Predicting the Inversive Generator

Simon R. Blackburn[1], Domingo Gomez-Perez[2],
Jaime Gutierrez[3] and Igor E. Shparlinski[4]

[1] Department of Mathematics, Royal Holloway, University of London
Egham, Surrey, TW20 0EX, UK
s.blackburn@rhul.ac.uk
[2] Faculty of Science, University of Cantabria
E-39071 Santander, Spain
gomezd@unican.es
[3] Faculty of Science, University of Cantabria
E-39071 Santander, Spain
jaime.gutierrez@unican.es
[4] Department of Computing, Macquarie University,
NSW 2109, Australia
igor@comp.mq.edu.au

Abstract. Let p be a prime and let a and b be integers modulo p. The inversive congruential generator (ICG) is a sequence (u_n) of pseudorandom numbers defined by the relation $u_{n+1} \equiv au_n^{-1} + b \bmod p$. We show that if b and sufficiently many of the most significant bits of three consecutive values u_n of the ICG are given, one can recover in polynomial time the initial value u_0 (even in the case where the coefficient a is unknown) provided that the initial value u_0 does not lie in a certain small subset of exceptional values.

1 Introduction

For a prime p, denote by $\mathbb{F}_p$ the field of p elements and always assume that it is represented by the set $\{0, 1, \ldots, p-1\}$. Accordingly, sometimes, where obvious, we treat elements of $\mathbb{F}_p$ as integer numbers in the above range.

For fixed $a, b \in \mathbb{F}_p^*$, let $\psi_{a,b}$ be the permutation of $\mathbb{F}_p$ defined by

$$\psi_{a,b}(w) = \begin{cases} aw^{-1} + b, & \text{if } w \neq 0, \\ b, & \text{if } w = 0. \end{cases}$$

We refer to the coefficients a and b as the *multiplier* and *shift*, respectively.

We define the *inversive generator* (u_n) of elements of $\mathbb{F}_p$ by the recurrence relation

$$u_{n+1} = \psi_{a,b}(u_n), \qquad n = 0, 1, \ldots, \tag{1}$$

where u_0 is the *initial value*.

This generator has proved to be extremely useful for Quasi-Monte Carlo type applications, and in particular exhibits very attractive uniformity of distribution

and nonlinearity properties, see [11–14] for surveys or recent results. This paper concentrates on the cryptographic properties of the inversive generator.

In the cryptographic setting, the initial value u_0 and the constants a and b are assumed to be the secret key, and we want to use the output of the generator as a stream cipher. Of course, if several consecutive values u_n are revealed, it is easy to find u_0, a and b. So in this setting, we output only the most significant bits of each u_n in the hope that this makes the resulting output sequence difficult to predict. In a recent paper [2], we have shown that not too many bits can be output at each stage: the inversive generator is unfortunately polynomial time predictable if sufficiently many bits of its consecutive elements are revealed, so long as a small number of secret keys are excluded. However, most of the results of [2] only hold after excluding a small set of pairs (a, b). If this small set is not excluded, the algorithm for finding the secret information may fail. An optimist might hope that by deliberately choosing the pair (a, b) to lie in this excluded set, one can generate cryptographically stronger sequences. This paper aims to show that this strategy is unlikely to succeed. Namely we introduce some modifications and additions to the method of [2] which allow us to attack the generators no matter how the values of a and b are chosen. We demonstrate our approach in the special case when b is public. Of course, the assumption that b is public reduces the relevance of the problem to cryptography. But we believe that the extra strength of the result we obtain makes this situation of interest in its own right. We also believe this approach can be extended to the case when both a and b are secret.

Assume that the sequence (u_n) is not known but, for some n, approximations w_j of 3 consecutive values u_{n+j}, $j = 0, 1, 2$, are given. We show that if b is public, the values u_{n+j} and a can be recovered from this information in polynomial time if the approximations w_j are sufficiently good and if a certain small set of initial values u_0 are excluded. (The results in [2] exclude a small set of pairs (a, b) in addition to values of u_0, and so in this sense our result here is stronger.)

Throughout the paper the term polynomial time means polynomial in $\log p$. Our results involve another parameter Δ which measures how well the values w_j approximate the terms u_{n+j}. This parameter is assumed to vary independently of p subject to satisfying the inequality $\Delta < p$ (and is not involved in the complexity estimates of our algorithms).

We should emphasise that this paper is concerned with rigorous results (see [2] for a discussion of both rigorous and heuristic methods).

The remainder of the paper is structured as follows.

We start with a short outline of some basic facts about lattices in Section 2.1 and rational functions Section 2.2. In Section 3.1 we formulate our main results and outline the plan of the proof, which is given in Section 3.2. Finally, Section 4 makes some final comments and poses several open questions.

