

## Better polynomials for GNFS

Shi Bai, Cyril Bouvier, Alexander Kruppa, Paul Zimmermann

► **To cite this version:**

Shi Bai, Cyril Bouvier, Alexander Kruppa, Paul Zimmermann. Better polynomials for GNFS. Mathematics of Computation / Mathematics of Computation, American Mathematical Society, 2016, 85, pp.12. <10.1090/mcom3048>. <hal-01089507>

**HAL Id: hal-01089507**

**<https://hal.inria.fr/hal-01089507>**

Submitted on 1 Dec 2014

**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.

## BETTER POLYNOMIALS FOR GNFS

SHI BAI, CYRIL BOUVIER, ALEXANDER KRUPPA, AND PAUL ZIMMERMANN

ABSTRACT. The general number field sieve (GNFS) is the most efficient algorithm known for factoring large integers. It consists of several stages, the first one being polynomial selection. The quality of the selected polynomials can be modelled in terms of size and root properties. We propose a new kind of polynomials for GNFS: with a new degree of freedom, we further improve the size property. We demonstrate the efficiency of our algorithm by exhibiting a better polynomial than the one used for the factorization of RSA-768, and a polynomial for RSA-1024 that outperforms the best published one.

### 1. INTRODUCTION TO GNFS

The general number field sieve [11] is the most efficient algorithm known for factoring large integers. It has been used in many record factorizations such as RSA-704 [3] and RSA-768 [18]. GNFS consists of several stages including polynomial selection, sieving, filtering, linear algebra and finding square roots.

Let  $n$  be the integer to be factored. In polynomial selection, we want to choose two irreducible and coprime polynomials  $f(x)$  and  $g(x)$  over  $\mathbb{Z}$  which share a common root  $m$  modulo  $n$ . In practice, the homogenized polynomials  $F(x, y)$  and  $G(x, y)$  are often used. We want to find many coprime pairs  $(a, b) \in \mathbb{Z}^2$  such that the polynomials values  $F(a, b)$  and  $G(a, b)$  are simultaneously smooth with respect to some bounds  $B_1$  and  $B_2$ . An integer is smooth with respect to bound  $B$  (or  $B$ -smooth) if none of its prime factors are larger than  $B$ . The line sieving and lattice sieving [17] are commonly used to identify such pairs  $(a, b)$ .

The running-time of sieving depends on the quality of the chosen polynomials in polynomial selection, hence many polynomial pairs will be generated and optimized in order to produce a good one.

This paper proposes a new kind of polynomials for the number field sieve, together with a corresponding algorithm to generate such polynomials, and exhibits better polynomials found by this algorithm for RSA challenge numbers.

### 2. POLYNOMIAL SELECTION

For large integers, most methods for polynomial selection [6, 9, 10, 12, 13] in GNFS use a linear polynomial for  $g(x)$  and a nonlinear one for  $f(x)$  (degree 6 in latest records). The standard method to generate such polynomial pairs is to expand  $n$  in base- $(m_1, m_2)$  so  $n = \sum_{i=0}^d c_i m_1^i m_2^{d-i}$ . The polynomial pair is given by  $f(x) = \sum_{i=0}^d c_i x^i$  and  $g(x) = m_2 x - m_1$ .

---

2010 *Mathematics Subject Classification.* Primary 11Y05, 11Y16.

The running-time of sieving depends on the smoothness of the polynomial values  $|F(a, b)|$  and  $|G(a, b)|$ . Let  $\Psi(z, z^{1/u})$  be the number of  $z^{1/u}$ -smooth integers below  $z$  for some  $u > 0$ . The Dickman-de Bruijn function  $\rho(u)$  [7] is often used to estimate the density of smooth numbers  $\Psi(z, z^{1/u})$ . It can be shown that

$$\lim_{z \rightarrow \infty} \frac{\Psi(z, z^{1/u})}{z} = \rho(u).$$

The Dickman-de Bruijn function satisfies the differential-difference equation

$$u\rho'(u) + \rho(u-1) = 0, \quad \rho(u) = 1 \quad \text{for } 0 \leq u \leq 1.$$

It may be shown that  $\rho$  satisfies the asymptotic estimate

$$\log(\rho(u)) = -(1 + o(1))u \log u \quad \text{as } u \rightarrow \infty.$$

For practical purposes, the frequency of smooth numbers can be approximated by the Canfield-Erdős-Pomerance theorem, which can be stated as follows (Corollary 1.3 from [8]):

**Theorem 2.1.** *For any fixed  $\epsilon > 0$ , we have*

$$\Psi(z, z^{1/u}) = zu^{-u(1+o(1))}$$

as  $z^{1/u}$  and  $u$  tend to infinity, uniformly in the region  $z \geq u^{u/(1-\epsilon)}$ .

We want to choose polynomials that produce many smooth polynomial values across the sieve region. This heuristically requires that the size of polynomial values is small in general. In addition, one can choose an algebraic polynomial  $f(x)$  which has many roots modulo small prime powers. Then the polynomial values are likely to be divisible by small prime powers. This may increase the smoothness probability of polynomial values. In the rest of this section, we recall some methods [9, 13] to estimate and compare the quality of polynomials.

**2.1. Quality of polynomials.** The quality of the chosen polynomials in polynomial selection can be modelled in terms of size and root properties [13].

