Scalable Practical Byzantine Fault Tolerance with Short-Lived Signature Schemes

Xinxin Fan
IoTeX
Menlo Park, CA, USA
xinxin@iotex.io

ABSTRACT

The Practical Byzantine Fault Tolerance (PBFT) algorithm is a popular solution for establishing consensus in blockchain systems. The execution time of the PBFT consensus algorithm has an important effect on the blockchain throughput. Digital signatures are extensively used in PBFT to ensure the authenticity of messages during the different phases. Due to the round-based and broadcast natures of PBFT, nodes need to verify multiple signatures received from their peers, which incurs significant computational overhead and slows down the consensus process. To address this issue, we propose an efficient short-lived signature based PBFT variant, which utilizes short-length cryptographic keys to sign/verify messages in PBFT for a short period of time and blockchain-aided key distribution mechanisms to update those keys periodically. We also present efficient algorithms for accelerating the software implementation of the BLS threshold signature scheme. Our extensive experiments with three elliptic curves and two signature schemes demonstrate the efficacy of using short-lived signature schemes for improving the scalability of PBFT significantly.

KEYWORDS
Practical Byzantine Fault Tolerance, Scalability, Short-Lived Public Key, ECDSA, BLS Threshold Signature

ACM Reference Format:

1 INTRODUCTION

A blockchain constitutes a distributed ledger that records transactions in an append-only fashion across a network of random agents, nodes or peers. At the center of the blockchain operation is the maintenance of the consistent global view of the information recorded on the blockchain, which forms the backbone for enabling users to interact with each other in a trustless and decentralized manner. The security of the consensus model is perhaps the most critical aspect for a blockchain platform. The two key properties of a consensus model are: i) Safety/Consistency: All honest nodes produce the same output and the outputs produced by the honest nodes are valid; and ii) Liveness: All honest nodes in consensus eventually produce a value. A secure and robust consensus protocol needs to tolerate a wide variety of Byzantine behaviors, including, but not limited to, failures of network nodes, partition of the network, message delay, out-of-order and corruption, and reach consensus in the present of Byzantine nodes as long as the number of those nodes within the system is limited.

The seminal Practical Byzantine Fault Tolerance (PBFT) algorithm, proposed by Castro and Liskov [8], provides a practical Byzantine state machine replication [36] that tolerates Byzantine failures with low overhead. Essentially, a group of replicas running the PBFT algorithm are ordered in a sequence with one replica being the leader and others the validators. The leader node is changed in round-robin manner and in every round all the nodes communicate with each other asynchronously through a three-phase protocol. The final result is that all honest nodes reach an agreement on the order of the record by either accepting or rejecting it. PBFT provides safety and liveness guarantees provided that the number of Byzantine nodes is less than one third of the overall nodes in the system for any given window of vulnerability. In other words, to tolerate \( f \) Byzantine faults, the group needs at least \( 3f + 1 \) members. In practice, PBFT and its variants have been adopted by a handful of blockchain platforms such as Hyperledger [21], Tendermint [41], Zilliqa [45], IoTeX [22], just to name a few.

Although PBFT offers the salient property like instant action finality, it only works well in its classical form with small consensus group size due to the broadcast nature of the three-phase protocol. More specifically, upon receiving the block proposal from the leader, all the validators need to exchange their opinions in the subsequent phases by broadcasting and verifying multiple signed messages. Using digital signatures at high security levels (e.g., 128-bit), while providing strong security guarantees, incurs significant computational and communication overhead in PBFT, thereby limiting its scalability and throughput. We note that a typical round of PBFT consensus process in blockchain systems is rather fast (i.e., in a few tens of seconds) for a group of tens of nodes. Therefore, it is possible to utilize digital signatures at low or medium security levels (e.g., 56- or 80-bit) in PBFT and update the short-length key pairs periodically, which enables us to maintain the throughput and improve the scalability of PBFT simultaneously. The extensive experiments with two short-lived signature schemes demonstrate that the verification time for multiple signatures received from peers can be reduced significantly in PBFT. We also propose novel key updating techniques to refresh the short-length key pairs with the aid of the blockchain as well as efficient algorithms for accelerating the software implementation of the BLS threshold signature scheme at the medium security level.
The rest of the paper is organized as follows. Section 2 gives an overview of elliptic curve and pairing-based cryptography as well as the ECDSA and (threshold) BLS signature schemes, followed by the detailed description of the short-lived signature-based PBFT variant and the corresponding blockchain-aided key distribution schemes in Section 3. In Section 4, we describe various optimization techniques for efficient software implementations of two short-lived signature schemes. Section 5 presents our experimental results for the performance of the short-lived signature schemes under the PBFT setting, which is followed by a summary of the related work in Section 6. Finally, we conclude this paper in Section 7.

2 PRELIMINARIES

2.1 Elliptic Curve Cryptography and ECDSA

2.1.1 Elliptic Curve Cryptography. Let \( \mathbb{F}_q \) be a finite field with \( q = p^l \) elements, where \( p \) is a prime and \( l \) is a positive integer. An elliptic curve \( E(\mathbb{F}_q) \) is the set of solutions \((x, y)\) over \( \mathbb{F}_q \) satisfying an equation of the form

\[
E(\mathbb{F}_q) : y^2 + a_1 x + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6,
\]

where \( a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q \) and the curve discriminant \( \Delta \neq 0 \), together with an additional point \( \text{point-at-infinity} \), denoted by \( O \). The points on an elliptic curve form an additive Abelian group, where \( O \) is the identity element and the group operation is defined by the well-known chord-and-tangent rule [20]. The order of \( E(\mathbb{F}_q) \) is denoted by \( \#E(\mathbb{F}_q) \) and the order of a point \( P \in E(\mathbb{F}_q) \) is defined as the smallest non-negative integer \( n \) such that \( nP = O \), where \( n \neq E(\mathbb{F}_q) \) and \( nP \) is the addition of the point \( P \) to itself \( n - 1 \) times. Let \( G = \langle P \rangle \) be a cyclic subgroup of \( E(\mathbb{F}_q) \) generated by the point \( P \), such that the Elliptic Curve Discrete Logarithm Problem (ECDLP) is intractable. For more details about elliptic curve cryptography, the interested reader is referred to [20].

2.1.2 Elliptic Curve Digital Signature Algorithm (ECDSA). The ECDSA is the elliptic curve analogue of the Digital Signature Algorithm (DSA). Let \( G = \langle P \rangle \) be a cyclic subgroup of \( E(\mathbb{F}_q) \) generated by the point \( P \) with prime order \( n \) and identity element \( O \), and \( H : \{0,1\}^* \to \mathbb{F}_q^* \) be a collision-resistant hash function. The ECDSA works as follows:

- **KeyGen:** Choose a random \( d^* \xleftarrow{\$} \mathbb{Z}_n \), set \( Q = d^*P \) and output a private/public key pair \((sk, pk) = (d^*, Q)\).
- **Sign:** Given a private key \( sk \) and a message \( m \in \{0,1\}^* \), the signer does the following:
  1. Select a random \( k' \xleftarrow{\$} \mathbb{Z}_n \) and compute \( R = k'P = (x_1, y_1) \).
  2. Compute \( r = x_1 \mod n \). If \( r = 0 \) then go to step 1.
  3. Compute \( s = k'^{-1}(H(m) + d^* \cdot r) \mod n \). If \( s = 0 \) then go to step 1.
  4. Output the signature \((r,s)\).
- **Verify:** Given a public key \( pk \), a message \( m \), and the signature \((r,s)\), the verifier does the following:
  1. Check that \( r, s \in [1, n-1] \). If any verification fails, reject the signature.
  2. Compute \( w = s^{-1} \mod n \), \( u_1 = H(m) \cdot w \mod n \) and \( u_2 = r \cdot w \mod n \).
  3. Compute \( R' = u_1P + u_2Q = (x_1', y_1') \). If \( R' = O \), reject the signature. Otherwise, compute \( v = x_1' \mod n \).
  4. Accept the signature if and only if \( v = r \).

