Public Key Identification Based on the Equivalence of Quadratic Forms

Rupert J. Hartung and Claus-Peter Schnorr

Johann Wolfgang Goethe Universität Frankfurt a. M.
Postfach 11 19 32; Fach 238
60054 Frankfurt a. M., Germany
{schnorr,hartung}@mi.informatik.uni-frankfurt.de

Abstract. The computational equivalence problem for quadratic forms is shown to be NP-hard under randomized reductions, in particular for indefinite, ternary quadratic forms with integer coefficients. This result is conditional on a variant of the Cohen-Lenstra heuristics on class numbers. Our identification scheme proves knowledge of an equivalence transform.

1 Introduction

The arithmetic theory of quadratic forms has a long history. Algorithmic problems on lattices and quadratic forms, however, have long been neglected; their study has been significantly pushed by the LLL-algorithm for lattice basis reduction [19]. Recently definite forms, or lattices, gave rise to cryptographic protocols related to the NP-hard problems of finding a shortest, respectively, closest lattice vector; see [22], [20] for hardness results and [2], [16], [17], [15], [14] for the applications. Cryptographic protocols based on NP-hard problems seem to withstand attacks by quantum computers. However, lattice cryptography requires lattices of high dimension. This yields long cryptographic keys and slow protocols.

By contrast, we show that quadratic form cryptography is possible in dimension three. We prove conditional NP-hardness of the equivalence and representation problems of indefinite, ternary forms over the integers using randomized reductions. We build on the work of Adleman and Manders [21] who proved NP-hardness of deciding solvability of inhomogeneous binary quadratic equations over the integers.

In Sect. 3, we present an identification scheme that proves knowledge of an equivalence transform of quadratic forms. This scheme is statistical zero-knowledge under reasonable heuristics. It allows short keys and performs merely one LLL-reduction and a few arithmetic steps per round, but its security requires many independent rounds.

2 The Equivalence Problem for Quadratic Forms

Quadratic Forms. An n-ary *quadratic form* (or simply *form*) f over $\mathbb{Z}$ is a homogeneous quadratic polynomial $f = \sum_{i,j=1}^{n} a_{i,j} x_i x_j = \boldsymbol{x}^t A \boldsymbol{x}$ with integer coef-

ficients $a_{i,j} = a_{j,i} \in \mathbb{Z}$, $A = (a_{i,j})$ and $\boldsymbol{x} = (x_1, ..., x_n)^t$. Then f has *determinant* $\det A$ and *dimension* n.

Equivalence classes. Let $f = \boldsymbol{x}^t A \boldsymbol{x}$ be an n-ary form. For $T \in \mathbb{Z}^{n \times n}$ let fT denote the form $\boldsymbol{x}^t T^t A T \boldsymbol{x}$. The forms f, fT are called *equivalent* if $T \in \mathrm{GL}_n(\mathbb{Z})$, i.e., if $|\det T| = 1$, notation $f \sim fT$. The equivalence class of f is simply called the *class* of f. Obviously $\det(fT) = (\det T)^2 \det f = \det f$.

Analogously, we define the equivalence of integral forms over $\mathbb{Z}_p$, the ring of p-adic integers. Two forms f, g over $\mathbb{Z}$ belong to the same *genus* if they are equivalent over $\mathbb{Z}_p$ for all primes p, and over $\mathbb{Z}_\infty = \mathbb{R}$. Clearly, every genus is a disjoint union of classes.

We study the equivalence problem of forms f having certain properties $\mathcal{P}$. Relevant properties are: f is *regular* if $\det f \neq 0$; f is *indefinite* if $f(x)$ takes both positive and negative values, otherwise f is *definite* (definite forms correspond to the Gram matrices $A = B^t B$ of lattice bases B); $f = \boldsymbol{x}^t A \boldsymbol{x}$ is *properly primitive* if $\gcd(a_{ii}, 2a_{ij} \mid i \neq j) = 1$; f is *isotropic* if $f(\boldsymbol{u}) = 0$ holds for some nonzero $\boldsymbol{u} \in \mathbb{Z}^n$, otherwise f is *anisotropic*. Every regular isotropic form is necessarily indefinite.

The computational equivalence problem, CEP

INPUT: equivalent forms f, g satisfying certain properties $\mathcal{P}$.
OUTPUT: $T \in \mathrm{GL}_n(\mathbb{Z})$ such that $g = fT$.

The concept of LLL-reduction [19] extends in a natural way from lattice bases and definite forms to anisotropic indefinite forms. LLL-reduced forms $f = \boldsymbol{x}^t A \boldsymbol{x}, A = (a_{i,j})$ satisfy $a_{1,1}^2 \leq 2^{2/n} \det A^{2/n}$. There is a polynomial time LLL-algorithm that transforms f into an LLL-form fT with $T \in \mathrm{GL}_n(\mathbb{Z})$, see [26], [18], and [25].

Outline. Section 3 presents an identification scheme based on the equivalence problem. Section 4 proves a variant of the equivalence problem to be NP-hard under randomized reductions. This is shown for indefinite, anisotropic forms of dimension $n = 3$ using a number theoretic assumption that guarantees class number 1 in real quadratic fields.

3 Identification Based on the Equivalence Problem

Key generation. Pick a random, indefinite, anisotropic, ternary LLL-form f_1 and a random $T \in \mathrm{GL}_3(\mathbb{Z})$ following **CT**. LLL-reduce $f_1 T$ to $f_0 := f_1 T T'$.
The *public key* is f_0, f_1, the *secret key* is $S := TT'$. S is uniquely determined by f_0, f_1 up to small automorphisms of f_0, f_1.

