Pythagoras number of quartic orders containing $\sqrt{2}$

Zilong HE and Yong HU

Abstract

Let $K$ be a quartic number field containing $\sqrt{2}$ and let $O \subseteq K$ be an order such that $\sqrt{2} \in O$. We prove that the Pythagoras number of $O$ is at most 5. This confirms a conjecture of Krárens ký, Raška and Sgallová. The proof makes use of Beli’s theory of bases of norm generators for quadratic lattices over dyadic local fields.

1 Introduction

Let $R$ be a commutative ring and let $\Sigma R^2$ denote the set of all elements that can be written as a sum of finitely many squares of elements of $R$. The Pythagoras number $P(R)$ of $R$ is defined as the smallest positive integer $m$ or $\infty$ such that every element in $\Sigma R^2$ can be written as a sum of at most $m$ squares of elements in $R$. The Pythagoras number of fields has been extensively studied in the literature. The case when $R$ is an order in a number field $K$ is another case of interest to many researchers. In this case, obtaining an upper bound for $P(R)$ is generally much more difficult than for $P(K)$. The interested readers are referred to [Pfi95, Chapter 7], [KRS21], [Tin21], [Kr´ a22] and the references therein for more information on this topic.

In a recent paper [KRS21], Krárens ký, Raška and Sgallová studies the Pythagoras number of the maximal order $O_K$ in a totally real biquadratic number field $K$. They prove, among others, that $P(O_K) \geq 5$ with at most seven exceptions and that $P(O_K) \leq 5$ if $\sqrt{5} \in O_K$. Based on results of numerical calculations with the help of computers and by a theoretical analysis of analogy with the case $\sqrt{5} \in O_K$, they conjecture that $P(O_K) \leq 5$ when $\sqrt{2} \in O_K$ ([KRS21, Conjecture 1.6 (2)]). The goal of this note is to confirm this conjecture. More precisely, we prove the following analog of [KRS21, Theorem 1.3]:

**Theorem 1.1.** Let $K$ be a quadratic extension of $\mathbb{Q}(\sqrt{2})$ and let $O \subseteq K$ be an order such that $\sqrt{2} \in O$. Then $P(O) \leq 5$.

A strategy for proving Theorem 1.1 has already been suggested in [KRS21, § 8.3]. The suggestion is mainly motivated by the fact that the genus of the quadratic form $I_4 = (1, 1, 1, 1)$ over $\mathbb{Z}[\sqrt{2}]$ consists of a single class, just as over the maximal order of $\mathbb{Q}(\sqrt{5})$. By a generalization of [KY21, Cor. 3.3] (see [KRS21, Prop. 7.5]), a key step
turns out to be the determination of the invariant $g_R(2)$ for $R = \mathbb{Z}[\sqrt{2}]$. (We will recall the definition of this invariant in § 4.) For the maximal order $R$ of the field $\mathbb{Q}(\sqrt{5})$, the invariant $g_R(2)$ is known from Sasaki’s work [Sas05]. When working with $\mathbb{Q}(\sqrt{2})$ instead of $\mathbb{Q}(\sqrt{5})$, the main difficulty results from the ramification of the prime number 2, which makes it less easy to analyze integral quadratic forms at the dyadic place.

To determine $g_R(2)$ for $R = \mathbb{Z}[\sqrt{2}]$, we need a characterization of binary forms that are representable as a sum of 4 linear forms over $R$. In §§ 2 and 3, we derive this missing ingredient from the representation theory of integral quadratic forms over general dyadic fields. This is made possible by the theory of bases of norm generators developed by Beli (see e.g. [Bel01] and [Bel19]). We feel that this method, though important and powerful, has not been widely used in the literature.

As expected, the result is $g_R(2) = 5$ for $R = \mathbb{Z}[\sqrt{2}]$ (Theorem 4.1). From this Theorem 1.1 follows by [KRS21, Prop. 7.5].

