An Efficient and Parallel R-LWE Cryptoprocessor


Published in:
*IEEE Transactions on Circuits and Systems II: Express Briefs*

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2020 IEEE. This work is made available online in accordance with the publisher’s policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person’s rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
An Efficient and Parallel R-LWE Cryptoprocessor

Yuqing Zhang, Chenghua Wang, Dur E Shahwar Kundi, Member, IEEE, Ayesha Khalid, Member, IEEE, Máire O’Neill, Senior Member, IEEE, and Weiqiang Liu, Senior Member, IEEE

Abstract—Lattice-based cryptography (LBC) is a promising and efficient public key cryptography scheme whose theoretical foundation usually lies in Learning with Error (LWE) problem and its variant such as Ring-LWE (R-LWE) is the most studied cryptosystem which allows for more efficient implementation while maintaining the hardness of an original problem. Polynomial multiplication is the bottleneck of R-LWE, that can either be done using Number Theoretic Transform (NTT) or schoolbook polynomial multiplication (SPM) algorithm. The use of SPM is wider and possible for all parameters of R-LWE schemes. This work proposes an efficient and parallel strategy for SPM in R-LWE; by successfully reducing its time complexity from \( n^2 \) to \( n^2/4 \) (making it 1.8× faster and 1.4× hardware efficient). Furthermore, by adjusting the bit width for the error terms, the polynomial multiplication and addition blocks are reused for both encryption and decryption modules resulting in 14% reduced area and 1.7× better throughput in comparison to state-of-art SPM based R-LWE designs.

Index Terms—Lattice-based cryptography (LBC), Schoolbook polynomial multiplication (SPM), Ring Learning with Errors (R-LWE)

I. INTRODUCTION

Due to the algorithm proposed by Shor [1], classic public key cryptography (PKC) algorithms such as elliptic curve cryptography (ECC) and RSA have been proved to be vulnerable to attacks performed by quantum computers. In the round 2 of National Institute of Standards and Technology (NIST) post-quantum cryptography (PQC) standardization process, lattice-based cryptosystems (LBC) are the most promising candidates and account for 53% of the total of public key encryption (PKE) schemes [2].

Polynomial multiplication is a core module in LBC and is the bottleneck of the cryptosystem, which can be typically implemented by schoolbook and NTT algorithm. To speed up the operation of polynomial multiplication, usually NTT algorithm is utilized. However, it can only be used with selected parameters of cryptography and needs complex operations, including pre-computation, array reordering and post-computation.

In fact, half of the round 2 LBC candidates such as LAC, FRODO-KEM, Round5, Saber, Threebears and NTRUPrime in the NIST PQC initiative [2] have not used NTT method. In spite of the time complexity of schoolbook algorithm being larger than that of NTT algorithm, its simple implementation and low resource consumption makes it attractive for resource-constraint applications. Various architectures for NTT and schoolbook algorithm are analyzed and implemented comprehensively in [3]. NTT based Compact R-LWE cryptoprocessor are proposed in [4], [5], whereas NTT based high-throughput R-LWE cryptoprocessor are proposed in [6], [7].

In schoolbook polynomial multiplication (SPM), very little research is focused on the trade-off between performance and resource consumption. A lightweight design for SPM was proposed in [8], but it does not optimize the iteration clock cycle, as a result the throughput is not very high. On the other hand, [9] proposed an optimized SPM that reduces half of the iterative clock cycle. However, since its Gaussian sampling is signed data, some modules in encryption and decryption cannot be shared.

In this paper, we proposed a highly efficient SPM based R-LWE architecture by exploiting the features of polynomial multiplication. In addition, the bit width of Gaussian sampled data is adjusted appropriately, so that the polynomial multiplication module can be reused. In this method, only one module of polynomial multiplication is needed for the full R-LWE cryptosystem, which ensures reduced consumption of resources.

The paper is organized as follows: Section II reviews the background for R-LWE public key scheme. Section III presents the design of schoolbook polynomial multiplier and describes the entire structure of R-LWE. Section IV presents the hardware implementation results of the proposed designs, and comparison with other works. Section V concludes this paper.

