# Polynomial Selection for Number Field Sieve in Geometric View 

Min Yang ${ }^{1}$, Qingshu Meng ${ }^{2}$, zhangyi Wang ${ }^{2}$, Lina Wang ${ }^{2}$, and Huanguo Zhang ${ }^{2}$<br>${ }^{1}$ International school of software, Wuhan University, Wuhan, China,<br>${ }^{2}$ Computer school, Wuhan University, Wuhan China


#### Abstract

Polynomial selection is the first important step in number field sieve. A good polynomial not only can produce more relations in the sieving step, but also can reduce the matrix size. In this paper, we propose to use geometric view in the polynomial selection. In geometric view, the coefficients' interaction on size and the number of real roots are simultaneously considered in polynomial selection. We get two simple criteria. The first is that the leading coefficient should not be too large or some good polynomials will be omitted. The second is that the coefficient of degree $d-2$ should be negative and it is better if the coefficients of degree $d-1$ and $d-3$ have opposite sign. These criteria tell us where to find them and how to efficiently find them. Using these new criteria, the computation can be reduced while we can get good polynomials. Many experiments on large integers show the effectiveness of our conclusion.


Keywords: cryptography, number field sieve, polynomial optimization

## 1 Introduction

The general number field sieve $[1,2]$ is known as the fastest algorithm for factoring general large integers. It is based on the fact that if $a^{2}=b^{2} \bmod N$ and $a \neq b, \operatorname{gcd}(a-b, N)$ will give a proper factor of $N$ with at least a half chance. The number field sieve starts by choosing two irreducible and coprime polynomials $f(x)$ and $g(x)$ over $Z$, which share a common root $m$ modulo $N$. Let $F(x, y)=y^{d_{1}} f(x / y)$ and $G(x, y)=y^{d_{2}} g(x / y)$ be the homogenized polynomials corresponding to $f(x)$ and $g(x)$ respectively, where $d_{1}$ and $d_{2}$ are the degree of $f(x)$ and $g(x)$ respectively. We want to find many coprime pairs $(a, b) \in Z^{2}$ such that the polynomials values $F(a, b)$ and $G(a, b)$ are simultaneously smooth with respect to some upper bound B and the pair $(a, b)$ is called a relation. An integer is smooth with respect to bound $B$ (or $B$-smooth) if none of its prime factors are larger than $B$. If we find enough number of relations, by finding linear dependency $[3,4]$ we can construct:

$$
\prod_{(a, b) \in S}\left(a-b \alpha_{1}\right)=\beta_{1}^{2}, \text { where } f\left(\alpha_{1}\right)=0, \beta_{1} \in Z\left[\alpha_{1}\right]
$$

$$
\prod_{(a, b) \in S}\left(a-b \alpha_{2}\right)=\beta_{2}^{2}, \text { where } g\left(\alpha_{2}\right)=0, \beta_{2} \in Z\left[\alpha_{2}\right]
$$

As there exist maps such that $\varphi_{1}\left(\alpha_{1}\right)=m \bmod N$ and $\varphi_{2}\left(\alpha_{2}\right)=m \bmod N$, we have $\varphi_{1}\left(\beta_{1}^{2}\right)=\varphi_{2}\left(\beta_{2}^{2}\right)$. We can obtain the square root $\beta_{1}$ and $\beta_{2}$ from $\beta_{1}^{2}$ and $\beta_{2}^{2}$ respectively using method in [5]. If we let $\varphi_{1}\left(\beta_{1}\right)=x$ and $\varphi_{2}\left(\beta_{2}\right)=y$, then $y^{2}=x^{2} \bmod N$, and we have constructed a congruent squares and so may attempt to factor $N$ by computing $\operatorname{gcd}(x-y, N)$.

In order to obtain enough relations, selecting a pair of polynomials $f(x), g(x)$ with high probability of being smooth is very important. A good pair of polynomials not only can decrease sieving time, but also can reduce the expected matrix size[6]. The polynomial selection is now a hot research area. Based on base-m method and with translation and rotation technique[6], non-skewed or skewed polynomial can be constructed, where one polynomial $f(x)$ is nonlinear and the other $g(x)$ is monic and linear. If the linear polynomial is nonmonic, the size of nonlinear polynomial can be greatly reduced[7,1]. The two methods above are called linear method. Montgomery[9] proposed the nonlinear method, where the two polynomials are both nonlinear. Recently several papers[10-12] discuss the nonlinear polynomial construction problem. Most of recently factored large integers[13-15] use Kleinjung's polynomial selection method[8].

In this paper we propose to use the geometric view in polynomial selection. In geometric view, if a nonlinear polynomial is good, its graph should be flat and near the $x$-axis. To be a good polynomial, the polynomial's leading coefficient should not be too large and the coefficient of degree $d-2$ should be negative and it is better if the coefficients of degree $d-1$ and $d-3$ have opposite sign. The first requirement tells where to find good polynomials and the second requirement tells how to find them efficiently. Many experiments on large integers of size from 129 to 210 digits show the effectiveness of our conclusion.

