



HAL
open science

Q-adic Transform revisited

Jean-Guillaume Dumas

► **To cite this version:**

| Jean-Guillaume Dumas. Q-adic Transform revisited. 2007. hal-00173894v4

HAL Id: hal-00173894

<https://hal.science/hal-00173894v4>

Preprint submitted on 21 Nov 2007 (v4), last revised 23 Jun 2008 (v6)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Q-adic Transform revisited

Jean-Guillaume Dumas*

November 21, 2007

Abstract

We present an algorithm to perform fast modular polynomial multiplication. The idea is to convert the X -adic representation of modular polynomials, with X an indeterminate, to a q -adic representation where q is a prime power larger than the field characteristic. With some control on the different involved sizes it is then possible to perform some of the q -adic arithmetic directly with machine integers or floating points. Depending also on the number of performed numerical operations one can then convert back to the q -adic or X -adic representation and eventually mod out high residues. In this note we present a new version of both conversions: more tabulations and a way to reduce the number of divisions involved in the process are presented. The polynomial multiplication is then applied to arithmetic in small finite field extensions.

1 Introduction

The FFLAS/FFPACK project has demonstrated the need of a wrapping of cache-aware routines for efficient small finite field linear algebra [3, 4].

A conversion between a modular representation of prime fields and e.g. floating points used exactly is natural. It uses the homomorphism to the integers. Now for extension fields (isomorphic to polynomials over a prime field) such a conversion is not direct. In [3] we propose to transform the polynomials into a q -adic representation where q is a prime power larger than the field characteristic. We call this transformation DQT for Discrete Q-adic Transform. With some care, in particular on the size of q , it is possible to map the operations in the extension field into the floating point arithmetic realization of this q -adic representation and convert back using an inverse DQT.

In this note we propose some implantation improvements: we propose to use a tabulated Zech logarithm for the DQT and we give a trick to reduce the number of machine divisions involved in the inverse. This thus gives rise to an improved DQT which we thus call FQT (Fast Q-adic Transform). Therefore we recall in section 2 the previous conversion algorithm. We then present our new

*Laboratoire J. Kuntzmann, Université de Grenoble, umr CNRS 5224. BP 53X, 51, rue des Mathématiques. F38041 Grenoble, France. Jean-Guillaume.Dumas@imag.fr

reduction in section 3. We then show in section 4 how a time-memory trade-off can make this reduction very fast. This can then be applied on modular polynomial multiplication with small prime fields in section 5 as well as for small extension field arithmetic and fast matrix multiplication in section 6.

2 Q-adic representation of polynomials

We follow here the presentation of [3] of the idea of [7]: polynomial arithmetic is performed a q -adic way, with q a sufficiently big prime or power of a single prime.

Suppose that $a = \sum_{i=0}^{k-1} \alpha_i X^i$ and $b = \sum_{i=0}^{k-1} \beta_i X^i$ are two polynomials in $\mathbb{Z}/p\mathbb{Z}[X]$. One can perform the polynomial multiplication ab via q -adic numbers. Indeed, by setting $\tilde{a} = \sum_{i=0}^{k-1} \alpha_i q^i$ and $\tilde{b} = \sum_{i=0}^{k-1} \beta_i q^i$, the product is computed in the following manner (we suppose that $\alpha_i = \beta_i = 0$ for $i > k - 1$):

$$\tilde{a}\tilde{b} = \sum_{j=0}^{2k-2} \left(\sum_{i=0}^j \alpha_i \beta_{j-i} \right) q^j \quad (1)$$

Now if q is large enough, the coefficient of q^i will not exceed q . In this case, it is possible to evaluate a and b as machine numbers (e.g. floating point or machine integers), compute the product of these evaluations, and convert back to polynomials by radix computations (see e.g. [5, Algorithm 9.14]). There just remains then to perform modulo p reductions on every coefficient. We call DQT the evaluation of polynomials modulo p at q and DQT inverse the radix conversion of a q -adic development followed by a modular reduction, as shown in algorithm 1.