2.2 Bilinear Pairing and BLS Signature

2.2.1 Bilinear Pairing. Let \( n \) be a positive integer and \( G_1 \) and \( G_2 \) be additively-written groups of order \( n \) with identity \( O \), and let \( G_T \) be a multiplicatively-written group of order \( n \) with identity 1. A bilinear pairing \( \langle \cdot, \cdot \rangle \) or pairing for short is a computable, non-degenerate function \( e : G_1 \times G_2 \to G_T \) satisfying the additional properties. The most important property for cryptographic applications is so-called bilinearity, namely:

\[
e(aP, bQ) = e(P, bQ)^a = e(P, Q)^ab,
\]

for all \( P \in G_1, Q \in G_2 \) and \( a, b \in \mathbb{Z}_n \). In practice, the groups \( G_1 \) and \( G_2 \) are subgroups or quotient groups of an elliptic curve defined over a finite field \( \mathbb{F}_q \) or one of its extensions and \( G_T \) is a subgroup or quotient group \( \mathbb{F}_q^* \), where \( k \) is called the embedded degree. The security of pairing-based cryptography requires that the discrete logarithm problem on \( G_1, G_2 \) and \( G_T \) is sufficiently difficult. For more details about the bilinear pairing and its cryptographic applications, the interested reader is referred to [16, 32].

2.2.2 The BLS Signature Scheme. In [6], Boneh et al. described a simple, deterministic short signature scheme, namely the BLS short signature. It works in any Gap Diffie-Hellman (GDH) group and requires a hash function from the message space onto the group. Let \( q_1, q_2 \) and \( q_T \) be an arbitrary generator of \( G_1, G_2 \) and \( G_T \), respectively, and \( H_1 : \{0,1\}^* \to G_1 \) be a hash function with values in \( G_1 \). The BLS signature scheme consists of the following three algorithms:

- **KeyGen:** Choose a random \( \alpha \xleftarrow{\$} \mathbb{Z}_n \), set \( h \xleftarrow{\$} q_2^\alpha \in G_2 \) and output a private/public key pair \((sk, pk) = (\alpha, h)\).
- **Sign:** Given a private key \( sk \) and a message \( m \in \{0,1\}^* \), output \( \sigma = H_1(m)^sk \in G_1 \). Note that the signature \( \sigma \) is a single group element.
- **Verify:** Given a public key \( pk \), a message \( m \), and a signature \( \sigma \), check whether \( e(\sigma, q_T) = e(H_1(m), pk) \) and output "accept" or "reject" accordingly.

Due to the simple mathematical structure, the BLS signature scheme supports a variety of extensions [4, 5], including threshold signatures, multisignatures, aggregate signatures, and blind signatures.

2.2.3 The BLS Threshold Signature Scheme. In [5], Boneh et al. showed that one can build a non-interactive threshold signature scheme based on the plain BLS signature scheme and the Shamir’s secret sharing [40]. For a system with \( N \) entities \( P_1, \ldots, P_N \), any of which may be corrupted, and a trusted dealer TD, the BLS threshold signature consists of the following five algorithms:

- **KeyGen:** TD chooses a random \( \alpha \xleftarrow{\$} \mathbb{Z}_n \) and set \( h \xleftarrow{\$} q_2^\alpha \in G_2 \). The system private/public key pair is \((sk, pk) = (\alpha, h)\). TD also generates a random polynomial \( f(z) = (z) \in \mathbb{Z}_n \) of degree \( t \), such that \( f(0) = sk \). TD then computes \( N \) private key shares \( sk_i = f(id_i) \) and public key shares \( pk_i = q_2^{sk_i} \) for \( i \in [1, N] \). The private key share \( sk_i \) as well as the public key share \( pk_i \) are distributed to the system’s members.
At the end of the three phases, all the honest nodes in the consensus group have the same view regarding the state of blockchain by either accepting or rejecting the block proposal, thereby achieving the instant transaction finality. In PBFT, both the Prepare and Commit phases involve intensive interactions among nodes within the consensus group. Given that a consensus group of N players can tolerate f = ⌊(N − 1)/3⌋ Byzantine nodes in PBFT, each honest node needs to collect and verify 2f + 1 digital signatures for both phases, respectively, which limits the application of PBFT to small consensus groups due to significant communication and computational overhead. For improving the scalability of PBFT and reducing the computational cost, it is crucial to speeding up the signature verification process. One key observation is that one round of PBFT within a small consensus group can be completed in as little as a few seconds. As a result, in lieu of utilizing digital signatures at the high security levels (e.g., 128-bit), nodes can sign the PREPARE and COMMIT messages with short-lived private keys at the low or medium security levels (e.g., 56- or 80-bit) consecutively for a number of rounds before refreshing those keys. Introducing short-lived key pairs into PBFT brings the following benefits:

- The signature size is smaller, which reduces the communication overhead;
- The signature verification is faster, which reduces the computational overhead.

The lifetime of a short-lived key pair depending on the difficulty of breaking underlying cryptosystems with the state-of-the-art cryptanalysis techniques and appropriate security margins should be kept when using short-lived key pairs in practice.

### 3 SCALABLE PBFT USING SHORT-LIVED SIGNATURE SCHEMES

#### 3.1 Key Observation on the PBFT Algorithm

PBFT runs in rounds with one node acting as a leader and others as validators in each round. Every round of PBFT consists of the following three phases as illustrated in Figure 1:

1. **Pre-prepare phase**: The leader initiates the consensus process by sending a signed PRE-PREPARE message that is a block proposal containing a certain number of transactions.
2. **Prepare phase**: Upon receiving the PRE-PREPARE message, every node in the consensus group checks the correctness and validity of the block and multicasts a signed PREPARE message (i.e., ‘YES/NO’) to all the other nodes.
3. **Commit phase**: Based on the analysis of the received PREPARE messages, each node multicasts a signed COMMIT message (i.e., ‘YES/NO’) to the consensus group. The block proposal is committed to the blockchain only if a sufficient number of nodes in the consensus group agree on it.

![Figure 1: Overview of Three Phases in PBFT](image_url)

At the end of the three phases, all the honest nodes in the consensus group have the same view regarding the state of blockchain by
3.3 Blockchain-Aided Key Distribution

One remaining challenge for using short-lived signature schemes in PBFT is to refresh the key pair after a certain number of rounds. We note that a blockchain does provide a reliable communication channel for handling key distribution in a dynamic and distributed network environment where entities might behave arbitrarily. Based on this observation, we describe efficient, blockchain-aided key distribution techniques for the ECDSA and BLS threshold signature schemes in this subsection. We assume that node $i$ starts using a new, short-lived key pair $(sk_i, pk_i)$ at round $L$ for the consecutive $T$ rounds, during which each node in the consensus group acts as the leader exactly $K$ times (i.e., $K = T/N$).