II. BACKGROUND

The hardware implementation of R-LWE mainly includes a discrete Gaussian sampler and a polynomial multiplier. The entire process of encryption and decryption is shown in Alg. 1. In the algorithm, all operations are based on polynomials on the ring \( \mathbb{Z}_q[x]/(x^n + 1) \), where \( n \) is a power-of-two integer representing lattice dimension while \( q \) is prime modular [10]. \( U \) is a uniform distribution, and \( D_\sigma \) is a Gaussian distribution with \( \mu = 0 \) and a variance of \( \sigma \). ‘\( \times \)’ and ‘\( \leftarrow \)’ represent polynomial multiplication and polynomial addition, respectively.

We use the Cumulative Distribution Table (CDT) based scheme for discrete Gaussian sampling. CDT sampling is easier to be implemented in hardware than other sampling methods, it promises low resource consumption and high-throughput [11]. In addition, CDT sampling has been proven to
Algorithm 1 The Encryption and Decryption of R-LWE

**Input:** Public key: $a \leftarrow U$ and $p$; Secret key: $r_2 \leftarrow D_{\sigma}$; Plaintext: $m \leftarrow \{0,1\}$.

**Output:** Ciphertext: $c_1, c_2$; Decrypted plaintext: $m'$.

1. Sampling to generate $e_1, e_2, e_3 \leftarrow D_{\sigma}$;
2. Encoding plaintext $\hat{m} = encode(m)$;
3. Ciphertext is calculated $c_1 = a \times e_1 + e_2$, $c_2 = p \times e_1 + e_3 + \hat{m}$;
4. Decrypting $c = c_1 \times r_2 + c_2$;
5. Decoding plaintext $m' = decode(c)$;
6. return $c_1, c_2, m'$.

be resistant to side-channel attack (SCA), scalable in nature [12], and adaptable to deter fault-attacks [13]. The product $c(x)$ of two integer polynomial $a(x)$ and $b(x)$ can be computed by SPM algorithm, as expressed by (1) [3]:

$$a(x) \cdot b(x) = \sum_{i+j=n} a_i b_j x_i \mod (x^n + 1)$$  \hfill (1)

The sign bit of the coefficient in polynomial $c(x)$ is determined by $(-1)^{(i+j)/n}$ as follows:

$$\text{sign bit} = \begin{cases} 0 & i + j < n \\ 1 & n \leq i + j \leq 2n - 2 \end{cases}$$  \hfill (2)

The method of shift-addition-multiplication-subtraction-addition (SAMS2) [14] is taken up to accelerate the implementation of modular reduction operation in hardware. A well-optimized structure for SPM is proposed in [9]. For discrete Gaussian noise distribution, sampling with signed numbers is proposed for the first time, which reduces resources consumption. The formula of signed modular multiplication is expressed by (3):

$$a \times (-e) \mod q = q - \left[ a \times e \mod q \right]$$  \hfill (3)

Where, $a$ represents an unsigned number, $e$ is a signed Gaussian noise and $q$ is the prime modulus. Thus a 13-bit sampled signed number can be written as a 6-bit signed number for $q = 7681$ and the original $13 \times 13$ bit modular multiplication becomes $13 \times 5$ bit modular multiplication and one subtraction. Furthermore, as a result of this reduction, a single digital signal processing slice (i.e. DSP48) in Xilinx 7 series FPGA is configured to accommodate 2 simultaneous multiplications [9], which in turn reduces the complexity of SPM by half (from $n^2$ to $n^2/2$).

III. PROPOSED OPTIMIZED STRUCTURE OF R-LWE

In this section, we present our efficient SPM design and later the hardware sharing structure for overall implementation of R-LWE hardware. In this work, a medium security level R-LWE parameter set $(n = 256, q = 7681, s = 11.31)$ is selected [15].

