A New Prime-Norm Ideal Sampling for SQIsign

Li-Jie Jian

Academia Sinica, Taipei, Taiwan
jcuyo613@gmail.com

Abstract. SQIsign samples a random ideal of odd prime norm at every key generation and every signature through **PrimeNormIdealInertSampling**. We revisit this step from a new perspective: the inert ideals of a given norm form a circle, and drawing lines through a fixed point of this circle turns a single random parameter into an ideal, hitting each ideal once.
Compared with the reference implementation [AAA+26], the new algorithm requires one modular inversion and a handful of additions and multiplications. It achieves 13.76× speed-up at NIST security level I in SQIsign while producing an uniform output.

Keywords: SQIsign, Isogeny-based cryptography

1 Introduction

SQIsign uses **PrimeNormIdealInertSampling** to randomly generate an inert ideal I, which is a non-principal ideal of norm N with small bias. Algorithm 2 and Algorithm 1 comes from [AAA+26] as the reference.

Algorithm 1 is the only source of randomness for the ideals SQIsign works with. Specifically, key generation calls it to obtain the secret ideal I_{sk} (secret key), and every signature calls the same function again to generate the commitment ideal I_{com}. In both cases, the first line of that routine is a call to Algorithm 2 with $N = D_{\text{mix}} \approx p \cdot 2^{2\lambda}$. Therefore, its cost is also paid once per key generation and once per signature.

Algorithm 2 is intuitive and simple, but computationally expensive. It requires some quadratic residuosity tests, a modular square root and a CRT reconstruction. Since the test succeeds with the probability about $\frac{1}{2}$, the loop is executed twice on average. Moreover, the residuosity test and the square root are both exponentiations modulo N, so their algorithms need many multiplications. It seems worthwhile to find whether the same output distribution can be obtained at a lower cost.

Algorithm 1 Uniform$\mathcal{O}_0$ClassSampling

Input: M : Int, prng : PRNG
Output: I : InertIdeal, where the class of I is sampled with small bias among the $\mathcal{O}_0$-ideal classes different from the class of $\mathcal{O}_0$, and IdealNorm(I) is coprime to M if $M \neq 0$ and prime otherwise

1: $I \leftarrow \mathsf{PrimeNormIdealInertSampling}(D_{\text{mix}}, \text{prng})$
2: $I, _ \leftarrow \mathsf{EquivalentCoprimeIdeal}(I, M, \text{BitLenDegMix}, \text{prng})$
3: **return** I

Algorithm 2 PrimeNormIdealInertSampling

Input: N : positive odd prime
prng : a stateful pseudorandom number generator
Output: I : InertIdeal, a non-principal $\mathcal{O}_0$-ideal of norm N sampled with small bias

while `true` **do**
  $x \leftarrow \mathsf{SampleFromInterval}(0,\ N-1,\ \text{prng})$
  if $-4x^2 - p$ is a square mod $4N$ **then**
    $t \leftarrow -4x^2 - p \bmod 4N$
    $r \leftarrow \mathsf{ModularSQRT}(t \bmod N,\ N)$
    $r_4 \leftarrow t \bmod 2$
    $Y \leftarrow \mathsf{CRT}(r_4,\ 4,\ r,\ N)$
    $b \leftarrow \mathsf{SampleFromInterval}(0,\ 1,\ \text{prng})$
    $Y \leftarrow (-1)^b Y \bmod 2N$
    $y \leftarrow (Y-1)/2$
    return $(N, (x, y))$
  end if
end while

1.1 Our contribution

This paper presents a new perspective for sampling a random inert ideal and improves Algorithm 2. Specifically, we show that the solutions of $(2x)^2 + (2y+1)^2 \equiv -p \bmod 4N$ form a circle over $\mathbb{F}_N$, and that every point of this circle can be produced by a uniformly random number with only a single inversion, additions, and multiplications. The new algorithm therefore is simple and efficient. Moreover, we prove that the output distribution is uniform.

1.2 Technical overview

Set a prime $p \equiv 3 \bmod 4$. Let N be odd primes with $N > p$, and set $D \equiv -p \bmod N$. This means $D \neq 0$, so we apply to $\mathcal{C} = \{(x, y) \in \mathbb{F}_N^2 : x^2 + y^2 = D\}$, mapping a uniformly random number m to a uniformly random point of $\mathcal{C}$.

