Non-applicability of the Gaborit&Aguilar-Melchor patent to Kyber and Saber

Vadim Lyubashevsky\textsuperscript{1} and Damien Stehlé\textsuperscript{2,3}

\textsuperscript{1} IBM Research Europe – Zurich, Switzerland
\textsuperscript{2} ENS de Lyon, LIP (U. Lyon, CNRS, ENSL, INRIA, UCBL), France
\textsuperscript{3} Institut Universitaire de France

08 October 2021

Abstract. In the context of the NIST post-quantum cryptography project, there have been claims that the Gaborit&Aguilar-Melchor patent could apply to the Kyber and Saber encryption schemes. In this short note, we argue that these claims are in contradiction with the potential validity of the patent.

1 Introduction

In 2015, NIST announced its intention to standardize post-quantum cryptographic primitives (encryption schemes, key exchange mechanisms and signatures). For this purpose, it set up a post-quantum cryptography project,\textsuperscript{4} based on submissions of candidate schemes. At the time this note is written, we are at the third round of selection, with 7 finalists and 8 so-called alternates. As the project moved forward, the question of applicability of patents to candidates became more pertinent. This note concerns the Gaborit&Aguilar-Melchor patent [GAM10] owned by CNRS, and claims about its applicability to Kyber [BDK\textsuperscript{+}18, ABD\textsuperscript{+}21] and Saber [DKRV18, BMD\textsuperscript{+}20]. In the rest of the note, we restrict the discussion to Kyber, as the differences between Kyber and Saber are irrelevant for the question under scope.

CNRS, which owns the patent, has made its position available online.\textsuperscript{5} (As this webpage changes over time, we provide its current version in appendix.) Although it does not mention the Gaborit&Aguilar-Melchor patent nor Kyber and Saber explicitly, no other CNRS-owned patent is known that would apply to the NIST project third round finalists. CNRS could possibly claim rights for the NTRU LPRime scheme of the NTRU Prime alternate [BBC\textsuperscript{+}20]: we do not cover the case of NTRU LPRime in this note. CNRS could possibly claim rights for the BIKE [ABB\textsuperscript{+}21] and HQC [MAB\textsuperscript{+}21] alternates, but it has lifted its intellectual property claims for these.\textsuperscript{6} This targeting of the lattice-based candidates was confirmed by Dustin Moody in an invited talk at the PQCrypto conference.\textsuperscript{7} Finally, the threat of this patent to Kyber and Saber was also mentioned in the documentation of the NTRU Prime candidate [BBC\textsuperscript{+}20].

Contribution. In this note, our aim is to clarify that this patent applicability claim to Kyber and Saber is baseless. The patent considers a commutative algebraic setup, as insisted upon by its owner, as otherwise it would be invalidated by the prior work, including that of Lyubashevsky, Palacio and Segev [LPS10] in the non-commutative case. Due to this algebraic setup, it cannot apply to the Kyber and Saber encryption schemes, which use the non-commutative setup of [LPS10]. Hence the patent cannot both claim novelty and apply to Kyber and Saber. This said, we do not discuss here whether the prior work, including [LPS10], invalidates the Gaborit&Aguilar-Melchor patent or not – we only focus on the invalidity of the applicability of the patent to cover Kyber and Saber.

In Section 2, we first describe the LPS scheme, which is a version of Regev’s LWE scheme [Reg09] in which the public key and the first ciphertext component are symmetrically formed. We then give the scheme

\textsuperscript{4} https://csrc.nist.gov/Projects/post-quantum-cryptography/
\textsuperscript{5} https://www.cnrsinnovation.com/?lang=en
\textsuperscript{6} See p. 26 and p. 15 of the BIKE and HQC IP statements, available at
\textsuperscript{7} See 1:51:06 and 2:03:20 of https://www.youtube.com/watch?v=Fd0KWktBLeU
from the GA patent, and then finally describe Kyber. This will illustrate how the structure of Kyber mimics LPS, while the scheme from the GA patent is essentially a dimension 1 instance of LPS. In Section 3, we give relevant quotes from the proceedings between Keltie LLP and CNRS, in which Keltie tried to invalidate the GA patent based on prior art. The CNRS defense was that by being a dimension 1 instance, the scheme became commutative, and this commutativity was a crucial element of the scheme that was not present in others. Most importantly for Kyber and Saber, CNRS then insisted that their patent does not stand in the way of non-commutative versions of the scheme being patented later by others.

Notations. Matrices are in bold upper-case. Vectors are in bold lower-case. The transpose of a vector is denoted $s^T$. Unless transposed, a vector is always a column vector. The notation $\mathbb{Z}_q$ refers to the set of integers modulo $q$.

