Almost Lossless Analog Compression without Phase Information – Complex Case

Georg Tauböck
Austrian Academy of Sciences, 1040 Vienna, Austria
georg.tauboeck@oeaw.ac.at

Erwin Riegler
ETH Zurich, 8092 Zurich, Switzerland
eriegler@nari.ee.ethz.ch

***Abstract*—We extend the recently proposed information-theoretic framework for phase retrieval [1] to the complex case. Specifically, we consider the problem of recovering an unknown random vector $\mathrm{x} \in \mathbb{C}^n$ up to an overall phase factor from $\lfloor Rn \rfloor$ phaseless measurements with compression rate R and derive a general achievability bound for R. Although phase retrieval is known not to extend straightforwardly from the real to the complex case, our bound on the compression rate turns out to be conceptually similar to the one derived for real-valued signals [1]. For x being s-sparse our results imply that $2s$ phaseless measurements are sufficient to recover x up to an overall phase factor irrespectively of the specific distribution of x. The best known recovery threshold for deterministic complex-valued s-sparse vectors is $4s-2$ so far.**

I. Introduction

In many different areas of science, physical limitations make it impossible to measure the phase of a signal but obtaining amplitudes is relatively easy. Well known examples are X-ray crystallography, astronomy, or diffraction imaging [2]–[4]. The problem of retrieving a signal up to a global phase factor from intensity measurements is often referred to as *phase retrieval*. More formally, let $\mathbb{C}^n_{\sim}$ be the set of equivalence classes $[\mathrm{x}] = \{e^{i\varphi}\mathrm{x} \mid \varphi \in [0, 2\pi)\}$ with $\mathrm{x} \in \mathbb{C}^n$. Phase retrieval is the problem of recovering $[\mathrm{x}] \in \mathbb{C}^n_{\sim}$ from m phaseless measurements of the form[1] $\mathrm{y} = |\mathrm{Ax}| \in \mathbb{R}^m$ with measurement matrix $\mathrm{A} \in \mathbb{C}^{m \times n}$.

It is by no means clear how large m has to be to allow for recovery of $[\mathrm{x}] \in \mathbb{C}^n_{\sim}$ from m phaseless measurements. Thus from the very beginning, there have been a number of works regarding recovery conditions for this problem in the context of specific applications [5]. This question has also been studied in more abstract terms, asking for the smallest number m of phaseless measurements that is required to make the mapping $[\mathrm{x}] \mapsto |\mathrm{Ax}|$ injective without imposing structural assumptions on A. It has also been observed that phase retrieval for complex-valued signals is different from phase retrieval for real-valued signals: while in the complex case $m \geq 4n-4$ measurements are generically sufficient to guarantee injectivity [6], the real case requires only $m \geq 2n-1$ measurements [7]. Furthermore, it was shown that semidefinite programming can be used to recover $[\mathrm{x}]$ if $\mathbf{A}$ is random with i.i.d. complex Gaussian entries or with i.i.d. rows that are uniformly distributed on the complex sphere, as long as $m \geq c_0 n$ for a sufficiently large constant c_0 [8].

Recently, there has been also interest in *sparse phase retrieval*, where the number s of nonzero coefficients of the vector x is much smaller than n. This a-priori knowledge about x can be used to reduce the number of measurements significantly: if the rows of the measurement matrix A are a generic choice of vectors in $\mathbb{C}^n$, injectivity of the mapping $[\mathrm{x}] \mapsto |\mathrm{Ax}|$ is guaranteed provided that $m \geq 4s-2$ [9].

Contributions: Inspired by the information-theoretic analysis of compressed sensing in [10], we formulate phase retrieval for complex-valued signals as an analog source coding problem, thereby complementing the recent work about phase retrieval for real-valued signals [1]. We assume that the unknown vector $\mathbf{x} \in \mathbb{C}^n$ is random[2] and derive recovery results for $[\mathbf{x}]$. Our results hold for Lebesgue almost all[2] (a.a.) measurement matrices A and are in terms of probability of error (with respect to the distribution of $\mathbf{x}$). In Theorem 1 we show that we can recover $[\mathbf{x}]$ from $m = \lfloor Rn \rfloor$ phaseless measurements with arbitrarily small probability of error for a.a. measurement matrices A, provided that n is sufficiently large and the *compression rate* R is larger than the (lower) Minkowski dimension compression rate (see Definition 4) of $\mathbf{x}$. It is remarkable that the obtained result is analogous to the corresponding result in the real case [1]; note, however, that here the (lower) Minkowski dimension is defined over a complex vector space rather than over a real vector space. In Theorem 2 we show that any s-sparse random vector $\mathbf{x} \in \mathbb{C}^n$ can be recovered with arbitrarily small probability of error, provided that $m \geq 2s$, which improves upon the best known recovery threshold $m \geq 4s-2$ for deterministic s-sparse vectors [9].

Notation: Roman letters $\mathrm{A}, \mathrm{B}, \ldots$ and $\mathrm{a}, \mathrm{b}, \ldots$ designate deterministic matrices and vectors, respectively. Boldface letters $\mathbf{A}, \mathbf{B}, \ldots$ and $\mathbf{a}, \mathbf{b}, \ldots$ denote random matrices and random vectors, respectively. For the distribution[2] of a random matrix $\mathbf{A}$ and a random vector $\mathbf{a}$, we write $\mu_{\mathbf{A}}$ and $\mu_{\mathbf{a}}$, respectively. The ith component of the vector u (random vector $\mathbf{u}$) is u_i (u_i). The superscripts $^{\mathsf{T}}$ and $^{\mathsf{H}}$ stand for transposition and Hermitian transposition, respectively. For a matrix A, $\mathrm{tr}(\mathrm{A})$ denotes its trace. The identity matrix of suitable size is denoted by I. For a vector u, we write $\|\mathrm{u}\| = \sqrt{\mathrm{u}^{\mathsf{H}}\mathrm{u}}$ for its Euclidean norm and $\mathrm{supp}(\mathrm{u})$ for its support. For the Euclidean space $(\mathbb{C}^k, \|\cdot\|)$,

This work was supported in part by the WWTF under grant VRG 12-009 and by the FWF under grant Y 551-N13.