Acknowledgment. The authors would like to thank Harald Niederreiter for his interest and helpful discussions. This paper was written during visits of I.S. to the University of Cantabria (partially supported by MECD grant SAB2000-0260) and to Royal Holloway, University of London (supported by an

EPSRC Visiting Fellowship). D. G.-P. and J.G. were partially supported by Spanish Ministry of Science grant BFM2001-1294. The support and hospitality of all these organisations are gratefully acknowledged.

2 Lattices and Rational Functions

2.1 Background on Lattices

Here we collect several well-known facts about lattices which form the background to our algorithms.

We review several related results and definitions on lattices which can be found in [3]. For more details and more recent references, we also recommend consulting [1, 4, 5, 8–10].

Let $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ be a set of linearly independent vectors in $\mathbb{R}^r$. The set

$$\mathcal{L} = \{\mathbf{z} \ : \ \mathbf{z} = c_1\mathbf{b}_1 + \ldots + c_s\mathbf{b}_s, \quad c_1, \ldots, c_s \in \mathbb{Z}\}$$

is called an *s-dimensional lattice* with *basis* $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$. If $s = r$, the lattice L is of *full rank*.

To each lattice $\mathcal{L}$ one can naturally associate its *volume*

$$\mathrm{vol}\,(\mathcal{L}) = \left(\det\left(\langle \mathbf{b}_i, \mathbf{b}_j\rangle\right)_{i,j=1}^{s}\right)^{1/2},$$

where $\langle \mathbf{a}, \mathbf{b}\rangle$ denotes the inner product, which does not depend on the choice of the basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$.

For a vector $\mathbf{u}$, let $\|\mathbf{u}\|$ denote its *Euclidean norm*. The famous Minkowski theorem, see Theorem 5.3.6 in Section 5.3 of [3], gives the upper bound

$$\min\{\|\mathbf{z}\|: \ \mathbf{z} \in \mathcal{L} \setminus \{\mathbf{0}\}\} \le s^{1/2}\,\mathrm{vol}\,(\mathcal{L})^{1/s} \tag{2}$$

on the shortest nonzero vector in any s-dimensional lattice $\mathcal{L}$ in terms of its volume. In fact $s^{1/2}$ can be replaced by the *Hermite constant* $\gamma_s^{1/2}$, for which we have

$$\frac{1}{2\pi e}s + o(s) \le \gamma_s \le \frac{1.744}{2\pi e}s + o(s), \qquad s \to \infty.$$

The Minkowski bound (2) motivates a natural question: how to find the shortest vector in a lattice. The celebrated *LLL algorithm* of Lenstra, Lenstra and Lovász [7] provides a desirable solution in practice, and the problem is known to be solvable in deterministic polynomial time (polynomial in the bit-size of the basis of $\mathcal{L}$) provided that the dimension of $\mathcal{L}$ is fixed (see Kannan [6, Section 3], for example). The lattices in this paper are of fixed dimension. (Note that there are several indications that the shortest vector problem is **NP**-complete when the dimension grows.)

In fact, in this paper we consider only very special lattices. Namely, only lattices which are consisting of integer solutions $\mathbf{x} = (x_0, \ldots, x_{s-1}) \in \mathbb{Z}^s$ of the

system of congruences

$$\sum_{i=0}^{s-1} a_{ij} x_i \equiv 0 \bmod q_j, \qquad j = 1, \ldots, m,$$

modulo some integers $q_1, \ldots, q_m$. Typically (although not always) the volume of such a lattice is the product $Q = q_1 \ldots q_m$. Moreover all the aforementioned algorithms, when applied to such a lattice, become polynomial in $\log Q$.

2.2 Zeros of Rational Functions

Our second basic tool is essentially the theorem of Lagrange which asserts that a non-zero polynomial of degree N over any field has no more than N zeros in this field. In fact we apply it to rational functions which require only obvious adjustments.

The rational functions we consider belong to a certain family of functions parametrised by small vectors in a certain lattice, thus the size of the family can be kept under control. Zeros of these rational functions correspond to potentially "bad" initial values of the inversive generator (1). Thus, if all rational functions in this family are not identical to zero modulo p then we have an upper bound on the number of such "bad" initial values. Hence, a crucial part of our approach is to study possible vanishing of functions in the above family and to show that this may happen only for very few values of the coefficients of the generator (1). To establish this property we repeatedly use the fact that non-trivial linear combinations of rational functions with pairwise distinct poles do not vanish identically.

3 Predicting the Inversive Generator with Unknown Multiplier

3.1 Formulation of the Main Result and Plan of Proof

Assume the multiplier a of the inversive generator is unknown, but shift b is given to us. We show that we can recover u_0 and a for all but $O(\Delta^5)$ values of u_0 when given approximations to three consecutive values u_n, u_{n+1}, u_{n+2} produced by the inversive generator, except when u_0 lies in a small set of exceptional values. To simplify the notation, we assume that $n = 0$ from now on.