3.3.1 Blockchain-Aided Key Distribution for the ECDSA Signature Scheme. To update the short-lived key pair $(sk_i, pk_i)$ in ECDSA, node $i$ in the consensus group does the following:

1. Once node $i$ becomes the leader for the first time, it generates a new key pair $(sk'_i, pk'_i)$ as well as a secret key $k_i$.
2. Node $i$ encrypts the new public key $pk'_i$ using a symmetric-key algorithm (e.g., AES) and adds the ciphertext $Enc_{k_i}(pk'_i)$ into the ‘extra data’ field in its block proposal.
3. If the block fails to be appended to the blockchain for some reason, node $i$ will repeat the Step 2 next time it acts as the leader.
4. When node $i$ works as the leader for the $K_1$-th time ($K_1 < K$), it aims to release the secret key $k_i$ to the public by adding it into the ‘extra data’ field in its block proposal.
5. If the block cannot pass the consensus process, node $i$ will repeat the Step 4 next time it becomes the leader. This step continues until the block containing the secret key $k_i$ has been appended to the blockchain successfully.
6. After the block proposal containing the secret key $k_i$ occurs in the blockchain, other nodes in the consensus group are able to decrypt the ciphertext $Enc_{k_i}(pk'_i)$ and retrieve the updated public key $pk'_i$ of node $i$.
7. Node $i$ will be replaced by another node in the system provided that it fails to distribute the updated public key $pk'_i$ to its peers during the lifetime of the key pair $(sk_i, pk_i)$.

This step continues until the block containing the ciphertext $Enc_{k_i}(pk'_i)$ has been committed in the blockchain successfully.

The above two-phase key distribution process is illustrated in Figure 2a. Essentially every node in the consensus group first generates a new short-lived key pair and commits the encrypted public key to the blockchain, followed by the release of the secret key for decryption towards the end of life of the old key pair. If a node fails to distribute the updated public key, it is highly likely that the node is under attack and should be evicted from the consensus group.

3.3.2 Blockchain-Aided Key Share Distribution for the BLS Threshold Signature Scheme. Due to the lack of a trusted dealer in blockchain networks, the key updating for the BLS threshold signature scheme involves a distributed key generation (DKG) protocol such as the
one proposed by Pedersen in [34]. The Pedersen’s DKG scheme requires every node in the consensus group to distribute the shares of a randomly generated secret to all its peers. To update the short-lived key pair share \((sk_i, pk_i)\) in the BLS threshold signature scheme, node \(i\) in the consensus group does the following:

1. Once node \(i\) becomes the leader for the first time, it chooses at random a polynomial \(f_i(z)\) of degree \(t\) over \(\mathbb{Z}_n^i:\)
   \[
   f_i(z) = a_0 + a_1z + \cdots + a_tz^t,
   \]
   where \(f_i(0) = a_0 = \bar{x}_i\) is a random secret that node \(i\) selects. Node \(j\) then computes the shares \(\bar{x}_{ij} = f_j(\bar{x}_j) \pmod n\) for \(j \in \{1, N\}, j \neq i\), where \(\bar{x}_j\) is the identifier of node \(j\).

2. For \(j \in \{1, N\}, j \neq i\), node \(i\) encrypts the shares \(\bar{x}_{ij}\) using a public-key algorithm and obtains a set of ciphertexts \(\{\text{Enc}_{pk_j}(\bar{x}_{ij})\}\).

3. If the block fails to be appended to the blockchain for some reason, node \(i\) will repeat the Step 2 next time it acts as the leader. This step continues until the block containing the ciphertext and witness sets has been committed in the blockchain successfully.

4. When all the peers in the consensus group complete the above steps, node \(i\) is able to compute the new key pair share \((sk_i', pk_i')\) by following the Pedersen’s DKG scheme.

5. Node \(i\) will be replaced by another node in the system provided that it is marked by its peers as ‘disqualified’ in the DKG scheme or fails to commit the block containing the ciphertext and witness sets during the lifetime of the key share pair \((sk_i, pk_i)\).

The above key share distribution process is illustrated in Figure 2b. Basically every node in the consensus group acts as the trusted dealer and commits the encrypted shares of a randomly selected secret to the blockchain. By collecting the corresponding shares from the blockchain, each node is able to calculate the new short-lived key pair share. If a node provides incorrect shares or fails to distribute the shares, it will be evicted from the consensus group due to potential attacks.

4 OPTIMIZATION TECHNIQUES FOR IMPLEMENTING SHORT-LIVED SIGNATURE SCHEMES

In this section, we discuss various optimization techniques used in our implementations of the short-lived ECDSA and BLS threshold signature schemes.

4.1 Efficient Implementation of the ECDSA Signature Scheme

The parameters of the two elliptic curves secp112r1 and sect163k1 are summarized in Appendix A.1 and A.2, respectively. While secp112r1 is an elliptic curve defined over a 112-bit prime field, sect163k1 is a Koblitz curve defined over a 163-bit binary field. For efficient implementation of elliptic curve arithmetic, we utilize various optimization techniques proposed in the literature such as the lambda coordinates [33], the window-Non-Adjacent Form (NAF) method [20], the pre-computation [43], etc. Furthermore, BLAKE2 hash function [35] is used in our implementation due to its high performance on different computing platforms.

For speeding up the verification of multiple ECDSA signatures in the PBFT algorithm, we implement the efficient batch verification technique proposed in [10], which is applicable to the modified ECDSA scheme in [1]. The security of the modified ECDSA scheme is equivalent to the standard ECDSA, but the modified signature is the pair \((R, s)\) instead of \((r, s)\) in the original scheme.

4.2 Efficient Implementation of the BLS Threshold Signature Scheme

The parameters of the MNT curve with embedding degree \(k = 6\) is summarized in Appendix A.3. To improve the performance of the threshold BLS signature scheme, we propose a multibase variant of Miller’s algorithm for computing the Tate pairing and generalize the method of encapsulating the computation of the line function with the group operations described in [9] to the multibase case, as detailed in the following subsections.

4.2.1 Tate Pairing on Elliptic Curves. Tate pairing is defined in terms of divisors of a rational function. For our purpose, a divisor is a formal sum \(\sum_{P \in \mathbb{P}} a_P P\) of points on the curve \(E(F_{q^k})\). The degree of a divisor \(D\) is the sum \(\deg(D) = \sum_{P \in \mathbb{P}} a_P\). The set of divisors forms an Abelian group by the addition of corresponding coefficients in their formal sums. Let \(f : E(F_{q^k}) \to E(F_{q^k})\) be a rational function on the elliptic curve, then the divisor of \(f\) is \(\text{div}(f) = \sum_{P \in \mathbb{P}} \text{ord}_P(f)(P)\), where \(\text{ord}_P(f)\) is the order of the zero or pole of \(f\) at \(P\). Let \(D\) be a divisor of degree zero, the evaluation of the rational function \(f\) at \(D\) is defined as \(f(D) = \prod_{P \in \mathbb{P}} f(P)^{a_P}\). A divisor \(D\) is called a principal divisor if \(D = \text{div}(f)\) for some rational function \(f\). A divisor \(D = \sum_{P \in \mathbb{P}} a_P P\) is principal if and only if \(\deg(D) = 0\) and \(\sum_{P \in \mathbb{P}} a_P P = O\). Two divisors \(D_1\) and \(D_2\) are equivalent (i.e., \(D_1 \sim D_2\)) if their difference \(D_1 - D_2\) is a principal divisor. Let \(P \in E(F_{q^k})[n]\) where \(n\) is coprime to \(q\) and \(Q \in E(F_{q^k})\). Let \(D_P\) be a divisor equivalent to \((P) - (Q)\) and therefore there is a rational function \(f_{n,p} \in F_q(\mathbb{P})^*\) such that \(\text{div}(f_{n,p}) = nD_P = n(P) - n(O)\). Let \(D_Q\) be a divisor equivalent to \((Q) - (O)\) with its support disjoint from \(\text{div}(f_{n,p})\). The Tate pairing [15] is a well-defined, non-degenerate, bilinear map \(e_n : E(F_{q^k})(n) \times E(F_{q^k})(n) \to \mathbb{F}_{q^k}^*/(\mathbb{F}_{q^k})^*\) given by \(e_n(P,Q) = f_{n,p}(D_Q)\). The computation of \(f_{n,p}(D_Q)\) is achieved by an application of Miller’s algorithm [30], whose output is only defined up to \(n\)-th powers in \(\mathbb{F}_{q^k}^*\) which is usually undesirable in practice since many pairing-based protocols require a unique pairing value. According to Theorem 1 in [2], one can define reduced Tate pairing as \(e(P, Q) = f_{n,p}(Q)^{q^k-1/n}\). The extra powering required to compute the reduced pairing is referred to as the final exponentiation.

