Quaternion Gaussian matrices satisfy the RIP

Agnieszka Badeńska
badenska@mini.pw.edu.pl
1 Faculty of Mathematics
and Information Science
Warsaw University of Technology
ul. Koszykowa 75
00-662 Warszawa, Poland

Łukasz Błaszczyk
l.blaszczyk@mini.pw.edu.pl
1,2
1 Institute of Radioelectronics
and Multimedia Technology
Warsaw University of Technology
ul. Nowowiejska 15/19
00-665 Warszawa, Poland

Keywords: quaternion Gaussian random matrix, restricted isometry property, sparse signals.

Abstract

We prove that quaternion Gaussian random matrices satisfy the restricted isometry property (RIP) with overwhelming probability. We also explain why the restricted isometry random variables (RIV) approach is not appropriate for drawing conclusions on restricted isometry constants.

1 Introduction

One of the conditions which guarantees exact reconstruction of a sparse signal (real, complex or quaternion) – by $\ell_1$-norm minimization – from a few number of its linear measurements is that the measurement matrix satisfies the celebrated restricted isometry property (RIP) with a sufficiently small constant. The notion of restricted isometry constants was introduced by Candès and Tao in [4] and repeatedly considered afterwards. The concept was also generalized to quaternion signals [1, 2].

Definition 1.1. Let $\Phi \in \mathbb{H}^{m \times n}$ and $s \in \{1, \ldots, n\}$. The $s$-restricted isometry constant of $\Phi$ is the smallest number $\delta_s$ with the property that

$$(1 - \delta_s) \|x\|^2 \leq \|\Phi x\|^2 \leq (1 + \delta_s) \|x\|^2$$

(1.1)

for all $s$-sparse quaternion vectors $x \in \mathbb{H}^n$.

Recall that we call a vector (signal) $s$-sparse if it has at most $s$ nonzero coordinates and $\mathbb{H}$ denotes the quaternion algebra.

It is known that e.g. real Gaussian and Bernoulli random matrices, also partial Discrete Fourier Transform matrices satisfy the RIP (with overwhelming probability) [5], however, until recently there were no known examples of quaternion matrices satisfying this condition. In [1, Lemma 3.2] we proved that if a real matrix $\Phi \in \mathbb{R}^{m \times n}$ satisfies the inequalities (1.1) for real $s$-sparse vectors $x \in \mathbb{R}^n$, then it also satisfies it – with the same constant $\delta_s$ – for $s$-sparse quaternion vectors $x \in \mathbb{H}^n$. This was a first step towards developing theoretical background of compressed sensing methods in the quaternion algebra, since we also showed that it is possible to reconstruct sparse quaternion vectors from a small number of their linear measurements if an appropriate restricted isometry constant of a real measurement matrix $\Phi \in \mathbb{R}^{m \times n}$ is sufficiently small [1, Corollary 5.1].

Later we extended the above-mentioned result to the full quaternion setting – with the expected assumption that a quaternion measurement matrix $\Phi \in \mathbb{H}^{m \times n}$ satisfies
the RIP with a sufficiently small constant \[2\text{, Corollary 5.1}\]. The question of existence of quaternion matrices satisfying the RIP was, however, still open. Let us also mention a very interesting recent result of N. Gomes, S. Hartmann and U. Kähler concerning the quaternion Fourier matrices – arising in color representation of images \[6\]. They showed that with high probability such matrices allow a sparse reconstruction (again by \(\ell_1\)-norm minimization) \[3\text{, Theorem 3.2}\]. Their proof is straightforward and does not use the RIP notion, however, they also point out that it is not clear whether quaternion sampling matrices fulfill the RIP condition.

It has been believed that quaternion Gaussian random matrices satisfy the RIP and, therefore, they have been widely used in numerical experiments \[3, 7, 17\] but there was a lack of theoretical justification for this conviction. In this article we prove that this hypothesis is true, i.e. quaternion Gaussian random matrices satisfy the RIP, and we provide estimates on matrix sizes that guarantee the RIP with overwhelming probability (Theorem 4.2). The existence of quaternion matrices satisfying the RIP, together with the main results of \[2\], constitute the theoretical foundation of the classical compressed sensing methods in the quaternion algebra.

Apart from the aforementioned numerical experiments, our main motivation was the article \[2\], which revealed the need for quaternion matrices satisfying the RIP. In our research we got interested in the articles \[8, 9, 10\], the authors of which claim they had found a simpler proof of the RIP for real and complex Gaussian random matrices. This brought us to study the ratio random variable, so-called Rayleigh quotient

\[
\mathcal{R} = \frac{\|\Phi \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2}
\]

for \(\Phi \in \mathbb{H}^{m \times n}\) and non-zero, \(s\)-sparse vector \(\mathbf{x} \in \mathbb{H}^n\), and eventually finding its distribution, i.e. \(\Gamma(2m, 2m)\), which does not depend on \(\mathbf{x}\), (Lemma 3.1). Its mean (expectation) equals 1 and its variance equals \(\frac{1}{2m}\). Recall that in the real case \(\mathcal{R}\) has distribution \(\Gamma\left(\frac{m}{2}, \frac{m}{2}\right)\) with the same mean but four times bigger variance \(\frac{2}{m}\) \[8\]. It clearly explains higher rate of correct sparse signals reconstructions in the quaternion case in compare with the real case of the same size (see \[2\, Section 6\]).

Having proved Lemma 3.1 we were initially hoping to mimic the approach from \[8, 9, 10\] to obtain the proof of the RIP for quaternion Gaussian matrices, however, in our humble opinion it contains certain inaccuracies, which we want to briefly point out before proceeding further (we refer to \[8\] for more details). The authors call a vector \(\mathbf{x}\) \(s\)-sparse if it has exactly \(s\) non-zero coordinates and introduce left and right \(s\)-restricted isometry constants (RIC) as

\[
\delta^L_s = 1 - \min_{\# \text{supp } \mathbf{x} = s} \frac{\|\Phi \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2}, \quad \delta^R_s = \max_{\# \text{supp } \mathbf{x} = s} \frac{\|\Phi \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} - 1, \tag{1.2}
\]

where \#\text{supp } \mathbf{x} denotes the cardinality of the support set of \(\mathbf{x}\) (i.e. the number of its nonzero coordinates). Note that it differs substantially from our approach from Definition 1.1 (\(\max(\delta^L_s, \delta^R_s)\) might be smaller than \(\delta_s\) and we loose monotonicity of \(\delta_s\) with respect to \(s\)). Moreover, the set of vectors with a fixed cardinality of the support does not have the structure of a module over the ring \(\mathbb{H}\) anymore, hence the above variables may not be well-defined (one might need to use inf / sup instead).