Algorithm 2 High Efficiency SPM for R-LWE via Splitting

**Input:** $a(x), b(x)$ (polynomial in $\mathbb{Z}_q[x]/(x^n + 1)$); $q$ prime modular.

**Output:** $c(x)$ (polynomial in $\mathbb{Z}_q[x]/(x^n + 1)$);

1: for $i = 0:4:n-4$ do
2: for $j = 0:n-1$ do
3: $\text{coeff}_0 \leftarrow a[j];$
4: $\text{coeff}_1 \leftarrow a[j+2]$ mod $n$;
5: $\text{coeff}_2 \leftarrow b[i-j]$ mod $n$;
6: $\text{coeff}_4 \leftarrow b[i-j+1]$ mod $n$;
7: $\text{me}_b \leftarrow \{\text{coeff}_0[4:0], 13'b0, \text{coeff}_1[4:0]\}$;
8: $\text{temp}_ab0 \leftarrow \text{coeff}_0 \times \text{me}_b$;
9: $\text{temp}_ab2 \leftarrow \text{coeff}_1 \times \text{me}_b$;
10: $\text{temp}_01 \leftarrow \text{temp}_ab1[17:0]$ mod $q$;
11: $\text{temp}_2 \leftarrow \text{temp}_ab1[35:18]$ mod $q$;
12: $\text{temp}_3 \leftarrow \text{temp}_ab2[17:0]$ mod $q$;
13: $\text{temp}_4 \leftarrow \text{temp}_ab2[35:18]$ mod $q$;
14: $\text{sign}1 \leftarrow (i < j) ? 1 : 0$;
15: $\text{sign}2 \leftarrow (i + 1 < j) ? 1 : 0$;
16: $\text{sign}3 \leftarrow (i + 2 < (j + 2)$ mod $n) ? 1 : 0$;
17: $\text{sign}4 \leftarrow (i + 3 < (j + 2)$ mod $n) ? 1 : 0$;
18: if $\text{sign}1 \oplus \text{coeff}_0[5] == 1$ then
19: $\text{temp}_1 \leftarrow q - \text{temp}_1$;
20: end if
21: if $\text{sign}2 \oplus \text{coeff}_1[5] == 1$ then
22: $\text{temp}_2 \leftarrow q - \text{temp}_2$;
23: end if
24: if $\text{sign}3 \oplus \text{coeff}_0[5] == 1$ then
25: $\text{temp}_3 \leftarrow q - \text{temp}_3$;
26: end if
27: if $\text{sign}4 \oplus \text{coeff}_1[5] == 1$ then
28: $\text{temp}_4 \leftarrow q - \text{temp}_4$;
29: end if
30: $c[i] \leftarrow \text{temp}_1 + c[i]$;
31: $c[i+1] \leftarrow \text{temp}_2 + c[i+1]$;
32: $c[i+2] \leftarrow \text{temp}_3 + c[i+2]$;
33: $c[i+3] \leftarrow \text{temp}_4 + c[i+3]$;
34: end for
35: end for
36: return $c(x)$.
polynomial multiplication with dimensions of \( n \) requires to compute \( n^2 \) times modular multiplications and \((n-1)^2\) times modular additions/subtractions. Our strategy is to improve the parallelism in the design of SPM to reduce its time complexity and thereby improving its performance. As shown in Alg. 1, two operations of polynomial multiplication are needed in encryption, while one in decryption.

Public key, secret key and errors require 2 Block RAMs (BRAMs) for storage. Based on the operational features of the SPM, a scheme of efficient memory accessing is proposed. For simplicity, toy setting with \( n = 4 \) is considered as an example to illustrate the polynomial multiplication. In order to intuitively display the operation of \( c(x) = a(x) \times b(x) \), it is represented in the form of a matrix as shown in Fig. 1. In the first clock cycle, the corresponding coefficients of \( a(x) \) and \( b(x) \) as highlighted by red box are accessed from the memory and after their retrieval, the strategy of full utilization of FPGA DSP Blocks in [9] is used to compute 4 partial products of \( c(x) \) at once. Then, in second clock cycle the next coefficients of \( a(x) \) and \( b(x) \) as highlighted by blue shade are accessed to generate the next 4 partial products of \( c(x) \), and so on. Hence, the iteration period of the polynomial multiplication is reduced from 16 to 4 for \( n = 4 \).