Let $X = 2x$ and $Y = 2y + 1$. Since $p \equiv 3 \bmod 4$ and $\gcd(4, N) = 1$, the condition of Algorithm 2 becomes

$$Y^2 \equiv -4x^2 - p \pmod{4N} \qquad \Longleftrightarrow \qquad X^2 + Y^2 \equiv -p \pmod{N},$$

which also been viewed as a circle equation $\mathcal{C}$:

$$X^2 + Y^2 = D, \text{ for } D \equiv -p \bmod N.$$

We state that sampling an inert ideal of norm N is same as sampling a point $P = (X, Y) \in \mathcal{C}$.

First, we fix a public point $P_0 = (X, Y) = (a_0, b_0) \in \mathcal{C}$ and construct a line with the slope m which is randomly generated. Since the line is determined by the random number, its second intersection point P_1 meeting with the circle is also random. This is to say, finding the second intersection point is equal to solving $P_1 = (a_0 + t, b_0 + mt), t \in \mathbb{F}_N$:

$$(a_0 + t)^2 + (mt + b_0)^2 = D, \tag{1}$$

$$a_0^2 + b_0^2 = D. \tag{2}$$

Next, we subtract Equation 2 from Equation 1 and get $t = -\frac{2(a_0 + mb_0)}{1+m^2}$ and $P_1 = (a_0 + t, b_0 + mt) = (X, Y) = (2x, 2y + 1)$. This process finally samples a random point $(x, y) = (\frac{(a_0+t)}{2}, \frac{(b_0-1)+mt}{2})$. Note that $(b_0 - 1)$ and $2^{-1} \bmod N$ are precomputed.

However, sampling a random number m as the slope from $\{0, \ldots, N-1\}$ does not give a uniform point. For example, it misses the vertical direction, $m = \infty$, which maps to $(a_0, -b_0)$. Moreover, sampling may cause the line meeting the circle at P_0 alone or may lead to $1 + m^2 = 0$, where t is undefined.

To resolve these issues, let $Z = \{m \in \mathbb{F}_N \mid 1 + m^2 = 0\}$, we define an extended domain $M = (\mathbb{F}_N \setminus Z) \cup \{\infty\}$, and our main theorem proves that the second-intersection map is a bijection from M to $\mathcal{C}$. Therefore, a uniform $m \in M$ gives an exactly uniform point of $\mathcal{C}$.

We discuss the implementation and its details in Section 4.

2 Preliminaries

Let N be an odd prime and a be an integer. We write $\left(\frac{\cdot}{N}\right)$ for the Legendre symbol, so that $\left(\frac{a}{N}\right) \in \{+1, 0, -1\}$.

Fix $D \in \mathbb{F}_N$ with $D \neq 0$ and let

$$\mathcal{C} \;=\; \{(x, y) \in \mathbb{F}_N^2 \mid x^2 + y^2 = D\}$$

be the solution set of the circle equation $x^2 + y^2 = D$. Let $P_0 = (a_0, b_0) \in \mathcal{C}$ be a fixed point.

For $m \in \mathbb{F}_N$, we define

$$d(m) \;=\; 1 + m^2 \quad (\text{the } \textit{denominator}), \qquad n(m) \;=\; a_0 + b_0 m \quad (\text{the } \textit{numerator}).$$

Definition 1 (Isotropic solutions)**.** Let $G(x, y) = x^2 + y^2$. The form G is *isotropic* over $\mathbb{F}_N$ if it has a nontrivial zero such that $G(x, y) = 0$ for some $(x, y) \neq (0, 0)$. We denote by

$$Z \;=\; \{\, m \in \mathbb{F}_N \mid d(m) = 1 + m^2 = 0 \,\},$$

the set of isotropic solutions of G.

