A Unified Framework for Small Secret Exponent Attack on RSA*

Noboru Kunihiro[1], Naoyuki Shinohara[2], and Tetsuya Izu[3]

[1] The University of Tokyo, Japan
kunihiro@k.u-tokyo.ac.jp
[2] NICT, Japan
[3] Fujitsu Labs, Japan

Abstract. We address a lattice based method on small secret exponent attack on RSA scheme. Boneh and Durfee reduced the attack into finding small roots of a bivariate modular equation: $x(N+1+y)+1 \equiv 0(\mod e)$, where N is an RSA moduli and e is the RSA public key. Boneh and Durfee proposed a lattice based algorithm for solving the problem. When the secret exponent d is less than $N^{0.292}$, their method breaks RSA scheme. Since the lattice used in the analysis is not full-rank, the analysis is not easy. Blömer and May gave an alternative algorithm. Although their bound $d \le N^{0.290}$ is worse than Boneh–Durfee result, their method used a full rank lattice. However, the proof for their bound is still complicated. Herrmann and May gave an elementary proof for the Boneh–Durfee's bound: $d \le N^{0.292}$. In this paper, we first give an elementary proof for achieving the bound of Blömer–May: $d \le N^{0.290}$. Our proof employs unravelled linearization technique introduced by Herrmann and May and is rather simpler than Blömer–May's proof. Then, we provide a unified framework to construct a lattice that are used for solving the problem, which includes two previous method: Herrmann–May and Blömer–May methods as a special case. Furthermore, we prove that the bound of Boneh–Durfee: $d \le N^{0.292}$ is still optimal in our unified framework.

Keywords: LLL algorithm, small inverse problem, RSA, lattice-based cryptanalysis

1 Introduction

1.1 Background

RSA cryptosystem is the widely used cryptosystem [12]. Let N be an RSA moduli and d be an RSA secret key. The small secret exponent d is often used to speed up the decryption or signature generation in some cryptographic applications. However, it is well known that RSA scheme is easily broken if secret exponent d is small.

* This is the full version of [9]

In 1990, Wiener [14] showed that RSA scheme is broken by using continued fraction expansion when $d < \frac{1}{3}N^{1/4}$. In 1999, Boneh and Durfee reduced the small secret exponent attack into finding small roots of a bivariate modular equation:

$$x(A + y) \equiv 1 (\text{mod } e)$$

and then proposed two algorithms for solving the problem [2]. They referred to the problem as the small inverse problem. Their algorithms are based on Coppersmith's approach [3–5]. Their first algorithm breaks RSA scheme when $d \leq N^{0.284}$. Then, they presented another algorithm for solving the small inverse problem and improved the bound to $d \leq N^{0.292}$. It employed a non-full rank lattice for improving the bound. Evaluation of a volume of non-full rank lattice was needed in evaluating the bound, which is not so easy task in general. To overcome this difficulty, they introduced a concept of "Geometrically Progressive Matrix" and succeeded to evaluate an upper bound of its volume [2]. However, its proof is rather complicated.

In 2001, Blömer and May proposed another algorithm for solving the small inverse problem [1]. When $d \leq N^{0.290}$, their method solves the small inverse problem. One of good properties is that the lattice used in their method is full rank. However, the analysis for bound is still complicated. In 2010, Herrmann and May [7] presented another algorithm which achieves Boneh–Durfee's improved bound: $d \leq N^{0.292}$. In their proof, they employed unravelled linearization technique introduced in Asiacrypt2009 [6]. As opposed to the Boneh–Durfee's method, their method used a full rank lattice.

1.2 Our Contributions

In this paper, we first give a novel method for achieving the bound of Blömer–May by using unravelled linearization technique, which is also used in the proof of Herrmann–May. We use the same set of shift-polynomials as Blömer–May's and show that our method achieves the same bound as that of Blömer–May: $d \leq N^{0.290}$. Nevertheless, our proof is rather simpler than Blömer–May's original proof. Next, we provide a unified framework which includes two previous methods: Herrmann–May's and Blömer–May's as a special case. Our framework captures well the lattice structure in the previous methods. Then, we derive a condition such that the small inverse problem can be solved in polynomial time and make an optimization in our framework. Since our framework includes Herrmann–May's method, we have a chance to go beyond the Boneh–Durfee's bound:

$d \leq N^{0.292}$. Unfortunately, that does not happen. We prove that the bound $d \leq N^{0.292}$ is still optimal in our framework (Theorem 3). Then, we present a hybrid method which enjoys the both advantages of Herrmann–May's and Blömer–May's methods. Finally, we generalize to the case when the upper bound of solution y is much smaller than $e^{1/2}$. We show that Blömer–May's method can be superior to Boneh–Durfee's method and is optimal in our framework (Theorem 4).

1.3 Organization

Section 2 gives preliminaries and reviews previous known results. In Section 3, we present an elementary proof for Blömer–May's bound: $d \leq N^{0.290}$. In Section 4, we present a unified framework which includes Herrmann–May's proof and Blömer–May's proof as a special case. Then, we show a condition that the problem is solvable in polynomial time. Then, we prove that the Boneh–Durfee's bound: $d \leq N^{0.292}$ is optimal in our framework. In Section 5, we extend to more general situation and discuss its optimal bound in our framework. Section 6 concludes our paper.

2 Preliminaries

First, we briefly recall the LLL algorithm and Howgrave-Graham's lemma. Then, we review the small secret exponent attack on RSA cryptosystem [2] and introduce the "small inverse problem." Then, we explain previous algorithms for solving the small inverse problem.

2.1 The LLL Algorithm and Howgrave-Graham's Lemma

For a vector $\boldsymbol{b}$, $||\boldsymbol{b}||$ denotes the Euclidean norm of $\boldsymbol{b}$. For an n-variate polynomial $h(x_1, \ldots, x_n) = \sum h_{j_1,\ldots,j_n} x_1^{j_1} \cdots x_n^{j_n}$, define the norm of a polynomial as $||h(x_1, \ldots, x_n)|| = \sqrt{\sum h_{j_1,\ldots,j_n}^2}$. That is, $||h(x_1, \ldots, x_n)||$ denotes the Euclidean norm of the vector which consists of coefficients of $h(x_1, \ldots, x_n)$.

Let $B = \{a_{ij}\}$ be a non-singular $w \times w$ square matrix of integers. The rows of B generate a lattice L, a collection of vectors closed under addition and subtraction; in fact the rows forms a basis of L. The lattice L is also represented as follows. Letting $\boldsymbol{a_i} = (a_{i1}, a_{i2}, \ldots, a_{iw})$, the lattice L spanned by $\langle \boldsymbol{a_1}, \ldots, \boldsymbol{a_w} \rangle$ consists of all integral linear combinations of $\boldsymbol{a_1}, \ldots, \boldsymbol{a_w}$, that is:

$$L = \left\{ \sum_{i=1}^{w} n_i \boldsymbol{a_i} | n_i \in \mathbb{Z} \right\}.$$

The volume of full-rank lattice is defined by $\mathrm{vol}(L) = |\det(B)|$.

The LLL algorithm outputs short vectors in a lattice L:

Proposition 1 (LLL [10]). *Let $B = \{a_{ij}\}$ be a non-singular $w \times w$ matrix of integers. The rows of B generate a lattice L. Given B, the LLL algorithm finds vectors $\boldsymbol{b_1}, \boldsymbol{b_2} \in L$ such that*

$$||\boldsymbol{b_1}|| \leq 2^{(w-1)/4}(\mathrm{vol}(L))^{1/w}, ||\boldsymbol{b_2}|| \leq 2^{w/4}(\mathrm{vol}(L))^{1/(w-1)}$$

in time polynomial in $(w, \max\log_2|a_{ij}|)$.

To convert the modular equation into an equation over the integers, we use the following lemma.

Lemma 1 (Howgrave-Graham [8]). *Let $\hat{h}(x, y, z) \in \mathbb{Z}[x, y, z]$ be a polynomial, which is a sum of at most w monomials. Let m be a positive integer and X, Y, Z and ϕ be some positive integers. Suppose that*

1. *$\hat{h}(\bar{x}, \bar{y}, \bar{z}) = 0 \bmod \phi^m$, where $\bar{x}, \bar{y}$ and $\bar{z}$ are integers such that $|\bar{x}| < X, |\bar{y}| < Y, |\bar{z}| < Z$.*
2. *$||\hat{h}(xX, yY, zZ)|| < \phi^m/\sqrt{w}$.*

Then $\hat{h}(\bar{x}, \bar{y}, \bar{z}) = 0$ holds over integers.

2.2 Small Inverse Problem [2]

Let (N, e) be a public key in RSA cryptosystem, where $N = pq$ is the product of two distinct primes. For simplicity, we assume that $\gcd(p-1, q-1) = 2$. A secret key d satisfies that $ed = 1 \bmod (p-1)(q-1)/2$. Hence, there exists an integer k such that $ed+k((N+1)/2-(p+q)/2) = 1$. Writing $s = -(p+q)/2$ and $A = (N+1)/2$, we have $k(A+s) = 1 \pmod e$.