Theorem. *Let p be a prime number and let Δ be an integer such that $p > \Delta \geq 1$. Let $a, b \in \mathbb{F}_p^*$. There exists a set $\mathcal{U}(\Delta; a, b) \subseteq \mathbb{F}_p$ of cardinality $\#\mathcal{U}(\Delta; a, b) = O(\Delta^5)$ with the following property. Whenever $u_0 \notin \mathcal{U}(\Delta; a, b)$ then, given approximations $|w_j - u_j| \leq \Delta$, $j = 0, 1, 2$ to three consecutive values u_0, u_1, u_2 produced by the inversive generator (1), and given the value of b, one can recover u_0 and a in deterministic polynomial time.*

An outline of the algorithm given in the proof of this Theorem goes as follows. The algorithm is divided into six stages.

Stage 1: We assume that the two exceptional values 0 and $-a/b$ lie in $\mathcal{U}(\Delta; a, b)$. We construct a certain lattice $\mathcal{L}$ (see (5) below) of dimension five; this lattice depends on the approximations w_0, w_1, w_2 and the integer b. We also show that a certain vector $\mathbf{e}$ directly related to missing information about u_0, u_1, u_2 is a very short vector in this lattice. A shortest nonzero vector $\mathbf{f} = (f_0, \ldots, f_4)$ in $\mathcal{L}$ is found; see [6] for the appropriate algorithm.

Stage 2: We show that $\mathbf{f}$ provides some valuable information about $\mathbf{e}$ for all initial values u_0 except for u_0 from a certain exceptional set $\mathcal{V}(\Delta; a, b) \subseteq \mathbb{F}_p$ of cardinality $\#\mathcal{V}(\Delta; a, b) = O(\Delta^5)$ (which is defined as a set of zeros of a certain parametric family of rational functions).

Stage 3: We show that if $f_0 \neq 0$ then recovering $\mathbf{e}$ (and hence the secret information u_0 and a) from $\mathbf{f}$ is straight forward. If $f_0 \neq 0$, the algorithm terminates at this stage.

Stage 4: We show that if $f_0 = 0$ then the vector $\mathbf{f}$ enables us to compute small integers r and s such that $b = r/s \bmod p$. (In fact these integers can be found independently by the continued fraction algorithm.) The algorithm uses this information, together with the integers w_0, w_1, w_2 and b, to compute a second lattice $\mathcal{L}'$ of dimension four. There is a short vector $\mathbf{e}'$ in $\mathcal{L}$, and again this vector is closely related to the secret information u_0 and a.

Stage 5: We show that all short vectors in $\mathcal{L}'$ are parallel to $\mathbf{e}'$ for all initial values u_0 except for u_0 from another exceptional set $\mathcal{V}'(\Delta; a, b) \subseteq \mathbb{F}_p$ of cardinality $\#\mathcal{V}'(\Delta; a, b) = O(\Delta^5)$ (which is also defined as a set of zeros of a certain parametric family of rational function).

Stage 6: We find a shortest nonzero vector $\mathbf{f}'$ in $\mathcal{L}'$ and show that if $u_0 \notin \mathcal{U}(\Delta; a, b)$, where

$$\mathcal{U}(\Delta; a, b) = \{0, -a/b\} \cup \mathcal{V}(\Delta; a, b) \cup \mathcal{V}'(\Delta; a, b)$$

then recovering $\mathbf{e}'$ (and thus finding the secret information) from $\mathbf{f}$ and $\mathbf{f}'$ is now straightforward.

3.2 Proof of the Main Result

The theorem is trivial when $\Delta^5 \geq p$, and so we assume that $\Delta^5 < p$. Let us fix $a, b \in \mathbb{F}_p^*$. We assume that $u_0 \in \mathbb{F}_p$ is chosen so as not to lie in a certain subset $\mathcal{U}(\Delta; a, b)$ of $\mathbb{F}_p^*$. This subset is of cardinality $O(\Delta^5)$, but as its definition is fairly complicated we define it gradually as we move through the proof.

Stage 1: Building the lattice $\mathcal{L}$. We begin by defining a lattice $\mathcal{L}$, and showing how knowing a short vector in $\mathcal{L}$ usually leads to the recovery of the secret information.