Since $d(m) = 0$ if and only if $m^2 = -1$, the set Z consists of the square roots of -1 in $\mathbb{F}_N$. It therefore has two elements when -1 is a square modulo N, that is when $N \equiv 1 \pmod 4$, and is empty when $N \equiv 3 \pmod 4$. In both cases, we note its size is

$$|Z| \;=\; 1 + \left(\tfrac{-1}{N}\right).$$

In addition to the slopes $m \in \mathbb{F}_N$, we consider the vertical direction for $m = \infty$, which corresponds to the line $x = a_0$. These help construct a parameter set that exactly collects those directions for which the line passing through P_0 meets the circle at a second point. Specifically, isotropic directions must be discarded because they cause the line to intersect the circle redundantly at P_0; whereas the vertical line must be considered since the line $x = a_0$ always intersects the circle at a point $P_1 = (a_0, -b_0)$. Therefore, we define the set

$$M \;=\; \big(\mathbb{F}_N \setminus Z\big) \cup \{\infty\}$$

to serve as the domain of the map defined in Section 3. Its size is

$$|M| \;=\; \big(N - |Z|\big) + 1 \;=\; N - \left(\tfrac{-1}{N}\right).$$

3 Main Theorem

In this section, we construct a map F and prove that it is a bijection onto $\mathcal{C}$. The main idea is to fix a public point $P_0 \in \mathcal{C}$ and draw a line passing through it. The line meets a circle in two points, so once P_1 is discounted, the direction of the line determines a single further point P_1 that is in $\mathcal{C}$.

Since P_0 already lies on $\mathcal{C}$, the quadratic equation describing the intersection has $t = 0$ as a known root. The second intersection point is then the root of a *linear* equation, and costs one inversion instead of a square root. Lemma 1 shows this idea.

Lemma 1 (Intersection equation)**.** *Let $m \in \mathbb{F}_N$, construct a line passing through P_0 with the slope m, and describe it as the set $L = \{(a_0 + t,\ b_0 + mt) \mid t \in \mathbb{F}_N\}$. A point $P_1 \in L$ other than P_0 corresponds to a nonzero value of t. The condition that P_1 lies on the circle $x^2 + y^2 = D$ is a quadratic equation in t. One of its roots is $t = 0$, which gives back P_0, so the other root implies to solve a linear equation. Then the following hold.*

(i) $d(m) = (1 + m^2)$ *and* $n(m) = (a_0 + b_0 m)$ *never be zero simultaneously.*

(ii) $(a_0 + t)^2 + (b_0 + mt)^2 = D$ *if and only if* $t\big(d(m)\,t + 2n(m)\big) = 0$.

(iii) If $m \notin Z$, the roots are $t = 0$, which gives back P_0, and $t = -2n(m)/d(m)$; they coincide if and only if $n(m) = 0$.

(iv) If $m \in Z$, the equation reduces to $2n(m)\,t = 0$ with $n(m) \neq 0$, so $t = 0$ is its only root. A line whose slope is an isotropic solution therefore meets $\mathcal{C}$ in P_0 alone.

Proof. (i) Suppose $d(m) = n(m) = 0$. Then $m^2 = -1$ and $a_0 = -b_0 m$,

$$D = a_0^2 + b_0^2 = b_0^2 m^2 + b_0^2 = b_0^2(m^2 + 1) = 0,$$

which contradicts $D \neq 0$.

(ii) Due to $a_0^2 + b_0^2 = D$, we substract it from $(a_0 + t)^2 + (b_0 + mt)^2 = D$ and leaves $2a_0 t + t^2 + 2b_0 mt + m^2 t^2 = 0$, that is $t\big((1 + m^2)t + 2(a_0 + b_0 m)\big) = 0$. Then, replacing $(1 + m^2)$ and $(a_0 + b_0 m)$ with $d(m)$ and $n(m)$ gets $t\big(d(m)\,t + 2n(m)\big) = 0$.

(iii) and (iv) follow by solving the equation: $t\big(d(m)\,t + 2n(m)\big) = 0$. Recall from Section 2 that Z is the isotropic solution set of $G(x, y) = x^2 + y^2$ over $\mathbb{F}_N$. If $d(m) \neq 0$, that is $m \notin Z$, t has the root $-2n(m)/d(m)$ and zero. Additionally, if $d(m) = 0$, that is if $m \in Z$, then $n(m) \neq 0$ by (i), so $2n(m)$ is the nonzero constant and the equation reduces to $2n(m)\,t = 0$. Hence, $t = 0$ is its only root. □

For a well-defined direction, part (iii) of Lemma 1 is the source of the map, which leads to the second intersection point by an algorithm involving one inversion and a handful of additions and multiplications. Part (i) explains what makes the formula of part (iii) invertible. Moreover, part (iv) is the reason the isotropic solutions have to be discarded. We now define the map.