2 Encryption schemes

In this section, we recall the public-key encryption schemes from Lyubashevsky, Palacio and Segev [LPS10], Gaborit and Aguilar-Melchor [GAM10] and Kyber [BDK+18, ABD+21]. We focus on the aspects relevant to their comparison.

The LPS encryption scheme. As discussed in [LPS10, Section 1] and [Gol10, Section 7], the LPS encryption scheme from [LPS10, Section 3] can equivalently be described either in terms of the subset-sum problem modulo an integer $q^m$, or with $m$-dimensional square matrices and vectors modulo $q$. Here, we choose the second formalism.

KEYGEN: The secret key $sk$ is a vector $s \in \mathbb{Z}_q^m$ that is small, i.e., whose entries have absolute values that are small compared to $q$. The public key $pk = (A, t)$ consists of a matrix $A \in \mathbb{Z}_q^{n \times m}$ and a vector $t = A \cdot s + e \in \mathbb{Z}_q^n$, where $e$ is a small vector (in the sense above). Note that $A$ is independent of $sk$ and can be considered as a public parameter rather than as part of $pk$.

ENC: To encrypt a bit $z \in \{0, 1\}$, one first samples a small vector $r \in \mathbb{Z}_q^n$. Then one computes $c_1^T = r^T \cdot A + e_1^T \in \mathbb{Z}_q^n$ and $c_2 = r^T \cdot t + e_2 + (q - 1)/2 \cdot z$, where the coordinates of $e_1$ and $e_2$ have small absolute values compared to $q$. Finally, one returns the ciphertext $ct = (c_1, c_2) \in \mathbb{Z}_q^m \times \mathbb{Z}_q$.

DEC: To decrypt a well-formed ciphertext $ct = (c_1, c_2)$ with the secret key $sk = s$, one computes

$$c_2 - c_1^T \cdot s = r^T \cdot (A \cdot s + e) + e_2 + \frac{q - 1}{2} \cdot z - (r^T \cdot A + e_1^T) \cdot s$$

$$= \frac{q - 1}{2} \cdot z + (r^T \cdot e + e_2 - e_1^T \cdot s).$$

The term $r^T \cdot e + e_2 - e_1^T \cdot s$ having a small absolute value when reduced modulo $q$ (which is ensured by setting $q$ appropriately), the message $z$ can be recovered by checking whether $c_2 - c_1^T \cdot s$ is closer to 0 or to $(q - 1)/2$.

In [LPS10, Section 3], several bits $z_1, \ldots, z_k$ can be encrypted at once, by using $k$ vectors $s_i$ for the secret key and $k$ vectors $t_i = A \cdot s_i + e_i$ in the public key. In [LPS10, Section 3], the small vectors $s$ and $r$ are chosen binary and the error terms $e, e_1, e_2$ are deterministically determined by the other quantities: this is due to the subset sum formulation of the scheme (these terms correspond to carries). The scheme can obviously handle small vectors that are not binary and randomized error terms (as observed for example in [Gol10]). Overall, the key equation providing correctness is

$$r^T \cdot (A \cdot s) - (r^T \cdot A) \cdot s = 0.$$  (1)

A reader from the area may notice the resemblance between the LPS encryption scheme and Regev’s encryption scheme from [Reg09]. The main difference lies in the symmetry between $c_1$ and $t$, which in particular allows to choose a square matrix $A$. 

2
The **GA encryption scheme.** The Gaborit&Aguilar-Melchor patent [GAM10] provides both a key exchange mechanism and a public key encryption scheme. They are equivalent, and we choose here the encryption formalism, to ease the comparisons. For the same reason, we adapt the notations of [GAM10] to those of the prior work [LPS10]. The scheme relies on a ring $\mathcal{R}$ for which there exists a notion of smallness, and on a map $f : \mathcal{R} \to \mathcal{R}$ such that for all $x, y \in \mathcal{R}$, if $x, y$ are small compared to $f(x), f(y)$ then $x \cdot f(y) - y \cdot f(x)$ is small. The map $f$ is a public parameter.

**KeyGen:** The secret key $sk$ is a small ring element $s \in \mathcal{R}$. The public key $pk$ is a ring element $t = f(s) + e$, where $e \in \mathcal{R}$ is small.

**Enc:** To encrypt a message $z$, one first samples $r, e_1, e_2 \in \mathcal{R}$ small and computes $c_1 = f(r) + e_1$ and $c_2 = r \cdot t + G \cdot z + e_2$, where $G \in \mathcal{R}$ is a public parameter. The ciphertext is $ct = (c_1, c_2) \in \mathcal{R} \times \mathcal{R}$.