We may assume that $u_0 u_1 \not\equiv 0 \bmod p$, for clearly there are at most two values of u_0, namely $u_0 \equiv 0 \pmod p$ and $u_0 \equiv -a/b \pmod p$ for which this does not

hold, and we place these two values in $\mathcal{U}(\Delta; a, b)$. From

$$u_1 \equiv a u_0^{-1} + b \bmod p \qquad \text{and} \qquad u_2 \equiv a u_1^{-1} + b \bmod p$$

we derive

$$u_1 u_0 \equiv a + b u_0 \bmod p \qquad \text{and} \qquad u_1 u_2 \equiv a + b u_1 \bmod p. \tag{3}$$

Therefore,

$$u_1(u_2 - u_0) \equiv b(u_1 - u_0) \bmod p. \tag{4}$$

For $j \in \{0, 1, 2\}$, define $\varepsilon_j = u_j - w_j$. We have that $|\varepsilon_j| \le \Delta$. Now (4) becomes

$$(w_1 + \varepsilon_1)(w_2 - w_0 + \varepsilon_2 - \varepsilon_0) \equiv b(w_1 - w_0) + b(\varepsilon_1 - \varepsilon_0) \bmod p.$$

Writing

$$\begin{aligned} &A \equiv (w_1(w_2 - w_0) - b(w_1 - w_0))\, \Delta^{-2} \bmod p, \quad B_0 \equiv -(w_1 + b)\Delta^{-1} \bmod p, \\ &B_1 \equiv (w_2 - w_0 - b)\Delta^{-1} \bmod p, \quad B_2 \equiv w_1 \Delta^{-1} \bmod p \quad \text{and} \quad C \equiv 1 \bmod p, \end{aligned}$$

we obtain

$$A\Delta^2 + B_0 \Delta \varepsilon_0 + B_1 \Delta \varepsilon_1 + B_2 \Delta \varepsilon_2 + C \varepsilon_1 (\varepsilon_2 - \varepsilon_0) \equiv 0 \bmod p.$$

Therefore the lattice $\mathcal{L}$ consisting of solutions $\mathbf{x} = (x_0, x_1, x_2, x_3, x_4) \in \mathbb{Z}^5$ of the congruences

$$\begin{aligned} A x_0 + B_0 x_1 + B_1 x_2 + B_2 x_3 + C x_4 &\equiv 0 \bmod p, \\ x_0 &\equiv 0 \bmod \Delta^2, \\ x_1 \equiv x_2 \equiv x_3 &\equiv 0 \bmod \Delta, \end{aligned} \tag{5}$$

contains a vector

$$\mathbf{e} = \left(\Delta^2 e_0, \Delta e_1, \Delta e_2, \Delta e_3, e_4\right) = \left(\Delta^2, \Delta\varepsilon_0, \Delta\varepsilon_1, \Delta\varepsilon_2, \varepsilon_1(\varepsilon_2 - \varepsilon_0)\right).$$

We have

$$e_0 = 1, \qquad |e_1|, |e_2|, |e_3| \le \Delta, \qquad |e_4| \le 2\Delta^2$$

thus the Euclidean norm $\|\mathbf{e}\|$ of $\mathbf{e}$ satisfies the inequality

$$\|\mathbf{e}\| \le \left(\Delta^4 + \Delta^4 + \Delta^4 + \Delta^4 + 4\Delta^4\right)^{1/2} \le 3\Delta^2.$$

Let $\mathbf{f} = (\Delta^2 f_0, \Delta f_1, \Delta f_2, \Delta f_3, f_4) \in \mathcal{L}$ be a shortest nonzero vector in $\mathcal{L}$. So $\|\mathbf{f}\| \le \|\mathbf{e}\| \le 3\Delta^2$. We have

$$|f_0| \le \|\mathbf{f}\| \Delta^{-2} \le 3, \qquad |f_1|, |f_2|, |f_3| \le \|\mathbf{f}\| \Delta^{-1} \le 3\Delta, \qquad |f_4| \le \|\mathbf{f}\| \le 3\Delta^2.$$

Note that we may compute $\mathbf{f}$ in polynomial time from the information we are given.

Stage 2: Defining the first exceptional set $\mathcal{V}(\Delta; a, b)$. The vector $\mathbf{d}$ defined by $f_0\mathbf{e} - e_0\mathbf{f}$ lies in $\mathcal{L}$ and has first component 0. We might hope that $\mathbf{e}$ and $\mathbf{f}$ are always parallel, in which case $\mathbf{d}$ would be the zero vector. Sadly, this is not always the case. So we claim that something weaker is true: namely that $d_2 = 0$ and $d_3 - d_1 = 0$ unless u_0 belongs to the set $\mathcal{V}(\Delta; a, b)$ which we define below. Before we establish this claim, we prove some facts about the vector $\mathbf{d}$.