Definition 2 (Second intersection map)**.** For $m \in M$, and a fixed, public point $P_0 = (a_0, b_0) \in \mathcal{C}$ on the circle $x^2 + y^2 = D$, $t(m) = -\frac{2n(m)}{d(m)}$ is well-defined for $d(m) \neq 0$ and determines a point of $\mathcal{C}$ on the line of slope m. The **second intersection map** is

$$F : M \to \mathcal{C}, \qquad F(m) = \big(a_0 + t(m),\ b_0 + m \cdot t(m)\big), \qquad F(\infty) = (a_0,\ -b_0).$$

The definition covers three situations. First, when $t(m) \neq 0$, $F(m)$ is a point of $\mathcal{C}$ distinct from P_0. This represents a secant line with the slope m through the circle. Next, if $n(m) = 0$, this means $t(m) = 0$ and $F(m) = P_0$. The line is tangent at P_0, and it has been considered rather than an exception. Finally, the vertical line $x = a_0$ meets the circle in $(a_0, \pm b_0)$, so its second intersection point is $(a_0, -b_0)$.

Theorem 1. *Let N be an odd prime, $D \in \mathbb{F}_N^\times$, and $P_0 = (a_0, b_0) \in \mathcal{C}$. Let F be as in Definition 2 and let m be sampled uniformly at random from*

$$M = \big(\mathbb{F}_N \setminus Z\big) \cup \{\infty\}.$$

Then F is a bijection, so $F(m) = P_1 = (a_1, b_1)$ is uniformly distributed on $\mathcal{C}$, each point occurring with probability

$$\frac{1}{|\mathcal{C}|} = \frac{1}{N - \left(\frac{-1}{N}\right)}.$$

Proof. We first check that F takes its values in $\mathcal{C}$, that every point of $\mathcal{C}$ is reached by $m \in M$, and then that $F(m)$ is determined by only one m. All of them show that the map F is bijection. The size of $\mathcal{C}$ therefore is computed by the definition of bijection. Therefore we conclude $F(m) = P_1$ is uniformly distributed on $\mathcal{C}$.

Step 1: The image of F is well-defined. Let $m \in M$. If $m \neq \infty$ and $m \notin Z$, $t(m)$ is a root of the intersection equation with Lemma 1 and $F(m) \in \mathcal{C}$. If $m = \infty$, then $a_0^2 + (-b_0)^2 = a_0^2 + b_0^2 = D$, so $F(\infty) \in \mathcal{C}$. Thus $F : M \to \mathcal{C}$.

Step 2: F is surjective. Let $t = a_1 - a_0$ and $m = \frac{b_1 - b_0}{a_1 - a_0}$. They imply $P_1 = (a_0 + t,\ b_0 + mt) \in \mathcal{C}$. There are three different cases.

Case A: $a_1 \neq a_0$. Since $P_1 \in \mathcal{C}$ and $t \neq 0$, Lemma 1(ii) gives

$$d(m)\, t + 2n(m) = 0. \tag{3}$$

If $d(m) = 0$, then (3) gives $n(m) = 0$, contradicting Lemma 1(i). Hence $d(m) \neq 0$ and $m \notin Z$. Moreover, since $d(m) \neq 0$, (3) implies $t = -\frac{2n(m)}{d(m)} = t(m)$, and therefore $F(m) = (a_0 + t,\ b_0 + mt) = P_1$ for $m \in \mathbb{F}_N \setminus Z \subseteq M$.

Case B: $a_1 = a_0$ and $b_1 \neq b_0$. From $a_1^2 + b_1^2 = D = a_0^2 + b_0^2$ we get $b_1^2 = b_0^2$, so $b_1 \in \{b_0, -b_0\}$. As $b_1 \neq b_0$, we must have $b_1 = -b_0 \neq 0$. Hence $P_1 = (a_0, -b_0) = F(\infty)$ with $\infty \in M$.