4.2.2 Multibase Number Representation. Various multibase number representations (MSRs) have been proposed to accelerate elliptic curve scalar multiplication, see [12, 13, 25, 28] for examples. The basic idea behind MSR is to record a scalar in a very compact and sparse form and therefore significantly reduce the number of point additions during the computation of scalar multiplication. In [28],
Longa and Miri introduced the following generic multibase representation for a scalar $n$:

$$n = \sum_{i=1}^{m} n_i \prod_{j=1}^{f_i} a_i^{c_i(j)},$$

where (1) bases $a_1 \neq a_2 \neq \cdots \neq a_f$ are positive primes ($a_1$ is called the main base) and $m$ is the length of the expansion; (2) $n_i$ are signed digits from a given set $D$; (3) $c_1(j) \geq c_2(j) \geq \cdots \geq c_m(j) \geq 0$ for each $j$ from 2 to $f$; and (4) $c_1(1) > c_2(1) > \cdots > c_m(1) > 0$. Based on the above MSR, Longa et al. [25, 28] proposed Multibase Non-Adjacent Form (mbNAF) and its window-based variants including Window-$w$ Multibase Non-Adjacent Form (wmbNAF) and Fractional Window-$w$ Multibase Non-Adjacent Form (Frac-wmbNAF). Combined with optimized composite operations and precomputation schemes [27], the multibase methods have set new speed records for computing the elliptic curve scalar multiplication [25].

### 4.2.3 A Multibase Variant of Miller’s Algorithm

We propose a multibase variant of Miller’s algorithm and show how to efficiently extract the required rational functions in this case. Considering that one inversion is required for each group operation in the process of computing Tate pairings, and the calculation of the inversion of an element in large characteristic is usually quite expensive, we use Jacobian coordinates to represent points on an elliptic curve instead of affine coordinates. Moreover, we also employ the Frac-wmbNAF method [25] to record the scalar, which has been shown to achieve the highest performance among window-based methods for standard elliptic curves in Weierstrass form.

Let $\mathcal{A} = \{a_1, a_2, \ldots, a_f\}$ be a set of bases, where $a_1 = 2$ and $a_2 \neq a_3 \neq \cdots \neq a_f$ are positive odd primes. Let $D = \{0, \pm 1, \pm 3, \ldots, \pm t\}$ be a digit set, where $t \geq 3$ is an odd integer. Let $P \in E(\mathbb{F}_q)[n]$ and $Q \in E(\mathbb{F}_q^k)$, where $n$ is a prime. Assume that $n$ is represented by the Frac-wmbNAF method as $(n_1^{(b_1)}, \ldots, n_t^{(b_t)}, 0_0^{(b_0)})$, where $n_i^{(b_i)} \in D$ is the $i$-th digit and the superscript $b_i \in \mathcal{A}$ denotes the base associated to the corresponding digit for $0 \leq i \leq l$. With the above notations, the multibase variant of Miller’s algorithm is described in Algorithm 1. The algorithm includes two phases: the precomputation phase that calculates the required points and evaluates the corresponding rational functions at the image point $Q$, and the main loop that uses the results of the precomputation phase to compute the Tate pairing efficiently.

We now explain how to perform the precomputation and the main loop efficiently. Let $I$, $M$ and $S$ denote an inversion, a multiplication and a squaring in $\mathbb{F}_q$, and let $M_k$ and $S_k$ denote a multiplication and a squaring in the large field $\mathbb{F}_q^k$. We also assume that an elliptic curve $E$ over $\mathbb{F}_q$ admits a twist $E'$ of degree $w$ with $w \mid k$. Letting $Q = D$ or $W$, we can take $Q$ to be a point on the twist curve $E'(\mathbb{F}_q^k)$ in this case, which allows us to apply the efficient denominator elimination technique due to Barreto, Lynn and Scott [3]. Using twist curves, we can work within the groups $E(\mathbb{F}_q)[n]$ and $E'(\mathbb{F}_q^k)$ at all times except when a pairing is being evaluated, where we use the twist map and operate in $\mathbb{F}_q^k$.

#### 4.2.4 Precomputation Method

The precomputed points $P_i$ and function values $f_i$ are extensively used to accelerate the computation of Tate pairing in the above multibase variant of Miller’s algorithm. Longa and Miri [27] proposed a highly efficient precomputation scheme for elliptic curve cryptosystems over prime fields, which is based on the combination of the traditional chain $P \rightarrow 2P \rightarrow 3P \rightarrow \cdots \rightarrow tP$ and the special point addition formula with the same $z$-coordinate introduced by Meloni [29]. In this subsection, we show how to efficiently obtain the function values $f_{si} = f_{si, p}(Q)$ for $i \in \{1, 3, \ldots, t\}$ and $Q \in E'(\mathbb{F}_p)$ based on the precomputation scheme in [27]. Considering that the precomputed points $(P, 2P, \ldots, tP)$ are stored in affine coordinates, it is not hard to find that $f_2$ can be computed more efficiently in affine coordinates instead of projective coordinates. Assuming that $J^P = (x_p, y_p)$ for any $j \in \mathbb{Z}$ and $Q = (x_Q, y_Q)$, we first have $f_1 = f_{1, p}(Q) = 1$ and $f_{-1} = f_{-1, p}(Q) = \frac{x_Q - x_p}{y_Q}$. Using the parabola method proposed by Eisentraeger et al. [14], we can calculate $f_{3, 5}$ simultaneously as follows:

$$f_3 = f_{3, p}(Q) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} f_{3, p}(Q) = \frac{(x_Q - x_p)^2 + (x_Q - x_p) + (x_Q - x_p) + (x_Q - x_p) + (x_Q - x_p)}{x_Q - x_p} = \frac{(x_Q - x_p)^2 + (x_Q - x_p) + (x_Q - x_p) + (x_Q - x_p) + (x_Q - x_p)}{x_Q - x_p}.$$

where $\lambda_1$ and $\lambda_2$ are the slopes of the lines $\ell_p, \ell_{2p}, \ell_{3p}, \ell_{4p}, \ell_{5p}$, respectively. Using the fact that $-J^P = (x_{-p}, y_{-p})$ for any $j \in \mathbb{Z}$, we can obtain $f_{-3}$ virtually for free by reusing the intermediate results during the computation of $f_3$:

$$f_{-3} = \frac{(x_Q - x_p)^2 + (x_Q - x_p) + (x_Q - x_p) + (x_Q - x_p) + (x_Q - x_p)}{x_Q - x_p}.$$

#### Algorithm 1: A Multibase Variant of Miller’s Algorithm for Computing Tate Pairing

**Input:** $n = (n_1^{(b_1)}, \ldots, n_t^{(b_t)}, 0_0^{(b_0)})$, where $n_i^{(b_i)} \in D$ and $b_i \in \mathcal{A}$

$P = (x_P, y_P) \in E(\mathbb{F}_q)[n]$ and $Q = (x_Q, y_Q) \in E(\mathbb{F}_q^k)$.\[Main Loop\] $f(Q)$

**Output:** $e(P, Q) = f_{n, p}(Q)q^{(t-1)/n}$

**Precomputation:**

1. Compute $P_1 = iP$ for $i \in \{1, 3, \ldots, t\}$, $t \geq 3$ is an odd integer.
2. Compute $f_{si} = f_{si, p}(Q)$ for $i \in \{1, 3, \ldots, t\}$, where $d(iP) = i(P) - (i-1)(O)$.