In the protocol $(\mathcal{P}, \mathcal{V})$ the prover $\mathcal{P}$ proves to the verifier $\mathcal{V}$ knowledge of S.

Identification scheme, $(\mathcal{P}, \mathcal{V})$

1. $\mathcal{P}$ computes an LLL-form $g := f_0 T$ via **CT**, and sends g,
2. $\mathcal{V}$ sends a random one-bit challenge $b \in_R \{0, 1\}$,
3. $\mathcal{P}$ sends $R := S^b T$, and $\mathcal{V}$ checks that $f_b R = g$.

CT: *Computation of* $T = (T_{i,j}) \in \mathrm{GL}_3(\mathbb{Z})$. Let $\|T\| = \max_{i,j} |T_{i,j}|$ be the *norm*. Set $r := 2^{100}\|S\|$. Pick the $T_{i,j} \in_R [-r, r]$ at random for $j \neq 1$. Compute $T_{1,1}, ..., T_{3,1} \in \mathbb{Z}$ by applying the extended Euclidean algorithm to $T_{1,1}^{adj}, ..., T_{3,1}^{adj}$ in order to achieve that $\det T = \pm 1$.

Note that $\det T = \sum_{i=1}^{3} \pm T_{i,1}\, T_{i,1}^{adj}$, where the $T_{i,1}^{adj}$ are values of homogeneous, quadratic polynomials in the $T_{i,j}$ with $j \neq 1$. Make sure that $gcd(T_{1,1}^{adj}, ..., T_{3,1}^{adj}) = 1$ by picking, if necessary, some new $T_{i,j}$. The Euclidean algorithm yields $|T_{i,1}| \leq \max_i |T_{i,1}^{adj}| \leq 4r^2$.

Finally LLL-reduce $f_0 T$ into $f_0 T T'$ and replace $T := TT'$. This balances the large $T_{i,1}$ with the smaller $T_{i,j}$, $j > 1$. The leading 100 bits and the least significant 100 bits of the $T_{i,j}$ are nearly random. Think of $f_0 T$ to be a random LLL-form out of a "sphere" of "radius" $\Theta(r^2)$ centered at f_0.

Completeness. The true prover $\mathcal{P}$ withstands the test $f_b R = g$.

Proof of knowledge. Consider a fraudulent $\widetilde{\mathcal{P}}$ that sends arbitrary $\bar{g}, \bar{R}$. The trivial $\widetilde{\mathcal{P}}$ guesses b in step 1 with probability $\frac{1}{2}$, sends the LLL-form $\bar{g} := f_b \bar{T}_b$ and the reply $\bar{R}_b := \bar{T}_b$. Then $\widetilde{\mathcal{P}}$ withstands the verification with probability $\frac{1}{2}$. The probability $\frac{1}{2}$ cannot be increased or else $\widetilde{\mathcal{P}}$ obtains an cryptographically equivalent secret key $S' := \bar{R}_0^{-1} \bar{R}_1$ such that $f_0 = f_1\, S'$.

Suppose an arbitrary $\widetilde{P}$ withstands the verification for the same $\bar{g}$ and both challenges $b = 0, 1$ replying $\bar{R}_b$. Then $\bar{R}_b$ transforms f_b into $\bar{g}$ and thus $f_0 \bar{R}_0^{-1} \bar{R}_1 = f_1$.
More precisely, letting $|(\widetilde{\mathcal{P}}, \mathcal{V})|$ bound the number of steps of $(\widetilde{\mathcal{P}}, \mathcal{V})$ we have:

Theorem 1. *A fraudulent prover* $\widetilde{P}$ *that withstands* k *independent executions of* $(\widetilde{\mathcal{P}}, \mathcal{V})$ *with probability* $\varepsilon > 2^{-k}$*, obtains an "equivalent" secret key in expected time* $|(\widetilde{\mathcal{P}}, \mathcal{V})|/(\varepsilon - 2^{-k})$.

Statistical zeroknowledge. The protocol $(\mathcal{P}, \mathcal{V})$ is by definition *statistical zeroknowledge* if for every probabilistic poly-time verifier $\widetilde{\mathcal{V}}$, there is a probabilistic poly-time simulator $\mathcal{S}$, which produces randomized strings whose distribution is statistical close to the communication of $(\mathcal{P}, \widetilde{\mathcal{V}})$ ($\|\cdot\|_1$-distance $\leq 2^{-100}$ suffices in practice). The simulator $\mathcal{S}$ has resettable black-box access to $\widetilde{\mathcal{V}}$ but does not know the secret key.

The simulator $\mathcal{S}$ mimics $\widetilde{\mathcal{P}}$ replying $\bar{R}_b = \bar{T}_b$ in step 3 whereas the true prover replies $R = S^b T$. The distributions of $S^b T, \bar{T}_b$ must be statistical close for both $b = 0, 1$.

Note that $S^b T$, $\bar{T}_b$ are determined, up to small automorphisms of f_0, f_1, as the isomorphisms from f_b to g, resp., from f_b to $\bar{g}$. Consider $g, \bar{g}$ as random LLL-forms out of spheres of radius $\Theta(r^2)$ centered at f_b, resp., $\bar{f}_b$. Small deviations

of r should have a negligible impact. In particular, we assume that ST reveals negligible information about S, and the LLL-reduction reveals merely an upper bound of $\|S\|$.