[1]For a vector $\mathrm{u} \in \mathbb{C}^k$, we define the element-wise absolute value operation as $|\mathrm{u}| = (|u_1|,\ldots,|u_k|)^{\mathsf{T}}$.

[2] We follow the usual convention to identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ and $\mathbb{C}^{m\times n}$ with $\mathbb{R}^{2m\times 2n}$, cf. also Remark 1.

978-1-4673-7852-9/15/$31.00 ©2015 IEEE

we denote the open ball of radius r centered at $\mathrm{u} \in \mathbb{C}^k$ by $\mathcal{B}_k(\mathrm{u}, r)$, $V(k, r)$ stands for its volume (measured by the Lebesgue measure). For a set $\mathcal{S}$, $|\mathcal{S}|$ denotes its cardinality. The Borel sigma algebra on $\mathbb{C}$ is denoted by $\mathscr{B}_{\mathbb{C}}$. We write $\mathbb{R}_{\geq}$ for the set of nonnegative real numbers with Borel sigma algebra $\mathscr{B}_{\mathbb{R}_{\geq}}$. For $\mathrm{u}, \mathrm{v} \in \mathbb{C}^k$, $\mathrm{u} \sim \mathrm{v}$ means that there exists a $\varphi \in [0, 2\pi)$ s.t. $\mathrm{u} = e^{i\varphi}\mathrm{v}$ and we write for the corresponding equivalence classes $[\mathrm{u}] = \{\mathrm{v} \in \mathbb{C}^k | \mathrm{v} \sim \mathrm{u}\}$ with $\mathrm{u} \in \mathbb{C}^k$. For a set $\mathcal{S} \subseteq \mathbb{C}^k$, $\mathcal{S}_{\sim} = \{[\mathrm{u}] \mid \mathrm{u} \in \mathcal{S}\}$. The indicator function on a set $\mathcal{U}$ is denoted by $\chi_{\mathcal{U}}$.

II. Main Results

We start by formulating phase retrieval as a source coding problem.

Definition 1. *(Source vector) Let $(\mathrm{x}_i)_{i \in \mathbb{N}}$ be a stochastic process on $(\mathbb{C}^{\mathbb{N}}, \mathscr{B}_{\mathbb{C}}^{\otimes \mathbb{N}})$. Then, for $n \in \mathbb{N}$, the source vector $\mathbf{x}$ of length n is given by $\mathbf{x} = (\mathrm{x}_1, \ldots, \mathrm{x}_n)^\mathsf{T} \in \mathbb{C}^n$.*

Definition 2. *(Code, achievable rate) For $\mathbf{x}$ as in Definition 1 and $\varepsilon > 0$, an (n, m) code consists of*

(i) measurements $|\mathrm{A} \cdot | : \mathbb{C}^n \to \mathbb{R}_{\geq}^m$;

(ii) a decoder $g : \mathbb{R}_{\geq}^m \to \mathbb{C}^n$ that is measurable with respect to $\mathscr{B}_{\mathbb{R}_{\geq}}^{\otimes m}$ and $\mathscr{B}_{\mathbb{C}}^{\otimes n}$.

We call R with $0 < R \leq 1$ an ε-achievable rate if there exists an $N(\varepsilon) \in \mathbb{N}$ and a sequence of $(n, \lfloor Rn \rfloor)$ codes with decoders g such that

$$\mathrm{P}[g(|\mathrm{A}\mathbf{x}|) \not\sim \mathbf{x}] \leq \varepsilon$$

for all $n \geq N(\varepsilon)$.

Next, we introduce the Minkowski dimension compression rate for source vectors.

Definition 3. *(Minkowski dimension) Let $\mathcal{U}$ be a nonempty bounded set in $\mathbb{C}^n$. The lower Minkowski dimension of $\mathcal{U}$ is defined as*

$$\underline{\dim}_{\mathrm{B}}(\mathcal{U}) = \liminf_{\rho \to 0} \frac{\log N_{\mathcal{U}}(\rho)}{\log \frac{1}{\rho}}$$

and the upper Minkowski dimension of $\mathcal{U}$ is defined as

$$\overline{\dim}_{\mathrm{B}}(\mathcal{U}) = \limsup_{\rho \to 0} \frac{\log N_{\mathcal{U}}(\rho)}{\log \frac{1}{\rho}}$$

where $N_{\mathcal{U}}(\rho)$ is the covering number of $\mathcal{U}$ given by

$$N_{\mathcal{U}}(\rho) = \min \Big\{ k \in \mathbb{N} \mid \mathcal{U} \subseteq \bigcup_{i \in \{1, \ldots, k\}} \mathcal{B}_n(\mathrm{u}_i, \rho),\ \mathrm{u}_i \in \mathbb{C}^n \Big\}.$$

If $\underline{\dim}_{\mathrm{B}}(\mathcal{U}) = \overline{\dim}_{\mathrm{B}}(\mathcal{U})$, we write $\dim_{\mathrm{B}}(\mathcal{U})$.

Definition 4. *(Minkowski dimension compression rate) For $(\mathrm{x}_i)_{i \in \mathbb{N}}$ from Definition 1 and $\varepsilon > 0$, we define the lower Minkowski dimension compression rate as*

$$\underline{R}_{\mathrm{B}}(\varepsilon) = \limsup_{n \to \infty} \underline{a}_n(\varepsilon), \quad \text{where}$$

$$\underline{a}_n(\varepsilon) = \inf \Big\{ \frac{\underline{\dim}_{\mathrm{B}}(\mathcal{U})}{n} \ \Big|\ \mathcal{U} \subset \mathbb{C}^n,\ \mathrm{P}[\mathbf{x} \in \mathcal{U}] \ \geq 1 - \varepsilon \Big\}.$$

and the upper Minkowski dimension compression rate as

$$\overline{R}_{\mathrm{B}}(\varepsilon) = \limsup_{n \to \infty} \overline{a}_n(\varepsilon), \quad \text{where}$$

$$\overline{a}_n(\varepsilon) = \inf \Big\{ \frac{\overline{\dim}_{\mathrm{B}}(\mathcal{U})}{n} \ \Big|\ \mathcal{U} \subset \mathbb{C}^n,\ \mathrm{P}[\mathbf{x} \in \mathcal{U}] \ \geq 1 - \varepsilon \Big\}.$$

The sets $\mathcal{U}$ in the definitions for $\underline{a}_n(\varepsilon)$ and $\overline{a}_n(\varepsilon)$ are assumed to be nonempty and bounded.

