No Silver Bullet: Optimized Montgomery Multiplication on Various 64-bit ARM Platforms

Hwajeong Seo[1][0000−0003−0069−9061], Pakize Sanal[2], Reza Azarderakhsh[2,3], and Wai-Kong Lee[4]

[1]IT Department, Hansung University, Seoul, South Korea, hwajeong84@gmail.com
[2]Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, FL, USA, {psanal2018,razarderakhsh}@fau.edu
[3]PQSecure Technologies, LLC
[4]Department of Computer Engineering, Gachon University, Seongnam, Incheon (13120), Korea, waikonglee@gachon.ac.kr

Abstract. In this paper, we firstly presented optimized implementations of Montgomery multiplication on 64-bit ARM processors by taking advantages of Karatsuba algorithm and efficient multiplication instruction sets for ARM64 architectures. The implementation of Montgomery multiplication can improve the performance of (pre-quantum and post-quantum) public key cryptography (e.g. CSIDH, ECC, and RSA) implementations on ARM64 architectures, directly. Last but not least, the performance of Karatsuba algorithm does not ensure the fastest speed record on various ARM architectures, while it is determined by the clock cycles per multiplication instruction of target ARM architectures. In particular, recent Apple processors based on ARM64 architecture show lower cycles per instruction of multiplication than that of ARM Cortex-A series. For this reason, the schoolbook method shows much better performance than the sophisticated Karatsuba algorithm on Apple processors. With this observation, we can determine the proper approach for multiplication of cryptography library (e.g. Microsoft-SIDH) on Apple processors and ARM Cortex-A processors.

Keywords: Montgomery Multiplication · ARM64 Processor · Public Key Cryptography · Software Implementation.

1 Introduction

The modular reduction is the fundamental building block of conventional public key cryptography (e.g. RSA [15], El-Gamal [5], and ECC [10, 13]) to post-quantum cryptography (e.g. RLWE [12], SIDH [4], and CSIDH [2]). One of the most well-known modular reduction techniques is Montgomery algorithm [14]. This approach replaces the complicated division operation for the modular reduction in relatively simple multiplication operations. For this reason, efficient implementations of Montgomery multiplication on target processors have been

actively studied. In this paper, we firstly introduce the optimized Montgomery multiplication on ARM64 processors and show the impact on public key cryptography protocols (i.e. CSIDH)[1]. Furthermore, we found that recent Apple processors provide the multiplication instruction with very low latency. This nice feature leads to the new direction to implement the multiplication on Apple processors (i.e. simple schoolbook approach rather than sophisticated Karatsuba algorithm with additional routines). With this observation, we can improve the performance of cryptography libraries based on Karatsuba algorithm (e.g. Microsoft-SIDH) on recent Apple processors by replacing the multiplication to the schoolbook method[2].

1.1 Contribution

- **Optimized Montgomery multiplication on ARM64 architecture** This paper presents the optimized implementation of Montgomery multiplication by taking advantages of Karatsuba algorithm and efficient multiplication instruction sets of ARM64 architectures. The proposed implementation is evaluated on both ARM Cortex-A and Apple A processors for benchmarking tests.
- **Efficient implementation of CSIDH on ARM64 architecture** The implementation of CSIDH is accelerated through the proposed Montgomery multiplication. Full protocols of CSIDH-P511 are evaluated on 64-bit ARM Cortex-A and Apple A processors. In order to prevent the timing attack on the protocol, the implementation ensures the constant timing.
- **In-depth performance evaluation on ARM64 architecture** Depending on ARM64 architectures (i.e. ARM Cortex-A and Apple A), same program code leads to different performance results. We analyze the performance of Montgomery multiplication and CSIDH on both architectures. Finally, we show that Karatsuba algorithm is not the best solution across all ARM64 architectures, since it introduces a number of addition operations than school-book methods.

The remainder of the paper is structured as follows: We review the related work on Montgomery multiplication, Karatsuba algorithm, and ARM64 processors in Section 2. We present the optimized implementation of Montgomery multiplication on ARM64 processors and its use cases (i.e. CSIDH) in Section 3. In Section 4, we present results on various 64-bit ARM platforms (ARM Cortex-A and Apple A processors). We end with conclusions in Section 5.