We set $f(x, y) = x(A + y) + 1$. Note that the solution of $f(x, y) \equiv 0 (\mathrm{mod}\ e)$ is $(x, y) = (-k, s)$. If one can solve a bivariate modular equation: $f(x, y) = x(A+y)+1 = 0 \pmod e$, one has k and s and knows the prime factors p and q of N by solving an equation: $v^2 + 2sv + N = 0$. Suppose that the secret key satisfies $d \leq N^\delta$. Further assume that $e \approx N$. To summarize, the secret key will be recovered by finding the solution $(x, y) = (\bar{x}, \bar{y})$ of the equation:

$$f(x, y) = x(A + y) + 1 \equiv 0 \pmod e,$$

where $|\bar{x}| < e^\delta$ and $|\bar{y}| < e^{1/2}$. They referred to this as the *small inverse problem.*

2.3 Known Algorithms for Solving Small Inverse Problem

Boneh and Durfee proposed a lattice-based algorithm for solving the small inverse problem [2]. First, we briefly recall the algorithm though we use different symbols from the original description.

They define the polynomials $g_{[i,j]}(x,y) := x^i f(x,y)^j e^{m-j}$ and $h_{[i,u]}(x,y) := y^i f(x,y)^u e^{m-u}$. The $g_{[i,j]}$ polynomials are referred as x-shifts and the $h_{[i,u]}$ polynomials are referred as y-shifts. Let $\mathcal{F}_{\mathrm{BD}}(m;\tau)$ be a set of shift-polynomials. The set $\mathcal{F}_{\mathrm{BD}}(m;\tau)$ is given by

$$\begin{aligned}\mathcal{G}_{\mathrm{BD}}(m) &:= \{g_{[u-i,i]} | u = 0,\ldots,m; i = 0,\ldots,u\},\\ \mathcal{H}_{\mathrm{BD}}(m;\tau) &:= \{h_{[i,u]} | u = 0,\ldots,m; i = 1,\ldots,\tau m\} \text{ and}\\ \mathcal{F}_{\mathrm{BD}}(m;\tau) &:= \mathcal{G}_{\mathrm{BD}}(m) \cup \mathcal{H}_{\mathrm{BD}}(m;\tau).\end{aligned}$$

They achieved a bound: $d \le N^{0.284}$ using $\mathcal{F}_{\mathrm{BD}}(m;\tau)$. We refer to this method as Boneh–Durfee's weaker method. Then, Boneh and Durfee improved the bound to $d \le N^{0.292}$ by removing y-shift polynomials whose coefficient of leading term exceeds e^m. The resulting lattice is not full rank and computing its volume is not easy. To overcome this difficulty, they introduced a concept of “Geometrically Progressive Matrix” and succeeded to obtain an upper bound of the volume. The analysis for its bound, especially its volume evaluation, is rather complicated.

Blömer and May [1] presented another algorithm. Although the bound: $d \le N^{0.290}$ is worse than Boneh–Durfee's bound, their method has several interesting features. The first is that it requires a smaller lattice dimension for solving the problem. The second is that the involved lattice is full rank and the analysis for the bound is simpler than Boneh–Durfee's. However, the evaluation of bound is still complicated.

Herrmann and May [7] proposed a novel method which achieves the bound: $d \le N^{0.292}$ by employing unravelled linearization technique. We briefly recall Herrmann–May's method. Note that we use different notation from the original description of [7]. First, $f(x,y)$ is transformed into $f(x,y) = x(A+y)+1 = (xy+1)+Ax$. The first step of their method is to perform a linearization of $f(x,y)$ into $\bar{f}(x,z) := z + Ax$ by setting $xy+1 = z$. In a second step of analysis, xy is back-substituted by $xy = z-1$ for each occurrence of xy. They define the polynomials as $\bar{g}_{[i,j]}(x,z) := x^i \bar{f}(x,z)^j e^{m-j}$ and $\bar{h}_{[i,u]}(x,y,z) := y^i \bar{f}(x,z)^u e^{m-u}$. Let τ be an optimization parameter with $0 < \tau \le 1$. Let $\mathcal{F}_{\mathrm{HM}}(m;\tau)$ be a set of shift-polynomials. The $\mathcal{F}_{\mathrm{HM}}(m;\tau)$ is given by

$$\begin{aligned}\mathcal{G}_{\mathrm{HM}}(m) &:= \{\bar{g}_{[u-i,i]} | u = 0,\ldots,m; i = 0,\ldots,u\},\\ \mathcal{H}_{\mathrm{HM}}(m;\tau) &:= \{\bar{h}_{[i,u]} | u = 0,\ldots,m; i = 1,\ldots,\tau u\} \text{ and}\\ \mathcal{F}_{\mathrm{HM}}(m;\tau) &:= \mathcal{G}_{\mathrm{HM}}(m) \cup \mathcal{H}_{\mathrm{HM}}(m;\tau).\end{aligned}$$

They achieved the bound: $d \leq N^{0.292}$ using $\mathcal{F}_{\mathrm{HM}}(m;\tau)$. Note that its lattice is also full rank.

3 A New Proof for Bound of Blömer–May: $d \leq N^{0.290}$

Blömer and May [1] presented the algorithm which achieves the bound: $d \leq N^{0.290}$. Although this bound is worse than the result of Boneh–Durfee, it has a desirable property. Since it uses full-rank lattice, the analysis for bound is rather easy. On the other hand, Herrmann and May [7] presented the algorithm which achieves $d \leq N^{0.292}$ by using unravelled linearization technique. In this section, we provide a new proof for the bound of Blömer–May: $d \leq N^{0.290}$ by using unravelled linearization technique as like as the proof of Herrmann–May.

3.1 A Set of Shift-Polynomials

First, we transform $f(x,y) = x(A+y)+1$ into $f(x,y) = (xy+1)+Ax$. We define $z = xy+1$ and

$$\bar{f}(x,z) := z + Ax$$

as well as Herrmann and May method [7]. Note that the term xy will be replaced by $xy = z-1$ for each occurrence of xy in the consequent analysis.

We define shift-polynomials as follows. For x-shifts, we define

$$\bar{g}_{[i,k]}(x,z) := x^i \bar{f}(x,z)^k e^{m-k}.$$

Let $\bar{z} = \bar{x}\bar{y}+1$. It is easy to see that $\bar{g}_{[i,k]}(\bar{x},\bar{z}) = 0(\text{mod } e^m)$ for any non-negative integers i and k. The upper bound of $|\bar{z}|$ is given by $XY+1$ and then we define $Z = XY+1$.

For y-shifts, we set

$$\bar{h}_{[i,k]}(x,y,z) := y^i \bar{f}(x,z)^k e^{m-k}.$$

It is easy to see that $\bar{h}_{[i,k]}(\bar{x},\bar{y},\bar{z}) = 0(\text{mod } e^m)$ for any non-negative integers i and k.

Remark 1. From the definition, it holds that $\bar{g}_{[0,u]}(x,z) = \bar{h}_{[0,u]}(x,y,z)$.

Next, we fix a set of indexes for shift-polynomials. Let t be a parameter which is optimized later with $0 \leq t \leq m$. Let $\mathcal{F}_{\mathrm{BM}}(m;t)$ be a set of shift-polynomials. The set $\mathcal{F}_{\mathrm{BM}}(m;t)$ is given by

$$\begin{aligned}
&\mathcal{G}_{\mathrm{BM}}(m;t) := \{\bar{g}_{[u-i,i]} | u = m-t, \ldots, m; i = 0, \ldots, u\}, \\
&\mathcal{H}_{\mathrm{BM}}(m;t) := \{\bar{h}_{[i,u]} | u = m-t, \ldots, m; i = 1, \ldots, t-(m-u)\} \text{ and} \\
&\mathcal{F}_{\mathrm{BM}}(m;t) := \mathcal{G}_{\mathrm{BM}}(m;t) \cup \mathcal{H}_{\mathrm{BM}}(m;t).
\end{aligned}$$

Then, we define a polynomial order $\preceq$ in $\mathcal{F}_{\mathrm{BM}}(m;t)$ as follows:

– $\bar{g}_{[i,j]} \preceq \bar{h}_{[i',u]}$ for any i, j, i', u
– $\bar{g}_{[i,j]} \preceq \bar{g}_{[i',j']}$ if $(i+j < i'+j')$ or $(i+j = i'+j'$ and $j \leq j')$
– $\bar{h}_{[i,u]} \preceq \bar{h}_{[i',u']}$ if $(u < u')$ or $(u = u'$ and $i \leq i')$

We write $a \prec b$ if $a \preceq b$ and $a \neq b$.

Regarding the set $\mathcal{F}_{\mathrm{BM}}(m;t)$ for shift-polynomials and the above polynomial order, we have the following two lemmas.

Lemma 2. *If* $\bar{g}_{[u-j,j]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$ *for* $j \geq 1$, *then* $\bar{g}_{[u-j+1,j-1]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$ *and* $\bar{g}_{[u-j+1,j-1]} \prec \bar{g}_{[u-j,j]}$.

