J Cryptol (2018) 31:1164–1182
https://doi.org/10.1007/s00145-018-9298-8

Journal of
CRYPTOLOGY

Fast Multi-precision Multiplication for Public-Key Cryptography on Embedded Microprocessors*

Michael Hutter
Rambus Cryptography Research Division, 425 Market Street, 11th Floor, San Francisco, CA 94105, USA
Institute for Applied Information Processing and Communications (IAIK), Graz University of Technology, Inffeldgasse 16a, 8010 Graz, Austria
michael.hutter@cryptography.com

Erich Wenger
Institute for Applied Information Processing and Communications (IAIK), Graz University of Technology, Inffeldgasse 16a, 8010 Graz, Austria
wenger.erich@gmail.com

Communicated by Mitsuru Matsui.

Received 3 October 2012 / Revised 21 September 2017
Online publication 28 June 2018

Abstract. Multi-precision multiplication is one of the most fundamental operations on microprocessors to allow public-key cryptography such as RSA and elliptic curve cryptography (ECC). In this paper, we present a novel multiplication technique that increases the performance of multiplication by sophisticated caching of operands. Our method significantly reduces the number of needed *load* instructions which is usually one of the most expensive operations on modern processors. We evaluate our new technique on an 8-bit ATmega128 and a 32-bit ARM7TDMI microcontroller and compare the results with existing solutions. For the ATmega128, our implementation needs only 2395 clock cycles for a 160-bit multiplication. The number of required *load* instructions is reduced from 167 (needed for the best known hybrid multiplication) to only 80. On the ARM7TDMI, our implementation needs only 281 clock cycles as opposed to 357. For both platforms, the proposed technique outperforms related work by a factor of about 10–23%. We also show that the method scales very well even for larger Integer sizes (required for RSA) and limited register sets. It fully complies with existing multiply–accumulate instructions that are integrated in most of the available processors.

Keywords. Multi-precision arithmetic, Microprocessors, Elliptic curve cryptography, RSA, Embedded devices.

1. Introduction

Multiplication is one of the most important arithmetic operations in public-key cryptography. It engrosses most of the resources and execution time of modern microprocessors

*This work was done while the authors were at Graz University of Technology (IAIK).

(up to 80% for elliptic curve cryptography (ECC) and RSA implementations [8]). In order to increase the performance of multiplication, most effort has been put by researchers and developers to reduce the number of instructions or minimize the amount of memory-access operations.

Common multiplication methods are the schoolbook or Comba [5] technique which are widely used in practice. They require at least $2n^2$ *load* instructions to process all operands and to calculate the necessary partial products. In 2004, Gura et al. [8] presented a new method that combines the advantages of these methods (hybrid multiplication). They reduced the number of *load* instructions to only $2\lceil n^2/d \rceil$ where the parameter d depends on the number of available registers of the underlying architecture. They reported a performance gain of about 25% compared to the classical Comba multiplication. Their 160-bit implementation needs 3106 clock cycles on an 8-bit ATmega128 microcontroller. Since then, several authors applied this method [10,15,21,22,24] and proposed various enhancements to further improve the performance. Most of the related work reported between 2593 and 2881 clock cycles on the same platform.

In this paper, we present a novel multiplication technique that reduces the number of needed *load* instructions to only $2n^2/e$ where $e > d$. We propose a new way to process the operands which allows efficiently caching of required operands. In order to evaluate the performance, we implemented a 160-bit multiplication on an ATmega128 and an ARM7TDMI microcontroller and compare the results with related work. For the ATmega128, only 2395 clock cycles are required which is an improvement by a factor of 10% compared to the best reported implementation of Scott et al. [21] (which need 2651 clock cycles) and by a factor of about 23% compared to the work of Gura et al. [8]. For the ARM7TDMI, only 281 clock cycles are needed which is a reduction by about 20% compared to the best reported solution in the literature so far. We further compare our solution with different Integer sizes (160, 192, 256, 512, 1024, and 2048) and a parameter e that depends on the number of available registers ($e = 2$, 4, 8, 10, and 20). It shows that our solution needs about 15% less clock cycles for any chosen Integer size. Our solution also scales very well for different register sizes without significant loss of performance. Besides this, the method fully complies with common architectures that support multiply–accumulate instructions using a (Comba-like) triple-register accumulator.

The paper is organized as follows. In Sect. 2, we describe related work on that topic and give performance numbers for different multiplication techniques. Section 3 describes different multi-precision multiplication techniques used in practice. We describe the operand-scanning, product-scanning, and the hybrid methods and compare them with our solution. In Sect. 4, we present the results of our evaluations. We describe the ATmega128 and ARM7TDMI architecture and give details about the implementation. Summary and conclusions are given in Sect. 5.

2. Related Work

In this section, we describe related work on multi-precision multiplication over prime fields. Most of the work given in the literature makes use of the hybrid multiplication technique [8] which provides best performance on most microprocessors. This technique

was first presented at CHES 2004 where the authors reported a speed improvement of up to 25% compared to the classical Comba multiplication technique [5] on 8-bit platforms. Their implementation requires 3106 clock cycles for a 160-bit multiplication on an ATmega128 [1]. Several authors adopted the idea and applied the method for different devices and environments, e.g., sensor nodes. Wang et al. [25] and Ugus et al. [23] made use of this technique and implemented it on the MICAz motes which featured an ATmega128 microcontroller. Results for the same platform have been also reported by Liu et al. [14] and Szczechowiak et al. [22] in 2008 who provide software libraries (TinyECC and NanoECC) for various sensor-mote platforms. One of the first who improved the implementation of Gura has been due to Uhsadel et al. [24]. They have been able to reduce the number of needed clock cycles to only 2881. Further improvements have been also reported by Scott et al. [21]. They introduced additional registers (so-called *carry catchers*) and could increase the performance to 2651 clock cycles. Note that they fully unrolled the execution sequence to avoid additional clock cycles for loop instructions. Similar results have been also obtained by Kargl et al. [10] in 2008 which reported 2593 clock cycles for an unrolled 160-bit multiplication on the ATmega128.

In 2009, Lederer et al. [12] showed that the needed number of addition and move instructions can be reduced by simply rearranging the instructions during execution of the hybrid multiplication method. Similar findings have been also reported recently by Liu et al. [15] who reported the fastest looped version of the hybrid multiplication needing 2865 clock cycles in total.