Terminology and notation. In the rest of the paper, we use the geometric language of quadratic spaces and lattices in the study of quadratic forms. We refer the reader to O’Meara’s book [O’M00] for standard terminology and notation about them. In particular, the scale and the norm of a quadratic lattice $M$ are denoted by $s(M)$ and $n(M)$ respectively.

Unless otherwise stated, all quadratic spaces and lattices are assumed to be nonsingular.

Given nonzero elements $a_1, \ldots, a_n$ in a field $F$, we denote by $[a_1, \ldots, a_n]$ the quadratic space defined by the diagonal quadratic form $a_1x_1^2 + \cdots + a_nx_n^2$ on the vector space $F^\oplus n := F \oplus \cdots \oplus F$. Similarly, when a Dedekind domain $R$ is fixed, let $\langle a_1, \ldots, a_n \rangle$ denote the quadratic lattice defined by the same quadratic form on the module $R^\oplus n$. For any integer $m \geq 1$, let $I_m$ denote the rank $m$ lattice $(1, \ldots, 1)$.

If $M, N$ are quadratic lattices over $R$, we write $N \twoheadrightarrow M$ if $N$ is represented by $M$. Similarly for representations of quadratic spaces.

## 2 A representability criterion over dyadic fields

The main result in this section, Theorem 2.5, is a criterion for the representability of a binary lattice as a sum of linear forms over a general dyadic local field. This result will be deduced from a general representation theorem in the theory of bases of norm generators, as developed by Beli in a series of papers [Bel01, Bel03, Bel06, Bel10, Bel19].

Let us briefly review some key definitions and facts that will be used in this paper. The reader is referred to Beli’s papers for any unexplained notation and definition.

Throughout this section, let $F$ be an arbitrary dyadic local field, i.e. a finite extension of the field $\mathbb{Q}_2$ of 2-adic numbers. Let $\text{ord} : F \to \mathbb{Z} \cup \{\infty\}$ denote the normalized discrete valuation on $F$. We write $\mathcal{O}_F$ for the valuation ring of $F$ and put $c = \text{ord}(2)$. For any $c \in F^\times := F \setminus \{0\}$, let $\mathcal{O}(c) = \bigcap_{x \in F}(c - x^2)\mathcal{O}_F$. The function

$$d : F^\times \longrightarrow \mathbb{N} \cup \{\infty\} ; c \mapsto d(c) := \min\{\text{ord}(c^{-1}z) | z \in \mathcal{O}(c)\}$$
is the called the order of relative quadratic defect ([Bel03, p.127, Definition 1]). It is well known that
\[ d(F^x) = \{0, 2e, \infty\} \cup \{1, 3, 5, \ldots, 2e - 1\} \, . \]

**Lemma 2.1.** With notation as above, we have \( d(-1) \geq e \).

**Proof.** From the definition, we have \( d(-1) = d(1 + (-2)) \subseteq 2\mathcal{O}_F \). Hence \( d(-1) \geq \text{ord}(2) = e \). \qed

(2.2) Let \( M \) be a quadratic \( \mathcal{O}_F \)-lattice with associated quadratic form \( Q \). A vector \( x \in M \) is called a norm generator of \( M \) if \( nM = Q(x)\mathcal{O}_F \). A sequence of vectors \( x_1, \ldots, x_m \) in \( FM \) is called a Basis Of Norm Generators (BONG) for \( M \) if \( x_1 \) is a norm generator for \( M \) and \( x_2, \ldots, x_m \) is a BONG for \( \text{pr}_{x_1^\perp} M \), where \( \text{pr}_{x_1^\perp} \) denotes the projection from \( FM \) to \( (Fx_1)^\perp \), the orthogonal complement of \( Fx_1 \) in \( FM \).

A BONG \( x_1, \ldots, x_m \) is said to be good, if \( \text{ord}(Q(x_i)) \leq \text{ord}(Q(x_{i+2})) \) for all \( 1 \leq i \leq m-2 \). By [Bel03, Corollary 2.6], a lattice is uniquely determined by a BONG. Also, every lattice possesses a good BONG (see [Bel03, Lemma 4.6] for a proof and [Bel06, §7] for an algorithm).

