM 86, Springer-Verlag, 1982.

A Several Proofs

In this section we prove several propositions and lemmas described in this paper. Here, we review the following Lemma to prove the uniqueness of ϕ-GNAF and ϕ-rNAF. For detail, refer [12].

Lemma 3 *Let $\alpha = a + b\phi \in \mathbb{Z}[\phi]$ $(a, b \in \mathbb{Z})$, then we have*

$$\phi | \alpha \;\Leftrightarrow\; q | a.$$

In particular, for rational integer $a \in \mathbb{Z}$, we have $\phi | a \Leftrightarrow q | a$.

Proof of Proposition 2.

(1) We suppose that $d \in \mathbb{Z}[\phi]$ has two such representations

$$d = \sum e_i \phi^i = \sum e'_i \phi^i.$$

If i_0 is the smallest number such that $e_i = e'_i$ for $0 \leqslant i \leqslant i_0 - 1$ and $e_{i_0} \neq e'_{i_0}$, then we can replace d by $(d - \sum_{i=0}^{i_0-1} e_i\phi^i)/\phi^{i_0} = \sum_{i=i_0}^{\ell-1} e_i\phi^{i-i_0}$, and we have representations

$$d = \sum_{i=0}^{\ell-1} e_i\phi^i = \sum_{i=0}^{\ell-1} e'_i\phi^i, \quad e_0 \neq e'_0,$$

where ℓ is the maximum of length of the two representations. Then from Lemma 3, we have $q|(e_0 - e'_0)$, and by the assumption, $|e_0|, |e'_0| \leqslant q-1$, hence it satisfies that $|e_0 - e'_0| \leqslant |e_0| + |e'_0| \leqslant 2q-2$. Therefore we have $e_0 - e'_0 = \pm q$. In the following, we will give a proof for the case $e_0 - e'_0 = q$. For another case, we can prove similarly, thus we omit it.

We have $q = (e_0 - e'_0) = \sum_{i=1}^{\ell-1}(e'_i - e_i)\phi^i$ and by the characteristic polynomial of ϕ, we have $\phi - \phi^2 = \sum_{i=1}^{\ell-1}(e'_i - e_i)\phi^i$. Thus we can see

$$(e_1 - e'_1 + 1) = (e'_2 - e_2 + 1)\phi + \sum_{i=3}^{\ell-1}(e'_i - e_i)\phi^{i-1}.$$

Hence, from Lemma 3, we also have $q|(e_1 - e'_1 + 1)$. It follows that $e'_1 \in \{e_1 + 1 - q, e_1 + 1, e_1 + 1 + q\}$. Then, by the same way as [22], we can easily induce the contradiction (For details, refer [22], pp.177).

(2) At first, we show that $C_{\ell+1} = (q-2)C_\ell$, $C_1 = 2(q-1)$. We fix an element $d = (e_{\ell-1}, \cdots, e_0)_\phi$ of $\mathbb{Z}[\phi]$ which is a ϕ-GNAF representation such that each $e_i \neq 0$. For the fixed d, we count the number of e which satisfy $0 < |e| < q-1$, and $(e_{\ell-1}, \cdots, e_0, e)_\phi$ is a ϕ-GNAF representation. We can assume that $e_0 > 0$. If $e_0 e > 0$, e satisfy $0 < e < q - 1 - e_0$ and otherwise $0 < e < e_0$. Thus we have that the number of e is $q-2$, hence the recurrence equation for C_ℓ. It is trivial that $C_1 = 2(q-1)$.

