NEON-SIDH: Efficient Implementation of Supersingular Isogeny Diffie-Hellman Key Exchange Protocol on ARM

Reza Azarderakhsh[1], Brian Koziel[1], Amir Jalali[1], Mehran Mozaffari Kermani[2] and David Jao[3]

[1] Computer Engineering Department
[2] Electrical and Microelectrical Engineering Department
Rochester Institute of Technology, Rochester, NY 14623, USA
{rxaeec, bck650, aj2628@, mmkeme}@rit.edu.
[3]Combinatorics and Optimization, University of Waterloo, CANADA
djao@math.uwaterloo.ca

Abstract. In this paper, we investigate the efficiency of implementing a post-quantum key exchange protocol over isogenies (PQCrypto 2011) on ARM-powered embedded platforms. This work proposes to employ new primes to speed up constant-time finite field arithmetic and perform isogenies quickly. Montgomery multiplication and reduction are employed to produce a speedup of 3 over the GNU Multiprecision Library. We analyze the recent projective isogeny formulas presented in Costello et al., ePrint 2016/413 and conclude that affine isogeny formulas are much faster in ARM devices. We provide fast affine SIDH libraries over 512, 768, and 1024-bit primes. We provide timing results for emerging embedded ARM platforms using the ARMv7A architecture for 85-, 128-, and 170-bit quantum security levels. Our assembly-optimized arithmetic cuts the computation time for the protocol by 50% in comparison to our portable C implementation and performs approximately 3 times faster than the only other ARMv7 results found in the literature. The goal of this paper is to show that isogeny-based cryptosystems can be implemented further and be used as an alternative to classical cryptosystems on embedded devices.

Keywords: Elliptic curve cryptography, post-quantum cryptography, isogeny-based cryptosystems, ARM embedded processors, finite-field arithmetic, assembly

1 Introduction

Post-quantum cryptography (PQC) refers to research on cryptographic primitives (usually public-key cryptosystems) that are not efficiently breakable using quantum computers more than classical computer architectures. Notably, Shor's algorithm [1] can be efficiently performed with a quantum computer to break standard Elliptic Curve Cryptography (ECC) and RSA cryptosystems. There are some alternatives to be secure against quantum computer threats like the McEliece cryptosystem, lattice-based cryptosystems, code-based cryptosystems, multivariate public key cryptography, and the like. Recently, in [2], [3], [4], and [5], efficient implementations of quantum-safe cryptosystems have been implemented on embedded systems. None of these works consider making the current cryptosystems based on elliptic curves to be quantum-resistant. Hence, they introduce and implement new cryptosystems with different performance metrics.

To avoid quantum computing concerns, an elliptic curve based alternative to Elliptic Curve Diffie-Hellman (ECDH) which is not susceptible to Shor's attack is the Supersingular Isogeny Diffie-Hellman (SIDH) key exchange. Isogeny computations construct an algebraic map between elliptic curves, which appear resistant to quantum attacks. Thus, this system improves upon traditional ECC and is a strong candidate for quantum-resistant cryptography [6]. Faster isogeny

constructions would speed up such cryptosystems, increase the viability of existing proposals, and make new designs feasible. In [6], the use of isogenies to create new and existing cryptographic protocols with quantum-resistance is presented. However, their implementations on emerging embedded devices have not been investigated yet. It is expected that the use of mobile devices, such as smartphones, tablets, and emerging embedded systems, will become further widespread in the coming years for increasingly sensitive applications. In this work, we further investigate the applicability of advances in theoretical quantum-resistant algorithms on real-world applications by several efficient implementations on emerging embedded systems. Our goal is to improve the performance of isogeny-based cryptosystems to the point where deployment is practical.

In a recent announcement at PQC 2016 [7], NIST announced a preliminary plan to start the gradual transition to quantum-resistant protocols. As such, there is a tremendous need to discover and implement new proposed methods that are resistant to both classical computers and quantum computers. NIST will evaluate these PQC schemes based on security, speed, size, and tunable parameters. Isogeny-based cryptography provides a suitable replacement for standard ECC or RSA protocols because it provides small key sizes, provides forward secrecy, and has a Diffie-Hellman key exchange available. A protocol is forward secure if the compromise of long-term keys does not compromise past session keys [8]. Furthermore, isogeny-based cryptography utilizes standard ECC point multiplication schemes, but take it a step further by computing large isogenies to provide quantum-resistance.

Our contribution:

- We provide efficient libraries for the key exchange protocol presented in [6] to highly optimized C and ASM.
- We present fast and secure prime candidates for 85-bit, 128-bit, and 170-bit quantum security levels.
- We provide hand-optimized finite field arithmetic computations over various ARM-powered processors to produce constant-time arithmetic that is three times as fast as GMP's.
- We analyze the effectiveness of projective [9] and affine [10] isogeny schemes.
- We provide implementation results for embedded devices running a Cortex-A8 and a Cortex-A15. For the latter, an entire quantum-resistant key exchange with 85-bit quantum security operates in approximately a tenth of a second. Further, our Cortex-A15 assembly optimized results are 3 times faster than [11], the fastest results available in the literature.

2 SIDH Protocol

This serves as a quick introduction to the Supersingular Isogeny Diffie-Hellman key exchange. For a full mathematical background of the protocol, we point the reader to the original works proposing it in [6,10] or [12] for a complete look at elliptic curve theory.

2.1 Isogenies on Elliptic Curves

Isogeny-based cryptography utilizes unique algebraic maps between elliptic curves that satisfy group homomorphism. The idea of isogeny-based cryptography was first introduced by Rostovtsev and Stolbunov in [13]. This original work detailed a Diffie-Hellman-like cryptosystem based on the difficulty of computing isogenies between ordinary elliptic curves. Originally, this was thought to be quantum-resistant until Childs, Jao, and Stolbunov [14] discovered a quantum algorithm that could compute isogenies between ordinary curves in subexponential time. This

algorithm assumes the Generalized Riemann Hypothesis and abuses the commutative group structure of the endomorphism ring of isogenies between ordinary curves. To defend against this attack, Jao and De Feo adapted the isogeny-based key exchange protocol to be based on the difficulty of computing isogenies between supersingular elliptic curves, which does not have a commutative endomorphism ring [6]. There is currently no known quantum algorithm that can compute isogenies between supersingular elliptic curves in subexponential. Over a field of characteristic p, the best known attack is based on solving the claw problem with complexity $O(p^{1/4})$ and $O(p^{1/6})$ for classical and quantum computers, respectively [10].

For two elliptic curves over a finite field to be isogenous, they must have the same number of points [15] and have the same j-invariant. We define an isogeny over $\mathbb{F}_q$ to be $\phi : E \to E'$ as a non-constant rational map defined over $\mathbb{F}_q$ such that ϕ satisfies group homomorphism from $E(\mathbb{F}_q)$ to $E'(\mathbb{F}_q)$. The degree of an isogeny, $\deg\{\phi\}$, is its degree as an algebraic map. We are particularly interested in computing isogenies of high degree.