**2.1.1. Size property.** Let  $(a, b)$  be pairs of relatively prime integers in the sieving region  $\Omega$ . Since  $G$  is a linear polynomial, we may assume that  $\log(|G(a, b)|)$  does not vary much across the sieving region. We thus only consider the size of the nonlinear polynomial  $F$ , which is modelled by the circular logarithmic  $L^2$ -norm (the smaller the better):

$$(2.1) \quad \frac{1}{2} \log \left( s^{-d} \int_0^{2\pi} \int_0^1 F^2(s \cos \theta, \sin \theta) r^{2d+1} dr d\theta \right).$$

For the norm defined in Equation (2.1), one should not only be able to estimate accurately that norm for a given skewness  $s$ , but find the optimal skewness that gives the minimal norm.

For a sextic polynomial, the logarithmic  $L^2$ -norm in Equation (2.1) can be expressed as follows, where  $\tilde{c}_i = c_i s^{i-d/2}$ :

$$(2.2) \quad \frac{1}{2} \log \left( \frac{\pi}{7168} \left( 231 \tilde{c}_0^2 + 42 \tilde{c}_0 \tilde{c}_2 + 14 \tilde{c}_0 \tilde{c}_4 + 10 \tilde{c}_0 \tilde{c}_6 + 21 \tilde{c}_1^2 + 14 \tilde{c}_1 \tilde{c}_3 \right. \right. \\ \left. \left. + 10 \tilde{c}_1 \tilde{c}_5 + 7 \tilde{c}_2^2 + 10 \tilde{c}_2 \tilde{c}_4 + 14 \tilde{c}_2 \tilde{c}_6 + 5 \tilde{c}_3^2 + 14 \tilde{c}_3 \tilde{c}_5 \right. \right. \\ \left. \left. + 7 \tilde{c}_4^2 + 42 \tilde{c}_4 \tilde{c}_6 + 21 \tilde{c}_5^2 + 231 \tilde{c}_6^2 \right) \right).$$

For a degree- $d$  polynomial  $f(x)$ , locating the extrema with respect to  $s$  of the term inside the logarithm simply reduces to finding the roots of a degree- $d$  polynomial, and keeping the root giving the smallest norm. We often say “the skewness of  $f$ ” for this optimal skewness  $s$ .

2.1.2. *Root property.* If a polynomial  $f(x)$  has many roots modulo small prime powers, the polynomial values may behave more smooth than random integers of about the same size. Boender, Brent, Montgomery and Murphy [5, 12, 13, 14] described some quantitative measures of this effect (root property). Let  $n_{p^k}$  be the number of roots of  $f \pmod{p^k}$  for  $k \geq 1$ . The expected  $p$ -valuation  $\nu_p(f)$  is defined to be  $\nu_p(f) = \sum_{k=1}^{\infty} n_{p^k} / (p^{k-1}(p+1))$ . The root property can be quantified by

$$\alpha(F) = \sum_{\substack{p \leq B \\ p \text{ prime}}} \left( \frac{1}{p-1} - \nu_p(F) \right) \log p,$$

which compares the cumulative expected  $p$ -valuation of polynomial values to random integers of similar size. We refer to [2, 4, 13] for more details.

2.1.3. *Combined score.* The logarithmic  $L^2$ -norm in Equation (2.1) can be modified to take the root property into account. Assuming the polynomial value  $F(x, y)$  behaves — for the smoothness — like a random integer around  $F(x, y) e^{\alpha(F)}$ , the combined score is defined by adding  $\alpha(F)$  to the logarithmic  $L^2$ -norm. The combined score is only a rough estimate to compare polynomials. In practice, it is only trustful when the differences between polynomials are large.

2.1.4. *Murphy’s E score.* Murphy’s E score [13] is a (relatively) reliable ranking function to identify the best polynomials without test sieving. Taking the root property into account, the number of sieving reports (coprime pairs that lead to smooth polynomial values) can be approximated by

$$(2.3) \quad \frac{6}{\pi^2} \int_{\Omega} \rho \left( \frac{\log|F(x, y)| + \alpha(F)}{\log B_1} \right) \rho \left( \frac{\log|G(x, y)| + \alpha(G)}{\log B_2} \right) dx dy.$$

For comparison, one can ignore the constant multiplier  $6/\pi^2$ . To approximate the integral, Murphy used a summation over a set of  $K$  sample points  $(x_i, y_i)$ :

$$E(F, G) = \sum_{i=1}^K \rho \left( \frac{\log|F(x_i, y_i)| + \alpha(F)}{\log B_1} \right) \rho \left( \frac{\log|G(x_i, y_i)| + \alpha(G)}{\log B_2} \right).$$

Over a circular region of radius  $r$ , we can sample  $x_i = r \cos \theta_i$  and  $y_i = r \sin \theta_i$ . The angles  $\theta_i$  sample the points on (the boundary of) the circular region. The Dickman-de Bruijn function  $\rho(x)$  does not admit a closed form solution. An asymptotic expansion can be used to approximate its values. Murphy’s E score is a better ranking function and we use it in §3.5 to compare polynomials.

2.2. **Optimizing the quality of polynomials.** Polynomial selection can be divided into three steps: polynomial generation, size optimization and root optimization. In polynomial generation, we generate many raw polynomials whose size is admissible. We further reduce the size of the raw polynomials in size optimization. Many polynomials can have comparable size after size optimization. We produce and choose the best polynomials in terms of root properties in root optimization.