After noticing that all ratio random variables \(\frac{\|\Phi \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2}\) have the same distribution, which does not depend on the choice of \(\mathbf{x}\), the aforementioned authors only consider one vector \(\mathbf{x}^S\) with \(\text{supp } \mathbf{x}^S = S\) for every support set \(S\) with \(#S = s\) and define the following variable,
called restricted isometry random variable (RIV)

$$\Delta_s^R = \max_{S \neq S'} \left\{ \frac{\|\Phi x^S\|^2}{\|x^S\|^2} : \text{for some } x^S \text{ with } \text{supp } x^S = S \right\} - 1$$

($\Delta_s^L$ analogously), that is omitting max / min over uncountable sets of vectors supported on appropriate sets $S$. That is true that for a fixed vector and full ensemble of Gaussian matrices all Rayleigh quotients $R$ have the same distribution, yet the vector that realizes the max / min on a fixed support set $S$ depends on the realization matrix in general, therefore the distribution of $\max \left\{ \frac{\|\Phi x^S\|^2}{\|x^S\|^2} : \text{supp } x = S \right\}$ very likely differs from the distribution of a single $R$ for a fixed vector. Therefore, one cannot draw any conclusions concerning $\delta_s^L$, $\delta_s^R$ or $\delta_s$ based on the upper estimates of $\Delta_s^L$, $\Delta_s^R$ since obviously

$$\Delta_s^L \leq \delta_s^L \leq \delta_s \quad \text{and} \quad \Delta_s^R \leq \delta_s^R \leq \delta_s.$$

Empirical distributions of RICs and RIVs are presented in section 5 in Fig. 1 and 2.

Finally, the authors of [8] claim to have derived the precise distribution functions of $\Delta_s^R$ and $\Delta_s^L$ using, what they call, an 'i.i.d. representation', even though they admit that the Rayleigh quotients for different supports are not independent if the support sets intersect. In fact, they suggest a lower estimate of the cumulative distribution functions (CDF) of the RIVs by the CDF of a maximum of $(n)$ (the number of support sets) i.i.d. variables of the same distribution as $R$ [8 Appendix E, Step 1], however, their proof is very unsatisfactory due to the lack of theoretical justification for the presented inequalities. For the case of two random variables what they really use is so-called positive dependence, proved for certain distributions (including Gamma and Chi-square) in more general case, i.e.

$$\mathbb{P}(X_1 \in A, X_2 \in A) \geq \mathbb{P}(X_1 \in A) \cdot \mathbb{P}(X_2 \in A)$$

for all measurable sets $A \subset \mathbb{R}$ (see e.g. [11]). For the bigger number of variables with Gamma distribution this result was extended to a certain extent [13] but – to our best knowledge – there is no proof for the general case.

To sum up, it might be somewhat interesting to study the RIVs in the compressed sensing analysis, however, it brings no information on the upper estimate of the restricted isometry constants. Therefore, this approach is not useful for attempts of proving the RIP. In fact, the eigenvalue approach, i.e. analysis of moduli of eigenvalues of Hermitian matrices $\Phi^*_S \Phi_S$, criticized by the authors of [8], is the right direction – it is well known that the restricted isometry constants can be expressed in terms of these eigenvalues (also in the quaternion case – cf. [2, Lemma 2.1]). Currently however, we do not know distribution of the eigenvalues of $\Phi^*_S \Phi_S$ in the quaternion case.

That is why our proof of the RIP of quaternion Gaussian random matrices refers to the classical approach for real sub-Gaussian random matrices [5 section 9.1], which uses the notion of sub-exponential (locally sub-Gaussian) random variables and a $\gamma$-covering of a unit ball. We are aware that our result (Theorem [12]) is not optimal, and in the current form only applies to the case of big $n$ and very small $s$, but it is a first step towards further thorough analysis of the quaternion random matrices and their applications in compressed sensing.

The article is organized as follows. First, we briefly recall basic facts about the quaternion algebra and quaternion matrices. Third section is devoted to quaternion random variables and matrices. We define the quaternion Gaussian random variable with mean zero and variance $\sigma^2$, denoted by $X \sim \mathcal{N}_H(0, \sigma^2)$, where in particular we always assume
independence of its components – note that this aspect was not clear in [17]. We also provide distribution of the Rayleigh quotient $R$ for quaternion Gaussian random matrices and justify that it is sub-exponential with appropriate parameters. In the fourth section we prove the main result of the article, i.e. the RIP for quaternion Gaussian random matrices. Finally, we present outcomes of numerical simulations illustrating the theoretical considerations and we conclude with a short résumé of the obtained results and our further research perspectives.

2 Quaternions

Denote by $\mathbb{H}$ the algebra of (Hamilton or real) quaternions

$$q = a + bi + cj + dk,$$

where $a, b, c, d \in \mathbb{R}$

endowed with the standard norm

$$|q| = \sqrt{\bar{q}q} = \sqrt{a^2 + b^2 + c^2 + d^2},$$

where $\bar{q} = a - bi - cj - dk$ is the conjugate of $q$. Recall that multiplication is associative but in general not commutative in the quaternion algebra and is defined by the following rules

$$i^2 = j^2 = k^2 = ijk = -1 \quad \text{and} \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j.$$ 

Multiplication is distributive with respect to addition and has a neutral element $1 \in \mathbb{H}$, hence $\mathbb{H}$ forms a ring, which is usually called a noncommutative field.

In this article we consider vectors with quaternion coordinates, i.e. elements of $\mathbb{H}^n$. Algebraically $\mathbb{H}^n$ is a module over the ring $\mathbb{H}$, usually called the quaternion vector space (although it is not a vector space, since $\mathbb{H}$ is not a field). For any $n \in \mathbb{N}$ we consider the Hermitian form $\langle \cdot, \cdot \rangle: \mathbb{H}^n \times \mathbb{H}^n \to \mathbb{H}$ with quaternion values

$$\langle x, y \rangle = y^* x = \sum_{i=1}^n y_i x_i, \quad \text{for} \quad x = (x_1, \ldots, x_n)^T, \ y = (y_1, \ldots, y_n)^T \in \mathbb{H}^n,$$

where $y^* = \overline{y}^T$, which satisfies axioms of an inner product (in terms of the right quaternion vector space, i.e. considering the right scalar multiplication) and the Cauchy-Schwarz inequality with respect to the norm

$$\|x\|_2 = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^n |x_i|^2}, \quad \text{for any} \quad x = (x_1, \ldots, x_n)^T \in \mathbb{H}^n.$$ 

See [2] Section 2] for the details.