Algorithm 1 Polynomial multiplication by DQT

Require: Two polynomials v_1 and v_2 in $\mathbb{Z}/p\mathbb{Z}[X]$ of degree less than k .

Require: a prime power q .

Ensure: $R \in \mathbb{Z}/p\mathbb{Z}[X]$, with $R = v_1.v_2$.

Polynomial to q -adic conversion

- 1: Set \tilde{v}_1 and \tilde{v}_2 to the floating point vectors of the evaluations at q of the elements of v_1 and v_2 . {Using e.g. Horner's formula}

One computation

- 2: Compute $\tilde{r} = \tilde{v}_1 \tilde{v}_2$

Building the solution

- 3: $\tilde{r} = \sum_{i=0}^{2k-2} \tilde{\mu}_i q^i$. {Using radix conversion, see e.g. [5, Algorithm 9.14]}
 - 4: For each i , set $\mu_i = \tilde{\mu}_i \pmod p$
 - 5: set $R = \sum_{i=0}^{2k-2} \mu_i X^i$
-

Depending on the size of q , the results can still remain exact:

Theorem 1. [3] Let m be the number of available mantissa bits within the machine numbers and n_q be the number of machine products $v_1.v_2$ accumulated, e.g. in a dot product, before the re-conversion. If

$$q > n_q k (p - 1)^2 \text{ and } (2k - 1) \log[2](q) < m, \quad (2)$$

then Algorithm 1 is correct.

Note that the prime power q can be chosen to be a power of 2. Then the Horner like evaluation of the polynomials at q (line 1 of algorithm 1) is just a left shift. One can then compute this shift with exponent manipulations in floating point arithmetic and use native shift operator (e.g. the \ll operator in C) as soon as values are within the 32 (or 64 when available) bit range.

In the following we will thus always consider that q is a power of two.

It is shown on [3, Figures 5 & 6] that this wrapping is already a pretty good way to obtain high speed linear algebra over some small extension fields. Indeed we were able to reach high peak performance, quite close to those obtained with prime fields, namely 420 Mop/s on a PIII, 735 MHz, and more than 500 Mop/s on a 64-bit DEC alpha 500 MHz. This is roughly 20 percent below the pure floating point performance and 15 percent below the prime field implementation.

3 REDQ: modular reduction in the DQT domain

The first improvement we propose to the DQT is to replace the costly modular reduction of the polynomial coefficients by a single division by p (or, better, by a multiplication by its inverse) followed by several shifts. The idea is summarized in algorithm 2 (note that when q is a power of 2 division by q^i and flooring are a single operation, a right shift).

Algorithm 2 REDQ

Require: a prime p and a prime power q satisfying the conditions (2).

Require: $\tilde{r} = \sum_{i=0}^d \tilde{\mu}_i q^i \in \mathbb{Z}$.

Ensure: $\rho \in \mathbb{Z}$, with $\rho = \sum_{i=0}^d \mu_i q^i$ where $\mu_i = \tilde{\mu}_i \pmod{p}$.

- 1: $rop = \left\lfloor \frac{\tilde{r}}{p} \right\rfloor$;
 - 2: **for** $i = 0$ to d **do**
 - 3: $u_i = \left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor - p \left\lfloor \frac{rop}{q^i} \right\rfloor$;
 - 4: **end for**
 - 5: $\mu_d = u_d$
 - 6: **for** $i = 0$ to $d - 1$ **do**
 - 7: $\mu_i = u_i - q u_{i+1} \pmod{p}$;
 - 8: **end for**
 - 9: Return $\rho = \sum_{i=0}^d \mu_i q^i$;
-

Theorem 2. *Algorithm REDQ is correct.*

We first need the following lemma:

Lemma 1. *For $r \in \mathbb{N}$ and $a, b \in \mathbb{N}^*$,*

$$\left\lfloor \frac{\left\lfloor \frac{r}{b} \right\rfloor}{a} \right\rfloor = \left\lfloor \frac{\left\lfloor \frac{r}{a} \right\rfloor}{b} \right\rfloor$$