We will write \( M \cong \langle a_1, \ldots, a_m \rangle \) to mean that \( M \) has a good BONG \( x_1, \ldots, x_m \) such that \( a_i = Q(x_i) \). Using such a good BONG, we define the \( R \)-invariants
\[ R_i(M) := \text{ord}(a_i) = \text{ord}(Q(x_i)), \quad 1 \leq i \leq m \]
and the \( \alpha \)-invariants
\[ \alpha_i(M) := \min \left\{ \left\{ \frac{R_{i+1} - R_i}{2} + e \right\} \cup \left\{ R_{i+1} - R_j + d(-a_ja_{j+1}) \mid 1 \leq j \leq i \right\} \right. \\
\left. \cup \left\{ R_{j+1} - R_i + d(-a_ja_{j+1}) \mid i \leq j < m \right\} \right\} \]
for \( 1 \leq i \leq m - 1 \). These invariants are independent of the choice of the good BONG ([Bel10, Thm. 3.1]).

**Example 2.3.** Let \( m \) be a positive integer and \( M = I_m \). It is easy to see that \( M \cong \langle a_1, \ldots, a_m \rangle \) with \( a_1 = \cdots = a_m = 1 \). Note that \( e \leq d(-1) \) by Lemma 2.1. Hence the \( R \)-invariants and the \( \alpha \)-invariants of \( M \) are given by
\[ R_1 = R_2 = \cdots = R_m = 0 \quad \text{and} \quad \alpha_1 = \cdots = \alpha_{m-1} = \min\{e, d(-1)\} = e \, . \]

Let \( N \) be a binary \( \mathcal{O}_F \)-lattice and assume \( N \cong \langle b_1, b_2 \rangle \). We put \( S_i = R_i(N) \) and \( \beta_1 = \alpha_1(N) \). Then \( S_i = \text{ord}(b_i) \) and
\[ \beta_1 = \min \left\{ \frac{S_2 - S_1}{2} + e, S_2 - S_1 + d(-b_1b_2) \right\} \, . \]

(2.4) To a pair of \( \mathcal{O}_F \)-lattices \( M \) and \( N \) with \( n = \text{rank}(N) \leq m = \text{rank}(M) \), one can associate the \( A \)-invariants \( A_i(M, N) \) for \( 1 \leq i \leq \min\{m - 1, n\} \) as in [Bel06,
Definition 4.3]. We do not repeat the definition in the general case, but only give explicit formulas in a special case needed in this paper.

Let $N \cong \langle b_1, b_2 \rangle$ and $M = I_m$ with $m \geq 4$. With the same notation as in Example 2.3, assume further that $S_1 \geq 0$ and $S_1 + S_2 \geq 0$.

Then the two invariants $A_1 = A_1(M, N)$ and $A_2 = A_2(M, N)$ are given by

$$A_1 = \min \left\{ -\frac{S_1}{2} + e, -S_1 + e, -S_1 + d(-1) \right\} = e - S_1$$

and if $m = 4$,

$$A_2 = \min \left\{ -\frac{S_2}{2} + e, -S_2 + \min\{d(-b_1), \alpha_3, \beta_1\} \right\}$$

$$= \min \left\{ e - \frac{S_2}{2}, \min\{d(-b_1), e\} - S_2, e - \frac{S_1 + S_2}{2}, d(-b_1 b_2) - S_1 \right\}$$

$$= \min \left\{ d(-b_1) - S_2, e - S_2, e - \frac{S_1 + S_2}{2}, d(-b_1 b_2) - S_1 \right\}$$

or if $m \geq 5$,

$$A_2 = \min \left\{ -\frac{S_2}{2} + e, -S_2 + \min\{d(-b_1), \alpha_3, \beta_1\}, -S_1 - S_2 + \alpha_4 \right\}$$

$$= \min \left\{ d(-b_1) - S_2, e - S_2, e - \frac{S_1 + S_2}{2}, d(-b_1 b_2) - S_1, \alpha_4 - S_1 - S_2 \right\}$$

$$= \min \left\{ d(-b_1) - S_2, d(-b_1 b_2) - S_1, e - S_1 - S_2 \right\}.$$ 

We need the following special case of [Bel06, Theorem 4.5].