Translation and rotation are useful to optimize the size and root properties. Let  $f(x) = \sum_{i=0}^d c_i x^i$  and  $g(x) = m_2 x - m_1$  where  $m_1/m_2 \pmod{n}$  is the common root.

Translation of  $f(x)$  by  $k$  gives a new polynomial  $f(x+k)$ ; the linear polynomial becomes  $g(x) + km_2$ , and the common root is changed to  $m_1/m_2 - k \pmod{n}$ . Translation does not alter the root properties.

Rotation by a polynomial  $\lambda(x)$  gives a new polynomial  $f(x) + \lambda(x)g(x)$ ; the linear polynomial and the common root are unchanged.  $\lambda(x)$  is often a linear or quadratic polynomial, depending on  $n$  and on the skewness of  $f(x)$ . Rotation can affect both size and root properties.

The contributions of this paper are threefold: (i) we introduce a new kind of polynomials for GNFS; (ii) we propose an algorithm which generates such polynomials, assuming a good translation  $k$  is known; and (iii) we propose an algorithm to find such a good translation.

### 3. SIZE OPTIMIZATION

Polynomial generation (e.g., using Kleinjung's methods [9, 10]) gives many raw polynomials with small leading coefficients. The raw polynomials have very small  $|c_d|, |c_{d-1}|$  and small  $|c_{d-2}|$ . The coefficients  $|c_{d-3}|, \dots, |c_0|$  are comparable to  $m_1 \approx (n/c_d)^{1/d}$ . In size optimization, we want to produce polynomials with smaller logarithmic  $L^2$ -norm (e.g., Equation (2.1)) by changing the skewness, translating and rotating.

Assuming no cancellation occurs, we can approximate  $|F(a, b)| \approx \sum_{i=0}^d |c_i a^i b^{d-i}|$ ; this approximation is maximal at the corner of the sieve region  $a = U\sqrt{s}, b = U/\sqrt{s}$ , where  $|F(a, b)| \approx U^d \sum_{i=0}^d |c_i s^{i-d/2}|$ . For quintic polynomials,  $|c_5|, |c_4|$  and  $|c_3|$  are small when using Kleinjung's algorithm. The next non-controlled coefficient is  $c_2$ . As  $s \geq 1$ , the dominant term is  $|c_2|s^{-1/2}U^5$ , so the contribution of  $c_2$  on the polynomial value is already reduced by a factor of  $s^{-1/2}$ .

For sextic polynomials, the approximate polynomial values are dominated by the term  $|c_3|U^6$  in the regions where no cancellation occurs. Here,  $c_3$  is not controlled in the polynomial generation step, and we do not get a reduction in size like the  $s^{-1/2}$  factor for quintic polynomials. Therefore, it is important to size-optimize sextic polynomials before trying to optimize the root properties.

Sextic polynomials are of main interest since they have been used in record factorizations such as RSA-768 [18] and should be used for future records. Murphy [13] shows that the running time of the number field sieve, to factor an integer  $n$  with a degree- $d$  polynomial, is about

$$\exp\left(\left(1 + o(1)\right)\left(d \log d + \sqrt{(d \log d)^2 + 4 \log(n^{1/(d+1)}) \log \log(n^{1/(d+1)})}\right)\right).$$

With numerical calculations for various  $n$  and  $d$ , this would indicate that sextic polynomials are preferable for numbers between 220 and 360 decimal digits. The two challenge numbers RSA-896 (270 digits) and RSA-1024 (309 digits) are thus suitable for using sextic polynomials.

Let  $f(x)$  be a sextic polynomial. We can use quadratic rotations since  $c_3, \dots, c_0$  have order  $(n/c_6)^{1/6}$ . A quadratic rotation is defined for integers  $u, v, w$  by:

$$(3.1) \quad f_{u,v,w}(x) = f(x) + (ux^2 + vx + w)g(x).$$

**3.1. Some “classical” methods.** We describe here some of the state-of-the-art size-optimization ideas, some of which are not due to the authors of this article, but since they are not published elsewhere, we mention them for completeness.

We want to produce a polynomial with small  $L^2$ -norm by translation and rotation. We focus on degree 6 here for sake of clarity, it is straightforward to generalize to other degrees.

In the raw polynomial,  $c_0, c_1, c_2, c_3$  have similar size and are much larger than  $c_4, c_5, c_6$ . In Equation (2.2),  $\tilde{c}_0, \tilde{c}_1, \tilde{c}_2$  are bounded by  $\tilde{c}_3$ . Therefore, the  $L^2$ -norm can be controlled by terms involving  $\tilde{c}_3, \tilde{c}_4, \tilde{c}_6$  (since  $|c_6| \approx |c_5| \ll |c_4|$ ). Assuming no cancellation occurs in Eq. (2.2), a lower bound, not depending on skewness, is the term  $\tilde{c}_3^2 = c_3^2$ . Hence, a small  $c_3$  is a necessary condition for a small  $L^2$ -norm. The idea is to minimize  $c_3$  by translation. Translation by  $k$  gives a polynomial in  $x$  whose coefficients are functions of  $k$ :

$$f(x+k) = c_6x^6 + (6c_6k + c_5)x^5 + (15c_6k^2 + 5c_5k + c_4)x^4 \\ + (20c_6k^3 + 10c_5k^2 + 4c_4k + c_3)x^3 + \dots$$

Let  $c_i(k)$  be the coefficients of the  $i$ -th term in the translated polynomial.  $c_3(k)$  of  $f(x+k)$  is a cubic polynomial in  $k$ . The coefficients  $c_0(k), c_1(k), c_2(k)$  will increase due to translation. We can use rotation to reduce them, if needed.