Proof. We proceed by splitting the possible values of r into intervals $kab \leq r < (k+1)ab$. Then $kb \leq \frac{r}{a} < (k+1)b$ and since kb is an integer we also have that $kb \leq \left\lfloor \frac{r}{a} \right\rfloor < (k+1)b$. Thus $k \leq \frac{\left\lfloor \frac{r}{a} \right\rfloor}{b} < k+1$ and $\left\lfloor \frac{\left\lfloor \frac{r}{a} \right\rfloor}{b} \right\rfloor = k$. Obviously the same is true for the left hand side which proves the lemma. \square

PROOF of theorem 2. First we need to prove that $0 \leq u_i < p$. By definition of the truncation, we have $\frac{\tilde{r}}{q^i} - 1 < \left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor \leq \frac{\tilde{r}}{q^i}$ and $\frac{\tilde{r}}{pq^i} - 1 - \frac{1}{q^i} < \left\lfloor \frac{rop}{q^i} \right\rfloor \leq \frac{\tilde{r}}{pq^i}$. Thus $-1 < u_i < p + \frac{p}{q^i}$, which is $0 \leq u_i \leq p$ since u_i is an integer. We now consider the possible case $u_i = p$ and show that it does not happen. $u_i = p$ means that $\left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor = p(1 + \left\lfloor \frac{rop}{q^i} \right\rfloor) = pg$. This means that $pgq^i \leq r < pgq^i + q^i$. So that in turns $gq^i \leq rop \leq \frac{\tilde{r}}{p} < gq^i + \frac{q^i}{p}$. Thus $g \leq \frac{rop}{q^i} < g + \frac{1}{p}$ so that $\left\lfloor \frac{rop}{q^i} \right\rfloor = g$. But then from the definition of g we have that $g = g - 1$ which is absurd. Therefore $0 \leq u_i \leq p - 1$.

Second we show that $u_i = \sum_{j=i}^d \mu_j q^{j-i} \pmod{p}$. Well $u_i = \left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor - p \left\lfloor \frac{\left\lfloor \frac{\tilde{r}}{p} \right\rfloor}{q^i} \right\rfloor$

and thus lemma 1 gives that $u_i = \left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor - p \left\lfloor \frac{\left\lfloor \frac{\tilde{r}}{p} \right\rfloor}{p} \right\rfloor$. The latter is $u_i = \left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor \pmod{p}$. Now, since $\tilde{r} = \sum_{j=0}^{2k-2} \widetilde{\mu}_j q^j$, we have that $\left\lfloor \frac{\tilde{r}}{q^i} \right\rfloor = \sum_{j=i}^{2k-2} \widetilde{\mu}_j q^{j-i}$. Therefore, as $\mu_j = \widetilde{\mu}_j \pmod{p}$, the equality is proven. \square

In the algorithm, the computation of rop has to be a division in exact arithmetic. The following lemma gives bounds for which this division can be performed by a floating point multiplication with the precomputed inverse of p , as is done e.g. in NTL [8].

Lemma 2. *For a prime p and $r, t \in \mathbb{N}^*$ with $r < tp$, then*

$$\left\lfloor \frac{r}{p} \right\rfloor = \left\lfloor r \left(\frac{1}{p} + \epsilon \right) \right\rfloor \quad \text{as long as } 0 \leq \epsilon < \frac{1}{p^2(t+1-1/p)}.$$

Proof. Consider $up \leq r < up + i$ with u, i positive integers and $i < p$. Then $\left\lfloor \frac{r}{p} \right\rfloor = u$ and $r(\frac{1}{p} + \epsilon) = u + \frac{i}{p} + (up+i)\epsilon$. Then $\frac{i}{p} + (up+i)\epsilon < 1$ if $\epsilon < \frac{p-i}{p(up+i)}$. The latter right hand side is decreasing in i and is therefore minimal at $i = p-1$ for which it is $\frac{1}{p^2(u+1-1/p)}$. \square

Therefore, it is possible to perform the division by a multiplication by the precomputed inverse of the prime number. Since entries are already in floating point format this is a potential significant speed-up. Remark that for ϵ to be positive, the precomputed inverse has just to be rounded towards $+\infty$. Then since the division is correctly rounded, $\epsilon < 2^{-m}$, where m is the available floating point mantissa, and the non reduced r can be as large as

$$r < \frac{2^m}{p} + 1$$

while remaining correct.