**Theorem 2.5.** Let $N \cong \langle b_1, b_2 \rangle$ and $M = I_m$ with $m \geq 4$ as in (2.4).

Then $N \rightarrow M$ if and only if $F N \rightarrow F M$ and the following conditions hold:

1. $S_1 \geq 0$ and $S_1 + S_2 \geq 0$.
2. $d(-b_1) \geq e - S_1$ and $d(-b_1 b_2) \geq e - S_2$.
3. If $m = 4$, $S_2 < 0$, $d(-b_1) + d(-b_1 b_2) > 2e + S_2$ and $2d(-b_1 b_2) > 2e + S_1$, then the binary space $[b_1, b_2]$ is represented by the ternary space $[1, 1, 1]$.

**Proof.** First, it is easy to see that in our situation condition (i) in [Bel06, Theorem 4.5] is the same as our condition (1). Condition (ii) of [Bel06, Theorem 4.5] reads

$$\min\{d(b_1), \alpha_1, \beta_1\} \geq A_1 \quad \text{and} \quad \min\{d(b_1 b_2), \alpha_2\} \geq A_2.$$ 

As we have seen in (2.4), we have $\alpha_1 = \alpha_2 = e$ and $A_1 = e - S_1$. Assuming (1), we have $\alpha_1 \geq A_1$ and

$$A_2 = \min \left\{ d(-b_1) - S_2, d(-b_1 b_2) - S_1, e - S_1 - S_2 \right\} \leq e - S_1 - S_2 \leq \alpha_2 = e.$$ 

4
Thus, the first inequality in (2.5.1) is equivalent to the two inequalities $d(b_1) \geq e - S_1$ and $\beta_1 = \min\left\{ \frac{S_2 - S_1}{2} + e, S_2 - S_1 + d(-b_1 b_2) \right\} \geq e - S_1$. Since $S_1 + S_2 \geq 0$, we have $\beta_1 \geq e - S_1$ if and only if $d(-b_1 b_2) \geq e - S_2$.

Since $d(-1) \geq e \geq \max\{A_2, e - S_1\}$, by the domination principle for the function $d$ ([Bel03, Lemma 1.1]), the inequalities $d(b_1) \geq e - S_1$ and $d(b_1 b_2) \geq A_2$ are equivalent to $d(-b_1) \geq e - S_1$ and $d(-b_1 b_2) \geq A_2$ respectively. So (2.5.1) holds if and only if

$$d(-b_1) \geq e - S_1, \quad d(-b_1 b_2) \geq e - S_2 \quad \text{and} \quad d(-b_1 b_2) \geq A_2.$$ 

In fact, the first two inequalities here imply the third, because they imply that

$$A_2 = \min \left\{ d(-b_1) - S_2, d(-b_1 b_2) - S_1, e - S_1 - S_2 \right\} = e - S_1 - S_2 \leq e - S_2.$$ 

Thus, we see that condition (ii) of [Bel06, Theorem 4.5] can be translated into condition (2) in our theorem (when assuming (1)).

Note that the inequality $0 > S_1$ does not hold by (1). By [Bel20, p.6, Remarks (1)], condition (iii) of [Bel06, Theorem 4.5] can be rephrased as follows: If $S_2 < 0$ and $d[-a_{13} b_{11}] + d[-a_{14} b_{12}] > 2e + S_2$, then $[b_1, b_2] \rightarrow [1, 1, 1]$, where in our context

$$d[-a_{13} b_{11}] = \min \left\{ d(-b_1), e, e + \frac{S_2 - S_1}{2}, S_2 - S_1 + d(-b_1 b_2) \right\}$$ 

and

$$d[-a_{14} b_{12}] = \begin{cases} d(-b_1 b_2) & \text{if } m = 4, \\ \min\{d(-b_1 b_2), e\} & \text{if } m \geq 5. \end{cases}$$ 