Example 1. An i.i.d. stochastic process $(\mathrm{x}_i)_{i \in \mathbb{N}}$ on $(\mathbb{C}^{\mathbb{N}}, \mathscr{B}_{\mathbb{C}}^{\otimes \mathbb{N}})$ has a mixed discrete-continuous distribution if for each $i \in \mathbb{N}$ the random variables x_i are distributed according to

$$\mu_{\mathrm{x}_i} = (1 - \lambda)\mu_d + \lambda \mu_c$$

where $0 \leq \lambda \leq 1$ is the mixing parameter, μ_c is a distribution on $(\mathbb{C}, \mathcal{B}_{\mathbb{C}})$, absolutely continuous with respect to Lebesgue measure, and μ_d is a discrete distribution. Then, following [10, Th. 15],

$$\underline{R}_{\mathrm{B}}(\varepsilon) = 2\lambda, \quad 0 < \varepsilon < 1. \tag{1}$$

Remark 1. The factor two in (1) compared to the corresponding result for the real case [10, Th. 15] comes from the fact that for an open nonempty and bounded set $\mathcal{M} \subseteq \mathbb{C}^n$ we have $\dim_{\mathrm{B}}(\mathcal{M}) = 2n$. This follows from the isometry

$$\psi : \mathbb{C}^n \to \mathbb{R}^{2n},\ \mathrm{z} \mapsto (\mathrm{Re}(z_1), \mathrm{Im}(z_1), \ldots, \mathrm{Re}(z_n), \mathrm{Im}(z_n))^\mathsf{T} \tag{2}$$

which maps open nonempty and bounded sets in $\mathbb{C}^n$ to open nonempty and bounded sets in $\mathbb{R}^{2n}$, and using [11, Sec. 3.2, Property (i)].

Since in phase retrieval we can only hope to recover x up to a global phase factor, it is in general sufficient to consider only one specific element of $[\mathrm{x}]$. For a set, this motivates the following definition of a reduced set.

Definition 5. *(Reduced set) Let $\emptyset \neq \mathcal{U} \subseteq \mathbb{C}^n$. A set $\mathcal{U}_{\mathrm{red}} \subseteq \mathcal{U}$ is called reduced set of $\mathcal{U}$, if for each $[\mathrm{u}] \in \mathcal{U}_{\sim}$, $\mathcal{U}_{\mathrm{red}}$ contains exactly one $\mathrm{u}_* \in [\mathrm{u}]$.*

Note that $\mathcal{U}_{red} \subseteq \mathcal{U}$ implies $\underline{\dim}_{\mathrm{B}}(\mathcal{U}_{\mathrm{red}}) \leq \underline{\dim}_{\mathrm{B}}(\mathcal{U})$. The following result shows that $R > \underline{R}_{\mathrm{B}}(\varepsilon)$ is an ε-recovery guarantee.

Theorem 1. *Let $0 < \varepsilon < 1$ and $(\mathrm{x}_i)_{i \in \mathbb{N}}$ as in Definition 1. Then, for Lebesgue a.a. matrices $\mathrm{A} \in \mathbb{C}^{m \times n}$ with $m = \lfloor Rn \rfloor$, R is an ε-achievable rate provided that $R > \underline{R}_{\mathrm{B}}(\varepsilon)$.*

Proof: Since $R > \underline{R}_{\mathrm{B}}(\varepsilon)$ and $m = \lfloor Rn \rfloor$, Definition 4 implies that there exists a sequence of nonempty bounded sets $\mathcal{U}_n \subseteq \mathbb{C}^n$ and an $N(\varepsilon) \in \mathbb{N}$ such that

$$\underline{\dim}_{\mathrm{B}}(\mathcal{U}) < m \quad \text{and} \quad \mathrm{P}\big[\mathbf{x} \in \mathcal{U}\big] \geq 1 - \varepsilon \tag{3}$$

for all $\mathcal{U} = \mathcal{U}_n$ with $n \geq N(\varepsilon)$. In the remainder of the proof we assume that n is sufficiently large for (3) to hold. The claim now follows from $\underline{\dim}_{\mathrm{B}}(\mathcal{U}_{\mathrm{red}}) \leq \underline{\dim}_{\mathrm{B}}(\mathcal{U})$ and Proposition 1 below. ■

Proposition 1. *Let $\varepsilon \geq 0$, $\mathbf{x} \in \mathbb{C}^n$ a random vector, and $\mathcal{U} \subseteq \mathbb{C}^n$ a nonempty bounded set with $\mathrm{P}[\mathbf{x} \in \mathcal{U}] \geq 1 - \varepsilon$. Then, for Lebesgue a.a. matrices $\mathrm{A} \in \mathbb{C}^{m \times n}$, there exists a decoder g with $\mathrm{P}[g(|\mathrm{A}\mathbf{x}|) \not\sim \mathbf{x}] \leq \varepsilon$ provided that $\underline{\dim}_{\mathrm{B}}(\mathcal{U}_{\mathrm{red}}) < m$ for a reduced set $\mathcal{U}_{\mathrm{red}}$ of $\mathcal{U}$.*

Proof: See Section III. ■

Remark 2. It is worth noting that the proof of Proposition 1 shares several similarities with the corresponding proof in the real case [1]. However, the complex case forces us to develop a novel concentration of measure result (Lemma 3) that is based on different techniques.

For the mixed discrete-continuous case we obtain the following corollary to Theorem 1.

Corollary 1. *Let $0 < \varepsilon < 1$ and $(\mathrm{x}_i)_{i\in\mathbb{N}}$ be distributed according to the mixed discrete-continuous distribution in Example 1 with mixing parameter λ. Then, for Lebesgue a.a. matrices $\mathrm{A} \in \mathbb{C}^{m\times n}$ with $m = \lfloor Rn \rfloor$, R is ε-achievable provided that $R > 2\lambda$.*

Proposition 1 also allows us to derive a sufficient condition for the number of measurements that guarantee recovery of s-sparse vectors with arbitrarily small error probability (irrespectively of the specific distribution). To that end, let us denote by $\mathcal{V}_s \subseteq \mathbb{C}^n$ the set of all $\mathrm{x} \in \mathbb{C}^n$ with at most s nonzero entries.