For instance, take the polynomial $R = 1234X^3 + 5678X^2 + 9123X + 4567$, the prime $p = 23$ and use $q = 10^6$ to see what is happening. The trick is that by dividing only once by 23 all the coefficients of R are divided at once: $rop = \lfloor 1234005678009123004567/23 \rfloor = 53652420783005348024$. Then $rop \times 23 = 1234005678009123004552$ so that $u_0 = 15$. Then we shift to get 1234005678009123 and $53652420783005 \times 23 = 1234005678009115$ which gives $u_1 = 8$. We shift and multiply twice to get $u_2 = 18$ and $u_3 = \mu_3 = 15$. Now $-q = -10^6 \pmod{23} = 17$ and thus we have to compute

$$\mu = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 17 & 1 & 0 & 0 \\ 0 & 17 & 1 & 0 \\ 0 & 0 & 17 & 1 \end{bmatrix} u \pmod{p}$$

to get the final result, $R = 15X^3 + 20X^2 + 15X + 13 \pmod{23}$.

As is, there is no benefit to use this method when compared to direct remaindering by p . The trick is that this last step from the u_i to the μ_i can be tabulated since u is of much smaller size than R .

4 Time-Memory trade-off in REDQ

Indeed, there is a bijection between the u_i and the μ_i since

$$Q_d = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ -q & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -q & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ q & \ddots & \ddots & & \vdots \\ q^2 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ q^d & \dots & q^2 & q & 1 \end{bmatrix}^{-1}$$

If this operation is fully tabulated, it requires a table of size at least p^{d+1} . But, due to the nature of Q_d , we have the relations of figure 1.

Therefore, it is very easy to tabulate with a table of size p^k only and perform $\frac{d-1}{k-1}$ table accesses. The time memory trade-off, after the computation of the u_i , which requires 1 div & $(d+1)$ mul & $2d$ shifts, is thus as shown on table 1.

$$Q_{2d-1} = \begin{bmatrix} \boxed{Q_d} & 0 \\ & \boxed{1} \\ 0 & \boxed{Q_d} \end{bmatrix} \quad Q_{2d} = \begin{bmatrix} \boxed{Q_{d+1}} & 0 \\ & \boxed{1} \\ 0 & \boxed{Q_d} \end{bmatrix}$$

Figure 1: Recurring relations on the Q_d matrices.

Memory	Extra time
0	d (mul,add,mod)
p^2	d accesses
p^k	$\left\lceil \frac{d}{k-1} \right\rceil$ accesses
p^{d+1}	1 access

Table 1: Time-Memory trade-off in REDQ

In practice, indexing by a t -tuple of integers mod p is made by evaluating at p , as $\sum u_i p^i$. If a few more memory space is available, one can also directly index in the binary format using $\sum u_i (2^{\lceil \log_2(p) \rceil})^i$. On the one hand all the multiplications by p are replaced by binary shifts. On the other hand, this makes the table grow a little bit, from p^k to $2^{\lceil \log_2(p) \rceil k}$.

5 Comparison with delayed reduction for polynomial multiplication

The classical alternative to algorithm 1 to perform modular polynomial multiplication is to use delayed reductions e.g. as in [1]: the idea is to accumulate products of the form $\sum_i a_i b_{k-i}$, without reductions, while the sum does not overflow. Thus, if we use for instance a centered representation modulo p (integers from $\frac{1-p}{2}$ to $\frac{p-1}{2}$), it is possible to accumulate at least n_d products as long as

$$n_d(p-1)^2 < 2^{m+1} \quad (3)$$

The modular reduction can be made by many different ways (e.g. classical division, floating point multiplication by the inverse, Montgomery reduction, etc.), we just call the best one REDC here. It is at most equivalent to 1 machine division.

