# Efficient Implementation of Bilinear Pairings on ARM Processors 

Gurleen Grewal ${ }^{1}$, Reza Azarderakhsh ${ }^{1}$, Patrick Longa ${ }^{2, \star}$, Shi $\mathrm{Hu}^{3}$, and David Jao ${ }^{1}$<br>${ }^{1}$ Dept. of Combinatorics \& Optimization University of Waterloo, Waterloo ON, N2L 3G1, Canada<br>grewal.gurleen@ymail.com, \{razarder,djao\}@math.uwaterloo.ca<br>${ }^{2}$ Microsoft Research, 1 Microsoft Way, Redmond, WA 98052, USA<br>plonga@microsoft.com<br>${ }^{3}$ Stanford University, Stanford, California, USA<br>s3hu@stanford.edu


#### Abstract

As hardware capabilities increase, low-power devices such as smartphones represent a natural environment for the efficient implementation of cryptographic pairings. Few works in the literature have considered such platforms despite their growing importance in a postPC world. In this paper, we investigate the efficient computation of the Optimal-Ate pairing over Barreto-Naehrig curves in software at different security levels on ARM processors. We exploit state-of-the-art techniques and propose new optimizations to speed up the computation in the tower field and curve arithmetic. In particular, we extend the concept of lazy reduction to inversion in extension fields, analyze an efficient alternative for the sparse multiplication used inside the Miller's algorithm and reduce further the cost of point/line evaluation formulas in affine and projective homogeneous coordinates. In addition, we study the efficiency of using M-type sextic twists in the pairing computation and carry out a detailed comparison between affine and projective coordinate systems. Our implementations on various mass-market smartphones and tablets significantly improve the state-of-the-art of pairing computation on ARM-powered devices, outperforming by at least a factor of 3.7 the best previous results in the literature.


Keywords: Optimal-Ate pairing, Barreto-Naehrig curves, ARM processor, pairing implementation.

## 1 Introduction

In the past decade, bilinear pairings have found a range of constructive applications in areas such as identity-based encryption and short signatures. Naturally, implementing such protocols requires efficient computation of the pairing function. Considerable work has been done to compute fast pairings on PCs 36|8/1517]. Most

[^0]recently, Aranha et al. [3 have computed the O-Ate pairing at the 128-bit security level in under 2 million cycles on various 64 -bit PC processors. In contrast, relatively few articles [110 have considered efficient software implementations of pairings on ARM-based platforms such as hand-held smartphones and tablets. These platforms are widely predicted to become a dominant computing platform in the near future. Therefore, efficient implementation of pairing-based protocols for these devices is crucial for deployment of pairing-based cryptography in a mobile world and represents a natural area of research.

In this paper, we investigate efficient pairing computations at multiple security levels across different generations of ARM-based processors. We extend the work of Aranha et al. [3] to different BN curves and higher security levels. In addition, we make several further optimizations and analyze different options available for implementation at various stages of the pairing computation. We summarize our contributions as follows:

- Firstly, we extend the concept of lazy reduction employed by Aranha et al. [3] (see also Longa [13, Chapter 6]) to inversion in extension fields. We also optimize the sparse multiplication algorithm in the degree 12 extension.
- We examine different choices of towers for extension field arithmetic over various prime fields including BN-254, BN-446, and BN-638 [8]. We determine the most efficient implementation of extension fields in the context of pairing computation over BN-curves from the various choices available.
- The M-type sextic twist [16] has been largely ignored for use in pairing computations, most likely due to the inefficient untwisting map. We demonstrate that by computing the pairing on the twisted curve, we can bypass the inefficient untwisting. As a result, for the purposes of optimization one can use either M-type or D-type twists, thus roughly doubling the available choice of curves.
- Finally, we implement the proposed algorithms for computing the O-Ate pairing over BN curves on different ARM-based platforms and compare our measured timing results to their counterparts in the literature. Our experimental results are 3 to 5 times faster than the fastest available in prior literature, depending on the security level.

Acar et al. [1] have recently raised the question of whether affine coordinates or projective coordinates are a better choice for curve arithmetic in the context of software implementation of pairings. Their conclusion is that affine coordinates are faster at all security levels at or above 128 bits on the ARM platform. In contrast, our results (Section 6.2) demonstrate a clear advantage for homogeneous projective coordinates at the 128-bit security level, although affine coordinates remain faster at the 256 -bit security level. We believe that our findings are more reliable since they represent a more realistic amount of optimization of the underlying field arithmetic implementation. We stress that, except for the work described in Section 6.1 our code does not contain any hand-optimized assembly or any overly aggressive optimizations that would compromise portability or maintainability.

The rest of this paper is organized as follows. In Section 2 we provide some background on the O-Ate pairing. In Section 3, we discuss the representation of extension fields. In Section 4 we describe arithmetic on BN curves including point addition and doubling presented in different coordinates. We provide operation counts for our algorithms in Section 5. In Section 6, we present the results of our implementation of the proposed scheme for computing O-Ate pairings on different ARM processors, and compare them with prior work.

## 2 Preliminaries

Barreto and Naehrig [4] describe a family of pairing friendly curves $E: y^{2}=x^{3}+b$ of order $n$ with embedding degree 12 defined over a prime field $\mathbb{F}_{q}$ where $q$ and $n$ are given by the polynomials:

$$
\begin{align*}
q & =36 x^{4}+36 x^{3}+24 x^{2}+6 x+1 \\
n & =36 x^{4}+36 x^{3}+18 x^{2}+6 x+1 \tag{1}
\end{align*}
$$

for some integer $x$ such that both $q$ and $n$ are prime and $b \in \mathbb{F}_{q}^{*}$ such that $b+1$ is a quadratic residue.

Let $\Pi_{q}: E \rightarrow E$ be the $q$-power Frobenius. Set $\mathbb{G}_{1}=E[n] \cap \operatorname{ker}\left(\Pi_{q}-[1]\right)$ and $\mathbb{G}_{2}=E[n] \cap \operatorname{ker}\left(\Pi_{q}-[q]\right)$. It is known that points in $\mathbb{G}_{1}$ have coordinates in $\mathbb{F}_{q}$, and points in $\mathbb{G}_{2}$ have coordinates in $\mathbb{F}_{q^{12}}$. The Optimal-Ate or O-Ate pairing [18] on $E$ is defined by:

$$
\begin{equation*}
a_{\mathrm{opt}}: \mathbb{G}_{2} \times \mathbb{G}_{1} \rightarrow \mu_{n},(Q, P) \rightarrow f_{6 x+2, Q}(P) \cdot h(P) \tag{2}
\end{equation*}
$$

where $h(P)=l_{[6 x+2] Q, q Q}(P) l_{[6 x+2] Q+q Q,-q^{2} Q}(P)$ and $f_{6 x+2, Q}(P)$ is the appropriate Miller function. Also, $l_{Q_{1}, Q_{2}}(P)$ is the line arising in the addition of $Q_{1}$ and $Q_{2}$ at point $P$. This function can be computed using Miller's algorithm [14]. A modified version of the algorithm from [14] which uses a NAF representation of $x$ is given in Algorithm 1 .

Let $\xi$ be a quadratic and cubic non-residue over $\mathbb{F}_{q^{2}}$. Then the curves $E^{\prime}: y^{2}=$ $x^{3}+\frac{b}{\xi}$ (D-type) and $E^{\prime \prime}: y^{2}=x^{3}+b \xi$ (M-type) are sextic twists of $E$ over $\mathbb{F}_{q^{2}}$, and exactly one of them has order dividing $n$ [3]. For this twist, the image $\mathbb{G}_{2}^{\prime}$ of $\mathbb{G}_{2}$ under the twisting isomorphism lies entirely in $E^{\prime}\left(\mathbb{F}_{q^{2}}\right)$. Instead of using a degree 12 extension, the point $Q$ can now be represented using only elements in a quadratic extension field. In addition, when performing curve arithmetic and computing the line function in the Miller loop, one can perform the arithmetic in $\mathbb{G}_{2}^{\prime}$ and then map the result to $\mathbb{G}_{2}$, which considerably speeds up operations in the Miller loop.