We are particularly interested in matrices with quaternion entries with usual multiplication rules. A matrix $\Phi \in \mathbb{H}^{m \times n}$ defines a $\mathbb{H}$-linear transformation $\Phi : \mathbb{H}^n \to \mathbb{H}^m$ (in terms of the right quaternion vector space) which acts by the standard matrix-vector multiplication. By $\Phi^*$ we denote the adjoint matrix, i.e. $\Phi^* = \overline{\Phi}^T$, which defines the adjoint $\mathbb{H}$-linear transformation, i.e.

$$\langle x, \Phi^* y \rangle = (\Phi^* y)^* x = y^* \Phi x = \langle \Phi x, y \rangle \quad \text{for} \quad x \in \mathbb{H}^n, \ y \in \mathbb{H}^m.$$ 

Recall also that a linear transformation (matrix) $\Psi \in \mathbb{H}^{n \times n}$ is called Hermitian if $\Psi^* = \Psi$. Obviously, $\Phi^* \Phi$ is Hermitian for any $\Phi \in \mathbb{H}^{m \times n}$.

Below we recall certain property of quaternion Hermitian matrices [2 Lemma 2.1], which will be used later on.
Lemma 2.1. Suppose $\Psi \in \mathbb{H}^{n \times n}$ is Hermitian. Then

$$\|\Psi\|_{2\to 2} = \max_{x \in \mathbb{H}^n \setminus \{0\}} |\langle \Psi x, x \rangle| = \max_{x \in \mathbb{H}^n} \frac{|\langle \Psi x, x \rangle|}{\|x\|_2^2},$$

where $\|\cdot\|_{2\to 2}$ is the standard operator norm in the right quaternion vector space $\mathbb{H}^n$ endowed with the norm $\|\cdot\|_2$, i.e.

$$\|\Psi\|_{2\to 2} = \max_{x \in \mathbb{H}^n \setminus \{0\}} \frac{\|\Psi x\|_2}{\|x\|_2} = \max_{x \in \mathbb{H}^n, \|x\|_2 = 1} \|\Psi x\|_2.$$ 

3 Quaternion Gaussian random matrices

For a real random variable $X$ we will denote its expectation (mean) by $E X$ and its variance by $\text{Var} X$. For Gamma distribution $\Gamma(\alpha, \beta)$ with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$, i.e. random variable $X$ with the probability density function $\gamma_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ for $x \in (0, +\infty)$, we have that

$$E X = \frac{\alpha}{\beta} \quad \text{and} \quad \text{Var} X = \frac{\alpha}{\beta^2}.$$ 

Recall also that a sum of squares of $k$ independent standard Gaussian random variables $\mathcal{N}(0, 1)$ has Chi-square distribution with $k$ degrees of freedom, i.e. $\chi^2(k) = \Gamma \left( \frac{k}{2}, \frac{1}{2} \right)$.

Quaternion random variables have not been studied so far as thoroughly as their real or complex counterparts. However, during last two decades they attracted the attention of researchers both in theoretical and applied sciences [12, 15]. Quaternion random variable $X$ is defined by four real random variables $X = X_0 + X_1 i + X_2 j + X_3 k$

and as such can be associated with the four-dimensional real random vector $(X_1, X_2, X_3, X_4)$. There are several definitions of a quaternion Gaussian random variable [15]. The most general (so-called $R$-Gaussian) calls the quaternion variable $X$ Gaussian if $(X_1, X_2, X_3, X_4)$ is a Gaussian random vector in $\mathbb{R}^4$.

In this article we will only consider quaternion $R$-Gaussian random variables with independent components. More precisely, we will assume that

$$X_i \sim \mathcal{N} \left( 0, \frac{\sigma_i^2}{4} \right), \quad i = 1, \ldots, 4, \quad \text{and} \quad X_i \text{ are pairwise independent.}$$

Such variables $X = X_0 + X_1 i + X_2 j + X_3 k$ will be called Gaussian with mean zero and variance $\sigma^2$ and denoted by $X \sim \mathcal{N}_\mathbb{H} (0, \sigma^2)$.

In what follows we will consider quaternion random matrices with independent entries sampled from quaternion Gaussian distribution, which has been defined above. Let us emphasize once again that we will always assume independence of components of quaternion Gaussian random variables.

Lemma 3.1. Let $\Phi = (\phi_{ij}) \in \mathbb{H}^{m \times n}$ be a quaternion Gaussian random matrix whose entries are independent random variables with the distribution $\mathcal{N}_\mathbb{H} \left( 0, \frac{1}{m} \right)$ and let $0 \neq \mathbf{z} \in \mathbb{H}^n$. Then the random variable

$$\mathcal{R} = \frac{\|\Phi \mathbf{z}\|_2^2}{\|\mathbf{z}\|_2^2}$$

has Gamma distribution $\Gamma(2m, 2m)$ and it does not depend on $\mathbf{z}$. In particular, $E \mathcal{R} = 1$ and $\text{Var} \mathcal{R} = \frac{1}{2m}$. 

5
Proof. Since \( \Phi \frac{x}{\|x\|_2} = \Phi \left( \frac{x}{\|x\|_2} \right) \), without loss of generality we can assume that \( \|x\|_2 = 1 \) and hence \( \mathcal{R} = \|\Phi x\|_2^2 \). Let us decompose the matrix \( \Phi \) into its components:

\[
\Phi = \Phi_r + \Phi_i i + \Phi_j j + \Phi_k k,
\]

where \( \Phi_r, \Phi_i, \Phi_j, \Phi_k \in \mathbb{R}^{m \times n} \), and analogously every matrix entry can be written as

\[
\phi_{ij} = \phi_{r, ij} + \phi_{i, ij} i + \phi_{j, ij} j + \phi_{k, ij} k \quad \text{with} \quad \phi_{e, ij} \in \mathbb{R} \quad \text{for} \quad e \in \{r, i, j, k\}.
\]

We will use the same indices to denote components of the vector \( x = (x_1, \ldots, x_n)^T \).