Now the idea of the FQT (Fast Q -adic Transform) is to represent modular polynomials of the form $P = \sum_{i=0}^N a_i X^i$ by $P = \sum_{i=0}^{N/k} P_i X^i$ where the P_i are degree k polynomials stored in a single integer in the q -adic way. Therefore, a product PQ has the form $\sum (\sum P_i Q_{t-i}) X^t$. There, each multiplication $P_i Q_{t-i}$ is made by algorithm 1 on a single machine integer. The reduction is made by

a tabulated REDQ and can also be delayed now as long as conditions (2) are guaranteed.

For the complexity, table 2 gives the respective complexities of both strategies.

Complexity	Multiplications	Reductions
Delayed	N^2	$\frac{N^2}{n_d}$ REDC
FQT	$\left(\frac{N}{k}\right)^2$	$\frac{1}{n_q} \left(\frac{N}{k}\right)^2$ REDQ

Table 2: Modular polynomial multiplication complexities.

For instance, with $p = 3$, $N = 100$, if we choose a double floating point representation and a degree 4 DQT (i.e. $k = 4$), the fully tabulated FQT with boils down to 1024 multiplications and 57 divisions. For the same parameters, the classical polynomial multiplication algorithm requires 10^4 multiplications and only 1 remaindering, which is roughly 10 times more operations as shown on figure 2.

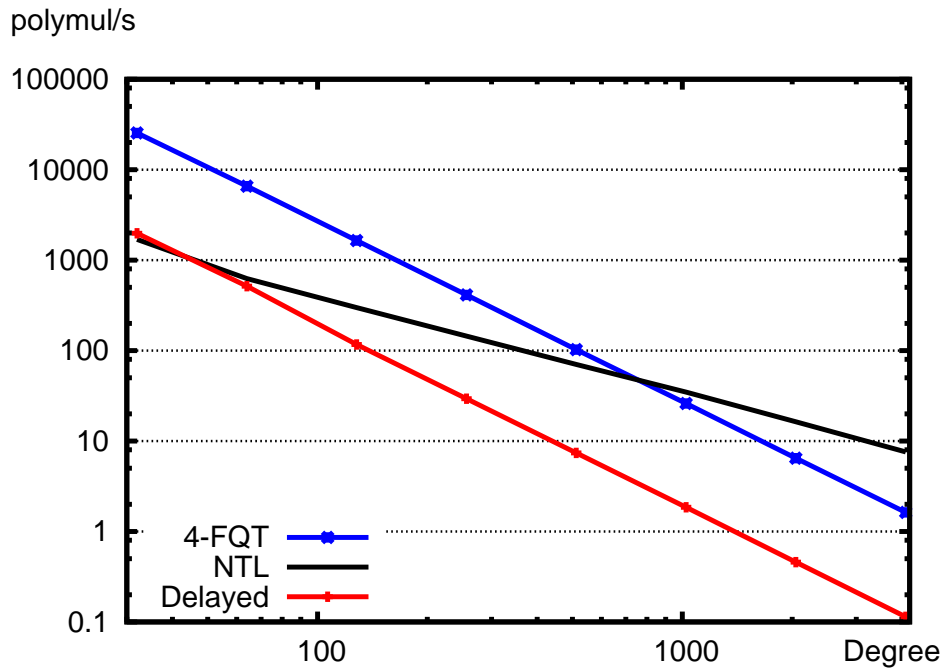


Figure 2: Polynomial multiplications modulo 3 per second on a Xeon 3.6 GHz

Even by switching to a larger mantissa, say e.g. 128 bits, so that the DQT multiplications are roughly 4 times costlier than double floating point opera-

tions, this can still be useful: take $p = 1009$ and choose $k = 3$, gives 1521 multiplications over 128 bits and 73 divisions. This should still be faster than the delayed.

On figure 2, we compare also our two implementations with that of NTL [8]. We see that the FQT is faster than NTL as long as better algorithms are not used. Indeed the change of slope in NTL's curve reflects the use of Karatsuba's algorithm for polynomial multiplication. Our strategy is thus very useful for small degrees and small primes. Furthermore, we have not implemented the FQT as the base case of faster recursive algorithms such as Karatsuba, Toom-Cook, etc. The figure shows that these recursive algorithms together with the FQT could be the fastest-s.