### 2.1 Notations and Definitions

Throughout this paper, lower case variables denote single-precision integers, upper case variables denote double-precision integers. The operation $\times$ represents

```
Algorithm 1. Miller's Algorithm for the O-Ate Pairing [14]
Input: Points \(P, Q \in E[n]\) and integer \(n=\left(n_{l-1}, n_{l-2}, \cdots, n_{1}, n_{0}\right)_{2} \in \mathbb{N}\).
Output: \(f_{n, P}(Q)^{\frac{q^{k}-1}{n}}\).
    \(T \leftarrow P, f \leftarrow 1\)
    for \(i=l-2\) down to 0 do
        \(f \leftarrow f^{2} \cdot l_{T, T}(Q)\)
        \(T \leftarrow 2 T\)
        if \(l_{i} \neq 0\) then
            \(f \leftarrow f \cdot l_{T, P}(Q)\)
            \(T \leftarrow T+P\)
        end if
    end for
    \(f \leftarrow f^{\frac{q^{k}-1}{n}}\)
    return \(f\)
```

multiplication without reduction, and $\otimes$ represents multiplication with reduction. The quantities $m, s, a, i$, and $r$ denote the times for multiplication, squaring, addition, inversion, and modular reduction in $\mathbb{F}_{q}$, respectively. Likewise, $\tilde{m}, \tilde{s}, \tilde{a}, \tilde{i}$, and $\tilde{r}$ denote times for multiplication, squaring, addition, inversion, and reduction in $\mathbb{F}_{q^{2}}$, respectively, and $m_{u}, s_{u}, \tilde{m}_{u}$, and $\tilde{s}_{u}$ denote times for multiplication and squaring without reduction in the corresponding fields. Finally, $m_{b}, m_{i}, m_{\xi}$, and $m_{v}$ denote times for multiplication by the quantities $b, i, \xi$, and $v$ from Section 3.

## 3 Representation of Extension Fields

Efficient implementation of the underlying extension fields is crucial to achieve fast pairing results. The IEEE P1363.3 standard [9] recommends using towers to represent $\mathbb{F}_{q^{k}}$. For primes $q$ congruent to $3 \bmod 8$, we employ the following construction of Benger and Scott [5] to construct tower fields:

Property 1. For approximately $2 / 3$ rds of the BN -primes $q \equiv 3 \bmod 8$, the polynomial $y^{6}-\alpha, \alpha=1+\sqrt{-1}$ is irreducible over $\mathbb{F}_{q^{2}}=\mathbb{F}_{q}(\sqrt{-1})$.

This gives the following towering scheme:

$$
\begin{cases}\mathbb{F}_{q^{2}}=\mathbb{F}_{q}[i] /\left(i^{2}-\beta\right), & \text { where } \beta=-1 \\ \mathbb{F}_{q^{6}}=\mathbb{F}_{q^{2}}[v] /\left(v^{3}-\xi\right), & \text { where } \xi=1+i \\ \mathbb{F}_{q^{12}}=\mathbb{F}_{q^{6}}[w] /\left(w^{2}-v\right) . & \end{cases}
$$

Based on this scheme, multiplication by $i$ requires one negation over $\mathbb{F}_{q}$, and multiplication by $\xi$ requires only one addition over $\mathbb{F}_{q^{2}}$. For primes congruent to $7 \bmod 8$, we use the following construction which can be proven using the same ideas as those in Benger and Scott [5]:

Property 2. For approximately $2 / 3$ rds of the BN -primes $q \equiv 7 \bmod 8$, the polynomial $y^{6}-\alpha, \alpha=1+\sqrt{-2}$ is irreducible over $\mathbb{F}_{q^{2}}=\mathbb{F}_{q}(\sqrt{-2})$.

This gives the following towering scheme:

$$
\begin{cases}\mathbb{F}_{q^{2}}=\mathbb{F}_{q}[i] /\left(i^{2}-\beta\right), & \text { where } \beta=-2 \\ \mathbb{F}_{q^{6}}=\mathbb{F}_{q^{2}}[v] /\left(v^{3}-\xi\right), & \text { where } \xi=1+i \\ \mathbb{F}_{q^{12}}=\mathbb{F}_{q^{6}}[w] /\left(w^{2}-v\right) . & \end{cases}
$$

All things being equal, the towering scheme derived from Property 1 is slightly faster for a given bit size. However, in practice, desirable BN-curves are rare, and it is sometimes necessary to use primes $q \equiv 7 \bmod 8$ in order to optimize other aspects such as the Hamming weight of $x$. In particular, the curves BN-446 and BN-638 [8] have $q \equiv 7 \bmod 8$. In such cases, Property 1 does not apply, so we use the towering scheme derived from Property 2. We also considered other approaches to construct tower extensions as suggested in [8], but found the above schemes consistently resulted in faster pairings compared to the other options.