Based on the proposed memory access method in Fig. 1, a scheme of parallel SPM is proposed, which mainly includes 6 steps to determine \( a(x) \times b(x) \), detailed outline in Alg. 2:

1) Reading coefficients \( a[j], a[(j + 2) \mod n] \) from \( RAM_A \) that stores \( a(x) \) and coefficients \( b[(i - j) \mod n], b[(i - j + 1) \mod n] \) from \( RAM_B \) that stores \( b(x) \) in one clock cycle (line 3–6).
2) Merging data bit of \( b[(i - j) \mod n] \) and \( b[(i - j + 1) \mod n] \) into a single 23-bit number (line 7).
3) Performing a modular multiplication operation of 23-bit \( \times \) 13-bit (line 8–9).
4) Separating the 18-bit data after the 36-bit product (line 10–13).
5) Determining the sign of separated data to carry out in-place reduction as well as subtraction from modulus \( q \) based on (3) (line 14–28). The value of \( sign1 \) to \( sign4 \) represents the condition for in-place reduction (line 14–17) whereas MSB of \( coeff_{b0}[5] \& coeff_{b1}[5] \) provides the coefficients sign of polynomial \( b(x) \). Both are XORed, if equal to “1” then \( temp1 \ldots temp4 \) are to be subtracted from the modulus \( q \) (line 18–28).
6) Final accumulation of generated result (line 30–33).

B. Structure of Polynomial Multiplier (PM)

A high-performance and parallel structure of polynomial multiplier is proposed based on Alg. 2, is shown in Fig. 2. The polynomial multiplier has polynomials \( a(x) \) and \( b(x) \) as multiplicand and multiplier. Their values are accessed every clock cycle. After data bit of \( coeff_{b0} \) and \( coeff_{b1} \) are merged, operation of multiplication is performed. The proposed structure requires only 2 DSPs, while a single DSP block is supporting a 25 \( \times \) 18 bit multiplication in Xilinx Kintex-7 FPGA. Resultant product is split into two halves and reduced via Barret Reduction (BR) module. The data is split into a total of four numbers, \( temp1 \ldots temp4 \). The most significant bit (MSB) of both coefficients \( coeff_{b0} \) and \( coeff_{b1} \) are input to the control unit before the multiplication operation, and sign of the final reduced data is determined by control unit. Finally, the generated 4 data are used as inputs to 4 modules of accumulation (Acc.), respectively. After completing all iterations, the 4 coefficients of polynomial \( c(x) \) are output, which are \( coeff_{c0} \), \( coeff_{c1} \), \( coeff_{c2} \) and \( coeff_{c3} \).

C. Reuse of Polynomial Multiplication and Addition in both R-LWE Encryption/Decryption.

Further to enhance the design efficiency, we exploit a smart and novel resource-sharing possibility for a combined R-LWE

![Fig. 2. The proposed structure of the efficient polynomial multiplier (PM).](image-url)
cryptoprocessor design. It stems from a simple observation that encryption and decryption operations have the same operation rules, both of which are \( d(x) = a(x) \times b(x) + e(x) \), as can be seen in Alg. 1 (line 3-4). Hence decryption module can reuse the polynomial multiplication as well as the polynomial addition in encryption module. However, error items \( e_2 \) and \( e_3 \) are 6-bit signed samples from a narrow discrete Gaussian distribution, hence to enable the reuse of polynomial addition hardware, the bit width of \( e_2 \) and \( e_3 \) is changed from 6 bits to 13 bits as presented by Alg. 3. If \( e(x) \) is negative (i.e. \( e[5] = 1 \)), the data bits (last 5 bits) of this number should be subtracted from the modulus \( q \) (line 4). Otherwise, MSB bits of \( e(x) \) is padded by “0” to make it 13 bits (line 6).