In view of the ARM7TDMI, most of the related work focused on ECC implementations and reported numbers mainly for scalar multiplication, e.g., the work of Aydos et al. [3], Medwed et al. [16], or Xu and Batina [26]. Pelzl et al. [19] presented a hyper-elliptic curve implementation on the ARM7TDMI in 2003. For a GF(2^{191}) field multiplication, they reported 50.7 µs for an implementation that runs at a clock frequency of 80 MHz. This corresponds to about 4056 clock cycles. A paper that discusses multi-precision multiplication on that platform has been presented by Scott et al. [21]. For a 192-bit multiplication, they need 580 clock cycles for the Comba and 487 clock cycles for the hybrid multiplication. They make use of a single carry-catcher register that efficiently handles carry propagation on the ARM7TDMI which allowed them to lower the execution time compared to related work.

3. Multi-precision Multiplication Techniques

In the following subsections, we describe common multiplication techniques that are often used in practice. We describe the operand scanning, product scanning, and hybrid multiplication method.[1] The methods differ in several ways how to process the operands and how many *load* and *store* instructions are necessary to perform the calculation. Most of these methods lack in the fact that they load the same operands not only once but

[1] Note that we do not consider multiplication methods such as Karatsuba–Ofman or FFT in this paper since they are considered to require more resources and memory accesses on common microcontrollers than the given methods [11].

several times throughout the algorithm which results in additional and unnecessary clock cycles. We present a new multiplication technique that improves existing solutions by efficiently reducing the *load* instructions through sophisticated caching of operands.

Throughout the paper, we use the following notation. Let a and b be two m-bit large Integers that can be written as multiple-word array structures $A = (A[n-1], \ldots, A[2], A[1], A[0])$ and $B = (B[n-1], \ldots, B[2], B[1], B[0])$. Further let W be the word size of the processor (e.g., 8, 16, 32, or 64 bits) and $n = \lceil m/W \rceil$ the number of needed words to represent the Integers a or b. We denote the result of the multiplication by $c = ab$ and represent it in a double-size word array $C = (C[2n-1], \ldots, C[2], C[1], C[0])$.

3.1. *Operand-Scanning Method*

The simplest way to perform large Integer multiplication is the operand-scanning method (or often referred as *schoolbook* or *row-wise* multiplication method). The multiplication can be implemented using two nested loop operations. The outer loop loads the operand $A[i]$ at index $i = 0 \ldots n-1$ and keeps the value constant inside the inner loop of the algorithm. Within the inner loop, the multiplicand $B[j]$ is loaded word by word and multiplied with the operand $A[i]$. The partial product is then added to the intermediate result of the same column which is usually buffered in a register or stored in data memory. Figure 1 shows the structure of the algorithm on the left side. The individual row levels can be clearly discerned. On the right side of the figure, all n^2 partial products are displayed in form of a rhombus. Each point in the rhombus represents a multiplication $A[i] \times B[j]$. The most right-sided corner of the rhombus starts with the lowest indices $i, j = 0$, and the most left-sided corner ends with the highest indices $i, j = n-1$. By following all multiplications from the right to the lower-mid corner of the rhombus, it can be observed that the operand $A[i]$ keeps constant for any index $i \in [0, n)$. The same holds true for the operand $B[j]$ and $j \in [0, n)$ by following all multiplications from right to the upper-mid corner of the rhombus. Note that this is also valid for the left-handed side of the rhombus.

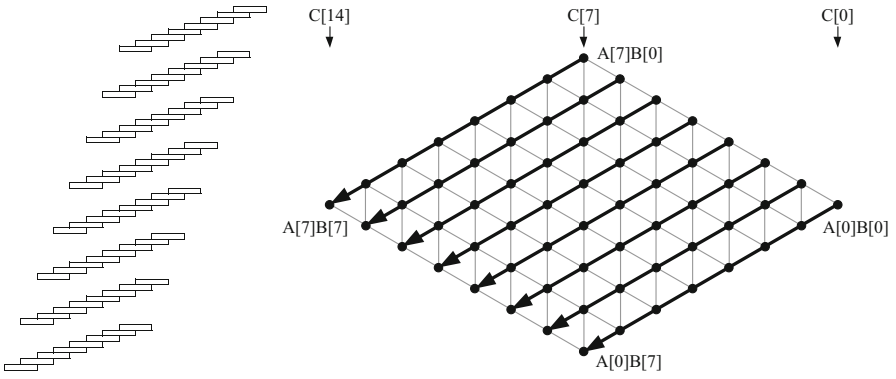

Fig. 1. Operand-scanning multiplication of 8-word large Integers a and b.

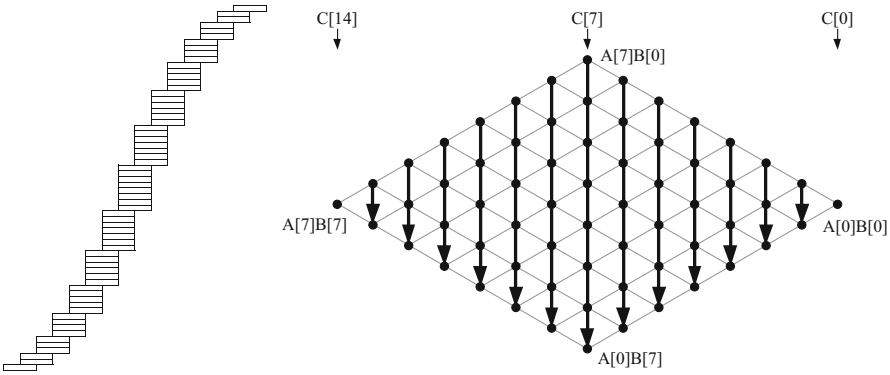

Fig. 2. Product-scanning multiplication of 8-word large Integers a and b.

For the operand-scanning method, it can be seen that the partial products are calculated from the upper-right side to the lower-left side of the rhombus (we marked the processing of the partial products with a black arrow). In each row, n multiplications have to be performed. Furthermore, $2n$ *load* operations and n *store* operations are required to load the multiplicand and the intermediate result $C[i+j]$ and to store the result $C[i+j] \leftarrow C[i+j] + A[i] \times B[j]$. Thus, $3n^2 + 2n$ memory operations are necessary for the entire multi-precision multiplication. Note that this number decreases to $n^2 + 3n$ for architectures that can maintain the intermediate result in available working registers.