### 3.1 Finite Field Operations and Lazy Reduction

Aranha et al. 3] proposed a lazy reduction scheme for efficient pairing computation in tower-friendly fields and curve arithmetic using projective coordinates. We extensively exploit their method and extend it to field inversion and curve arithmetic over affine coordinates. The proposed schemes using lazy reduction for inversion are given in Algorithms 2, 3and 4 for $\mathbb{F}_{q^{2}}, \mathbb{F}_{q^{6}}$ and $\mathbb{F}_{q^{12}}$, respectively. The total savings with lazy reduction vs. no lazy reduction are one $\mathbb{F}_{q}$-reduction in $\mathbb{F}_{q^{2}}$-inversion, and $36 \mathbb{F}_{q^{-}}$-reductions in $\mathbb{F}_{q^{12}}$-inversion (improving upon [3] by $16 \mathbb{F}_{q}$-reductions). Interestingly enough, if one applies the lazy reduction technique to the recent $\mathbb{F}_{q^{12}}$ inversion algorithm of Pereira et al. 8], it replaces two $\tilde{m}_{u}$ by two $\tilde{s}_{u}$ but requires five more $\tilde{r}$ operations, which ultimately makes it slower in practice in comparison with the proposed scheme.

```
Algorithm 2. Inversion over \(\mathbb{F}_{q^{2}}\) employing lazy reduction technique
Input: \(a=a_{0}+a_{1} i ; a_{0}, a_{1} \in \mathbb{F}_{q} ; \beta\) is a quadratic non-residue over \(\mathbb{F}_{q}\)
Output: \(c=a^{-1} \in \mathbb{F}_{q^{2}}\)
    \(T_{0} \leftarrow a_{0} \times a_{0}\)
    \(T_{1} \leftarrow-\beta \cdot\left(a_{1} \times a_{1}\right)\)
    \(T_{0} \leftarrow T_{0}+T_{1}\)
    \(t_{0} \leftarrow T_{0} \bmod p\)
    \(t_{0} \leftarrow t_{0}^{-1} \bmod p\)
    \(c_{0} \leftarrow a_{0} \otimes t_{0}\)
    \(c_{1} \leftarrow-\left(a_{1} \otimes t_{0}\right)\)
    return \(c=c_{0}+c_{1} i\)
```

```
Algorithm 3. Inversion over \(\mathbb{F}_{q^{6}}\) employing lazy reduction technique
Input: \(a=a_{0}+a_{1} v+a_{2} v^{2} ; a_{0}, a_{1}, a_{2} \in \mathbb{F}_{q^{2}}\)
Output: \(c=a^{-1} \in \mathbb{F}_{q^{6}}\)
    \(T_{0} \leftarrow a_{0} \times a_{0}\)
    \(t_{0} \leftarrow \xi a_{1}\)
    \(T_{1} \leftarrow t_{0} \times a_{2}\)
    \(T_{0} \leftarrow T_{0}-T_{1}\)
    \(t_{1} \leftarrow T_{0} \bmod p\)
    \(T_{0} \leftarrow a_{2} \times a_{2}\)
    \(T_{0} \leftarrow \xi T_{0}\)
    \(T_{1} \leftarrow a_{0} \times a_{1}\)
    \(T_{0} \leftarrow T_{0}-T_{1}\)
    \(t_{2} \leftarrow T_{0} \bmod p\)
    \(T_{0} \leftarrow a_{1} \times a_{1}\)
    \(T_{1} \leftarrow a_{0} \times a_{2}\)
    \(T_{0} \leftarrow T_{0}-T_{1}\)
    \(t_{3} \leftarrow T_{0} \bmod p\)
    \(T_{0} \leftarrow t_{0} \times t_{3}\)
    \(T_{1} \leftarrow a_{0} \times t_{1}\)
    \(T_{0} \leftarrow T_{0}+T_{1}\)
    \(t_{0} \leftarrow \xi a_{2}\)
    \(T_{1} \leftarrow t_{0} \times t_{2}\)
    \(T_{0} \leftarrow T_{0}+T_{1}\)
    \(t_{0} \leftarrow T_{0} \bmod p\)
    \(t_{0} \leftarrow t_{0}^{-1}\)
    \(c_{0} \leftarrow t_{1} \otimes t_{0}\)
    \(c_{1} \leftarrow t_{2} \otimes t_{0}\)
    \(c_{2} \leftarrow t_{3} \otimes t_{0}\)
    return \(c=c_{1}+c_{2} v+c_{3} v^{2}\)
```

```
Algorithm 4. Inversion over \(\mathbb{F}_{q^{12}}\) employing lazy reduction technique
Input: \(a=a_{0}+a_{1} w ; a_{0}, a_{1} \in \mathbb{F}_{q^{6}}\)
Output: \(c=a^{-1} \in \mathbb{F}_{q^{12}}\)
    \(T_{0} \leftarrow a_{0} \times a_{0}\)
    \(T_{1} \leftarrow v \cdot\left(a_{1} \times a_{1}\right)\)
    \(T_{0} \leftarrow T_{0}-T_{1}\)
    \(t_{0} \leftarrow T_{0} \bmod p\)
    \(t_{0} \leftarrow t_{0}^{-1} \bmod p\)
    \(c_{0} \leftarrow a_{0} \otimes t_{0}\)
    \(c_{1} \leftarrow-a_{1} \otimes t_{0}\)
    return \(c=c_{0}+c_{1} w\)
```

The line function in the Miller loop evaluates to a sparse $\mathbb{F}_{q^{12}}$ element containing only three of the six basis elements over $\mathbb{F}_{q^{2}}$. Thus, when multiplying the line function output with $f_{i, Q}(P)$, one can utilize the sparseness property to avoid full $\mathbb{F}_{q^{12}}$ arithmetic (Algorithm (5). For the BN-254 curve [8], our sparse multiplication algorithm requires $13 \tilde{m}$ and $44 \tilde{a}$ when a D-type twist is involved.

```
Algorithm 5. D-type sparse-dense multiplication in \(\mathbb{F}_{q^{12}}\)
Input: \(a=a_{0}+a_{1} w+a_{2} v w ; a_{0}, a_{1}, a_{2} \in \mathbb{F}_{q^{2}}, b=b_{0}+b_{1} w ; b_{0}, b_{1} \in \mathbb{F}_{q^{6}}\)
Output: \(a b \in \mathbb{F}_{q^{12}}\)
    \(A_{0} \leftarrow a_{0} \times b_{0}[0], A_{1} \leftarrow a_{0} \times b_{0}[1], A_{2} \leftarrow a_{0} \times b_{0}[2]\)
    \(A \leftarrow A_{0}+A_{1} v+A_{2} v^{2}\)
    \(B \leftarrow \mathrm{Fq} 6 \mathrm{SparseMul}\left(a_{1} w+a_{2} v w, b_{1}\right)\)
    \(c_{0} \leftarrow a_{0}+a_{1}, c_{1} \leftarrow a_{2}, c_{2} \leftarrow 0\)
    \(c \leftarrow c_{0}+c_{1} v+c_{2} v^{2}\)
    \(d \leftarrow b_{0}+b_{1}\)
    \(E \leftarrow \mathrm{Fq} 6 \operatorname{SparseMul}(c, d)\)
    \(F \leftarrow E-(A+B)\)
    \(G \leftarrow B v\)
    \(H \leftarrow A+G\)
    \(c_{0} \leftarrow H \bmod p\)
    \(c_{1} \leftarrow F \bmod p\)
    return \(c=c_{0}+c_{1} w\)
```

```
Algorithm 6. Fq6SparseMul, used in Algorithm [5
Input: \(a=a_{0}+a_{1} v ; a_{0}, a_{1} \in \mathbb{F}_{q^{2}}, b=b_{0}+b_{1} v+b_{2} v^{2} ; b_{0}, b_{1}, b_{2} \in \mathbb{F}_{q^{2}}\)
Output: \(a b \in \mathbb{F}_{q^{6}}\)
    \(A \leftarrow a_{0} \times b_{0}, B \leftarrow a_{1} \times b_{1}\)
    \(C \leftarrow a_{1} \times b_{2} \xi\)
    \(D \leftarrow A+C\)
    \(e \leftarrow a_{0}+a_{1}, f \leftarrow b_{0}+b_{1}\)
    \(E \leftarrow e \times f\)
    \(G \leftarrow E-(A+B)\)
    \(H \leftarrow a_{0} \times b_{2}\)
    \(I \leftarrow H+B\)
    return \(D+G v+I v^{2}\)
```

A similar dense-sparse multiplication algorithm works for M-type twists, and requires an extra multiplication by $v$. We note that our approach requires 13 fewer additions over $\mathbb{F}_{q^{2}}$ compared to the one used in [3] (lazy reduction versions).

### 3.2 Mapping from the Twisted Curve to the Original Curve

Suppose we take $\xi$ (from the towering scheme) to be the cubic and quadratic nonresidue used to generate the sextic twist of the BN -curve $E$. After manipulating points on the twisted curve, they need to be mapped to the original curve. In the case of a D-type twist, the untwisting isomorphism is given by:

$$
\begin{equation*}
\Psi:(x, y) \rightarrow\left(\xi^{\frac{1}{3}} x, \xi^{\frac{1}{2}} y\right)=\left(w^{2} x, w^{3} y\right) \tag{3}
\end{equation*}
$$

where both $w^{2}$ and $w^{3}$ are basis elements, and hence the untwisting map is almost free. If one uses a M-type twist the untwisting isomorphism is given as follows:

$$
\begin{equation*}
\Psi:(x, y) \rightarrow\left(\xi^{-\frac{2}{3}} x, \xi^{-\frac{1}{2}} y\right)=\left(\xi^{-1} w^{4} x, \xi^{-1} w^{3} y\right) \tag{4}
\end{equation*}
$$

Untwisting of (4) is not efficient as the one given in (3). However, if we compute the pairing value on the twisted curve instead of the original curve, then we do not need to use the untwisting map. Instead, we require the inverse map which is almost free. Therefore, we compute the pairing on the original curve $E$ when a D-type twist is involved, and on the twisted curve $E^{\prime}$ when an M-type twist is involved. Using this approach, we have found that both twist types are equivalent in performance up to point/line evaluation. The advantage of being able to consider both twist types is the immediate availability of many more useful curves for pairing computation.