2 Related Works

2.1 Montgomery Multiplication

Montgomery multiplication consists of multiplication and reduction parts. The multiplication can be implemented in different ways by altering the order of

[1] The proposed method is applicable to RSA, and ECC, as well.
[2] `https://github.com/microsoft/PQCrypto-SIDH/tree/master/src`

Algorithm 1 Montgomery Reduction

Require: An m-bit modulus M, Montgomery radix $R = 2^m$, two m-bit operands A and B, and m-bit pre-computed constant $M' = -M^{-1} \bmod R$
Ensure: Montgomery product ($Z = (A \cdot B) \cdot R^{-1} \bmod M$)
1: $T \leftarrow A \cdot B$
2: $Q \leftarrow T \cdot M' \bmod R$
3: $Z \leftarrow (T + Q \cdot M)/R$
4: **if** $Z \geq M$ **then** $Z \leftarrow Z - M$ **end if**
5: **return** Z

operands and intermediate results. The operand-scanning method performs a multiplication in a row-wise manner. This approach is suitable for processors with many registers to retain long intermediate results as many as possible. The alternative approach is the Comba (i.e. product-scanning) method [3]. Partial products are computed in a column-wise manner and only small number of registers is required to maintain the intermediate result. Furthermore, since all partial products of each word of the result are computed and added, consecutively, the final result word is obtained directly and no intermediate results have to be stored or loaded in the algorithm In [6], a hybrid-scanning method combining two aforementioned methods is presented. Afterward, several optimized implementations for multiplication were suggested by caching operands [7, 16]. However, the complexity of partial products for N-word multiplication is N^2.

Karatsuba algorithm computes a N-word multiplication with only three $N/2$-word partial products compared to four $N/2$-word that are required by aforementioned multiplication methods [9]. The number of partial products is estimated by $N^{log_2 3}$, which is a great improvement compared to N^2 of the standard multiplication. The previous multiplication on ARM64 processors mainly utilized the Karatsuba algorithm for optimal performance since it reduces the number of complicated multiplication.

Montgomery algorithm was firstly proposed in 1985 [14]. Montgomery algorithm avoids the division in modular multiplication by introducing simple shift operations (i.e. division by 2). Given two integers A and B and the modulus M, to compute the product $P = A \cdot B \bmod M$ in Montgomery method, operands A and B are firstly converted into Montgomery domain (i.e. $A' = A \cdot R \bmod M$ and $B' = B \cdot R \bmod M$). For efficient computations, Montgomery residue R is selected as a power of 2 and constant $M' = -M^{-1} \bmod 2^n$ is pre-computed. To compute the product, following three steps are conducted: (1) compute $T = A \cdot B$; (2) perform $Q = T \cdot M' \bmod 2^n$; (3) calculate $Z = \frac{(T+Q \cdot M)}{2^n}$ and (4) compute final reduction $Z \leftarrow Z - M$ if $Z \geq M$. Detailed descriptions of Montgomery reduction is available in Algorithm 1.

There are two approaches, including separated and interleaved ways, for Montgomery multiplication. The interleaved approach reduces the number of memory access for the intermediate result, but many implementations on ARM64 selected the separated approach. The separated implementation can employ the most optimal approach for multiplication and reduction operations each. In this

Implementation	Multiplication	Montgomery Reduction	Application
Liu et al. [11]	Karatsuba	–	–
Seo et al. [17]	Karatsuba	Product Scanning	SIKE
Seo et al. [19]	Karatsuba	Product Scanning	SIKE
Jalali et al. [8]	Operand Scanning	Operand Scanning	CSIDH, RSA, ECC
This work	Karatsuba	Karatsuba	CSIDH, RSA, ECC

Table 1. Comparison of Montgomery multiplication on 64-bit ARMv8 Cortex-A processors.

paper, we also selected the separated approach. Both multiplication and Montgomery reduction parts are optimized with Karatsuba algorithm. In particular, we utilized the Karatsuba based Montgomery reduction by [18].

In Table 1, the comparison of Montgomery multiplication on 64-bit ARMv8 Cortex-A processors is given. Previous works [11, 17, 19] target for Montgomery friendly prime and it's application is limited to only SIKE due to the special prime form. On the other hand, the proposed Montgomery multiplication is targeting for random prime and it's applications are CSIDH, RSA, and ECC.