A curve's endomorphism ring is defined as the ring of all isogenies from a curve to itself, under point addition and functional composition. A curve is considered supersingular if this endomorphism ring has $\mathbb{Z}$-rank equal to 4. Supersingular curves can be defined over $\mathbb{F}_{p^2}$ or $\mathbb{F}_p$. Therefore, a common field that includes all isogenous curves is $\mathbb{F}_{p^2}$. Supersingular curves have the property that for every prime $\ell \neq p$, there exist $\ell+1$ isogenies of degree ℓ originating from a given supersingular curve. An isogeny can be computed over a kernel, κ, such that $\phi : E \to E/\langle\kappa\rangle$ by using Vélu's formulas [16]. By specifying curves of a smooth order, a large number of isogenies are available that can be computed efficiently [10].

2.2 Computing Large Degree Isogenies

The degree of an isogeny is its degree as an algebraic map. As shown in [17], isogeny computations can be done iteratively. Given an elliptic curve E and a point R of order ℓ^e, we compute $\phi\colon E \to E/\langle R\rangle$ by decomposing ϕ into a chain of degree ℓ isogenies, $\phi = \phi_{e-1} \circ \cdots \circ \phi_0$, as follows. Set $E_0 = E$ and $R_0 = R$, and define

$$E_{i+1} = E_i/\langle \ell^{e-i-1} R_i\rangle \qquad \phi_i : E_i \to E_{i+1} \qquad R_{i+1} = \phi_i(R_i).$$

Essentially, point additions are used to compute the kernel at each iteration and Vélu's formulas are used to compute ϕ_i and E_{i+1}. An optimal strategy to compute these isogenies relies on walking a path of the least cost on a large directed acyclic graph in the shape of a pyramid to the leaves, which is shown in Figure 1. For this graph, performing a multiplication by ℓ results in walking left and evaluating an $\ell-$isogeny results in walking right. Upon finding the point $[\ell^{e-i-1}]R_i$, we can compute the kernel of ϕ_i using $O(\ell)$ point additions and apply Vélu's formulas to compute ϕ_i and E_{i+1}. An optimal strategy is determined by comparing the cost of point multication by ℓ and and cost of evaluating an $\ell-$isogeny. The process is broken down into combining sub-strategies on smaller sets of the graph until an optimal strategy for the entire graph is determined. Figure 1 also illustrates one fast strategy for computing an isogeny of degree ℓ^7. Refer to [10] for more information regarding the optimal strategy.

2.3 Key Exchange Protocol Based on Isogenies

Two parties, Alice and Bob, want to exchange a secret key over an insecure channel in the presence of malicious third-parties. They agree on a smooth isogeny prime p of the form $\ell_A^a \ell_B^b \cdot f \pm 1$ where ℓ_A and ℓ_B are small primes, a and b are positive integers, and f is a small cofactor to make the number prime. They define a supersingular elliptic curve, $E_0(\mathbb{F}_q)$ where $q = p^2$. Lastly, they agree on four points on the curve that form two independent bases. Over a starting

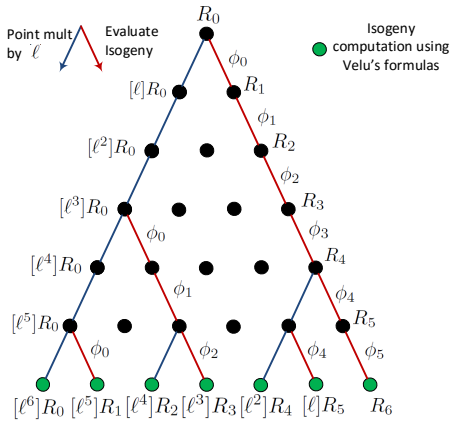

Fig. 1. Isogeny graph computation structure. This figure also demonstrates one efficient strategy for computing an isogeny of degree ℓ^7.

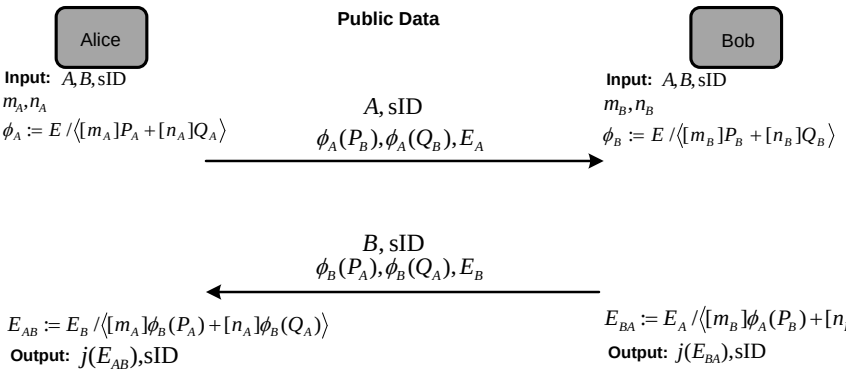

Fig. 2. Key exchange protocol using isogenies on supersingular curves. Note that $E[\ell_A^a] = \langle P_B, Q_B \rangle$ and $E[\ell_B^b] = \langle P_A, Q_A \rangle$. The secret data is $R_A = m_A P_A + n_A Q_A$ and $R_B = m_B P_B + n_B Q_B$. The public data is $E/\langle R_A \rangle, \phi_A(P_B), \phi_A(Q_B)$ and $E/\langle R_B \rangle, \phi_B(Q_A), \phi_B(Q_A)$. We denote A as Alice, B as Bob, and sID as session ID.

supersingular curve E_0, these are a basis $\{P_A, Q_A\}$ and $\{P_B, Q_B\}$ which generate $E_0[\ell_A^{e_A}]$ and $E_0[\ell_B^{e_B}]$, respectively, such that $\langle P_A, Q_A \rangle = E_0[\ell_A^{e_A}]$ and $\langle P_B, Q_B \rangle = E_0[\ell_B^{e_B}]$.

As first noted in [18], consider a graph of all supersingular elliptic curves of a fixed isogeny graph under $\mathbb{F}_{p^2}$. In this graph, the vertices represent each isomorphism class of supersingular elliptic curves and the edges represent the degree-ℓ isogenies of a particular isomorphism class. Essentially, each party takes seemingly random walks in the graph of supersingular isogenies of degree ℓ_A^a and ℓ_B^b to both arrive at supersingular elliptic curves with the same isomorphism class and j-invariant, similar to a Diffie-Hellman key exchange. In a graph of supersingular isogenies, the infeasibility to discover a path that connects two particular vertices provides security for this protocol.

Alice chooses two private keys $m_A, n_A \in \mathbb{Z}/\ell_A^a\mathbb{Z}$ with the stipulation that both are not divisible by ℓ_A^a. On the other side, Bob chooses two private keys $m_B, n_B \in \mathbb{Z}/\ell_B^b\mathbb{Z}$, where both private keys are not divisible by ℓ_B^b. From there, the key exchange protocol can be broken down into two rounds of the following:

1. Compute $R = \langle [m]P + [n]Q \rangle$ for points P, Q.
2. Compute the isogeny $\phi : E \to E/\langle R \rangle$ for a supersingular curve E.
3. Compute the images $\phi(P)$ and $\phi(Q)$ for the basis of the opposite party for the first round.