Theorem 2. *Let $\mathbf{x} \in \mathbb{C}^n$ be a random vector with $\mathrm{P}[\mathbf{x} \in \mathcal{V}_s] = 1$. Then, for Lebesgue a.a. matrices $\mathrm{A} \in \mathbb{C}^{m\times n}$ and all $\varepsilon > 0$, there exists a decoder g with $\mathrm{P}[g(|\mathrm{A}\mathbf{x}|) \not\sim \mathbf{x}] \le \varepsilon$ provided that $m \ge 2s$.*

Proof: Let $\mathcal{U} = \mathcal{V}_s \cap \mathcal{B}_n(0, L)$ with L sufficiently large, s.t. $\mathrm{P}[\mathbf{x} \in \mathcal{U}] \ge 1 - \varepsilon$ holds. Observe that

$$\mathcal{U}_{\text{red}} = \bigcup_{[\mathrm{u}]\in\mathcal{U}_\sim\setminus\{[0]\}} \{\mathrm{u}_* \in [\mathrm{u}] \mid \bar{u}_* = |\bar{u}_*|\} \cup \{0\}, \qquad (4)$$

where $\bar{u}_*$ denotes the first nonzero component of u_*, is a reduced set of $\mathcal{U}$. Since $\mathcal{U}_{\text{red}} \subseteq \mathcal{V}_s$, we can write $\mathcal{U}_{\text{red}}$ as the finite union

$$\mathcal{U}_{\text{red}} = \bigcup_{\mathcal{S}\subseteq\{1,\dots,n\},|\mathcal{S}|=s} \mathcal{U}_{\text{red}}(\mathcal{S})$$

with $\mathcal{U}_{\text{red}}(\mathcal{S}) = \{\mathrm{u} \in \mathcal{U}_{\text{red}} \mid \operatorname{supp}(\mathrm{u}) \subseteq \mathcal{S}\}$. By the identification ψ in (2) and [11, Sec. 3.2, Properties (i)–(iii)],

$$\begin{aligned}\underline{\dim}_{\mathrm{B}}(\mathcal{U}_{\text{red}}) &\le \overline{\dim}_{\mathrm{B}}(\mathcal{U}_{\text{red}})\\ &= \max_{\mathcal{S}\subseteq\{1,\dots,n\},|\mathcal{S}|=s} \overline{\dim}_{\mathrm{B}}(\mathcal{U}_{\text{red}}(\mathcal{S}))\\ &= 2s - 1,\end{aligned}$$

where in the last step we used that for all $\tilde{\mathrm{u}} \in \psi(\mathcal{U}_{\text{red}}(\mathcal{S}))$ we have $\operatorname{supp}(\tilde{\mathrm{u}}) \subseteq \mathcal{R}$ for a set $\mathcal{R} \subseteq \{1, \dots, 2n\}$ with $|\mathcal{R}| = 2s - 1$ (according to (4) the first component in $\mathcal{S}$ of all $\mathrm{u} \in \mathcal{U}_{\text{red}}(\mathcal{S})$ is real-valued). ∎

III. Proof of Proposition 1

We define the decoder $g : \mathbb{R}^m_{\ge} \to \mathbb{C}^n$ by

$$g(\mathrm{y}) = \begin{cases} \mathrm{u}_0, & \text{if } \{\mathrm{u} \in \mathcal{U}_{\text{red}} \mid |\mathrm{A}\mathrm{u}| = \mathrm{y}\} = \{\mathrm{u}_0\} \\ \mathrm{e}, & \text{else} \end{cases}$$

where e is some fixed vector in the complement of $\mathcal{U}$ (used to declare a decoding error). Then,

$$\begin{aligned}&\mathrm{P}\big[g(|\mathrm{A}\mathbf{x}|) \not\sim \mathbf{x}\big]\\ &= \mathrm{P}\big[g(|\mathrm{A}\mathbf{x}|) \not\sim \mathbf{x}, \mathbf{x} \in \mathcal{U}\big] + \mathrm{P}\big[g(|\mathrm{A}\mathbf{x}|) \not\sim \mathbf{x}, \mathbf{x} \notin \mathcal{U}\big]\\ &\le \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U} \big| \mathrm{u} \not\sim \mathbf{x}, |\mathrm{A}\mathrm{u}| = |\mathrm{A}\mathbf{x}|, \mathbf{x} \in \mathcal{U}\big] + \varepsilon. \qquad (5)\end{aligned}$$

Fix an arbitrary $r > 0$. Suppose that we can show that

$$\mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U} \text{ with } \mathrm{u} \not\sim \mathrm{x}, |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big] = 0 \text{ for all } \mathrm{x} \in \mathcal{U} \quad (6)$$

where $\mathbf{A} \in \mathbb{C}^{m\times n}$ has independent rows that are uniformly distributed on $\mathcal{B}_n(0, r)$. Then,

$$\begin{aligned}&\int_{\mathcal{A}(r)} \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U} \big| \mathrm{u} \not\sim \mathbf{x}, |\mathrm{A}\mathrm{u}| = |\mathrm{A}\mathbf{x}|, \mathbf{x} \in \mathcal{U}\big]\, \mathrm{d}\mu_{\mathbf{A}}\\ &= \int_{\mathcal{U}} \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U} \text{ with } \mathrm{u} \not\sim \mathrm{x}, |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big]\, \mathrm{d}\mu_{\mathbf{x}}\\ &= 0 \qquad (7)\end{aligned}$$

where we used Fubini's Theorem and set $\mathcal{A}(r) = \mathcal{B}_n(0, r) \times \dots \times \mathcal{B}_n(0, r)$. Since r is arbitrary, (7) implies that

$$\mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U} \big| \mathrm{u} \not\sim \mathbf{x}, |\mathrm{A}\mathrm{u}| = |\mathrm{A}\mathbf{x}|, \mathbf{x} \in \mathcal{U}\big] = 0 \qquad (8)$$

for Lebesgue a.a. matrices A. Hence, combining (5) and (8) proves the Proposition provided that we can show that (6) holds, which is done in Section IV, Lemma 2. ∎

IV. Proof of (6)

We first present an auxiliary lemma.