In the case $S_2 < 0$, from (1) and (2) we get $e + \frac{S_2 - S_1}{2} < e$ and $d(-b_1 b_2) \geq e - S_2 > e$, whence

$$d[-a_{13} b_{11}] = \min \left\{ d(-b_1), e + \frac{S_2 - S_1}{2}, S_2 - S_1 + d(-b_1 b_2) \right\}$$ 

and

$$d[-a_{14} b_{12}] = \begin{cases} d(-b_1 b_2) & \text{if } m = 4, \\ e & \text{if } m \geq 5. \end{cases}$$ 

Note that $S_1 + S_2 \geq 0$ by (1). Thus, if $S_2 < 0$ and $m \geq 5$, we have

$$d[-a_{13} b_{11}] + d[-a_{14} b_{12}] \leq e + \frac{S_2 - S_1}{2} + e = 2e + S_2 - \frac{S_1 + S_2}{2} \leq 2e + S_2.$$ 

So there is no need to check condition (iii) of [Bel06, Theorem 4.5] if $m \geq 5$.

If $S_2 < 0$ and $m = 4$, the inequality $d[-a_{13} b_{11}] + d[-a_{14} b_{12}] > 2e + S_2$ means

$$\min \left\{ d(-b_1), e + \frac{S_2 - S_1}{2}, S_2 - S_1 + d(-b_1 b_2) \right\} + d(-b_1 b_2) > 2e + S_2.$$ 

If $S_2 - S_1 + d(-b_1 b_2) + d(-b_1 b_2) > 2e + S_2$, we have $d(-b_1 b_2) > e + \frac{S_2}{2}$ and thus

$$e + \frac{S_2 - S_1}{2} + d(-b_1 b_2) > e + \frac{S_2 - S_1}{2} + e + \frac{S_1}{2} = 2e + \frac{S_2}{2} > 2e + S_2.$$
From this we see that condition (iii) of [Bel06, Theorem 4.5] is equivalent to our condition (3).

Since the inequality \( R_{i+2} > R_{i+1} + 2e \) does not hold in our context, there is no need to check condition (iv) of [Bel06, Theorem 4.5]. The theorem is thus proved. \(\square\)

**Corollary 2.6.** Let \( N \cong b_1, b_2 \succ \) and \( S_i = \text{ord}(b_i) \).

Then the following assertions are equivalent:

1. \( N \rightarrow I_5 \).
2. \( N \rightarrow I_m \) for some \( m \geq 2 \).
3. \( N \rightarrow I_m \) for some \( m \geq 5 \).
4. The following two conditions hold:
   
   \[
   \begin{align*}
   (a) & \quad S_1 \geq 0 \text{ and } S_1 + S_2 \geq 0. \\
   (b) & \quad d(-b_1) \geq e - S_1 \text{ and } d(-b_1b_2) \geq e - S_2.
   \end{align*}
   \]

**Proof.** Trivially, (1)⇒(2). Since \( I_m \rightarrow I_{m+3} \), we have (2)⇒(3). A quadratic space of dimension at least 5 represents all binary spaces ([O’M00, 63:21]). So the condition \( FN \rightarrow FM \) holds automatically if \( M = I_m \) with \( m \geq 5 \). Thus, from Theorem 2.5 we see that (3)⇒(4)⇒(1). \(\square\)

**Remark 2.7.** To test representability by \( I_m \) over any dyadic local field, one can also use the Third Main Theorem in Riehm’s work [Rie64], which relies only on the classical invariants of lattices as presented in [O’M00]. However, we feel that our criterion in Theorem 2.5 (especially in the case \( m = 4 \)), obtained by using the theory of BONGs, is more convenient for our purpose.

## 3 Sums of 4 squares of linear forms over \( \mathbb{Z}_2[\sqrt{2}] \)

In this section, let \( F = \mathbb{Q}_2(\sqrt{2}) \).