Lemma 3. *If* $\bar{h}_{[j,u]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$, *then* $\bar{h}_{[j-1,u]}$ *and* $\bar{h}_{[j-1,u-1]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$. *Furthermore, it holds that* $\bar{h}_{[j-1,u]} \prec \bar{h}_{[j,u]}$ *and* $\bar{h}_{[j-1,u-1]} \prec \bar{h}_{[j,u]}$.

Proof of Lemma 3 It is clear that $\bar{h}_{[j-1,u]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$. Note that we can $\bar{g}_{[0,u]}$ instead of $\bar{h}_{[0,u]}$ since $\bar{h}_{[0,u]}$ and $\bar{g}_{[0,u]}$ are identical from Remark 1. Since $\bar{h}_{[j,u]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$, it holds that $1 \leq j \leq u+t-m$. Then, $0 \leq j-1 \leq (u-1)+t-m$. Hence, it holds that $\bar{h}_{[j-1,u-1]} \in \mathcal{F}_{\mathrm{BM}}(m;t)$. □

3.2 Expansions of Shift-Polynomials

First, we introduce some definitions.

Definition 1. *We denote by* $\mathcal{S}(f)$ *a set of monomials appearing in expansion of* f.

Note that a monomial $x^i y^j z^k$ with $i, j \geq 1$ never appears in $\mathcal{S}(h_{[i,j]}(x,y,z))$ since we replace xy by $xy = z-1$. Hence, only the terms $x^i z^k$ and $y^j z^k$ appear in the expansion of shift-polynomials.

Definition 2. *We say* $f(x,y,z) \cong g(x,y,z)$ *if* $\mathcal{S}(f) = \mathcal{S}(g)$.

A lattice basis is constructed by using the coefficient vectors of shift-polynomials in $\mathcal{F}_{\mathrm{BM}}(m;t)$ as basis vectors. Note that the coefficient vectors of the shift-polynomials $g_{[u-i,i]}(xX, zZ)$ and $h_{[i,u]}(xX, yY, zZ)$ are written as row vectors. Let $B_{\mathrm{BM}}(m;t)$ be a matrix, where all rows of $B_{\mathrm{BM}}(m;t)$ are the coefficient vectors of shift-polynomials according to the ordering of $\mathcal{F}_{\mathrm{BM}}(m;t)$.

Theorem 1. *Let m and t be integers with $t \leq m$. A lattice basis matrix $B_{\mathrm{BM}}(m;t)$ is triangular for any m and t.*

Before giving a proof, we give three lemmas, whose proofs are given in Appendix A.1.

Lemma 4. *If $0 \leq u \leq m$, $\mathcal{S}(\bar{g}_{[u,0]} - e^m x^u) = \emptyset$.*

Lemma 5. *If $0 \leq u \leq m$ and $1 \leq j \leq u$, $\mathcal{S}(\bar{g}_{[u-j,j]} - e^{m-j}x^{u-j}z^j) = \mathcal{S}(\bar{g}_{[u-j+1,j-1]})$.*

Lemma 6. *If $1 \leq u \leq m$ and $i \geq 1$, $\mathcal{S}(\bar{h}_{[i,u]} - e^{m-u}y^i z^u) \subseteq \mathcal{S}(\bar{h}_{[i-1,u-1]}) \cup \mathcal{S}(\bar{h}_{[i-1,u]})$.*

Proof of Theorem 1 We show that the number of monomials newly appearing in expansion of shift-polynomial is one for any shift-polynomials in $\mathcal{F}_{\mathrm{BM}}(m;t)$. In this proof, we abbreviate $\mathcal{F}_{\mathrm{BM}}(m;t)$ as $\mathcal{F}$. We define $\mathcal{F}^f := \{g \in \mathcal{F} | g \prec f\}$ and $\mathcal{S}(\mathcal{F}^f) := \bigcup_{g\in\mathcal{F}^f} \mathcal{S}(g)$. It is enough for proving Theorem 1 to show that for any polynomial $f \in \mathcal{F}$ there exist a monomial m_f such that

- $\mathcal{S}(f - m_f) \subseteq \mathcal{S}(\mathcal{F}^f)$ and
- $m_f \notin \mathcal{S}(\mathcal{F}^f)$.

From Lemmas 2–3 and 4–6, for any $f \in \mathcal{F}$, there exists m_f such that $\mathcal{S}(f - m_f) \subseteq \mathcal{S}(\mathcal{F}^f)$. We can easily verify that $m_f \notin \mathcal{S}(\mathcal{F}^f)$. Then, the lattice basis matrix is triangular. □

We show an example for $m = 2$. We consider $\bar{g}_{[1,2]}(x, z)$. The expansion of $\bar{g}_{[1,2]}(x, z)$ is given by $x^1(z + Ax)^2 = xz^2 + 2Ax^2z + A^2x^3$. Since $\bar{g}_{[1,2]}(x, z) - xz^2 = 2Ax^2z + A^2x^3$, it holds that $\mathcal{S}(\bar{g}_{[1,2]} - xz^2) = \{x^2z, x^3\}$. On the other hand, since $\bar{g}_{[2,1]} = ex^2(z + Ax) = ex^2z + eAx^3$, it holds that $\mathcal{S}(\bar{g}_{[2,1]}) = \{x^2z, x^3\}$. Then, $\mathcal{S}(\bar{g}_{[1,2]} - xz^2) = \mathcal{S}(\bar{g}_{[2,1]})$ and Lemma 5 holds. We'll show another example. We consider $\bar{h}_{[2,2]}(x, y, z)$. The expansion of $\bar{h}_{[2,2]}(x, y, z)$ is given by $y^2(z+Ax)^2 = y^2z^2 + 2Axy^2z + A^2(xy)^2 = y^2z^2 + 2Ay(z-1)z + A^2(z-1)^2 = y^2z^2 + 2Ayz^2 - 2Ayz + A^2z^2 - 2A^2z + A^2$. Then, we have $\mathcal{S}(\bar{h}_{[2,2]} - y^2z^2) = \{yz^2, yz, z^2, z, 1\}$. On the other hand, since $\bar{h}_{[1,1]} = ey(z + Ax) = eyz + Aexy = eyz + Ae(z - 1) = eyz + Aez - Ae$, we have $\mathcal{S}(\bar{h}_{[1,1]}) = \{yz, z, 1\}$. Furthermore, we have $\bar{h}_{[1,2]} = y(z + Ax)^2 = y(z^2 + 2Axz + A^2x^2) = yz^2 + 2Axyz + A^2x^2y = yz^2 + 2A(z - 1)z + A^2x(z - 1) = yz^2 + 2Az^2 - 2Az + A^2xz - A^2x$. Hence, we have $\mathcal{S}(\bar{h}_{[1,2]}) = \{yz^2, z^2, z, xz, x\}$. Then, it holds that $\mathcal{S}(\bar{h}_{[2,2]} - y^2z^2) = \{yz^2, yz, z^2, z, 1\} \subseteq \mathcal{S}(\bar{h}_{[1,1]}) \cup \mathcal{S}(\bar{h}_{[1,2]}) = \{yz^2, yz, z^2, z, xz, x, 1\}$ and Lemma 6 holds.

3.3 Deriving the Bound of Blömer–May: $d \le N^{0.290}$

A lattice basis is constructed by using coefficient vectors of x-shifts $\bar{g}_{[i,k]}(xX, zZ)$ in $\mathcal{G}_{\mathrm{BM}}(m;t)$ and y-shifts $\bar{h}_{[j,u]}(xX, yY, zZ)$ in $\mathcal{H}_{\mathrm{BM}}(m;t)$. We denote the number of shift-polynomials used in x-shifts and y-shifts by w_x and w_y, respectively. We also denote contributions in x-shifts and y-shifts to lattice volume by $\mathrm{vol}(L_X)$ and $\mathrm{vol}(L_Y)$, respectively. The total number of shift-polynomials w is given by $w = w_x + w_y$ and a lattice volume $\mathrm{vol}(L)$ is given by $\mathrm{vol}(L) = \mathrm{vol}(L_X)\mathrm{vol}(L_Y)$.