**Algorithm 3 Adjusting Bit Width (ABW) of Error Terms**

- **Input**: \( e[N] \) (each element is a 6-bit signed integer)
- **Output**: \( f[N] \) (each element is a 13-bit unsigned integer)

```
1: for i = 0 : N - 1 do
2:    reg_temp ← e[i];
3:    if reg_temp.get_bit(0) == 1 then
4:        f[i] ← q - reg_temp;
5:    else
6:        f[i] ← reg_temp × (1 << 8) + reg_temp;
7:    end if
8: end for
9: return f[N].
```

The proposed polynomial multiplier and adder can be utilized by the R-LWE cryptosystem, including key generation, encryption and decryption. The block diagram of the combined hardware-sharing structure is shown in Fig. 3, comprising of the discrete Gaussian sampler, Encryption and Decryption modules. As can be seen from this proposed structure, the polynomial multiplier is utilized by all the three operation of key-generation, encryption and decryption. The execution sequence is controlled by the control unit. Implementation of the Adjusting Bit Width (ABW) unit follows Alg. 3.

This thrifty reuse of schoolbook polynomial multiplier and the polynomial adder (PA) results in the reduction of area consumption of hardware resources, which has not been undertaken in the reported implementations.

**IV. HARDWARE IMPLEMENTATION RESULTS**

The proposed structures are synthesized and implemented on a Kintex-7 (KC705) FPGA using Xilinx Vivado 2018.3. Table I lists the detailed resource consumption, performance results (frequency, execution time and throughput) and hardware efficiency of FPGA based designs in terms of Throughput per Slice (TPS) from the post place and route (post-PAR) results. TPS is a trade off between throughput and the utilized hardware resources to achieve that throughput, thus higher value of TPS represents a good balanced design. To calculate the TPS for each implementation, a Slice equivalency is taken in account for dedicated/embedded resources of FPGA such as BRAM and DSP [9]. For a fair comparison of our proposed R-LWE design, we selected SPM based hardware implementations only. To facilitate the comparison, first the encryption and decryption architectures of proposed R-LWE design incorporating efficient polynomial multiplier is provided separately and then in the combined one.

The R-LWE design proposed in [8] is a lightweight design, but takes more clock cycles hence results in very low throughput as well as TPS. The design is based on parameters set of (256, 4092/4093, 8.35) that not only provided low security but also the structure of modular multiplication consumed less area resources. Moreover it utilized Bernoulli approach for rejection sampling that is more vulnerable to SCA. The design presented in [16], is not only based on better parameters set (256, 7681, 11.31) to achieve medium security but also utilized the CDT Gaussian sampling which is more resistant to timing attacks. Whereas, an optimized R-LWE design in [9] made...
use of signed bit sampling and full utilization of FPGA DSP Blocks and achieved the highest throughput of 1,119.84 Kbps and 2,255.49 Kbps for encryption and decryption, respectively. According to the features of the SPM, we devised strategic memory accessing from the BRAM and used 2 DSPs with 100% utilization which enable us to perform 4 multiplication operations in one clock cycle. Hence, reduce the overall latency of the design by 4 and achieved highest throughput values. Compared with the encryption and decryption module numbers of [9], our proposed separate designs have $1.7 \times$ throughput and $1.4 \times$ better hardware efficiency (i.e. TPS) at a cost of just 25.35% increased hardware.

Further, adjusting the bit width for the error terms helps us to improve the area consumption by re-using the hardware of SPM and PA for both the encryption and decryption processes of R-LWE. Our proposed combine architecture resulted in reduced area consumption for the overall cryptosystem i.e. utilized 796 Equivalent Slices (479+2×102.4+2×56). In comparison to the combined area resources of both encryption and decryption designs in [9], we saved 14% area and still have better throughput figures i.e. $1.7 \times$ of the state-of-art fastest available SPM based R-LWE design [9].

V. CONCLUSION

This work proposes an efficient and parallel SPM structure for the most critical operation of polynomial multiplication in R-LWE. We exploit the operational features of polynomial multiplication to implement a parallel design, which reduced iteration clock cycle by factor of 4 and results in $1.8 \times$ speedup and $1.4 \times$ TPS. Furthermore, an idea of re-using the most critical parts of the R-LWE encryption and decryption hardware is presented in this work. The combined R-LWE encryption/decryption architecture have 14% reduced area with $1.7 \times$ throughput.

REFERENCES