(3.1) Let us recall some useful facts about the field \( F \).

Clearly, a uniformizer in \( F \) is \( \sqrt{2} \) and the ramification index \( e = \text{ord}(2) = 2 \). Note that \(-1 = (1 + \sqrt{2})^2 - (4 + 2\sqrt{2}) \) and \( \text{ord}(4 + 2\sqrt{2}) = 3 < 4 = 2e \). So we have \( d(-1) = 3 \) by [O’M00, 63:5].

Since \([F : \mathbb{Q}_2]\) is even, \(-1\) is a sum of two squares in \( F \) ([Pfl95, Chapter 3, 1.2(6)]). Thus, the ternary space \([1, 1, 1]\) is isotropic and the quaternary space \([1, 1, 1, 1]\) is hyperbolic over \( F \). In particular, \([1, 1, 1, 1]\) represents all binary spaces over \( F \).

The unique quadratic unramified extension of \( F \) is \( F(\sqrt{5}) \). So by [O’M00, 63:3 and 63:4], an element \( c \in F^\times \) satisfies \( d(c) = 4 = 2e \) if and only if \( c \in 5F^\times^2 \), i.e. \( c/5 \) is a square in \( F \).
Lemma 3.2. Let $Q$ be a binary quadratic form over $\mathcal{O}_F$. Suppose that $Q$ is a sum of finitely many squares of linear forms over $\mathcal{O}_F$. Then $Q$ is not a sum of 4 squares of linear forms over $\mathcal{O}_F$ if and only if $Q$ is equivalent to the form $G(x, y) := 2\sqrt{2}(x^2 + xy + y^2)$.

Proof. Let $N$ be the lattice defined by $Q$ on the module $\mathcal{O}_F \oplus \mathcal{O}_F$. Assume $N \cong \langle b_1, b_2 \rangle$ and put $S_i = \text{ord}(b_i)$. As we have mentioned in (3.1), $FN$ is represented by $[1, 1, 1, 1]$. Here we have assumed that $N$ is represented by $I_m$ for some (large) $m$. So by Theorem 2.5 (and Corollary 2.6), $N$ is not represented by $I_4$ if and only if the following conditions hold:

1. $S_1 \geq -S_2 > 0$.
2. $d(-b_1) \geq 2 - S_1$ and $d(-b_1 b_2) \geq 2 - S_2$.
3. $(d(-b_1) + d(-b_1 b_2)) > 4 + S_2$.
4. $2d(-b_1 b_2) > 4 + S_1$.
5. The binary space $[b_1, b_2]$ is not represented by $[1, 1, 1]$.

We claim that the above 5 conditions are equivalent to

$$S_1 = 3, \quad S_2 = -1 \quad \text{and} \quad d(-b_1 b_2) = 4.$$ 

We will use the fact that $d(F^\times) = \{0, 4, \infty, 1, 3\}$. As we have said in (3.1), the equality $d(-b_1 b_2) = 4$ means that $-b_2 \in 5b_1 F^\times$. When $S_1 = 3$, this implies that the Hilbert symbol $(-b_1, -b_2)_F$ is equal to $(-b_1, 5)_F = -1 \ (\text{[O'M00, 63:11a]}).$ But the binary space $[b_1, b_2]$ is represented by $[1, 1, 1] \cong [b_1, -b_1, -1]$ if and only if the Hilbert symbol $(-b_1, -b_2)_F$ equals 1. From this the sufficiency part of the claim follows easily.

Now let us prove the necessity part of the claim.

If $S_1 = 1$, then we have $d(-b_1) = 0$, contradicting the first inequality in (2). So we must have $S_1 \geq 2$. This combined with (4) yields $d(-b_1 b_2) > 3$. On the other hand, since $[1, 1, 1]$ is isotropic, condition (5) implies that $-b_1 b_2$ is not a square in $F$, or equivalently $d(-b_1 b_2) < \infty$. Hence $d(-b_1 b_2) = 4$.