The key exchange protocol is shown in Figure 2. For a better illustration, we provide a step-by-step example of this protocol in Section **??**. Alice performs the double point multiplication with her private keys to obtain a kernel, $R_A = \langle [m_A]P + [n_A]Q \rangle$ and computes an isogeny $\phi_A : E_0 \to E_A = E_0/\langle [m_A]P + [n_A]Q \rangle$. She performs the large degree isogeny efficiently by performing many small isogenies of degree ℓ_A. She then computes the projection $\{\phi_A(P_B), \phi_A(Q_B)\} \subset E_A$ of the basis $\{P_B, Q_B\}$for $E_0[\ell_B^b]$ under her secret isogeny ϕ_A, which can be done efficiently by pushing the points P_B and Q_B through each isogeny of degree ℓ_A. Over a public channel, she sends these points and curve E_A to Bob. Likewise, Bob performs his own double-point multiplication and computes his isogeny over the supersingular curve E with $\phi_B : E_0 \to E_B = E_0/\langle [m_B]P + [n_B]Q \rangle$. He also computes his projection $\{\phi_B(P_A), \phi_B(Q_A)\} \subset E_B$ of the basis $\{P_A, Q_A\}$for $E_0[\ell_A^a]$ under his secret isogeny ϕ_B and sends these points and curve E_B to Alice. For the second round, Alice performs the double point multiplication to find a second kernel, $R_{AB} = \langle [m_A]\phi_B(P_A) + [n_A]\phi_B(Q_A) \rangle$, to compute a second isogeny $\phi'_A : E_B \to E_{AB} = E_B/\langle [m_A]\phi_B(P_A) + [n_A]\phi_B(Q_A) \rangle$. Bob also performs a double point multiplication and computes a second isogeny $\phi'_B : E_A \to E_{BA} = E_A/\langle [m_B]\phi_A(P_B) + [n_B]\phi_A(Q_B) \rangle$. Alice and Bob now have isogenous curves and can use the common j-invariant as a shared secret key.

$$\begin{aligned} E_{AB} = \phi'_B(\phi_A(E_0)) = \phi'_A(\phi_B(E_0)) = \\ = E_0/\{[m_A]P_A + [n_A]Q_A, [m_B]P_B + [n_B]Q_B\}, \\ j(E_{AB}) \equiv j(E_{BA}). \end{aligned}$$

2.4 Protocol Optimizations

There are many optimizations that have been proposed in [10] and [9]. Notably, all arithmetic is on Montgomery curves [19] as they have been shown to have fast scalar point multiplication and fast isogeny formulas. We refer the reader to the Explicit Formulas Database (EFD) [20] for the fastest operation counts on elliptic curves. The Kummer representation for Montgomery curves provides extremely fast curve arithmetic by performing operations on the curve's Kummer line [19]. Points are represented as $(X : Z)$, where $x = X/Z$. Under this scheme, there is no difference between points P and $-P$. The EFD provides explicit formulas for differential addition and point doubling. Let $\tilde{M}$ and $\tilde{S}$ refer to a multiplication and squaring in $\mathbb{F}_{p^2}$, respectively. Differential addition computes $P+Q$ with knowledge of $P-Q$ in $4\tilde{M}+2\tilde{S}$ or $3\tilde{M}+2\tilde{S}$ when the Z-coordinate for $P-Q$ is scaled to 1. Point doubling computes $2P$ in $3\tilde{M}+2\tilde{S}$ or $2\tilde{M}+2\tilde{S}$ when the point's input coordinate is scaled to 1. It is noted that P and $-P$ generate the same subgroup of points on the elliptic curve, so isogenies can be evaluated correctly on the Kummer line. Lastly, the optimal path to compute large-degree isogenies involves finding an optimal strategy of point multiplications and isogeny evaluations. The general trend has been to use isogeny graphs of base 2 and 3, since fast isogenies between Montgomery curves and fast scalar point multiplications can be performed over these isogeny graphs.

Our implementation style closely follows the methods of [10]. We use a 3-point Montgomery differential ladder (also presented in [10]) for a constant set of operations for double point multiplcations and their "affine" isogeny formulas for computing and evaluating large degree isogenies. We note that [10] does not scale the Z-coordinates of the inputs to the ladder to 1. This would decrease the cost of a 3-point step by 2 multiplications per step. [9] recently proposed "projective" isogeny formulas that represent the curve coefficients of a Montgomery

curve in projective space (i.e. a numerator and denominator), so that isogeny calculations do not need inversion until the very end of a round of a key exchange. We also note that [9] proposes sending isogenies evaluated over the points P, Q, and PQ in Kummer coordinates to the other party in the first round and that isogenies of degree 4 have been shown to be faster than isogenies of degree 2.

3 Proposed Choice of SIDH-Friendly Primes

The primes used in the key exchange protocol are the foundation of the underlying arithmetic. Since supersingular curves are used, it is necessary to generate primes to allow the curve to have smooth order so that the isogenies can be computed quickly. For this purpose, smooth isogeny primes of the form $p = \ell_A^a \ell_B^b \cdot f \pm 1$ are selected. Within that group of primes, [10] and [9] specifically chose isogeny-based cryptosystem parameters of $\ell_A = 2$ and $\ell_B = 3$. These isogeny graph bases provides efficient formulas for isogenies of degree 2 and 3, as shown in [10] and [9].

Smooth isogeny primes do not feature the distinct shape of a Mersenne prime (e.g. $2^{521} - 1$) or pseudo-Mersenne prime, but the choice of $\ell_A = 2$ does provide for several optimizations to finite-field arithmetic, which will be covered in more detail in Section 4.

The security of the underlying isogeny-based cryptosystem is directly related to the relative magnitude of ℓ_A^a and ℓ_B^b, or rather $\min(\ell_A^a, \ell_B^b)$. Whichever isogeny graph spanned by the prime is smaller is easier to attack. Therefore, a prime should be chosen where these isogeny graphs are approximately equal. As was demonstrated in [10], the classical security of the prime is approximately its size in bits divided by 4 and quantum security of a prime is approximately its size in bits divided by 6. Based on this security assessment, the SIDH protocol over a 512-bit, 768-bit, and 1024-bit prime feature approximately 85, 128, and 170 bits of quantum security, respectively.

3.1 Proposed Prime Search

Primes were searched for by setting balanced isogeny orders ℓ_A^a and ℓ_B^b for $\ell_A = 2$ and $\ell_B = 3$ and searching for factors f that produce a prime ± 1. However, using $+1$ in the form of the prime produced a prime where $-1 \bmod p$ was a quadratic residue. This was a result of using $\ell_A = 2$ as the first isogeny graph and is not optimal for the extension field $\mathbb{F}_{p^2}$. Thus, primes of the form $p = 2^a 3^b \cdot f - 1$ were primarily investigated. The primes were found by using a Sage script that changes f to find such primes. The prime number theorem in arithmetic progressions in [21] holds that the density of such primes is sufficient. We did not search for primes with an f value greater than 100. The primes that we discovered were compared and selected based on the following parameters:

- **Security**: The relative security of SIDH over a prime is based on $\min(\ell_A^a, \ell_B^b)$. Therefore, the prime should have balanced isogeny graphs and a small f term.
- **Size**: These primes are designed to be used in ARM processors, some that are limited in memory. These primes should feature a size slightly less than a power of 2 to allow for some speed optimizations such as lazy reduction and carry cancelling, while still featuring a high quantum security.
- **Speed**: These primes efficiently use space to reduce the number of operations per field arithmetic, but also have nice properties for the field arithmetic. Notably, all primes of the form $p = 2^a \ell_B^b \cdot f - 1$ will have the Montgomery friendly property [22] because the least significant half of the prime will have all bits set to '1'.