In particular, the FQT already improves the speed of small finite field extension's arithmetic as shown next.

6 Application to small finite field extensions

We use here the isomorphism between finite field to see any finite field extension as the set of polynomials modulo a prime p and an irreducible polynomial \mathcal{P} of degree k . Clearly we can convert any finite field element to its q -adic expansion ; perform the FQT between two elements and then reduce the obtained polynomial modulo \mathcal{P} . In this paper we propose to improve all the conversion steps of [3, algorithm 4.1] in order to approach the performance of the prime field wrapping also for several extension fields:

1. Replace the Horner evaluation of the polynomials, to form the q -adic expansion, by a single table lookup recovering directly the floating point representation.
2. Replace the radix conversion and the costly modular reductions of each polynomial coefficient, by a single half REDQ operation.
3. Replace the polynomial division by two table lookups and a single field operation.

Indeed, suppose the internal representation of the extension field is already by Zech logarithms and uses conversion tables from polynomial to index representations. See e.g. [1] for more details. Then we choose a time-memory trade-off for the REDQ operation of the same order of magnitude, that is to say p^k . The overall memory required by these new tables only doubles and the REDQ requires only 2 accesses. Moreover, in the small extension, the polynomial multiplication must also be reduced by an irreducible polynomial, \mathcal{P} . We show next that this reduction can be precomputed in the REDQ table lookup and is therefore almost free.

More precisely, we propose algorithm 3.

A variant of REDQ is used in algorithm 3, but u_i still satisfies $u_i = \sum_{j=i}^{2k-2} \mu_j q^{j-i} \pmod p$ as shown in theorem 2. Therefore the representations of $\sum \mu_i X^i$ in the

Algorithm 3 Fast Dot product over Galois fields via FQT and FQT inverse

Require: a field $\mathbf{GF}(p^k)$ with elements represented as exponents of a generator of the field.

Require: Two elements v_1 and v_2 of $\mathbf{GF}(p^k)$.

Require: a prime power q .

Ensure: $R \in \mathbf{GF}(p^k)$, with $R = v_1 \cdot v_2$.

Tabulated q -adic conversion

{Use conversion tables from exponent to floating point evaluation}

- 1: Set \tilde{v}_1 and \tilde{v}_2 to the floating point vectors of the evaluations at q of the elements of v_1 and v_2 .

The floating point computation

- 2: Compute $\tilde{r} = \tilde{v}_1 \tilde{v}_2$;

Computing a radix decomposition

- 3: $r = \lfloor \tilde{r} \rfloor$; { $r = \tilde{r}$ but we might need a conversion to an integral type}
- 4: $rop = \lfloor \frac{\tilde{r}}{p} \rfloor$;
- 5: **for** $i = 0$ to $2k - 2$ **do**
- 6: $u_i = \lfloor \frac{r}{q^i} \rfloor - p \lfloor \frac{rop}{q^i} \rfloor$;
- 7: **end for**

Tabulated radix conversion to exponents of the generator

{ μ_i is such that $\mu_i = \tilde{\mu}_i \pmod p$ for $\tilde{r} = \sum_{i=0}^{2k-2} \tilde{\mu}_i q^i$ }

- 8: Set $L = \text{representation}(\sum_{i=0}^{k-2} \mu_i X^i)$.
- 9: Set $H = \text{representation}(X^{k-1} \times \sum_{i=k-1}^{2k-2} \mu_i X^{i-k+1})$.

Reduction in the field

- 10: Return $R = H + L \in \mathbf{GF}(p^k)$;
-

field can be precomputed and stored in a table where the indexing will be made by (u_0, \dots, u_{k-1}) and $(u_{k-1}, \dots, u_{2k-2})$ and not by the μ_i 's. Note also that the representation of X^{k-1} can be just $k - 1$ if the irreducible polynomial used to build $\mathbf{GF}(p^k)$ is primitive and X has been chosen as the generator.