Lemma 1. *Let $r > 0$, $\emptyset \neq \mathcal{S} \subseteq \mathcal{B}_n(0, L)$, $\rho > 0$, $\mathrm{x} \in \mathcal{B}_n(0, L)$, and $\mathbf{A} \in \mathbb{C}^{m\times n}$ with independent rows that are uniformly distributed on $\mathcal{B}_n(0, r)$. Then, there exist $\mathrm{s}_l(\rho) \in \mathcal{S}$, $l = 1,\dots, N_{\mathcal{S}}(\rho)$ with $N_{\mathcal{S}}(\rho)$ being the covering number of $\mathcal{S}$, such that*

$$\begin{aligned}&\mathrm{P}\big[\exists \mathrm{u} \in \mathcal{S} \text{ with } \big\||\mathbf{A}\mathrm{u}| - |\mathbf{A}\mathrm{x}|\big\| \le \rho\big]\\ &\le \sum_{l=1}^{N_{\mathcal{S}}(\rho)} \mathrm{P}\big[\big||\mathbf{a}^{\mathsf{H}}\mathrm{s}_l(\rho)|^2 - |\mathbf{a}^{\mathsf{H}}\mathrm{x}|^2\big| \le 2Lr(2r+1)\rho\big]^m \qquad (9)\end{aligned}$$

where $\mathbf{a}$ is uniformly distributed on $\mathcal{B}_n(0, r)$.

Proof: Identical to the proof of [1, Lem. 2] with $\mathbb{R}$ replaced by $\mathbb{C}$ and the superscript $^{\mathsf{T}}$ replaced by $^{\mathsf{H}}$. ∎

Lemma 2. *Let $r > 0$, $\emptyset \neq \mathcal{U} \subseteq \mathcal{B}_n(0, L)$ with $\underline{\dim}_{\mathrm{B}}(\mathcal{U}_{\text{red}}) < m$, $\mathrm{x} \in \mathcal{U}$, and $\mathbf{A} \in \mathbb{C}^{m\times n}$ with independent rows that are uniformly distributed on $\mathcal{B}_n(0, r)$. Then,*

$$\mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U} \text{ with } \mathrm{u} \not\sim \mathrm{x}, |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big] = 0.$$

Proof: First note that the claim is equivalent to

$$P(\mathrm{x}) = \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U}_{\text{red}} \text{ with } \mathrm{u} \not\sim \mathrm{x}, |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big] = 0.$$

Hence, in case $\mathrm{x} = 0$ it is sufficient to show $\mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U}_{\text{red}}\setminus\{0\} \mid \mathbf{A}\mathrm{u} = 0\big] = 0$, which is accomplished by repeating the steps of the corresponding proofs of [12, Lem. 3, Prop. 1] for the real case (omitted here). In the following we therefore assume $\mathrm{x} \neq 0$. We can upper-bound $P(\mathrm{x}) \le P_1(\mathrm{x}) + P_2(\mathrm{x})$ with

$$\begin{aligned}P_1(\mathrm{x}) &= \mathrm{P}\big[\exists \mathrm{u} \in \bar{\mathcal{U}}(\mathrm{x}) \text{ with } |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big],\\ \bar{\mathcal{U}}(\mathrm{x}) &= \{\mathrm{u} \in \mathcal{U}_{\text{red}} \mid \operatorname{rank}(\mathrm{x}, \mathrm{u}) = 2\}\end{aligned}$$

and

$$P_2(\mathrm{x}) = \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{U}_{\text{red}} \text{ with } \mathrm{u} = c\,\mathrm{x}, |c| \neq 1, |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big]$$

$$\leq \mathrm{P}\big[\exists c \in \mathbb{C} \text{ with } |c| \neq 1, |c||\mathbf{A}\mathrm{x}| = |\mathbf{A}\mathrm{x}|\big]$$
$$= \mathrm{P}\big[\mathbf{A}\mathrm{x} = 0\big]$$
$$= 0.$$

It remains to show that $P_1(\mathrm{x}) = 0$. To this end, we define the sets $\mathcal{T}^{(j)}(\mathrm{x})$ by

$$\mathcal{T}^{(j)}(\mathrm{x}) = \left\{ \mathrm{u} \in \bar{\mathcal{U}}(\mathrm{x}) \,\middle|\, \|\mathrm{u}\|^2\|\mathrm{x}\|^2 - |\mathrm{u}^{\mathsf{H}}\mathrm{x}|^2 > \frac{1}{j} \right\}, \quad j \in \mathbb{N}.$$

Since

$$P_1(\mathrm{x}) \leq \sum_{j\in\mathbb{N}} \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{T}^{(j)}(\mathrm{x}) \text{ with } |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big],$$

it is sufficient to prove that

$$P^{(j)}(\mathrm{x}) = \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{T}^{(j)}(\mathrm{x}) \text{ with } |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big] = 0$$

for all $j \in \mathbb{N}$. Suppose, by contradiction, that there exists a $j \in \mathbb{N}$ such that $P^{(j)}(\mathrm{x}) > 0$. Then, $\mathcal{T}^{(j)}(\mathrm{x}) \neq \emptyset$ and

$$\underline{\dim}_{\mathrm{B}}\big(\mathcal{T}^{(j)}(\mathrm{x})\big) \leq \underline{\dim}_{\mathrm{B}}(\bar{\mathcal{U}}(\mathrm{x})) < m.$$