Table 1 contains a list of strong primes for 512, 768, and 1024-bit SIDH implementations. Each of these primes feature approximately balanced isogeny graphs. Each prime requires the least number of total bits for a quantum security level. We provide a prime with the f term to be 1 for each security level, but that is not a requirement.

Table 1. Proposed smooth isogeny primes

Security Level	Prime Size (bits)	$p = \ell_A^a \ell_B^b \cdot f \pm 1$	$\min(\ell_A^a, \ell_B^b)$	Classical Security	Quantum Security
p_{512}	499	$2^{251}3^{155}5 - 1$	3^{155}	123	82
	503	$2^{250}3^{159} - 1$	2^{250}	125	83
	510	$2^{252}3^{159}37 - 1$	2^{252}	126	84
p_{768}	751	$2^{372}3^{239} - 1$	2^{372}	186	124
	758	$2^{378}3^{237}17 - 1$	3^{237}	188	125
	766	$2^{382}3^{238}79 - 1$	3^{238}	189	126
p_{1024}	980	$2^{493}3^{307} - 1$	3^{307}	243	162
	1004	$2^{499}3^{315}49 - 1$	2^{499}	249	166
	1008	$2^{501}3^{316}41 - 1$	3^{316}	250	167
	1019	$2^{508}3^{319}35 - 1$	3^{319}	253	168

We provide several primes within each security level to give tunable parameters for an SIDH implementation. [9] proposes using the prime $2^{372}3^{239} - 1$ for a 768-bit implementation. This prime is actually 751 bits, allowing for 17 bits of freedom for speed optimizations in systems using 32 or 64-bit words. However, as Table 1 shows, the prime $2^{378}3^{237}17 - 1$ is a 758-bit prime that gives 1 more bit of quantum security and still has 10 bits of freedom to allow for speed optimizations. We find it useful to have several strong primes to work with, which could allow for a variety of speed techniques.

For our design, we chose to implement over the primes:

$$\begin{aligned} p_{512} &= 2^{250}3^{159} - 1 \\ p_{768} &= 2^{372}3^{239} - 1 \\ p_{1024} &= 2^{501}3^{316}41 - 1 \end{aligned}$$

4 Proposed Finite-Field Arithmetic

For any cryptosystem featuring large finite-fields, the finite-field arithmetic lies at the heart of the computations. This work is no exception. The critical operations are finite-field addition, squaring, multiplication, and inversion. The abundance of these operations throughout the entire key exchange protocol calls for numerous optimizations to the arithmetic, even at the assembly level. This work targets the ARMv7-A architectures. All operations are done in the Montgomery domain [23] to take advantage of the extremely fast Montgomery reduction for the primes above.

4.1 Field Addition

Finite-field addition performs $A + B = C$, where $A, B, C \in \mathbb{F}_p$. Essentially, this just means that there is a regular addition of elements A and B to produce a third element C. If $C \geq p$, then $C = C - p$. For ARMv7, this can be efficiently done by using the *ldmia* and *stmia* instructions, which load and store multiple registers at a time, incrementing the address each time. The operands are loaded into multiple registers and added with the carry bit. If the resulting value is larger than the prime for a field, then a subsequent subtraction by the prime occurs. For a constant-time implementation, the conditional flags are used to alter a mask that is applied to the prime as the subtraction occurs. In the case that the value is not larger than the prime, the masked prime becomes 0. Finite-field subtraction is nearly identical to addition, but subtraction with borrow is used and if the borrow flag is set at the end of the subtraction, then the prime is added to the resulting value.

4.2 Field Multiplication and Squaring

Finite-field multiplication performs $A \times B = C$, where $A, B, C \in \mathbb{F}_p$. This equates to a regular multiplication of A and B to produce a third element C. However, if elements A and B are both m-bits, then the result, C, is $2m$-bits. A reduction must be made so that the result is still within the field. Montgomery multiplication and reduction [23] was chosen because of its fast reduction method. Introduced in [9], smooth isogeny primes of the form $2^a \ell^b f - 1$ feature a fast reduction based on simplifying the Montgomery reduction formula [23]:

$$c = (a + (aM' \bmod R)p)/R = (a - aM' \bmod R)/R + ((p+1)(aM' \bmod R))$$

where m is slightly larger than the size of the prime (e.g. $R = 2^{512}$ for p_{512}), a is a result of a multiplication and less than $2m$ bits long, $M' = -p^{-1} \bmod 2^m$, and $c = a \bmod p$. In this equation, $p+1$ has many least-signficant limbs of '0', since approximately half of the least-significant limbs of p are all '1'. Thus, many partial products can be avoided for reduction over this scheme. An alternative to the above scheme is to leave the Montgomery reduction in its standard form, but perform the first several partial products as subtractions since $0\text{xFF} \times A = A \times 2^8 - A$ and the least significant limbs are all '1'.

The typical scheme for Montgomery multiplication is to use $M' = -p^{-1} \bmod 2^w$, where w is the word size. We note that the form of the prime $2^a \ell^b f - 1$ guarantees that $M' = 1$ as long as $2^a > 2^{64}$, for our ARMv7 implementation. This reduces the complexity of Montgomery reduction from $k^2 + k$ to k^2 single-precision multiplication operations, where k is the number of words of an element within the field that must be multiplied.

We utilize the ARM-NEON vector unit to perform the multiplications because it can hold many more registers and parallelize the multiplications. We adopt the multiplication and squaring scheme of [24] to perform large multiplications efficiently. This scheme utilizes a transpose of individual registers within NEON to reduce data dependency stalls. This same technique was employed in this work to perform the multiplication for 512-bit multiplication with the Cascade Operand Scanning (COS) method, as shown in Figure 3. By using a transposed quad register in NEON, the partial products can be determined out of order and the carries applied later, reducing data dependencies in the multiplication sequence. Figure 3 demonstrates an example of a 32×256 bit multiplication, which is applied several times to produce a 512×512 multiplication. A separated reduction scheme was used. A 1024-bit multiplication is composed of three 512×512 multiplications, based on a 1 level additive Karatsuba multiplication. Squaring can reuse the input operands and several partial products for multiplication and requires approximately 75% of the cycles for a multiplication.