### 3.3 Final Exponentiation Scheme

We use the final exponentiation scheme proposed in [7], which represents the current state-of-the-art for BN curves. In this scheme, first $\frac{q^{12}-1}{n}$ is factored into $q^{6}-1, q^{2}+1$, and $\frac{q^{4}-q^{2}+1}{n}$. The first two factors are easy to exponentiate. The remaining exponentiation $\frac{q^{4}-q^{2}+1}{n}$ can be performed in the cyclotomic subgroup. Using the fact that any fixed non-degenerate power of a pairing is a pairing, we raise to a multiple of the remaining factor. Recall that $q$ and $n$ are polynomials in $x$, and hence so is the final factor. We denote this polynomial as $d(x)$. In 7 ] it is shown that

$$
\begin{align*}
2 x\left(6 x^{2}+3 x+1\right) d(x)= & \lambda_{3} q^{3}+\lambda_{2} q^{2}+\lambda_{1} q+\lambda_{0} 1+6 x+12 x^{2}+12 x^{3} \\
& +\left(4 x+6 x^{2}+12 x^{3}\right) p(x)+\left(6 x+6 x^{2}+12 x^{3}\right) p(x)^{2} \\
& +\left(-1+4 x+6 x^{2}+12 x^{3}\right) p(x)^{3} \tag{5}
\end{align*}
$$

where

$$
\begin{align*}
& \lambda_{3}(x)=-1+4 x+6 x^{2}+12 x^{3} \\
& \lambda_{2}(x)=6 x+6 x^{2}+12 x^{3} \\
& \lambda_{1}(x)=4 x+6 x^{2}+12 x^{3} \\
& \lambda_{0}(x)=1+6 x+12 x^{2}+12 x^{3} \tag{6}
\end{align*}
$$

To compute (6), the following exponentiations are performed:

$$
\begin{equation*}
f \mapsto f^{x} \mapsto f^{2 x} \mapsto f^{4 x} \mapsto f^{6 x} \mapsto f^{6 x^{2}} \mapsto f^{12 x^{2}} \mapsto f^{12 x^{3}} . \tag{7}
\end{equation*}
$$

The cost of computing (17) is 3 exponentiations by $x, 3$ squarings and 1 multiplication. We then compute the terms $a=f^{12 x^{3}} f^{6 x^{2}} f^{6 x}$ and $b=a\left(f^{2 x}\right)^{-1}$, which require 3 multiplications. The final pairing value is obtained as

$$
\begin{equation*}
a f^{6 x^{2}} f b^{p} a^{p^{2}}\left(b f^{-1}\right)^{p^{3}}, \tag{8}
\end{equation*}
$$

which costs 6 multiplications and 6 Frobenius operations. In total, this method requires 3 exponentiations by $x, 3$ squarings, 10 multiplications, and 3 Frobenius operations. In comparison, the technique used in [3] requires 3 additional multiplications and an additional squaring, and thus is slightly slower.

## 4 Curve Arithmetic

In this section, we discuss our optimizations to curve arithmetic over affine and homogeneous projective coordinates. We also evaluated other coordinate systems such as Jacobian coordinates but found that none were faster than homogeneous coordinates for our application.

### 4.1 Affine Coordinates

Let the points $T=(x, y)$ and $Q=\left(x_{2}, y_{2}\right) \in E^{\prime}\left(\mathbb{F}_{q}\right)$ be given in affine coordinates, and let $T+Q=\left(x_{3}, y_{3}\right)$ be the sum of the points $T$ and $Q$. When $T=Q$ we have

$$
\begin{aligned}
m & =\frac{3 x^{2}}{2 y} \\
x_{3} & =m^{2}-2 x \\
y_{3} & =(m x-y)-m x_{3}
\end{aligned}
$$

For D-type twists, the secant or tangent line evaluated at $P=\left(x_{P}, y_{P}\right)$ is given by:

$$
\begin{equation*}
l_{2 \Psi(T)}(P)=y_{P}-m x_{P} w+(m x-y) w^{3} . \tag{9}
\end{equation*}
$$

To compute the above, we precompute $\bar{x}_{P}=-x_{P}$ (to save the cost of computing the negation on-the-fly) and use the following sequence of operations which requires $1 \tilde{i}, 3 \tilde{m}, 2 \tilde{s}, 7 \tilde{a}$, and $2 m$ if $T=Q$. In comparison, the doubling formula in Lauter et al. [12] costs 3 additional $\tilde{a}$.

$$
\begin{gathered}
A=\frac{1}{2 y} \quad B=3 x^{2} \quad C=A B \quad D=2 x \quad x_{3}=C^{2}-D \\
E=C x-y \quad y_{3}=E-C x_{3} \quad F=C \bar{x}_{P} \\
l_{2 \Psi(T)}(P)=y_{P}+F w+E w^{3}
\end{gathered}
$$

Similarly, when $T \neq Q$ we use the following sequence of operations which requires $1 \tilde{i}, 3 \tilde{m}, 1 \tilde{s}, 6 \tilde{a}$, and $2 m$ - saving $2 \tilde{a}$ compared to the addition formula in Lauter et al. [12].

$$
\begin{gathered}
A=\frac{1}{y_{2}-y} \quad B=x_{2}-x \quad C=A B \quad D=x+x_{2} \quad x_{3}=C^{2}-D \\
E=C x-y \quad y_{3}=E-C x_{3} \quad F=C \bar{x}_{P} \\
l_{2 \Psi(T)}(P)=y_{P}+F w+E w^{3}
\end{gathered}
$$

In an M-type twist, the tangent line evaluated at $\Psi(P)=\left(x_{P} w^{2}, y_{P} w^{3}\right)$ is given by:

$$
\begin{equation*}
l_{2 T}(\Psi(P))=y_{P} w^{3}-m x_{P} w^{2}+(m x-y) \tag{10}
\end{equation*}
$$

and can be computed in a similar way.

### 4.2 Homogeneous Coordinates

During the first iteration of the Miller loop, the $Z$-coordinate of the point $Q$ has value equal to 1 . We use this fact to eliminate a multiplication and three squarings using a special first doubling routine in the first iteration. Recently, Arahna et al. [3], presented optimized formulas for point doubling/line evaluation. We note that the twisting point $P$ given by $\left(x_{P} / w^{2}, y_{P} / w^{3}\right)$ is better represented by $\left(x_{P} w, y_{P}\right)$, which is obtained by multiplying by $w^{3}$. This eliminates the multiplication by $\xi$ and gives the following revised formula. Let $T=(X, Y, Z) \in E^{\prime}\left(\mathbb{F}_{q^{2}}\right)$ be in homogeneous coordinates. Then $2 T=\left(X_{3}, Y_{3}, Z_{3}\right)$ is given by:

$$
\begin{aligned}
X_{3} & =\frac{X Y}{2}\left(Y^{2}-9 b^{\prime} Z^{2}\right) \\
Y_{3} & =\left[\frac{1}{2}\left(Y^{2}+9 b^{\prime} Z^{2}\right)\right]^{2}-27 b^{\prime 2} Z^{4} \\
Z_{3} & =2 Y^{3} Z
\end{aligned}
$$

In the case of a D-type twist, the corresponding line function evaluated at $P=$ $\left(x_{P}, y_{P}\right)$ is given by:

$$
l_{2 \Psi(T)}(P)=-2 Y Z y_{P}+3 X^{2} x_{P} w+\left(3 b^{\prime} Z^{2}-Y^{2}\right) w^{3}
$$

We compute this value using the following sequence of operations.

$$
\begin{gathered}
A=\frac{X Y}{2} \quad B=Y^{2} \quad C=Z^{2} \quad E=3 b^{\prime} C \quad F=3 E \quad X_{3}=A \cdot(B-F) \\
G=\frac{B+f}{2} \quad Y_{3}=G^{2}-3 E^{2} \quad H=(Y+Z)^{2}-(B+C) \quad Z_{3}=B \cdot H \\
l_{2 \Psi(T)}(P)=H \bar{y}_{P}+3 X^{2} x_{P} w+(E-B) w^{3}
\end{gathered}
$$