**3.1.1. Minimizing  $c_3(k)$ .** The cubic polynomial  $c_3(k)$  has either one or three real roots. (This is not particular to degree 6: the coefficient  $c_{d-3}(k)$  is always cubic in  $k$ .) For each real root  $r$ , we choose  $k_0$  to be either  $\lceil r \rceil$  or  $\lfloor r \rfloor$ , whichever minimizes  $|c_3(k)|$ . We translate  $f(x)$  by  $k_0$ . Then we can further optimize the polynomial by a local descent method.

We give an example with the raw polynomial of  $A_{768}$  from §4.1. The coefficient of  $x^3$  in  $f(x+k)$  is

$$1810534320k^3 + 2410120740k^2 - 1189404920661858930542720k \\ + 7294790451575028477050464058865868764.$$

This cubic polynomial has a real root near  $k_0 = -1591376391$ . We translate  $f(x)$  by  $k_0$ , which reduces the coefficient  $c_3$  from  $7.3 \cdot 10^{36}$  to  $-1.4 \cdot 10^{27}$ . After reducing the coefficients of degree 2, 1, 0 of  $f(x+k_0)$  by rotation, we obtain a logarithmic  $L^2$ -norm of 77.36, compared to 81.82 for  $f(x)$ . We then apply a local optimization method.

This method works better on average than a local optimization used alone. Table 1 (see also Figure 1) shows that on our RSA-768 data set of  $10^5$  polynomials, it reduces the average logarithmic  $L^2$ -norm to 70.34 (with standard deviation 0.60).

**3.2. A new class of GNFS polynomials.** Previous polynomial selection algorithms that generate a nonlinear polynomial  $f(x)$  and a linear one  $g(x)$  all produced polynomials such that  $\text{Res}(f, g) = \pm n$  where  $n$  is the number to factor.

We propose a more general class, such that  $\text{Res}(f, g) = \ell n$ , where  $\ell$  is a small integer, which we call the “multiplier”, as in MPQS polynomial selection [19].

**3.3. An algorithm to find such polynomials.** For sake of simplicity we describe the algorithm for a degree-6 polynomial, but it applies to any degree  $d$ -polynomial, as long as  $d \geq 3$ . Given a raw degree-6 polynomial  $f = c_6x^6 + \dots + c_1x + c_0$ , and a linear polynomial  $g = m_2x - m_1$  — for example found by Kleinjung’s algorithm [9, 10] —, we want to find a polynomial of small  $L^2$ -norm, as defined in Eq. (2.1),

by using either translation or rotation. Choose a positive integer  $s$  (corresponding to the skewness of  $f$ ) and form the following lattice from the coefficients of  $f$  (each column representing a vector):

$$L = \begin{pmatrix} s^6 c_6 & 0 & 0 & 0 & 0 \\ s^5 c_5 & 0 & 0 & 0 & 0 \\ s^4 c_4 & s^4 m_2 & 0 & 0 & 0 \\ s^3 c_3 & -s^3 m_1 & s^3 m_2 & 0 & 0 \\ s^2 c_2 & 0 & -s^2 m_1 & s^2 m_2 & 0 \\ s c_1 & 0 & 0 & -s m_1 & s m_2 \\ c_0 & 0 & 0 & 0 & -m_1 \end{pmatrix}.$$

Then reduce this lattice with the Lenstra-Lenstra-Lovász algorithm (LLL), one obtains short vectors of the form:

$$\begin{pmatrix} s^6(\ell c_6) \\ s^5(\ell c_5) \\ s^4(\ell c_4 + t m_2) \\ s^3(\ell c_3 - t m_1 + u m_2) \\ s^2(\ell c_2 - u m_1 + v m_2) \\ s(\ell c_1 - v m_1 + w m_2) \\ \ell c_0 - w m_1 \end{pmatrix}$$

which correspond to  $\ell f(x) + (tx^3 + ux^2 + vx + w)g(x)$ , after dividing (exactly) the coefficients by  $s^6, s^5, \dots$ . Then one applies some local optimization procedure. If one wants the first vector of  $L$  to be not much larger than the last one in norm, one will take  $s \leq (m_1/c_6)^{1/6}$ .

In practice the polynomial  $\ell f(x) + (tx^3 + ux^2 + vx + w)g(x)$  corresponding to the shortest vector is the same for a large range of skewness  $s$ ; it thus suffices to consider a few values of  $s$ , say  $s = 10^3, 10^4, 10^5, 10^6$  for example.

This LLL-based algorithm finds a good rotation  $\ell f(x) + (\dots + ux^2 + vx + w)g(x)$ ; to find a good translation we apply this algorithm on  $f(x+k), g(x+k)$  for several values of  $k$ , and keep the overall best polynomials (in §3.4, we describe how to find good values of the translation  $k$ ).

Previous work has focused on  $\ell = 1$  only; here we allow a *multiplier*  $\ell > 1$ . If  $\text{Res}(f, g) = \pm n$  where  $n$  is the number to factor, then with a multiplier we get  $\text{Res}(\ell f + \dots, g) = \pm \ell n$ , thus the resultant is increased by a factor  $\ell$ , but nonetheless we might find better polynomials (as demonstrated experimentally below).