First, we derive w_x and $\mathrm{vol}(L_X)$. The lattice dimension w_x is given by $w_x = \sum_{l=m-t}^{m} \sum_{k=0}^{l} 1$. The volume $\mathrm{vol}(L_X)$ is given by

$$\mathrm{vol}(L_X) = \prod_{l=m-t}^{m} \prod_{k=0}^{l} X^{l-k} Z^k e^{m-k} = e^{mw_x} \prod_{l=m-t}^{m} \prod_{k=0}^{l} X^{l-k} (Z/e)^k.$$

Let $\mathrm{vol}(L_X) = e^{mw_x} X^{s_{XX}} (Z/e)^{s_{XZ}}$. Each s_{XX} and s_{XZ} is explicitly given as follows:

$$s_{XX} = \sum_{l=m-t}^{m} \sum_{k=0}^{l} l - k = \frac{m^3 - (m-t)^3}{6} + o(m^3) = \frac{1 - (1-\eta)^3}{6} m^3 + o(m^3)$$

$$s_{XZ} = \sum_{l=m-t}^{m} \sum_{k=0}^{l} k = \frac{m^3 - (m-t)^3}{6} + o(m^3) = \frac{1 - (1-\eta)^3}{6} m^3 + o(m^3),$$

where $\eta := t/m$. Then, we have

$$\mathrm{vol}(L_X) = e^{mw_x} X^{(1-(1-\eta)^3)m^3/6} (Z/e)^{(1-(1-\eta)^3)m^3/6}.$$

Second, we derive w_y and $\mathrm{vol}(L_Y)$. The lattice dimension w_y is given by $w_y = \sum_{l=0}^{t} \sum_{j=1}^{l} 1$. The volume $\mathrm{vol}(L_Y)$ is given by

$$\mathrm{vol}(L_Y) = \prod_{l=0}^{t} \prod_{j=1}^{l} Y^j Z^{l+m-t} e^{m-l-m+t} = e^{mw_y} \prod_{l=0}^{t} \prod_{j=1}^{l} Y^j (Z/e)^{l+m-t}.$$

Let $\mathrm{vol}(L_Y) = e^{mw_y} Y^{s_{YY}} (Z/e)^{s_{YZ}}$. Each s_{YY} and s_{YZ} is explicitly given as follows:

$$s_{YY} = \sum_{l=0}^{t} \sum_{j=1}^{l} j = \sum_{l=0}^{t} \frac{l(l+1)}{2} = \frac{t^3}{6} + o(m^3) = \eta^3 \frac{m^3}{6} + o(m^3)$$

$$s_{YZ} = \sum_{l=0}^{t} \sum_{j=1}^{l} l + (m-t) = \frac{t^3}{3} + (m-t)\frac{t^2}{2} + o(m^3) = \eta^2(3-\eta)\frac{m^3}{6} + o(m^3).$$

Then, we have

$$\mathrm{vol}(L_Y) = e^{mw_y} Y^{\eta^3 m^3/6} (Z/e)^{\eta^2(3-\eta)m^3/6}.$$

Summing up the above discussion, we have

$$\begin{aligned}\mathrm{vol}(L) &= \mathrm{vol}(L_X)\mathrm{vol}(L_Y) \\ &= e^{mw} X^{(1-(1-\eta)^3)m^3/6} Y^{\eta^3 m^3/6} (Z/e)^{\eta m^3/2}.\end{aligned} \tag{1}$$

By combining Proposition 1 and Lemma 1, the condition that the problem can be solved in polynomial time is given by $2^{w/4}\mathrm{vol}(L)^{1/(w-1)} \leq e^m/\sqrt{w}$. By ignoring small terms, we have the condition: $\mathrm{vol}(L) \leq e^{mw}$. From Eq. (1), we have the condition:

$$X^{3-3\eta+\eta^2} Y^{\eta^2} Z^3 \leq e^3. \tag{2}$$

By substituting $Z = XY + 1 \leq 2XY$ and $Y = e^{1/2}$ into Eq. (2) and neglecting small terms which don't depend on e, we have the following inequality about X:

$$X < e^{\frac{3-\eta^2}{2(6-3\eta+\eta^2)}}.$$

The maximum value of the exponent part in the right hand side is given by $(\sqrt{6}-1)/5 \approx 0.290$ when $\eta = 3 - \sqrt{6} \approx 0.55$. This is exactly the same as the bound of Blömer–May [1].

4 A Unified Framework for Solving Small Inverse Problem

As we showed in previous section, the Blömer–May method [1] can be explained by unravelled linearization technique. It is natural to think that Herrmann–May method [7] and Blömer–May method [1] have some kinds of relation. In this section, we will present an explicit relation and an interpolation between two methods. First, we present a unified framework for solving small inverse problem, which includes Herrmann–May method and Blömer–May method as a special case by adequately setting parameters. Then, we show that Boneh–Durfee's improved bound: $d \leq N^{0.292}$ is still optimal in our framework. Finally, we propose a hybrid method by interpolating two methods, which enjoys the both advantages of two methods.

4.1 A Set of Shift-Polynomials

We define $\bar{g}_{[i,k]}(x,z) := x^i \bar{f}(x,z)^k e^{m-k}$ for x-shifts and $\bar{h}_{[i,u]}(x,y,z) := y^i \bar{f}(x,z)^u e^{m-u}$ for y-shifts, respectively. The above are the same shift-polynomials described in Section 3. However, we use a different set of index for shift-polynomials. Let τ and η be parameters which are optimized later with $0 < \tau \leq 1$ and $0 < \eta \leq 1$.

We define sets $\mathcal{G}(m;\eta), \mathcal{H}(m;\tau,\eta)$ and $\mathcal{F}(m;\tau,\eta)$ of shift-polynomials as follows:

$$\mathcal{G}(m;\eta) := \{\bar{g}_{[u-i,i]} | u = \lceil m(1-\eta) \rceil, \ldots, m; i = 0, \ldots, u\}$$
$$\mathcal{H}(m;\tau,\eta) := \{\bar{h}_{[i,u]} | u = \lceil m(1-\eta) \rceil, \ldots, m; i = 1, \ldots, \lceil \tau(u - m(1-\eta)) \rceil\} \text{ and}$$
$$\mathcal{F}(m;\tau,\eta) := \mathcal{G}(m;\eta) \cup \mathcal{H}(m;\tau,\eta)$$

We define a polynomial order $\preceq$ in $\mathcal{F}(m;\tau,\eta)$ as follows:

- $\bar{g}_{[i,j]} \preceq \bar{h}_{[i',u]}$ for any i, j, i', u
- $\bar{g}_{[i,j]} \preceq \bar{g}_{[i',j']}$ if $(i + j < i' + j')$ or $(i + j = i' + j'$ and $j \leq j')$
- $\bar{h}_{[i,u]} \preceq \bar{h}_{[i',u']}$ if $(u < u')$ or $(u = u'$ and $i \leq i')$

Regarding the set $\mathcal{F}(m;\tau,\eta)$ for shift-polynomials and the above polynomial order, we have the following two lemmas.

Lemma 7. *Suppose that* $0 < \tau \leq 1$. *If* $\bar{g}_{[u-j,j]} \in \mathcal{F}(m;\tau,\eta)$ *for* $j \geq 1$, *then* $\bar{g}_{[u-j+1,j-1]} \in \mathcal{F}(m;\tau,\eta)$ *and* $\bar{g}_{[u-j+1,j-1]} \prec \bar{g}_{[u-j,j]}$.

Lemma 8. *Suppose that* $0 < \tau \leq 1$. *If* $\bar{h}_{[j,u]} \in \mathcal{F}(m;\tau,\eta)$, *then* $\bar{h}_{[j-1,u]}$ *and* $\bar{h}_{[j-1,u-1]} \in \mathcal{F}(m;\tau,\eta)$. *Furthermore, it holds that* $\bar{h}_{[j-1,u]} \prec \bar{h}_{[j,u]}$ *and* $\bar{h}_{[j-1,u-1]} \prec \bar{h}_{[j,u]}$.

Proof of Lemma 8 It is clear that $\bar{h}_{[j-1,u]} \in \mathcal{F}(m;\tau,\eta)$. Since $\bar{h}_{[j,u]} \in \mathcal{F}(m;\tau,\eta)$, it holds that $1 \leq j \leq \tau(u - m(1-\eta))$. Then, $0 \leq j - 1 \leq \tau(u - m(1-\eta)) - 1$. Since $\tau \leq 1$ from the setting, it holds that $\tau(u - m(1-\eta)) - 1 \leq \tau(u - m(1-\eta)) - \tau = \tau((u-1) - m(1-\eta))$. Then, $\bar{h}_{[j-1,u-1]} \in \mathcal{F}(m;\tau,\eta)$. □

Remark 2. If $\tau > 1$, Lemma 8 does not always hold.

4.2 Our Framework Includes Previous Works as Special Cases

We show that our framework includes previous works as special cases. First, we show that our Framework includes Herrmann–May's work [7] as a special case. We gave the set of shift-polynomials $\mathcal{F}_{\mathrm{HM}}(m;\tau)$ for Herrmann–May's method in Section 2.3. From the definition, it holds that

$$\mathcal{F}_{\mathrm{HM}}(m;\tau) = \mathcal{F}(m;\tau,1).$$

Then, Herrmann–May's method is obtained by setting $\eta = 1$ in our unified framework. Next, we show that our Framework includes Blömer–May's work [1] as a special case. We gave the set of shift-polynomials $\mathcal{F}_{\mathrm{BM}}(m;t)$ for Blömer–May's method in Section 3.1. From the definition, it holds that

$$\mathcal{F}_{\mathrm{BM}}(m;t) = \mathcal{F}(m;1,t/m).$$

Note that $t/m \leq 1$ from the definition. Then, Blömer–May's method is obtained by setting $\tau = 1$ in our unified framework.