2.2 64-bit ARMv8 Processors

ARMv8 is a 64-bit architecture for high-performance embedded applications. The 64-bit ARMv8 processors support both 32-bits (AArch32) and 64-bits (AArch64) architectures. It provides 31 general purpose registers which can hold 32-bit values in registers `w0-w30` or 64-bit values in registers `x0-x30`. ARMv8 processors started to dominate the smartphone market soon after the release in 2011 and nowadays they are widely used in various smart phones (e.g. iPhone and Samsung Galaxy series) and laptop (e.g. MacBook Air and MacBook Pro). Since the processor is used primarily in embedded systems, smart phones and laptop computers, efficient and compact implementations are of special interest in real world applications. ARMv8 processors support powerful 64-bit wise unsigned integer multiplication instructions. Proposed implementation of modular multiplication uses the AArch64 architecture and makes extensive use of the following multiply instructions:

- `MUL` (unsigned multiplication, low part):
  `MUL X0, X1, X2` computes `X0` $\leftarrow$ (`X1` $\times$ `X2`) mod 2^{64}.
- `UMULH` (unsigned multiplication, high part):
  `UMULH X0, X1, X2` computes `X0` $\leftarrow$ (`X1` $\times$ `X2`)$/2^{64}$.

The two instructions above are required to compute a full 64-bit multiplication of the form 128-bit $\leftarrow 64 \times 64$-bit, namely, the `MUL` instruction computes the lower 64-bit half of the product while `UMULH` computes the higher 64-bit half.

For the addition and subtraction operations, `ADDS` and `SUBS` instructions ensure 64-bit wise results, respectively. Detailed descriptions are as follows:

- `ADDS` (unsigned addition):
  `ADDS X0, X1, X2` computes {`CARRY,X0`} ← (`X1` + `X2`).
- `SUB` (unsigned subtraction):
  `SUBS X0, X1, X2` computes {`BORROW,X0`} ← (`X1` − `X2`).

Further details of ARMv8-A architecture can be found in official documents [1].

3 Proposed Implementations

3.1 Optimization of Montgomery multiplication

One of the most expensive operation for public key cryptography is modular multiplication. In this paper, we present the optimal modular multiplication implementation in the separated way for 64-bit ARM architectures.

First, the multi-precision multiplication is performed by following the Karatsuba algorithm. 2-level Karatsuba computations are performed for 512-bit multiplication on 64-bit ARM architectures (i.e. 128 → 256 → 512) [17]. Karatsuba's method reduces a multiplication of two N-limb operands to three multiplications, which have a length of $\frac{N}{2}$-limb. These three half-size multiplications can be performed with any multiplication techniques that we covered before (e.g., operand scanning method or product scanning method). Taking the multiplication of N-limb operand A and B as an example, we represent the operands as $A = A_H \cdot 2^{\frac{N}{2}} + A_L$ and $B = B_H \cdot 2^{\frac{N}{2}} + B_L$. The multiplication $P = A \cdot B$ can be computed according to the following equation by using additive Karatsuba's method:

$$A_H \cdot B_H \cdot 2^n + [(A_H + A_L)(B_H + B_L) - A_H \cdot B_H - A_L \cdot B_L] \cdot 2^{\frac{n}{2}} + A_L \cdot B_L \quad (1)$$

We further optimized the memory access by using general purpose registers to retain operands as many as possible since the access speed of register is much faster than that of memory. In particular, Karatsuba multiplication needs to update operands but these operands are used in following computations. In this case, we keep operands in general purpose registers to avoid frequent memory loading operations where the memory access incurs read-write-dependency problems.

Multiplication and reduction operations are implemented in one function to avoid the function call and register `push`/`pop` instructions. Furthermore, the intermediate result of 512-bit multiplication is 1024-bit wise and this can be stored in 16 general purpose registers (i.e. 16 = 1024/64) in the ideal case. By maintaining the whole intermediate result in general purpose registers, 16 memory load and 16 memory store operations are optimized away.

However, part of intermediate results are stored in `STACK` memory since Karatsuba approach requires a number of registers to retain operands and intermediate results. The stored result is directly used in the following modular reduction operation.

For the computation of Montgomery reduction, the product ($Q \leftarrow T \cdot M' \bmod R$) is performed (Step 2 of Algorithm 1) in ordinary way. Afterward, the product

Modulus M	Quotient Q	Temporal registers	Constant M'
8	4	15	1

Table 2. Register utilization for Montgomery reduction on ARM64.