It should be noted that one does not try to reduce the coefficient  $c_5$ , since with Kleinjung's algorithm, it is of the same order of magnitude as  $c_6$ .

This method works better on average than the previous methods. Table 1 shows that it reduces the average logarithmic  $L^2$ -norm to 68.42 (with standard deviation 0.72) on our RSA-768 data set of  $10^5$  polynomials.

**3.4. Finding a good translation.** In this section we describe two methods to generate translations that can be used as input for the algorithm described in Section 3.3.

**3.4.1. Minimizing  $c_{d-2}(k)$  and  $c_{d-3}(k)$  simultaneously.** This method is derived from the one described in §3.1.1.

After translation and the algorithm described in §3.3, if only the rotation of highest degree is considered, the new polynomial will look like  $\ell f(x+k) + tx^{d-3}g(x +$

$k$ ), where  $\ell$  and  $t$  are integers. Let  $q = t/\ell$ , the problem reduces to find rational values of  $q$  for which there exists values of  $k$  that make  $\hat{c}_{d-2}(k)$  and  $\hat{c}_{d-3}(k)$  small in  $\hat{f} = f(x+k) + qx^{d-3}g(x+k) =: \sum_{i=0}^d \hat{c}_i(k)x^i$ . As we want to minimize  $\hat{c}_{d-2}(k)$  and  $\hat{c}_{d-3}(k)$  simultaneously, we try to find values of  $q$  for which  $\text{Res}_k(\hat{c}_{d-3}(k), \hat{c}_{d-2}(k))$  has a root or is small (in absolute value).

For degree 6,  $\hat{c}_{d-2}(k)$  and  $\hat{c}_{d-3}(k)$  have the following form, with  $g = m_2x - m_1$ :

$$\begin{aligned} \hat{c}_4(k) &= 15c_6k^2 + 5c_5k + c_4 + qm_2, \\ \hat{c}_3(k) &= 20c_6k^3 + 10c_5k^2 + 4c_4k + qm_2k + c_3 - qm_1. \end{aligned}$$

The formula of  $\text{Res}(\hat{c}_3(k), \hat{c}_4(k))$  shows that it is a polynomial of degree 3 in  $q$ , with integer coefficients. So it has either 1 or 3 real roots.

In the case where the resultant has 3 real roots, each root can be approximated with a rational number, using continued fractions, Farey approximations or trying all rational fractions with bounded denominator. For each of these rational approximations, roots of  $\hat{c}_4(k)$  and  $\hat{c}_3(k)$  can be computed (they are close, by choice of  $q$ ), and can be rounded to obtain values for the translation  $k$ .

In the case where the resultant has 1 real root, experiments show that it is necessary to consider also extrema of the resultant and not only the root. Then, as in the other case, rational approximations of these values of  $q$  are used to find values of  $k$ .

In both cases, the values of  $k$  found provide pairs of translated polynomials  $(f(x+k), g(x+k))$  that can be used as input for the algorithm described in §3.3.

3.4.2. *Translations of the form  $i \times 10^j$ .* In addition to the above direct method, one can try values of  $k$  of the form  $i \cdot 10^j$  for small  $i, j$ . In practice, this helps to locate a few good values of  $k$  which are missed by the above direct algorithm.

3.5. **Experiments.** We examine a data set consisting of  $10^5$  raw sextic polynomials for RSA-768. Those polynomials were generated by CADO-NFS [1] and Msieve [16, 15] using Kleinjung’s algorithm [10]. Table 1 (left) shows the average logarithmic  $L^2$ -norm for the raw and optimized polynomials. Figure 1 shows the normalized discrete density distribution of logarithmic  $L^2$ -norm for the raw and optimized polynomials.

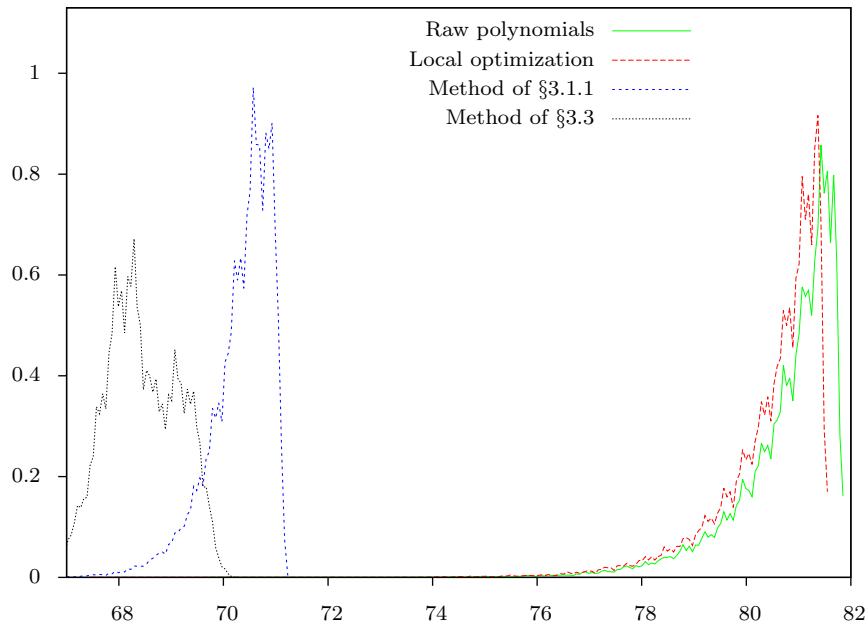
TABLE 1. Comparison of size optimization methods on  $10^5$  raw polynomials for RSA-768 (left) and 5795 raw polynomials for RSA-512 (right). Columns  $\log(L^2)$  and  $\text{std. } \log(L^2)$  record the average logarithmic  $L^2$ -norm of polynomials and its standard deviation.

