Error Bounds for FDD Massive MIMO Channel Covariance Conversion with Set-Theoretic Methods

Renato L. G. Cavalcante,[†] L. Miretti,[‡] S. Stańczak[†]

[†] Fraunhofer Heinrich Hertz Institute and Technical University of Berlin, Germany

[‡] EURECOM, France

Abstract—We derive novel bounds for the performance of algorithms that estimate the downlink covariance matrix from the uplink covariance matrix in frequency division duplex (FDD) massive multiple-input multiple-output (MIMO) systems. The focus is on algorithms that use estimates of the angular power spectrum as an intermediate step. Unlike previous results, the proposed bounds follow from simple arguments in possibly infinite dimensional Hilbert spaces, and they do not require strong assumptions on the array geometry or on the propagation model. Furthermore, they are suitable for the analysis of set-theoretic methods that can efficiently incorporate side information about the angular power spectrum. This last feature enables us to derive simple techniques to enhance set-theoretic methods without any heuristic arguments. In particular, we show that the performance of a simple algorithm that requires only a simple matrix-vector multiplication cannot be improved significantly in some practical scenarios, especially if coarse information about the support of the angular power spectrum is available.

I. Introduction

Owing to the lack of channel reciprocity in the frequency division duplex (FDD) mode, most proposals for FDD massive multiple-input multiple-output (MIMO) systems envision some sort of feedback from the user equipment for downlink channel state acquisition at the base stations [1]–[7]. If traditional channel feedback mechanisms are employed, and the number of antennas is large, downlink channel estimation in FDD massive MIMO systems may incur a prohibitive feedback overhead. Therefore, approaches for reducing this overhead have received a great deal of attention in recent years. Many promising approaches for reducing this overhead exploit information about the downlink covariance matrix [1], [6], which is estimated from the uplink covariance matrix [2]–[5]. In this study, we address this estimation problem, hereafter called the uplink-downlink conversion problem.

Although channel reciprocity is lost in FDD massive MIMO systems, estimating the downlink covariance matrix from the uplink covariance matrix is possible by exploiting a different form of reciprocity, the so-called reciprocity of the angular power spectrum [3]–[6]. The basic assumption behind this form of reciprocity is that the average receive/transmit power at a unit of angle (hereafter called *angular power spectrum*) at an antenna array is frequency invariant because, if the frequency separation is not too large, the scattering environments are the same for both the uplink and downlink channels. Building upon this characteristic of wireless channels, researchers have proposed to estimate the angular power spectrum from the uplink covariance matrix, with the intent to use this estimate in models of the antenna array to recover the downlink covariance matrix [3]–[5] .

In particular, the algorithms in [3], [4] estimate the angular power spectrum with set-theoretic methods that can easily include side information expressed in terms of closed convex sets in a Hilbert space. Despite working in possibly infinite dimensional spaces, one of the approaches in [3], [4] have shown that good uplink-downlink conversion performance can be obtained with a very simple matrix-vector multiplication. In that scheme with remarkably low computational complexity, the matrix is computed only once for the entire system lifetime, and the vector is constructed by rearranging the components of the uplink covariance matrix to be converted. If additional information about the angular power spectrum is used, the studies in [3], [4] have also shown that the conversion performance of set-theoretic approaches can be further improved with simple fixed point algorithms that do not appeal to finite-dimensional approximations of the physical models. These approaches can easily take into account the polarization of antennas and real-world impairments (e.g., dissimilarities of the antennas in the array), but performance bounds on the conversion performance have not been considered in [3], [4].

More recently, by using ideal antenna models in uniform linear arrays, the study in [5] has proved that, in uplink-downlink channel covariance conversion based on algorithms that first estimate the angular power spectrum (such as those in [3], [4]), some of the components of the downlink covariance matrix can be reliably reconstructed. Based on this observation, that study has derived a scheme in which the angular power spectrum is first estimated by using solvers for nonnegative least square problems (this first step can be interpreted as a finite-dimensional approximation of a particular case of [3, Algorithm 2]). This estimate is then used to reconstruct the downlink covariance matrix, and, based on a formal analysis of the reliability of the reconstruction, the authors of [5] have proposed to set to zero the components of the downlink covariance matrix that are not guaranteed to be reliably estimated, in an approach called *truncation*. However, setting to zero these components is a somewhat heuristic approach that can actually decrease the performance of the reconstruction in some scenarios, as the simulations in [5] have already shown. Furthermore, the reliability analysis does not seem easy to extend to realistic propagation and antenna models such as those in [4], or to cases where the antenna array response is measured to mitigate modeling errors.