**Dec:** To decrypt a well-formed ciphertext $ct = (c_1, c_2)$ with a secret key $sk = s$, one computes

\[
    c_2 - c_1 \cdot s = r \cdot (f(s) + e) + G \cdot z + e_2 - (f(r) + e_1) \cdot s \\
    = G \cdot z + (re - e_1s + e_2) + (r \cdot f(s) - f(r) \cdot s).
\]

The term $re - e_1s + e_2$ is small as it is a combination of small elements, and the term $r \cdot f(s) - f(r) \cdot s$ is small by assumption on $f$. If $G$ is set properly, then the term $G \cdot z$ dominates (unless $z = 0$) and one may be able to recover $z$.

Note that the key equation providing correctness of the GA scheme is

\[
    r \cdot f(s) - f(r) \cdot s \approx 0. \tag{2}
\]

Several instantiations are provided in [GAM10]. If one wants to compare with the LPS scheme, the relevant one is to set $\mathcal{R} = \mathbb{Z}_q$, $f : x \mapsto a \cdot x$ for some public parameter $a \in \mathcal{R}$ and $G = (q - 1)/2$. One then exactly recovers the LPS encryption scheme as presented above, with $m = 1$. In particular, in that case, Equation (2) with an equality is exactly Equation (1). Note that one cannot recover [LPS10] for $m \geq 2$, as it involves matrices and vectors (over a ring), which do not commute: to recover the GA scheme, one would need a set of matrices that forms a commutative ring.

The **Kyber encryption scheme.** Kyber [BDK+18, ABD+21] relies on a polynomial ring $R_q = \mathbb{Z}_q[x]/(x^{256} + 1)$. We describe here a simplified version of the CPA-secure public-key encryption scheme version of Kyber [BDK+18, Section 3].

**KeyGen:** The secret key $sk$ is a vector $s \in R_q^m$ that is small, i.e., whose entries are polynomials with coefficients that have absolute values that are small compared to $q$. The public key $pk = (A, t)$ consists of a matrix $A \in R_q^{m \times m}$ and a vector $t = A \cdot s + e \in R_q^m$, where $e$ is small.

**Enc:** To encrypt a polynomial $z \in R_q$ with binary coefficients, one first samples a small vector $r \in R_q^m$.

Then one computes $c_1^T = r^T \cdot A + e_1^T \in R_q^m$ and $c_2 = r^T \cdot t + e_2 + \lfloor q/2 \rfloor \cdot z$, where $e_1$ and $e_2$ are small.

Finally, one returns a ciphertext $ct = (c_1, c_2) \in R_q^m \times R_q$.

**Dec:** To decrypt a well-formed ciphertext $ct = (c_1, c_2)$ with a secret key $sk = s$, one computes

\[
    c_2 - c_1^T \cdot s = r^T \cdot (A \cdot s + e) + e_2 + \lfloor q/2 \rfloor \cdot z - (r^T \cdot A + e_1^T) \cdot s \\
    = \lfloor q/2 \rfloor \cdot z + (r^T \cdot e + e_2 - e_1^T \cdot s) .
\]

The scheme correctness is argued exactly as in the LPS scheme. Actually, in terms of equations, this simplified version of the Kyber encryption scheme exactly matches with the above description of the LPS scheme.
3 On the applicability of the GA patent to Kyber

In the descriptions of the three schemes above, it is important to note that the GA scheme requires commutativity of the ring $\mathcal{R}$, to which belong the secret key $sk$, the public key $pk$, and the two ciphertext components $c_1$ and $c_2$. Indeed, Equation (1) makes a crucial use of transposition when applied to matrices and vectors rather than ring elements. This is similarly used for the matrices and vectors occurring in the Kyber encryption scheme. To make the equation work without transposition requires that the elements all belong to a commutative ring, as in the GA scheme. Now, if commutativity is what makes the GA scheme novel compared to prior work, including [LPS10], it cannot be claimed that Kyber derives from the GA scheme.

This basic impossibility is well-understood by the owner of the patent itself. The comparison between the encryption schemes from [GAM10] and prior art came under scrutiny in Keltie LLP’s opposition to the patent, at the European Patent Office.\(^8\)

Keltie claimed that Claim 1 of the patent, the GA encryption scheme presented in Section 2, is too general and fails when the ring $\mathcal{R}$ is non-commutative. CNRS replied that the fact that Claim 1 only applies to commutative rings is obvious to any expert. Here is their reply (the translation from French – with some aid from “google translate”— the content of square brackets, and the emphasis are ours):

(1) “Commutative” character: The opponent [Keltie] argues that the commutativity of the ring $\mathcal{R}$ would be presented as indispensable in the description and should therefore appear in Claim 1.