**Main Loop:**

3. for $i \leftarrow l - 1$ downto 0 do

4. if $b_i = 2$, then $f_t \leftarrow f_t \cdot \frac{l_{r_t}(Q)}{v_{q_t}(Q)}, T \leftarrow 2T$

5. else $f_t \leftarrow f_t \cdot \frac{l_{b_i}(Q)}{v_{q_t}(Q)}, T \leftarrow l_{b_i}T$

6. for $i \leftarrow 0$ downto 0 do

7. if $n_i^{(b_i)} \neq 0$ then

8. if $n_i^{(b_i)} > 0$, then $P_t \leftarrow P_t \cdot Q_{n_i^{(b_i)}}$; else $P_t \leftarrow -P_t$.

9. $f_t \leftarrow f_t \cdot f_{n_i^{(b_i)}}, T \leftarrow T + P_t$

10. return $f_t (q^{(t-1)/n})$
Note that \((3x_2^2 + a)\) and \(2y_2^2\) can be obtained from the computation of \(2P\) and the denominator \((x_P - x_{2P})(x_Q - x_{3P})\) can be eliminated by the final exponentiation. Therefore, we only need to compute the numerators of \(f_{3s}\) and the computation cost is \(dM\), with a precomputation of \(15d + (d + k)M\).

Similarly, for an odd integer \(s\) satisfying \(5 \leq s \leq t\), \(f_{3s}\) can be computed simultaneously as follows:

\[
f_{3s} = f_{3s}, p(Q) = f_2 \cdot f_{s-2} \cdot \frac{l_{2P, (s-2)}p(Q)}{v_2p(Q)} = f_2 \cdot f_{s-2} \cdot \frac{(y_Q - y_{2P}) - \lambda_s(x_Q - x_{2P})}{v_2p(Q)} = f_2 \cdot f_{s-2} \cdot \frac{(x_{(s-2)p} - x_{2P})(y_Q - y_{2P}) - (y_{(s-2)p} - y_{2P})(x_Q - x_{2P})}{(x_{(s-2)p} - x_{2P})(x_Q - x_{s_1})},
\]

where \(\lambda_s\) is the slope of the line \(l_{2P, (s-2)}p\). Hence, given \(f_{3(s-2)}\), we need \(2M_k + (k + d + 1)M\) to calculate the numerators of \(f_{3s}\) simultaneously.

We summarize the precomputation scheme in Algorithm 2. Again, we only consider the computations of the numerators of \(f_{3i}\) due to the efficient denominator elimination technique [3]. Note that \(P_i, i \in \{1, 3, \cdots, t\}\) can be calculated with \(1f + \frac{R_i + 1}{2}M + (t + 5)S\) using the precomputation method (Scheme 2) in [27]. Moreover, we also need to cost another \(4M\) to recover the affine coordinates of \(2P\) that are used in the computation of \(f_{3i}\). Therefore, the total cost of the precomputation scheme in Algorithm 2 is \((t-3)M_k + 1M_d + 1f + \left\{ (5t - 2) + \frac{(k + d)(t - 1)}{2} \right\} M + (t + 5)S\).

**Algorithm 2. Precomputation Scheme for the Multibase Variant of Miller’s Algorithm**

**Input:** \(P = (x_P, y_P) \in E(\mathbb{F}_q)[n]\) and \(Q = (x_Q, y_Q) \in E(\mathbb{F}_q)[k]\)  
**Output:** \(P_i = iP\) and the numerators \(f_{3i} = f_{3s}, p(Q)\) for \(i \in \{1, 3, \cdots, t\}\), where \(t \geq 3\) is an odd integer and \(\text{div}(f_{3i}) = \text{div}(P) = (ip) - (i - 1)Q\).

1. Compute \(P_i = iP = (x_{ip}, y_{ip})\) for \(i \in \{1, 3, \cdots, t\}\) and \(2P = (x_{2P}, y_{2P})\) using the precomputation method proposed in [27] (see pp. 238-240 of [27]).
2. \(f_1 = 1, f_{-1} = 1\)
3. \(T_1 \leftarrow \left\{ 2y_{2P} + 3(x_P^2 + a)(x_Q - x_P) \right\}, T_2 \leftarrow 2y_{2P}y_Q, f_{-1} \leftarrow T_2 + T_3, f_{-2} \leftarrow T_1 \leftarrow T_2 + T_3\)
4. \(T_1 \leftarrow x_Q - x_{2P}, T_2 \leftarrow T_1 \left\{ (x_P - x_{2P})y_1 + (3x_P^2 + a) + 2y_{2P} \right\}, T_3 \leftarrow 2y_{2P}y_Q, f_{s-3} \leftarrow T_2 + T_3, f_{s-2} \leftarrow T_2 + T_3\)
5. **for** \(s = 5, 7, 9, 11, \cdots, t\) **do**
6. \(T_1 \leftarrow x_{(s-2)p} - x_{2P}, T_2 \leftarrow T_1 y_Q, T_3 \leftarrow T_1 y_{2P} + (y_{(s-2)p} - y_{2P})(x_Q - x_{2P})\)
7. \(f'_{s-3} \leftarrow f'_{s-1} \left\{ (T_2 - T_3), f_s = f'_{s-2} \cdot f'_{(s-2)}, \cdot (T_2 + T_3) \right\} s \leftarrow s + 2\)
8. **return** \(P_i = f'_{3i}, i \in \{1, 3, \cdots, t\}\)

### 4.2.5 Encapsulated Composite Operations and Line Computations

In [9], Chatterjee et al. introduced the idea of encapsulating elliptic curve group operations (i.e., point addition and point doubling) and line computations, and applied this idea to improve the computation of Tate pairing for two families of pairing-friendly curves with embedding degree 2. It is also straightforward to adapt this idea to other pairing-friendly curves which admit twists of degree \(w\) with \(w \mid k\). Moreover, fast doubling and mixed-addition formulae proposed in [26] can be used to further improve Chatterjee et al.’s algorithms. In this subsection, we show how to efficiently perform the main loop in the multibase variant of Miller’s algorithm by encapsulating fast composite operations and line computations.

**Encapsulated Point Mixed-Addition and Line Computation:**

Let \(T = (X_1, Y_1, Z_1)\) and \(P' = (x_2, y_2)\) be two points in Jacobian and affine coordinates, respectively, on the elliptic curve \(E(\mathbb{F}_q)\). The mixed addition \(T + P' = (X_3, Y_3, Z_3)\) can be computed efficiently as follows [26]:

\[
X_3 = a_2^2 - 4b_2^3 - 8x_1b_2^2,
Y_3 = a(4X_2b_2^2 - X_1) - 8Y_1b_3^3,
Z_3 = (Z_1 + b_2^2)(Z_3),
\]

where \(a = 2(Z_1^2y_2 - Y_1)\) and \(b = Z_1^2x_2 - X_1\). In this case, the evaluation of the line function \(l_{T, P'}\) at the image point \(Q\) can be calculated as follows:

\[
l_{T, P'}(Q) = \frac{y_Q - y_2}{z_3 - z_2} \cdot (x_Q - x_2) = Z_3y_Q - ax_Q + (ax_2 - z_3y_2),
\]

where \(Z_3\) and \(a\) are available from the point mixed-addition. The encapsulated point mixed-addition and line computation are given in Algorithm 3, which requires \((9 + d + k)M + 4S\). Note that in the case of \(d = 1\) (i.e., \(x_Q \in \mathbb{F}_q\)) we can save one more field multiplication by combining the terms involving \(a\) in the above equation (1).