4.3 Deriving a Condition for Solving Small Inverse Problem in our Framework

A lattice basis is constructed by using the coefficient vectors of shift-polynomials in $\mathcal{F}(m;\tau,\eta)$ as basis vectors. Note that the coefficient vectors of the shift-polynomials $\bar{g}_{[u-i,i]}(xX,zZ)$ and $\bar{h}_{[i,u]}(xX,yY,zZ)$ are written as row vectors. Let $B(m;\tau,\eta)$ be a matrix, where all rows of $B(m;\tau,\eta)$ are the coefficient vectors of shift-polynomials according to the ordering of $\mathcal{F}(m;\tau,\eta)$.

Theorem 2. *Let m be an integer. Let τ and η be parameters with $0 < \tau \leq 1$ and $0 \leq \eta \leq 1$. A lattice basis matrix $B(m;\tau,\eta)$ is triangular for any m, τ and η.*

Proof of Theorem 2 We show that the number of monomials newly appearing in expansion of shift-polynomial is one for any shift-polynomials in $\mathcal{F}(m;\tau,\eta)$. In this proof, we abbreviate $\mathcal{F}(m;\tau,\eta)$ as $\mathcal{F}$. We define $\mathcal{F}^f := \{g \in \mathcal{F} | g \prec f\}$ and $\mathcal{S}(\mathcal{F}^f) := \bigcup_{g\in\mathcal{F}^f} \mathcal{S}(g)$. It is enough for proving Theorem 2 to show that for any polynomial $f \in \mathcal{F}$ there exist a monomial m_f such that

- $\mathcal{S}(f - m_f) \subseteq \mathcal{S}(\mathcal{F}^f)$ and
- $m_f \notin \mathcal{S}(\mathcal{F}^f)$.

From Lemmas 4–6 and 7–8, for any $f \in \mathcal{F}$, there exists m_f such that $\mathcal{S}(f - m_f) \subseteq \mathcal{S}(\mathcal{F}^f)$. We can easily verify that $m_f \notin \mathcal{S}(\mathcal{F}^f)$. Then, the lattice basis matrix is triangular. □

We show a small example for $m = 3, \tau = 1/2$ and $\eta = 1/3$. We have

$$\begin{aligned} \mathcal{G}(3; 1/3) &= \{g_{[u-i,i]} | u = 2, 3; i = 0, \dots, u\} \text{ and} \\ \mathcal{H}(3; 1/2, 1/3) &= \{h_{[i,u]} | u = 2, 3; i = 1, \dots, u/2 - 1\}. \end{aligned}$$

or we explicitly have

$$\mathcal{G}(3; 1/3) = \{\bar{g}_{[2,0]}, \bar{g}_{[1,1]}, \bar{g}_{[0,2]}, \bar{g}_{[3,0]}, \bar{g}_{[2,1]}, \bar{g}_{[1,2]}, \bar{g}_{[0,3]}\} \text{ and } \mathcal{H}(3; 1/2, 1/3) = \{\bar{h}_{[1,3]}\}.$$

A lattice basis is constructed by using the coefficients vectors x-shifts $\bar{g}_{[i,j]}(xX, zZ)$ in $\mathcal{G}(3; 1/3)$ and y-shifts $\bar{h}_{[i,u]}(xX, yY, zZ)$ in $\mathcal{H}(3; 1/2, 1/3)$.

$$\begin{array}{c|cccccccc}
 & x^2 & xz & z^2 & x^3 & x^2z & xz^2 & z^3 & yz^3 \\ \hline
\bar{g}_{[2,0]} & X^2e^3 & & & & & & & \\
\bar{g}_{[1,1]} & Ae^2X^2 & e^2XZ & & & & & & \\
\bar{g}_{[0,2]} & A^2eX^2 & 2eAXZ & eZ^2 & & & & & \\
\bar{g}_{[3,0]} & & & & X^3e^3 & & & & \\
\bar{g}_{[2,1]} & & & & AX^3e^2 & e^2X^2Z & & & \\
\bar{g}_{[1,2]} & & & & eA^2X^3 & e2AX^2Z & eXZ^2 & & \\
\bar{g}_{[0,3]} & & & & A^3X^3 & 3A^2X^2Z & 3AXZ^2 & Z^3 & \\ \hline
\bar{h}_{[1,3]} & -A^3X^2 & -3A^2XZ & -3AZ^2 & 0 & A^3X^2Z & 3A^2XZ^2 & 3AZ^3 & YZ^3
\end{array}$$

Note that if we expand $\bar{h}_{[1,3]}$ by x and y instead of x and z, many monomials appears. The determinant of the above matrix is given by the product of diagonal elements: $e^{12}X^9Y^1Z^{12}$.

For the following asymptotic analysis, we omit roundings in setting of $\mathcal{F}(m; \tau, \eta)$ as their contribution is negligible for sufficiently large m. We denote by w_x and w_y the number of shift-polynomials used in x-shifts and y-shifts, respectively. And we denote by $\mathrm{vol}(L_X)$ and $\mathrm{vol}(L_Y)$ contributions in x-shifts and y-shifts to a lattice volume, respectively. The total number of shift-polynomials w is given by $w = w_x + w_y$ and a lattice volume $\mathrm{vol}(L)$ is given by $\mathrm{vol}(L) = \mathrm{vol}(L_X)\mathrm{vol}(L_Y)$.

First, we derive w_x and $\mathrm{vol}(L_X)$. The lattice dimension w_x is given by $w_x = \sum_{l=m(1-\eta)}^{m} \sum_{k=0}^{l} 1$. The volume $\mathrm{vol}(L_X)$ is given by

$$\mathrm{vol}(L_X) = \prod_{l=m(1-\eta)}^{m} \prod_{k=0}^{l} X^{l-k} Z^k e^{m-k} = e^{mw_x} \prod_{l=m(1-\eta)}^{m} \prod_{k=0}^{l} X^{l-k} \left(\frac{Z}{e}\right)^k.$$

Let $\mathrm{vol}(L_X) = e^{mw_x} X^{s_{XX}} (Z/e)^{s_{XZ}}$. Each s_{XX} and s_{XZ} is explicitly given as follows:

$$s_{XX} = \sum_{l=m(1-\eta)}^{m} \sum_{k=0}^{l} l - k = \frac{1-(1-\eta)^3}{6} m^3 + o(m^3) \text{ and}$$
$$s_{XZ} = \sum_{l=m(1-\eta)}^{m} \sum_{k=0}^{l} k = \frac{1-(1-\eta)^3}{6} m^3 + o(m^3).$$

Then, we have

$$\mathrm{vol}(L_X) = e^{mw_x} X^{(1-(1-\eta)^3)m^3/6} \left(\frac{Z}{e}\right)^{(1-(1-\eta)^3)m^3/6}.$$

Second, we derive w_y and $\mathrm{vol}(L_Y)$. The lattice dimension w_y is given by $w_y = \sum_{l=0}^{\eta m} \sum_{j=1}^{\tau l} 1$. The volume $\mathrm{vol}(L_Y)$ is given by

$$\mathrm{vol}(L_Y) = \prod_{l=0}^{\eta m} \prod_{j=1}^{\tau l} Y^j Z^{l+m(1-\eta)} e^{m-l-m(1-\eta)} = e^{mw_y} \prod_{l=0}^{\eta m} \prod_{j=1}^{\tau l} Y^j \left(\frac{Z}{e}\right)^{l+m(1-\eta)}.$$

Let $\mathrm{vol}(L_Y) = e^{mw_y} Y^{s_{YY}} (Z/e)^{s_{YZ}}$ Each s_{YY} and s_{YZ} is explicitly given as follows:

$$s_{YY} = \sum_{l=0}^{\eta m} \sum_{j=1}^{\tau l} j = \eta^3 \tau^2 \frac{m^3}{6} + o(m^3) \text{ and}$$
$$s_{YZ} = \sum_{l=0}^{\eta m} \sum_{j=1}^{\tau l} l + (1-\eta)m = \tau\eta^3 \frac{m^3}{3} + \tau(1-\eta)m\frac{\eta^2 m^2}{2} = \tau\eta^2(3-\eta)\frac{m^3}{6} + o(m^3).$$

Then, we have

$$\mathrm{vol}(L_Y) = e^{mw_y} Y^{\eta^3\tau^2 m^3/6} \left(\frac{Z}{e}\right)^{\tau\eta^2(3-\eta)m^3/6}.$$

Summing up the above discussion, we have

$$\begin{aligned} \mathrm{vol}(L) &= \mathrm{vol}(L_X)\mathrm{vol}(L_Y) \\ &= e^{mw} X^{\eta(3-3\eta+\eta^2)m^3/6} Y^{\eta^3\tau^2 m^3/6} \left(\frac{Z}{e}\right)^{(\eta(3-3\eta+\eta^2)+\tau\eta^2(3-\eta))m^3/6} \end{aligned} \tag{3}$$

Remember that the condition that the problem can be solved in polynomial time is given by $\mathrm{vol}(L) \leq e^{mw}$ by ignoring small terms. From Eq. (3), we have the condition:

$$X^{3-3\eta+\eta^2} Y^{\tau^2\eta^2} \left(\frac{Z}{e}\right)^{(3-3\eta+\eta^2)+\tau(3\eta-\eta^2)} \leq 1. \tag{4}$$

As described in previous subsection, we obtain the same set as those of Herrmann–May or Blömer–May if we set $\eta = 1$ or τ=1. Deriving bounds for each case are described in Appendix B.