Case C: $P_1 = P_0$. If $b_0 \neq 0$ and $m = -a_0/b_0$, m is well-defined. Then $n(m) = a_0 + b_0 m = 0$, so $d(m) \neq 0$ by Lemma 1(i), which implies $m \in M$. Furthermore, $t(m) = -\frac{2n(m)}{d(m)} = 0$ and hence $F(m) = P_0 = P_1$. Suppose now that $b_0 = 0$. It means $P_0 = (a_0, 0) = (a_0, -b_0) = F(\infty) = P_1$, so $\infty \in M$ is a preimage of $P_1 = P_0$.

In all cases, P_1 is one of the elements of the image of F, so $F(M) = \{F(m) \mid m \in M\} = \mathcal{C}$.

Step 3: F is injective. We show that the $m \in M$ in Step 2 is the only preimage. Fix $P_1 = (a_1, b_1) \in \mathcal{C}$ and suppose $F(m) = P_1$. There are also three different cases.

Case A: $a_1 \neq a_0$. The point $F(\infty) = (a_0, -b_0)$ has first coordinate a_0, so it forces $m \neq \infty$. We now get $m \in \mathbb{F}_N \setminus Z$ and $F(m) = \big(a_0 + t(m),\ b_0 + m\, t(m)\big)$. Comparing this with $P_1 = (a_1, b_1)$, we obtain $t = a_1 - a_0$ and $m \cdot t = b_1 - b_0$. Since $t \neq 0$, $m = \frac{b_1 - b_0}{a_1 - a_0}$ is well-defined. Note that the right-hand side of the equation only depends on $P_0 = (a_0, b_0)$ and $P_1 = (a_1, b_1)$. Thus, m is uniquely determined by P_1.

Case B: $a_1 = a_0$ and $b_1 \neq b_0$. For $m \in \mathbb{F}_N \setminus Z$ and $F(m) = \big(a_0 + t,\ b_0 + mt\big)$, comparing first coordinates gives $t = a_1 - a_0 = 0$, and then the second coordinate is $b_1 = b_0 + m \cdot 0 = b_0$. This contradicts $b_1 \neq b_0$. Hence no finite slope maps to $P_1 = (a_1, b_1)$, and $m = \infty$ is its only preimage.

Case C: $P_1 = P_0$. If $P_1 = P_0$, we need $t = a_1 - a_0 = t(m) = 0$, equivalently $n(m) = a_0 + b_0 m = 0$. When $b_0 \neq 0$, this has the unique solution $m = -a_0/b_0$. When $b_0 = 0$, we have $n(m) = a_0 \neq 0$ for every m, so there is no finite preimage and $m = \infty$ is the unique one.

Therefore, F maps distinct elements to distinct elements.

Step 4: The size of $\mathcal{C}$. Based on Steps 2 and 3, $F : M \to \mathcal{C}$ is a bijection between two finite sets, so they have the same number of elements. Recall from Section 2 that Z is the set of square roots of -1 in $\mathbb{F}_N$, which shows $|Z| = 1 + \left(\frac{-1}{N}\right)$ and M contains the $N - |Z|$ finite slopes together with $\{\infty\}$. We define

$$|\mathcal{C}| \;=\; |M| \;=\; \big(N - |Z|\big) + 1 \;=\; N - \left(\frac{-1}{N}\right),$$

which by Euler's criterion is $N - 1$ when $N \equiv 1 \pmod 4$ and $N + 1$ when $N \equiv 3 \pmod 4$.

Step 5: Conclusion. Since F is a bijection, the preimage $F^{-1}(P_1)$ is a single element for every $P_1 \in \mathcal{C}$. Therefore, if m is uniform on M,

$$\Pr\left[F(m) = P_1\right] \;=\; \Pr\left[m \in F^{-1}(P_1)\right] \;=\; \frac{1}{|M|} \;=\; \frac{1}{|\mathcal{C}|} \;=\; \frac{1}{N - \left(\frac{-1}{N}\right)}.$$

That is, $F(m)$ is uniformly distributed on $\mathcal{C}$. □

4 Implementation

In this section, we focus the implementation for the odd prime $N \equiv 3 \bmod 4$. Since $N \equiv 3 \bmod 4$, -1 is not a square modulo N, leading to the isotropic solution set Z is empty. No slope is forbidden, and $1 + m^2$ is never zero. Thus, $M = (\mathbb{F}_N \setminus Z) \cup \{\infty\}$ has $N + 1$ elements. This aligns with Definition 2 and ensures that the vertical direction has been considered and $F(m)$ is always valid.