## 2 Elements related to smoothness of a polynomial

An integer is said to be B-smooth if the integer can be factored into factors bounded by B. By Dickman function, given the smooth bound B, the less the integer is, the more likely the integer is B-smooth. In number field sieve, we need the homogenous form $F(x, y)=a_{d} x^{d}+\cdots+a_{1} x y^{d-1}+a_{0} y^{d}$ of the polynomial $f(x)=a_{d} x^{d}+\cdots+a_{1} x+a_{0}$ to be small. In [6], the size and root property are used to describe the quantity. By size we refer to the magnitude of the values taken by $F(x, y)$. By root property we refer to the distribution of the roots of $F(x, y)$ modulo small $p^{k}$, for p prime and $k \geq 1$. If $F(x, y)$ has many roots modulo small $p^{k}$, values taken by $F(x, y)$ "behave" as if they are smaller than they actually are. That is, on average, the likelihood of $F(x, y)$ values being smooth is increased. It has always been well understood that size affects the yield of $F(x, y)$. In [16], the number of real roots, the order of Galois group of $f g$ were taken into account. By the number of real roots, if $a / b$ is near a real root, the value $F(a, b)$ will be small and will be smooth with high chance. By
the order of Galois group of $f g$, it is better to chose polynomial for which the order of Galois group of $f g$ are small, because they provide more free relations.

If the coefficients of $f(x)$ are small, $F(x, y)$ would have good size property. In order to obtain polynomial with small coefficients, we can search extensively, or let the linear polynomial be nonmonic as suggested in $[1,7]$. However, the interaction of coefficients on size is not fully or directly considered. In order to obtain good root property, we can increase the projective roots by requiring that the leading coefficient contains many small prime as its factors[6]. The paper[17, 19] used the translation and rotation technique to improve the root property. The methods in paper[18, 19] are implemented in CADO-NFS, and are used to factor RSA704[14]. As for the number of the real roots, it is left as random.

In this paper, we will take the interaction of coefficients on size and the number of real roots into consideration to select good polynomials.

## 3 The geometric view on polynomial selection

In this section, we will study the polynomial selection in geometric view. First we give some basics on the graphs of pow functions.

### 3.1 The graph of function $a x^{b}$

A function $f(x)=a x^{b}$ is called power function. The parameter $a$ serves as a simple scaling factor, moving the values of $x^{b}$ up or down as $a$ increases or decreases, respectively and the parameter $b$, called either the exponent or the power, determines the function's rates of growth or decay. Depending on whether it is positive or negative, a whole number or a fraction, $b$ will also determine the function's overall shape and behavior.

More so than other simple families like lines, exponentials, and logs, members of the power family can exhibit many distinctive behaviors. For example, when $\mathrm{b}=0$, the function simplifies to $f(x)=a$, a constant function with an output of $a$ for every input. When $b>0, f(0)=a 0^{b}=0$. That is, every power function with a positive exponent passes through $(0,0)$. When $b<0, f(0)$ is undefined. However, we mainly focus on functions of type $a x^{b}$, where $a$ is an integer and $b$ is a positive integer.

If $b$ is an even positive integer like $\mathrm{b}=2,4,6$, etc., then for any input $x$ we will have $f(-x)=a(-x)^{b}=a(x)^{b}=f(x)$. The function has a certain symmetry: its outputs for any $x$ are exactly the same as its outputs for $-x$. Any function with this behavior is called an even function, with even powers serving as the archetype.

If $b$ is an odd positive integer like $\mathrm{b}=1,3,5$, etc., then for any input $x$ we will have $f(-x)=a(-x)^{b}=a(-1)(x)^{b}=-f(x)$. The function has a certain anti-symmetry: its outputs for any $x$ are exactly the opposite of its outputs for $-x$. Any function with this behavior is called an odd function, with odd powers serving as the archetype.

In short, as shown in Fig. 1, when $a$ is positive and $b$ is even, the graph of $f(x)=a x^{b}$ is similar to the graph of $f(x)=x^{2}$. When $a$ is positive and $b$ is odd, the graph is similar to the graph of $f(x)=x^{3}$. When $a<0$ and $b$ is even, the graph of $f(x)=a x^{b}$ is similar to the graph of $f(x)=-x^{2}$. When $a<0$ and $b$ is odd, the graph of $f(x)=a x^{b}$ is similar to the graph of $f(x)=-x^{3}$.


Fig. 1. The graph of $y=x^{3}($ blue $), y=-x^{3}$ (orange), $y=x^{2}$ (red), $y=-x^{2}$ (green)

Now we consider functions of form $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+\cdots+a_{1} x+a_{0}$ step by step, where $a_{i} \in Z$ for $i=0,1, \cdots, d$ and $d$ is a fixed positive integer. First consider a function of form $f(x)=a_{d} x^{d}$. Obviously as $\left|a_{d}\right|$ gets bigger, the value $|f(x)|=\left|a_{d} x^{d}\right|$ will get bigger or in geometric view the graph of $f(x)$ will get steeper.