4.4 Optimal Bound in our Framework

We have seen that the optimal bound of X is $e^{1-\sqrt{1/2}}$ if $\eta = 1$ or $\tau = 1$. Hence, we have a chance to go beyond the Boneh–Durfee's bound. Unfortunately, the following theorem shows that $d \leq N^{0.292}$ is still optimal in our framework.

Theorem 3. *Suppose that $Y = e^{1/2}$. The maximal bound of X in our framework is $e^{1-\sqrt{1/2}}$.*

Proof of Theorem 3 By substituting $Z = XY + 1$ and $Y = e^{1/2}$ into Eq. (4) and ignoring small terms, Eq. (4) is transformed into

$$X \leq e^{\frac{1}{2}\frac{(3-3\eta+\eta^2)+(3\eta-\eta^2)\tau-\eta^2\tau^2}{2(3-3\eta+\eta^2)+\tau(3\eta-\eta^2)}}. \tag{5}$$

Let $\mathcal{P}$ and $\bar{\mathcal{P}}$ be sets such that $\mathcal{P} = \{(\tau,\eta) \mid 0 < \tau < 1,\ 0 < \eta < 1\}$ and $\bar{\mathcal{P}} = \{(\tau,\eta) \mid 0 < \tau \leq 1,\ 0 < \eta \leq 1\}$. In order to obtain the maximal value of the right side of Eq. (5) in $\bar{\mathcal{P}}$, we firstly consider the extremal values of the following function $\Psi(\tau,\eta)$ in $\mathcal{P}$:

$$\Psi(\tau,\eta) := \frac{(3-3\eta+\eta^2)+(3\eta-\eta^2)\tau-\eta^2\tau^2}{2(3-3\eta+\eta^2)+(3\eta-\eta^2)\tau}.$$

Let $Num(\tau,\eta)$ and $Den(\tau,\eta)$ be the numerator and denominator of $\Psi(\tau,\eta)$ respectively. Here, we show that $Den(\tau,\eta) \neq 0$ in $\mathcal{P}$. If $Den(\tau,\eta) = 0$, then we have

$$0 < \tau = \frac{2(3-3\eta+\eta^2)}{\eta^2-3\eta} = 2\frac{(\eta-3/2)^2+\frac{3}{4}}{(\eta-3)\eta}.$$

However, this contradicts the condition $0 < \eta < 1$. Therefore, the rational function $\Psi(\tau,\eta) \in \mathbb{Q}(\tau,\eta)$ is obviously differentiable in $\mathcal{P}$. By solving the algebraic equation $\frac{\partial \Psi}{\partial \tau} = \frac{\partial \Psi}{\partial \eta} = 0$, we show that there are no extremal values of $\Psi(\tau, \eta)$ in $\mathcal{P}$. Let $\Phi_\tau(\tau, \eta), \Phi_\eta(\tau, \eta)$ be polynomials such that

$$\Phi_\tau(\tau, \eta) := \frac{\partial \Psi}{\partial \tau} \cdot Den(\tau, \eta)^2, \ \Phi_\eta(\tau, \eta) := \frac{\partial \Psi}{\partial \eta} \cdot Den(\tau, \eta)^2.$$

Note that both Φ_τ and Φ_η are in $\mathbb{Z}[\tau, \eta]$, and we solve the algebraic equation $\Phi_\tau = \Phi_\eta = 0$ by introducing Gröbner basis. Let G be the Gröbner basis for the ideal generated by Φ_τ, Φ_η with respect to the lexicographic order $\prec_{\text{LEX}}$ such that $\eta \prec_{\text{LEX}} \tau$. Then G contains three polynomials in $\mathbb{Z}[\tau, \eta]$, and one of them is $m(\eta)$ such that

$$m(\eta) = \eta(\eta - 1)(\eta - 3)(\eta^2 - 3\eta + 3)\{3(\eta - 1)^2 + 2(\eta - 3)^2\}.$$

This fact implies that, for every extremal value $\Psi(\tau_0, \eta_0)$ where $(\tau_0, \eta_0) \in \mathbb{R}^2$, η_0 is a root of $m(\eta)$ over $\mathbb{R}$. Since $m(\eta)$ does not have its root in the real interval $(0, 1)$, there are no extremal values of $\Psi(\tau, \eta)$ in $\mathcal{P}$.

Hence, we only have to check the maximal values of $\Psi(0, \eta), \Psi(1, \eta)$ for $0 \leq \eta \leq 1$ and $\Psi(\tau, 0), \Psi(\tau, 1)$ for $0 \leq \tau \leq 1$, and furthermore the two cases $\tau = 1$ and $\eta = 1$ are discussed above. The maximal value of the right side of Eq. (5) for $\tau = 0$ or $\eta = 0$ is $e^{1/4}$ since $\Psi(0, \eta) = \Psi(\tau, 0) = 1/2$, and thus the maximal value of the right side of Eq. (5) in $\bar{\mathcal{P}}$ is $e^{1-\sqrt{1/2}}$. □

4.5 A Hybrid Method

It has been known that Blömer–May method: $(\tau, \eta) = (1, 3 - \sqrt{6})$ has an advantage because their method requires a smaller lattice dimension. On the other hands, Herrmann–May method: $(\tau, \eta) = (\sqrt{2} - 1, 1)$ has an advantage because it achieves a higher bound. We present a simple hybrid method which enjoys both of advantages by interpolating two methods. Letting t be a parameter with $0 \leq t \leq 1$, we set $\tau(t)$ and $\eta(t)$ by

$$(\tau(t), \eta(t)) = (1 - (2 - \sqrt{2})t, (\sqrt{6} - 2)t + (3 - \sqrt{6}))$$

and use the parameter $(\tau(t), \eta(t))$ for our framework. The setting $t = 0$ corresponds to Blömer–May's method: $(\tau(0), \eta(0)) = (1, 3 - \sqrt{6})$ and the setting $t = 1$ corresponds to Herrmann–May's method: $(\tau(1), \eta(1)) = (\sqrt{2} - 1, 1))$. We define $\bar{\Psi}(t) := \Psi(\tau(t), \eta(t))$. We can easily see that $\bar{\Psi}(t)$is monotonically increasing function in the interval $0 \leq t \leq 1$. Then, there is a trade-off between a lattice dimension and an achievable bound. That is, the choice of a bigger t implies a higher bound but less efficiency and the choice of a smaller t implies more efficiency but a lower bound. Our hybrid method makes it possible to choose the best lattice construction for a practical attack.

5 Extension to Cryptanalysis of Arbitrary $Y = e^{\alpha}$

In previous section, we discussed only the case of $Y = e^{1/2}$. In this section, we extend our results to arbitrary $Y = e^{\alpha}$. Sarkar et al. presented the small secret exponent attack under the situation that a few MSBs of the prime p is known [13]. Suppose that some estimate p_0 of p is known such that $|p-p_0| < N^{\alpha}$. Let q_0 be an estimation of q. Letting $A = N+1-p_0-q_0$, a solution of the modular equation $x(A+y)+1 = 0 (\text{mod } e)$ is given by $(x, y) = (k, p_0+q_0-p-q)$. Note that $k < e^{\delta}$ and $|p_0+q_0-p-q| < e^{\alpha}$. They showed that the barrier $d < N^{0.292}$ can be broken through if α is strictly less than 1/2. In this section, we focus on the problem: $x(A+y)+1 = 0(\text{mod } e)$ with upper bound of solution: $X = e^{\delta}$ and $Y = e^{\alpha}$. They showed extensions of three algorithms: two algorithms from Boneh and Durfee's paper [2], and one algorithm from Blömer and May's paper [1] into arbitrary α [13]. Although α should be $1/4 < \alpha \leq 1/2$ in this attack scenario[4], we show an analysis for $0 < \alpha < 1$.

It is important to point out that the discussion in Sections 3 and 4 (except Sections 4.4 and 4.5) is valid for an arbitrary α, which implies that a set of indexes $\mathcal{F}(m; \tau, \eta)$ of shift-polynomials and the determinant calculation of the volume are also valid. From the same analysis, we have the same condition as Eq. (4). Letting $X = e^{\delta}$ and $Y = e^{\alpha}$, we have the following theorem. A proof is given in Appendix A.2.

Theorem 4. *Suppose that $Y = e^{\alpha}$ and $X = e^{\delta}$. The maximal bound of δ in our framework is given by*

$$\delta < \begin{cases} 1 - \sqrt{\alpha} & \text{if } \alpha \geq 1/4, \\ \dfrac{2}{5}(\sqrt{4\alpha^2 - \alpha + 1} - 3\alpha + 1) & \text{if } 0 < \alpha < 1/4. \end{cases}$$

We will present a hybrid method for arbitrary α in Appendix C. Theorem 4 shows that Blömer–May like method ($\tau = 1$) is superior to Herrmann–May like method ($\eta = 1$) if $\alpha < 1/4$. Interestingly, if α is extremely small ($\alpha < 3/35$), Herrmann–May like and Blömer–May like methods are not best-known algorithms. We show the details in Appendix D. We also another extension in Appendix E.