Let \( \Phi x = y = (y_1, \ldots, y_m)^T \), then

\[
y_k = \sum_{\ell=1}^{n} \phi_{k, \ell} x_\ell = \sum_{\ell=1}^{n} \left( \phi_{r, \ell} + \phi_{i, \ell} i + \phi_{j, \ell} j + \phi_{k, \ell} k \right) \cdot \left( x_{r, \ell} + x_{1, \ell} i + x_{j, \ell} j + x_{k, \ell} k \right)
\]

\[
= \sum_{\ell=1}^{n} \left( \phi_{r, \ell} x_{r, \ell} - \phi_{i, \ell} x_{1, \ell} \right) + \sum_{\ell=1}^{n} \left( \phi_{i, \ell} i + \phi_{j, \ell} j + \phi_{k, \ell} k \right) \cdot x_{r, \ell}
\]

\[
+ \sum_{\ell=1}^{n} \left( \phi_{j, \ell} j + \phi_{k, \ell} k \right) \cdot x_{1, \ell}
\]

\[
+ \sum_{\ell=1}^{n} \left( \phi_{k, \ell} k \right) \cdot x_{j, \ell}
\]

\[
= y_{r,k} + y_{i,k} i + y_{j,k} j + y_{k,k} k.
\]

Recall that

\[
\phi_{e, ij} \sim \mathcal{N} \left( 0, \frac{1}{4m} \right) \quad \text{for} \quad e \in \{r, i, j, k\}, \quad i \in \{1, \ldots, m\}, \quad j \in \{1, \ldots, n\}.
\]

Hence for each \( e \in \{r, i, j, k\} \) and every \( k \in \{1, \ldots, m\} \), random variables \( y_{e,k} \) are Gaussian (as linear combinations of Gaussian random variables) with \( \mathbb{E} y_{e,k} = 0 \) and \( \text{Var} y_{e,k} = \frac{1}{4m} \) – since all \( \phi_{e,k,\ell} \) are independent and

\[
\text{Var} y_{r,k} = \sum_{\ell=1}^{n} \left( x_{r,\ell}^2 \text{Var} \phi_{r,\ell} + x_{1,\ell}^2 \text{Var} \phi_{i,\ell} + x_{j,\ell}^2 \text{Var} \phi_{j,\ell} + x_{k,\ell}^2 \text{Var} \phi_{k,\ell} \right)
\]

\[
= \frac{1}{4m} \sum_{\ell=1}^{n} \left( x_{r,\ell}^2 + x_{1,\ell}^2 + x_{j,\ell}^2 + x_{k,\ell}^2 \right) = \frac{\|x\|_2^2}{4m} = \frac{1}{4m}.
\]

For the remaining components analogously.

Independence of the variables \( \phi_{e,k,\ell} \) implies also independence of \( y_{e,k} \) and \( y_{e,\ell} \) for every fixed \( e \in \{r, i, j, k\} \) and all pairs \( k, \ell \in \{1, \ldots, m\} \), \( k \neq \ell \). In order to verify independence
of \( y_{r,k} \) and \( y_{i,k} \), \( k \in \{1, \ldots, m\} \), observe that

\[
\text{Cov}(y_{r,k}, y_{i,k}) = \mathbb{E}(y_{r,k} \cdot y_{i,k}) - \mathbb{E} y_{r,k} \cdot \mathbb{E} y_{i,k}
\]

\[
= \mathbb{E} \left[ \left( \sum_{\ell = 1}^{n} (\phi_{r,k\ell} x_{r,\ell} - \phi_{i,k\ell} x_{i,\ell}) \right) \cdot \left( \sum_{p=1}^{n} (\phi_{i,kp} x_{i,p} - \phi_{r,kp} x_{r,p} + \phi_{j,kp} x_{j,p}) \right) \right]
\]

\[
= \sum_{\ell = 1}^{n} \mathbb{E} \left( (\phi_{r,k\ell} x_{r,\ell} - \phi_{i,k\ell} x_{i,\ell}) \cdot (\phi_{i,k\ell} x_{i,\ell} - \phi_{r,k\ell} x_{r,\ell}) \right)
\]

\[
= \sum_{\ell = 1}^{n} (x_{r,\ell} x_{i,\ell} \mathbb{E} \phi_{r,k\ell}^2 - x_{i,\ell} x_{r,\ell} \mathbb{E} \phi_{i,k\ell}^2 - x_{j,\ell} x_{k,\ell} \mathbb{E} \phi_{j,k\ell}^2 + x_{k,\ell} x_{j,\ell} \mathbb{E} \phi_{k,\ell}^2)
\]

\[
= \frac{1}{4m} \sum_{\ell = 1}^{n} (x_{r,\ell} x_{i,\ell} - x_{i,\ell} x_{r,\ell} - x_{j,\ell} x_{k,\ell} + x_{k,\ell} x_{j,\ell}) = 0,
\]

since \( \mathbb{E} \phi_{r,k \ell}^2 = \text{Var} \phi_{r,k \ell} = \frac{1}{4m} \). The same way one argues that covariance of the remaining pairs is also zero. Recall that real Gaussian random vectors have independent components if and only if their covariance equals zero.

We get, therefore, that all \( y_{e,k}, e \in \{r, i, j, k\} \) and \( k \in \{1, \ldots, m\} \), are independent (real) random variables with distribution \( \mathcal{N}(0, \frac{1}{4m}) \) and hence \( \sqrt{4m} y_{e,k} \sim \mathcal{N}(0, 1) \). This implies that

\[
4m \mathcal{R} = \left\| \sqrt{4m} y \right\|^2 = \sum_{\ell = 1}^{m} \left( \left( \sqrt{4m} y_{r,\ell} \right)^2 + \left( \sqrt{4m} y_{i,\ell} \right)^2 + \left( \sqrt{4m} y_{j,\ell} \right)^2 + \left( \sqrt{4m} y_{k,\ell} \right)^2 \right)
\]

is a sum of 4m squares of independent standard Gaussian random variables \( \mathcal{N}(0, 1) \) and consequently, \( 4m \mathcal{R} \) has Chi-square distribution \( \chi^2(4m) = \Gamma(2m, \frac{1}{2}) \). We conclude that \( \mathcal{R} \) has distribution \( \Gamma(2m, 2m) \), independently of \( x \). This random variable has mean \( \frac{2m}{(2m)^2} = \frac{1}{2m} \).