3.2. *Product-Scanning Method*

Another way to perform a multi-precision multiplication is the product-scanning method (also referred as *Comba* [5] or *column-wise* multiplication method). There, each partial product is processed in a column-wise approach. This has several advantages. First, since all operands of each column are multiplied and added consecutively (within a multiply–accumulate approach), a final word of the result is obtained for each column. Thus, no intermediate results have to be stored or loaded throughout the algorithm. In addition, the handling of carry propagation is very easy because the carry can be simply added to the result of the next column using a simple register-copy operation. Second, only five working registers are needed to perform the multiplication: two registers for the operand and multiplicand and three registers for accumulation.[2] This makes the method very suitable for low-resource devices with limited registers.

Figure 2 shows the structure of the product-scanning method. By having a look at the rhombus, it shows that by processing the partial products in a column-wise instead of a row-wise approach, only one *store* operation is needed to store the final word of the result. For the entire multi-precision operation, $2n^2$ *load* operations are necessary to load the operands $A[i]$ and $B[j]$ and $2n$ *store* operations are needed to store the result. Therefore, $2n^2 + 2n$ memory operations are needed.

[2] We assume the allocation of three registers for the accumulator register, whereas $2 + \lceil log_2(n)/W \rceil$ registers are actually needed to maintain the sum of partial products.

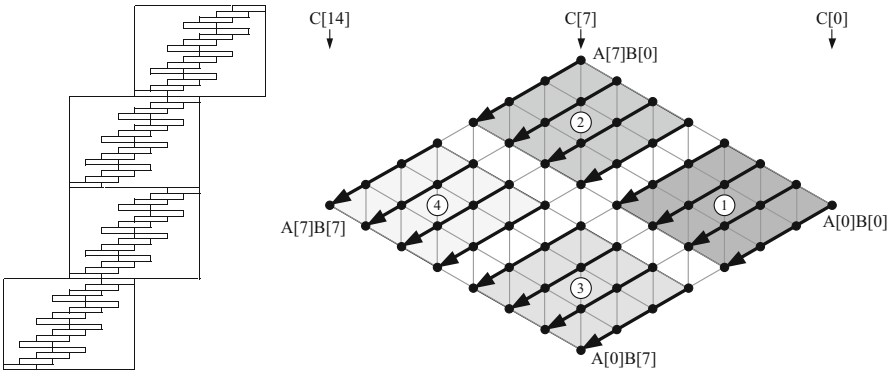

Fig. 3. Hybrid multiplication of 8-word large Integers a and b ($d = 4$).

3.3. *Hybrid Method*

The hybrid multiplication method [8] combines the advantages of the operand-scanning and product-scanning method. It can be implemented using two nested loop structures where the outer loop follows a product-scanning approach and the inner loop performs a multiplication according to the operand-scanning method.

The main idea is to minimize the number of *load* instructions within the inner loop. For this, the accumulator has to be increased to a size of $2d+1$ registers. The parameter d defines the number of rows within a processed block. Note that the hybrid multiplication is equal to the product-scanning method if parameter d is chosen as $d = 1$ and it is equal to the operand-scanning method if $d = n$.

Figure 3 shows the structure of the hybrid multiplication for $d = 4$. It shows that the partial products are processed in the form of individual blocks. (We marked the processing sequence of the blocks from 1 to 4.) Within one block, all operands are processed row by row according to the operand-scanning approach. Note that these blocks use operands with a very limited range of indices. Thus, several *load* instructions can be saved in cases where enough working registers are available. However, the outer loop of the hybrid method processes the blocks in a column-wise approach. So between two consecutive blocks no operands can be shared and all operands have to be loaded from memory again. This becomes clear by having a look at the processing of Block 1–3. Block 2 and 3 do not share any operands that possess the same indices. Therefore, all operands that have already been loaded for Block 1 and that can be reused in Block 3 have to be loaded again after processing of Block 2 which requires additional and unnecessary *load* instructions. However, in total, the hybrid method needs $2\lceil n^2/d \rceil + 2n$ memory-access instructions which provides good performances on devices that have many registers.

3.4. *Operand-Caching Method*

We present a new method to perform multi-precision multiplication. The main idea is to reduce the number of memory accesses to a minimum by efficiently caching of operands. We show that by spending a certain amount of *store* operations, a significant amount

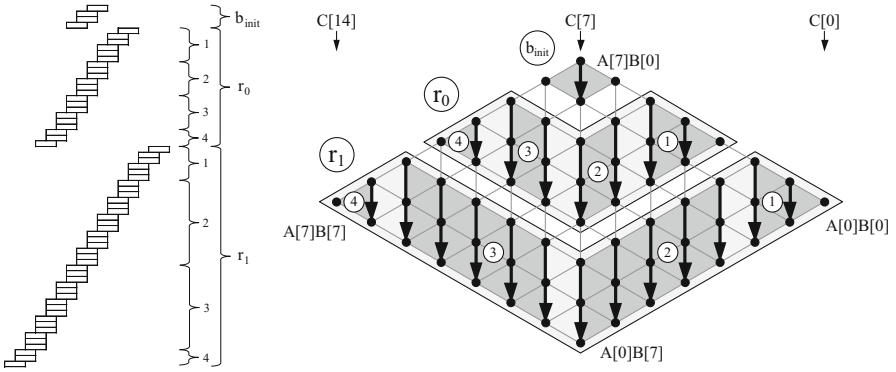

Fig. 4. Operand-caching multiplication of 8-word large Integers a and b ($e = 3$).

of *load* instructions can be saved by reusing operands that have been already loaded in working registers.

The method basically follows the product-scanning approach but divides the calculation into several rows. In fact, the product-scanning method provides best performance if all needed operands can be maintained in working registers. In such a case, only $2n$ *load* instructions and $2n$ *store* instructions would be necessary. However, the product-scanning method becomes inefficient if not enough registers are available or if the Integer size is too large to cache a significant amount of operands. Hence, several *load* instructions are necessary to reload and overwrite the operands in registers.

In the light of this fact, we propose to separate the product-scanning method into individual rows $r = \lfloor n/e \rfloor$. The size e of each row is chosen in a way that all needed words of one operand can be cached in the available working registers. Figure 4 shows the structure of the proposed method for parameter $e = 3$. That means, 3 registers are reserved to store 3 words of operand a and 3 registers are reserved to store 3 words of operand b. Thus, we assume $f = 2e + 3 = 9$ available registers including a triple-word accumulator. The calculation is now separated into $r = \lfloor 8/3 \rfloor = 2$ rows, i.e., r_0 and r_1, and consists of one remaining block which we further denote as initialization block b_{init}. This block calculates the partial products which are not processed by the rows.

All rows are further separated into four parts. Parts 1 and 4 use the classical product-scanning approach. Parts 2 and 3 perform an efficient multiply–accumulate operation of already cached operands.