**Encapsulated Point Doubling and Line Computation:**

Let \(T = (X_1, Y_1, Z_1)\) be a point in Jacobian coordinates on the elliptic curve \(E(\mathbb{F}_q)\). The point doubling \(2T = (X_3, Y_3, Z_3)\) can be computed efficiently as follows [26]:

\[
X_3 = a_2^2 - 2b_1,
Y_3 = a(\beta - X_3) - 8Y_1^4,
Z_3 = (Y_1 + Z_1^2)^2 - Y_1 - Z_1^2,
\]

where \(a = 3X_1^2 + aZ_1^2\) and \(\beta = 2 \left( (X_1 + Y_1^2)^2 - X_1^2 - Y_1^4 \right)\) for a general \(a\), and \(a = 3(X_1 + Z_1^2)^4(X_1 - Z_1^2)\) for a general \(a\), and \(b = 4X_1Y_1^2\) if \(a = -3\). In this case, the evaluation of the line function \(l_{T, T'}\) at the image point \(Q\) can be calculated as follows:

\[
l_{T, T'}(Q) = \frac{y_Q - y_1}{z_3 - z_1} \cdot 3X_1^2 + aZ_1^2 \cdot \frac{x_Q - X_1}{Z_1^2} = (Z_3Z_1^2y_Q - ((2Y_1^2 - ax_1) + (aZ_1^2)x_Q)) \cdot Z_3Z_1^2,
\]

where \(Z_3\) and \(z_3\) are available from the point mixed-addition.
where $Z_3, Z_1^2, Y_2^2$ and $a$ are available from the point doubling. The encapsulated point doubling and line computation are given in Algorithm 3, where some intermediate results of the point doubling, as shown in the boxes, can be reused for the line computation. Algorithm 4 requires $(5 + d + k)M + 8S$ to compute the point doubling and evaluate the line function. In the case $a$ is small, this cost is $(4 + d + k)M + 8S$ and for the case $a = -3$, this cost is $(6 + d + k)M + 8S$. Note that in the case of $d = 1$ (i.e., $x_Q \in \mathbb{F}_q$) we can save two more field multiplications by combining the terms involving $\alpha$ in the above equation (2).

Encapsulated Point Tripling and Line Computation:
Let $T = (X_1, Y_1, Z_1)$ be a point in Jacobian coordinates on the elliptic curve $E/\mathbb{F}_q$. The point tripling $3T = (X_3, Y_3, Z_3)$ can be computed efficiently as follows [26]:

$$
X_3 = 16Y_1^2(2\beta - 2\alpha) + 4X_1\omega^2,
Y_3 = 8Y_1[(2\alpha - 2\beta)(4\beta - 2\alpha) - \omega^2],
Z_3 = (Z_1 + \omega^2 - Z_1^2 - \omega^2),
$$

where $2\alpha = (\theta + \omega^2 - \theta^2 - \omega^2, 2\beta = 16Y_1^4, \theta = 3X_1^2 + aZ_1^2$ and $\omega = 6\{X_1 + Y_1^2\}^2 - X_1^2 - Y_1^4\} - 2\theta$ for a general $a$. For the case when $a = -3$, $\theta$ and $\omega$ can be computed more efficiently as $\theta = 3(X_1 + Z_1^2)(X_1 - Z_1^2)$ and $\omega = 12X_1Y_1^2 - \theta^2$. Similar to the precomputation procedure, we can apply the parabola method [14] again to simplify the required line function $f_{3,T}$. However, in the main loop of Algorithm 2 the points $T$ and $Q$ are represented in Jacobian and affine coordinates, respectively. Letting $\lambda'_1$ and $\lambda'_2$ be the slopes of the lines $l_{T,T}$ and $l_{2T,T}$, we can obtain $\lambda'_1 = \frac{\theta}{Z_1^2}$ and $\lambda'_2 + \lambda'_2 = \frac{Y_1^3}{Z_1^{3\alpha}}$. In this case, the evaluation of the line function $f_{3,T}$

---

**Algorithm 3. Encapsulated Point Mixed-Addition and Line Computation**

**Input:** $T = (X_1, Y_1, Z_1)$ in Jacobian coordinates, $P' = (x_2, y_2)$ in affine coordinates on $E/\mathbb{F}_q$ and $Q = (x_Q, y_Q) \in E(\mathbb{F}_q^*)$

**Output:** $T + P' = (X_3, Y_3, Z_3)$ in Jacobian and the numerator $I_{ADD}$ of $I_{T,P'}(Q)$

**Point Mixed-Addition**

1. $T_1 \leftarrow Z_1^2$
2. $T_2 \leftarrow x_1\cdot T_1$
3. $T_3 \leftarrow X_1 + T_1$
4. $T_4 \leftarrow X_1 - T_1$
5. $T_5 \leftarrow Z_1$
6. $T_6 \leftarrow T_5\cdot T_1$
7. $T_7 \leftarrow T_6\cdot T_1$
8. $T_8 \leftarrow T_7\cdot T_1$
9. $T_9 \leftarrow T_8\cdot T_1$
10. $T_{10} \leftarrow T_9\cdot T_1$
11. $T_{11} \leftarrow T_{10}\cdot T_1$
12. $T_{12} \leftarrow T_{11}\cdot T_1$
13. $T_{13} \leftarrow T_{12}\cdot T_1$
14. $T_{14} \leftarrow T_{13}\cdot T_1$
15. $T_{15} \leftarrow T_{14}\cdot T_1$
16. $T_{16} \leftarrow T_{15}\cdot T_1$
17. $T_{17} \leftarrow T_{16}\cdot T_1$
18. $T_{18} \leftarrow T_{17}\cdot T_1$
19. $T_{19} \leftarrow T_{18}\cdot T_1$
20. $T_{20} \leftarrow T_{19}\cdot T_1$
21. $T_{21} \leftarrow T_{20}\cdot T_1$

**Line Computation**

1. $T_2 \leftarrow x_2\cdot T_1$
2. $T_3 \leftarrow x_2 - T_1$
3. $T_4 \leftarrow y_Q\cdot Z_3$
4. $T_5 \leftarrow T_4\cdot T_1$

**Return** $(X_3, Y_3, Z_3)$ and $I_{ADD}$

---

**Algorithm 4. Encapsulated Point Doubling and Line Computation**

**Input:** $T = (X_1, Y_1, Z_1)$ in Jacobian coordinates on $E/\mathbb{F}_q$ and $Q = (x_Q, y_Q) \in E(\mathbb{F}_q^*)$. 

**Output:** $2T = (X_3, Y_3, Z_3)$ in Jacobian coordinates and the numerator $I_{DBL}$ of $I_{T,T}(Q)$

**Point Doubling**

<table>
<thead>
<tr>
<th>$a = -3$</th>
<th>General $a$</th>
</tr>
</thead>
</table>
| 1. $T_1 \leftarrow Z_1^2$ | 9. $Z_3 \leftarrow Z_1^2$
| 2. $T_2 \leftarrow x_1 + T_1$ | 10. $Z_3 \leftarrow Z_1 - T_1$
| 3. $T_3 \leftarrow X_1^2$ | 11. $T_3 \leftarrow a - T_2$
| 4. $T_4 \leftarrow T_2 - T_3$ | 12. $T_4 \leftarrow X_1^2$
| 5. $T_5 \leftarrow T_1$ | 13. $T_5 \leftarrow 3\cdot T_1$
| 6. $T_6 \leftarrow T_2\cdot T_4$ | 14. $T_6 \leftarrow T_2\cdot T_4$
| 7. $T_7 \leftarrow T_3\cdot T_6$ | 15. $T_7 \leftarrow T_3\cdot T_6$
| 8. $T_8 \leftarrow T_3\cdot T_6$ | 16. $T_8 \leftarrow T_3\cdot T_6$
| 9. $T_9 \leftarrow T_3\cdot T_6$ | 17. $T_9 \leftarrow T_3\cdot T_6$
| 10. $T_{10} \leftarrow T_3\cdot T_6$ | 18. $T_{10} \leftarrow T_3\cdot T_6$