Using the first congruence in (5), we find that

$$B_0\Delta d_1 + B_1\Delta d_2 + B_2\Delta d_3 + Cd_4 \equiv 0 \bmod p \tag{6}$$

where we define

$$d_i = f_0e_i - e_0f_i = f_0e_i - f_i \text{ for } i \in \{0, 1, 2, 3\}. \tag{7}$$

Note that $|d_i| \leq 3|e_i| + |f_i|$ and hence

$$|d_1|, |d_2|, |d_3| \leq 6\Delta \qquad \text{and} \qquad |d_4| \leq 9\Delta^2. \tag{8}$$

Using the definitions of B_0, B_1, B_2 and C, we find that

$$-(w_1 + b)d_1 + d_2(w_2 - w_0 - b) + d_3w_1 + d_4 \equiv 0 \bmod p,$$

and after the substitutions $w_i = u_i - \varepsilon_i$ we find

$$(d_3 - d_1)u_1 + d_2u_2 - d_2u_0 \equiv b(d_2 + d_1) + E \bmod p \tag{9}$$

where

$$E = -d_4 - \varepsilon_0d_2 + \varepsilon_1(d_3 - d_1) + \varepsilon_2d_2.$$

The bound (8) implies that $|E| \leq 33\Delta^2$. We now write this equality as a rational function of u_0. Setting

$$\Psi_1(u) = \frac{bu + a}{u} \qquad \text{and} \qquad \Psi_2(u) = \frac{(a + b^2)u + ab}{a + bu},$$

we have that $u_i = \Psi(u_0)$ for $i \in \{1, 2\}$. So (9) becomes

$$(d_3 - d_1)\Psi_1(u_0) + d_2\Psi_2(u_0) - d_2u_0 \equiv b(d_2 + d_1) + E \bmod p. \tag{10}$$

Let us consider the rational function

$$\Phi_{\mathbf{d}}(u) = (d_3 - d_1)\Psi_1(u) + d_2\Psi_2(u) - d_2u$$

corresponding to the left hand side of (10). Clearly, $\Phi_{\mathbf{d}}(u)$ can be written as the quotient of a polynomial of degree at most 3 and a polynomial of degree at most 2.

We assert that if $d_2 \neq 0$ or $d_3 - d_1 \neq 0$ then $\Phi_{\mathbf{d}}(u)$ is not a constant function. We prove the contrapositive implication. So assume that $\Phi_{\mathbf{d}}(u)$ is constant. Now $\Psi_1(u)$ is not constant, since $a \not\equiv 0 \bmod p$. So $\Psi_1(u)$ has a pole at 0 (and has no other poles). Similarly, $\Psi_2(u)$ is not constant and so has a pole at $-a/b$ (and no

other poles). The functions $\Psi_1(u)$ and $\Psi_2(u)$ have poles at distinct places and u has no finite poles at all, so the only way that $\Phi_{\mathbf{d}}(u)$ can be the constant function is if $d_2 \equiv 0 \bmod p$ and $d_3 - d_1 \equiv 0 \bmod p$. But our bounds (8) on the size of $\mathbf{d}$ now imply that $d_2 = 0$ and $d_3 - d_1 = 0$. This establishes our assertion about $\Phi_{\mathbf{d}}(u)$.

Suppose that $d_2 \neq 0$ or $d_3 - d_1 \neq 0$. Since $\Phi_{\mathbf{d}}(u)$ is a nonconstant quotient of two polynomials of degree at most 3, the congruence (10) can be satisfied for at most 3 values of u_0 once d_1, d_2, d_3 and E have been chosen. There are $O(\Delta)$ choices for each of d_1, d_2 and d_3, by (8). There are $O(\Delta^2)$ choices for E since $|E| \leq 33\Delta^2$. Hence there are only $O(\Delta^5)$ values of u_0 that satisfy some congruence of the form (10) where $\mathbf{d}$ and E satisfy the appropriate bounds. We place these $O(\Delta^5)$ values of u_0 in $\mathcal{V}(\Delta; a, b)$, and once this is done we see that the case when $d_2 \neq 0$ or $d_3 - d_1 \neq 0$ cannot occur (for then (10) would imply that $u_0 \in \mathcal{V}(\Delta; a, b)$).

This establishes the claim we made in the first paragraph of Stage 2, so we may assume that $d_2 = 0$ and $d_3 - d_1 = 0$.

Stage 3: Predicting the generator when $f_0 \neq 0$. Suppose that $f_0 \neq 0$. The definition (7) of d_2 shows that $0 = d_2 = f_0\varepsilon_1 - f_2$. Thus $\varepsilon_1 \equiv f_2/f_0 \bmod p$ and so we may compute the secret information ε_1. To obtain the remainder of the secret information, we note that the following three congruences hold:

$$\begin{aligned} a + b(\varepsilon_0 + w_0) &\equiv (\varepsilon_0 + w_0)(\varepsilon_1 + w_1) \bmod p, \\ a + b(\varepsilon_1 + w_1) &\equiv (\varepsilon_1 + w_1)(\varepsilon_2 + w_2) \bmod p, \\ f_0\varepsilon_0 - f_1 &\equiv f_0\varepsilon_2 - f_2 \bmod p. \end{aligned} \tag{11}$$