Furthermore, we have

$$0 = \liminf_{\rho\to 0} \frac{\log P^{(j)}(\mathrm{x})}{\log \frac{1}{\rho}}$$
$$= \liminf_{\rho\to 0} \frac{\log \mathrm{P}\big[\exists \mathrm{u} \in \mathcal{T}^{(j)}(\mathrm{x}) \text{ with } |\mathbf{A}\mathrm{u}| = |\mathbf{A}\mathrm{x}|\big]}{\log \frac{1}{\rho}}$$
$$\leq \liminf_{\rho\to 0} \frac{\log\Big(\sum_{l=1}^{N_{\mathcal{T}^{(j)}(\mathrm{x})}(\rho)} \mathrm{P}\Big[\big||\mathbf{a}^{\mathsf{H}}\mathrm{s}_l^{(j)}(\rho,\mathrm{x})|^2 - |\mathbf{a}^{\mathsf{H}}\mathrm{x}|^2\big| \leq \tilde{\rho}\Big]^m\Big)}{\log \frac{1}{\rho}} \tag{10}$$
$$\leq \liminf_{\rho\to 0} \frac{\log\Big(\tilde{\rho}^m \sum_{l=1}^{N_{\mathcal{T}^{(j)}(\mathrm{x})}(\rho)} (f(\tilde{\rho}, r, \mathrm{s}_l^{(j)}(\rho,\mathrm{x}), \mathrm{x}))^m\Big)}{\log \frac{1}{\rho}} \tag{11}$$
$$\leq \liminf_{\rho\to 0} \frac{\log\big(\tilde{\rho}^m N_{\mathcal{T}^{(j)}(\mathrm{x})}(\rho)(\tilde{f}(\tilde{\rho}, r, L, j))^m\big)}{\log \frac{1}{\rho}} \tag{12}$$
$$= \underline{\dim}_{\mathrm{B}}(\mathcal{T}^{(j)}(\mathrm{x})) - m + m \underbrace{\lim_{\rho\to 0} \frac{\log \tilde{f}(\tilde{\rho}, r, L, j)}{\log \frac{1}{\rho}}}_{=0} \tag{13}$$
$$< 0, \tag{14}$$

where in (10) we applied Lemma 1 with $\mathcal{S} = \mathcal{T}^{(j)}(\mathrm{x})$ and set $\tilde{\rho} = 2Lr(2r+1)\rho$, (11) follows from Lemma 3 below with $\mathrm{u} = \mathrm{s}_l^{(j)}(\rho,\mathrm{x})$, $\mathrm{v} = \mathrm{x}$, and $\delta = \tilde{\rho}$ where f is defined in (15), and in (12) we used that

$$f\big(\tilde{\rho}, r, \mathrm{s}_l^{(j)}(\rho,\mathrm{x}), \mathrm{x}\big)$$
$$\leq \tilde{f}(\tilde{\rho}, r, L, j)$$
$$= \frac{2(2r)^{2(n-2)}\pi^2 j}{V(n,r)}\big(\tilde{\rho} + 2L^2r^2\big), \quad l = 1,\dots,N_{\mathcal{T}^{(j)}(\mathrm{x})}(\rho),$$

which follows from $\mathrm{s}_l^{(j)}(\rho,\mathrm{x}) \in \mathcal{T}^{(j)}(\mathrm{x})$, $l = 1,\dots,N_{\mathcal{T}^{(j)}(\mathrm{x})}(\rho)$, hence, $\|\mathrm{s}_l^{(j)}(\rho,\mathrm{x})\| \leq L$, and $\|\mathrm{x}\| \leq L$. Now (14) is a contradiction to $P^{(j)}(\mathrm{x}) > 0$. This implies that $P(\mathrm{x}) = P_1(\mathrm{x}) + P_2(\mathrm{x}) = 0$ and concludes the proof of the lemma. ∎

V. Concentration of Measure Result

Lemma 3. *Let $r > 0$, $\mathbf{a}$ be uniformly distributed on $\mathcal{B}_n(0,r)$, $\mathrm{C} = \mathrm{u}\mathrm{u}^{\mathsf{H}} - \mathrm{v}\mathrm{v}^{\mathsf{H}}$ with linearly independent vectors $\mathrm{u},\mathrm{v} \in \mathbb{C}^n$, and $\delta > 0$. Then*

$$\mathrm{P}\big[|\mathbf{a}^{\mathsf{H}}\mathrm{C}\mathbf{a}| \leq \delta\big] \leq \delta f(\delta, r, \mathrm{u}, \mathrm{v})$$

with

$$f(\delta, r, \mathrm{u}, \mathrm{v}) = \frac{2(2r)^{2(n-2)}\pi^2\Big(\delta + (\|\mathrm{u}\|^2 + \|\mathrm{v}\|^2)r^2\Big)}{(\|\mathrm{u}\|^2\|\mathrm{v}\|^2 - |\mathrm{u}^{\mathsf{H}}\mathrm{v}|^2)V(n,r)} \tag{15}$$

Proof: We have

$$\mathrm{P}\big[|\mathbf{a}^{\mathsf{H}}\mathrm{C}\mathbf{a}| \leq \delta\big]$$
$$= \frac{1}{V(n,r)} \int_{\mathcal{B}_n(0,r)} \chi_{\left\{\mathrm{a}\in\mathbb{C}^n \,\middle|\, |\mathrm{a}^{\mathsf{H}}\mathrm{C}\mathrm{a}| < \delta\right\}} \,\mathrm{da}$$
$$= \frac{1}{V(n,r)} \int_{\mathcal{B}_n(0,r)} \chi_{\left\{\mathrm{a}\in\mathbb{C}^n \,\middle|\, |\mathrm{a}^{\mathsf{H}}\mathrm{W}\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}\mathrm{W}^{\mathsf{H}}\mathrm{a}| < \delta\right\}} \,\mathrm{da} \tag{16}$$
$$= \frac{1}{V(n,r)} \int_{\mathcal{B}_n(0,r)} \chi_{\left\{\mathrm{b}\in\mathbb{C}^n \,\middle|\, |\mathrm{c}^{\mathsf{H}}\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}\mathrm{c}| < \delta\right\}} \,\mathrm{db} \tag{17}$$
$$\leq \frac{(2r)^{2(n-2)}}{V(n,r)} \int_{\mathcal{B}_2(0,r)} \chi_{\left\{\mathrm{c}\in\mathbb{C}^2 \,\middle|\, |\mathrm{c}^{\mathsf{H}}\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}\mathrm{c}| < \delta\right\}} \,\mathrm{dc} \tag{18}$$

where (16) follows from Lemma 4 with R and J defined in (24) and W defined in (25) and (17) follows from changing variables to $\mathrm{a} = \bar{\mathrm{W}}\mathrm{b}$ with $\bar{\mathrm{W}} = (\mathrm{W}, \mathrm{Z}) \in \mathbb{C}^{n\times n}$ where $\mathrm{Z} \in \mathbb{C}^{n\times(n-2)}$ is chosen in such a way that $\bar{\mathrm{W}}\bar{\mathrm{W}}^{\mathsf{H}} = \mathrm{I}$ and $\mathrm{c} = (c_1, c_2)^{\mathsf{T}}$ with $c_1 = b_1$ and $c_2 = b_2$.