As we previously remarked, in the real case \( \mathcal{R} \) has distribution \( \Gamma \left( \frac{m}{2}, \frac{m}{2} \right) \), that is with four times bigger variance [8]. It explains the aforementioned better results of quaternion sparse vectors reconstruction in compare with the real case [2, Fig. 1 and Fig. 2 (a)], since a quaternion Gaussian random matrix statistically has smaller restricted isometry constant than its real counterpart.

Let us now proceed with the tools needed for the proof of the RIP. Recall that a real random variable \( X \) is called sub-exponential (locally sub-Gaussian) [10] if there exist \( \sigma^2 > 0 \) and \( \delta > 0 \) such that

\[
\mathbb{E} \left( e^{t(X - \mathbb{E} X)} \right) \leq \exp \left( \frac{\sigma^2 t^2}{2} \right) \text{ for } |t| \leq \delta.
\]

We will denote it \( X \sim \text{SubExp}(\sigma^2, \delta) \). Equivalently, one may write that

\[
M(t) = \mathbb{E} \left( e^{tX} \right) \leq \exp \left( t \mathbb{E} X + \frac{\sigma^2 t^2}{2} \right) \text{ for } |t| \leq \delta,
\]

where \( M(t) \) is the moment generating function.
It is known that a random variable with Gamma distribution is sub-exponential with any \( \sigma^2 > \text{Var } X \) for some \( \delta > 0 \) \cite{16}. Below we recall a simple proof of this fact, with \( \sigma^2 \) and \( \delta \) chosen for our purposes, in which we shall use the following form of the moment generating function of the Gamma distribution \( \Gamma(\alpha, \beta) \)

\[
M(t) = \frac{1}{\left(1 - \frac{t}{\beta}\right)^\alpha} \quad \text{for } t < \beta.
\]

**Lemma 3.2.** Let \( X \) has distribution \( \Gamma(\alpha, \beta) \). Then \( X \sim \text{SubExp} \left( \frac{5}{2} \text{Var } X, \frac{\delta}{5} \right) \).

**Proof.** Indeed, take \(|t| \leq \frac{\beta}{5}\), which equivalently means that \( \frac{|t|}{\beta} \leq \frac{1}{5} \). Then

\[
\mathbb{E}(e^{t(X - \frac{\alpha}{\beta})}) = \left(1 + \frac{t}{\beta} + \left(\frac{t}{\beta}\right)^2\right) \cdot e^{-t^2/2 \sigma^2} \\
\leq \exp\left(\alpha \cdot \left(\frac{t}{\beta} + \frac{5}{\beta} \left(\frac{t}{\beta}\right)^2\right)\right) \cdot \exp\left(-t^2/2 \sigma^2\right) \\
= \exp\left(\frac{1}{2} \cdot \frac{5}{\sigma^2} \cdot t^2\right),
\]

where we used well known estimation \(1 + x \leq e^x\) for \( x \in \mathbb{R} \). Hence \( \sigma^2 = \frac{5}{2} \frac{\alpha}{\beta^2} = \frac{5}{2} \text{Var } X \) and \( \delta = \frac{\delta}{5} \).

In what follows we will also use the following known fact \cite{16}: if \( X \sim \text{SubExp}(\sigma^2, \delta) \), then

\[
\mathbb{P}(\|X - \mathbb{E}X\| \geq t) \leq 2 \exp\left(-\frac{t^2}{2 \sigma^2}\right) \quad \text{for } 0 \leq t \leq \sigma^2 \delta.
\]

**Corollary 3.3.** The random variable \( R \sim \Gamma(2m, 2m) \) from Lemma 3.1 is sub-exponential with parameters \( \sigma^2 = \frac{5}{2} \cdot \frac{1}{2m} = \frac{5}{4m} \) and \( \delta = \frac{2m}{5} \). Hence

\[
\mathbb{P}(\|R - 1\| \geq t) \leq 2 \exp\left(-\frac{t^2}{2 \sigma^2}\right) \quad \text{for } 0 \leq t \leq \frac{1}{2},
\]

and therefore

\[
\forall_0 \neq x \in \mathbb{H}^n \quad \forall_{0 \leq t \leq \frac{1}{2}} \quad \mathbb{P}(\|\Phi x\|_2^2 - \|x\|_2^2 \geq t \|x\|_2^2) \leq 2 \exp\left(-\frac{t^2}{2 \sigma^2}\right). \quad (3.1)
\]

### 4 Proof of the RIP

As it was stated in the introduction, we say that a matrix \( \Phi \in \mathbb{H}^{m \times n} \) satisfies the \( s \)-restricted isometry property (for quaternion vectors) with a constant \( \delta_s \geq 0 \) if the inequalities \( (1.1) \) holds for all \( s \)-sparse quaternion vectors \( x \in \mathbb{H}^n \). The smallest number \( \delta_s \geq 0 \) with this property is called the \( s \)-restricted isometry constant. Without loss of generality one can only consider \( s \)-sparse unit vectors, i.e. \( \|x\|_2 = 1 \). Moreover, in \cite{2} Lemma 3.2] we proved that the \( s \)-restricted isometry constant \( \delta_s \) of \( \Phi \in \mathbb{H}^{m \times n} \) equivalently equals

\[
\delta_s = \max_{S \subseteq \{1, \ldots, n\}, \#S \leq s} \|\Phi_S^* \Phi_S - \text{Id}\|_{2 \rightarrow 2}, \quad (4.1)
\]
where $\Phi_S$ is the submatrix of $\Phi$ consisting of columns with indices from $S \subset \{1, \ldots, n\}$.

We begin with the following result in which we fix the support $S \subset \{1, \ldots, n\}$.

**Lemma 4.1.** Let $\Phi \in \mathbb{H}^{m \times n}$ be a quaternion Gaussian matrix whose entries $\phi_{ij}$ are independent quaternion random variables with distribution $\mathcal{N}_\mathbb{H}(0, \frac{1}{m})$. Moreover, let the set $S \subset \{1, \ldots, n\}$ be such that $\#S = s \leq n$. For any $\delta \in \left(0, \frac{1}{\sqrt{s}}\right)$ and $\varepsilon \in (0, 1)$, if

$$m \geq \frac{10}{3} \delta^{-2} \left(14s + \ln \left(\frac{2}{\varepsilon}\right)\right),$$

then with probability at least $1 - \varepsilon$ we have that

$$\|\Phi_S^* \Phi_S - \text{Id}\|_{2\to2} < \delta.$$

**Proof.** Fix a set $S \subset \{1, \ldots, n\}$ with $\#S = s$ and denote

$$\mathcal{A}_S = \{x \in \mathbb{H}^n : \text{supp } x \subset S \land \|x\|_2 = 1\}.$$  

This set can be associated with the unit sphere $S^{4s-1}$ in $\mathbb{R}^{4s}$.