Aranha et al. [3] observe $\tilde{m}-\tilde{s} \approx 3 \tilde{a}$ and hence computing $X Y$ directly is faster than using $(X+Y)^{2}, Y^{2}$ and $X^{2}$ on a PC. However, on ARM processors, we have $\tilde{m}-\tilde{s} \approx 6 \tilde{a}$. Thus, the latter technique is more efficient on ARM processors. The overall cost of point doubling and line evaluation is $2 \tilde{m}, 7 \tilde{s}, 22 \tilde{a}$, and $4 m$, assuming that the cost of division by two and multiplication by $b^{\prime}$ are equivalent to the cost of addition. Similarly, we compute point addition and line function evaluation using the following sequence of operations which uses $11 \tilde{m}, 2 \tilde{s}, 8 \tilde{a}$, and $4 m$ (saving $2 \mathbb{F}_{q^{2}}$ additions over [3]). Note that $\bar{x}_{P}$ and $\bar{y}_{P}$ are precomputed to save again the cost of computing $-x_{P}$ and $-y_{P}$.

$$
\begin{gathered}
A=Y_{2} Z \quad B=X_{2} Z \quad \theta=Y-A \quad \lambda=X-B \\
D=\lambda^{2} \quad E=\lambda^{3} \quad F=Z C \quad G=X D \quad H=E+F-2 G \\
X_{3}=\lambda H \quad I=Y E \quad Y_{3}=\theta(G-H)-I \quad Z_{3}=Z E \quad J=\theta X_{2}-\lambda Y_{2} \\
l_{\Psi(T+Q)}(P)=\lambda \bar{y}_{P}+\theta \bar{x}_{P} w+J w^{3}
\end{gathered}
$$

In the case of an M-type twist the corresponding line computation can be computed using the same sequences of operations as above. As in [3], we also use lazy reduction techniques to optimize the above formulae (see Table 11).

Table 1. Operation counts for 254 -bit, 446 -bit, and 638 -bit prime fields

| $E^{\prime}\left(\mathbb{F}_{p^{2}}\right)$ Arith. | 254-bit | 446-bit/638-bit |
| :---: | :---: | :---: |
| Doubl/Eval (Proj) | $2 \tilde{m}_{u}+7 \tilde{s}_{u}+8 \tilde{r}+25 \tilde{a}+4 m$ | $2 \tilde{m}_{u}+7 \tilde{s}_{u}+8 \tilde{r}+34 \tilde{a}+a+4 m$ |
| Doubl/Eval (Affi) | $\tilde{i}+3 \tilde{m}_{u}+2 \tilde{s}_{u}+5 \tilde{r}+7 \tilde{a}+2 m$ | $\tilde{i}+3 \tilde{m}_{u}+2 \tilde{s}_{u}+5 \tilde{r}+7 \tilde{a}+2 m$ |
| Add./Eval (Pro) | $11 \tilde{m}_{u}+2 \tilde{s}_{u}+11 \tilde{r}+10 \tilde{a}+4 m$ | $11 \tilde{m}_{u}+2 \tilde{s}_{u}+11 \tilde{r}+10 \tilde{a}+4 m$ |
| Add/Eval (Affi) | $\tilde{i}+3 \tilde{m}_{u}+\tilde{s}_{u}+4 \tilde{r}+6 \tilde{a}+2 m$ | $\tilde{i}+2 \tilde{m}_{u}+\tilde{s}_{u}+3 \tilde{r}+6 \tilde{a}+2 m$ |
| First doubl./Eval | $3 \tilde{m}_{u}+4 \tilde{s}_{u}+7 \tilde{r}+14 \tilde{a}+4 m$ | $3 \tilde{m}_{u}+4 \tilde{s}_{u}+7 \tilde{r}+23 \tilde{a}+a+4 m$ |
| $p$-power Frob. | $2 \tilde{m}+2 a$ | $8 \tilde{m}+2 a$ |
| $p^{2}$ - power Frob. | $4 m$ | $16 \tilde{m}+4 a$ |
| $\mathbb{F}_{p^{2}}$ Arith. | 254-bit | 446-bit/638-bit |
| Add/Subtr./Nega. | $\tilde{a}=2 a$ | $\tilde{a}=2 a$ |
| Mult. | $\tilde{m}=\tilde{m}_{u}+\tilde{r}=3 m_{u}+2 r+8 a$ | $\tilde{m}=\tilde{m}_{u}+\tilde{r}=3 m_{u}+2 r+10 a$ |
| Squaring | $\tilde{s}=\tilde{s}_{u}+\tilde{r}=2 m_{u}+2 r+3 a$ | $\tilde{s}=\tilde{s}_{u}+\tilde{r}=2 m_{u}+2 r+5 a$ |
| Mult. by $\beta$ | $m_{b}=a$ | $m_{b}=2 a$ |
| Mult. by $\xi$ | $m_{\xi}=2 a$ | $m_{\xi}=3 a$ |
| Inversion | $i=i+2 m_{u}+2 s_{u}+3 r+3 a$ | $i=i+2 m_{u}+2 s_{u}+3 r+5 a$ |
| $\mathbb{F}_{p^{12}}$ Arith. | 254-bit | 446-bit/638-bit |
| Multi. | $18 \tilde{m}_{u}+110 \tilde{a}+6 \tilde{r}$ | $18 \tilde{m}_{u}+117 \tilde{a}+6 \tilde{r}$ |
| Sparse Mult. | $13 \tilde{m}_{u}+6 \tilde{r}+48 \tilde{a}$ | $13 \tilde{m}_{u}+6 \tilde{r}+54 \tilde{a}$ |
| Sparser Mult. | $6 \tilde{m}_{u}+6 \tilde{r}+13 \tilde{a}$ | $6 \tilde{m}_{u}+6 \tilde{r}+14 \tilde{a}$ |
| Affi. Sparse Mult. | $10 \tilde{m}_{u}+6 \tilde{r}+47 \tilde{a}+6 m_{u}+a$ | $10 \tilde{m}_{u}+53 \tilde{a}+6 \tilde{r}+6 m_{u}+a$ |
| Squaring | $12 \tilde{m}_{u}+6 \tilde{r}+73 \tilde{a}$ | $12 \tilde{m}_{u}+6 \tilde{r}+78 \tilde{a}$ |
| Cyclotomic Sqr. | $9 \tilde{s}_{u}+46 \tilde{a}+6 \tilde{r}$ | $9 \tilde{s}_{u}+49 \tilde{a}+a+6 \tilde{r}$ |
| Simult. Decomp. | $9 \tilde{m}+6 \tilde{s}+22 \tilde{a}+\tilde{i}$ | $\begin{gathered} 9 \tilde{m}+6 \tilde{s}+24 \tilde{a}+\tilde{i}(\mathrm{BN}-446) \\ 16 \tilde{m}+9 \tilde{s}+35 \tilde{a}+\tilde{i}(\mathrm{BN}-638) \end{gathered}$ |
| $p$-power Frob. | $5 \tilde{m}+6 a$ | $5 \tilde{m}+6 a$ |
| $p^{2}$-power Frob. | $10 m+2 \tilde{a}$ | $10 m+2 \tilde{a}$ |
| Expon. by $x$ | $\begin{aligned} & 45 \tilde{m}_{u}+378 \tilde{s}_{u}+ \\ & 275 \tilde{r}+2164 \tilde{a}+\tilde{i} \end{aligned}$ | $\begin{gathered} 45 \tilde{m}_{u}+666 \tilde{s}_{u}+467 \tilde{r}_{u}+ \\ 3943 \tilde{a}+\tilde{i}(\mathrm{BN}-446) \\ \hline \end{gathered}$ |
|  |  | $\begin{gathered} 70 \tilde{m}+948 \tilde{s}+675 \tilde{r}+ \\ 5606 \tilde{a}+158 a+\tilde{i}(\mathrm{BN}-638) \end{gathered}$ |
| Inversion | $25 \tilde{m}_{u}+9 \tilde{s}_{u}+16 \tilde{r}+121 \tilde{a}+i$ | $25 \tilde{m}_{u}+9 \tilde{s}_{u}+18 \tilde{r}+138 \tilde{a}+i$ |
| Compressed Sqr. | $6 \tilde{s}_{u}+31 \tilde{a}+4 \tilde{r}$ | $6 \tilde{s}_{u}+33 \tilde{a}+a+4 \tilde{r}$ |

## 5 Operation Counts

We provide here detailed operation counts for our algorithms on the $\mathrm{BN}-254$, BN-446, and BN-638 curves used in Acar et. al [1] and defined in [8. Table 1 provides the operation counts for all component operations. Numbers for BN-446 and BN-638 are the same except where indicated.