The algorithm starts with the calculation of b_{init} and processes the individual rows afterward (starting from the smallest to the largest row, i.e., from the top to the bottom of the rhombus). Furthermore, all partial products are generated from right to left. In the following, we describe the algorithm in more detail. A pseudocode algorithm of the method is given in "Appendix A."

Initialization Block b_{init}. This block (located in the upper-mid of the rhombus) performs the multiplication according to the classical product-scanning method. The Integer size of the b_{init} multiplication is $(n - re)$, i.e., $8 - 6 = 2$ in our example, which is by definition smaller than e. Because of that, all operands can be loaded and maintained within the

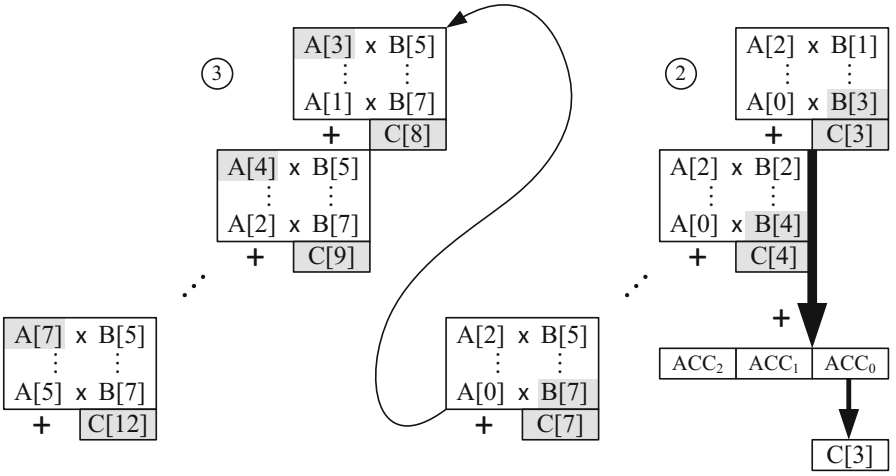

Fig. 5. Processing of Parts 2 and 3 of the row r_1.

available registers resulting in only $4(n - re)$ memory-access operations. Note that the calculation of b_{init} is only required if there exist remaining partial products, i.e., n mod $e \neq 0$. If n mod $e = 0$, the calculation of b_{init} is skipped. Furthermore, consider the special case when $n < e$ where only b_{init} has to be performed skipping the processing of rows (trivial case).

Processing of Rows. In the following, we describe the processing of each row $p = r - 1 \ldots 0$. Each row consists of four parts.

Part 1. This part starts with a product-scanning multiplication. All operands for that row are first loaded into registers, i.e., $A[i]$ with $i = pe \ldots e(p + 1) - 1$ and $B[j]$ with $j = 0 \ldots e - 1$. The sum of all partial products $A[i] \times B[j]$ is then stored as intermediate result to the memory location $C[i]$ (same index range as $A[i]$). Therefore, $2e$ *load* instructions and e *store* instructions are needed.

Part 2. The second part processes $n - e(p + 1)$ columns using a multiply–accumulate approach. Since all operands of $A[i]$ were already loaded and used in Part 1, only one word $B[j]$ has to be loaded from one column to the next. The operands $A[i]$ are kept constant throughout the processing of Part 2. Next to the needed *load* instructions for $B[j]$, we have to load and update the intermediate result of Part 1 with the result obtained in Part 2. Thus, $2(n - e(p + 1))$ *load* and $n - e(p + 1)$ *store* instructions are required for that part.

Part 3. The third part performs the same operation as described in Part 2 except that the already loaded operands $B[j]$ are kept constant and that one word $A[i]$ is loaded for each column. Figure 5 shows the processing of Parts 2 and 3 of row r_1 ($p = 0$). For each column, two *load* instructions are necessary (marked in gray). All other operands have been loaded and cached in previous parts. Operands which are not required for further processing are overwritten by new operands, e.g., $B[1] \ldots B[4]$ in Part 2 of our example.

Table 1. Memory-access complexity of b_{init} and each part of row $p = 0 \ldots r-1$.

Component	Load instr.	Store instr.	Total
b_{init}	$2(n-re)$	$2(n-re)$	$4(n-re)$
Part 1	$2e$	e	$3e$
Part 2	$2(n-e(p+1))$	$n-e(p+1)$	$3(n-e(p+1))$
Part 3	$2(n-e(p+1))$	$n-e(p+1)$	$3(n-e(p+1))$
Part 4	0	e	e

Table 2. Memory-access complexity of different multiplication techniques.

Method	Load instructions	Store instructions	Memory instructions
Operand scanning	$2n^2+n$	n^2+n	$3n^2+2n$
Product scanning [5]	$2n^2$	$2n$	$2n^2+2n$
Hybrid [8]	$2\lceil n^2/d \rceil$	$2n$	$2\lceil n^2/d \rceil + 2n$
Operand caching	$2n^2/e$	$n^2/e+n$	$3n^2/e+n$

Part 4. The last part calculates the remaining partial products. In contrast to Part 1, no *load* instructions are required since all operands have been already loaded in Part 3. Hence, only e memory-access operations are needed to store the remaining words of the (intermediate) result c.

Table 1 summarizes the memory-access complexity of the initialization block and the individual parts of a row p. By summing up all *load* instructions, we get

$$2(n-re) + \sum_{p=0}^{r-1}(4n-4pe-2e) = 2n + 4rn - 2er^2 - 2er \le \frac{2n^2}{e}. \tag{1}$$

The total number of *store* operations can be evaluated by

$$2(n-re) + \sum_{p=0}^{r-1}(2n-2pe) = 2n + 2rn - er^2 - er \le \frac{n^2}{e} + n. \tag{2}$$

Table 2 lists the complexity of different multi-precision multiplication techniques. It shows that the hybrid method needs $2\lceil \frac{n^2}{d} \rceil$ *load* instructions, whereas the operand-caching technique needs about $\frac{2n^2}{e}$. Since the total number of available registers f equals to $2e+3$ for the operand-caching technique ($2e$ registers for the operand registers and three registers for the accumulator) and $3d+2$ for the hybrid method ($d+1$ registers for the operands and $2d+1$ registers for the accumulator), we obtain

$$f = 2e+3 = 3d+2 \implies e = \frac{3d-1}{2} \quad \text{and} \quad e > d. \tag{3}$$

If we compare the total number of memory-access instructions for the hybrid and the operand-caching method and express both runtimes using f, we get

$$2\left\lceil \frac{n^2}{\left\lfloor \frac{f-2}{3} \right\rfloor} \right\rceil + 2n > \frac{6n^2}{f-3} + n \tag{4}$$

Note that there are more parameters to consider. The number of additions of the operand-caching method is $3n^2$, and the number of additions of the hybrid method is $n^2(2 + d/2)$ (upper bound). Also the pseudocode of Gura et al. [8] for the hybrid multiplication method is inefficient in the special case of $n \bmod d \neq 0$.