The bound (27) on the determinant of the matrix $\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}$ implies that one eigenvalue of $\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}$, say λ_1, is positive and the other eigenvalue of $\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}$, say $-\lambda_2$, is negative. We can assume without loss of generality that $\lambda_1 \geq \lambda_2$. Using the eigendecomposition $\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}} = \mathrm{U}\,\mathrm{diag}(\lambda_1, -\lambda_2)\mathrm{U}^{\mathsf{H}}$, where $\mathrm{U} \in \mathbb{C}^{2\times 2}$ with $\mathrm{U}\mathrm{U}^{\mathsf{H}} = \mathrm{I}$, and changing variables to $\mathrm{c} = \mathrm{U}\mathrm{d}$, we can further upper bound (18) by

$$\frac{(2r)^{2(n-2)}}{V(n,r)} \int_{\mathcal{B}_2(0,r)} \chi_{\left\{\mathrm{c}\in\mathbb{C}^2 \,\middle|\, |\mathrm{c}^{\mathsf{H}}\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}\mathrm{c}| < \delta\right\}} \,\mathrm{dc}$$
$$= \frac{(2r)^{2(n-2)}}{V(n,r)} \int_{\mathcal{B}_2(0,r)} \chi_{\left\{\mathrm{d}\in\mathbb{C}^2 \,\middle|\, |\lambda_1|d_1|^2 - \lambda_2|d_2|^2| < \delta\right\}} \,\mathrm{dd}$$
$$= \frac{(2r)^{2(n-2)}}{\lambda_1\lambda_2 V(n,r)} \int_{\mathbb{C}^2} \chi_{\left\{\mathrm{t}\in\mathbb{C}^2 \,\middle|\, \frac{|t_1|^2}{\lambda_1} + \frac{|t_2|^2}{\lambda_2} \leq r^2\right\}} \times \chi_{\left\{\mathrm{t}\in\mathbb{C}^2 \,\middle|\, ||t_1|^2 - |t_2|^2| < \delta\right\}} \,\mathrm{dt} \tag{19}$$
$$\leq \frac{(2r)^{2(n-2)}}{\lambda_1\lambda_2 V(n,r)} \int_{\mathbb{C}^2} \chi_{\left\{\mathrm{t}\in\mathbb{C}^2 \,\middle|\, |t_1|^2 \leq \lambda_1 r^2, |t_2|^2 \leq \lambda_2 r^2\right\}} \times \chi_{\left\{\mathrm{t}\in\mathbb{C}^2 \,\middle|\, ||t_1|^2 - |t_2|^2| < \delta\right\}} \,\mathrm{dt}$$
$$\leq \frac{(2r)^{2(n-2)}}{\lambda_1\lambda_2 V(n,r)} \int_{\mathbb{C}^2} \chi_{\left\{\mathrm{t}\in\mathbb{C}^2 \,\middle|\, |t_1|^2 + |t_2|^2 \leq \delta + 2\lambda_2 r^2\right\}}$$

$$\times \chi_{\left\{t\in\mathbb{C}^2 \middle| ||t_1|^2-|t_2|^2|<\delta\right\}} \mathrm{dt} \quad (20)$$

$$= \frac{(2r)^{2(n-2)}\pi^2}{\lambda_1\lambda_2 V(n,r)} \int_{\mathbb{R}^2_{\geq}} \chi_{\left\{z\in\mathbb{R}^2 \middle| z_1+z_2\leq\delta+2\lambda_2 r^2\right\}} \times \chi_{\left\{z\in\mathbb{R}^2 \middle| |z_1-z_2|<\delta\right\}} \mathrm{dz} \quad (21)$$

$$\leq \frac{(2r)^{2(n-2)}\pi^2}{\lambda_1\lambda_2 V(n,r)} \int_{\mathbb{R}^2_{\geq}} \chi_{\left\{z\in\mathbb{R}^2 \middle| z_1\leq\delta+2\lambda_2 r^2\right\}} \times \chi_{\left\{z\in\mathbb{R}^2 \middle| |z_1-z_2|<\delta\right\}} \mathrm{dz}$$

$$\leq \frac{(2r)^{2(n-2)}\pi^2}{\lambda_1\lambda_2 V(n,r)} \int_0^{\delta+2\lambda_2 r^2} 2\delta \,\mathrm{d}z_1$$

$$= \frac{2\delta(2r)^{2(n-2)}\pi^2}{\lambda_1\lambda_2 V(n,r)}\left(\delta+2\lambda_2 r^2\right) \quad (22)$$

where in (19) we changed variables to $\mathrm{t} = \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2})\mathrm{d}$, in (20) we used that $|\mathrm{t}_2|^2 \leq \lambda_2 r^2$ and $||t_1|^2 - |t_2|^2| < \delta$ imply $|t_1|^2 + |t_2|^2 \leq \delta + 2\lambda_2 r^2$, and in (21) we changed variables to $t_1 = \sqrt{z_1}e^{i\varphi_1}$ and $t_2 = \sqrt{z_2}e^{i\varphi_2}$. Combining (18) with (22) and using the expressions (27) and (29) from Lemma 4 gives (15). ■