In algebra, a commutative ring is a ring whose multiplication law is commutative. Commutativity is one of the main properties of rings. This emerges, for example, from an algebra course intended for undergraduate students in Mathematics (document P1), in which Chapter 3, Section 1.1.1, Page 37 proposes a definition of the word “ring” followed immediately by its two main properties: the “commutative” character and the “unitary” character. Thus the most natural example of a ring is a commutative ring.

In the description of the patent, the examples of rings, i.e., the rings $F_2[x]/(x - 1)$, $\mathbb{Z}/p\mathbb{Z}$, $(\mathbb{Z}/p\mathbb{Z})[x]/(x^n - 1)$ (Paragraph 53), are all commutative rings.

In fact, $P_A$ and $P_B$ [with the notations of the GA description of Section 2, these are $c_1s$ and $rt$] are presented in Claim 1 respectively in the form $P_A = Y_A X_B + Y_A f(Y_B)$ and $P_B = Y_B X_A + Y_B f(Y_A)$ [in Section 2, these are $sc_1 + sf(r)$ and $re + rf(s)$]. According to the description, it is deduced that $P_A$ and $P_B$ then only differ by the value $Y_A X_B - Y_B X_A$ [in Section 2, this is $se_1 - re$], which implies that $Y_A f(Y_B) - Y_B f(Y_A)$ is zero or is at the very least of small norm. **This supposes in particular that the used ring is commutative (beyond the choice of the function $f$).** If the ring is not commutative, Claim 1 should have been reformulated to take account of the non-negligible difference $Y_A f(Y_B) - Y_B f(Y_A)$.

Those skilled in the art would therefore have recognized that, even if the commutativity is not explicitly specified, it is an implicit characteristic of Claim 1.

Since it is an implicit characteristic, it is not necessary to include it in Claim 1, since as stated in Guidelines F-IV-4.5.3, “it is not necessary to include all the details of the invention in the independent claim”.

For all practical purposes, to make precise this implicit characteristic, several auxiliary requests have been filed specifying that the ring $\mathcal{R}$ is “commutative” (AR1, AR3, AR5 and AR7).

In the same document, CNRS went further than just saying that commutativity is implicit: commutativity is actually crucial to separating their scheme from the prior art of Regev’s LWE encryption scheme.\(^9\) The parameters $(S_A$ and $S_B$, which correspond to $t$ and $c_1$ with the notations of Section 2) are of the same form in the GA scheme, i.e., each is an element of the ring $\mathcal{R}$. In Regev’s scheme, and in LPS and Kyber, these two ciphertext elements are crucially different (i.e., they are vectors or matrices). Because they are different, there is no commutativity in Equation (1). The claimed novelty in the GA scheme seems to be that everything is an element of $\mathcal{R}$.

---

\(^{8}\) All documents available at https://register.epo.org/application?number=EP11712927

\(^{9}\) See Page 3 of https://register.epo.org/application?documentId=E0V2M3NP1191DSU&number=EP11712927&lng=en&npl=false

\(^{10}\) See Page 7 of https://register.epo.org/application?documentId=E0V2M3NP1191DSU&number=EP11712927&lng=en&npl=false
It should also be noted that the parameters ‘$P$’ and ‘$u$’ of document E1 [See Algorithm 5 on Page 19 of https://cims.nyu.edu/~regev/papers/pqc.pdf – $P$ corresponds to $t$ and $u$ corresponds to $c_1$ in Section 2], which the opponent claims to correspond to the syndromes $S_A$ and $S_B$, are not of the same nature. There is thus no symmetry of calculation in the document E1. Claim 1 does not specifically mention that the $S_A$ and $S_B$ syndromes must be of the same nature, but this is made implicit by the identical calculation formulas of $S_A$ and $S_B$. This symmetry allows, during reconciliation, to have a difference $P_A - P_B$ which is of small norm. It should also be noted that the dimensions of the parameters ‘$P$’ and ‘$u$’ of the document E1 depend on three parameters $n$, $m$ and $l$. This is a much more general teaching than Claim 1 of the patent, in which $n = m = l = 1$. In other words, to at least partially achieve Claim 1, it would be necessary to choose the values of three parameters. However, a multiple selection among three parameters is necessarily new (see paragraph L.C.6.3.3 of the Case Law of the Boards of Appeal).