4. Results

In order to demonstrate the performance of our method, we implemented all multiplication techniques described in Sect. 3 on two different platforms. The first platform is an 8-bit ATmega128 microcontroller, and the second platform is a 32-bit ARM7TDMI microcontroller. In order to facilitate the evaluation, we implemented a code generator (based on Java language) that generates the Assembler source code for all multiplication methods. The code generator allows flexible adjustment of individual settings such as operand sizes, available registers, and used compiler (supported compilers are the avrgcc [18], Crossworks for AVR [20], and the IAR compiler [9]). Furthermore, it can be adjusted if the multiplication should be implemented using a loop (to limit code size) or if the instructions should be unrolled in Assembler (to increase speed). It also allows to generate code for devices without a dedicated hardware multiplier which is the case for the ATtiny family of microcontrollers from Atmel, for instance.[3]

4.1. *Performance Results for the 8-bit ATmega128*

The ATmega128 is part of the megaAVR family from Atmel [1]. It has been widely used in embedded systems, automotive environments, and sensor-node applications. It is based on a RISC architecture and provides 133 instructions [2]. The maximum operating frequency is 16 MHz. The device features 128 kB of flash memory and 4 kB of internal SRAM. There exist 32 8-bit general-purpose registers (R0 to R31). Three 16-bit register pairs can be used for memory addressing, i.e., R26:R27, R28:R29, and R30:R31 which are often denoted as X, Y, and Z. Note that the processor also allows pre-decrement and post-increment functionalities that can be used for efficient addressing of operands. The ATmega128 further provides a hardware multiplier that performs an 8×8-bit multiplication within two clock cycles. The 16-bit result is stored in the registers R0 (lower word) and R1 (higher word).

For the ATmega128, we used register R22 to store a zero value. Furthermore, we reserved R23, R24, and R25 as accumulator registers. Thus, 20 registers, i.e., R2…R21, can be used to store and cache the words of the operands ($e = 10$ registers for each operand a and b).

[3] The code generator is available from http://www.iaik.tugraz.at/content/research/sesys/tools/mulopcache/.

Table 3. Unrolled instruction counts for a 160-bit multiplication on the ATmega128.

Method	Instruction						Clock cycles
	LD	ST	MUL	ADD	MOVW	Others	
Operand scanning	820	440	400	1600	2	464	5427
Product scanning	800	40	400	1200	2	159	3957
Hybrid ($d = 4$)	200	40	400	1250	202	109	2904
Operand caching ($e = 10$)	80	60	400	1240	2	68	2395

As a first comparison, we decided to focus on a 160×160-bit multiplication as it has been done by most of the related work. Note that for RSA and ECC, larger Integer sizes are recommended in practice [13,17]. The Standards for Efficient Cryptography (SEC) already removed the recommended secp160r1 elliptic curve from their standard since SEC version 2 of 2010 [4].

Table 3 summarizes the instruction counts for the operand scanning, product scanning, hybrid, and operand-caching implementation. The operand-scanning and product-scanning methods have been implemented without using all the available registers (as it usually would be implemented). For hybrid multiplication, we applied $d = 4$ because it allows a better optimization regarding necessary addition operations compared to a multiplication with $d = 5$. The carry propagation problem has been solved by implementing a similar approach as proposed by Liu et al. [15]. Thus, 200 `MOVW` instructions have been necessary to handle the carry propagation accordingly. For a fair comparison, all methods have been optimized for speed and provide unrolled instruction sequences. Furthermore, we implemented all accumulators as ring buffers to reduce necessary `MOV` instructions. After each partial-product generation, the indices of the accumulator registers are shifted so that no `MOV` instructions are necessary to copy the carry.

Best results have been obtained for the operand-caching technique. By trading additional 20 *store* instructions, up to 120 *load* instructions could be saved when we compare the result with the best reference values (hybrid implementation). Note that *load*, *store*, and *multiply* instructions on the ATmega128 are more expensive than other instructions since they require two clock cycles instead of only one. For operand-caching multiplication, almost the same amount of *load* and *store* instructions are required. In total 2395 clock cycles are needed to perform the multiplication. Compared to the hybrid implementation, a speed improvement of about 18% could be achieved.

We also compare the performance of the implemented multi-precision methods for different Integer sizes. Table 4 shows the result for Integer sizes from 160 up to 2048 bits.[4] The operand-caching technique provides the best performance for any Integer size. It is therefore also well suited for large Integer sizes such as it is in the case of RSA. On average, a speed improvement of about 15% could be achieved compared to the hybrid method. Figure 6 shows the appropriate performance chart in a double logarithmic scale.

Table 5 and Fig. 7 show the performance for different Integer sizes in relation to parameter e. The parameter e is defined by the number of available registers to store

[4] Note that a fully unrolled implementation using such large Integer multiplications might be impractical due to the huge amount of code.

Table 4. Comparison of multiplication methods for different Integer sizes.

Size (bit)	Op. scan.	Prod. scan.	Hybrid method	Operand caching
160	5427	3957	2904	2395
192	7759	5613	4144	3469
256	13,671	9789	7284	6123
512	53,959	38,013	28,644	24,317
1024	214,407	149,757	113,604	96,933
2048	854,791	594,429	452,484	387,195

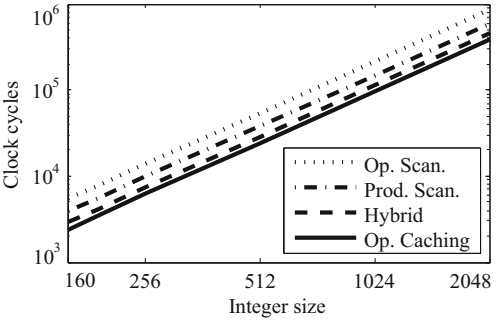

Fig. 6. Comparison chart.

Table 5. Performance of operand-caching multiplication for different Integer sizes and available registers.