Take a number $0 < \gamma < \frac{1}{2}$ (the exact value of $\gamma$ will be specified later). By [5, Proposition C.3] there exists a $\gamma$-covering $\mathcal{A}_\gamma$ of $\mathcal{A}_S$ such that

$$\#\mathcal{A}_\gamma \leq \left(1 + \frac{2}{\gamma}\right)^{4s}.$$  

For any $0 \leq \tilde{\delta} \leq \frac{1}{2}$, using (3.1), we get that

$$\mathbb{P}\left(\exists y \in \mathcal{A}_\gamma, \ |\|\Phi y\|_2^2 - \|y\|_2^2| \geq \tilde{\delta} \|y\|_2^2\right) = \mathbb{P}\left(\bigcup_{y \in \mathcal{A}_\gamma} \left\{|\|\Phi y\|_2^2 - \|y\|_2^2| \geq \tilde{\delta} \|y\|_2^2\right\}\right)$$

$$\leq \sum_{y \in \mathcal{A}_\gamma} \mathbb{P}\left(|\|\Phi y\|_2^2 - \|y\|_2^2| \geq \tilde{\delta} \|y\|_2^2\right)$$

$$\leq \#\mathcal{A}_\gamma \cdot 2 \exp\left(-\frac{\tilde{\delta}^2}{2\sigma^2}\right)$$

$$\leq 2 \left(1 + \frac{2}{\gamma}\right)^{4s} \exp\left(-\frac{\tilde{\delta}^2}{2\sigma^2}\right)$$

with $\sigma^2 = \frac{5}{4m}$. This implies that

$$\mathbb{P}\left(\forall y \in \mathcal{A}_\gamma, |\|\Phi y\|_2^2 - \|y\|_2^2| < \tilde{\delta} \|y\|_2^2\right) \geq 1 - 2 \left(1 + \frac{2}{\gamma}\right)^{4s} \exp\left(-\frac{\tilde{\delta}^2}{2\sigma^2}\right).$$

Since $\mathcal{A}_\gamma \subset \mathcal{A}_S$, we obviously have that

$$\forall y \in \mathcal{A}_\gamma, |\|\Phi y\|_2^2 - \|y\|_2^2| < \tilde{\delta} \|y\|_2^2 \iff \forall y \in \mathcal{A}_\gamma, |\|\Phi_S y_S\|_2^2 - \|y_S\|_2^2| < \tilde{\delta} \|y_S\|_2^2 \quad (4.2)$$

and

$$|\|\Phi_S y_S\|_2^2 - \|y_S\|_2^2| = |\langle (\Phi_S^* \Phi_S - \text{Id}) y_S, y_S \rangle|.$$
For a matrix $\Phi$ satisfying (4.2) denote $\Psi = \Phi^*_S \Phi_S - \text{Id}$. Since all vectors in $A$, are unit and supported on $S$, (4.2) implies that

$$\forall y \in A, \quad |(\Psi y, y_S)| < \tilde{\delta} \|y\|_2^2 = \tilde{\delta} = \tilde{\delta} = \tilde{\delta}.$$ 

By the definition of a $\gamma$-covering, for every $x \in A_s$ there is some $y \in A$, such that $\|x - y\|_2 \leq \gamma < \frac{1}{2}$. Since both $x$ and $y$ are unit and supported on $S$, using properties of the Hermitian norm and quaternion Cauchy-Schwarz inequality, we get that

$$|(\Psi x, x_S)| = |(\Psi y, y_S) + (\Psi x_S, x_S - y_S) + (\Psi (x_S - y), y_S)|$$

$$\leq |(\Psi y, y_S)| + \|\Psi\|_{2 \rightarrow 2} \|x\|_2 \|x - y\|_2 + \|\Psi\|_{2 \rightarrow 2} \|x - y\|_2 \|y\|_2$$

$$< \tilde{\delta} + 2\gamma \|\Psi\|_{2 \rightarrow 2}.$$ 

In view of Lemma 2.1, since $\Psi$ is Hermitian, taking supremum over all $x \in A_s$, we obtain

$$\|\Psi\|_{2 \rightarrow 2} < \tilde{\delta} + 2\gamma \|\Psi\|_{2 \rightarrow 2} \Rightarrow \|\Psi\|_{2 \rightarrow 2} < \frac{\delta}{1 - 2\gamma}.$$ 

Denoting $\delta = \frac{\tilde{\delta}}{1 - 2\gamma} \leq \frac{1}{2} \cdot \frac{1}{1 - 2\gamma}$, we get that

$$\mathbb{P}(\|\Phi^*_S \Phi_S - \text{Id}\|_{2 \rightarrow 2} < \delta) \geq \mathbb{P}((\forall y \in A, \quad \|\Phi y\|_2^2 - \|y\|_2^2 < \tilde{\delta} \|y\|_2^2)$$

$$\geq 1 - 2^{4s} \exp(-\frac{\delta^2}{2\sigma^2})$$

$$= 1 - 2^{4s} \exp(-\frac{2}{5} \delta^2 (1 - 2\gamma)^2 m).$$

It implies that if

$$m \geq \frac{5}{2} \cdot \frac{\delta^2}{(1 - 2\gamma)^2} \left(4s \cdot \ln \left(1 + \frac{2}{\gamma}\right) + \ln \left(\frac{2}{\varepsilon}\right)\right),$$

then

$$\mathbb{P}(\|\Phi^*_S \Phi_S - \text{Id}\|_{2 \rightarrow 2} < \delta) \geq 1 - \varepsilon.$$ 

(4.3)

Taking $\gamma = \frac{2}{\sqrt{2\sigma^2}} \approx 6.23 \cdot 10^{-2}$, for which $\frac{1}{(1 - 2\gamma)^2} \leq \frac{4}{3}$ and $\ln \left(1 + \frac{2}{\gamma}\right) = \frac{7}{2}$, we finally obtain that for any positive $\delta \leq \frac{1}{2} \cdot \frac{2}{\sqrt{3}} = \frac{1}{\sqrt{3}}$, if

$$m \geq \frac{10}{3} \cdot \delta^2 \left(14s + \ln \left(\frac{2}{\varepsilon}\right)\right),$$

then (4.3) holds, which concludes the proof. \square

We are ready to prove the main result.