Secondly consider functions of form $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}$. If $a_{d} x^{d}$ is symmetric then $a_{d-1} x^{d-1}$ will be anti-symmetric or if $a_{d} x^{d}$ is anti-symmetric then $a_{d-1} x^{d-1}$ will be symmetric. The graph of $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}$ in one side of $y$-axis becomes steeper while on the other side of $y$-axis the graph becomes flatter and nearer the $x$-axis than the graph of $f(x)=a_{d} x^{d}$. The item $a_{d-1} x^{d-1}$ may make the function $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}$ have one more real root.

Thirdly consider function of form $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}$. If the sign of $a_{d}$ is the same as the sign of $a_{d-2}$, the graph of $f(x)=a_{d} x^{d}+$ $a_{d-1} x^{d-1}+a_{d-2} x^{d-2}$ will become steeper than the graph of $f(x)=a_{d} x^{d}+$ $a_{d-1} x^{d-1}$. If the sign of $a_{d}$ is opposite to the sign of $a_{d-2}$, the graph of $f(x)=$ $a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}$ will become flatter and nearer the $x$-axis than the graph of $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}$. The item $a_{d-2} x^{d-2}$ may make $f(x)=$ $a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}$ have one or two more real roots.

Next, consider function of form $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}+$ $a_{d-3} x^{d-3}$. If the sign of $a_{d-1}$ is the same as the sign of $a_{d-3}$, the graph of $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}+a_{d-3} x^{d-3}$ will become steeper than the graph of $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}$. If the sign of $a_{d-1}$ is opposite to the sign of $a_{d-3}$, the graph of $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+a_{d-2} x^{d-2}+a_{d-3} x^{d-3}$ will become flatter and nearer the $x$-axis than the graph of $f(x)=a_{d} x^{d}+$ $a_{d-1} x^{d-1}+a_{d-2} x^{d-2}$. The item $a_{d-3} x^{d-3}$ may make $f(x)=a_{d} x^{d}+a_{d-1} x^{d-1}+$ $a_{d-2} x^{d-2}+a_{d-3} x^{d-3}$ have one or two more real roots. The case in other degree can be similarly discussed.

### 3.2 Requirements on polynomial coefficients in geometric view

In Kleinjung's method[7], one polynomial,say $g(x)$, is linear and $F(a, b)$ is much larger than $G(a, b)$, therefore selecting the nonlinear function $f(x)$ is the focus. As said in [16], first the maximal value of $F(a, b)$ should be small, making them more likely to be smooth and Secondly, when a polynomial has many real roots, more ratios $a / b$ will be near a real root and more values $F(a, b)$ are expected to be small. To obtain the two objectives above, the graph of function $f(x)$ should be flat and near the $x$-axis. To make the graph flat, the coefficients of higher degree, especially the leading coefficient, should be small. To make the graph near the $x$-axis or have more real roots, the coefficients $a_{d}$ and $a_{d-2}$ should have opposite sign and it is better if $a_{d-1}$ and $a_{d-2}$ also have opposite sign.

Remark 1:The requirement on leading coefficient just means that the chance for a polynomial being good gets less as the leading coefficient gets bigger. Therefore, if we have much computation capability we can search with a smaller leading coefficient increment instead of searching in a larger range.

Remark 2: To reduce the computation in evolutionary selection of polynomial, paper[20] in algebraic view discussed the coefficients conditions for a function to have more real roots and try to build the correlation between good polynomials and their coefficients. Stimulated by its idea to increase the number of real roots, we also aim to increase real roots, but in the geometric view. We not only can increase the number of real roots, but also we improve the size property. Further, we build the correlation between good polynomials and their individual coefficient in a simple way.

Table 1. Comparison of Murphy E for polynomial of degree 5 between two different increments

| integer | rsa129 | rsa130 | rsa140 | rsa150 | c151 | rsa155 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| increment $=30030$ | $8.01 \mathrm{e}-11$ | $6.46 \mathrm{e}-11$ | $1.84 \mathrm{e}-11$ | $4.59 \mathrm{e}-12$ | $3.99 \mathrm{e}-12$ | $2.36 \mathrm{e}-12$ |
| increment $=210$ | $8.80 \mathrm{e}-11$ | $7.89 \mathrm{e}-11$ | $1.90 \mathrm{e}-11$ | $5.88 \mathrm{e}-12$ | $4.35 \mathrm{e}-12$ | $2.70 \mathrm{e}-12$ |

Table 2. Comparison of Murphy E for polynomial of degree 6 between two different increments

| integer | b2042 | b204 (3) b2044 | b2045 | b2046 |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| increment $=720720$ | $2.23 \mathrm{e}-15$ | $2.10 \mathrm{e}-15$ | $2.04 \mathrm{e}-15$ | $2.20 \mathrm{e}-15$ | $1.89 \mathrm{e}-15$ |
| increment $=60$ | $2.64 \mathrm{e}-15$ | $2.45 \mathrm{e}-15$ | $2.27 \mathrm{e}-15$ | $2.44 \mathrm{e}-15$ | $2.23 \mathrm{e}-15$ |

### 3.3 Experiments

Based on the criteria above and the polynomial selection program polyselect2l.c of CADO-NFS project, some modifications are made to the polynomial selection program. With the modified polyselect2l.c, we make three kinds of experiments. The notation for larger integers is the same as in [21].