Inspired by the findings in [5], we derive novel bounds on the reconstruction error of each component of the downlink covariance matrix. Unlike previous results, the proposed performance bounds do not assume any particular antenna

array or propagation model, so they can be easily applied to the realistic models in [4] (without any changes) or to cases where the array response is measured. In addition, if the angular power spectrum is normalized (as assumed in [5]), a possible value for the only unknown multiplying constant in the bounds is trivial to determine. The bounds are based on elementary arguments in Hilbert spaces, and they also provide insights to improve set-theoretic algorithms. In particular, if information about the support of the angular power spectrum is used appropriately, we show that the simplest of the algorithms for uplink-downlink conversion in [3], [4] can be so effective that performance improvements obtained with more computationally demanding conversion mechanisms are marginal at best (NOTE: the proposed bounds can also be used in the analysis of these complex mechanisms). This result is particularly appealing because that simple algorithm has no parameters to be tuned, and all steps of the enhancements we propose are justified by rigorous arguments.

This study is structured as follows. In Sect. II we introduce the main mathematical concepts, and we prove a simple result (Proposition 1) that is the main mathematical tool used to derive the novel bounds. The general result in Sect. II is specialized to the problem of uplink-downlink covariance matrix conversion in Sect. III, which also discusses how to exploit the proposed bounds to enhance set-theoretic methods (see Sect. III-C). To keep the presentation as general as possible, we do not assume any particular antenna array or propagation model in Sect. III. A concrete application of the theory developed here is shown in Sect. IV.

II. MATHEMATICAL PRELIMINARIES

In the following, we use $\mathbb{R}_+$ to indicate nonnegative reals. The (coordinate-wise) real and imaginary components of complex vectors or matrices are given by, respectively, $\mathrm{Real}(\cdot)$ and $\mathrm{Imag}(\cdot)$, and i is the imaginary unit, which is the solution to $i^2 = -1$. By $(\cdot)^T$ we denote the transpose of a matrix or vector. Given a matrix $\boldsymbol{M} \in \mathbb{R}^{Q \times N}$, $\mathrm{vec}(\boldsymbol{M}) \in \mathbb{R}^{QN}$ is the vector obtained by stacking the columns of $\boldsymbol{M}$, and $\boldsymbol{M}^\dagger$ is the Moore-Penrose (pseudo-)inverse of $\boldsymbol{M}$. We denote by $\mathcal{H}$ a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$. Given $x \in \mathcal{H}$ and a set $C \subset \mathcal{H}$, we define $x + C := \{ h + x \in \mathcal{H} \mid h \in C \}$. A *linear variety* $V \subset \mathcal{H}$ is a set that can be expressed as $V = x + M$ for a vector $x \in \mathcal{H}$ and a subspace $M \subset \mathcal{H}$; i.e., V is a translation of the subspace M. If $C \subset \mathcal{H}$ is a nonempty closed convex set, the projection $P_C : \mathcal{H} \to C$ maps a vector $x \in \mathcal{H}$ to the uniquely existing vector $y_0 \in C$ satisfying $(\forall y \in C) \ \|x - y_0\| \le \|x - y\|$. The orthogonal complement of a subset $C \subset \mathcal{H}$ is the closed subspace given by $C^\perp := \{ y \in \mathcal{H} \mid (\forall x \in C) \ \langle x, y \rangle = 0 \}$, and note that $M \cup M^\perp = \mathcal{H}$ and $(M^\perp)^\perp = M$ for any closed subspace M. The closure of a set $C \subset \mathcal{H}$ is denoted by $\overline{C}$. A set $S = \{ x_1, \ldots, x_N \} \subset \mathcal{H}$ is called *linear independent* (respectively, dependent) if the vectors $x_1, \ldots, x_N$ are linearly independent (respectively, dependent). Given a function $f : \Omega \to \mathbb{R}$ with $\Omega \subset \mathbb{R}^N$, we define its support to be the set $\mathrm{Supp}(f) = \{ x \in \Omega \mid f(x) \ne 0 \}$. By $L^2(\Omega)$, $\Omega \subset \mathbb{R}^N$, we denote the space of real-valued square-integrable functions $f : \Omega \to \mathbb{R}$ with respect to the standard Lebesgue measure.

Below is a summary of standard results in convex analysis that we use throughout this study. The proof can be found in most standard references on convex and functional analysis (e.g., [8], [9]).