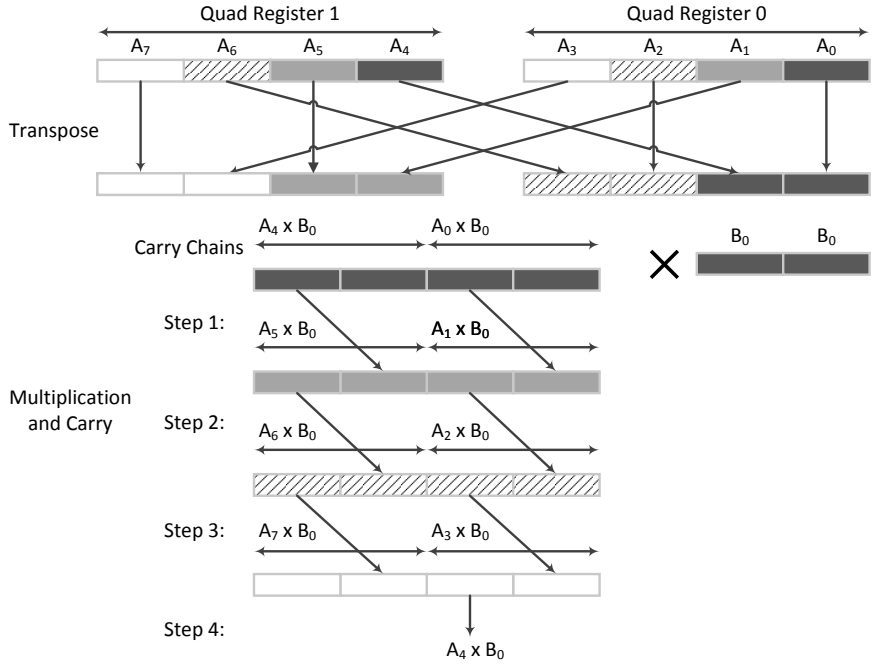

Fig. 3. Finite-field Multiplication using NEON

4.3 Field Inversion

Finite-field inversion finds some A^{-1} such that $A \cdot A^{-1} = 1$, where $A, A^{-1} \in \mathbb{F}_p$. There are many schemes to perform this efficiently. Fermat's little theorem exponentiates $A^{-1} = A^{p-2}$. This requires many multiplications and squarings, but is a constant set of operations. The Extended Euclidean Algorithm (EEA) has a significantly lower time complexity of $O(\log^2 n)$ compared to $O(\log^3 n)$ for Fermat's little theorem. EEA uses a greatest common divisor algorithm to compute the modular inverse of elements a and b with respect to each other, $ax + by = \gcd(a, b)$. Based on the analysis presented in Section 5, the EEA was chosen because it made affine SIDH much faster than projective SIDH. The GMP library already employs a highly optimized version of EEA for various architectures. EEA performs an inversion quickly, but does leak some information about the value being inverted from the timing information. Therefore, to take advantage of this fast inversion and provide some protections against simple power analysis and timing attacks, a random value was multiplied to the element before and after the inversion, effectively obscuring what value was initially being inverted. This requires two extra multiplications, but the additional defense against timing and simple power analysis attacks is necessary for a secure key exchange protocol.

4.4 Extension Field Arithmetic

Since isogenies can be defined over $\mathbb{F}_{p^2}$, a reduction modulus must be defined to simplify the multiplication between elements of $\mathbb{F}_{p^2}$. With the prime choice of $p = 2^a \ell_B^b \cdot f - 1$, -1 is never a quadratic residue of the prime and $x^2 + 1$ can be used as a modulus for the extension field. With this, we propose reduced arithmetic in $\mathbb{F}_{p^2}$ based on fast arithmetic in $\mathbb{F}_p$. These equations were made in a Karatsuba-like fashion to reduce the total number of multiplications and squarings.

The lazy reduction technique was also employed for inversion to minimize computational cost. For the equations below, assume $A = (A_0, A_1)$, $B = (B_0, B_1) \in \mathbb{F}_{p^2}$.The results of operations in $\mathbb{F}_{p^2}$ are $C = (C_0, C_1)$

$$\begin{array}{rc} A + B = & (A_0 + B_0, A_1 + B_1) \\ A - B = & (A_0 - B_0, A_1 - B_1) \\ A \times B = & (A_0B_0 - A_1B_1, (A_0 + A_1), \\ & (B_0 + B_1) - A_0B_1 - A_1B_0) \\ A^2 = & ((A_0 + A_1)(A_0 - A_1), 2A_0A_1) \\ A^{-1} = & (A_0(A_0^2 + A_1^2)^{-1}, -A_1(A_0^2 + A_1^2)^{-1}) \end{array}$$

Addition/subtraction in $\mathbb{F}_{p^2}$ require 2 additions in $\mathbb{F}_p$, squaring in $\mathbb{F}_{p^2}$ requires 2 multiplications and 3 additions in $\mathbb{F}_p$, multiplication in $\mathbb{F}_{p^2}$ requires 3 multiplications and 5 additions in $\mathbb{F}_p$, and inversion in $\mathbb{F}_{p^2}$ requires an inversion, 2 multiplications, 2 squarings, and 2 additions in $\mathbb{F}_p$. This arithmetic required 3 temporary registers in $\mathbb{F}_p$. Two extra multiplications by a random value were added to finite-field inversion to provide side-channel resistance. An inversion over the Montgomery domain also has an additional multiplication ($\mathrm{MM}(\frac{1}{ar}, r^3) = \frac{r}{a}$) in $\mathbb{F}_p$ to keep the result in the Montgomery domain.

5 Affine or Projective Isogenies

Here, we analyze the complexity of utilizing the new "projective" isogeny formulas presented by Costello et al. in [9] to the "affine" isogeny formulas presented by De Feo et al. in [10]. Notably, the projective formulas allow for constant-time inversion implementations without greatly increasing the total time of the protocol. However, in terms of non-constant inversion, we will show that the affine isogeny formulas are still much faster for ARMv7 devices. For cost comparison between these formulas, let I, M, and S refer to inversion, multiplication, and squaring in $\mathbb{F}_p$, respectively. A tilde above the letter indicates that the operation is in $\mathbb{F}_{p^2}$.

We introduce the idea of the inversion/multiplication ratio, or for SIDH over $\mathbb{F}_{p^2}$, $\tilde{I}/\tilde{M}$, as a metric to compare the relative cost of inversion and multiplication and decide between the effectiveness of affine or projective formulas. This inversion/multiplication ratio is dependent on the size of elements in $\mathbb{F}_p$, the processor, as well as the inversion used. For a constant-time inversion using Fermat's little theorem, the ratio is most likely several hundred since it requires several hundred multiplications and squarings for the inversion exponentiation. However, for non-constant time inversion, such as EEA or Kaliski's almost inverse [25], the ratio is much smaller. For instance, on ARMv7 platforms, as in [26], the ratio ranges from approximately 10 for a 254-bit number to approximately 7 for a 638-bit number, both over $\mathbb{F}_p$. For personal computers, the ratio is much larger for non-constant inversion, typically greater than 20 for optimized arithmetic.

In Table 2, we compare the relative computational costs of affine isogeny formulas presented in [10] and projective isogeny formulas presented in [9] over isogenies of degree 3 and 4. Point multiplications by ℓ are over Kummer coordinates with affine or projective curve coefficients. Isogeny computations compute the map between two points and isogeny evaluations push a point through the mapping, both of these are of degree ℓ. Affine isogeny computations cost more than their projective counterpart because certain calculations are performed that are reused across each affine isogeny evaluation.