Theorem 2. *The identification scheme $(\mathcal{P}, \mathcal{V})$ is statistical zeroknowledge under reasonable heuristics.*

The zero-knowledge property extends to independent sequential executions of $(\mathcal{P}, \mathcal{V})$. Since $(\mathcal{P}, \mathcal{V})$ is restricted to one-bit challenges, a security level 2^{100} requires to run 100 independent executions of $(\mathcal{P}, \mathcal{V})$. To allow a poly-time simulator $\mathcal{S}$ these executions must be sequential. It is open to extend $(\mathcal{P}, \mathcal{V})$ to long challenges.

4 NP-Hardness

4.1 Introduction and Results

In this section, we prove randomized NP-completeness of a decisional variant of the **CEP**, as well as the related representation problem. We will use a number-theoretic assumption, the special Cohen-Lenstra Heuristics (sCLH), which is discussed in Sect. 4.2.

A quadratic form f of dimension n is said to *represent* a number $m \in \mathbb{Z}$ if there exists $u \in \mathbb{Z}^n \setminus \{0\}$ such that $f(u) = m$; here u is called a *representation* of m by f. The representation is said to be *primitive* if $\gcd(u_1, \ldots, u_n) = 1$. It is natural to ask for an algorithm which, given f and m, computes such a vector u. We will consider the following version of this problem.

Decisional Interval Representation Problem, DIRepr
PARAMETERS: Set $\mathcal{P}$ of properties of quadratic forms.
INPUT: $n \in \mathbb{N}$, quadratic form f of dimension n satifying all properties from $\mathcal{P}$, integer m, vectors $v, w \in (\mathbb{Z} \cup \{\pm\infty\})^n$, factorization of $\det f$.
DECIDE: Whether there is $x \in \mathbb{Z}^n$, $v_i \leq x_i \leq w_i$ for all i s. t. $f(x) = m$.

We can also define the computational problem **IRepr** with the same parameters and inputs as **DIRepr**, but where a representation in the given interval is to be computed. But then by a straightforward divide-and-conquer algorithm, **IRepr**$(\mathcal{P})$ and **DIRepr**$(\mathcal{P})$ are polynomial-time equivalent.

Note that if $v = w = (\infty, \ldots, \infty)^t$, then either a representation of polynomial size exists, or none at all, by [8]. We might have restricted the intervals in the definitions to be origin-symmetric, and to restrict at most one component, as will turn out from the proofs.

We denote by $\operatorname{gen} f$ the genus of a form f, by $\operatorname{cls} f$ its class, and by $\operatorname{cls}^+ f$ its *proper class*, i. e. the set of forms $f\,U$ with $\det U = +1$.

Theorem 3. *Let $M \in \mathbb{N}$. Let $\mathcal{P}'_M$ consist of the properties*

$$\dim f = 3, \qquad f \text{ indefinite anisotropic}, \qquad \operatorname{gen} f = \operatorname{cls}^+ f,$$

$$f \text{ properly primitive}, \qquad \text{and} \qquad (\det f, M) = 1$$

for a quadratic form f. If the sCLH holds true, then $\mathbf{DIRepr}(\mathcal{P}'_M)$ *is NP-hard under randomized reductions with one-sided error; more precisely:*

$$NP \subseteq RP^{\mathbf{DIRepr}(\mathcal{P}'_M)}.$$

For the security of our identification scheme, it is important that the extraction of the secret key from public parameters is not feasible; here that means that given equivalent forms f, g, the computation of an equivalence transformation $S \in \mathrm{GL}_n(\mathbb{Z})$ is hard. Again we consider interval constraints on the desired solution.

Decisional Interval Equivalence Problem, DIEP
PARAMETERS: Set $\mathcal{P}$ of properties of quadratic forms.
INPUT: $n \in \mathbb{N}$, n-ary quadratic forms f, g satifying all properties from $\mathcal{P}$, matrices $A, B \in (\mathbb{Z} \cup \{\pm\infty\})^{n \times n}$, factorization of $\det f$.
DECIDE: Whether there exists $T \in \mathrm{GL}_n(\mathbb{Z})$, $A_{ij} \leq T_{ij} \leq B_{ij}$ for all i, j s. t. $f\,T = g$.

As with representations, the problem of computing such a transformation is polynomial-time equivalent to **DIEP**, and both are NP-hard:

Theorem 4. *Let $M \in \mathbb{N}$. Let $\mathcal{P}'_M$ consist of the properties*

$$\dim f = 3, \qquad f \text{ indefinite anisotropic}, \qquad \operatorname{gen} f = \operatorname{cls}^+ f,$$
$$f \text{ properly primitive}, \qquad \text{and} \qquad (\det f, M) = 1$$

for a quadratic form f. If the sCLH holds true, then $\mathbf{DIEP}(\mathcal{P}'_M)$ *is NP-hard under randomized reductions with one-sided error; precisely:*

$$NP \subseteq RP^{\mathbf{DIEP}(\mathcal{P}'_M)}.$$

We now explain our assumption in detail and then give proof sketches, while some of its details are elaborated in the appendix.

4.2 The Special Cohen-Lenstra Heuristics

In [5], Cohen and Lenstra suggested a very general heuristic framework for the prediction of the average behavior of the class group of a number field K. Based on the thought experiment that, roughly speaking, all properties of class groups which are not determined a priori (e. g. by the factorization of the discriminant),

develop according to a certain random model, they obtain a corresponding very comprehensive conjecture on the distribution of such properties on large sets of discriminants. The CLH intends to give a convincing link between several seemingly independent observations from calculations of class groups. Though still unproven, it has enjoyed a vivid reception, and is thought of as a realistic way of thinking about long-run development of class numbers and groups.