$Q \cdot M$ is computed in a hybrid way (Step 3) [18]. The complexity of N-word original Montgomery reduction is N^2+N word-wise multiplications, while the hybrid Montgomery reduction is $\frac{7N^2}{8} + N$ word-wise multiplications. In particular, the hybrid Montgomery reduction consists of two 256-bit Montgomery reduction in product-scanning approach and two 256-bit (1-level) Karatsuba multiplication operations (i.e. $128 \rightarrow 256$).

28 out of 31 registers are utilized for 512-bit Montgomery reduction on the ARM64 architecture. The detailed register utilization is given in Table 2. When we perform the hybrid Montgomery reduction ($\frac{T+Q \cdot M}{2^n}$), computations ($Q_L \cdot M_L$ and $Q_H \cdot M_L$) are performed in sub-Montgomery reduction and others ($Q_L \cdot M_H$ and $Q_H \cdot M_H$) are performed in Karatsuba multiplication, where L and H represent lower and higher parts of operand.

In Algorithm 2, the optimized implementation of 256-bit sub-Montgomery reduction on ARM64 processors is given. Partial products of reduction are performed in the product-scanning way. Three ARM64 instructions (MUL, UMULH, and ADD) are mainly utilized for partial products. Since multiplication operations require 6 clock cycles in Cortex-A series, the utilization of result directly incurs pipeline stalls [17]. In Line 1 of Algorithm 2, the quotient (`Q0`) is generated with 64-bit wise. This is simply performed with single `mul` instruction. In Line 2∼5, the quotient (`Q0`) is directly utilized, which incurs pipeline stalls. In Line 6∼7, the result of multiplication (`T0` and `T1`), which is computed in Line 2∼3, is accumulated to the intermediate result. This approach avoids the read-write dependency. In Line 10, the register is initialized with `xzr` instruction and the carry is obtained in `C0` register. In Line 11, the quotient (`Q1`) is generated but it is not utilized directly. In particular, the accumulation step (Line 12∼15) is performed with partial products in previous steps. This does not incur the read-write dependency. Following computations (after Line 16 to end) are performed in the similar way (i.e. read-write dependency free).

After the sub-Montgomery reduction, the remaining part is performed with the Karatsuba algorithm [17]. The additive Karatsuba algorithm performs the addition on the operand. This updates operands which cannot be used again. In the proposed implementation, we cached the operand in registers and this avoids the memory access for the operand re-loading. Afterward, one sub-Montgomery reduction and one Karatsuba multiplication are performed. Lastly, the final reduction is performed in the masked way in order to ensure the constant timing implementation. Firstly the intermediate result is subtracted by the modulus. When the borrow bit is captured, it sets the masked modulus. Otherwise, the modulus is set to zero. The result is subtracted by the masked modulus.

In conclusion, the proposed 512-bit Montgomery multiplication is performed in two steps as follows:

Algorithm 2 Optimized implementation of 256-bit sub-Montgomery reduction on ARM64 processors.

Input: Modulus (M0-M3), intermediate results (C0-C3), constant (M_INV), temporal registers (T0-T3).
Output: Intermediate results (C0-C3), quotient (Q0-Q3).

```
1: mul Q0, C0, M_INV                          { Q0 ← C0 × M' }

2: mul T0, Q0, M0                             { T ← Q × M }
3: umulh T1, Q0, M0
4: mul T2, Q0, M1
5: umulh T3, Q0, M1

6: adds C0, C0, T0                            { Accumulation of intermediate result }
7: adcs C1, C1, T1
8: adcs C2, C2, xzr
9: adcs C3, C3, xzr
10: adc C0, xzr, xzr

11: mul Q1, C1, M_INV                         { Q1 ← C1 × M' }
12: adds C1, C1, T2
13: adcs C2, C2, T3
14: adcs C3, C3, xzr
15: adc C0, C0, xzr

16: ...                                       { Omit }

17: mul T2, Q3, M3
18: umulh T3, Q3, M3
19: adds C1, C1, T0                           { Accumulation of intermediate result }
20: adcs C2, C2, T1
21: adc C3, C3, xzr
22: adds C2, C2, T2
23: adcs C3, C3, T3
```

Company	ARM	
Platform	Odroid-C2	Raspberry-pi4
Core	Cortex-A53(@1.5GHz)	Cortex-A72(@1.5GHz)
OS	Ubuntu 16.04	Ubuntu 20.10
Released	2014	2015
Revision	ARMv8.0-A	ARMv8.0-A
Decode	2-wide	3-wide
Pipeline depth	8	15
Out-of-order	×	O
Branch prediction	Conditional	O
Execution ports	2	8

Table 3. Comparison of ARMv8-A cores on ARM Cortex-A processors.