For BN-254, using the techniques described above, the projective pairing Miller loop executes one negation in $\mathbb{F}_{q}$, one first doubling with line evaluation, 63 point doublings with line evaluations, 6 point additions with line evaluations, one $p$-power Frobenius in $E^{\prime}\left(\mathbb{F}_{p^{2}}\right)$, one $p^{2}$-power Frobenius in $E^{\prime}\left(\mathbb{F}_{p^{2}}\right), 66$

Table 2. Cost of the computation of O-Ate pairings using various coordinates

| Curve | Coordinates | Cost |
| :---: | :---: | :---: |
| BN-254 | Proj. Miller loop | $1841 \tilde{m}_{u}+457 \tilde{s}_{u}+1371 \tilde{r}+9516 \tilde{a}+284 m+3 a$ |
|  | Aff. Miller loop | $70 i+1658 \tilde{m}_{u}+134 \tilde{s}_{u}+942 \tilde{r}+8292 \tilde{a}+540 m+132 a$ |
|  | Final exp. | $386 \tilde{m}_{u}+1164 \tilde{s}_{u}+943 \tilde{r}+4 \tilde{i}+7989 \tilde{a}+30 m+15 a$ |
| BN-446 | Proj. Miller loop | $3151 \tilde{m}_{u}+793 \tilde{s}_{u}+2345 \tilde{r}+18595 \tilde{a}+472 m+117 a$ |
|  | Aff. Miller loop | $118 \tilde{i}+2872 \tilde{m}_{u}+230 \tilde{s}_{u}+1610 \tilde{r}+15612 \tilde{a}+920 m+230 a$ |
|  | Final exp. | $386 \tilde{m}_{u}+2034 \tilde{s}_{u}+1519 \tilde{r}+4 i+13374 \tilde{a}+30 m+345 a$ |
| BN-638 | Proj. Miller loop | $4548 \tilde{m}_{u}+1140 \tilde{s}_{u}+3557 \tilde{r}+27198 \tilde{a}+676 m+166 a$ |
|  | Aff. Miller loop | $169 i+4143 \tilde{m}_{u}+330 \tilde{s}_{u}+2324 \tilde{r}+22574 \tilde{a}+1340 m+333 a$ |
|  | Final exp. | $436 \tilde{m}_{u}+2880 \tilde{s}_{u}+2143 \tilde{r}+4 i+18528 \tilde{a}+30 m+489 a$ |

sparse multiplications, 63 squarings in $\mathbb{F}_{p^{2}}, 1$ negation in $E^{\prime}\left(\mathbb{F}_{p^{2}}\right), 2$ sparser (i.e. sparse-sparse) multiplications [3], and 1 multiplication in $\mathbb{F}_{p^{12}}$. Using Table 1 we compute the total number of operations required in the Miller loop using homogeneous projective coordinates to be

$$
\begin{aligned}
\text { ML254P }= & a+3 \tilde{m}_{u}+7 \tilde{r}+14 \tilde{a}+4 m+63\left(2 \tilde{m}_{u}+7 \tilde{s}_{u}+8 \tilde{r}+25 \tilde{a}+4 m\right)+ \\
& 6\left(11 \tilde{m}_{u}+2 \tilde{s}_{u}+11 \tilde{r}+10 \tilde{a}+4 m\right)+2 \tilde{m}+2 a+4 m+ \\
& 66\left(\tilde{m}_{u}+6 \tilde{r}+48 \tilde{a}\right)+63\left(12 \tilde{m}_{u}+6 \tilde{r}+73 \tilde{a}\right)+\tilde{a}+ \\
& 2\left(6 \tilde{m}_{u}+6 \tilde{r}+13 \tilde{a}\right)+18 \tilde{m}_{u}+110 \tilde{a}+6 \tilde{r} \\
= & 1841 \tilde{m}_{u}+457 \tilde{s}_{u}+1371 \tilde{r}+9516 \tilde{a}+284 m+3 a .
\end{aligned}
$$

Similarly, we also compute the Miller loop operation costs for BN-446 and BN638 and for projective and affine coordinates, and give the results in Table 2

We also compute the operation count for the final exponentiation. For BN254 , the final exponentiation requires 6 conjugations in $\mathbb{F}_{p^{12}}$, one negation in $E^{\prime}\left(\mathbb{F}_{p^{2}}\right)$, one inversion in $\mathbb{F}_{p^{12}}, 12$ multiplications in $\mathbb{F}_{p^{12}}$, two $p$-power Frobenius in $\mathbb{F}_{p^{12}}, 3 p^{2}$-power Frobenius in $\mathbb{F}_{p^{12}}, 3$ exponentiations by $x$, and 3 cyclotomic squarings. Based on these costs, we compute the total number of operations required in the final exponentiation and give the result in Table 2 Similarly, we also compute the final exponentiation operation cost for BN-446 and BN-638. The total operation count for the pairing computation is the cost of the Miller loop plus the final exponentiation.

## 6 Implementation Results

To evaluate the performance of the proposed schemes for computing the O-Ate pairing in practice, we implemented them on various ARM processors. We used the following platforms in our experiments.

- A Marvell Kirkwood 6281 ARMv5 CPU processor (Feroceon 88FR131) operating at 1.2 GHz . In terms of registers it has 1632 -bit registers $\mathrm{r}_{0}$ to $\mathrm{r}_{15}$ of which two are for the stack pointer and program counter, leaving only 14 32 -bit registers for general use.
- An iPad 2 (Apple A5) using an ARMv7 Cortex-A9 MPCore processor operating at 1.0 GHz clock frequency. It has 16 128-bit vector registers which are available as 3264 -bit vector registers, as these registers share physical space with the 128 -bit vector registers.
- A Samsung Galaxy Nexus (1.2 GHz TI OMAP 4460 ARM Cortex-A9). The CPU microarchitecture is identical to the Apple A5. We included it to examine whether different implementations of the Cortex-A9 core have comparable performance in this application.

Our software is based on version 0.2.3 of the RELIC toolkit [2], with the GMP 5.0.2 backend, modified to include our optimizations. Except for the work described in Section 6.1 all of our software is platform-independent C code, and the same source package runs unmodified on all the above ARM platforms as well as both x86 and x86-64 Linux and Windows PCs. Our implementation also supports and includes BN curves at additional security levels beyond the three presented here. For each platform, we used the standard operating system and development environment that ships with the device, namely Debian Squeeze (native C compiler), XCode 4.3.0, and Android NDK (r7c) for the Kirkwood, iPad, and Galaxy Nexus respectively.