From this table, we created optimal strategies for traversing the large-degree isogeny graphs, visualized in Figure 1. The affine and projective optimal strategy differed because the ratio of

Table 2. Affine isogeny formulas vs. projective isogenies formulas

Computation	Affine Cost [10]	Projective Cost [9]
Point Mult-by-3	$7\tilde{M} + 4\tilde{S}$	$8\tilde{M} + 5\tilde{S}$
Iso-3 Computation	$1\tilde{I} + 5\tilde{M} + 1\tilde{S}$	$3\tilde{M} + 3\tilde{S}$
Iso-3 Evaluation	$4\tilde{M} + 2\tilde{S}$	$6\tilde{M} + 2\tilde{S}$
Point Mult-by-4	$6\tilde{M} + \tilde{S}$	$8\tilde{M} + 4\tilde{S}$
Iso-4 Computation	$1\tilde{I} + 3\tilde{M}$	$5\tilde{S}$
Iso-4 Evaluation	$6\tilde{M} + 4\tilde{S}$	$9\tilde{M} + 1\tilde{S}$

Table 3. Relative costs of computing large-degree isogenies based on affine vs. projective isogeny formulas

Prime	$\#3P$	#3eval	#3comp	LargeIso3Cost	$\#4P$	#4eval	#4comp	LargeIso4Cost
Affine Isogeny Computations								
p_{512}	496	698	159	$159\tilde{I} + 9417\tilde{M}$	457	410	124	$124\tilde{I} + 6966\tilde{M}$
p_{768}	780	1176	239	$239\tilde{I} + 15163\tilde{M}$	771	638	185	$185\tilde{I} + 11215\tilde{M}$
p_{1024}	1123	1568	316	$316\tilde{I} + 21005\tilde{M}$	1061	942	250	$250\tilde{I} + 15974\tilde{M}$
Projective Isogeny Computations								
p_{512}	500	691	159	$11525\tilde{M}$	423	441	124	$9182\tilde{M}$
p_{768}	811	1124	239	$18623\tilde{M}$	638	771	185	$14865\tilde{M}$
p_{1024}	1129	1558	316	$25792\tilde{M}$	981	1013	250	$21076\tilde{M}$

point multiplication over isogeny evaluation differed. Similar to the method proposed by [10] and also implemented in [9], we created an optimal strategy to traverse the graph. We based the cost of traversing the graph with the relationship $\tilde{S} = 0.66\tilde{M}$, since there are 2 multiplications in $\mathbb{F}_p$ for $\tilde{S}$ and 3 multiplications in $\mathbb{F}_p$ for $\tilde{M}$. We performed this experiment for our selected primes in the 512-bit, 768-bit, and 1024-bit categories, shown in Table 3. In Table 3, we count the total number of point multiplications by ℓ as $\#\ell P$, the total number of $\ell-$isogeny evaluations as $\#\ell$eval, and the total number of $\ell-$isogeny computations as $\#\ell$comp. From the cost of these operations in affine or projective coordinates, shown in Table 2, we calculated the total cost of the large-degree isogeny in terms of multiplications and inversions in $\mathbb{F}_{p^2}$ under LargeIsoℓCost.

We note that the difference in performance is also much greater for the first round of the SIDH protocol, as the other party's basis points are pushed through the isogeny mapping. This includes 3 additional isogeny evaluations per isogeny computation, as P, Q, and $P - Q$ are pushed through the isogeny. In Table 4, we compare the break-even points for when the cost of affine and projective isogenies are the same. If the ratio is smaller than the break-even point, then the large-degree isogeny computation is faster with affine isogeny formulas. Alice operates over degree 4 isogenies and Bob operates over degree 3 isogenies. We utilize $\tilde{I} = I + 3.33\tilde{M}$ to get the break-even points for operations in $\mathbb{F}_p$ since we used a Karatsuba-based inversion. Thus, $I/M = 3(\tilde{I}/\tilde{M} - 3.33)$. As an example, the break-even point for Alice's round 1 isogeny is $I = 53M$ at the 512-bit level. Thus, even with conservative estimates for the cost of using projective coordinates, affine isogenies trump projective coordinates for small I/M ratios.

6 Implementation Results and Discussion

In this section, we review the ARM architectures that were used as testing platforms, how we optimized the assembly code around them, and present our results.

Table 4. Comparison of break-even inversion/multiplication ratios for large-degree isogenies at different security levels. When the inversion over multiplication ratio is at the break-even point, affine isogenies require approximately the same cost as projective isogenies. Ratios smaller than these numbers are faster with affine formulas. The tilde indicates an operation in $\mathbb{F}_{p^2}$ and no tilde indicates an operation in $\mathbb{F}_p$.

Prime	Alice Round 1 Iso	Bob Round 1 Iso	Alice Round 2 Iso	Bob Round 2 Iso
p_{512}	$\tilde{I} = 20.87\tilde{M}$	$\tilde{I} = 19.26\tilde{M}$	$\tilde{I} = 17.87\tilde{M}$	$\tilde{I} = 13.26\tilde{M}$
p_{768}	$\tilde{I} = 22.73\tilde{M}$	$\tilde{I} = 20.48\tilde{M}$	$\tilde{I} = 19.73\tilde{M}$	$\tilde{I} = 14.48\tilde{M}$
p_{1024}	$\tilde{I} = 23.41\tilde{M}$	$\tilde{I} = 21.15\tilde{M}$	$\tilde{I} = 20.41\tilde{M}$	$\tilde{I} = 15.15\tilde{M}$
p_{512}	$I = 52.62M$	$I = 47.78M$	$I = 43.62M$	$I = 29.78M$
p_{768}	$I = 58.20M$	$I = 51.44M$	$I = 49.20M$	$I = 33.46M$
p_{1024}	$I = 60.23M$	$I = 53.46M$	$I = 51.23M$	$I = 35.46M$

6.1 ARM Architectures

As the name Advanced RISC Machines implies, ARM implements architectures that feature simple instruction execution. The architectures have evolved over the years, but this work will focus on the ARMv7-A. The ARMv7-A family employs a 32-bit architecture that uses 16 general-purpose registers, although registers 13, 14, and 15 are reserved for the stack pointer, link register, and program counter, respectively. ARM-NEON is a Single-Instruction Multiple-Data (SIMD) engine that provides vector instructions for the ARMv7 architecture. ARMv7's NEON features 32 registers that are 64-bits wide or alternatively viewed as 16 registers that are 128-bits wide. NEON provides nice speedups over standard register approaches by taking advantage of data paralellism in the large register sizes. This comes in handy primarily in multiplication, squaring, and reduction.

We benchmarked the following boards running various ARM architectures:

- A BeagleBoard Black running a single ARMv7 Cortex-A8 processor operating at 1.0 GHz. Conforming to the ARMv7-A architecture, this processor is among the more basic architectures within the family, supporting a decode width of 2, a pipeline depth of 13, a split Harvard L1 cache with 32 KB each, and an L2 cache of 512 KB.
- A Jetson TK1 running 4 ARMv7 Cortex-A15 cores operating at 2.3 GHz. This processor provides more performance than the Cortex-A8 with a decode width of 3, multiple depths of pipeline, and out-of-order execution. The L1 cache is the same as the Cortex-A8, but the L2 cache is shared among cores with up to 8 MB per chip.

We utilize loop unrolling, instruction re-ordering, register allocation, and multiple stores to hand-optimize our assembly used for finite-field arithmetic in the above boards.