Theorem 3. *Algorithm 3 is correct.*

Proof. There remains to prove that it is possible to compute L and H from the u_i . From the equality above, we see that $\mu_{2k-2} = u_{2k-2}$ and $\mu_i = u_i - qu_{i+1} \pmod p$, for $i = 0..(2k - 3)$. Therefore a p^k elements precomputed table, indexed by (u_0, \dots, u_{k-1}) , can provide the representation of

$$L = \sum_{i=0}^{k-2} (u_i - qu_{i+1} \pmod p) X^i.$$

Another table with p^k elements, indexed by $(u_{k-1}, \dots, u_{2k-2})$, can provide the

representation of

$$H = u_{2k-2}X^{2k-2} + \sum_{i=k-1}^{2k-3} (u_i - qu_{i+1} \pmod p)X^i.$$

Finally $R = X^{k-1} \times \sum_{i=k-1}^{2k-2} \mu_i X^{i-k+1} + \sum_{i=0}^{k-2} \mu_i X^i$ needs to be reduced modulo the irreducible polynomial used to build the field. But, if we are given the representations of H and L in the field, R is then equal to their addition inside the field, directly using the internal representations. \square

Table 3 recalls the respective complexities of conversion phase in the two presented algorithms.

Memory	Alg. 1 $3p^k$	Alg. 3 $6p^k$	Alg. 3 $4p^k + 2^{k \lceil \log_2 p \rceil + 1}$
Shift	$4k - 2$	$4k - 2$	$4k - 2$
Add	$4k - 4$	0	$2k - 1$
Axpy	0	$4k - 3$	$2k - 1$
Div	$2k - 1$	0	0
Table	0	3	3
Red	$\geq 5k$	4	4

Table 3: Complexity of the back and forth conversion between extension field and floating point numbers

Figure 3 shows only the speed of the conversion after the floating point operations. The log scales prove that for q ranging from 2^1 to 2^{26} (on a 32 bit Xeon) our new implantation is two to three times faster than the previous one.

Furthermore, these improvements e.g. allow the extension field routines to reach the speed of 7800 millions of GF(9) operations per second (on a XEON, 3.6 GHz, using Goto BLAS-1.09 as the numerical routines and FFLAS for the fast matrix multiplication) as shown on figure 4. The FFLAS routines are available within the LinBox 1.1.4 library [6] and the FQT is implemented in the `givgfqext.h` file of the Givaro 3.2.9 library [2].

With these new implementations, the obtained speed-up shown in figure 4 represents a reduction from the 15 percent overhead of the previous implementation to less than 4 percent now, when compared to GF(11).

References

- [1] Jean-Guillaume Dumas. Efficient dot product over finite fields. In Victor G. Ganzha, Ernst W. Mayr, and Evgenii V. Vorozhtsov, editors, *Proceedings of the seventh International Workshop on Computer Algebra in Scientific Computing, Yalta, Ukraine*, pages 139–154. Technische Universität München, Germany, July 2004.