Fact 1. (i) *Let $M \subset \mathcal{H}$ be a closed (linear) subspace. Then $(\forall x \in \mathcal{H}) \ x = P_M(x) + P_{M^\perp}(x)$. Furthermore, the projection $P_M : \mathcal{H} \to M$ onto M is a bounded linear operator with operator norm given by $\|P_M\|_\mathrm{o} := \sup_{\|x\|=1} \|P_M(x)\| \le 1$, and the equality is achieved if $M \ne \{0\}$.*

(ii) *Let $M \subset \mathcal{H}$ be a closed subspace. For a given $u \in \mathcal{H}$, consider the closed linear variety $V = u + M$. Then $(\forall x \in V) \ V = x + M$ and $x = P_V(0) + P_M(x)$.*

(iii) *Let $V = \cap_{k=1}^K \{ x \in \mathcal{H} \mid \langle x, v_k \rangle = b_k \} \ne \emptyset$, where $(v_k, \ b_k) \in \mathcal{H} \times \mathbb{R}$ for each $k \in \{1, \ldots, K\}$. Then we have $(\forall u \in V) \ V = u + M^\perp$, where $M = \mathrm{span}\{v_1, \ldots, v_K\}$.*

In the application described in the next section, we study the performance of algorithms producing an estimate $\widetilde{\rho} \in V$ of $\rho \in V$, where V is a closed linear variety generated by the translation of the orthogonal complement $M^\perp$ of a finite dimensional subspace M (NOTE: the closed subspace $M^\perp$ can be infinite dimensional). In that application, we are not directly interested in the error $\|\widetilde{\rho} - \rho\|$, but in the approximation of $\langle \rho, y \rangle$ by $\langle \widetilde{\rho}, y \rangle$ for a given $y \in \mathcal{H}$. The absolute error of the approximation is given by $e := |\langle \widetilde{\rho} - \rho, y \rangle|$, and we show in Proposition 1 elementary bounds for e that decouples into the product of two terms. The important aspect to note is that one of these terms only depends on the choice of the algorithm, but not on the estimand ρ.

Proposition 1. *Let $M \subset \mathcal{H}$ be a closed subspace, and consider the linear variety $V = \rho + M^\perp$ for a given $\rho \in \mathcal{H}$. Suppose that an algorithm produces an estimate $\widetilde{\rho} \in V$ of $\rho \in V$. Then each of the following holds:*

(i) $(\forall y \in \mathcal{H})$

$$|\langle \widetilde{\rho} - \rho, y \rangle| \le \|\widetilde{\rho} - \rho\| \ \|y - P_M(y)\|$$
$$= \|P_{M^\perp}(\widetilde{\rho}) - P_{M^\perp}(\rho)\| \ \|y - P_M(y)\| \quad (1)$$

(ii) *If $\widetilde{\rho} = P_V(0)$, then*

$$(\forall y \in \mathcal{H}) \ |\langle \widetilde{\rho} - \rho, y \rangle| \le \|\rho - P_M(\rho)\| \ \|y - P_M(y)\| \tag{2}$$

Proof. (i) Let $y \in \mathcal{H}$ be arbitrary. By assumption, both $\widetilde{\rho}$ and ρ are elements of the linear variety V, so we have

$$\widetilde{\rho} - \rho = P_{M^\perp}(\widetilde{\rho}) + P_V(0) - P_{M^\perp}(\rho) - P_V(0)$$
$$= P_{M^\perp}(\widetilde{\rho}) - P_{M^\perp}(\rho) = P_{M^\perp}(\widetilde{\rho} - \rho) \quad (3)$$

by Fact 1(i)-(ii). From $y = P_M(y) + P_{M^\perp}(y)$ (Fact 1(i)) and the definition of orthogonal complements, we obtain

$$|\langle \widetilde{\rho} - \rho, y \rangle| = |\langle P_{M^\perp}(\widetilde{\rho} - \rho), P_M(y) + P_{M^\perp}(y) \rangle|$$
$$= |\langle P_{M^\perp}(\widetilde{\rho} - \rho), P_{M^\perp}(y) \rangle| .$$

A direct application of the Cauchy-Schwartz inequality yields $|\langle \widetilde{\rho} - \rho, y \rangle| \le \|P_{M^\perp}(\widetilde{\rho} - \rho)\| \ \|P_{M^\perp}(y)\|$. The result in (1) now follows from $P_{M^\perp}(y) = y - P_M(y)$ (Fact 1(i)) and (3).