RSA-768	$\log(L^2)$	$\text{std. } \log(L^2)$	RSA-512	$\log(L^2)$	$\text{std. } \log(L^2)$
Raw polynomials	80.75	1.00	Raw polynomials	53.54	1.96
Method of §3.1.1	70.34	0.60	CADO-NFS 2.1	50.43	1.49
Method of §3.3	68.42	0.72	Method of §3.3	49.36	0.62

We performed a similar experiment with degree-5 raw polynomials produced by CADO-NFS for RSA-512 (see Table 1, right). Although the gain on the average logarithmic  $L^2$ -norm is smaller than for degree 6, it demonstrates the new algorithm is still valuable for degree 5.

In Table 2, we further consider some experiments to compare the final polynomials (after size and root optimization) obtained by the new methods (c.f. §3.3)



FIGURE 1. Distribution of logarithmic  $L^2$ -norm after size optimization.

and the local optimization method. For the RSA-768 data, we optimize the size of the raw polynomials in both ways and then run the root optimization on the size-optimized polynomials. Assuming that Murphy’s E score provides an accurate estimation of the yield rate, we can see that the new method produces polynomials with much higher yield rates on average.

TABLE 2. Comparison of two size optimization methods on  $10^5$  raw polynomials for RSA-768. Columns  $\log(L^2)$  and  $\alpha(F)$  record the average logarithmic  $L^2$ -norm and  $\alpha$ -score of polynomials after size and root optimization; Columns C and E are the average combined score (c.f. §2.1.3) and Murphy’s E score (c.f. §2.1.4) for polynomials after size and root optimization; Columns Top C and Top E are the average combined score and Murphy’s E score for the top 100 polynomials respectively. Here we use  $B_1 = 1.1 \times 10^9$ ,  $B_2 = 2.0 \times 10^8$  and the area  $2.362 \times 10^{18}$  for the domain  $\Omega$  in the computation of Murphy’s E score.

	$\log(L^2)$	$\alpha(F)$	C	Top C	E	Top E
Local optimization	80.65	-7.777	72.87	64.23	3.37e-14	1.53e-13
Method of §3.3	69.90	-6.812	63.08	59.05	1.10e-13	2.53e-13

#### 4. POLYNOMIALS FOR RSA CHALLENGE NUMBERS

4.1. **RSA-768.** Using our algorithm on a data-set kindly provided by Jason Pappadopoulos, we found several polynomials for RSA-768 that are better than the one used for its factorization, see Table 3.

TABLE 3. Better polynomials for RSA-768.

	Polynomial $A_{768}$ : $\log(L^2) = 65.40$ , $E = 4.42\text{e-}13$ .
$m_2$	3653258925429788683931
$m_1$	15447766910976513671275672403785068626
$c_6$	3258961776
$c_5$	288664131841057800
$c_4$	11030506237737074466307
$c_3$	-893188977600857037294644587163
$c_2$	94391058239630467884134336648314151
$c_1$	377093715995883343269663077625960978403307
$c_0$	9045161689950071726629005834738832965305854530
	Raw polynomial of $A_{768}$ : $\log(L^2) = 81.82$ .
$m_2$	3653258925429788683931
$m_1$	15447766964908044471905887663199854537
$c_6$	90526716
$c_5$	241012074
$c_4$	-297351230165464732635680
$c_3$	7294790451575028477050464058865868764
$c_2$	2834529958404715620819873213762675365
$c_1$	2885249650190088598028888005645211453
$c_0$	6518955908807555569064205871887548883
	Polynomial $B_{768}$ : $\log(L^2) = 63.39$ , $E = 4.52\text{e-}13$ .
$m_2$	5924452599136152496277
$m_1$	30571132577927123601048744672398340686
$c_6$	22604400
$c_5$	33946122310442580
$c_4$	74850428174211171973801
$c_3$	-253187194308237186134406064742
$c_2$	-81310927091457333751052543066797504
$c_1$	287725415794347943853989965924714064839458
$c_0$	-8118770924468304419104941835765577203531011465
	Raw polynomial of $B_{768}$ : $\log(L^2) = 82.44$ .
$m_2$	5924452599136152496277
$m_1$	30571134060766694419190738395978436148
$c_6$	1506960
$c_5$	-2418388
$c_4$	-1408315672283641690514244
$c_3$	-14265589093765299499016567124341772705
$c_2$	-9781453412190401648933585496938223891
$c_1$	6869814166916783294989812412963427466
$c_0$	11656834361594150492420981192388245365

The polynomials  $A_{768}$  and  $B_{768}$  correspond to a multiplier  $\ell = 36$  and  $15$  from their respective raw polynomials, thus could not have been found with previously known methods. The original polynomial used for the RSA-768 factorization has logarithmic- $L^2$  norm  $64.08$  and Murphy  $E$  score  $4.28\text{e-}13$ . Both polynomials  $A_{768}$  and  $B_{768}$  have better Murphy  $E$  score.