The first two of these congruences follow from (3), and the second follows from the fact that $d_1 = d_3$ together with the definition (7) of $\mathbf{d}$. But, since ε_1 is now known, the system (11) is linear in the variables a, ε_0 and ε_2. These equations have a unique solution if and only if $bf_0 \not\equiv 0 \bmod p$ (as can be seen by calculating the appropriate 3×3 determinant). Our assumption that $f_0 \neq 0$ together with our bound on $|f_0|$ shows that $f_0 \not\equiv 0 \bmod p$. Since $b \in \mathbb{F}_p^*$, we find that $bf_0 \not\equiv 0 \bmod p$ and so we may solve the system (11) to find ε_0, ε_2 and a. Finally, we compute u_0 from w_0 and ε_0 and so the algorithm terminates successfully in this case. So we are done when $f_0 \neq 0$.

Stage 4: Building the lattice $\mathcal{L}'$. We may now assume that $f_0 = 0$. So $d_i = -f_i$, $i = 1, 2, 3, 4$. We aim to show that b must have a special form.

The fact that $d_2 = 0$ and $d_3 - d_1 = 0$ means that the congruence (9) becomes

$$0 \equiv bd_1 + E \bmod p.$$

Using the definition of E, we find that $bd_1 \equiv d_4 \bmod p$, and so $bf_1 \equiv f_4 \bmod p$. It is easy to see that $f_1 \not\equiv 0 \bmod p$ (for the congruences $f_2 \equiv -d_2 \equiv 0 \bmod p$, $f_3 \equiv -d_3 \equiv -d_1 \equiv f_1 \bmod p$ and $f_4 \equiv bf_1 \bmod p$ would contradict the fact that $\mathbf{f}$ is a nonzero vector). Hence $b \equiv f_4/s_1 \bmod p$ and so we may write

$$b \equiv r/s \bmod p, \text{ where } r = f_4/\gcd(f_1, f_4) \text{ and } s = f_1/\gcd(f_1, f_4).$$

Note that r and s are coprime, $|r| \leq 3\Delta^2$ and $|s| \leq 3\Delta$. Moreover we know r and s since we have computed $\mathbf{f}$. Also note that r and s are determined by b, up to sign. To see this, suppose that r' and s' are coprime integers such that $|r'| \leq 3\Delta^2$, $|s'| \leq 3\Delta$ and $r'/s' \equiv b \equiv r/s \bmod p$. Then $rs' \equiv sr' \bmod p$ and since both rs' and sr' have absolute value at most $9\Delta^3$ we find that $rs' = sr'$. But since $\gcd(r,s) = \gcd(r',s') = 1$ we now find that $r = \sigma r'$ and $s = \sigma s'$ for some element $\sigma \in \{1, -1\}$.

We now consider a new lattice: the lattice $\mathcal{L}'$ consisting of solutions $\mathbf{x} = (x_0, x_1, x_2, x_3) \in \mathbb{Z}^4$ of the congruences

$$\begin{aligned} A'x_0 + B'x_1 + B_1'x_2 + C'x_4 &\equiv 0 \bmod p, \\ x_0 &\equiv 0 \bmod \Delta^3, \\ x_1 \equiv x_2 &\equiv 0 \bmod \Delta^2, \end{aligned} \tag{12}$$

where

$$A' \equiv sA\Delta^{-1} \bmod p, \quad B' \equiv sw_1\Delta^{-2} \bmod p,$$
$$B_1' \equiv s(w_2 - w_0)\Delta^{-2} \bmod p \quad \text{and} \quad C' \equiv 1 \bmod p.$$

It is easy to check that the lattice (12) contains the vector

$$\mathbf{e}' = \left(\Delta^3 e_0', \Delta^2 e_1', \Delta^2 e_2', e_3'\right),$$

where

$$\mathbf{e}' = \left(\Delta^3, \Delta^2(\varepsilon_2 - \varepsilon_0), \Delta^2\varepsilon_1, s\varepsilon_1(\varepsilon_2 - \varepsilon_0) - r(\varepsilon_1 - \varepsilon_0)\right).$$

We have

$$e_0' = 1, \qquad |e_1'| \leq 2\Delta, \qquad |e_2'| \leq \Delta, \qquad |e_3'| \leq 24\Delta^3$$

thus the Euclidean norm $\|\mathbf{e}'\|$ of $\mathbf{e}'$ satisfies the inequality

$$\|\mathbf{e}\| \leq \left(\Delta^6 + 4\Delta^6 + \Delta^6 + 576\Delta^6\right)^{1/2} \leq 25\Delta^3.$$

Stage 5: Defining the second exceptional set $\mathcal{V}'(\Delta; a, b)$. We now show that all short vectors in $\mathcal{L}'$ are parallel to $\mathbf{e}$ unless u_0 belongs to the set $\mathcal{V}'(\Delta; a, b)$ which we define below.