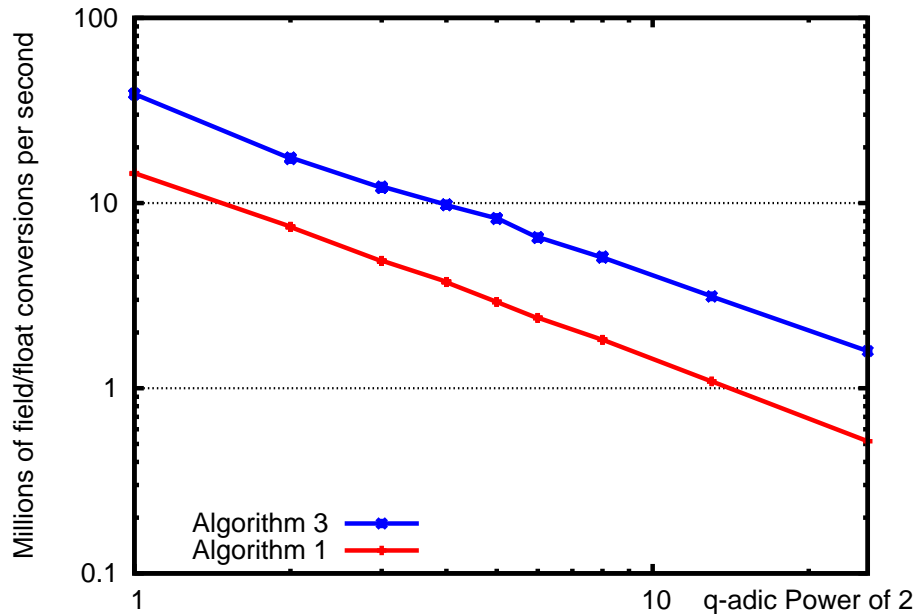


Figure 3: Small extension field conversion speed on a Xeon 3.6GHz

- [2] Jean-Guillaume Dumas, Thierry Gautier, Pascal Giorgi, Clément Pernet, Jean-Louis Roch, and Gilles Villard. Givaro 3.2.9: C++ library for arithmetic and algebraic computations, 2007. ljk.imag.fr/CASYS/LOGICIELS/givaro.
- [3] Jean-Guillaume Dumas, Thierry Gautier, and Clément Pernet. Finite field linear algebra subroutines. In Teo Mora, editor, *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation, Lille, France*, pages 63–74. ACM Press, New York, July 2002.
- [4] Jean-Guillaume Dumas, Pascal Giorgi, and Clément Pernet. FFPACK: Finite field linear algebra package. In Jaime Gutierrez, editor, *Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation, Santander, Spain*, pages 119–126. ACM Press, New York, July 2004.
- [5] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, New York, NY, USA, 1999.
- [6] The LinBox Group. Linbox 1.1.4: Exact computational linear algebra, 2007. www.linalg.org.
- [7] B. David Saunders. Personal communication, 2001.

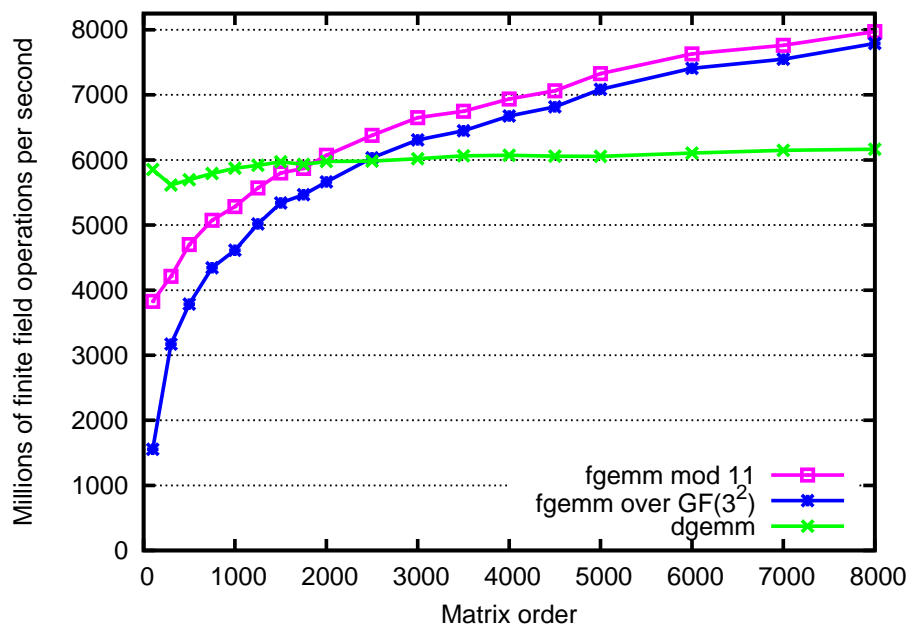


Figure 4: Speed of finite field Winograd matrix multiplication on a XEON, 3.6 GHz

- [8] Victor Shoup. NTL 5.4.1: A library for doing number theory, 2007.
www.shoup.net/ntl.