**Line Computation**

1. $T_4 \leftarrow Z_3\cdot T_1$
2. $T_5 \leftarrow T_1\cdot T_2$
3. $T_6 \leftarrow T_2\cdot T_3$
4. $T_7 \leftarrow T_3\cdot T_6$
5. $T_8 \leftarrow T_4\cdot T_1$
6. $T_9 \leftarrow T_4\cdot T_1$
7. $T_{10} \leftarrow T_5\cdot T_6$
8. $T_{11} \leftarrow T_6\cdot T_1$
9. $I_{DBL} \leftarrow T_k - T_d$

**Return** $(X_3, Y_3, Z_3)$ and $I_{DBL}$
at the image point $Q$ can be computed as follows:

\[ f_{3T}(Q) + \frac{I_{T, T}(Q)}{1 + I_{T, T}(Q)} \]

\[ = \left( x_Q - \frac{X}{Z^2} \right)^2 + (\theta_{1}^2 + \theta_{2}^2) \left( x_{Q} - \frac{X}{Z^2} \right) - \left( y_{Q} - \frac{Y}{Z^2} \right) \]

\[ = \frac{(\omega_3 Z^2 X^3 - X_3) \left[ (\omega_3 x_0 z_0^2 z - x_1 + 4Y^2 \theta^2 + 8Y^2 + (2Y_z Z_1)^2) y_0 \right]}{(\omega_3 X^3 Y_3^2 + Z_2^2 (4Y^2 \theta^2 - 2 \omega z_0 x_1) + x_1 (\omega_3 X_1 - 4Y^2 \theta^2)) + 8Y^2 + (2Y_z Z_1)^2} y_0 \]

\[ = \frac{a Z^2 (x_0 - x_{yT})}{a Z^2 (x_0 - x_{yT})}, \]

where $\theta_1, \theta_2, 4Y^2, Z^2, 8Y^4$ and $Z^2$ are available from the point tripling.

Due to space limitations, we only describe the encapsulated point tripling and line computation for the case $\alpha = -3$ in the following Algorithm 5, where some intermediate results of the point tripling, as shown in the boxes, can be reused for the line computation. However, it is also straightforward to derive the explicit formula for a general $\alpha$. Assuming that $x_0$ is precomputed with $15\Delta_2$, Algorithm 5 requires $(14 + 2d + k)M + 9S$ to calculate the point tripling and evaluate the line functions. In the case $a = 3$, this cost is $(12 + 2d + k)M + 11S$ and for a general $\alpha$, this cost is $(13 + 2d + k)M + 11S$.

Note that in the case of $d = 1$ or 2 (i.e., $x_0 \in F_p$ or $F_{p^2}$) it is more efficient to evaluate the line functions using equation (3) instead of (4), which can save $3M + 1S$ and $1M + 1S$, respectively.

we summarize the cost of our encapsulated composite operations and line computations as well as previous work for computing the Tate pairing in Table 1. When compared to the Chatterjee et al.’s work (where $k = 2$ and $d = 1$) [9], our formulae trade $2M$ for $2S$ in the encapsulated point doubling and line computation, and $1M$ for $1S$ in the encapsulated point mixed addition and line computation, respectively. We also give the computation cost of the encapsulated composite operations and line computations on general pairing-friendly curves.

### 4.2.6 Finite Field Arithmetic

For the MNT curve in our implementation, $p \equiv 1 \mod 3$ (where $\xi = -2$ is a cubic non-residue in $F_p$) and $y = -\sqrt{-2}$ is a quadratic non-residue in $F_{p^2}$. We can construct the extension field $F_{p^2}$ as a tower of finite extensions:

\[ F_{p^2} = \mathbb{F}_p[X]/(X^3 - \xi), \]
\[ F_{p^2} = \mathbb{F}_p[Y]/(Y^2 - y). \]

Hence an element $\eta \in F_{p^2}$ can be represented in any of the following three ways:

\[ \eta = \eta_0 + \eta_1 \sqrt[3]{\xi}, \]
\[ = \eta_0 + \eta_1 \sqrt[3]{\xi} + \eta_2 \sqrt[6]{\xi^2} + (\eta_1 + \eta_1 \sqrt[3]{\xi} + \eta_1 \sqrt[6]{\xi^2}) \sqrt[3]{\xi}, \]
\[ = \eta_0 + \eta_1 \sqrt[3]{\xi} + \eta_1 \sqrt[3]{\xi} + \eta_1 \sqrt[6]{\xi^2} + \eta_2 \sqrt[6]{\xi^2} + \eta_2 \sqrt[6]{\xi^2}. \]

where $\eta_0, \eta_1 \in F_{p^2}$ and $\eta_0, \eta_1, \eta_2, \eta_1, \eta_1, \eta_1, \eta_1, \eta_1 \in F_{p^2}$. With the above notations, the group homomorphism $\phi : E'(F_{p^2}) \rightarrow E(F_{p^2})$ defined as $\phi((x,y)) \mapsto (x,y\sqrt{\xi})$ can be computed at no cost. For the arithmetic in $F_{p^2}$, Karatsuba’s method reduces a multiplication and a squaring in $F_{p^2}$ to 6 multiplications and 6 squarings in $F_p$, respectively. Moreover, inversion in $F_{p^2}$ costs $1I + 9M + 3S$ in $F_p$ via the following direct inversion formula: if $\mu = \mu_0 + \mu_1 X + \mu_2 X^2 \in F_p$, then $\mu^{-1} = v_0 + v_1 X + v_2 X^2$ where $v_0 = (\mu_0^2 + 2\mu_1 \mu_2)A^{-1}, v_1 = -2(\mu_2 + 2\mu_1 \mu_2)A^{-1}, v_2 = (\mu_2 - 2\mu_1 \mu_2)A^{-1}$ and $A = \mu_0 (\mu_0^2 + 6\mu_1 \mu_2 + 2(2\mu_2^2 - \mu_1^2)).$

Since $F_{p^2}$ is a tower of cubic and quadratic extensions, it is more efficient to use Karatsuba over Karatsuba for multiplication, and use Complex over Karatsuba for squaring in $F_{p^2}$ based on the exhaustive testing described in [11]. Therefore, the computation cost for the multiplication and the squaring in $F_{p^2}$ is $18M$ and $12M$, respectively. Furthermore, inversion in $F_{p^2}$ can be reduced to 1 inversion, 2 multiplications, and 2 squarings in $F_p$.

### 4.2.7 Fast Hashing to $\mathbb{G}_2'$

To implement the BLS scheme on the MNT curve, we require a point $Q \in E(F_{p^2})$ of order $n$. Since
MNT curve has quadratic twists $E'(F_p)$, we only need to generate a point $Q' \in E'(F_p)$ of order $n$. The standard approach is to first generate a random point on $E'(F_p)$ and then multiply by the cofactor $c = \#E'(F_p)/n$ which is of a size in bits approximately the same as $p^2$. In [38], Scott et. al. proposed a fast method for the cofactor multiplication on $Q'_2 = \langle Q' \rangle$ using an efficiently computable homomorphism. Here, we adapt their idea to the MNT curve in question. Noting that $n = (z^2 - z + 1)/3, p = 2n + z$ and the trace of the Frobenius $x = 1, w$ can replace the cofactor $c$ in terms of $p$ and $z$ as follows:

$$c = \frac{\#E'(F_p)}{n} = \frac{p^3 + 1 + r^3 - 3rp}{n} = 2p^2 + (2z + 3)p - z - 2.$$

Let $\pi_p$ be the $p$-power Frobenius map on $E$ and $\phi$ be the group homomorphism defined above. Then $\psi = \phi^{-1}\pi_p\phi$ is an endomorphism of $E'$ such that $\psi : Q'_2 \rightarrow Q'_2$ and for any point $Q' \in Q'_2$ the identity $[n]Q' = [n]\psi(Q') = \psi^2(Q')$ holds [17]. Using this identity, we can calculate $cQ'$ as follows:

$$cQ' = [n]Q' = (2z + 3)p - z - 2.$$

Hence, the cost for computing $cQ'$ is one point doubling, one scalar multiplication by $2z$, three applications of the homomorphism $\psi$ and three point additions.