A sieving test in the special- $q$  range of length  $10^5$  starting at  $3,400,000,000$  shows that the polynomial  $A_{768}$  produces  $7\%$  more relations per special- $q$  than the original RSA-768 polynomial, and  $5\%$  more relations per second. For this test, we use the same binary than the one used for the factorization of RSA-768, with same parameters. A similar sieving test shows that  $B_{768}$  produces  $5\%$  more relations per special- $q$  than the original polynomial, and  $3\%$  more relations per second.

4.2. **RSA-896.** We consider a set of ten raw polynomials for RSA-896, which are deduced from  $c_6, m_2, m_1$  using the algorithm in [9, Lemma 2.1]:

#	$c_6$	$m_2$	$m_1$
1	120	30598679948073727114694567	388409740367611516819003632552473419494959325
2	120	39584252255977153653238969	388409740360914711183938301673437684112903927
3	180	24122614381393211892378929	363029208883450375335947639930852651396354800
4	240	127202957198327749843027	346033739807607082689932107732897888329286672
5	300	2057011251362034806314333	333400907514700794381936235452414684814163915
6	480	5403413584512371865850751	308281015227934361150655851241246900982158569
7	540	2017884993246970143225589	302288315235567244040480840949178516307608795
8	600	55808326686191457067	297026441131666391478870494345190939502152703
9	600	52318705091858802318954527	297026441133748328236580421265680417019345427
10	600	103526916061308104548087973	297026441135225710921799260036162267293068705

We give for each one the logarithmic  $L^2$ -norm of the raw polynomial, the best size-optimized polynomial found with CADO-NFS 2.1, and with our new algorithm:

#	1	2	3	4	5	6	7	8	9	10
raw lognorm	98.28	98.11	96.89	98.00	97.84	98.53	97.18	98.37	96.97	96.63
CADO-NFS 2.1	82.88	82.74	82.30	82.03	82.37	83.33	82.12	79.36	83.79	82.45
new algorithm	80.53	80.16	79.33	79.75	79.78	79.83	80.04	80.72	79.92	79.38

The new algorithm yields a logarithmic  $L^2$ -norm which is smaller by 2.40 on average (79.94 against 82.34), and always smaller, with the exception of #8.

4.3. **RSA-1024.** Let us consider the degree-6 polynomial for RSA-1024 from [20, Appendix A]. This polynomial has logarithmic  $L^2$ -norm 100.02 according to Eq. (2.1). We ran our algorithm to re-optimize this polynomial, and obtained the polynomial  $A_{1024}$  of logarithmic  $L^2$ -norm 94.91: Therefore we can expect polynomial values  $F(a, b)$  to be smaller by a factor about  $\exp(100.02 - 94.91) \approx 166$ .

Moreover, using our implementation in CADO-NFS, we found the polynomial  $B_{1024}$ . According to its Murphy E score of 7.26e-12, against 9.75e-13 for the polynomial from [20], we can expect a relation yield about 7.4 times larger. With same parameters ( $B_1 = B_2 = 10^{11}$ , area  $10^{18}$ ), we get a Murphy E score of 3.56e-09 for the polynomial used to factor RSA-768, which would give a sieving time for RSA-1024 about 490 times larger than for RSA-768 (instead of 1000 times as claimed in [18]). Note this polynomial was found in a few cpu hours only, we expect a much better polynomial with a real search of a few thousands cpu years. This polynomial is also better than the one from [21], which has Murphy E score 6.79e-12.

## 5. CONCLUSION

We introduced a new class of polynomials for GNFS, proposed new algorithms that produce such polynomials of small size, and demonstrated the efficiency of those algorithms on RSA challenge numbers. Those algorithms are implemented in CADO-NFS [1], an open-source implementation of the number field sieve.

## ACKNOWLEDGEMENTS

The authors are grateful to Richard Brent, Steven Galbraith, Jean-Charles Gilbert, Thorsten Kleinjung, Jason Papadopoulos and Emmanuel Thomé for helpful comments and discussions.

The RSA-768 data set provided by Jason Papadopoulos was obtained with the help of several contributors, among which Greg Childers, Alexander Kruppa, Paul

TABLE 4. Two polynomials for RSA-1024.

	Polynomial $A_{1024}$ : $\log(L^2) = 94.91$ .
$m_2$	1
$m_1$	6290428606355899027255723320027391722163288699413
$c_6$	1173597989242921482240
$c_5$	-43608157020293570037272873757855616
$c_4$	691958140341173987625035104743657545537
$c_3$	4505112021612087343709577481323301185973519
$c_2$	-17304452519439643403755585110507512764935257500
$c_1$	-28313100773851304238101962712925551719741165867633
$c_0$	24996329564944807789602917136794373782308959799485325
	Polynomial $B_{1024}$ : $\log(L^2) = 88.51$ , $E = 7.26\text{e-}12$ .
$m_2$	479811439908216194249453
$m_1$	1817512374557883928497495225851470909509236874401
$c_6$	464627366356612000080
$c_5$	256643558339948645996144349012
$c_4$	11295236008100805191037635141341719
$c_3$	814363912742552517860400155495479115704
$c_2$	9575799550890607050943258085937364328751539
$c_1$	-13432074198278226490408460913573383217263160484466
$c_0$	162150861938319067560802383359648011883310816524625024

Leyland, Lionel Debroux, Jayson King, Bruce Dodson, Jeff Gilchrist, Luigi Morelli, Ed Hall and a few other anonymous contributors.