First kind of experiments check the leading coefficient's effect on Murphy E value. Table 1 lists the comparison of Murphy E value for polynomials of degree 5 between the two leading coefficient increments 30030 and 210 for large integers of size from 129 to 155 digits. The polynomials for increment 30030 are given in Appendix A. 1 and the polynomials for increment 210 are given in Appendix A.2. In these experiments all these results can be obtained by running program polyselect2l in CADO-NFS version f78e49c. When increment $=30030$, set admax $=1 \mathrm{e} 9$ and when increment $=210$, set admax $=7 \mathrm{e} 6$. The number of different $a_{d}$ in the two increment cases are about equal. Table 2 lists the comparison of Murphy E value for polynomials of degree 6 between increment 720720 and increment 60 for some integers of size 204 digits. These integers are modified from integer B204, with leading coefficient replaced by $2,3,4,5$ and 6 . The version of polyselect2l.c is 039 f 906 and $\operatorname{admax}=5 \mathrm{e} 6$ when increment $=60$ and $\mathrm{admax}=6 \mathrm{e} 10$ when increment $=720270$. From these two tables, we can see that we stand more chances to select good polynomials for a smaller increment. In other words, as stated in Remark 1, if we have enough computation capability, we should search with smaller increment instead of searching in a larger range. The polynomials for increment 720720 are given in Appendix A. 3 and the polynomials for increment 60 are given in Appendix A.4. In addition, in the experiments we notice that a smaller leading coefficient costs less time than a larger leading coefficient.

The second kind of experiments check effect on running time caused by limiting the sign of coefficient of degree $d-2$. Table 3 lists the comparison for time between the two programs: one is the original program of CADO-NFS version $039 f 906$ and the other is a modified program that checks polynomials with $a_{d-2}$ negative. In two programs the parameter $\operatorname{admax}=1 \mathrm{e} 8$ and the other parameters are not changed. The result in Table 3 is not as we expect. Initially we expect that at least we can save about $1 / 2$ time because we estimate that about $1 / 2$ of $a_{d-2}$ will be negative. Later we find that in CADO-NFS only polynomial with lognorm below a threshold can be considered. That is to say, the limitation on sign of $a_{d-2}$ is similar to the limitation on lognorm or most of polynomials with positive $a_{d-2}$ can't pass the norm threshold. Table 4 lists the running time for a new threshold, 2 bigger than the initial threshold. From Table 4, more polyno-
mials with positive $a_{d-2}$ now pass the threshold and our modified program can save more time. How to set a exact threshold for lognorm is not trivial especially for large integers. If the threshold is relatively small, there exists risk that we can not find polynomial. Maybe it is a good choice to use the both limitations.

The third kind of experiments try to select polynomial with good Murphy E value for integers of size from 160 to 190 digits. The leading coefficients are multiple of 60 . The number of different $a_{d}$ we try is about equal to that used in CADO-NFS project. For example, for c160, the maximal $a_{d}$ is 1 e 9 and the increment is 30030 in CADO-NFS. In our modified program, the maximal $a_{d}$ is 2 e 6 and the increment is 60 . The selected polynomials are given in Appendix B.1. In these polynomials, their MurphyE scores are bigger than these of polynomials used in real factorization. Table 5 compares the Murphy E values.

We list the pair of polynomials used in factoring RSA210 as the last example, which was factored on September 26, 2013 by Ryan Propper[22]. The softwares he used are Msieve and GGNFS. The pair of polynomials are as follows.
y1: 63190692009226810471
y0: -8311128239923121259046301811046853
c6: 744120
c5: 44263602924186
c4: -1333072472407237353592
c3: -35317070927593920606305065701
c2: 415031002380786834672968277117654072
c1: 4926444336634688706035599320492329943566740
c0: -46373978032319633360321876974395396247530766893600
skew 21829368.04 , size $3.501 \mathrm{e}-15$, alpha -11.183 , combined $=1.204 \mathrm{e}-15$ roots $=6$

We mainly analyze the nonlinear polynomial. First its leading coefficient $c_{6}$ is relatively small. Secondly see the signs of its coefficients: one group $c_{4}$ negative, $c_{2}$ positive and $c_{0}$ negative and another group $c_{5}$ positive, $c_{3}$ negative and $c_{1}$ positive. Its graph must be flat and near the $x$-axis. In fact, it has 6 real roots. In geometric view, it is very ideal. Of course, this is very ideal situation. We do not mean to select polynomial with such strict sign limitation, or we may find no polynomial.