4.2.8 Final Exponentiation. For the MNT curve used in our implementation, the output of the multibase variant of Miller’s algorithm must be exponentiated to the power of $(p^6 - 1)/n$. The final exponentiation can be further broken down into the following three components:

$$(p^6 - 1)/n = (p^3 - 1) \cdot (p + 1) \cdot ((p^2 - p + 1)/n).$$

Let $f' = f_3' + f_3'\sqrt{p}$ be the output of Miller’s algorithm. Then exponentiating $f'$ to $p^3 - 1$ can be trivially computed with a conjugation with respect to $\mathbb{F}_{p^6}^*$, a multiplication and an inversion in $\mathbb{F}_{p^6}$. Using the basis described in Subsection 4.2.6, the conjugation of $f'$ with respect to $\mathbb{F}_{p^6}$ is calculated as $f'' = f_3' - f_3'\sqrt{p}$. The second part of the final exponentiation can be computed with an application of the Frobenius and one multiplication in $\mathbb{F}_{p^6}$. The remaining exponentiation to $(p^2 - p + 1)/n$ can be dealt with the method proposed in [37]. Noting that $n = (z^2 - z + 1)/3$ and $p = 2n + z$, we obtain the following relation:

$$p^2 - p + 1 = \frac{(2n + z)^2 - (2n + z) + 1}{n} = \frac{2n(2n + 2z - 1) + (2z^2 - z + 1)}{n} = 2(p + z) + 1.$$

Therefore, the last part of the final exponentiation can be calculated with an application of the Frobenius, one squaring, two multiplications, and a simple exponentiation to the power of $z$ in $\mathbb{F}_{p^6}$.

5 PERFORMANCE EVALUATION

In this section, we present the experimental results to validate the performance of the short-lived ECDSA and BLS threshold signature schemes under the PBFT setting. We implement three software libraries for the elliptic curves secp112r1, sect163k1 and the MNT curve, respectively. These libraries are written in C and compiled using the GNU C Compiler (GCC) on a MacBook Pro with an Intel Core i7 2.2 GHz CPU and 16GB RAM. We consider PBFT groups varying in size from 52 to 502. Recall from Section 3.1 that in both Prepare and Commit phases each node needs to verify $2f + 1$ signatures received from its peers, where $f$ is number of Byzantine nodes in PBFT. For the ECDSA signature scheme, we implement the batch verification approach [10] for verifying $2f + 1$ modified ECDSA signatures. On the other hand, for the BLS threshold signature scheme, we set the threshold as $2f$ for different PBFT group sizes. The performance of verifying multiple signatures using the short-lived signature schemes is illustrated in Figure 3, where $N$ is the PBFT group size and $S$ is the number of signatures that need to be verified for each node.
Scalable PBFT with Short-Lived Signature Schemes

6 RELATED WORK

ByzCoin [23] reduced the communication overhead of PBFT from \( O(n^2) \) to \( O(n) \) by replacing the MAC based all-to-all communication with a primitive called scalable collective signing (CoSi). Cosi is composed of four rounds of communication and a collective signature, which is structured as a tree of Schnorr signature, is generated and verified by all members of the group at the end. As a result, each node does not need to collect individual signatures form its peers any more.

In [24], Li et al. proposed Gosig, a new BFT protocol that combines crypto-based leader selection and multi-round voting schemes to improve the scalability under a permissioned blockchain setting. In particular, the author adopted an (insecure) BLS multi-signature scheme to reduce the storage and communication overhead and the PKI to rule out the potential attacks for the aggregated signatures.

In [18], Gueta et al. presented SBFT, a scalable decentralized trust infrastructure for blockchains. SBFT implemented a new BFT algorithm for addressing the scalability and decentralization issues. The BLS threshold signature scheme at the higher security level (i.e., 128-bit) has been extensively used in the SBFT design to reduce the communication overhead through signature aggregation.

7 CONCLUSIONS

In this paper we propose an efficient variant for the classical PBFT consensus algorithm, which takes advantage of the short-lived signature schemes to accelerate the signature verification process across the different phases of PBFT. When combining with novel blockchain-aided key distribution schemes, the scalability of the PBFT can be improved significantly. The extensive experiments with three elliptic curves and two signature schemes further highlight the advantages of the proposed approach for reducing the communicational and computation overhead of PBFT.

A PARAMETERS FOR ELLIPTIC CURVES

A.1 Elliptic Curve secp112r1

The elliptic curve secp112r1 [39] is given by the equation \( E(F_p) : y^2 = x^3 - 3x + b \) over a prime field \( F_p \), where

\[
\begin{align*}
    p &= db7c 2abf62e3 5e668076 bea208b \\
    b &= 659e f8b0439 16edef89 11702b22
\end{align*}
\]

The base point \( G = (x, y) \) is:

\[
\begin{align*}
    x &= 0948 7239995a 5ee7b5f 59f20f98 \\
    y &= a89c e5af8724 c0a20e0 08777500
\end{align*}
\]

The order \( n \) of \( G \) and the cofactor \( h \) are:

\[
\begin{align*}
    n &= db7c 2abf62e3 5e7628df ac6561c5 \\
    h &= 01
\end{align*}
\]

A.2 Elliptic Curve sect163k1

The Koblitz curve sect163k1 [39] is given by the equation \( E(F_{2^{163}}) : y^2 + xy = x^3 + x^2 + 1 \) over a binary field \( F_{2^{163}} \), where the binary field is defined by the irreducible polynomial \( f(x) = x^{163} + x^4 + x^3 + 1 \).

The base point \( G = (x, y) \) is:

\[
\begin{align*}
    x &= 02 fe13c053 7bb11ac aa87d793 de4e6d5e 5c94eef8 \\
    y &= 02 89070eb0 5038feff 321f2e80 0536d538 cccada3d9
\end{align*}
\]

The order \( n \) of \( G \) and the cofactor \( h \) are:

\[
\begin{align*}
    n &= 04 00000000 00000000 00020108 a2e0cc0d 99f8a5ef \\
    h &= 02
\end{align*}
\]
A.3 MNT Curve with Embedding Degree $k = 6$

The MNT curve with embedding degree $k = 6$ [31] is given by the equation $E(F_p) : y^2 = x^3 - 3x + b$ over a prime field $F_p$, where

$$p = 7dca613\ a2e3dbd1\ 749d195bf914cf\ 4462639\ b = 213cf3ac\ 786d1f1\ f9273dfb\ 823d6e5\ 7dc8fd4a$$

The order $n$ of $G$ and the cofactor $h$ are:

$$n = 33ee5389\ d17eed8\ ba4e12de\ f44414fd\ 17d369b7\ h = 62$$

The parameters $p$ and $n$ can be represented in terms of an 80-bit integer $z = dbd7d316\ aed514bb\ 8f95$ (in hexadecimal) such that

$$n = (z^2 + z + 1)/3$$

and

$$p = 2n + 1$$

References


Xinxin Fan