One famous special case will be of central interest to us: Namely, the class numbers of real quadratic fields (cf. [6]). The empirical findings that large class numbers, in particular class numbers with odd part larger than one are rare, have been one of the central motivations to formulate these heuristics. We are going to use this empirically noticeable trait of class numbers. However, we cannot draw upon proven statements here as the conjecture is still wide open; in particular, our variant would imply that there are infinitely many real quadratic fields with class number 1, which is still unknown for number fields in general.

To explain our assumption, we first analyze different parts of the class group and the class number. Let d be an odd squarefree positive integer (for simplicity). Recall that the set $\mathfrak{F}(d)$ of proper classes of integral binary quadratic forms of determinant d forms a group under Gauß composition. The unit element is given by the *principal class*, i. e. the unique class of forms f_0 representing 1.

Let $\mathfrak{Cl}(d)$ be the ideal class group of the real quadratic number field $\mathbb{Q}[\sqrt{d}]$. Then $\mathfrak{Cl}(d) \cong \mathfrak{F}(d)/I$, where I is the subgroup of order 1 or 2 generated by the unique class which represents -1 (see [4, sec. 5.2]). Gauß [11] showed that $\mathfrak{F}(d)^2$ equals the genus of f_0, whence the 2-rank of $\mathfrak{F}(d)$ equals the number of genera, which is $2^{\omega(d)}$. This power of two constitutes the 'deterministic part' of the class number: It is determined by the prime factorization of d.

Beyond that part determined by genus theory, class numbers seem to behave 'randomly'; and essentially, Cohen's and Lenstra's idea was to formulate this impression explicitly and give a probabilistic model to describe the effects observed. In their original paper, however, they consider only the odd part $h^{\not 2}$ of the class number to avoid interference with the genus structure. But the total exclusion of the prime 2 now seems to be overly careful since only the index of $\mathfrak{Cl}^2(d)$ is linked to the genus structure, but not the 2-part of $|\mathfrak{Cl}^2(d)|$. It was conjectured in [12], [13] and, in contrast to the large remaining part of the heuristics, it was proven in [10, thm. 3] that the 2-part of $|\mathfrak{Cl}^2(d)|$ behaves as random as conjectured for the odd part of the class number.

We now want to assume, first put informally, that the Cohen-Lenstra Heuristics is compatible with the theorem on the 2-part still if restricted to primes of certain residue classes, and that this convergence is not too slow. Precisely, we state:

Special Cohen Lenstra Heuristics (sCLH) 41 *There are $c, e > 0$ and a polynomial F s. t. the following holds:*

Let $B > 0$ and primes $p_1, \ldots, p_k$ be given, where $k \leq e \log B$. Then

$$\#\{q \leq F(B) \,|\, q \text{ prime}, \left(\frac{q}{p_i}\right) = -1 \,\forall\, i; \quad |\mathfrak{Cl}(D(q))^2| = 1\} \geq c \frac{F(B)}{B \log F(B)}.$$

Here $D(q)$ denotes the fundamental discriminant corresponding to q.

It should be noted that the restriction to primes, and further to primes in specific residue classes, which is well prepared by tables as [27], already pops up in the original publication (see [5, §9, II. C12]) and is explicitly encouraged in [7, sec. 3].

4.3 Sketch of proofs

To bridge between the classical NP-complete problems and quadratic forms, we use the following problem on binary Diophantine equations.

MS Modular Square Problem
PARAMETER: $M \in \mathbb{N}$.
INPUT: Integers $a, b, c \in \mathbb{Z}$ with $c > 0$, a odd, squarefree, such that $(ab, M) = 1$ and there is an odd prime p s. t. if $u^2 | b$, then u is a power of p; factorization of b.
DECIDE: Whether there is $x \in \mathbb{Z}$, $|x| \leq c$ s. t. $x^2 \equiv a \bmod b$.

Proposition 1 is similar to the result by Adleman and Manders [21]; they proved it for arbitrary integers a, b in the problem instance.

We denote a deterministic Karp reduction by $\preccurlyeq$, whereas a probabilistic Karp reduction with one-sided error is depicted by $\preccurlyeq_r$.

Proposition 1. *Let $M \in \mathbb{N}$ be arbitrary. Then $3SAT \preccurlyeq_r \mathbf{MS}(M)$.*

This will be proven in the appendix.- Abbreviate the form $a_1 x_1^2 + \ldots + a_n x_n^2$ by $\langle a_1, \ldots, a_n \rangle$. From genus theory, the following can be derived:

Lemma 1.
Let $p \equiv 1 \bmod 4$ be a prime satisfying $\operatorname{cls} \langle 1, -p \rangle = \operatorname{gen} \langle 1, -p \rangle$. Let $m \in \mathbb{Z}$ be odd and satisfy $\left(\frac{m}{p}\right) = 1$ and $\left(\frac{q}{p}\right) = -1 \quad \forall\, q$ prime, $q | m$. (In particular, $p \nmid m$.) Then $\langle 1, -p \rangle$ represents m primitively.