(ii) By Fact 1(i)-(ii) and $\widetilde{\rho} = P_V(0)$, we deduce $P_{M^\perp}(\widetilde{\rho}) = 0$. Now use the equality $\rho = P_{M^\perp}(\rho) + P_M(\rho)$ (Fact 1(i)) in (1) to obtain the desired result. $\square$

Proposition 1 has a very natural interpretation. If the estimation error $\|\widetilde{\rho} - \rho\|$ of $\rho \in V$ is bounded, then (1) shows that the error $e := |\langle \widetilde{\rho} - \rho, y \rangle|$ is small if there exists a vector $u \in M$ that is sufficiently close to $y \in \mathcal{H}$ with respect to the metric $d(u, y) = \|u - y\|$. In this case, the choice of the algorithm used to produce the estimate $\widetilde{\rho} \in V$ does not play a decisive role in the minimization of the error e. Proposition 1(ii) shows the guaranteed performance bound of a simple algorithm for estimating ρ. For this scheme, the error e is small if ρ or y, or both, can be well approximated by vectors in the subspace M.

III. ERROR BOUNDS FOR UPLINK-DOWNLINK CONVERSION IN FDD MIMO SYSTEMS

In this section, we apply the results of Sect. II to the problem of covariance matrix conversion in FDD massive MIMO systems, which, as mentioned in the introduction, we call the uplink-downlink conversion problem. We first describe the problem in Sect. III-A, and then we proceed to tailor the bounds in Proposition 1 to our particular application in Sect. III-B. In Sect. III-C, we show how the bounds can be used to improve the approaches in [3], [4]. To keep the discussion as general as possible, we do not assume any particular array geometry or propagation model in this section.

A. The uplink-downlink conversion problem

We consider a single-cell flat-fading wireless system, in which a base station equipped with N antennas exchanges data with a single-antenna user. In the uplink, the base station first estimates the uplink channel covariance matrix $\boldsymbol{R}_{\mathrm{u}} = E[\boldsymbol{h}_{\mathrm{u}} \boldsymbol{h}_{\mathrm{u}}^H] \in \mathbb{C}^{N \times N}$ from samples of the uplink channel $\boldsymbol{h}_{\mathrm{u}}$ and any prior knowledge of this covariance matrix. In the uplink-downlink conversion problem, samples of the downlink channel $\boldsymbol{h}_{\mathrm{d}}$ are not available at the base station, and the objective is to obtain an estimate of the downlink channel covariance matrix $\boldsymbol{R}_{\mathrm{d}} = E[\boldsymbol{h}_{\mathrm{d}} \boldsymbol{h}_{\mathrm{d}}^H] \in \mathbb{C}^{N \times N}$ directly from the estimate of $\boldsymbol{R}_{\mathrm{u}}$. The main challenge for the conversion in FDD MIMO systems is the lack of channel reciprocity. The uplink and downlink channels use different frequencies, so their statistics are also different, which in turn implies that $\boldsymbol{R}_{\mathrm{u}} \neq \boldsymbol{R}_{\mathrm{d}}$. However, $\boldsymbol{R}_{\mathrm{d}}$ and $\boldsymbol{R}_{\mathrm{u}}$ are related. In particular, estimating $\boldsymbol{R}_{\mathrm{d}}$ from $\boldsymbol{R}_{\mathrm{u}}$ is possible by using the so-called reciprocity of the angular power spectrum, which we now formally describe.

For typical frequency separation gaps, the real and imaginary parts of each component of $\boldsymbol{R}_{\mathrm{u}}$ and $\boldsymbol{R}_{\mathrm{d}}$ can be seen as the result of an inner product in an infinite dimensional real Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, with the vectors (functions) and the inner product taking a particular form that depends on system parameters such as the antenna polarization, array geometry, and the propagation model, among others [3], [4]. More precisely, let $r_{\mathrm{u},k} \in \mathbb{R}$ and $r_{\mathrm{d},k} \in \mathbb{R}$ denote one of the $k \in \mathcal{I} := \{1, \ldots, 2N^2\}$ components of, respectively, $[\mathrm{Real}(\boldsymbol{R}_{\mathrm{u}}) \ \mathrm{Imag}(\boldsymbol{R}_{\mathrm{u}})] \in \mathbb{R}^{N \times 2N}$ and $[\mathrm{Real}(\boldsymbol{R}_{\mathrm{d}}) \ \mathrm{Imag}(\boldsymbol{R}_{\mathrm{d}})] \in \mathbb{R}^{N \times 2N}$. It has been shown in [3], [4] that, for each $k \in \mathcal{I}$ and for a given inner product $\langle \cdot, \cdot \rangle$ that depends on the system model, we have

$$r_{\mathrm{u},k} = \langle \rho, g_{\mathrm{u},k} \rangle \qquad (4)$$

and

$$r_{\mathrm{d},k} = \langle \rho, g_{\mathrm{d},k} \rangle, \qquad (5)$$

where $\rho \in \mathcal{H}$ is the unknown frequency independent function called *angular power spectrum*, and $g_{\mathrm{u},k} \in \mathcal{H}$ and $g_{\mathrm{d},k} \in \mathcal{H}$ are known uplink and downlink functions related to the antenna array responses (see Sect. IV for a concrete example). Intuitively, the angular power spectrum is a function that shows the average angular power density that an array receives from the user at a given azimuth (and possibly elevation) angle. In the literature [3]–[6], it is assumed to be the same for both the uplink and downlink channels, which is the phenomenon we call reciprocity of the angular power spectrum.