During the oral proceedings, CNRS emphasised again that their claim only covers commutative rings and that their claim does not prevent non-commutative rings from being patented later.\footnote{https://register.epo.org/application?documentId=E2T6283H7805DSU&number=EP11712927&lng=en&mpl=false}

3.5 He [Keltie LLP] added that the owner [CNRS] did not respond to the fact that the ring must be commutative. He insists on the fact that to date, we do not know how to implement a non-commutative ring.
3.21 The owner indicates that in no case would the patent as granted prevent the protection of a development based on non-commutative rings.

And finally, the implicit commutativity figured into the decision to uphold the patent.\footnote{See 16.2 of https://register.epo.org/application?documentId=E2T64BFC1036DSU&number=EP11712927&lng=en&mpl=false}

The opponent submitted that Claim 1 and all claims in general are directed to a ring $R$ in a general manner while the description only provides examples of commutative and cyclic polynomial rings so that these two characteristics are indispensable.

The patent owner has indicated that this is a disguised clarity objection.

The opposition division is of the opinion that the description of the patent provides several examples of rings allowing to carry out the invention and that the patent satisfies article EPC83. It is of the opinion that the commutative and cyclic polynomial aspects are sufficiently described in the patent by the function of the operations to be carried out during the reconciliation stage.

Acknowledgments. The first author is supported by the EU H2020 ERC Project 101002845 PLAZA. The second author was supported in part by European Union Horizon 2020 Research and Innovation Program Grant 780701 and BPI-France in the context of the national project RISQ (P141580).

References


Tech transfer CNRS subsidiary - From Research to Innovation since 1992

45
Experts at your service

2200
Exploitation agreements

29
Shareholdings

130
Technology Mappings

© Patrick VENAIL/CNRS Photothèque

Disclosure of CNRS Patent Properties

CNRS, acting on its behalf and on behalf of other French research organizations funded by the French Government, is managing certain patents which may be relevant to one or more proposals submitted to the National Institute of Standards and Technology (NIST) in the Post-Quantum Cryptography Standardization process (the “Patents”).
CNRS is committed to enabling a broad development of efficient solutions based on this upcoming standard. CNRS is also determined to provide fair and reasonable compensation to the research organizations that contributed to the Patents. Such compensation is dedicated to be reallocated to public research laboratories in order to continue their fundamental research programs.

Should a standard be adopted by NIST as a result of the Post-Quantum Cryptography Standardization process (the “Standard”), and should any claims of the aforementioned Patents be declared to be essential and necessary for the implementation of the Standard, then any party would have the right to use such Patent claims to implement and fully comply with the Standard, according to the following terms:

- CNRS will grant royalty-bearing licenses under the Patents to any party making a commercial use of the Standard to make, sell, use, import, or offer for sale the products and services listed in the table below (the “Licensed Products”). Such license agreements will be granted under transparent, royalty-bearing fair, reasonable, non-discriminatory terms (“FRAND terms”).

The royalty rates are:

<table>
<thead>
<tr>
<th>Licensed Product</th>
<th>Definition</th>
<th>Royalty rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedicated Hardware Products</td>
<td>Hardware Products are ready for use products (even if a battery or the like needs to be added for use) and can be directly used by an end user (an end user can be an entity or a person).</td>
<td>• 1% of the revenues from sales or lease of Hardware Products implementing the Standard</td>
</tr>
<tr>
<td>Non-exhaustive examples: hardware security modules (HSM), cryptographic security modules</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key Management Services</td>
<td>Services for managing cryptographic keys, including key generation, key exchange, storage and replacement of keys (“KMS”)</td>
<td>• 1% of the revenues from KMS involving the Standard. This includes for example using the Standard to securely communicate data and keys with the client of the KMS</td>
</tr>
<tr>
<td>Non-exhaustive examples: Cloud-based key management services, key management as a service</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- For the sake of clarity, CNRS will not assert any of the Patents against any party making a commercial use of the Standard to make, sell, use, import, or offer for sale products and services which are not Licensed Products. It is also understood that this document is not intended to present the full terms and conditions of the license agreements, but to set force only its financial conditions.
- CNRS will not assert any of the Patents against any party for making a non-commercial use of the Standard. Non-commercial use shall include academic research, learning and non-revenue generating services and products.

In the case of Licensed Products, discussions between CNRS and the licensees might lead to adjustments in the royalty payment modalities such as, but not limited to, lump sum payments based on the licensee’s revenue forecast from the sales or lease of Licensed Products.