By the theory of the spinor norm (see [24]) the following can be proven:

Proposition 2. *Let b be odd, p an odd prime, and $p \nmid b$. Then the form $f := \langle 2, b, -pb \rangle$ satisfies $\operatorname{gen} f = \operatorname{cls}^+ f$.*

Proof sketch of theorem 3: Let Φ' be an instance of 3SAT, i. e. a boolean formula in 3-CNF. Denote by $\varphi := |\Phi'|$ the binary length of Φ. Then by Proposition 1, Φ is randomly mapped to an instance of $\mathbf{MS}(M)$. For the resulting problem instance proceed as follows:

input: **MS**-instance (a, b, c).
answer := `false`;
repeat polynomially many times
  select random $k \in [0, b]$;
  $a' := a + kb$;
  repeat polynomially many times
    select random prime $p \equiv 1 \bmod 4$, $p > \max\left(\left\lceil \frac{c+|2a'+b|}{|b|} \right\rceil, |b|\right)$,
      and $\left(\frac{-2b}{p}\right) = -1$;

    ask oracle $\left(2a'+b, \begin{pmatrix} -c \\ -\infty \\ -\infty \end{pmatrix}, \begin{pmatrix} c \\ \infty \\ \infty \end{pmatrix}, \langle 2, b, -b\,p \rangle\right)$

    answer := answer $\vee$ (oracle answer)
return answer.

Obviously this establishes a polynomial-time oracle algorithm. Let us examine its correctness for solving $\mathbf{MS}(M)$. At first, note that if it returns `true` then there are $|x| \leq c$, $z_1, z_2 \in \mathbb{Z}$ s. t. $2\,x^2 + bz_1^2 - bpz_2^2 = 2a' + b$, hence, putting $y := z_1^2 - pz_2^2$, we have, in particular, that there are x, y s. t. $2x^2 + by = 2a' + b$ and thus $x^2 \equiv a' \equiv a \bmod b$ since 2 is invertible modulo the odd integer b. Thus, the $\mathbf{MS}(M)$ instance has a solution (x, y) and so is a 'yes'-instance.

Conversely, if the algorithm returns `false`, but nevertheless (a, b, c) is a 'yes'-instance, then there is $|x| \leq c$ s. t. $x^2 \equiv a \bmod b$; and thus there is $y \in \mathbb{Z}$, necessarily odd, s. t. $2x^2 + by = 2a' + b$, but y is not represented by any of the binary quadratic forms $\langle 1, -p \rangle$. For each of these forms, one of two things may have happened: Either y is represented by the genus of $\langle 1, -p \rangle$, but this genus consists of several classes; or y is not even represented by the genus of $\langle 1, -p \rangle$.

First, the sCLH 41 gives us an upper bound on the probability that the first case applies if the second does not. The second case, however, implies that

$$\forall\,(x, y) \in \mathbb{Z}^2,\ |x| \leq c,\ x^2 + by = a, \quad \exists\, q | y \text{ prime: } \left(\frac{q}{p}\right) \neq -1$$

by Lemma 45. As the q are odd, the symbol $\left(\frac{q}{p}\right)$ takes the values $1, -1$ according to the uniform distribution and independently for different q as p is randomly chosen; hence if (x, y) is any solution of the **MS**instance, the probability that the second case applies is bounded by $1 - 2^{-\omega(y)}$ (where $\omega(y)$ counts the number of distinct prime factors of y). We now have to show that if we start with a 'yes'-instance of $\mathbf{MS}(M)$, then with high probability, in some iteration we obtain an instance of (a', b, c) which has solution (x, y) with y decomposing into only logarithmically many prime factors in the input length. Observe that for all

solutions (x, y), $|y|$ is bounded from above by $2(b+1)$. Assume that (a, b, c) is a 'yes'-instance with some solution (x_0, y_0). Then, for $k = 0, \ldots, b$, the problem instance $(a' = a + kb, b, c)$ necessarily has a solution, namely $(x_0, y_0 + k)$. The range over which y varies thus is an interval $[y_0, y_0 + b] \cap \mathbb{Z}$, where $y_0 < 2b$. As follows directly from a result of Erdős and Nicolas [9, prop. 3], it holds that for $B > 0$,

$$\#\{Y \leq B \,|\, \omega(Y) > 2 \ln \ln B\} < \frac{6}{\pi^{5/2}} \frac{B}{(\ln B)^{2\ln(2)-1}\sqrt{\ln \ln B}} \left(1 + \mathcal{O}\left(\frac{1}{\ln \ln B}\right)\right). \tag{1}$$

Combining these insights, we conclude that the innermost *repeat* loop produces at most

$$\mathcal{O}\left(\frac{b}{(\ln b)^{2\ln(2)-1}\sqrt{\ln \ln b}}\right)$$

different a' for which there exists *no* solution (x, y) with y having less than $2 \ln \ln y$ prime factors. This implies that after $\log b$ iterations, we have seen at least one instance with a solution of few prime divisors with exponentially large probability.

Now that we have established the occurrence of at least one solution in which y_0 has few prime divisors with high probability, we may conclude that for every choice of p, the probability of failure according to case two is in each iteration independently bounded from above by

$$1 - 2^{-\omega(y_0)} \leq 1 - 2^{-2 \ln \ln y_0} \leq 1 - \frac{1}{\ln^2 \left\lceil \frac{c^2+|a|}{|b|} \right\rceil},$$

which after special treatment of finitely many instances is bounded away from 1. Together with the sCLH in the first case, we have bounded the error probability away from 1, and hence this is a one-sided error probabilistic reduction.