We present the results of our experiments in Table 3. For ease of comparison we have also included the numbers from [1] in Table 3. Roughly speaking, our timings are over three times faster than the results appearing in [1], which itself represents the fastest reported times prior to our work. Specifically, examining our iPad results, which were obtained on an identical micro-architecture and clock speed, we find that our implementation is $3.7,3.7$, and 5.4 times faster on BN-254, BN-446, and BN-638, respectively. Some, but not all, of the improvement can be attributed to faster field arithmetic; for example, $\mathbb{F}_{q}$-field multiplication on the RELIC toolkit is roughly 1.4 times as fast on the iPad 2 compared to [1]. A more detailed comparison based on operation counts is difficult because [1] does not provide any operation counts, and also because our strategy and our operation counts rely on lazy reduction, which does not play a role in [1].

### 6.1 Assembly Optimization

In order to investigate the potential performance gains available from hand optimized machine code, we implemented the two most commonly used field arithmetic operations (addition and multiplication) for the BN-254 curve in ARM assembly instructions. Due to the curve-specific and platform-specific nature of this endeavor, we performed this work only for the BN-254 curve and only on the Linux platforms (Marvell Kirkwood and Galaxy Nexus).

The main advantage of assembly language is that it provides more control for lower level arithmetic computations. Although the available C compilers are quite good, they still produce inefficient code since in the C language it is infeasible to express instruction priorities. Moreover, one can use hand-optimized assembly code to decompose larger computations into small pieces suitable for vectorization. We employ the following techniques to optimize our implementation in assembly:

Table 3. Timings for affine and projective pairings on different ARM processors and comparisons with prior literature. Times for the Miller loop (ML) in each row reflect those of the faster pairing.

| Marvell Kirkwood (ARM v5) Feroceon 88FR131 at 1.2 GHz [This work] |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Field | Lang | Operation Timing |  |  |  |  |  |  |  |  | [ $\mu s$ ] |  |  |  |
| Size |  | $a$ | $m$ | $r$ | $i$ | I/M | a | $\tilde{m}$ | $\tilde{s}$ | \| $i$ | ML | FE | O-A(a) | $\mathrm{O}-\mathrm{A}(\mathrm{p})$ |
| 254-bit | ASM | 0.12 | 1.49 | 1.12 | 17.53 | 11.8 | 0.28 | 4.08 | 3.44 | 23.57 | 9,722 | 6,176 | 16,076 | 15,898 |
|  | C | 0.18 | 1.74 | 1.02 | 17.40 | 10.0 | 0.35 | 4.96 | 4.01 | 24.01 | 11,877 | 7,550 | 19,427 | 19,509 |
| 446-bit |  | 0.20 | 3.79 | 2.25 | 34.67 | 9.1 | 0.38 | 10.74 | 8.57 | 48.90 | 42,857 | 23,137 | 65,994 | 65,958 |
| 638-bit |  | 0.27 | 6.82 | 3.83 | 52.33 | 7.7 | 0.51 | 18.23 | 14.93 | 77.11 | 98,044 | 51,351 | 149,395 | 153,713 |
| iPad 2 (ARM v7) Apple A5 Cortex-A9 at 1.0 GHz [This work] |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Field | Lang | Operation Timing [ $\mu s$ ] |  |  |  |  |  |  |  |  |  |  |  |  |
| Size |  | $a$ | $m$ | $r$ | 2 | I/M | a | $\tilde{m}$ | $\tilde{s}$ | 亿 | ML | FE | O-A (a) | $\mathrm{O}-\mathrm{A}(\mathrm{p})$ |
| 254-bit | C | 0.16 1.28 <br> 0.16 2.92 <br> 0.20 5.58 |  | 0.93 | 13.44 | 10.5 | 0.25 | 3.48 | 2.88 | 19.19 | 8,338 | 5,483 | 14,604 | 13,821 |
| 446-bit |  |  |  | 1.62 | 27.15 | 9.3 | 0.26 | 8.03 | 6.46 | 37.95 | 32,087 | 17,180 | 49,365 | 49,267 |
| 638-bit |  |  |  | 2.92 | 43.62 | 7.8 | 0.34 | 15.07 | 12.09 | 64.68 | 79,056 | 40,572 | 119,628 | 123,410 |
| Galaxy Nexus (ARM v7) TI OMAP 4460 Cortex-A9 at 1.2 GHz [This work] |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Field Size | Lang | Operation Timing [ $\mu s$ ] |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | $a$ | $m$ | $r$ | $\imath$ | I/M | $\tilde{a}$ | $\tilde{m}$ | $\tilde{s}$ | \| i | ML | FE | O-A (a) | $\mathrm{O}-\mathrm{A}(\mathrm{p})$ |
| 254-bit | ASM | 0.05 | 0.93 | 0.55 | 9.42 | 10.1 | 0.10 | 2.46 | 2.07 | 13.79 | 6,147 | 3,758 | 10,573 | 9,905 |
|  | C | 0.07 | 0.98 | 0.53 | 9.62 | 9.8 | 0.13 | 2.81 | 2.11 | 14.05 | 6,859 | 4,382 | 11,839 | 11,241 |
| 446-bit |  | 0.12 | 2.36 | 1.27 | 23.08 | 9.8 | 0.22 | 6.29 | 5.17 | 32.27 | 25,792 | 13,752 | 39,886 | 39,544 |
| 638-bit |  | 0.19 | 4.87 | 3.05 | 38.45 | 7.9 | 0.45 | 12.20 | 10.39 | 56.78 | 65,698 | 33,658 | 99,356 | 99,466 |
| NVidia Tegra 2 (ARM v7) Cortex-A9 at 1.0 GHz [1] |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Field | Lang | Operation Timing [ $\mu s$ ] |  |  |  |  |  |  |  |  |  |  |  |  |
| Size |  | $a$ | $m$ | $r$ | $i$ | I/M | $\tilde{a}$ | $\tilde{m}$ | $\tilde{s}$ | \| i | ML | FE | O-A (a) | $\mathrm{O}-\mathrm{A}(\mathrm{p})$ |
| 254-bit | C | 0.67 | 1.72 | n/a | 18.35 | 10.7 | 1.42 | 8.18 | 5.20 | 26.61 | 26,320 | 24,690 | 51,010 | 55,190 |
| 446-bit |  | 1.17 | 4.01 | n/a | 35.85 | 8.9 | 2.37 | 17.24 | 10.84 | 54.23 | 97,530 | 86,750 | 184,280 | 195,560 |
| 638-bit |  | 1.71 | 8.22 | n/a | 56.09 | 6.8 | 3.48 | 31.81 | 20.55 | 91.92 | 236,480 | 413,370 | 649,850 | 768,060 |

- Loop unrolling: since the maximum number of bits of the operands are known, it makes sense to unroll all loops in order to provide us the ability to avoid conditional branches (which basically eliminates branch prediction misses in the pipeline), reorder the instructions, and insert carry propagate codes at desired points.
- Instruction re-ordering: by careful reordering of non-dependent instructions (in terms of data and processing units), it is possible to minimize the number of pipeline stalls and therefore execute the code faster. Two of the most frequent multi-cycle instructions used in our code are word multiplication and memory reads. Each 32-bit word multiplication takes between 3 and 6 cycles and each memory read needs 2 cycles. By applying loop unrolling, it is possible to load the data required for the next multiplication while the pipeline is performing the current multiplication. Also, lots of register cleanups and carry propagation codes are performed while the pipeline is doing a multiplication.
- Register allocation: all of the available registers were used extensively in order to eliminate the need to access memory for fetching the operands or store partial results. This improves overall performance considerably.
- Multiple stores: ARM processors are capable of loading and storing multiple words from or to the memory by one instruction. By storing the final result at once instead of writing a word back to memory each time when a new result is ready, we minimize the number of memory access instructions.

Also, we do some register clean-ups (cost-free) when the pipeline is performing the multiple store instruction. It is worth mentioning that while it was possible to write 8 words at once, only 4 words are written to memory at each time because the available non-dependent instructions to re-order after the multiple store instruction are limited.