**Theorem 4.2.** Let $\Phi \in \mathbb{H}^{m \times n}$ be a quaternion Gaussian matrix whose entries $\phi_{ij}$ are independent quaternion random variables with distribution $N_\mathbb{H}(0, \frac{1}{m})$. For any $\delta \in \left(0, \frac{1}{\sqrt{3}}\right)$ and $\varepsilon \in (0, 1)$, if

$$m \geq \frac{10}{3} \cdot \delta^2 \left(15s + \ln \left(\frac{2}{\varepsilon}\right) + s \ln \left(\frac{n}{s}\right)\right),$$

then with probability at least $1 - \varepsilon$ the $s$-restricted isometry constant $\delta_s$ of $\Phi$ satisfies $\delta_s < \delta$. \hfill \square
Proof. Using (4.1), the proof of Lemma 4.1 and well known estimates of the Newton’s symbol we get that
\[
\Pr(\delta_s \geq \delta) = \Pr\left(\max_{S: \#S=s} \|\Phi_S^* \Phi_S - \text{Id}\|_{2\rightarrow 2} \geq \delta\right)
= \Pr\left(\exists S: \#S=s \quad \|\Phi_S^* \Phi_S - \text{Id}\|_{2\rightarrow 2} \geq \delta\right)
= \Pr\left(\bigcup_{S: \#S=s} \{\|\Phi_S^* \Phi_S - \text{Id}\|_{2\rightarrow 2} \geq \delta\}\right)
\leq \sum_{S: \#S=s} \Pr(\|\Phi_S^* \Phi_S - \text{Id}\|_{2\rightarrow 2} \geq \delta)
\leq \binom{n}{s} \cdot 2 \left(1 + \frac{2\gamma}{7}\right)^{4s} \exp\left(-\frac{2}{5}\delta^2 (1-2\gamma)^2 m\right)
\leq 2 \left(\frac{en}{s}\right)^s \left(1 + \frac{2\gamma}{7}\right)^{4s} \exp\left(-\frac{2}{5}\delta^2 (1-2\gamma)^2 m\right).
\]
Therefore if
\[
m \geq \frac{5}{2} \cdot \frac{\delta^{-2}}{(1-2\gamma)^2} \left(s \ln\left(\frac{en}{s}\right) + 4s \cdot \ln\left(1 + \frac{2\gamma}{7}\right) + \ln\left(\frac{2}{\varepsilon}\right)\right),
\]
then \(\Pr(\delta_s < \delta) \geq 1 - \varepsilon\). Taking again \(\gamma = \frac{2}{\varepsilon/\delta^2 - 1}\) we get the thesis. \(\square\)

5 Numerical experiment

For the sake of illustrating our discussion regarding the RICs and the RIVs \cite{8,9,10}, we performed a numerical experiment of evaluating empirical distributions of the random variables \(\delta_s^R, \delta_s^L, \Delta_s^R, \Delta_s^L\) for the case of real Gaussian matrices \(\Phi \in \mathbb{R}^{m \times n}\). The experiment was carried out in MATLAB R2016a on a standard PC machine, with Intel(R) Core(TM) i7-4790 CPU (3.60GHz), 16GB RAM and with Microsoft Windows 10 Pro.

In order to efficiently derive empirical distributions, we set the size of Gaussian matrices \(\Phi \in \mathbb{R}^{m \times n}\) to \(m = 64, n = 8\), and the sparsity of unit \((\|x\|_2 = 1)\) vectors \(x \in \mathbb{R}^n\) to \(s = 5\). This choice of \(n\) and \(s\) gave us \(\binom{n}{s} = 56\) different support sets \(S \subset \{1, \ldots, n\}\) with \(#S = s\), and for each support \(S\) we drew \(10^3\) unit vectors \(x \in \mathbb{R}^n\) with \(\text{supp} x = S\). Hence, for every realization of the random variable \(\Phi\), we took maximum over the set consisting of \(56 \cdot 10^3\) elements to estimate distributions of \(\delta_s^R\) and \(\delta_s^L\). Estimating distributions of \(\Delta_s^R\) and \(\Delta_s^L\) required only one vector \(x\) with \(\text{supp} x = S\) for each support set \(S\). The matrix sample consisted of \(10^3\) realizations of the random variable \(\Phi\).

Results of this experiment, i.e. empirical probability and cumulative distribution functions, are shown in Fig.1 and 2. As one can see, knowledge of any upper bound of the variables \(\Delta_s^R\) and \(\Delta_s^L\) (with certain probability) brings no estimate on the upper bound of the RICs \(\delta_s^R\) and \(\delta_s^L\).

In the second experiment we derived empirical distribution of the restricted isometry constant \(\delta_s\) for quaternion Gaussian random matrices and compared it with the case of real Gaussian random matrices. We carried out the simulations for matrices \(\Phi \in \mathbb{K}^{m \times n}\), where \(\mathbb{K} = \mathbb{H}\) or \(\mathbb{K} = \mathbb{R}\), with parameters \(n = 256, m = 64\), and for unit \(s\)-sparse vectors \(x \in \mathbb{K}^n\) such that \(s = \frac{m}{2} = 32\).

Recall that the \(s\)-restricted isometry constant can be equivalently defined as
\[
\delta_s = \max \|\|\Phi x\|_2^2 - 1\|,
\]
Empirical probability density functions

Empirical cumulative distribution functions

Figure 1: Empirical distributions of \( \delta_R^s \) and \( \Delta_R^s \) for Gaussian random matrices \( \Phi \in \mathbb{R}^{m \times n} \), \( n = 8 \), \( m = 64 \), \( s = 5 \).