Next, we show that A_ℓ and S_ℓ satisfy the recurrence equations $A_1 = 2(q-1)$, $A_2 = 2(q-1)^2$, $A_{\ell+2} - (q-1)A_{\ell+1} - qA_\ell = 0$ $(\ell \geq 1)$, and $S_1 = 0$, $S_2 = 2(q-1)$,

$$S_{\ell+2} - (q-1)S_{\ell+1} - qS_\ell = 2(q-1)\left(1 + \frac{2(q-1)}{(q+1)}\left(\frac{q(q^\ell-1)}{q-1} - \frac{(-1)^\ell - 1}{2}\right)\right).$$

In Table 4, the symbol '$*$' means a non-zero digit. We explain how to see Table 4. The left column stand for forms of ℓ digits ϕ-GNAF representation, the center column stand for the number of each ϕ-GNAF representation correspond to each form of the left column, and the right column stand for each number of 0 digits correspond to each form of the left column. From Table 4, we can derive

$$\begin{aligned}
A_\ell &= \sum_{j=1}^{\ell-2} C_j \left(\sum_{i=1}^{\ell-1-j} A_i + 1\right) + C_{\ell-1} + C_\ell, \\
S_\ell &= \sum_{j=1}^{\ell-2} C_j \left(\sum_{i=1}^{\ell-1-j} S_i + (\ell-i-j)A_i\right) + \sum_{i=1}^{\ell-1}(\ell-i)C_i.
\end{aligned}$$

forms	# of representations	# of 0 digits
$(*0* \quad \cdots \quad)$	$C_1A_{\ell-2}$	$C_1A_{\ell-2}+C_1S_{\ell-2}$
$(*00* \quad \cdots \quad)$	$C_1A_{\ell-3}$	$2C_1A_{\ell-3}+C_1S_{\ell-3}$
$(*000* \quad \cdots \quad)$	$C_1A_{\ell-4}$	$3C_1A_{\ell-4}+C_1S_{\ell-4}$
$\vdots$	$\vdots$	$\vdots$
$(*0000 \ \cdots 0*)$	C_1A_2	$(\ell-3)C_1A_2+C_1S_2$
$(*0000 \ \cdots 00*)$	C_1A_1	$(\ell-2)C_1A_1+C_1S_1$
$(*0000 \ \cdots 000)$	C_1	$(\ell-1)C_1$
$(**0* \quad \cdots \quad)$	$C_2A_{\ell-3}$	$C_2A_{\ell-3}+C_2S_{\ell-3}$
$(**00* \quad \cdots \quad)$	$C_2A_{\ell-4}$	$2C_2A_{\ell-4}+C_2S_{\ell-4}$
$(**000* \cdots \quad)$	$C_2A_{\ell-5}$	$3C_2A_{\ell-5}+C_2S_{\ell-5}$
$\vdots$	$\vdots$	$\vdots$
$(**000 \ \cdots 0*)$	C_2A_2	$(\ell-4)C_2A_2+C_2S_2$
$(**000 \ \cdots 00*)$	C_2A_1	$(\ell-3)C_2A_1+C_2S_1$
$(**000 \ \cdots 000)$	C_2	$(\ell-2)C_2$
$\vdots$	$\vdots$	$\vdots$
$(*** \quad \cdots *0*)$	$C_{\ell-2}A_1$	$C_{\ell-2}A_1+C_{\ell-2}S_1$
$(*** \quad \cdots *00)$	$C_{\ell-2}$	$2C_{\ell-2}$
$(*** \quad \cdots **0)$	$C_{\ell-1}$	$C_{\ell-1}$
$(*** \quad \cdots ***)$	C_ℓ	0

Table 4. Forms of ℓ digits ϕ-GNAF representations

From these equations, we have the desired recurrence equations, and by elementary calculations, we have $A_\ell = 2(q-1)\left(q^\ell - (-1)^\ell\right)/(q+1)$, and

$$\begin{aligned}S_\ell = {} & \frac{2(q-1)}{q+1}\left(\frac{q^{\ell-1}-1}{q-1}+\frac{(-1)^{\ell-1}-1}{2}\right)+\frac{4(q-1)^2}{(q+1)^2}(\ell-1)\left(\frac{q^\ell}{q-1}-\frac{(-1)^\ell}{2}\right)\\ & -\frac{4(q-1)^2}{(q+1)^2}\left(\frac{q(q^{\ell-1}-1)}{(q-1)^2}+\frac{(-1)^{\ell-1}-1}{4}\right)+\frac{q^{\ell-1}-(-1)^{\ell-1}}{(q+1)^2}\\ & +\frac{(-1)^\ell(q-1)(1-(-q)^{\ell-1})+2((-1)^{\ell-1}-q^{\ell-1})}{(q+1)^3}.\end{aligned}$$

Thus $\lim_{\ell\to\infty} S_\ell/(\ell A_\ell) = 2/(q+1)$. Hence we have the desired result. □

Proof of Lemma 2.