It remains to be shown that the forms constructed here satisfy all the properties entailed on them. Obviously, all forms constructed here are indefinite, of dimension 3, and of determinant prime to M. Next consider anisotropy: By [3, sec. 4.2] and the Hasse principle, a ternary quadratic form f over $\mathbb{Z}$ is isotropic if and only if $c_q f = 1$ for all symbols q (see [3, ch. 4] for the definition of c_p). But we have chosen $\left(\frac{-2b}{p}\right) = -1$, hence it can be computed that $c_p\langle 2, b, -pb\rangle = -1$, so that our forms are anisotropic. Finally, we have to establish the one-proper-class property for all forms constructed above. But this follows directly from Proposition 2. □

Proof sketch of theorem 4: Let f, c, m be taken from an instance of $\mathbf{IRepr}(\mathcal{P}'_M)$ produced in the proof of theorem 3 above. Construct a form g in the genus of f satisfying $g((1,0,0)^t) = m$. This can be accomplished essentially by following the proof of the existence of genera from [3]; the main steps are the following: First write down p-adic forms g_p with $g_p((1,0,0)^t) = m$, for all $p|2d\infty$. From values of the g_p, construct an integer t which is primitively represented over $\mathbb{Z}_p$ by all g_p; to

this end, we have to select a prime probabilistically from an arithmetic progression. Then we can compute $U_p \in \mathrm{GL}_3(\mathbb{Z}_p)$ such that $t\, f_p\, U_p = t^2 x_1^2 + f_p^*(x_2, x_3)$. Then the algorithm calls itself recursively and returns a form f^* which is $\mathbb{Z}_p$-equivalent to each f_p^*. Now we obtain the desired f by an application of the Chinese Remainder Theorem to the U_p. If this fails then we know we have started from a 'no'-instance.

If it outputs 'yes', then a matrix T with $fT = g$ and $|T_{11}| \leq c$ exists if and only if there are $|x| \leq c$, z_1, z_2 s. t. $f((x, z_1, z_2)) = m$.

Finally, if the algorithm terminates replying 'no' incorrectly, then we have hit a wrong orbit of representations of m by f. After polynomially many independent random iterations this only occurs with negligible probability.

Altogether this shows $\mathbf{DIRepr}(\mathcal{P}'_M) \preccurlyeq_r \mathbf{DIEP}(\mathcal{P}'_M)$. □

References

1. M. Agrawal, N. Kayal, and N. Saxena, *PRIMES is in P*, Annals of Mathematics **160** (2004), no. 2, 781–793.
2. M. Ajtai and C. Dwork, *A public-key cryptosystem with worst-case/average-case equivalence*, Proceedings of the 29th annual ACM symposium on theory of computing, El Paso, TX, USA, May 4-6, 1997 (New York), Association for Computing Machinery, 1997, pp. 284–293.
3. J. W. S. Cassels, *Rational quadratic forms*, London Mathematical Society Monographs, no. 13, Academic Press, London, New York, San Francisco, 1978.
4. H. Cohen, *A course in computational algebraic number theory*, Graduate Texts in Mathematics, vol. 138, Springer-Verlag, Berlin, 1993.
5. H. Cohen and H. W. Lenstra jun., *Heuristics on class groups of number fields*, Number Theory, Proc. Journ. arith., Noordwijkerhout 1983 (Berlin), Lecture Notes in Computer Science, no. 1068, Springer-Verlag, 1984, pp. 33–62.
6. H. Cohen and J. Martinet, *Class groups of number fields: Numerical heuristics*, Mathematics of Computation **48** (1987), no. 177, 123–137.
7. ———, *Heuristics on class groups: Some good primes are no too good*, Mathematics of Computation **63** (1994), no. 207, 329–334.
8. R. Dietmann, *Small solutions of quadratic Diophantine equations*, Proceedings of the London Mathematical Society, III. Ser. **86** (2003), no. 3, 545–582.
9. P. Erdős and J.-L. Nicolas, *Sur la fonction: Nombre de facteurs premiers de n*, ensMath2 **27** (1981), 3–27.
10. É. Fouvry and J. Klüners, *On the 4-rank of class groups of quadratic number fields*, preprint, 2006.
11. C. F. Gauß, *Untersuchungen über höhere Arithmetik (Disquisitiones Arithmeticae)*, Springer-Verlag, Berlin, 1889, German translation by H. Maser.
12. F. Gerth III, *The 4-class ranks of quadratic fields*, Inventiones Mathematicae **77** (1984), no. 3, 489–515.
13. ———, *Extension of conjectures of Cohen and Lenstra*, Expositiones Mathematicae **5** (1987), no. 2, 181–184.
14. O. Goldreich, Shafi Goldwasser, and S. Halevi, *Public-key cryptosystems from lattice reduction problems*, Advances in Cryptology - CRYPTO '97, 17th annual international cryptology conference. Santa Barbara, CA, USA. (Berlin) (B. S. jun.