Empirical probability density functions

Empirical cumulative distribution functions

Figure 2: Empirical distributions of \( \delta_L^s \) and \( \Delta_L^s \) for Gaussian random matrices \( \Phi \in \mathbb{R}^{m \times n} \), \( n = 8 \), \( m = 64 \), \( s = 5 \).

where the maximum is taken over the set of all \( s \)-sparse vectors \( \mathbf{x} \in \mathbb{K}^n \) such that \( \| \mathbf{x} \|_2 = 1 \). To estimate the cumulative distribution function of the random variable \( \delta_i \) for each random matrix \( \Phi \) we drew \( 10^5 \) \( s \)-sparse unit vectors and took maximum over this set. The matrix sample consisted of \( 10^4 \) realizations of the random variable \( \Phi \) in both cases. Entries of the matrices were independently sampled from the normal distribution with mean zero and variance \( \frac{1}{m} \), i.e. \( \mathcal{N} \left( 0, \frac{1}{m} \right) \) in the real case and \( \mathcal{N}_{\mathbb{H}} \left( 0, \frac{1}{m} \right) \) in the quaternion case.

Results of this experiment – for real and quaternion random matrices – are shown in Fig. 3 and Fig. 4 respectively. In view of the previous theoretical considerations, we can see that (with certain probability) the \( s \)-restricted isometry constant of a quaternion
Gaussian random matrix is small. It proves, in particular, that there exist quaternion matrices satisfying the RIP. Moreover, notice that the $s$-restricted isometry constant of a quaternion random matrix $\Phi \in \mathbb{H}^{m \times n}$ is statistically smaller than its real counterpart for a matrix $\Phi \in \mathbb{R}^{m \times n}$ of the same size. As we previously mentioned, we think that the reason for this phenomenon is smaller variance of the Rayleigh random variable in the quaternion case (Lemma 3.1).

![Empirical probability density function](image1.png)  ![Empirical cumulative distribution function](image2.png)

(a) (b)

Figure 3: Empirical distributions of $\delta_s$ for Gaussian random matrices $\Phi \in \mathbb{R}^{m \times n}$, $n = 256$, $m = 64$, $s = 32$.

![Empirical probability density function](image3.png)  ![Empirical cumulative distribution function](image4.png)

(a) (b)

Figure 4: Empirical distributions of $\delta_s$ for Gaussian random matrices $\Phi \in \mathbb{H}^{m \times n}$, $n = 256$, $m = 64$, $s = 32$.

Finally, one might be interested in comparing these results with the formulation of The-
We should remember, however, that the estimate in this theorem is significant only for very large $n$ (more than few thousand) and $s \ll n$ (in practice, less than a dozen) and in our experiment, due to computational reasons, $n = 256$ and $s = 32$. As we also commented, we are aware that this result is not sharp and the numerical experiment additionally suggests that it can be improved in future.

6 Conclusion

This article brings positive answer to the question about existence of quaternion matrices satisfying the RIP. We confirm that – as it was conjectured – restricted isometry constants of Gaussian quaternion matrices are small with big probability (and typically smaller than their real counterparts). Together with our previous result, in which we proved that quaternion measurement matrices with small RIP constants allow exact reconstruction of sparse quaternion vectors, it explains success of compressed sensing based experiments in the quaternion algebra and brings hope for their wider applications.

The main result, however, in the current form is not sharp. One of the reasons is that we used techniques previously applied to the case of real subgaussian random matrices. It would be interesting in the future to study distribution of singular values of quaternion Gaussian matrices in order to improve our result. Other direction of the further research might be search for other quaternion matrices satisfying the RIP, in particular, those more desired from the application point of view, e.g. partial quaternion DFT matrix.

Acknowledgments

The research was supported in part by WUT grant No. 504/02499/1120. The work conducted by the second author was supported by a scholarship from Fundacja Wspierania Rozwoju Radiokomunikacji i Technik Multimedialnych.

References

[1] A. Badeńska, Ł. Błaszczyk, Compressed sensing for real measurements of quaternion signals, preprint (2015), arXiv:1605.07985.

[2] A. Badeńska, Ł. Błaszczyk, Compressed sensing in the quaternion algebra, preprint (2017), arXiv:1704.08202.

[3] Q. Barthélémy, A. Larue, J. Mars, Color Sparse Representations for Image Processing: Review, Models, and Prospects, IEEE Trans Image Process, (2015), 1–12.

[4] E. J. Candès, T. Tao, Decoding by linear programming, IEEE Trans. Inform. Theory 51 (12) (2005) 4203–4215.

[5] S. Foucart, H. Rauhut, A mathematical introduction to compressive sensing, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, New York (2013).

[6] N. Gomes, S. Hartmann, U. Kähler, Compressed Sensing for Quaternionic Signals, Complex Anal. Oper. Theory 11 (2017), 417–455.

[7] M. B. Hawes, W. Liu, A Quaternion-Valued Reweighted Minimisation Approach to Sparse Vector Sensor Array Design, Proceedings of the 19th International Conference on Digital Signal Processing, (2014), 426–430.
[8] O. James, H.-N. Lee, Restricted Isometry Random Variables: Probability Distributions, RIC Prediction and Phase Transition Analysis for Gaussian Encoders, *preprint* (2015), arXiv:1410.1956v4.

[9] O. James, On the RIP of Real and Complex Gaussian Sensing Matrices via RIV framework in Sparse Signal Recovery Analysis, *IEEE sponsored 2nd International Conference on Electronics and Communications System* (2015).

[10] O. James, Revisiting the RIP of Real and Complex Gaussian Sensing Matrices Through RIV Framework, *Wireless Pers Commun*, 87 (2016), no. 2, 513–526.

[11] D. R. Jensen, A note on positive dependence and the structure of bivariate distributions, *SIAM J. Appl. Math.* 20 (1971), 749–753.

[12] N. Le Bihan, The geometry of proper quaternion random variables, *Signal Processing* 138 (2017), 106–116.

[13] T. Royen, Some probability inequalities for multivariate gamma and normal distributions, *Far East J. Theor. Stat.* 51 (2015), no. 1-2, 17–36.

[14] J. F. Sturm, Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones, *Optimization Methods and Software* 11 (1999), no. 1-4, 625–653.

[15] N. N. Vakhania, G. Z. Chelidze, Quaternion Gaussian random variables, (Russian) *Teor. Veroyatn. Primen.* 54 (2009), no. 2, 337–344; translation in *Theory Probab. Appl.* 54 (2010), no. 2, 363–369.

[16] M. Wainwright, Basic tail and concentration bounds, *High-dimensionnal statistics: A non-asymptotic viewpoint* (Ed. M. Wainwright), manuscript (2016).

[17] J. Wu, X. Zhang, X. Wang, L. Senhadji, H. Shu, L1-norm minimization for quaternion signals, *Journal of Southeast University* 1 (2013) 33–37.