We apply the rule in the statement of Lemma 2 to $(c, c', c'', b, b')_\phi$. We denote the conversion of i-th digit and $i+1$-th digit by $\flat$ and the conversion of $i+2$-th digit and $i+3$-th digit by $\sharp$. We assume $a' \equiv 0 \bmod q, a' \neq 0$. From $|a'| \leqslant 2q-2$, we have $a' = q$ or $a' = -q$. It suffices to consider the case of $a' = q$ (In the same way as the case of $a' = q$, we can also the case of $a' = -q$). It is easy to understand the following : the situations which have the possibility of $a' \equiv 0 \bmod q, a' \neq 0$ is

$(\sharp, \flat) = ((1),(2)),((2),(1)),((2),(2)),((3),(2))$, where (1) means (Rule 1) in the statement of Lemma 2 and so on. From now on, we investigate each situation.

(Case 1) $(\sharp, \flat) = ((1),(2))$. It is easy to show that $a' = c + c' + (q-1)$, i.e. $c' = -c+1$. We apply the rule to $(c, c', c'', b, b')_\phi$, We have $e' = (-c+1)q + c'' + b + q^2$ and it is easy to show that $e' > (q^2-1)/2$. This is contrary to $|e'| \leqslant (q^2-1)/2$.

(Case 2) $(\sharp, \flat) = ((2),(1))$. It is easy to show that $a' = c + c' + 1$, i.e. $c = c' = (q-1)/2$. We apply the rule to $(c, c', c'', b, b')_\phi$, We have $e''' = bq + b' + q^2$ and $e' = c'' + b + (q^2+3q-2)/2$. It is easy to show that $b = -(q-3)/2, -(q-1)/2, -(q+1)/2$. It is also easy to see that $b \leqslant -q$. This is contrary to $|b| \leqslant (q+1)/2$.

(Case 3) $(\sharp, \flat) = ((2),(2))$. It is easy to show that $a' = c + c' + q$, i.e. $c' = -c$. We apply the rule to $(c, c', c'', b, b')_\phi$, We have $e''' = bq + b' + q^2$ and $e' = (-c+1)q + c'' + b + (q-1) + q^2$. It is easy to show that $b = -(q-3)/2, -(q-1)/2, -(q+1)/2$. It is also easy to see that $-cq + c'' < -(q^2-1)/2$. This is contrary to the ranges of c'', c.

(Case 4) $(\sharp, \flat) = ((3),(2))$. It is easy to show that $a' = c + c' + q - 2$, i.e. $c' = -c + 2$. We apply the rule to $(c, c', c'', b, b')_\phi$, We have $e''' = bq + b' - q^2$ and $e' = (-c+1)q + c'' + b - (q-1) + q^2$. It is easy to show that $b = (q-3)/2, (q-1)/2, (q+1)/2$. It is also easy to see that $-cq + c'' < -(q^2-1)/2$. This is contrary to the ranges of c'', c.

Therefore it always satisfy that $a' \not\equiv 0 \bmod q$ except for $a' = 0$. □

Proof of Proposition 3.

(1) By the same way as Proposition 2, we suppose $d \in \mathbb{Z}[\phi]$ has two such representations

$$d = \sum_{i=0}^{\ell-1} e_i \phi^i = \sum_{i=0}^{\ell-1} e'_i \phi^i, \quad e_0 \neq e'_0.$$