With the above explanations, we can summarize the set-theoretic approaches in [3], [4] (and some of the approaches in [5]) to the uplink-downlink conversion problem with the following two steps:

(i) We first obtain an estimate $\widetilde{\rho} \in \mathcal{H}$ of $\rho \in \mathcal{H}$ from the equations $(\forall k \in \mathcal{I}) \ r_{\mathrm{u},k} = \langle \rho, g_{\mathrm{u},k} \rangle$ and possibly known properties of $\rho \in \mathcal{H}$ by using set-theoretic methods.
(ii) With $\widetilde{\rho} \in \mathcal{H}$, we obtain an estimate $\widetilde{r}_{\mathrm{d},k} \in \mathbb{R}$ of $r_{\mathrm{d},k} \in \mathbb{R}$ for each $k \in \mathcal{I}$ by computing $\widetilde{r}_{\mathrm{d},k} = \langle \widetilde{\rho}, g_{\mathrm{d},k} \rangle$.

One of the contributions of the next section is to derive conditions guaranteeing that, given $k \in \mathcal{I}$, the estimate $\widetilde{r}_{\mathrm{d},k}$ of $r_{\mathrm{d},k}$ in step (ii) is accurate even if the estimate $\widetilde{\rho}$ of ρ in step (i) is inaccurate. These conditions will be used to derive simple techniques to improve set-theoretic methods addressing the problem in step (i).

B. Bounds for the error of UL-DL covariance conversion with general arrays

We now derive performance bounds for the set-theoretic approaches described above. To this end, we assume that the uplink covariance matrix $\boldsymbol{R}_{\mathrm{u}}$ (and hence $(r_{\mathrm{u},k})_{k \in \mathcal{I}}$) is perfectly estimated. By recalling that covariance matrices have structure (they are at least Hermitian), the number of different equations in (4) and (5) is strictly less than $|\mathcal{I}| = 2N^2$ ($|\mathcal{I}|$ denotes the cardinality of $\mathcal{I}$). Therefore, many repeating equations can be removed, but for brevity this simple operation is not considered in this section.

To proceed with the bounds, we define

$$S' = \{g_{\mathrm{u},1}, \ldots, g_{\mathrm{u},|\mathcal{I}|}\} \subset \mathcal{H} \qquad (6)$$

to be the set corresponding to the uplink functions in (4). The angular power spectrum $\rho \in \mathcal{H}$ is related to S' by the fact that, from (4) and Fact. 1(ii)-(iii), we have $\rho \in V' := \cap_{k \in \mathcal{I}} \{x \in \mathcal{H} \mid \langle x, g_{\mathrm{u},k} \rangle = r_{\mathrm{u},k}\} = \rho + \mathrm{span}(S')^{\perp}$. Intuitively, the linear variety $V' \subset \mathcal{H}$ is the set containing all angular power spectrum functions that produce the same uplink covariance matrix.

To include in the analysis any prior information about $\rho \in \mathcal{H}$ expressed in terms of closed linear varieties or closed subspaces, we assume that

$$\rho \in V'' := \cap_{k=1}^{Q} \{x \in \mathcal{H} \mid \langle x, v_k \rangle = b_k\} \qquad (7)$$

and that the tuples $\{(v_k, b_k)\}_{k=1,\ldots,Q} \subset \mathcal{H} \times \mathbb{R}$ used to construct the linear variety V'' (V'' is a subspace if $b_1 = \ldots = b_Q = 0$) are known. We now define $V := V' \cap V''$ and construct a new set $S \subset \mathcal{H}$ containing all vectors in S' in (6) and all vectors in $S'' := \{v_1, \ldots, v_Q\}$; i.e.,

$$S := S' \cup S'' \subset \mathcal{H}. \qquad (8)$$

By recalling that a nonempty intersection of closed linear varieties is a closed linear variety, the set $V = V' \cap V'' \ni \rho$ defined above is a closed linear variety V that can be equivalently written as $V = \rho + M^\perp$, where $M \subset \mathcal{H}$ is the closed subspace $M := \mathrm{span}(S)$. With these operations, we are now exactly in the setting of Proposition 1. Before we proceed with the specialization of this proposition to the problem of uplink-downlink conversion, we first show that the projections $P_M : \mathcal{H} \to M$ and $P_V(0) \in \mathcal{H}$ are easy to compute.