Company	Apple		
Platform	iPad mini5	iPhone SE2	iPhone12 mini
Core	A12(Vortex@2.49GHz)	A13(Lightning@2.65GHz)	A14(Firestorm@3.10GHz)
OS	iPadOS 14.4	iOS 14.4	iOS 14.4
Released	2018	2019	2020
Revision	ARMv8.3-A	ARMv8.4-A	ARMv8.4-A
Decode	7-wide	8-wide	8-wide
Pipeline depth	16	16	–
Out-of-order	O	O	–
Branch prediction	–	–	–
Execution ports	13	13	–

Table 4. Comparison of ARMv8-A cores on Apple A processors.

$2-level\ Karatsuba\ multiplication\ \rightarrow\ hybrid\ reduction\ (1-level\ Karasuba)$

3.2 Acceleration of Public Key Cryptography

The proposed implementation of Montgomery multiplication is efficiently optimized. We can directly apply the proposed Montgomery multiplication to the CSIDH library by [8]. CSIDH is a variant of isogeny-based cryptography that offers (conjecturally) post-quantum secure non-interactive key exchange with tiny public keys and practical performance.

We checked the improved CSIDH implementation based on the proposed Montgomery multiplication passed the CSIDH tests and public-key validations. Furthermore, the conventional public key cryptography based on random prime (RSA and ECC) can also take advantages of the proposed method.

Implementation	Odroid-C2 Timing [cc×10^6]			Raspberry-pi4 Timing [cc×10^6]		
	[8]	Opt	[8]/Opt	[8]	Opt	[8]/Opt
Montgomery multiplication	1,309 cc	1,044 cc	1.25	973 cc	792 cc	1.23
Alice key generation	14,374	12,392	1.16	11,892	9,864	1.21
Bob key generation	14,386	12,392	1.16	12,098	9,916	1.22
Validation of Bob's key	58	50	1.16	43	35	1.21
Validation of Alice's key	58	50	1.16	43	35	1.21
Alice shared key generation	14,252	12,628	1.13	11,570	10,099	1.15
Bob shared key generation	14,544	12,555	1.16	12,453	10,114	1.23
Alice total computations	28,684	25,070	1.14	23,504	19,998	1.18
Bob total computations	28,988	24,998	1.16	24,594	20,065	1.23

Table 5. Comparison of clock cycles ($\times 10^6$) for Montgomery multiplication (for 512-bit) and (constant-time) CSIDH-P511 on 64-bit Odroid-C2 and Raspberry-pi4.

4 Evaluation

The proposed implementation is evaluated on the various ARM64 architectures, which is largely divided into ARM Cortex-A and Apple A series. Detailed specifications for each processor are given in Table 3 and 4.

In Table 5 and 6, the comparison of clock cycles for 512-bit Montgomery multiplication and constant-time CSIDH-P511 on 64-bit ARM architectures is given. Proposed implementations of 512-bit modular multiplication achieved performance enhancements than the school-book method [8] by 1.25× and 1.23× on Odroid-C2 and Raspberry-pi4, respectively. Since the approach is a generic method, we can apply the proposed method to larger operand sizes without difficulties. Furthermore, Montgomery multiplication is the fundamental operation in PKC. For the case study, we ported the implementation of Montgomery multiplication to CSIDH implementation. The performance of key exchange is improved by 1.16× and 1.23× than previous works on Odroid-C2 and Raspberry-pi4, respectively.

On the other hand, proposed implementations on Apple platforms show the opposite performance result. The schoolbook method based 512-bit modular multiplication achieved performance enhancements than proposed method by 0.71×, 0.65× and 0.69× on iPad mini5, iPhone SE2, and iPhone12 mini, respectively. The performance of key exchange is degraded by 0.71×, 0.64×, and 0.72× than previous works on Odroid-C2 and Raspberry-pi4, respectively.