Then from Lemma 3, we have $q | (e_0 - e'_0)$, and by the assumption, $|e_0|, |e'_0| \leqslant (q^2-1)/2$, hence it satisfies that $|e_0 - e'_0| \leqslant |e_0| + |e'_0| \leqslant q^2 - 1$. Therefore we have $e_0 - e'_0 = \pm q, \pm 2q, \cdots, \pm(q-1)q$. We put $e_0 - e'_0 = aq$, where $a \in \{\pm 1, \pm 2, \cdots, \pm(q-1)\}$. We must have $e_0, {e_0}' \neq 0$. Because if $e_0 = 0$, ${e_0}' = -(e_0 - {e_0}')$ is divisible by q but ${e_0}' \in D_{q,(q^2-1)/2}$, i.e. $q \nmid {e_0}'$. This is contrary to $e_0 = 0$. Similarly, we can show that ${e_0}' \neq 0$. By the definition of ϕ-rNAF, we must have $e_1, {e_1}' = 0$. We have $aq = (e_0 - e'_0) = \sum_{i=1}^{\ell-1}(e'_i - e_i)\phi^i$ and by the characteristic polynomial of ϕ, we have $a\phi - a\phi^2 = \sum_{i=1}^{\ell-1}(e'_i - e_i)\phi^i$. Thus we can see

$$(e_1 - e'_1 + a) = (e'_2 - e_2 + a)\phi + \sum_{i=3}^{\ell-1}(e'_i - e_i)\phi^{i-1}.$$

Hence, from Lemma 3, we also have $q | (e_1 - e'_1 + a)$. It follows that $q | a$. This is contrary to $a \in \{\pm 1, \pm 2, \cdots, \pm(q-1)\}$.

(2) We will show that B_ℓ and T_ℓ satisfy the recurrence equations $B_1 = B_2 = q^2 - q$, $B_{\ell+2} - B_{\ell+1} - (q^2 - q)B_\ell = 0$ and $T_1 = 0$, $T_2 = q^2 - q$,

$$T_{\ell+2} - T_{\ell+1} - (q^2 - q)T_\ell = (q^2 - q)\left(\sum_{i=1}^{\ell} B_i\right).$$

forms	# of representations	# of 0 digits
$(*0* \ \cdots \ )$	$B_1B_{\ell-2}$	$B_1B_{\ell-2} + B_1T_{\ell-2}$
$(*00* \ \cdots \ )$	$B_1B_{\ell-3}$	$2B_1B_{\ell-3} + B_1T_{\ell-3}$
$(*000*\cdots \ )$	$B_1B_{\ell-4}$	$3B_1B_{\ell-4} + B_1T_{\ell-4}$
$\vdots$	$\vdots$	$\vdots$
$(*0000\cdots0*0)$	B_1B_2	$(\ell-3)B_1B_2 + B_1T_2$
$(*0000\cdots00*)$	B_1B_1	$(\ell-2)B_1B_1 + B_1T_1$
$(*0000\cdots000)$	B_1	$(\ell-1)B_1$

Table 5. Forms of ℓ digits ϕ-rNAF representations

In Table 5, the symbol '$*$' means a non-zero digit. We explain how to see Table 5. The left column stand for forms of ℓ digits ϕ-rNAF representation, the center column stand for the number of each ϕ-rNAF representation correspond to each form of the left column, and the right column stand for each number of 0 digits correspond to each form of the left column. From Table 5, we can derive $B_\ell = B_1\left(1 + \sum_{i=1}^{\ell-2} B_i\right)$, and

$$T_\ell = B_1\left(1 + \sum_{i=1}^{\ell-2} T_i\right) + B_1\left(\sum_{i=1}^{\ell-2}(\ell-1-i)B_i + (\ell-1)\right).$$

From these equations, we have the desired recurrence equations, and by elementary calculations, we have $B_\ell = q(q-1)(q^\ell - (1-q)^\ell)/(2q-1)$, and

$$T_\ell = \frac{q(q-1)(q^{\ell-1} - (1-q)^{\ell-1})}{2q-1} + \frac{(\ell-1)(q^{\ell+2}(q-1) - (1-q)^{\ell+2}q)}{(2q-1)^2} + \frac{q^\ell(1-q)^4 + q^4(1-q)^\ell}{q(2q-1)^2} + \frac{q^{\ell-1}(1-q)}{2q-1}\left(\frac{q^{\ell-1}-1}{q^{\ell-1}}\right),$$

thus we have $\lim_{\ell\to\infty} T_\ell/(\ell B_\ell) = q/(2q-1)$ as desired. $\square$