6 Concluding Remarks

We should point out the relationship between our results and the discussion in May's PhD thesis [11]. He presented the interpolation between

[4] Suppose that α is less than or equal to 1/4. The whole prime factor p can be found by Coppersmith's attack [4] since the upper half of p is known.

the results of Blömer–May and Boneh–Durfee by using a concept called *strictly decreasing pattern* in Section 7 of [11]. He also argued that Boneh–Durfee's stronger bound is optimal over all decreasing patterns. However, no *formal proof* of its optimality has been given in [11]. On the contrary to [11], we give a *strict proof* of the optimality within our framework in Section 4. Furthermore, we extend our results to arbitrary $Y = e^{\alpha}$, which has not been discussed in [11] and is also an advantage over [11].

It has been known that Blömer–May method has an advantage because their method requires a smaller lattice dimension than the Boneh–Durfee's lattice. Theorem 4 gives another view of their algorithm. Theorem 4 shows Blömer–May method has another advantage because it achieves a better bound in addition to less lattice dimension; Blömer–May method achieves a higher bound than Herrmann–May method (and Boneh–Durfee's method) if $\alpha \leq 1/4$.

For the usual small secret exponent attack on RSA, we *just* showed that $d \leq N^{0.292}$ is an optimal bound in our framework. Hence, the bound might be improved if we develop the other method outside of our framework, which is an open problem.

Acknowledgement

The first author was supported by KAKENHI 22700006.

References

1. J. Blömer and A. May, "Low Secret Exponent RSA Revisited," in Proc. of CaLC2001, LNCS 2146, pp. 4–19, 2001.
2. D.Boneh and G.Durfee, "Cryptanalysis of RSA with private key d less than $N^{0.292}$," IEEE Transactions on Information Theory, vol. 46, no. 4, pp. 1339–1349, 2000. (Firstly appeared in Eurocrypt'99).
3. D. Coppersmith, "Finding a Small Root of a Univariate Modular Equation,"in Proc. of Eurocrypt'96, LNCS 1070, pp. 155–165, 1996.
4. D. Coppersmith, "Finding a Small Root of a Bivariate Integer Equation; Factoring with High Bits Known," in Proc. of Eurocrypt'96, LNCS 1070, pp. 178–189, 1996.
5. D. Coppersmith, "Small Solutions to Polynomial Equations, and Low Exponent RSA Vulnerabilities," J. Cryptology 10(4): 233-260, 1997.
6. M. Herrmann and A. May, "Attacking Power Generators Using Unravelled Linearization: When do we Output Too Much?," in Proc. of Asiacrypt2009, LNCS5912, pp. 487–504, 2009.
7. M. Herrmann and A. May, "Maximizing Small Root Bounds by Linearization and Applications to Small Secret Exponent RSA," In Proc. of PKC2010, LNCS6056, pp. 53–69, 2010.
8. N. Howgrave-Graham, "Finding Small Roots of Univariate Modular Equations Revisited," IMA Int. Conf., pp.131–142, 1997.

9. N. Kunihiro, N. Shinohara and T. Izu, "A Unified Framework for Small Secret Exponent Attack on RSA," in Proc. of SAC2011, LNCS 7118, pp. 260–277, 2011.
10. A.K. Lenstra, H.W. Lenstra, L. Lovász, "Factoring polynomials with rational coefficients," Mathematische Annalen 261, pp.515–534, 1982.
11. A. May, "New RSA Vulnerabilities Using Lattice Reduction Methods," PhD thesis, University of Paderborn, 2003.
12. R. Rivest, A. Shamir and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," Communications of the ACM, vol. 21(2), pp. 120–126, 1978.
13. S. Sarkar, S. Maitra and S. Sarkar, "RSA Cryptanalysis with Increased Bounds on the Secret Exponent using Less Lattice Dimension," IACR ePrint Archive: Report 2008/315, 2008.
14. M. Wiener, "Cryptanalysis of Short RSA Secret Exponents," IEEE Transactions on Information Theory, vol. 36, pp. 553–558, 1990.

A Proofs

A.1 Proofs of Lemma 4–6

Proof of Lemma 4 The polynomial $\bar{g}_{[u,0]}$ is given by $\bar{g}_{[u,0]}(x,z) = e^m x^u$. Then, we have the lemma. □

Proof of Lemma 5 The expansion of $\bar{g}_{[u-j,j]}$ for $j \geq 1$ is given by

$$\begin{aligned}\bar{g}_{[u-j,j]}(x,z) &= e^{m-j}x^{u-j}(z+Ax)^j = \sum_{i=0}^{j} e^{m-j}x^{u-j}\binom{j}{i}z^i x^{j-i}A^{j-i}\\ &= e^{m-j}x^{u-j}z^j + \sum_{i=0}^{j-1} e^{m-j}A^{j-i}\binom{j}{i}x^{u-i}z^i.\end{aligned}$$

Then, we have

$$\begin{aligned}\bar{g}_{[u-j,j]}(x,z) - e^{m-j}x^{u-j}z^j &\cong \sum_{i=0}^{j-1} x^{u-i}z^i = x^{u-j+1}\sum_{i=0}^{j-1} x^{(j-1)-i}z^i\\ &\cong x^{u-j+1}(z+Ax)^{j-1} \cong \bar{g}_{[u-j+1,j-1]}.\end{aligned}$$

Then, we have $\mathcal{S}(\bar{g}_{[u-j,j]} - e^{m-j}x^{u-j}z^j) = \mathcal{S}(\bar{g}_{[u-j+1,j-1]})$. □

Proof of Lemma 6 The expansion of $\bar{h}_{[j,u]}$ for $j \geq 1$ is given as follows:

$$\begin{aligned}\bar{h}_{[j,u]}(x,y,z) &= y^j(z+Ax)^u e^{m-u} = e^{m-u}\sum_{i=0}^{u}\binom{u}{i}y^j z^i (Ax)^{u-i}\\ &= e^{m-u}y^j z^u + e^{m-u}\sum_{i=0}^{u-1}\binom{u}{i}A^{u-i}x^{u-i}y^j z^i.\end{aligned}$$

Then, we have

$$\begin{aligned}
&\bar{h}_{[j,u]}(x,y,z) - e^{m-u}y^j z^u \cong \sum_{i=0}^{u-1} x^{u-i}y^j z^i = y^{j-1}xy\sum_{i=0}^{u-1} x^{(u-1)-i}z^i \\
&\cong y^{j-1}(z-1)(z+Ax)^{u-1} \cong y^{j-1}(z+Ax)^{u-1}z + y^{j-1}(z+Ax)^{u-1} \\
&\cong \bar{h}_{[j-1,u-1]}z + \bar{h}_{[j-1,u-1]}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\mathcal{S}(\bar{h}_{[j,u]}(x,y,z) - e^{m-u}y^j z^u) = \mathcal{S}(\bar{h}_{[j-1,u-1]}z) \cup \mathcal{S}(\bar{h}_{[j-1,u-1]}) \\
&\subseteq \mathcal{S}(h_{[j-1,u-1]}(z+Ax)) \cup \mathcal{S}(\bar{h}_{[j-1,u-1]}) = \mathcal{S}(\bar{h}_{[j-1,u]}) \cup \mathcal{S}(\bar{h}_{[j-1,u-1]}).
\end{aligned}$$

Then, we have the lemma. □

A.2 Proof of Theorem 4

By substituting $Z = XY + 1$ and $Y = e^{\alpha}$ into Eq. (4) and ignoring small terms, Eq. (4) is transformed into

$$X \leq e^{\frac{(1-\alpha)((3-3\eta+\eta^2)+(3\eta-\eta^2)\tau)-\alpha\eta^2\tau^2}{2(3-3\eta+\eta^2)+(3\eta-\eta^2)\tau}}. \tag{6}$$

Let $\mathcal{P}$ and $\bar{\mathcal{P}}$ be the sets defined in the proof of Theorem 3. In order to obtain the maximal value of the right side of (6) in $\bar{\mathcal{P}}$, we firstly consider the extremal values of the following function $\Psi_\alpha(\tau,\eta)$ in $\mathcal{P}$:

$$\Psi_\alpha(\tau,\eta) = \frac{(1-\alpha)((3-3\eta+\eta^2)+(3\eta-\eta^2)\tau)-\alpha\eta^2\tau^2}{2(3-3\eta+\eta^2)+(3\eta-\eta^2)\tau}$$

Notice that the denominator of $\Psi_\alpha(\tau,\eta)$ is $Den(\tau,\eta)$ given in the proof of Theorem 3, and so $\Psi_\alpha(\tau,\eta)$ is also differentiable in $\mathcal{P}$.