Size	$e = 2$	$e = 4$	$e = 8$	$e = 10$	$e = 20$
160	3915	2965	2513	2395	2205
192	5611	4255	3577	3469	3207
256	9915	7531	6339	6123	5671
512	39,291	29,915	25,227	24,317	22,451
1024	156,411	119,227	100,635	96,933	89,529
2048	624,123	476,027	401,979	387,195	357,581

words of one operand, i.e., $e = \frac{f-3}{2}$, where $f = 2e + 3$ denotes the number of available registers in total (including the triple-size register for the accumulator). It shows that for $e > 10$ no significant improvement in speed is obtained. The performance decreases for smaller e and higher Integer sizes. However, if we compare our solution (160-bit multiplication with smallest parameter $e = 2 \rightarrow f = 7$ registers) with the product-scanning method (needing $f = 5$ registers), we obtain 3915 clock cycles for the operand-caching method and 3957 clock cycles for the product-scanning method. It therefore provides a good performance even for a smaller set of available registers. For the special case $e = 20$, where all 20 words of one 160-bit operand can be maintained in registers (ideal case for product scanning), it shows that the number of clock cycles reaches nearly the optimum of 2160 clock cycles, i.e., $4n = 80$ memory-access instructions, $n^2 = 400$ multiplications, and $3n^2 = 1200$ additions.

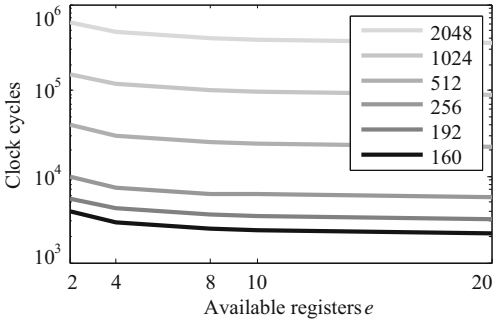

Fig. 7. Performance chart.

Table 6. Comparison with related work.

Method	Instruction						Clock cycles
	LD	ST	MUL	ADD	MOVW	Others	
Hybrid							
Gura et al. [8] ($d = 5$)	167	40	400	1360	355	197	3106
Uhsadel et al. [24] ($d = 5$)	238	40	400	986	355	184	2881
Scott et al. [21] ($d = 4$)[b]	200	40	400	1263	70	38	2651
Liu et al. [15] ($d = 4$)	200	40	400	1194	212	179	2865
Operand caching							
With looping[a,c] ($e = 9$)	92	66	400	1252	41	276	2685
Unrolled[b,c] ($e = 10$)	80	60	400	1240	2	68	2395

[a] b_{init}, Part 1, and Part 4 unrolled. Part 2 and Part 3 looped.
[b] Fully unrolled implementation without overhead of loop instructions.
[c] w/o `PUSH`/`POP`/`CALL`/`RET`

We compare our result with related work in Table 6. For a fair comparison, we also implemented an operand-caching version that does not unroll the algorithm but includes additional loop instructions. It shows that the operand-caching method provides best performance. Compared to Gura et al. [8] 23% less clock cycles are needed for a 160-bit multiplication. A 10% improvement could be achieved compared to the best solution reported in the literature [21]. Note that most of the related work needs to be between 167 and 238 *load* instructions which mostly explains the higher amount of needed clock cycles.

4.2. *Performance Results for the 32-bit ARM7TDMI*

The ARM7TDMI (ARM7 **T**humb **D**ebug **M**ultiplier **I**CE) was introduced by ARM in 1994 and has been used in a wide range of applications, e.g., mobile devices (e.g., produced by Nokia, Sony-Ericsson, Motorola, …), Apple's iPod, video game consoles (e.g., integrated by companies like Nintendo, SEGA, or Sony), routers, and automobile systems. It is a 32-bit RISC microcontroller that has been especially designed for low area and low power embedded systems. For the device, there exist two different instruction

sets: one that makes use of high-performance 32-bit instructions (the ARM set), and one that uses only 16-bit instructions (the THUMB set which is essentially a subset of the ARM instruction set). The latter configuration, however, has the main advantage that it significantly reduces the power consumption and the code size (approximately twice the density of a standard ARM code) and is therefore often applied in resource-constrained environments.

The controller essentially features a three-stage pipeline architecture (fetch, decode, and execute) and provides a Barrel shifter and a 32×8-bit hardware multiplier. For the standard ARM operating mode, 16 general-purpose registers are available to users R0…R15 whereas R13, R14, and R15 are special registers (program counter, link register, and stack pointer) that might not be used within custom applications. In THUMB mode, only 8 registers are available, i.e., R0…R7 which in general limits the applicability for many cryptographic algorithms. In the following, we therefore used the standard ARM operating mode since at least 10 registers are needed to perform a multi-precision multiplication.

We used R0, R1, and R2 as pointer registers to point to the memory location of the two operands and the final result. Furthermore, we reserved R3:R4:R5 as a triple accumulator register. R6 and R7 are used to store the result of partial products. Hence, 6 registers remain that can be used to hold (and cache) operands during multiplication. In order to improve performance, we also reused the link register R14 in subroutines by caching the value on the stack.

A single multiply–accumulate step basically requires four instructions. First, the operands are loaded using the `LDR` instruction. The loading of one operand needs 3 clock cycles. Then, a single 32×32-bit multiplication is performed using the `UMULL` instruction. This instruction needs up to 4 clock cycles to calculate a 64-bit result (depending on the value of the operands[5]). After that, three additional add-with-carry instructions (`ADDS`, `ADCS`, and `ADC`) are needed that add the partial product to the accumulator register. Storing (`STR`) of a 32-bit value needs 2 clock cycles. This results in up to 12 clock cycles in total.

For the hybrid multiplication method, we chose $d = 2$ which is the highest possible parameter on that platform. Similar to [21], we also decided to implement one carry-catcher register that stores two carries, i.e., in the lower and higher byte of the register. This can be done by using the `ADDCS` instruction which allows to add a carry into a specific byte of a register. The extraction of the carries can then be performed using 4 clock cycles. We further used 5 accumulator registers and shifted the values by simple code rearranging.

Table 7 shows the results for the implemented multiplication methods and different Integer sizes up to 256 bits. It shows that operand-caching multiplication needs only 281 clock cycles for 160-bit operands. 392 clock cycles are needed for 192-bit operands. Compared to Scott et al. [21], who implemented a 192-bit multiplication on the same platform, the results could be improved by a factor of 19.5%: the authors reported 580 clock cycles for product-scanning multiplication and 487 clock cycles for the hybrid multiplication.

[5] The early-terminating multiplier of the ARM7TDMI is susceptible to side-channel attacks that exploit the data-dependent runtime of the multiplier as shown in [6].

Table 7. ARM7TDMI results for different Integer sizes.