For implementation, the simple solution for the vertical direction is to let ∞ be one extra value. We sample m uniformly from $\{0, \ldots, N\}$, which has exactly $N + 1$ elements, and see $m = N$ as the vertical direction. Since we fix a public point $P_0 = (a_0, b_0)$, the image of $F(\infty) = (a_0, -b_0)$ is a precomputed constant without extra variable-time computing. With $A = a_0 \cdot 2^{-1}$, $B = (b_0 - 1) \cdot 2^{-1}$, and $B_\infty = -(b_0 + 1) \cdot 2^{-1}$, we just need to consider two cases:

$$\begin{gathered} d(m) = 1 + m^2, \quad n(m) = a_0 + b_0 m, \quad t(m) = -\frac{2n(m)}{d(m)}, \quad w(m) = -\frac{n(m)}{d(m)} \\ (x, y) = \left(\frac{(a_0 + t)}{2}, \frac{(b_0 - 1) + mt}{2}\right) = (A + w,\ B + mw), \quad (x, y)\big|_{m=N} = (A,\ B_\infty). \end{gathered} \tag{4}$$

Since the second-intersection map is a bijection from M to $\mathcal{C}$, the output is uniform on $\mathcal{C}$. There thus is no failure event to account for. Algorithm 3 presents the overall computation.

Algorithm 3 NewPrimeNormIdealInertSampling (N, P_0, prng)

Input: N : positive prime with $N \equiv 3 \bmod 4$
$P_0 = (a_0, b_0)$: a public point such that $a_0^2 + b_0^2 \equiv -p \mod N$
$A = a_0 \cdot 2^{-1}$, $B = (b_0 - 1) \cdot 2^{-1}$, $B_\infty = -(b_0 + 1) \cdot 2^{-1} \bmod N$: precomputed
prng : a stateful pseudorandom number generator
Output: I : an inert ideal with norm N

1: $m \leftarrow \mathsf{SampleFromInterval}(0, N, \text{prng})$
2: infty $\leftarrow \mathsf{Compare}(m, N)$ ▷ 1 for the vertical direction, 0 otherwise
3: inv $\leftarrow (1 + m^2)^{-1} \mod N$ ▷ $(1 + m^2) \neq 0$ because $N \equiv 3 \bmod 4$
4: $w \leftarrow -(a_0 + mb_0) \cdot \text{inv} \bmod N$
5: $x \leftarrow \mathsf{Select}(\text{infty},\ A + w,\ A)$
6: $y \leftarrow \mathsf{Select}(\text{infty},\ B + mw,\ B_\infty)$
7: **return** $I = (N, (x, y))$

Table 1: Performance of PrimeNormIdealInertSampling at NIST security level I on Cortex-A76.

Algorithm	Cycles	Improvement
Reference [AAA+26]	1,041,836	1×
This paper	75,738	13.76×

References

[AAA+26] Marius A. Aardal, Gora Adj, Diego F. Aranha, Andrea Basso, Giacomo Borin, Isaac Andrés Canales Martínez, Jorge Chávez-Saab, Maria Corte-Real Santos, Pierrick Dartois, Luca De Feo, Max Duparc, Thomas Espitau, Jonathan Komada Eriksen, Tako Boris Fouotsa, Décio Luiz Gazzoni Filho, Basil Hess, Riccardo Invernizzi, David Kohel, Antonin Leroux, Patrick Longa, Luciano Maino, Michael Meyer, Marzio Mula, Kohei Nakagawa, Hiroshi Onuki, Lorenz Panny, Sikhar Patranabis, Christophe Petit, Giacomo Pope, Krijn Reijnders, Damien Robert, Francisco Rodríguez-Henríquez, Sina Schaeffler, Frederik Vercauteren, Alexandre Wallet, Benjamin Wesolowski, and Wessel van Woerden. SQIsign. Technical report, National Institute of Standards and Technology, 2026.