Table 3 includes our measurements of pairing computation times using our assembly implementation alongside the results for the C implementation. We find that the BN-254 pairing using hand-optimized assembly code is roughly $20 \%$ faster than the C implementation. In all cases, the projective pairing benefits more than the affine pairing, because we did not hand-optimize the inversion routine in assembly.

### 6.2 Affine vs. Homogeneous Coordinates

Acar et al. [1] assert that on ARM processors, small inversion to multiplication $(\mathrm{I} / \mathrm{M})$ ratios over $\mathbb{F}_{q}$ render it more efficient to compute a pairing using affine coordinates. If we are using a prime $q$ congruent to $3 \bmod 8$, then compared to a projective doubling step, an affine doubling step costs an extra $\tilde{i}$ and an unreduced multiplication, and saves $5 \tilde{s}_{u}+3 \tilde{r}+16.5 \tilde{a}+2 m$. Compared to a first doubling, it costs an extra $\tilde{i}$ and saves $2 \tilde{s}_{u}+2 \tilde{r}+6.5 \tilde{a}+2 m$. An addition step costs an extra $\tilde{i}$ and saves $8 \tilde{m}_{u}+\tilde{s}_{u}+7 \tilde{r}+3 \tilde{a}+2 m$; and a dense-sparse multiplication needs an additional $6 m$ and saves $3 \tilde{m}_{u}+3 \tilde{r}+0.5 \tilde{a}$. Thus, the difference between an affine and projective pairing is $70 i-659 m_{u}-396 r-4417 a$.

From Table 3, we observe that the two pairings are roughly equal in performance at about the 446-bit field size. At this field size, we have $m \approx 1.5 r$ and $m \approx 15 a$. Plugging these estimates into the expression $70 i-659 m_{u}-396 r-4417 a$, we find that an affine pairing is expected to be faster than a projective pairing whenever the I/M ratio in the base field falls below about 10.0. The results of Table 3 indicate that our I/M ratios cross this point slightly above 446 bits. We observe that both affine and projective pairings achieve similar performance in our implementation on ARM processors, with an advantage in the range of $1 \%-6 \%$ for projective coordinates on BN-254, and a slight advantage for affine coordinates on BN-638, with slightly better results for projective coordinates on BN-446. These results overturn those of Acar et al. [1] which show too much advantage in favor of affine coordinates (well above $5 \%$ ). However, there are situations in which affine coordinates would be preferable even at the 128-bit security level (e.g., products of pairings). Different conclusions may hold for assembly-optimized variants at higher security levels, which are not included in our analysis.

## 7 Conclusions

In this paper, we present high speed implementation results of the Optimal-Ate pairing on BN curves for different security levels. We extend the concept of lazy
reduction to inversion in extension fields and optimize the sparse multiplication algorithm in the degree 12 extension. Our work indicates that D-type and M-type twists achieve equivalent performance for point/line evaluation computation, with only a very slight advantage in favor of D-type when computing sparse multiplications. In addition, we include an efficient method from [7] to perform final exponentiation and reduce its computation time. Finally, we measure the Optimal-Ate pairing over BN curves on different ARM-based platforms and compare the timing results to the leading ones available in the open literature. Our timing results are over three times faster than the previous fastest results appearing in [1]. Although the authors in [1] find affine coordinates to be faster on ARM in all cases, based on our measurements we conclude that homogeneous projective coordinates are unambiguously faster than affine coordinates for O-Ate pairings at the 128-bit security level when higher levels of optimization are used.

Acknowledgments. The authors would like to thank the reviewers for their comments. This work has been supported by NSERC CRD Grant CRDPJ 405857-10 and an Ontario Ministry of Research and Innovation Early Researcher Award.

## References

1. Acar, T., Lauter, K., Naehrig, M., Shumow, D.: Affine Pairings on ARM. IACR Cryptology ePrint Archive, 2011:243 (2011), Pairing 2012 (to appear)
2. Aranha, D.F., Gouvêa, C.P.L.: RELIC is an Efficient LIbrary for Cryptography, http://code.google.com/p/relic-toolkit/
3. Aranha, D.F., Karabina, K., Longa, P., Gebotys, C.H., López, J.: Faster Explicit Formulas for Computing Pairings over Ordinary Curves. In: Paterson, K.G. (ed.) EUROCRYPT 2011. LNCS, vol. 6632, pp. 48-68. Springer, Heidelberg (2011)
4. Barreto, P.S.L.M., Naehrig, M.: Pairing-Friendly Elliptic Curves of Prime Order. In: Preneel, B., Tavares, S. (eds.) SAC 2005. LNCS, vol. 3897, pp. 319-331. Springer, Heidelberg (2006)
5. Benger, N., Scott, M.: Constructing Tower Extensions of Finite Fields for Implementation of Pairing-Based Cryptography. In: Hasan, M.A., Helleseth, T. (eds.) WAIFI 2010. LNCS, vol. 6087, pp. 180-195. Springer, Heidelberg (2010)
6. Beuchat, J.-L., González-Díaz, J.E., Mitsunari, S., Okamoto, E., RodríguezHenríquez, F., Teruya, T.: High-speed software implementation of the optimal ate pairing over Barreto-Naehrig curves. In: Joye, et al. (eds.) [11], pp. 21-39
7. Fuentes-Castañeda, L., Knapp, E., Rodríguez-Henríquez, F.: Faster Hashing to $\mathbb{G}_{2}$. In: Miri, A., Vaudenay, S. (eds.) SAC 2011. LNCS, vol. 7118, pp. 412-430. Springer, Heidelberg (2012)
8. Pereira Geovandro, C.C.F., Simplício Jr., M.A., Naehrig, M., Barreto, P.S.L.M.: A family of implementation-friendly BN elliptic curves. Journal of Systems and Software 84(8), 1319-1326 (2011)
9. IEEE Std. 1363-3/D1. Draft Standard for Identity-based public key cryptography using pairings (January 2008)
10. Iyama, T., Kiyomoto, S., Fukushima, K., Tanaka, T., Takagi, T.: Efficient Implementation of Pairing on BREW Mobile Phones. In: Echizen, I., Kunihiro, N., Sasaki, R. (eds.) IWSEC 2010. LNCS, vol. 6434, pp. 326-336. Springer, Heidelberg (2010)
11. Joye, M., Miyaji, A., Otsuka, A. (eds.): Pairing 2010. LNCS, vol. 6487. Springer, Heidelberg (2010)
12. Lauter, K., Montgomery, P.L., Naehrig, M.: An analysis of affine coordinates for pairing computation. In: Joye, et al. (eds.) [11], pp. 1-20
13. Longa, P.: High-Speed Elliptic Curve and Pairing-Based Cryptography. PhD thesis, University of Waterloo (April 2011), http://hdl.handle.net/10012/5857
14. Miller, V.S.: The Weil pairing, and its efficient calculation. J. Cryptology 17(4), 235-261 (2004)
15. Naehrig, M., Niederhagen, R., Schwabe, P.: New Software Speed Records for Cryptographic Pairings. In: Abdalla, M., Barreto, P.S.L.M. (eds.) LATINCRYPT 2010. LNCS, vol. 6212, pp. 109-123. Springer, Heidelberg (2010)
16. Scott, M.: A note on twists for pairing friendly curves. Personal webpage (2009), ftp://ftp.computing.dcu.ie/pub/resources/crypto/twists.pdf
17. Scott, M.: On the Efficient Implementation of Pairing-Based Protocols. In: Chen, L. (ed.) IMACC 2011. LNCS, vol. 7089, pp. 296-308. Springer, Heidelberg (2011)
18. Vercauteren, F.: Optimal pairings. IEEE Transactions on Information Theory $56(1), 455-461$ (2010)

[^0]:    * This work was partially carried out while this author was a postdoctoral researcher in the Department of Combinatorics and Optimization at the University of Waterloo.