In the same manner as the proof of Theorem 3, we show that there are no extremal values of $\Psi_\alpha(\tau,\eta)$ in $\mathcal{P}$ for any $\alpha \in (0,1)$. Let $\Phi_\tau^{(\alpha)}(\tau,\eta)$, $\Phi_\eta^{(\alpha)}(\tau,\eta)$ be polynomials such that

$$\Phi_\tau^{(\alpha)}(\tau,\eta) = \frac{\partial\Psi_\alpha}{\partial\tau}\cdot Den(\tau,\eta)^2,\ \Phi_\eta^{(\alpha)}(\tau,\eta) = \frac{\partial\Psi_\alpha}{\partial\eta}\cdot Den(\tau,\eta)^2.$$

We solve the algebraic equation $\Phi_\tau^{(\alpha)} = \Phi_\eta^{(\alpha)} = 0$ by introducing Gröbner basis. Let G_α be the Gröbner basis under $0 < \alpha < 1$ for the ideal generated by $\Phi_\tau^{(\alpha)}, \Phi_\eta^{(\alpha)}$ with respect to the lexicographic order $\prec_{\mathrm{LEX}}$ such that $\eta \prec_{\mathrm{LEX}} \tau$. One of polynomials in G_α is $m_\alpha(\eta)$ such that

$$m_\alpha(\eta) = \eta(\eta-1)(\eta-3)(\eta^2-3\eta+3)\{3\alpha(\eta-1)^2+(\eta-3)^2\}.$$

This fact implies that, for every extremal value $\Psi_\alpha(\tau_0, \eta_0)$ where $(\tau_0, \eta_0) \in \mathbb{R}^2$, η_0 is a root of $m_\alpha(\eta)$ over $\mathbb{R}$. Since $m_\alpha(\eta)$ does not have its root in the real interval $(0, 1)$, there are no extremal values of $\Psi_\alpha(\tau, \eta)$ in $\mathcal{P}$.

Hence, we only have to check the maximal values of $\Psi_\alpha(0, \eta), \Psi_\alpha(1, \eta)$ for $0 \le \eta \le 1$ and $\Psi_\alpha(\tau, 0), \Psi_\alpha(\tau, 1)$ for $0 \le \tau \le 1$. If $\eta = 0$ or $\tau = 0$, then $\Psi_\alpha(\tau, 0) = \Psi_\alpha(0, \eta) = (1 - \alpha)/2$, and so the maximal value for $\eta = 0$ or $\tau = 0$ is $(1 - \alpha)/2$.

For $\eta = 1$, we have that

$$\Psi_\alpha(\tau, 1) = \frac{-\alpha\tau^2 + (1 - \alpha)(1 + 2\tau)}{2(\tau + 1)},$$

and so the maximal value for $\eta = 1$ is

$$\begin{cases} \frac{3}{4} - \alpha \quad (\tau = 1, 0 < \alpha < \frac{1}{4}) \\ 1 - \sqrt{\alpha}\ (\tau = \sqrt{1/\alpha} - 1, \frac{1}{4} \le \alpha < 1). \end{cases} \tag{7}$$

For $\tau = 1$, we have that

$$\Psi_\alpha(1, \eta) = \frac{3 - \alpha(\eta^2 + 3)}{6 - 3\eta + \eta^2},$$

and so the maximal value for $\tau = 1$ is

$$\frac{2}{5}(\sqrt{\alpha^2 - \alpha + 1} - 3\alpha + 1). \tag{8}$$

By comparing with the above values, we have the theorem. □

B Deriving Bounds for Each Case

B.1 Herrmann-May's case

First, we derive a bound for Herrmann-May's case. Substituting $\eta = 1$ into Eq. (4), we have $XY^{\tau^2}\left(\frac{Z}{e}\right)^{1+2\tau} \le 1$. By substituting $Z \le 2XY$ and $Y = e^{1/2}$ into the above inequality, we have

$$X \le e^{\frac{1+2\tau-\tau^2}{4(1+\tau)}}.$$

By maximizing X, the maximal value of X is given by $X \le e^{1-\sqrt{1/2}}$ by setting $\tau = \sqrt{2} - 1$. Note that $0 < \sqrt{2} - 1 \le 1$. This bound is equivalent to Boneh–Durfee's bound. We can easily extend to the case of $Y = e^{\alpha}$. In this case, we have the bound: $\delta < 1 - \sqrt{\alpha}$ by setting $\tau = \sqrt{1/\alpha} - 1$, which is equivalent to Theorem 4 in [13].

B.2 Blömer–May's case

Second, we derive a bound for Blömer–May's case. Substituting $\tau = 1$ into Eq. (4), we have $X^{3-3\eta+\eta^2}Y^{\eta^2}\left(\frac{Z}{e}\right)^3 \leq 1$. By substituting $Z \leq 2XY$ and $Y = e^{1/2}$ into the above inequality, we have

$$X \leq e^{\frac{3-\eta^2}{2(6-3\eta+\eta^2)}}.$$

By maximizing X, the maximal value of X is given by $X \leq e^{(\sqrt{6}-1)/5}$ by setting $\eta = 3 - \sqrt{6}$. Note that $0 \leq 3 - \sqrt{6} \leq 1$. This bound is equivalent to Blömer–May's bound [1]. We can easily extend to the case of $Y = e^{\alpha}$. In this case, we have the bound: $\delta < \frac{2}{5}(\sqrt{4\alpha^2 - \alpha + 1} - 3\alpha + 1)$ by setting $\eta = ((1+\alpha) - \sqrt{4\alpha^2 - \alpha + 1})/\alpha$, which is equivalent to Theorem 5 in [13].

C Hybrid Method for $\alpha \geq 1/4$

We can apply the hybrid method as described in Section 4.5 into $\alpha \geq 1/4$. Let t be a parameter with $0 \leq t \leq 1$. We set $\tau(t;\alpha)$ and $\eta(t;\alpha)$ by

$$(\tau(t;\alpha), \eta(t;\alpha)) = \left(1 - \left(2 - \sqrt{\frac{1}{\alpha}}\right)t, \frac{\sqrt{4\alpha^2 - \alpha + 1} - 1}{\alpha}t + \frac{\alpha + 1 - \sqrt{4\alpha^2 - \alpha + 1}}{\alpha}\right).$$

Set $\bar{\Psi}(t;\alpha) := \Psi(\tau(t;\alpha), \eta(t;\alpha))$. We can easily verify that $\bar{\Psi}(t;\alpha)$is monotonically increasing function in the interval $0 \leq t \leq 1$.

D Best Known Bound for an Arbitrary α

Sarkar et al. [13] proved that RSA is insecure by extending Boneh–Durfee's weaker method into arbitrary α if $\delta < (\alpha + 3 - 2\sqrt{\alpha(\alpha+3)})/3$. Note that this bound can be obtained by using $\mathcal{F}_{\mathrm{BD}}$ introduced in Section 2.3. Corollary 1 shows that the best bound in our framework and Sarkar et al.'s extension of Boneh–Durfee's result.

Corollary 1. *Suppose that $X = e^{\delta}$ and $Y = e^{\alpha}$. The maximal bound of δ in our framework and the Extension of Boneh-Durfee's weaker bound is given by*

$$\delta < \begin{cases} 1 - \sqrt{\alpha} & \text{if } \alpha \geq \frac{1}{4}, \\ \dfrac{2}{5}(\sqrt{4\alpha^2 - \alpha + 1} - 3\alpha + 1) & \text{if } \frac{3}{35} \leq \alpha < \frac{1}{4}, \\ \dfrac{\alpha + 3 - 2\sqrt{\alpha(\alpha+3)}}{3} & \text{if } 0 < \alpha < \frac{3}{35}. \end{cases}$$

E Extension to $x(A+y)+C = 0 (\mathrm{mod}\ e)$ for an Arbitrary Integer C

In Section 4, we discussed only the case of $x(A+y)+1 = 0(\mathrm{mod}\ e)$. In this section, we extend to $x(A+y)+C = 0(\mathrm{mod}\ e)$ for an arbitrary integer C. For simplicity, we assume that $0 < |C| < e$. In the discussion of Section 4, we set Z as $Z = XY + 1$. For general C, Z should be replaced into $Z = XY + |C|$. The value Z is upper bounded by $2\max(XY, |C|)$. We consider two typical cases.

Suppose that $|C|$ is small compared to XY for the first case. Concretely, suppose that $XY \geq |C|$. Since $Z \leq 2XY$, the discussion of Section 4 is valid and the same bound is obtained.

Suppose that $|C|$ is uniformly chosen from integers within the interval $(0, e)$. It is clear that $|C| \approx e$ with high probability. In this case, $Z \leq 2e$. By ignoring small terms, Eq. (4) is transformed into $X^{3-3\eta+\eta^2} Y^{\tau^2\eta^2} < 1$. Since X and Y are positive integers, there are no ranges for X and Y satisfying the above inequality. Then, we cannot solve the problem by using our framework[5] in this case.

[5] Boneh and Durfee's weaker method is valid even if $|C|$ is large.