In Table 7, the comparison of cycles per instruction on ARM64 is given. The timing is measured with the average cycles after performing 1,000 times of each iteration without read-write dependency. The timing also includes function call and push/pop instructions. In ARM Cortex-A series, ratios of $\frac{MUL}{ADD}$ and $\frac{UMULH}{ADD}$ are 2.62∼4.34 and 4.65∼6.20 for Odroid-C2 and Raspberry-pi4 boards, respectively. This shows that the multiplication operation is more expensive than the addition operation on the target ARM Cortex-A architecture. For this reason,

Implementation	iPad mini5			iPhone SE2			iPhone12 mini		
	Timing [cc×10^6]			Timing [cc×10^6]			Timing [cc×10^6]		
	[8]	Opt	[8]/Opt	[8]	Opt	[8]/Opt	[8]	Opt	[8]/Opt
Montgomery multiplication	167 cc	233 cc	0.71	154 cc	235 cc	0.65	150 cc	214 cc	0.69
Alice key generation	1,210	1,694	0.71	1,086	1,692	0.64	1,131	1,548	0.73
Bob key generation	1,212	1,681	0.72	1,086	1,693	0.64	1,132	1,557	0.73
Validation of Bob's key	8	11	0.71	7	11	0.65	7	10	0.72
Validation of Alice's key	8	11	0.71	7	11	0.66	7	10	0.72
Alice shared key generation	1,216	1,705	0.71	1,090	1,690	0.64	1,127	1,572	0.72
Bob shared key generation	1,214	1,681	0.72	1,084	1,706	0.64	1,135	1,546	0.73
Alice total computations	2,434	3,410	0.71	2,183	3,393	0.64	2,265	3,130	0.72
Bob total computations	2,433	3,373	0.72	2,177	3,409	0.64	2,274	3,113	0.73

Table 6. Comparison of clock cycles ($\times 10^6$) for Montgomery multiplication (for 512-bit) and (constant-time) CSIDH-P511 on 64-bit iPad mini5, iPhone SE2, and iPhone12 mini.

Platform	`MUL`	`UMULH`	`ADD`	`MUL/ADD`	`UMULH/ADD`
Odroid-C2	2.37	3.93	0.90	2.62	4.34
Raspberry-pi4	3.02	4.03	0.64	4.65	6.20
iPad mini5	0.57	0.57	0.42	1.34	1.34
iPhone SE2	0.49	0.49	0.37	1.31	1.31
iPhone12 mini	0.55	0.51	0.37	1.47	1.38

Table 7. Comparison of cycles per instruction on ARM64.

Karatsuba algorithm, which replaces multiplication operations into addition operations is effective on ARM Cortex-A series. On the other hand, ratios of $\frac{MUL}{ADD}$ and $\frac{UMULH}{ADD}$ on Apple A series are 1.34, 1.31, and 1.47∼1.38 for iPad mini5, iPhone SE2, iPhone12 mini, respectively. This is unique features of advanced ARM architectures. Unlike previous architectures, multiplication operations are efficiently performed and the complexity between addition and multiplication is narrow. This leads to different conclusion (i.e. best implementation technique) in Apple A series.

In Table 8, the number of instructions for Montgomery multiplication methods is given. The schoolbook method (i.e. operand-scanning) requires 297 addition, 265 multiplication, and 134 memory access instructions, respectively. Compared with the Karatsuba algorithm, 225 addition instructions are optimized away but it requires 51 multiplication and 64 memory access instructions, more than the Karatsuba approach. Due to the efficient multiplication instruction on Apple A processors, reducing the number of multiplication by sacrificing the addition operation is not effective. For this reason, even though processors based on ARM64 architecture, the implementation technique should be different depending on cycles per multiplication instruction. This observation is useful for optimization of public key cryptography on ARM64 architecture. For example, Montgomery multiplication of Microsoft-SIDH library is based on the Karatsuba algorithm. This library can be improved by using operand-scanning method on Apple A products.

Implementation	ADD/SUB	MUL/UMULH	LDR/STR
Schoolbook Method [8]	297	265	134
Karatsuba Algorithm	522	214	70

Table 8. Number of instructions for 512-bit Montgomery multiplication methods.