Table 3. Running time comparison with initial norm

| integer | rsa129 rsa130 rsa140 rsa150 c151 |  | rsa155 |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Cado-nfs 1016s | 2155 s | 1680 s | 2916 s | 3124 s | 3111 s |
| modified 834s | 1778 s | 1396 s | 3021 s | 2892 | 2990 |

Table 4. Running time comparison with initial norm plus 2

| integer | rsa129 rsa130 rsa140 rsa150 | c151 | rsa155 |
| :--- | :--- | :--- | :--- | :--- |
| Cado-nfs 16054s 22412s 9388 s | 5406 s | 5245 s 3784 s |  |
| modified 12461s 16618s 6234 s | 4367 s | 4246 s 3488 s |  |

Table 5. Murphy E of selected polynomial

| integer | c160 | c164 | rsa170 | c172 | c177 | rsa180 | c186 | rsa190 |
| :--- | :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| fact. poly. $1.08 \mathrm{e}-12$ | $7.00 \mathrm{e}-13$ | $2.27 \mathrm{e}-13$ | $2.85 \mathrm{e}-13$ | $1.11 \mathrm{e}-13$ | $7.22 \mathrm{e}-14$ | $3.11 \mathrm{e}-14$ | $1.55 \mathrm{e}-14$ |  |
| our poly. | $1.34 \mathrm{e}-12$ | $7.38 \mathrm{e}-13$ | $2.92 \mathrm{e}-13$ | $2.94 \mathrm{e}-13$ | $1.19 \mathrm{e}-13$ | $7.90 \mathrm{e}-14$ | $3.31 \mathrm{e}-14$ | $1.96 \mathrm{e}-14$ |

## 4 conclusion

Selecting a good polynomial is very important in number field sieve. A good polynomial not only can produce more relations, but also can reduce the matrix size. In this paper, we propose to use geometric view to select polynomial. In geometric view, the interaction of coefficients on size property and the number of real root are combined gracefully to select polynomials. To obtain the two properties simultaneously, it is required first that the leading coefficient should not be too large. Even given much computation power, we should search with a smaller leading coefficient increment instead of searching in a larger range. Secondly, it is required that the coefficient of degree $d-2$ should be negative and it is better if $a_{n-1}$ and $a_{n-3}$ have opposite sign. Using these criteria, the computation is reduced while we can get good polynomials. Many experiments on large integers of size from 129 to 210 digits show the effectiveness of our conclusion.

The criteria above also apply to polynomials generated by the nonlinear $\operatorname{method}[10-12]$ or by the base-m method. We hope this work not only can efficiently select good polynomials but also can lead to a new efficient sieving algorithm.

Acknowledge:This work is supported in part by the National Natural Science Foundation of China(No.61003267, No.61202385, No.21202386). The second author would thank scientists from CADO-NFS project for their help in installing CADO-NFS and thank Dr. Bai for much useful discussions on polynomial selection. We also thank Qi Shen and Mengyu Liu for their help in writing some programs. We like to thank Public Opinion Center, Internationals Software School, Wuhan University and High Performance Computing Center, Computer School, Wuhan University for allowing us to use their computers.

## References

1. J. P. Buhler, H.W. Lenstra, JR., C. Pomerance, Factoring Integers With The Number Field Sieve, in A. K. Lenstra and H. W. Lenstra, Jr. (eds.), The Development of the Number Field Sieve, LNCS 1554, 50-94, 1993.
2. C. Pomerance, The Number Field Sieve, Proceedings of Symposia in Applied Mathematics, Vol.48, 465-480, 1994.
3. P. L. Montgomery, A Block Lanczos Algorithm for Finding Dependencies over GF(2), Eurocrypt'95, LNCS921, 106-120, 1995.
4. D. Coppersmith, Solving Homogeneous Linear Equations over GF(2) via Block Wiedemann Algorithm, Mathematics of Computation. 62, 333-350, 1994.
5. P. Nguyen, A Montgomery-like Square Root for the Number Field Sieve, Proceedings ANTS III, Springer-Verlag, LNCS 1423,151-68, 1998.
6. B. Murphy, Polynomial Selection for the Number Field Sieve Integer Factorisation Algorithm, Ph.D. thesis, The Australian National University, 1999.
7. T. Kleinjung, On Polynomial Selection For The General Number Field Sieve, Mathematics of Computation, Vol.75, No.256, 2037-2047, 2006.
8. T. Kleinjung, Polynomial selection, CADO workshop on integer factorization, 2008 9. P. L. Montgomery, Small Geometric Progressions Modulo n, manuscript (1995).
9. N. Koo, G.H. Jo, and S. Kwon, On Nonlinear Polynomial Selection and Geometric Progression $(\bmod N)$ for Number Field Sieve. https://eprint.iacr.org/2011/292.pdf
10. T. Prest, P. Zimmermann, Non-linear Polynomial Selection For The Number Field Sieve, Journal of Symbolic Computation, Vol.47, Issue 4, 401-409, 2012.
11. R. S. Williams, Cubic Polynomials in the Number Field Sieve, Master Thesis, Texas Tech University, 2010.
12. S. A. Danilov, I. A. Popovyan, Factorization of RSA-180, http://eprint.iacr.org/2010/270, 2010.
13. S. Bai, E. Thomse,P. Zimmermann, Factorizaiton of RSA-704 With CADO-NFS, http://eprint.iacr.org/2012/ 369, 2012.
14. T. Kleinjung, K. Aoki, J. Franke, et al, Factorization of a 768 -bit RSA modulus, CRYPTO'2010, Proceedings of the 30th annual conference on Advances in cryptology, 333-350,2010.
15. M. Elkenbracht-Huizing,An Implementation of the Number Field Sieve, Experimental Mathematics, Vol.5, No.3,231-251, 1996.
16. J.E. Gower,Rotations and Translations of Number Field Sieve Polynomials, Advances in Cryptology - ASIACRYPT 2003, LNCS2894, 302-310, 2003.
17. S. BAI,P. ZIMMERMANN, Size optimization of sextic polynomials in the number field sieve, http://maths-people.anu.edu.au/ bai/paper/sopt.pdf, 2012.
18. S. Bai, P. Brent, E. Thom, Root Optimization of Polynomials in the Number Field Sieve. http://eprint.iacr.org/2012/691, 2012.
19. Min Yang, Qingshu Meng, Zhangyi Wang, Li Li and Huanguo Zhang, On the coefficients of polynomial for number field sieve, http://eprint.iacr.org/2012/599.
20. http://maths-people.anu.edu.au/ bai/proj_E/
21. http://www.mersenneforum.org/showpost.php?p=354259