From the second inequality in (2) we deduce that $S_2 \geq -2$, and from (4) we see $S_1 < 4$. We have already shown $S_1 \geq 2$. Thus by (1) we have $S_2 \in \{-1, -2\}$ and $S_1 \in \{2, 3\}$. As we mentioned in the proof of sufficiency, the condition $d(-b_1 b_2) = 4$ implies that $(-b_1, -b_2)_F = (-b_1, 5)_F$. If $S_1 = 2$, the Hilbert symbol $(-b_1, 5)_F$ equals 1, which leads to a contradiction to (5). Therefore, $S_1 = 3$.

If $S_2 = -2$, then $S_1 + S_2$ is odd, which implies $d(-b_1 b_2) = 0$, a contradiction. So we have $S_2 = -1$. This proves our claim.

Now, using [Bel03, Cor. 3.4 (iii)] we can conclude that $N$ has scale $s(N) = \sqrt{2}\mathcal{O}_F$ and norm $n(N) = 2\sqrt{2}\mathcal{O}_F$. Note that the space $FN$ is anisotropic. By [O'M00, 93:11], $N$ is isomorphic to the lattice represented by the matrix $\sqrt{2}A(2, 2) = \begin{pmatrix} 2\sqrt{2} & \sqrt{2} \\ \sqrt{2} & 2\sqrt{2} \end{pmatrix}$.

This is exactly the binary lattice corresponding to the binary form $2\sqrt{2}(x^2 + xy + y^2)$. The lemma is thus proved. $\square$
Proof of main result

For any commutative ring $R$ and any positive integer $k$, the invariant $g_R(k)$ is defined as the smallest positive integer $n$ (or $\infty$) such that every $k$-ary quadratic form that can be written as a finite sum of squares of linear forms over $R$ is a sum of $n$ squares of linear forms.

As we have said in the introduction, Theorem 1.1 follows from the following result, which is an analog of a theorem of Sasaki [Sas05] (see also [KRS21, Thm. 7.7]).

Theorem 4.1. We have $g_{\mathbb{Z}[\sqrt{2}]}(2) = 5$.

Proof. The binary form $f(x, y) := (16 + 2\sqrt{2})(x^2 + xy + y^2)$ is a sum of linear forms over $\mathbb{Z}[\sqrt{2}]$, in view of the identity

$$f = ((1 + \sqrt{2})x + y)^2 + (1 + \sqrt{2})y^2 + (x + y)^2 + 3(2x + y)^2 + 8y^2.$$

Over $\mathbb{Z}_2[\sqrt{2}]$ the form $f$ is equivalent to the form $G$ in Lemma 3.2, because $\frac{16 + 2\sqrt{2}}{2\sqrt{2}} = 1 + 4\sqrt{2}$ is a square in $\mathbb{Z}_2[\sqrt{2}]$ ([O’M00, 63:1]). This proves the inequality $g_{\mathbb{Z}[\sqrt{2}]}(2) \geq 5$.

To prove the inequality in the other direction, consider a binary quadratic form $Q(x, y)$ over $\mathbb{Z}[\sqrt{2}]$ and suppose it can be written as a finite sum $Q = \sum L_i^2$ where $L_i = a_i x + b_i y$ with $a_i, b_i \in \mathbb{Z}_2[\sqrt{2}]$. Over a non-dyadic completion, every binary form is a sum of four linear forms (see e.g. [HHX22, Prop. 3.3]). Since the $\mathbb{Z}[\sqrt{2}]$-lattice $I_4$ has class number 1 ([Dze60, p.272, Satz 24]), $Q$ is a sum of 4 squares of linear forms over $\mathbb{Z}[\sqrt{2}]$ if and only if it is so over $\mathbb{Z}_2[\sqrt{2}]$. So by Lemma 3.2, we may assume $Q$ is equivalent to the form $G = 2\sqrt{2}(x^2 + xy + y^2)$ over $\mathbb{Z}_2[\sqrt{2}]$. This means that $Q$ represents only elements that are divisible by $2\sqrt{2}$.