To simplify the notation, denote by $x_1, \ldots, x_L$ the $L = |\mathcal{I}| + Q$ vectors in the set S in (8). Without any loss of generality, we assume that $x_k = g_{\mathrm{u},k}$ for $k \in \{1, \ldots, |\mathcal{I}|\}$ and $x_k = v_{k-|\mathcal{I}|}$ for $k \in \{|\mathcal{I}| + 1, \ldots, |\mathcal{I}| + Q\}$. Define the following matrix:

$$G = \begin{bmatrix} \langle x_1, x_1 \rangle & \cdots & \langle x_1, x_L \rangle \\ \vdots & \ddots & \vdots \\ \langle x_L, x_1 \rangle & \cdots & \langle x_L, x_L \rangle \end{bmatrix} \in \mathbb{R}^{L \times L}. \quad (9)$$

With the above definitions, we can use arguments similar to those in [9, Ch. 3][1] to show that the projection from $y \in \mathcal{H}$ onto the closed subspace M is given by:

$$P_M : \mathcal{H} \to M : y \mapsto \sum_{k=1}^{L} \alpha_k x_k, \quad (10)$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_L]^T \in \mathbb{R}^L$ is any solution to $\boldsymbol{G}\boldsymbol{\alpha} = \boldsymbol{z}$, and $\boldsymbol{z} = [\langle x_1, y \rangle \ldots \langle x_L, y \rangle]^T \in \mathbb{R}^L$. In turn, the projection from 0 onto the linear variety V described above is given by

$$P_V(0) = \sum_{k=1}^{L} \beta_k x_k, \quad (11)$$

where $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_L]^T \in \mathbb{R}^L$ is any solution to $\boldsymbol{G}\boldsymbol{\beta} = [r_{\mathrm{u},1}, \ldots, r_{\mathrm{u},|\mathcal{I}|}, b_1, \ldots, b_Q]^T \in \mathbb{R}^L$.

As it will soon become clear, in the proposed performance bounds we are specially interested in the estimation error $\|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$ for each $k \in \mathcal{I}$, which can be easily computed as shown in the following standard result. We omit the proof for brevity, but it can be easily obtained by using (10).

Proposition 2. *Let* $P_M(y) \in M = \mathrm{span}(\{x_1, \ldots, x_L\})$ *be the approximation in the subspace M of an arbitrary vector $y \in \mathcal{H}$. Then the approximation error $\|y - P_M(y)\|$ is given by*

$$\|y - P_M(y)\| = \sqrt{(\|y\|^2 - \boldsymbol{z}^T \boldsymbol{G}^\dagger \boldsymbol{z})},$$

where $\boldsymbol{z} = \boldsymbol{G}\boldsymbol{\alpha} \in \mathbb{R}^L$, $\boldsymbol{\alpha} \in \mathbb{R}^L$, *and* $\boldsymbol{G} \in \mathbb{R}^{L \times L}$ *are as defined above.*

Now, let

$$\boldsymbol{Q} := \begin{bmatrix} \langle x_1, g_{\mathrm{d},1} \rangle & \cdots & \langle x_1, g_{\mathrm{d},|\mathcal{I}|} \rangle \\ \vdots & \ddots & \vdots \\ \langle x_L, g_{\mathrm{d},1} \rangle & \cdots & \langle x_L, g_{\mathrm{d},|\mathcal{I}|} \rangle \end{bmatrix} \in \mathbb{R}^{L \times |\mathcal{I}|}. \quad (12)$$

The proposed error bounds for uplink-downlink conversion are shown in the next corollary.

Corollary 1. *Denote by* $\boldsymbol{q}_k \in \mathbb{R}^L$ *the kth column of the matrix $\boldsymbol{Q}$ in (12), and let $\boldsymbol{G}$ be as defined in (9). Suppose*

[1]Here we do not assume S to be a linearly independent set.

that $\widetilde{\rho} \in V$ *is an estimate of the angular power spectrum* $\rho \in V$ *obtained by a given algorithm, where* $V = \rho + M^\perp$ *and* $M \subset \mathcal{H}$ *is the closed subspace* $M = \mathrm{span}(S)$ *with S as defined in (8). Further, assume that* $\|\rho\| \leq B$ *for some* $B \in \mathbb{R}$*. Let* $\langle \widetilde{\rho}, g_{\mathrm{d},k} \rangle = \widetilde{r}_{\mathrm{d},k}$ *be the estimate of the kth* $(k \in \mathcal{I})$ *component* $r_{\mathrm{d},k}$ *of the downlink covariance matrix* $\boldsymbol{R}_{\mathrm{d}}$*. Then the estimation error* $e_k := |\widetilde{r}_{\mathrm{d},k} - r_{\mathrm{d},k}|$ *for each* $k \in \mathcal{I}$ *satisfies the following:*