Size (bit)	Op. scan.	Prod. scan.	Hybrid method	Operand caching
160	493	357	–	281
192	699	506	424	392
256	1219	882	726	700

5. Conclusions

We presented a novel multiplication technique that is especially interesting for embedded microprocessors. The multiplication method significantly reduces the number of necessary *load* instructions through sophisticated caching of operands. Our solution is similar to the product-scanning method but divides the processing of operands into several parts/rows. This allows the scanning of sub-products where most of the operands are kept within the register-set throughout the algorithm.

In order to evaluate our solution, we implemented our and three other multiplication techniques using different Integer sizes on both the ATmega128 and ARM7TDMI microcontrollers. Using operand-caching multiplication, 2395 clock cycles are required for a 160-bit multiplication on the ATmega128. This outperforms the best reported solution by a factor of 10% [21]. Compared to the hybrid multiplication of Gura et al. [8], we achieved a speed up of 23%. Our evaluation further showed that our solution scales very well for different Integer sizes used for ECC and RSA. We obtained an improvement of about 15% for bit sizes between 256 and 2048 bits compared to a reference implementation of the hybrid multiplication. For the ARM7TDMI, operand caching requires only 281 clock cycles which improves the state-of-the-art performance by a factor of about 20%.

It is also worth to note that our multiplication method is perfectly suitable for processors that support multiply–accumulate (`MULACC`) instructions such as ARM9 or the dsPIC family of microcontrollers. It also fully complies with architectures which support instruction set extensions for `MULACC` operations such as proposed by Großschädl and Savaş [7].

Acknowledgements

The work has been supported by the European Commission through the ICT Program under Contract ICT-2007-216646 (European Network of Excellence in Cryptology—ECRYPT II) and under Contract ICT-SEC-2009-5-258754 (Tamper Resistant Sensor Node—TAMPRES).

A Algorithm for Operand-Caching Multiplication

The following pseudocode shows the algorithm for multi-precision multiplication using the operand-caching method. Variables that are located in data memory are denoted by

M_x where x represents the name of the Integer a or b. The parameter e describes the number of locally usable registers $R_a[e-1, \ldots, 0]$ and $R_b[e-1, \ldots, 0]$. The triple-word accumulator is denoted by $ACC = (ACC_2, ACC_1, ACC_0)$.

Require: word size n, parameter e, $n \geq e$, Integers $a, b \in [0, n)$, $c \in [0, 2n)$.
Ensure: $c = ab$.

$r = \lfloor n/e \rfloor$.
$R_A[e-1, \ldots, 0] \leftarrow M_A[n-1, \ldots, re]$.
$R_B[e-1, \ldots, 0] \leftarrow M_B[n-re-1, \ldots, 0]$.
$ACC \leftarrow 0$.
for $i = 0$ **to** $n - re - 1$ **do**
  for $j = 0$ **to** i **do**
    $ACC \leftarrow ACC + R_A[j] * R_B[i-j]$.
  end for
  $M_C[re+i] \leftarrow ACC_0$.
  $(ACC_1, ACC_0) \leftarrow (ACC_2, ACC_1)$.
  $ACC_2 \leftarrow 0$.
end for
for $i = 0$ **to** $n - re - 2$ **do**
  for $j = i+1$ **to** $n - re - 1$ **do**
    $ACC \leftarrow ACC + R_A[j] * R_B[n-re-j+i]$.
  end for
  $M_C[n+i] \leftarrow ACC_0$.
  $(ACC_1, ACC_0) \leftarrow (ACC_2, ACC_1)$.
  $ACC_2 \leftarrow 0$.
end for

} b_{init}

$M_C[2n-re-1] \leftarrow ACC_0$.
$ACC_0 \leftarrow 0$.
for $p = r-1$ **to** 0 **do**

} **Row Loop:**

  $R_A[e-1, \ldots, 0] \leftarrow M_A[(p+1)e-1, \ldots, pe]$.
  $R_B[e-1, \ldots, 0] \leftarrow M_B[e-1, \ldots, 0]$.
  for $i = 0$ **to** $e-1$ **do**
    for $j = 0$ **to** i **do**
      $ACC \leftarrow ACC + R_A[j] * R_B[i-j]$.
    end for
    $M_C[pe+i] \leftarrow ACC_0$.
    $(ACC_1, ACC_0) \leftarrow (ACC_2, ACC_1)$.
    $ACC_2 \leftarrow 0$.
  end for

} **Part 1**

for $i = 0$ **to** $n - (p+1)e - 1$ **do**
  $R_B[e-1, \ldots, 0] \leftarrow M_B[e+i]$, $R_B[e-2, \ldots, 1]$.
  for $j = 0$ **to** $e-1$ **do**
    $ACC \leftarrow ACC + R_A[j] * R_B[e-1-j]$.
  end for
  $ACC \leftarrow ACC + M_C[(p+1)e+i]$.
  $M_C[(p+1)e+i] \leftarrow ACC_0$.
  $(ACC_1, ACC_0) \leftarrow (ACC_2, ACC_1)$.
  $ACC_2 \leftarrow 0$.
end for } **Part 2**

for $i = 0$ **to** $n - (p+1)e - 1$ **do**
  $R_A[e-1, \ldots, 0] \leftarrow M_A[(p+1)e+i]$, $R_A[e-2, \ldots, 1]$.
  for $j = 0$ **to** $e-1$ **do**
    $ACC \leftarrow ACC + R_A[j] * R_B[e-1-j]$.
  end for
  $ACC \leftarrow ACC + M_C[(n+i]$.
  $M_C[n+i] \leftarrow ACC_0$.
  $(ACC_1, ACC_0) \leftarrow (ACC_2, ACC_1)$.
  $ACC_2 \leftarrow 0$.
end for } **Part 3**

for $i = 0$ **to** $e-2$ **do**
  for $j = i+1$ **to** $e-1$ **do**
    $ACC \leftarrow ACC + R_A[j] * R_B[e-j+i]$.
  end for
  $M_C[2n-(p+1)e+i] \leftarrow ACC_0$.
  $(ACC_1, ACC_0) \leftarrow (ACC_2, ACC_1)$.
  $ACC_2 \leftarrow 0$.
end for
$M_C[2n-1-pe] \leftarrow ACC_0$.
$ACC_0 \leftarrow 0$. } **Part 4**
end for
Return c.