The first author would like to thank the Australian National University, NeSI (New Zealand eScience Infrastructure) and the Centre for eResearch at the University of Auckland for providing funding and computing facilities.

Finally the authors are very grateful to the anonymous referee who asked for a major revision of the paper, in the process of which the new algorithm from §3.3 was discovered.

#### REFERENCES

1. S. Bai, C. Bouvier, A. Filbois, P. Gaudry, L. Imbert, A. Kruppa, F. Morain, E. Thomé, P. Zimmermann. CADO-NFS, an implementation of the number field sieve. Release 2.1, available from <http://cado-nfs.gforge.inria.fr>, 2014.
2. S. Bai, R. P. Brent, E. Thomé. Root optimization of polynomials in the number field sieve. *Mathematics of Computation* (to appear), 2014.
3. S. Bai, E. Thomé, P. Zimmermann. Factorisation of RSA-704 with CADO-NFS. Report, 2012. <http://eprint.iacr.org/2012/369.pdf>.
4. R. Barbulescu and A. Lachand. Some mathematical remarks on the polynomial selection in NFS. Report, 2014. <http://arxiv.org/abs/1403.0184>.
5. H. Boender. *Factoring large integers with the quadratic sieve*. PhD thesis, Leiden University, 1997.
6. J. Buhler, H. Lenstra, and C. Pomerance. Factoring integers with the number field sieve. In Lenstra and Lenstra [11], pages 50–94.
7. A. Granville. Smooth numbers: computational number theory and beyond. In *Proc. MSRI Conf. Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography*. MSRI Publications, Volume 44, 2008.
8. A. Hildebrand and G. Tenenbaum. Integers without large prime factors. *Journal de Théorie des Nombres de Bordeaux*, 5(2):411–484, 1993.
9. T. Kleinjung. On polynomial selection for the general number field sieve. *Mathematics of Computation*, 75(256):2037–2047, 2006.

10. T. Kleinjung. Polynomial selection. In *CADO workshop on integer factorization*, INRIA Nancy, 2008. <http://cado.gforge.inria.fr/workshop/slides/kleinjung.pdf>.
11. A. K. Lenstra and H. W. Lenstra, Jr., editors. *The Development of the Number Field Sieve*, volume 1554 of *Lecture Notes in Mathematics*. Springer, 1993.
12. B. A. Murphy. Modelling the Yield of Number Field Sieve Polynomials. In *Algorithmic Number Theory - ANTS III, LNCS 1443*, pages 137–147, 1998.
13. B. A. Murphy. *Polynomial selection for the number field sieve integer factorisation algorithm*. PhD thesis, The Australian National University, 1999.
14. B. A. Murphy and R. P. Brent. On quadratic polynomials for the number field sieve. In *Proceedings of the CATS '98*, volume 20 of *Australian Computer Science Communications*, pages 199–213. Springer, 1998.
15. J. Papadopoulos. Call for volunteers: RSA768 polynomial selection, 2011. <http://www.mersenneforum.org/showthread.php?t=15540>.
16. J. Papadopoulos. Msieve v1.48, 2011. <http://sourceforge.net/projects/msieve>.
17. J. M. Pollard. The lattice sieve. In Lenstra and Lenstra [11], pages 43–49.
18. T. Kleinjung, K. Aoki, J. Franke, A. K. Lenstra, E. Thomé, J. W. Bos, P. Gaudry, A. Kruppa, P. L. Montgomery, D. A. Osvik, H. J. J. te Riele, A. Timofeev, and P. Zimmermann. Factorization of a 768-bit RSA modulus. In *Proceedings of CRYPTO '10*, volume 6223 of *Lecture Notes in Computer Science*, pages 333–350. Springer, 2010.
19. R. D. Silverman. The Multiple Polynomial Quadratic Sieve. *Mathematics of Computation*, 48(177):329–339, 1987.
20. A. Lenstra, E. Tromer, A. Shamir, W. Kortsmit, B. Dodson, J. Hughes, and P. Leyland. Factoring Estimates for a 1024-Bit RSA Modulus. in *Advances in Cryptology - Asiacrypt, LNCS 2894*, pages 55–74, 2003.
21. T. Kleinjung. Cofactorisation strategies for the number field sieve and an estimate for the sieving step for factoring 1024 bit integers. <http://www.hyperelliptic.org/tanja/SHARCS/talks06/thorsten.pdf>, 2005.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF AUCKLAND, AUCKLAND, NEW ZEALAND.  
*E-mail address:* [shih.bai@gmail.com](mailto:shih.bai@gmail.com)

INRIA NANCY - GRAND EST, VILLERS-LÈS-NANCY, FRANCE.  
*E-mail address:* [cyril.bouvier@inria.fr](mailto:cyril.bouvier@inria.fr)

INRIA NANCY - GRAND EST, VILLERS-LÈS-NANCY, FRANCE.  
*E-mail address:* [alexander.kruppa@inria.fr](mailto:alexander.kruppa@inria.fr)

INRIA NANCY - GRAND EST, VILLERS-LÈS-NANCY, FRANCE.  
*E-mail address:* [paul.zimmermann@inria.fr](mailto:paul.zimmermann@inria.fr)