Assume, for a contradiction, that there is another vector

$$\mathbf{f}' = (\Delta^3 f_0', \Delta^2 f_1', \Delta^2 f_2', f_3') \in \mathcal{L}'$$

with $\|\mathbf{f}'\| \leq \|\mathbf{e}'\| \leq 25\Delta^3$ which is not parallel to $\mathbf{e}'$. The vector $\mathbf{d}'$ defined by $\mathbf{d}' = f_0'\mathbf{e}' - e_0'\mathbf{f}'$ lies in $\mathcal{L}'$ and has first component 0. Using the first congruence in (12), we find that

$$B'\Delta^2 d_1' + B_1'\Delta^2 d_2' + C'd_3' \equiv 0 \bmod p \tag{13}$$

where for $i \in \{1, 2, 3\}$ we define $d_i' = f_0'e_i' - e_0'f_i' = f_0'e_i' - f_i'$. Note that $|d_i'| \leq 25|e_i'| + |f_i'|$ and hence

$$|d_1'| \leq 75\Delta, \qquad |d_2'| \leq 50\Delta, \qquad |d_3'| \leq 25^2\Delta^3. \tag{14}$$

Using the definitions of B', B'_1 and C', we find that

$$sw_1 d'_1 + s(w_2 - w_0)d'_2 + d'_3 \equiv 0 \bmod p,$$

and after the substitutions $w_i = u_i - \varepsilon_i$ we find

$$u_1 s d'_1 + s(u_2 - u_0)d'_2 \equiv E' \bmod p \tag{15}$$

where

$$E' = -d'_3 + s\varepsilon_1 d'_1 - s(\varepsilon_0 - \varepsilon_2)d'_2.$$

The bounds (14) imply that $|E'| \leq 25^3 \Delta^3$. We now write this equality as a rational function of u_0. Then (15) becomes

$$sd'_1 \Psi_1(u_0) + sd'_2 \Psi_2(u_0) - sd'_2 u_0 \equiv E' \bmod p. \tag{16}$$

Let us consider the rational function

$$\Phi'_{\mathbf{d}'}(u) = sd'_1 \Psi_1(u) + sd'_2 \Psi_2(u) - d'_2 u$$

corresponding to the left hand side of (16). Clearly, $\Phi'_{\mathbf{d}'}(u)$can be written as the quotient of a polynomial of degree at most 3 and a polynomial of degree at most 2.

Now, $\Phi'_{\mathbf{d}'}(u)$ is a non-constant rational function of u. Suppose $\Phi'_{\mathbf{d}'}(u)$ is constant. Then (arguing as for $\Phi_{\mathbf{d}}(u)$ above) we must have that $d'_1 \equiv d'_2 \equiv 0 \bmod p$. But then (13) shows that $d'_3 \equiv 0 \bmod p$, and so our bounds (14) on the absolute value of d'_1, d'_2 and d'_3 imply that $d'_1 = d'_2 = d'_3 = 0$. This implies that $\mathbf{d}' = 0$ and so $\mathbf{e}'$ and $\mathbf{f}'$ are parallel. This contradicts our choice of $\mathbf{f}'$, and so we must have that $\Phi'_{\mathbf{d}'}(u)$ is a non-constant rational function of u.

Since $\Phi'_{\mathbf{d}'}(u)$ is of degree at most 3, the congruence (16) can be satisfied for at most 3 values of u_0 once s, d'_1, d'_2 and E' have been chosen. There are at most 2 choices for s (as s is determined up to sign by b). There are $O(\Delta)$ choices for each of d'_1, and d'_2, by (14). There are $O(\Delta^3)$ choices for E' since $|E'| \leq 25^3\Delta^3$. Hence there are only $O(\Delta^5)$ values of u_0 that satisfy some congruence of the form (16) where the d'_i and E' satisfy the appropriate bounds. We place these $O(\Delta^5)$ values of u_0 in $\mathcal{V}'(\Delta; a, b)$, and so we get a contradiction to our assumption that $\mathbf{f}'$ and $\mathbf{e}'$ are not parallel. So all short vectors in $\mathcal{L}'$ are parallel to $\mathbf{e}'$ whenever $u_0 \notin \mathcal{V}'(\Delta; a, b)$.