6.2 Testing Methodology

The key exchange was written in the standard C language. We used GMP version 6.1.0. The code was compiled using the standard operating system and development environment on the given device. A parameters file defining the agreed upon curve, basis points, and strategies for the key exchange was generated externally using Sage. The strictly C code with GMP is fairly portable and can be used with primes of any size, as long as it is provided with a valid parameters file. There are separate versions which include the 512-bit and 1024-bit assembly optimizations that only work with primes up to these sizes. The protocols are identical in both the C and ASM implementations. The primes that were used can be found in Table 3.

Table 5. Timing results of key exchange on Beagle Board Black ARMv7 device for different security levels

Beagle Board Black (ARM v7) Cortex-A8 at 1.0 GHz using C												
Field	$\mathbb{F}_p$ [cc]						$\mathbb{F}_{p^2}$ [cc]				Key Exchange [cc $\times 10^3$]	
Size	A	S	M	mod	I	I/M	$\tilde{A}$	$\tilde{S}$	$\tilde{M}$	$\tilde{I}$	Alice	Bob
p_{512}	115	1866	2295	3429	40100	7.0	1241	12229	14896	72400	483,968	514,786
p_{768}	142	3652	4779	6325	71500	6.4	1404	23167	28459	135400	1,406,381	1,525,215
p_{1024}	168	5925	8202	10150	111900	6.1	1558	38046	46891	211400	3,135,526	3,367,448
Beagle Board Black (ARM v7) Cortex-A8 at 1.0 GHz using ASM and NEON												
Field	$\mathbb{F}_p$ [cc]						$\mathbb{F}_{p^2}$ [cc]				Key Exchange [cc $\times 10^3$]	
Size	A	S	M	mod	I	I/M	$\tilde{A}$	$\tilde{S}$	$\tilde{M}$	$\tilde{I}$	Alice	Bob
p_{512}	70	718	953	962	40100	20.9	279	4445	6736	52756	216,503	229,206
p_{1024}	120	2714	3723	3956	111900	14.6	375	15714	23682	150795	1,597,504	1,708,383

6.3 Results and Comparison

The results for this experiment are presented in Table 5 and Table 6 for the BeagleBoard Black and Jetson TK1, respectively. This provides the timings, in clock cycles, of individual finite field operations in $\mathbb{F}_p$ and $\mathbb{F}_{p^2}$ as well as the total computation time of each party for the protocol. The expected time to run this protocol is roughly Alice or Bob's computation time and some transmission cost.

The Beagle Board Black achieved a speedup of 2.27 over the 512-bit primes and a speedup of 2.00 over 1024-bit primes when using our hand-optimized assembly code over our generic C code. The Jetson TK1 achieved a speedup of 1.94 for 512-bit primes and a speedup of 1.59 for 1024-bit primes when using the assembly code. These speedups came as a result of the optimized finite field arithmetic over $\mathbb{F}_p$. Addition is generally a fraction of the cost. Multiplication and squaring are almost twice as fast with the ASM. The most significant improvement is reduction around 3-3.5 times as fast with the ASM. Addition in $\mathbb{F}_{p^2}$ is approximately 5-7 times faster with assembly because the intermediate elements were guaranteed to be in the field, only requiring a subtraction with a mask as a modulus. With the assembly optimizations, the Beagle Board Black performs one party's computations in approximately 0.223 seconds and 1.65 seconds over 85-bit and 170-bit quantum security, respectively. The Jetson TK1 performs one party's computations in approximately 0.066 seconds and 0.491 seconds over 85-bit and 170-bit quantum security, respectively.

Our implementation follows the algorithms and formulas of the affine key exchange protocol given in [10]. Our implementation also includes side-channel resistance. Our finite-field arithmetic is constant-time, except for inversion which applies extra multiplications for protection, and we utilize a constant set of operations that deal with the secret keys. Lastly, our C implementation is portable because it only requires a C compiler and the GNU library.

The only other portable implementations of SIDH for ARMv7 are [11] and [9]. [9] only operates with projective isogeny formulas over the 751-bit prime, $2^{372}3^{239}-1$, and uses a generic, constant-time, implementation with Montgomery reduction. [11] uses the same affine formulas as our implementation, but uses primes that are not as efficient. Table 7 contains a comparison of these implementations for ARM Cortex-A15. We note that the assembly optimizations are not applied for our 768-bit version. Similarly, [9] has a generic implementation with Montgomery reduction. Our assembly optimized implementation is approximately 3 times faster than the implementation in [11] and the portable C implementation is about 5 times faster than the projective isogeny implementation in [9]. [11] does not consider side-channel attacks, but [9] is

Table 6. Timing results of key exchange on NVIDIA Jetson TK-1 ARMv7 device for different security levels

Jetson TK-1 Board (ARM v7) Cortex-A15 at 2.3 GHz using C												
Field	$\mathbb{F}_p$ [cc]						$\mathbb{F}_{p^2}$ [cc]				Key Exchange [cc $\times 10^3$]	
Size	A	S	M	mod	I	I/M	$\tilde{A}$	$\tilde{S}$	$\tilde{M}$	$\tilde{I}$	Alice	Bob
p_{512}	83	926	1152	2271	24302	7.1	877	7256	8776	42481	285,026	302,332
p_{768}	99	1679	2403	4024	39100	6.1	982	13467	16216	73922	783,303	848,461
p_{1024}	117	2955	4144	6053	59800	5.7	1122	21558	26286	115437	1,728,183	1,851,782
Jetson TK-1 Board (ARM v7) Cortex-A15 at 2.3 GHz using ASM and NEON												
Field	$\mathbb{F}_p$ [cc]						$\mathbb{F}_{p^2}$ [cc]				Key Exchange [cc $\times 10^3$]	
Size	A	S	M	mod	I	I/M	$\tilde{A}$	$\tilde{S}$	$\tilde{M}$	$\tilde{I}$	Alice	Bob
p_{512}	39	516	640	732	24302	17.7	158	3025	4579	34049	148,003	154,657
p_{1024}	73	1856	2464	2961	59800	11.0	273	11273	17007	97594	1,118,644	1,140,626

Table 7. Comparison of affine and projective isogeny implementations on ARM Cortex-A15 embedded processors. Our work and [9] was done on a Jetson TK1 and [11] was performed on an Arndale ARM Cortex-A15.

Work	Language	Field size [bits]	Quantum Security [bits]	Isogeny formulas	Timings [$cc \times 10^6$]				
					Alice R1	Bob R1	Alice R2	Bob R2	Total
Costello et al. [9][1]	C	751	124	Proj.	1,794	2,120	1,665	2,001	7,580
Azarderakhsh et al. [11]	C	521	85	Affine	N/A	N/A	N/A	N/A	1,069
	C	771	128		N/A	N/A	N/A	N/A	3,009
	C	1035	170		N/A	N/A	N/A	N/A	6,477
This work	ASM	503	83	Affine	83	87	66	68	302
	C	751	124		437	474	346	375	1,632
	ASM	1008	167		603	657	516	484	2,259

1. Targeted x86-64 architectures, but is portable on ARM. All arithmetic is in generic C.

a constant-time implementation, which is inherently protected by simple power analysis and timing attacks.