Appendix A.1: increment $=30030$ degree 5
c155
Y1: 11030979662144087
Y0: -118752915119517983657298384252
c5: 463302840

```
c4: -258352339064918
c3: 52529491812114301385
c2: 10706503798385550481283931
c1:-6872781839405756465476553137817
c0: -280553627695654846203247116060327085
# lognorm: 50.51, alpha: -7.52 (proj:-2.93), E: 42.99, nr: 3
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=2.36e-12
c151
Y1: 20358554612189839
Y0: -65424163360594437392360151980
c5: 3333330
c4: 2644591054915
c3: -20584552375993406802
c2: -21548343106118897914782276
c1: 11880805555204562583548483612768
c0:-800680230816519336716942744543783391
# lognorm: 49.07, alpha: -7.30 (proj:-1.94), E: 41.78, nr: 3
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=3.99e-12
c150
Y1: 16367832754391311
Y0: -15698922624130612481659780810
c5: 162642480
c4: 13751144613652
c3:-4217181371305427128
c2: -1021896268894194553512939
c1: 103172737433803327255675908654
c0:4024027982886495223075850745424093
# lognorm: 47.52, alpha: -6.37 (proj:-2.67), E: 41.15, nr: 1
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=4.59e-12
rsa140
Y1: 403415446853179
Y0:-157273349499220838438089404
c5: 221261040
c4: 11066485179066
c3: 258849189711230117
c2: -31047478090327840317339
c1: -102320911861190912728065405
c0: 7599594854795741192115028543225
# lognorm: 44.45, alpha: -6.50 (proj:-2.53), E: 37.95, nr: 3
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=1.84e-11
rsa130
Y1: 182993839904311
Y0: -1285110329900956329023995
c5: 515555040
c4: 6434889646864
```

c3: 33304708597634465
c2: -1403831336555872639610
c1: -1376944738189404658920336
c0: 4229924895115326513314489952
\# lognorm: 41.63, alpha: -6.53 (proj: -2.37), E: 35.09, nr: 3
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=6.46 \mathrm{e}-11$
rsa129
Y1: 57257978234827
Y0: -830992832442063416259303
c5: 288648360
c4: 7536745038561
c3: 81677458845202574
c2: -268081393893009639776
c1: -481327486952274874159712
c0: 38106114457527261962384480
\# lognorm: 40.58, alpha: -6.18 (proj: -2.37), E: 34.41, nr: 3
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=8.01 \mathrm{e}-11$

Appendix A.2:increment $=210$ degree 5
rsa155
Y1: 9306424547956003
Y0: -277190788480824063205171934974
c5: 6686400
c4: 6352633633700
c3: -3820200844339878773
c2: -12326474664994090812936777
c1: 419184484885735075874876623981
c0: 346812270639046078874549152718952285
\# lognorm: 48.40, alpha: -6.62 (proj: -2.18), E: 41.78, nr: 3
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=2.70 \mathrm{e}-12$
c151
Y1: 15943185982095047
Y0: -56790705716553062623117113429
c5: 6763680
c4: -4083840026166
c3: -14076746741596776799
c2: 1180102466173311242073200
c1: 1348212520143138969285289472428
c0: -53991082164777620240141253426992000
\# lognorm: 47.65, alpha: -6.78 (proj: -2.65), E: 40.88, nr: 5
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=4.35 \mathrm{e}-12$
rsa150
Y1: 899944572595289
Y0: -31271410389402649001843385752
c5: 5186160

```
c4: 5991728481660
c3: -1864703328174151400
c2: -286043673142010871648041
c1: 32395492235626956955264662
c0: 1506928827528734116655043221413395
# lognorm: 45.69, alpha: -6.17 (proj:-2.59), E: 39.52, nr: 3
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=5.88e-12
rsa140
Y1: 502577612448299
Y0:-475379178461037163550246649
c5: }87696
c4: 233892870956
c3: -221472147227053083
c2: -52153228709831212635694
c1: 5740550575780111315087839000
c0: -324977865080140427558896805892160
# lognorm: 43.95, alpha: -6.04 (proj:-2.07), E: 37.92, nr: 3
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=1.90e-11
rsa130
Y1: 6990882746809
Y0: -8726936313207531794957567
c5: 35700
c4: 64865488665
c3: -649412720971193
c2: -385591735192976648155
c1: 1055306508377476348824558
c0:-31630927963993174500497636760
# lognorm: 39.20, alpha: -5.61 (proj:-1.78), E: 33.60, nr: 3
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=7.89e-11
rsa129
Y1: 36811966796639
Y0: -2254673284011188300829143
c5:1963080
c4: 47737862216
c3: 333458029005527
c2: -242513555078196593286
c1: -1333846118342241679499322
c0: 28050040835137262443866649740
# lognorm: 39.36, alpha: -5.66 (proj:-1.85), E: 33.70, nr: }
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=8.80e-11
```