Kaliski, ed.), Lecture Notes in Computer Science, vol. 1294, Springer-Verlag, 1997, pp. 112–131.
15. J. Hoffstein, N. Howgrave-Graham, J. Pipher, J. H. Silverman, and W. Whyte, *NTRUSign: Digital signatures using the NTRU lattice*, Topics in cryptology – CT-RSA 2003. The cryptographers' track at the RSA conference 2003, San Francisco, CA, USA, April 13–17, 2003 (Berlin) (M. Joye, ed.), Lecture Notes in Computer Science, vol. 2612, Springer-Verlag, 2003, pp. 122–140.
16. J. Hoffstein, J. Pipher, and J. H. Silverman, *NTRU: A ring-based public key cryptosystem*, Algorithmic number theory. 3rd international symposium, ANTS-III, Portland, OR, USA, June 21–25, 1998 (Berlin) (J. P. Buhler, ed.), Lecture Notes in Computer Science, vol. 1423, Springer-Verlag, 1998, pp. 267–288.
17. ———, *NSS: an NTRU lattice-based signature scheme*, Advances in cryptology - EUROCRYPT 2001. 20th international conference on theory and application of cryptographic techniques, Innsbruck, Austria, May 6-10, 2001 (Berlin) (B. Pfitzmann, ed.), Lecture Notes in Computer Science, vol. 2045, Springer-Verlag, 2001, pp. 211–228.
18. Gábor Ivanyos and Ágnes Szánto, *Lattice basis reduction for indefinite forms and an application*, Journal on Discrete Mathematics **153** (1996), no. 1–3, 177–188.
19. H. W. Lenstra jun., A. K. Lenstra, and L. Lovász, *Factoring polynomials with rational coefficients*, Mathematische Annalen **261** (1982), 515–534.
20. S. Khot, *Hardness of approximating the shortest vector problem in lattices*, Journal of the ACM **52** (2005), no. 5, 789–808.
21. K. L. Manders and L. M. Adleman, *NP-complete decision problems for binary quadratics*, Journal of Computer and System Sciences **16** (1978), 168–184.
22. D. Micciancio and Shafi Goldwasser, *Complexity of lattice problems: a cryptographic perspective*, The Kluwer International Series in Engineering and Computer Science, vol. 671, Kluwer Academic Publishers, Boston, Massachusetts, March 2002.
23. D. S. Mitrinović, J. Sándor, and B. Crstici (eds.), *Handbook of number theory*, Mathematics and Its Applications, vol. 351, Kluwer Academic Publishers, Dordrecht, 1996.
24. O. T. O'Meara, *Introduction to quadratic forms*, Grundlehren der mathematischen Wissenschaften, vol. 117, Springer-Verlag, Berlin, 1963, reprinted in 2000.
25. C.-P. Schnorr, *Progress on LLL and lattice reduction*, Proceedings LLL+25, Caen, France, June 29–July 1, 2007, 2007, To appear.
26. D. Simon, *Solving quadratic equations using reduced unimodular quadratic forms*, Mathematics of Computation **74** (2005), no. 251, 1531–1543.
27. M. Tennenhouse and H. C. Williams, *A note on the class-number one in certain real quadratic and pure cubic fields*, Mathematics of Computation **46** (1986), no. 173, 333–336.

A Proof of proposition 1

Let Φ be a Boolean formula in 3-CNF which contains each possible clause at most once, and no clause of Φ contains any variable both complemented and uncomplemented. Let ℓ be the number of variables in Φ. Choose an enumeration $\sigma_1, \ldots, \sigma_m$ of all clauses in the variables $x_1, \ldots, x_\ell$ with exactly three literals containing no variable both complemented and uncomplemented, such that both

the bijection $i \mapsto \sigma_i$ and its inverse are polynomial-time (e. g. a suitable lexicographic enumeration). Denote by $\sigma \in \Phi$ the assertion that clause σ occurs in Φ, and by $x_j \in \sigma$ ($\bar{x}_j \in \sigma$) that the j-th variable occurs uncomplemented (complemented) in clause σ. Let $n = 2m + \ell$.

For a fixed assignment to the boolean variables x_i, we define $r_i = 1$ if $x_i =$`true` and $r_i = 0$ otherwise. Moreover, for a clause σ, define

$$W(\sigma, r) = \sum_{i:x_i \in \sigma} r_i + \sum_{i:\bar{x}_i \in \sigma} (1 - r_i). \tag{2}$$

For $k = 1, \ldots, m$, let furthermore

$$R_k := \begin{cases} y_k - W(\sigma_k, r) + 1 & \text{if } \sigma_k \in \Phi, \\ y_k - W(\sigma_k, r) & \text{if } \sigma_k \notin \Phi, \end{cases} \tag{3}$$

where y_k are new variables, for $k = 1, \ldots, m$. Since Φ is in 3-CNF, we have $W(\sigma_k, r) = 0$ if assignment r does not render clause σ true, and $1 \le W(\sigma_k, r) \le 3$ otherwise. Hence the equation system

$$R_k = 0, \qquad k = 1, \ldots, m \tag{4}$$

has a solution with

$$r \in \{0, 1\}^\ell, \qquad y \in \{0, 1, 2, 3\}^m \tag{5}$$

if and only if Φ is satisfiable. Now choose a prime $p \ge 11$ not dividing the M from the statement of the theorem. As $-3 \le R_k \le 4$ for all choices (5) of the variables, (4) is equivalent with

$$\sum_{k=1}^{m} R_k p^k = 0. \tag{6}$$

We may estimate $\left|\sum_{k=1}^{m} R_k p^k\right| \le 4 \sum_{k=1}^{m} p^k < p^{m+1} - 2$ as $p \ge 11$; hence (6) is equivalent with $\sum_{k=1}^{m} R_k p^k \equiv 0 \bmod p^{m+1}$, or, equivalently, as p is odd, with

$$\sum_{k=1}^{m} (2\, R_k) p^k \equiv 0 \bmod p^{m+1}. \tag{7}$$

Now replace the y_k, $k = 1, \ldots, m$ and the r_i, $i = \ell$, by new variables α_i, $i = 1, \ldots, n$, each ranging independently over $\{1, -1\}$, by the formula

$$\begin{aligned} y_k &= \frac{1}{2}\big((1 - \alpha_{2k-1}) + 2((1 - \alpha_{2k})\big), \\ r_i &= \frac{1}{2}(1 - \alpha_{2m+i}), \end{aligned} \tag{8}$$

which obviously induces a bijection between the sets over which the two sequences of variables range.