Stage 6: Predicting the generator for $f_0 = 0$. We apply a deterministic polynomial time algorithm for the shortest vector problem in a finite dimensional lattice to find a shortest nonzero vector $\mathbf{f}'$ in $\mathcal{L}'$, and this vector must be parallel to $\mathbf{e}'$. We recover $\mathbf{e}'$ by using the fact that $\mathbf{e}' = \mathbf{f}'/f'_0$. This gives us ε_1 which is used to calculate u_1. In order to compute u_0 we have to solve the following linear system of congruences in the unknowns ε_0 and ε_2:

$$\begin{aligned} f'_0(\varepsilon_2 - \varepsilon_0) &\equiv f'_1 \bmod p, \\ f'_0(s\varepsilon_1(\varepsilon_2 - \varepsilon_0) - r(\varepsilon_1 - \varepsilon_0)) &\equiv f'_3 \bmod p, \end{aligned} \tag{17}$$

which has a unique solution. Finally, a can be calculated by using the fact that $a \equiv u_0 u_1 - b u_0 \bmod p$. Defining

$$\mathcal{U}(\Delta; a, b) = \{0, -a/b\} \cup \mathcal{V}(\Delta; a, b) \cup \mathcal{V}'(\Delta; a, b)$$

which finishes the proof. □

4 Remarks and Open Questions

Obviously our result is nontrivial only for $\Delta = O(p^{1/5})$. Thus increasing the size of the admissible values of Δ (even at the cost of considering more consecutive approximations) is of prime importance.

One can presumably obtain a very similar result in the dual case, where a is given but the shift b is unknown.

As we have mentioned several other results about predictability of inversive and other nonlinear generators have recently been obtained in [2]. However, they are somewhat weaker than the present result because each of them excludes a certain small exceptional set of pairs of parameters (a, b). We believe that the approach of this work may help to eliminate this drawback. Certainly this question deserves further study.

We do not know how to predict the inversive (and other generators considered in [2]) in the case when the modulus p is secret as well. We remark that in the case of the linear congruential generator a heuristic approach to this problem has been proposed in [4]. However it is not clear how to extend this (even just heuristically) to the case of nonlinear generators.

References

1. M. Ajtai, R. Kumar and D. Sivakumar, 'A sieve algorithm for the shortest lattice vector problem', *Proc. 33rd ACM Symp. on Theory of Comput. (STOC 2001)*, Association for Computing Machinery, 2001, 601–610.
2. S. R. Blackburn, D. Gomez-Perez, J. Gutierrez and I. E. Shparlinski, 'Predicting nonlinear pseudorandom number generators', *Preprint*, 2003.
3. M. Grötschel, L. Lovász and A. Schrijver, *Geometric algorithms and combinatorial optimization*, Springer-Verlag, Berlin, 1993.
4. A. Joux and J. Stern, 'Lattice reduction: A toolbox for the cryptanalyst', *J. Cryptology*, **11** (1998), 161–185.
5. R. Kannan, 'Algorithmic geometry of numbers', *Annual Review of Comp. Sci.*, **2** (1987), 231–267.
6. R. Kannan, 'Minkowski's convex body theorem and integer programming', *Math. Oper. Res.*, **12** (1987), 415–440.
7. A. K. Lenstra, H. W. Lenstra and L. Lovász, 'Factoring polynomials with rational coefficients', *Mathematische Annalen*, **261** (1982), 515–534.
8. D. Micciancio and S. Goldwasser, *Complexity of lattice problems*, Kluwer Acad. Publ., 2002.

9. P. Q. Nguyen and J. Stern, 'Lattice reduction in cryptology: An update', in: W. Bosma (Ed), *Proc. ANTS-IV, Lect. Notes in Comp. Sci. Vol. 1838*, Springer-Verlag, Berlin, 2000, 85–112.
10. P. Q. Nguyen and J. Stern, 'The two faces of lattices in cryptology', in: J.H. Silverman (Ed), *Cryptography and Lattices Lect. Notes in Comp. Sci. Vol. 2146* , Springer-Verlag, Berlin, 2001, 146–180.
11. H. Niederreiter, 'New developments in uniform pseudorandom number and vector generation', in: H. Niederreiter and P.J. Shiue (Eds), *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, Lect. Notes in Statistics Vol. 106*, Springer-Verlag, Berlin, 1995, 87–120.
12. H. Niederreiter, 'Design and analysis of nonlinear pseudorandom number generators', in G.I. Schueller and P. D. Spanos (Eds) *Monte Carlo Simulation*, A.A. Balkema Publishers, Rotterdam, 2001, 3–9.
13. H. Niederreiter and I. E. Shparlinski, 'Recent advances in the theory of nonlinear pseudorandom number generators', in: K.-T. Fang, F.J. Hickernell and H. Niederreiter (Eds), *Proc. Conf. on Monte Carlo and Quasi-Monte Carlo Methods, 2000*, Springer-Verlag, Berlin, 2002, 86–102.
14. H. Niederreiter and I. E. Shparlinski, 'Dynamical systems generated by rational functions', in: Marc Fossorier, Tom Høholdt and Alain Poli (Eds), *Applied Algebra, Algebraic Algorithms and Error Correcting Codes – AAECC-15, Lect. Notes in Comp. Sci. Vol. 2643*, Springer-Verlag, Berlin, 2003, 6–17.