Appendix A.3:increment $=720720$ degree 6 b2046
Y1: 100365043016786149948901
Y0: -152446141482655037833234651302772

```
c6: 51412561200
c5: -4861258016374052
c4: 191012078761937461096
c3: -1232192043383404311741632339
c2: 27709004321074119391693157353917
c1: 28948797772441479943142736971754173511
c0: 275265528231818056175421955577503965001467
skew: 161600.000
# lognorm: 59.99, alpha: -9.29 (proj:-2.61), E: 50.70, nr: }
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=1.89e-15
b2045
Y1: 247508376663051903173565341
Y0: -144973507644730743695648895393852
c6: 15388092720
c5: 13526046743380968
c4: 9513563884835511003137
c3: -1240959825611120737955502794
c2: -176783252129219466230410671017568
c1: 12933374792506729122165034031755816466
c0: 478250604286750546802688573214915787374591
skew: 146496.000
# lognorm: 59.61, alpha: -9.69 (proj:-2.62), E: 49.92, nr: }
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=2.20e-15
b2044
Y1: 203832286460812823157733
Y0: -207643430855180966394033446026057
c6: 5570444880
c5: -15369818232120344
c4: 17658321323570831465722
c3: 2472164032181082014278260093
c2: -404762511073145496406749941587968
c1: 5069568070803749423330338143895755212
c0:-2604276577301591362513273658986211946028960
# lognorm: 60.63, alpha: -10.51 (proj: -2.60), E: 50.12, nr: }
# MurphyE(Bf=10000000,Bg=5000000,area=1.00e+16)=2.04e-15
b204(3)
Y1: 111208684750942131845509987
Y0:-133895241094704417326692345102443
c6: 51416885520
c5: 9593758899129981
c4: 1082213109766727358706
c3: -367942976934008799233604299
c2: -3436814917635922625249015621217
c1: 203332461663848634329446268365214804
c0: 302749346183938951659909505433458592000
```

skew: 43600.000
\# lognorm: 57.74, alpha: -7.87 (proj: -2.14), E: 49.86, nr: 4
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=2.10 \mathrm{e}-15$
b2042
Y1: 869681821260512544679
Y0: -179893642258905389576645535103764
c6: 7236749520
c5: 14780452929794934
c4: 12578621077716746458165
c3: -1919144068361180006990328618
c2: -290630858798447383065048098585444
c1: 23397134164352391863511815508850832820
c0: 1818776668992133670303723778709455121934815
skew: 186176.000
\# lognorm: 60.11, alpha: -9.98 (proj: -2.79), E: 50.13, nr: 4
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=2.23 \mathrm{e}-15$

Appendix A.4:increment $=60$ degree 6
b2046
Y1: 18156144145011285095707
Y0: -742861155652989661960085075960622
c6: 3839040
c5: 26145914834948
c4: 60513495217084917702
c3: -203234720173246118216693829
c2: -50045811644157539138812362067775
c1: 19216738856610694941318138186172709545
c0: 3940072299973891067218232754106787023092497
skew: 782080.000
\# lognorm: 57.24, alpha: -7.61 (proj: -1.87), E: 49.62, nr: 4
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=2.23 \mathrm{e}-15$
b2045
Y1: 12076655207945453272453
Y0: -731216865946632979000918109538380
c6: 3566940
c5: 18812232375757
c4: 42720263138353539735
c3: -41221797911712575605612836
c2: 56848878704995894508162321337842
c1: 10066431688234190371656061436819543439
c0: -10841548965861354284778809035131138546122877
skew: 1006336.000
\# lognorm: 56.97, alpha: -8.16 (proj: -1.60), E: 48.81, nr: 2
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=2.44 \mathrm{e}-15$
b2044