After this change of variables the left hand side of (7) is still integral, and thus the congruence notation makes sense. Collecting terms, (7) can be rephrased as

$$\sum_{j=1}^{n} c_j \alpha_j \equiv \tau' \bmod p^{m+1} \tag{9}$$

for some $c_j, \tau' \in \mathbb{Z}$; explicitly, we have

$$\begin{aligned} -\tau' &= \sum_{k=1}^{m} (5 - \sum_{i:x_i \in \sigma_k} 1 + \mathbf{1}_{\sigma_k \in \Phi}) p^k, \\ c_{2k-1} &= -p^k, \\ c_{2k} &= -4p^k, \\ c_{2m+i} &= \sum_{k=1}^{m} (\mathbf{1}_{x_i \in \sigma_k} - \mathbf{1}_{\bar{x}_i \in \sigma_k}) p^k, \end{aligned} \tag{10}$$

where $k = 1, \ldots, m$, $i = 1, \ldots, \ell$, and $\mathbf{1}_\Psi = 1$ if Ψ is true and $\mathbf{1}_\Psi = 0$ otherwise.

Without affecting solvability or the number of solutions, we may as well introduce an extra variable α_0, define $c_0 := 1$ and $\tau := \tau' + 1$, and write

$$\sum_{j=0}^{n} c_j \alpha_j \equiv \tau \bmod p^{m+1}. \tag{11}$$

Thus we have learnt that Φ was satisfiable if and only if (11) is solvable for $\alpha \in \{-1, 1\}^{n+1}$. For later use, we verify that $p \not| \tau$: Indeed, the constant term (i. e. independent of the α_i) of $W(\sigma_k, r)$, with the r_i replaced according to (8), equals $w_k := \sum_{x_i \in \sigma_k} 1 + \sum_{\bar{x}_i \in \sigma_k} 1$. Now τ' is obviously divisible by p; and thus

$$\tau = \tau' + 1 \equiv 1 \bmod p. \tag{12}$$

Now define p_0 to be some prime exceeding $4 \cdot p^{m+1}(n+1)$, and let p_j be some prime exceeding p_{j-1}, for $j = 1, \ldots, n$, both of polynomial size in p^m. They can be found by sampling integers uniformly at random in intervals of the shape $[N,\, 2\,N(\ln N)^2]$ and testing them for primality [1].

Choose θ_j, for $j = 1, \ldots, n$, as the smallest positive integer satisfying

$$\theta_j \begin{cases} \equiv c_j & \bmod \quad p^{m+1}, \\ \equiv 0 & \bmod \quad \prod_{i \neq j} p_i, \\ \not\equiv 0 & \bmod \quad p_j. \end{cases} \tag{13}$$

Finally, set $K := \prod_{j=0}^{n} p_j$ and $c := \sum_{j=0}^{n} \theta_j$. Then we can reformulate (11) as follows: Φ is satisfiable if and only if there is $\alpha \in \{1, -1\}^{n+1}$ s. t.

$$\sum_{j=0}^{n} \theta_j \alpha_j \equiv \tau \bmod p^{m+1}, \tag{14}$$

and the number of solutions still has not changed.

Now we claim: For $x \in \mathbb{Z}$, $|x| \leq c$ and $c^2 \equiv x^2 \bmod K$ hold if and only if

$$x = \sum_{j=0}^{n} \theta_j \alpha_j \tag{15}$$

for some $\alpha \in \{1, -1\}^{n+1}$.

The proof of this claim is analogous to a lemma in [21] and therefore omitted.

Combining (15) and (14), we obtain that the $3SAT$ formula Φ has a satisfying truth assignments if and only if there is a number $x \in \mathbb{Z}$, $|x| \leq c$ s. t.

$$\begin{aligned} c^2 - x^2 &\equiv 0 \bmod K, \\ x &\equiv \tau \bmod p^{m+1}. \end{aligned} \tag{16}$$

It is easily seen that $(\tau - \xi)(\tau + \xi) = \tau^2 - \xi^2 \equiv 0 \bmod p^{m+1}$ is equivalent with $\xi \equiv \tau \bmod p^{m+1}$ or $\xi \equiv -\tau \bmod p^{m+1}$. So Φ has a satisfying truth assignment if and only if there is an integer x with $|x| \leq c$ s. t.

$$\begin{aligned} c^2 - x^2 &\equiv 0 \bmod K, \\ \tau^2 - x^2 &\equiv 0 \bmod p^{m+1}. \end{aligned} \tag{17}$$

By the Chinese Remainder Theorem, the equations (17) are jointly equivalent to the equation $p^{m+1}(c^2 - x^2) + K(\tau^2 - x^2) \equiv 0 \bmod p^{m+1}K$. But as K is prime to p by the construction of the p_j, and $p^{m+1} + K$ is prime to K, we finally reach the equation

$$x^2 \equiv a \bmod b \tag{18}$$

where

$$a \equiv (p^{m+1} + K)^{-1}(K\tau^2 + p^{m+1}c^2) \bmod p^{m+1}K \tag{19}$$

and $b = p^{m+1}K$. Then (18) has a solution $x \in \mathbb{Z}$ with $|x| \leq c$ if and only if Φ had a satisfying truth assignment. Now by construction, K is odd and squarefree, and a is odd and coprime to b. Now it suffices to note that the arithmetic progression (19) contains sufficiently many squarefree numbers so that one of them can be selected randomly in expected polynomial time. By the Page-Siegel-Walfisz theorem (see [23, §VIII.6]), polynomially many random selections suffice to find a prime a in the arithmetic progression (19), and primes can be efficiently tested as above. Of course, a is then squarefree as well. By construction, it is no problem to output the prime factorization as well.