Lemma 4. *Let* $\mathrm{u}, \mathrm{v} \in \mathbb{C}^n$ *be linearly independent and* $\mathrm{C} = \mathrm{u}\mathrm{u}^{\mathsf{H}} - \mathrm{v}\mathrm{v}^{\mathsf{H}}$. *Then,*

$$\mathrm{C} = \mathrm{W}\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}\mathrm{W}^{\mathsf{H}} \quad (23)$$

with

$$\mathrm{J} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathrm{R} = \begin{pmatrix} \|\mathrm{u}\| & \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|} \\ 0 & \|\mathrm{v} - \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|^2}\mathrm{u}\| \end{pmatrix} \quad (24)$$

and

$$\mathrm{W} = \left(\tfrac{\mathrm{a}}{\|\mathrm{a}\|}, \tfrac{\mathrm{b}}{\|\mathrm{b}\|}\right) \quad (25)$$

where the orthonormal vectors $\mathrm{a}/\|\mathrm{a}\|$ *and* $\mathrm{b}/\|\mathrm{b}\|$ *are defined by*

$$\mathrm{a} = \mathrm{u}, \qquad \mathrm{b} = \mathrm{v} - \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|^2}\mathrm{a}. \quad (26)$$

Moreover,

$$\det(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) = |\mathrm{u}^{\mathsf{H}}\mathrm{v}|^2 - \|\mathrm{u}\|^2\|\mathrm{v}\|^2 < 0 \quad (27)$$

$$\mathrm{tr}(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) = \|\mathrm{u}\|^2 - \|\mathrm{v}\|^2 \quad (28)$$

$$\sigma_2(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) \leq \frac{1}{2}(\|\mathrm{u}\|^2 + \|\mathrm{v}\|^2) \quad (29)$$

where $\sigma_1(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) \geq \sigma_2(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}})$ *are the singular values of* $\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}$.

Proof: We can rewrite $\mathrm{C} = \mathrm{A}\mathrm{J}\mathrm{A}^{\mathsf{H}}$ with $\mathrm{A} = (\mathrm{u}, \mathrm{v})$. Hence, to prove (23), it is sufficient to show that $\mathrm{A} = \mathrm{W}\mathrm{R}$.

Using the definitions of the vectors a and b in (26), we can rewrite

$$\mathrm{A} = \left(\mathrm{a}, \tfrac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|^2}\mathrm{a} + \mathrm{b}\right) = (\mathrm{a}, \mathrm{b})\begin{pmatrix} 1 & \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|^2} \\ 0 & 1 \end{pmatrix} = \left(\tfrac{\mathrm{a}}{\|\mathrm{a}\|}, \tfrac{\mathrm{b}}{\|\mathrm{b}\|}\right)\begin{pmatrix} \|\mathrm{u}\| & \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|} \\ 0 & \|\mathrm{v} - \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|^2}\mathrm{u}\| \end{pmatrix} = \mathrm{W}\mathrm{R}$$

which proves (23).

The explicit form of the determinant in (27) follows from the fact that

$$\begin{aligned} \det(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) &= \det(\mathrm{R})\det(\mathrm{J})\det(\mathrm{R}^{\mathsf{H}}) \\ &= -|\det(\mathrm{R})|^2 \\ &= -\|\mathrm{u}\|^2\left\|\mathrm{v} - \frac{\mathrm{u}^{\mathsf{H}}\mathrm{v}}{\|\mathrm{u}\|^2}\mathrm{u}\right\|^2 \\ &= |\mathrm{u}^{\mathsf{H}}\mathrm{v}|^2 - \mathrm{u}^{\mathsf{H}}\mathrm{u}\mathrm{v}^{\mathsf{H}}\mathrm{v} \\ &< 0 \end{aligned} \quad (30)$$

where (30) follows from the Cauchy-Schwarz inequality [13, Sec. 0.6.3] and u and v being linearly independent. The expression for the trace (28) follows from $\mathrm{tr}(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) = \mathrm{tr}(\mathrm{C})$. Finally, (29) follows from

$$\begin{aligned} \sigma_2(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) &= \frac{1}{2}\big(\sigma_1(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) + \sigma_2(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}})\big) - \frac{1}{2}\big(\sigma_1(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}}) - \sigma_2(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}})\big) \\ &= \frac{1}{2}\sqrt{\mathrm{tr}(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}})^2 - 4\det(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}})} - \frac{1}{2}|\mathrm{tr}(\mathrm{R}\mathrm{J}\mathrm{R}^{\mathsf{H}})| \\ &= \frac{1}{2}\sqrt{(\|\mathrm{u}\|^2+\|\mathrm{v}\|^2)^2 - 4|\mathrm{u}^{\mathsf{H}}\mathrm{v}|^2} - \frac{1}{2}\big|\|\mathrm{u}\|^2-\|\mathrm{v}\|^2\big| \\ &\leq \frac{1}{2}(\|\mathrm{u}\|^2 + \|\mathrm{v}\|^2). \end{aligned}$$

■

REFERENCES

[1] E. Riegler and G. Tauböck, "Almost lossless analog compression without phase information," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, China, Jun. 2015.

[2] R. P. Millane, "Phase retrieval in crystallography and optics," *J. Opt. Soc. Am. A*, vol. 7, no. 3, pp. 394–411, 1990.

[3] J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin, "Hubble space telescope characterized by using phase-retrieval algorithms," *Appl. Opt.*, vol. 32, no. 10, pp. 1747–1767, 1993.

[4] H. M. Quiney, "Coherent diffractive imaging using short wavelength light sources," *J. Mod. Opt.*, vol. 57, no. 13, pp. 1109–1149, 2010.

[5] Y. M. Bruck and L. Sodin, "On the ambiguity of the image reconstruction problem," *Opt. Comm.*, vol. 30, no. 3, pp. 304–308, 1979.

[6] A. Conca, D. Edidin, M. Hering, and C. Vinzant, "An algebraic characterization of injectivity in phase retrieval," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 2, pp. 346–356, 2015.

[7] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Appl. Comput. Harmon. Anal.*, vol. 20, no. 3, pp. 345–356, 2006.

[8] E. J. Candès and X. Li, "Solving quadratic equations via PhaseLift when there are about as many equations as unknowns," *Found. Comput. Mathematics*, vol. 14, no. 5, pp. 1017–1026, 2012.

[9] M. Akçakaya and V. Tarokh, "New conditions for sparse phase retrieval," 2014, arXiv:1310.1351 [cs.IT].

[10] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3721–3748, Aug. 2010.

[11] K. Falconer, *Fractal Geometry*, 1st ed. New York, NY: Wiley, 1990.

[12] D. Stotz, E. Riegler, and H. Bölcskei, "Almost lossless analog signal separation," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 106–110.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.