(i) *If the algorithm used to produce the estimate* $\widetilde{\rho} \in V$ *also guarantees* $\|\widetilde{\rho}\| \leq B$*, then*

$$(\forall k \in \mathcal{I}) \; e_k \leq \|\rho - \widetilde{\rho}\| \; \|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$$
$$\leq 2B\sqrt{(\|g_{\mathrm{d},k}\|^2 - \boldsymbol{q}_k^T \boldsymbol{G}^\dagger \boldsymbol{q}_k)}.$$

(ii) *Using* $\widetilde{\rho} = P_V(0)$ *as the estimate of the angular power spectrum ρ, we have* $(\forall k \in \mathcal{I})$

$$e_k \overset{(a)}{\leq} \|\rho - P_M(\rho)\| \; \|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$$
$$\overset{(b)}{\leq} \|\rho\| \; \|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$$
$$\overset{(c)}{\leq} B\sqrt{(\|g_{\mathrm{d},k}\|^2 - \boldsymbol{q}_k^T \boldsymbol{G}^\dagger \boldsymbol{q}_k)}. \quad (13)$$

Proof. The proof of (i) is immediate from Proposition 1(i), Proposition 2, and the triangle inequality. To prove (ii), we note that the inequality in (a) follows from Proposition 1(ii), the inequality in (b) follows from $\|\rho - P_M(\rho)\| = \|P_{M^\perp}(\rho)\| \leq \|P_{M^\perp}\|_{\mathrm{o}} \|\rho\| \leq \|\rho\|$ [see Fact 1(i)], and the inequality in (c) follows from Proposition 2 and the assumption $\|\rho\| \leq B$. $\square$

C. Improving the performance of the conversion with information about the support of the angular power spectrum

One of the practical implications of Corollary 1 is that, for a given $k \in \mathcal{I}$, any algorithm producing an estimate $\widetilde{\rho} \in V$ of $\rho \in V$ is able to approximate reliably the components $r_{\mathrm{d},k}$ of the downlink covariance matrix $\boldsymbol{R}_{\mathrm{d},k}$ provided that the term $\|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$ is sufficiently small, regardless of how challenging the scenario for the estimation of ρ may be. By recalling that the projection $P_M(g_{\mathrm{d},k}) \in \mathcal{H}$ can be interpreted as the best approximation of $g_{\mathrm{d},k}$ in the closed subspace M, adding to the subspace M functions as similar as possible to $g_{\mathrm{d},k}$ is a natural idea to decrease the estimation error bound $2B\|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$ of $r_{\mathrm{d},k}$. In the discussion below, we show a simple technique to design M based on this simple principle.

The subspace M is by definition the span of $S = S' \cup S''$, where $S' = \{g_{\mathrm{u},1}, \ldots, g_{\mathrm{u},|\mathcal{I}|}\}$ is the set of uplink functions and $S'' = \{v_1, \ldots, v_Q\}$ is the set of functions resulting from any prior knowledge about ρ (see (7)). Therefore, a simple means of including in the subspace M functions that are close to each of the downlink functions $(g_{\mathrm{d},k})_{k \in \mathcal{I}}$ is to make the uplink functions $(g_{\mathrm{u},k})_{k \in \mathcal{I}}$ as similar as possible to the downlink functions $(g_{\mathrm{d},k})_{k \in \mathcal{I}}$. Alternatively, we can also include in the set S'' functions that are as similar as possible to the downlink functions $(g_{\mathrm{d},k})_{k \in \mathcal{I}}$ (NOTE: including $(g_{\mathrm{d},k})_{k \in \mathcal{I}}$ directly while guaranteeing $\rho \in V$ is difficult). The first approach, which corresponds to the design of uplink and downlink functions, may not be always possible because it typically entails changes in hardware (e.g., changes in the inter-antenna spacing) or

other modifications in standardized system parameters (e.g., operating frequencies). Therefore, here we focus on the second approach; namely, the construction of an appropriate set S'', or, equivalently, the corresponding linear variety or subspace V'' in (7). To derive the sets, we further assume the following:

(A1) The angular power spectrum $\rho \in \mathcal{H}$, the downlink functions $(g_{\mathrm{d},k})_{k \in \mathcal{I}} \subset \mathcal{H}$, and the uplink functions $(g_{\mathrm{u},k})_{k \in \mathcal{I}} \subset \mathcal{H}$ are functions in a Hilbert space of functions in $L^2(\Omega)$, or, as in [4], a Hilbert space $\mathcal{H}$ of tuples in $L^2(\Omega) \times L^2(\Omega)$ (NOTE: extensions to different Hilbert spaces is straightforward), [2] where $\Omega \subset \mathbb{R}^K$.