If all the $L_i^2$ represent only elements divisible by $2\sqrt{2}$, then $a_i^2$ and $b_i^2$ are divisible by $2\sqrt{2}$, hence 2 divides $a_i$ and $b_i$. But this would imply that the integral values represented by $Q$ are all divisible by 4. This contradicts the fact that $2\sqrt{2}$ is represented by $Q$ over $\mathbb{Z}_2[\sqrt{2}]$. Hence, one of $L_i^2$ represents an element not divisible by $2\sqrt{2}$. Then $Q - L_i^2$ is a sum of 4 linear forms over the dyadic completion by Lemma 3.2. Since $I_4$ has class number 1, this shows that $Q$ is a sum of 5 squares of linear forms over $\mathbb{Z}[\sqrt{2}]$.

Acknowledgements. We thank Prof. Fei Xu for helpful discussions. The authors were supported by a grant from the National Natural Science Foundation of China (no. 12171223) and the Guangdong Basic and Applied Basic Research Foundation (no. 2021A1515010396).

References

[Bel01] C. N. Beli. Integral spinor norm groups over dyadic local fields and representations of quadratic lattices. ProQuest LLC, Ann Arbor, MI, 2001. Thesis (Ph.D.)–The Ohio State University.

[Bel03] C. N. Beli. Integral spinor norm groups over dyadic local fields. J. Number Theory, 102(1):125–182, 2003.
[Bel06] C. N. Beli. Representations of integral quadratic forms over dyadic local fields. Electron. Res. Announc. Amer. Math. Soc., 12:100–112, 2006.

[Bel10] C. N. Beli. A new approach to classification of integral quadratic forms over dyadic local fields. Trans. Amer. Math. Soc., 362(3):1599–1617, 2010.

[Bel19] C. N. Beli. Representations of quadratic lattices over dyadic local fields. preprint available at arXiv:1905.04552, 2019.

[Bel20] C. N. Beli. Universal integral quadratic forms over dyadic local fields. preprint available at arXiv:2008.10113, 2020.

[Dze60] J. Dzewas. Quadratsummen in reellquadratischen Zahlkörpern. Math. Nachr., 21:233–284, 1960.

[HHX22] Z. He, Y. Hu, and F. Xu. On indefinite k-universal quadratic forms over number fields. arXiv:2201.10730, 2022.

[Krá22] J. Krášenský. A cubic ring of integers with the smallest Pythagoras number. Arch. Math. (Basel), 118(1):39–48, 2022.

[KRS21] J. Krášenský, M. Raška, and E. Sgallová. Pythagoras numbers of orders in biquadratic fields. available at arXiv:2105.08860v2, 2021.

[KY21] V. Kala and P. Yatsyna. Lifting problem for universal quadratic forms. Adv. Math., 377:107497, 2021.

[O’M00] O. T. O’Meara. Introduction to quadratic forms. Classics in Mathematics. Springer-Verlag, Berlin, 2000. Reprint of the 1973 edition.

[Pfi95] A. Pfister. Quadratic forms with applications to algebraic geometry and topology, volume 217 of London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1995.

[Rie64] C. Riehm. On the integral representations of quadratic forms over local fields. Amer. J. Math., 86:25–62, 1964.

[Sas05] H. Sasaki. Sums of squares of totally positive definite quadratic forms over real quadratic field $\mathbb{Q}(\sqrt{5})$ (in Japanese). Otemae Junior College Research Bulletin, 25, 2005.

[Tin21] M. Tinková. On the pythagoras numbers of the simplest cubic fields. available at arXiv:2101.11384, 2021.

Contact information of the authors:

Zilong HE
Department of Mathematics
Southern University of Science and Technology
Shenzhen 518055, China
Email: hezl6@sustech.edu.cn

Yong HU
Department of Mathematics
Southern University of Science and Technology
Shenzhen 518055, China
Email: huy@sustech.edu.cn