We were surprised to find that the performance of the SIDH library in [9] suffered on ARMv7, so we investigated this further. Table 8 compares the I/M ratio for different computer architectures over GMP computations. We note that with optimized multiplication, this would generally be higher, but it is an idea of the relative difference between I/M ratios for ARM architectures and x86 architectures. As Table 8 shows, the I/M ratio for a PC is much greater than ARM architectures, by a factor of 2. This shows that ARM implementations benefit much more from using affine coordinates . We primarily attribute the slowness experienced by the SIDH library in [9] to the generic arithmetic and low I/M ratios for ARMv7 devices.

There are several other popular post-quantum cryptosystems that have been implemented in the literature. The ones that consider embedded system have typically used FPGA's or 8-bit microcontrollers, such as the lattice-based system in [3], code-based system in [5], or McEllice system in [2] and [4]. The comparison with any of these works is difficult because the algorithms are extremely different and the implementations did not use ARM-powered embedded devices.

Table 8. Comparison of I/M ratios for various computer architectures based on GMP library

Architecture	Device	I/M ratio		
		p_{512}	p_{768}	p_{1024}
ARMv7 Cortex-A8	Beagle Board Black	7.0	6.4	6.1
ARMv7 Cortex-A15	Jetson TK1	7.1	6.1	5.9
ARMv8 Cortex-A53	Linaro HiKey	8.2	7.3	6.5
Haswell x86-64	i7-4790k	14.9	14.7	13.8

7 Conclusion

In this paper, we proved that isogeny-based key exchanges proposed in [6] can be implemented efficiently on emerging ARM embedded devices and represent a new alternative to classical cryptosystems. Both efficient primes and the impact of projective isogeny formulas were investigated. Without transmission overhead, a party can compute their side of the key exchange in fractions of a second. We hope that the initial investigation of this protocol on embedded devices will inspire other researchers to continue looking into isogeny-based implementations as a strong candidate for NIST's call for post-quantum resistant cryptosystems. As a future work, we plan to apply our assembly optimizations to the projective isogeny formulas presented in [9] for a constant-time implementation. We note that robust and high-performance implementations provide critical support for industry adoption of isogeny-based cryptosystems.

References

1. Shor, P.W.: Algorithms for quantum computation: Discrete logarithms and factoring. In: 35th Annual Symposium on Foundations of Computer Science (FOCS 1994). 124–134 (1994)
2. Eisenbarth, T., Güneysu, T., Heyse, S., Paar, C.: Microeliece: Mceliece for embedded devices. In Clavier, C., Gaj, K., eds.: 11th International Workshop Cryptographic Hardware and Embedded Systems - CHES 2009. Volume 5747 of Lecture Notes in Computer Science., Springer 49–64 (2009)
3. Güneysu, T., Lyubashevsky, V., Pöppelmann, T.: Practical lattice-based cryptography: A signature scheme for embedded systems. In Prouff, E., Schaumont, P., eds.: 14th International Workshop on Cryptographic Hardware and Embedded Systems - CHES 2012. Volume 7428 of Lecture Notes in Computer Science., Springer 530–547 (2012)
4. Heyse, S.: Implementation of mceliece based on quasi-dyadic goppa codes for embedded devices. In Yang, B.Y., ed.: 4th International Workshop on Post-Quantum Cryptography, PQCrypto 2011. Volume 7071 of Lecture Notes in Computer Science., Springer 143–162 (2011)
5. Heyse, S., von Maurich, I., Güneysu, T.: Smaller keys for code-based cryptography: Qc-mdpc mceliece implementations on embedded devices. In Bertoni, G., Coron, J.S., eds.: 15th International Workshop on Cryptographic Hardware and Embedded Systems - CHES 2013. Volume 8086 of Lecture Notes in Computer Science., Springer 273–292 (2013)
6. Jao, D., Feo, L.D.: Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In: Post-Quantum Cryptography–PQCrypto 2011. Volume 7071 of LNCS. 19–34 (2011)
7. Chen, L., Jordan, S.: Report on post-quantum cryptography, (2016)
8. Menezes, A., van Oorschot, P., Vanstone, S.: Handbook of Applied Cryptography, New York (1996)
9. Costello, C., Longa, P., Naehrig, M.: Efficient algorithms for supersingular isogeny diffie-hellman. Cryptology ePrint Archive, Report 2016/413 (2016)
10. De Feo, L., Jao, D., Plut, J.: Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. Journal of Mathematical Cryptology 8(3), 209–247 (September 2014)
11. Reza Azarderakhsh, Dieter Fishbein, D.J.: Efficient implementations of a quantum-resistant key-exchange protocol on embedded systems. Technical report, University of Waterloo (2014)

12. Silverman, J.H.: The Arithmetic of Elliptic Curves. Volume 106 of GTM. Springer, New York (1992)
13. Rostovtsev, A., Stolbunov, A.: Public-key cryptosystem based on isogenies (2006) http://eprint.iacr.org/2006/145/.
14. Childs, A., Jao, D., Soukharev, V.: Constructing elliptic curve isogenies in quantum subexponential time (2010)
15. Tate, J.: Endomorphisms of abelian varieties over finite fields. Inventiones Mathematicae 2, 134–144 (1966)
16. Vélu, J.: Isogénies entre courbes elliptiques. Comptes Rendus de l'Académie des Sciences Paris Séries A-B 273, A238–A241 (1971)
17. Couveignes, J.M.: Hard homogeneous spaces. Cryptology ePrint Archive, Report 2006/291 (2006)
18. Mestre, J.F.: La méthode des graphes. Exemples et applications. In: Proceedings of the international conference on class numbers and fundamental units of algebraic number fields (Katata, 1986), Nagoya, Nagoya Univ. 217–242 (1986)
19. Montgomery, P.: Speeding the Pollard and Elliptic Curve Methods of Factorization. Mathematics of computation, 243–264 (1987)
20. Bernstein, D.J., Lange, T.: Explicit-Formulas Database (2007) http://www.hyperelliptic.org/EFD/index.html.
21. Lagarias, J., Odlyzko, A.: Effective versions of the Chebotarev density theorem. In: Algebraic number fields: L-functions and Galois properties. Symposium Proceedings of the University of Durham 409–464 (1975)
22. Gueron, S., Krasnov, V.: Fast prime field elliptic-curve cryptography with 256-bit primes. Journal of Cryptographic Engineering 5(2), 141–151 (2014)
23. Montgomery, P.L.: Modular multiplication without trial division. Mathematics of Computation 44(170), 519–521 (1985)
24. Seo, H., Liu, Z., Grobschadl, J., Kim, H.: Efficient arithmetic on arm-neon and its application for high-speed rsa implementation. Cryptology ePrint Archive, Report 2015/465 (2015) http://eprint.iacr.org/.
25. Kaliski, B.S.: The montgomery inverse and its applications. IEEE Trans. Comput. 44(8), 1064–1065 (August 1995)
26. Grewal, G., Azarderakhsh, R., Longa, P., Hu, S., Jao, D.: Efficient Implementation of Bilinear Pairings on ARM Processors. In: Selected Areas in Cryptography: 19th International Conference, SAC 2012, Windsor, ON, Canada, August 15-16, 2012, Revised Selected Papers. Springer Berlin Heidelberg, Berlin, Heidelberg (2013) 149–165