References

[1] Atmel Corporation. 8-bit AVR microcontroller with 128K bytes in-system programmable flash (2007). http://www.atmel.com/dyn/resources/prod_documents/doc2467.pdf

[2] Atmel Corporation. 8-bit AVR instruction set (2008). http://www.atmel.com/dyn/resources/prod_documents/doc0856.pdf

[3] M. Aydos, T. Yanik, Ç. K. Koç, An high-speed ECC-based wireless authentication protocol on an ARM microprocessor, in *16th Annual Computer Security Applications Conference (ACSAC 2000), 11–15 December 2000, New Orleans, LA, USA* (IEEE, 2000), pp. 401–410

[4] Certicom Research. Standards for efficient cryptography, SEC 2: recommended elliptic curve domain parameters, version 2.0 (2010). http://www.secg.org/

[5] P. Comba, Exponentiation cryptosystems on the IBM PC. *IBM Syst. J.* **29**(4), 526–538 (1990)

[6] J. Großschädl, E. Oswald, D. Page, M. Tunstall, Side channel analysis of cryptographic software via early-terminating multiplications, in *Information, Security and Cryptology ? ICISC 2009, 12th International Conference, Seoul, Korea, December 2–4, 2009, Proceedings*, vol. 5984 of *Lecture Notes in Computer Science*, Seoul, Korea (Springer-Verlag, 2010), pp. 176–192

[7] J. Großschädl E. Savaş, Instruction set extensions for fast arithmetic in finite fields GF(p) and GF(2^m), in *CHES 2004, 6th International Workshop, Cambridge, MA, USA, August 11–13, 2004*, vol. 3156 of *LNCS* (Springer, 2004), pp. 133–147

[8] N. Gura, A. Patel, A. Wander, H. Eberle, S.C. Shantz, Comparing elliptic curve cryptography and RSA on 8-bit CPUs, in *CHES 2004, 6th International Workshop, Cambridge, USA, August 11–13, 2004* (Springer, 2004), pp. 119–132

[9] IAR Systems, IAR embedded workbench (2012). http://www.iar.com/

[10] A. Kargl, S. Pyka, H. Seuschek, Fast arithmetic on ATmega128 for elliptic curve cryptography. Cryptology ePrint Archive http://eprint.iacr.org/, Report 2008/442, (2008)

[11] Ç. K. Koç, High speed RSA implementation. Technical report, RSA Laboratories, RSA Data Security, Inc. 100 Marine Parkway, Suite 500 Redwood City (1994)

[12] C. Lederer, R. Mader, M. Koschuch, J. Großschädl, A. Szekely, S. Tillich, Energy-efficient implementation of ECDH key exchange for wireless sensor networks, in *3rd International Workshop in Information Security Theory and Practices—WISTP 2009, Brussels, Belgium, September 1–4, 2009*, vol. 5746 of *LNCS* (Springer, 2009), pp. 112–127

[13] A. Lenstra, E. Verheul, Selecting cryptographic key sizes. *J. Cryptol.* **14**(4), 255–293 (2001)

[14] A. Liu, P. Ning, TinyECC: a configurable library for elliptic curve cryptography in wireless sensor networks, in *International Conference on Information Processing in Sensor Networks—IPSN 2008, April 22–24, 2008, St. Louis, MO, USA*, St. Louis, MO (2008), pp. 245–256

[15] Z. Liu, J. Großschädl, I. Kizhvatov, Efficient and side-channel resistant RSA implementation for 8-bit AVR microcontrollers, in *Workshop on the Security of the Internet of Things—SOCIOT 2010, 1st International Workshop, November 29, 2010, Tokyo, Japan* (IEEE Computer Society, 2010)

[16] M. Medwed, E. Oswald, Template attacks on ECDSA, in K.-I. Chung, M. Yung, K. Sohn, editors, *9th International Workshop on Information Security Applications (WISA 2008), Jeju Island, Korea, September 23–25, 2008, Pre-Proceedings* (2008), pp. 14–27

[17] National Institute of Standards and Technology (NIST), SP800-57 Part 1: DRAFT recommendation for key management: part 1: general (2011). http://csrc.nist.gov/publications/drafts/800-57/Draft_SP800-57-Part1-Rev3_May2011.pdf

[18] nongnu.org. AVR Libc Home Page (2012). http://www.nongnu.org/avr-libc/

[19] J. Pelzl, T. Wollinger, J. Guajardo, C. Paar, Hyperelliptic curve cryptosystems: closing the performance gap to elliptic curves, in C.D. Walter, Ç.K. Koç, C. Paar, editors, *Cryptographic Hardware and Embedded Systems—CHES 2003, 5th International Workshop, Cologne, Germany, September 8–10, 2003, Proceedings*, vol. 2779 of *Lecture Notes in Computer Science* (2003), pp. 351–365

[20] Rowley. Crossworks for AVR (2012). http://www.rowley.co.uk/avr/

[21] M. Scott, P. Szczechowiak, Optimizing multiprecision multiplication for public key cryptography. Cryptology ePrint Archive http://eprint.iacr.org/, Report 2007/299 (2007)

[22] P. Szczechowiak, L.B. Oliveira, M. Scott, M. Collier, R. Dahab, NanoECC: testing the limits of elliptic curve cryptography in sensor networks, in R. Verdone, editor, *Wireless Sensor Networks 5th European Conference, EWSN 2008, Bologna, Italy, January 30-February 1, 2008*, vol. 4913 of *LNCS* (Springer, 2008), pp. 305–320

[23] O. Ugus, A. Hessler, D. Westhoff, Performance of additive homomorphic EC-ElGamal encryption for TinyPEDS, in *GI/ITG KuVS Fachgespr?ch ?Drahtlose Sensornetze?, RWTH Aachen, 2007* UbiSec (2007)

[24] L. Uhsadel, A. Poschmann, C. Paar, Enabling full-size public-key algorithms on 8-bit sensor nodes, in *Security and Privacy in Ad-hoc and Sensor Networks 4th European Workshop, ESAS 2007, Cambridge, UK, July 2–3, 2007* (2007)

[25] H. Wang, Q. Li, Efficient implementation of public key cryptosystems on mote sensors, in *Information and Communications Security 8th International Conference, ICICS 2006, Raleigh, NC, USA, December 4–7, 2006*, vol. 4307 of *LNCS* (Springer, 2006), pp. 519–528

[26] S.-B. Xu, L. Batina, Efficient implementation of elliptic curve cryptosystems on an ARM7 with hardware accelerator. In G.I. Davida, Y. Frankel, editors, *Information Security 4th International Conference, ISC 2001 Malaga, Spain, October 1–3, 2001 Proceedings*, vol. 2200 of *Lecture Notes in Computer Science* (Springer, 2001), pp. 266–279