5 Conclusion

In this paper, we presented optimized Montgomery multiplication implementations for the 64-bit ARM Cortex-A processors. Proposed implementations utilized the Karatsuba algorithm and ARMv8-A specific instruction sets. This work shows that proposed implementations on ARM Cortex-A platforms are more efficient than previous works.

However, the platform with low multiplication latency (e.g. Apple A processors) achieved the better performance with the schoolbook method. This is because the evaluation of previous works is usually conducted on ARM Cortex-A processors. With the observation on this paper, the implementation should be evaluated on various ARM platforms for fair comparison and practicality.

The obvious future work is improving the Microsoft-SIDH library on Apple A processors by utilizing the schoolbook method or other approaches (i.e. product-scanning and hybrid-scanning). Furthermore, we will investigate the multiplication method for various integer lengths. Lastly, the proposed implementation will be the public domain and other cryptography engineers can directly use them for their cryptography applications.

References

1. ARM: ARM architecture reference manual: ARMv8, for ARMv8-A architecture profile (2020)
2. Castryck, W., Lange, T., Martindale, C., Panny, L., Renes, J.: CSIDH: an efficient post-quantum commutative group action. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 395–427. Springer (2018)
3. Comba, P.G.: Exponentiation cryptosystems on the IBM PC. IBM systems journal **29**(4), 526–538 (1990)
4. Costello, C., Longa, P., Naehrig, M.: Efficient algorithms for supersingular isogeny diffie-hellman. In: Annual International Cryptology Conference. pp. 572–601. Springer (2016)
5. ElGamal, T.: A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE transactions on information theory **31**(4), 469–472 (1985)
6. Gura, N., Patel, A., Wander, A., Eberle, H., Shantz, S.C.: Comparing elliptic curve cryptography and RSA on 8-bit CPUs. In: International workshop on cryptographic hardware and embedded systems. pp. 119–132. Springer (2004)
7. Hutter, M., Wenger, E.: Fast multi-precision multiplication for public-key cryptography on embedded microprocessors. In: International Workshop on Cryptographic Hardware and Embedded Systems. pp. 459–474. Springer (2011)

8. Jalali, A., Azarderakhsh, R., Kermani, M.M., Jao, D.: Towards optimized and constant-time CSIDH on embedded devices. In: International Workshop on Constructive Side-Channel Analysis and Secure Design. pp. 215–231. Springer (2019)
9. Karatsuba, A.: Multiplication of multidigit numbers on automata. In: Soviet physics doklady. vol. 7, pp. 595–596 (1963)
10. Koblitz, N.: Elliptic curve cryptosystems. Mathematics of computation **48**(177), 203–209 (1987)
11. Liu, Z., Järvinen, K., Liu, W., Seo, H.: Multiprecision multiplication on ARMv8. In: 2017 IEEE 24th Symposium on Computer Arithmetic (ARITH). pp. 10–17. IEEE (2017)
12. Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 1–23. Springer (2010)
13. Miller, V.S.: Use of elliptic curves in cryptography. In: Conference on the theory and application of cryptographic techniques. pp. 417–426. Springer (1985)
14. Montgomery, P.L.: Modular multiplication without trial division. Mathematics of computation **44**(170), 519–521 (1985)
15. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM **21**(2), 120–126 (1978)
16. Seo, H., Kim, H.: Multi-precision multiplication for public-key cryptography on embedded microprocessors. In: International Workshop on Information Security Applications. pp. 55–67. Springer (2012)
17. Seo, H., Liu, Z., Longa, P., Hu, Z.: SIDH on ARM: faster modular multiplications for faster post-quantum supersingular isogeny key exchange. IACR Transactions on Cryptographic Hardware and Embedded Systems pp. 1–20 (2018)
18. Seo, H., Liu, Z., Nogami, Y., Choi, J., Kim, H.: Hybrid Montgomery reduction. ACM Transactions on Embedded Computing Systems (TECS) **15**(3), 1–13 (2016)
19. Seo, H., Sanal, P., Jalali, A., Azarderakhsh, R.: Optimized implementation of SIKE round 2 on 64-bit ARM Cortex-A processors. IEEE Transactions on Circuits and Systems I: Regular Papers **67**(8), 2659–2671 (2020)