Y1: 1787119143286307974861
Y0: -822608190100630943841284788767471
c6: 1437060
c5: -16119967003227
c4: 79738330640525006398
c3: 138398080617627405860884883
c2: -1486369386598515783150873933358417
c1: -545217385865459461498088907153996069110
c0: 21361339451237601640116678859198403897229600
skew: 2837504.000
\# lognorm: 58.55, alpha: -8.72 (proj: -1.17), E: 49.83, nr: 4
$\#$ MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=2.27 \mathrm{e}-15$
b204(3)
Y1: 451226887729588944826283
Y0: -848137671379965826799665283346099
c6: 936600
c5: -16984933213454
c4: 118986087064806350468
c3: 175752170564215014504698005
c2: -286192447570053542520826764681616
c1: -97084414340941131398057613230552602628
c0: 16416596710061755844708568535318993716043040
skew: 1422848.000
\# lognorm: 57.42, alpha: -8.41 (proj:-1.75), E: 49.01, nr: 4
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=2.45 \mathrm{e}-15$
b2042
Y1: 11936539193519779224973
Y0: -800868476432507880263006447949173
c6: 929280
c5: 18510308248492
c4: 160537251168110084276
c3: -335027720375899084454706675
c2: -357821614071745840990127267439462
c1: 41086387286697258345812460184959289272
c0: 58226725615720541001079746970759874529562560
skew: 1414656.000
\# lognorm: 57.99, alpha: -8.93 (proj: -2.01), E: 49.06, nr: 4
$\#$ MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=2.64 \mathrm{e}-15$

## APPENDIX B.1:

1. c160 Y1: 73343492551732367809

Y0: -7080281526284839070534171504274
c5: 526680
c4: -4030444723623
c3: -75690599220356786580
c2: 40635422193558944207973762
c1: 134098622407687656664586387063740
c0: 61212889900244498099909194446574429925
\# lognorm: 50.34, alpha: -7.17 (proj: -2.14), E: 43.16, nr: 3
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=1.34 \mathrm{e}-12$
2. c164

Y1: 123541940326963319
Y0: -54058513033974447762726882966342
c5: 128040
c4: -2686516299412
c3: -162718967946185600682
c2: 353878132668783479889456851
c1: 5469061015106861847185213406192552
c0: -6321641868496851373903801810309293538405
\# lognorm: 51.62, alpha: -6.92 (proj: -1.80), E: 44.69, nr: 5
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$, area $=1.00 \mathrm{e}+16)=7.38 \mathrm{e}-13$
3. rsa170

Y1: 618129780607871953447
Y0: -612140341462332393113951800416965
c5: 303240
c4: -16152203250443
c3: -1403038549878135914492
c2: 14439674549951367714202780312
c1: 905707366085815228292512942989335448
c0: -1256391901033019635504353361688172509696448
\# lognorm: 54.79, alpha: -7.23 (proj: -1.65), E: 47.56, nr: 5
\# MurphyE $(B f=10000000, B g=5000000$, area $=1.00 \mathrm{e}+16)=2.92 \mathrm{e}-13$
4. c172

Y1: 19515713805704501
Y0: -927258721840612331747139679189960
c5: 1707480
c4: -18352130541094
c3: -1488221951803881353773
c2: 3596074550388701304580450697
c1: 73593987424036765166088264291272877
c0: -32598403276253292635066815473294681202299
\# lognorm: 53.87, alpha: -7.48 (proj:-1.95), E: 46.40, nr: 5
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=2.94 \mathrm{e}-13$
5. c177

Y1: 67838180091015737
Y0: -10457469756684958676952550165914500
c5: 4929960
c4: -232065350284200
c3: -3668625602835038343869
c2: 59697630125848421845412632577
c1: 303023456406617938430549291614227109
c0: -834625261386359862990498681933225641142537
\# lognorm: 55.47, alpha: -6.86 (proj:-1.40), E: 48.61, nr: 5
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=1.19 \mathrm{e}-13$
6. rsa180

Y1: 49695697364496048887
Y0: -40050932905645131903608550770039276
c5: 1854840
c4: -143332330647392
c3: -8077906259480335330908
c2: -904290652357267129805452379751
c1: 4010646110165882699571654871969357900
c0: 9675282348653659427217705542108941551358455
\# lognorm: 57.96, alpha: -7.98 (proj: -2.08), E: 49.98, nr: 3
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=7.90 \mathrm{e}-14$
7. c186

Y1: 24496346005552593263
Y0: -787831582062046095748248502317510752
c5: 2170200
c4: 237135041708501
c3: -11983882555049745532740
c2: -512260184588829034819893656315
c1: 5316022562213358562571538145806360769
c0: 10891317972577135027201874924309077328681515
\# lognorm: 57.11, alpha: -5.96 (proj: -1.21), E: 51.16, nr: 5
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=3.31 \mathrm{e}-14$
8. rsa190

Y1: 156851698102734845483
Y0: -5214293802603225925060700435776690173
c5: 494880
c4: -175438603259948
c3: -110031941428018979808891
c2: 2781302632216237543639936846352
c1: 1944372801606040766389243235660639476364
c0: -31292433641184133044924980963461244101311782080
\# lognorm: 59.83, alpha: -7.15 (proj: -1.68), E: 52.67, nr: 5
\# MurphyE $(\mathrm{Bf}=10000000, \mathrm{Bg}=5000000$,area $=1.00 \mathrm{e}+16)=1.96 \mathrm{e}-14$