(A2) There exists a known non-null measurable set $C_{\mathrm{S}} \subset \Omega$ such that $\mathrm{Supp}(\rho) \subset C_{\mathrm{S}} \neq \emptyset$ for the angular power spectrum functions $\rho \in \mathcal{H}$ that can be observed in the system. Intuitively, the set C_{S} is a superset of $\mathrm{Supp}(\rho)$ for which $\theta \notin C_{\mathrm{S}}$ implies $\rho(\theta) = 0$.

Assumption A1 is very natural. It is satisfied in many realistic models representing the angular power spectrum in practical systems. In these models, Ω is the set of possible azimuth and elevation angles [3], [4]. Assumption A2 is system dependent, but it may be valid in scenarios where signals of users impinging on the antenna array are not likely to have any significant power at certain angles, which are used for the construction of C_{S}.

We now proceed to show how support information of ρ can be used to design the subspace M by using arguments that have a strong theoretical justification. To this end, consider the closed subspace

$$\mathcal{K} := \overline{\{x \in \mathcal{H} \mid (\forall \theta \in C_{\mathrm{S}}) \ x(\theta) = 0\}}.$$

The projection $P_{\mathcal{K}} : \mathcal{H} \to \mathcal{K}$ from $v \in \mathcal{H}$ onto $\mathcal{K}$ is the function given by (we omit the proof for brevity):

$$\mathcal{H} \ni P_{\mathcal{K}}(v) : \Omega \to \mathbb{R} : \theta \mapsto \begin{cases} 0, & \text{if } \theta \in C_{\mathrm{S}}, \\ v(\theta) & \text{otherwise.} \end{cases}$$

Since $P_{\mathcal{K}}(\rho) = 0$ from the assumption $\mathrm{Supp}(\rho) \subset C_{\mathrm{S}}$ and the definition of the subspace $\mathcal{K}$, we have $\rho \in \mathcal{K}^{\perp}$, and thus

$$(\forall v \in \mathcal{H}) \ \langle P_{\mathcal{K}}(v), \rho \rangle = 0. \tag{14}$$

In particular, using the downlink functions as the function v in (14) yields

$$(\forall k \in \mathcal{I}) \ \langle P_{\mathcal{K}}(g_{\mathrm{d},k}), \rho \rangle = 0. \tag{15}$$

We have now reached the point to show the closed subspace V'' we propose to represent the prior knowledge about the support of ρ. More precisely, in light of (15), we use

$$V'' = \cap_{k \in \mathcal{I}} \{x \in \mathcal{H} \mid \langle x, v_k \rangle = 0\} \ni \rho,$$

where $v_k := P_{\mathcal{K}}(g_{\mathrm{d},k})$ for each $k \in \mathcal{I}$. This choice is intuitively appealing because we add to the set S in (8) all vectors $(v_k)_{k \in \mathcal{I}}$ in $\mathcal{K}$ that best approximate (with respect to the metric $d(x,y) = \|x - y\|$) the downlink functions $(g_{\mathrm{d},k})_{k \in \mathcal{I}}$, and we recall from the above discussion that, for each $k \in \mathcal{I}$, the estimation error $|\widetilde{r}_{\mathrm{d},k} - r_{\mathrm{d},k}|$ decreases as

the ability to represent $g_{\mathrm{d},k}$ with functions in $M = \mathrm{span}(S)$ improves. Note that we could further improve the reliability of the conversion by repeating the above procedure to include in S additional functions of the form $P_{\mathcal{K}}(v)$ with $v \in \mathcal{H}$ [e.g., the functions $(P_{\mathcal{K}}(g_{\mathrm{u},k}))_{k \in \mathcal{I}}$]. Alternatively, we could also change the definition of inner products to consider only functions in $L^2(C_{\mathrm{S}})$. These approaches can be numerically unstable in large antenna arrays if the information about the support of ρ is erroneous and appropriate mitigation techniques are not applied, but we leave this discussion to a future study because of the space limitation.

All the above improvements are available for the simple approach using $\widetilde{\rho} = P_V(0)$ as the estimate of ρ. This approach is particularly interesting because, as shown in the study in [3], which has not considered the enhancements discussed above, the whole process of estimating the angular power spectrum and using this estimate to reconstruct the downlink covariance matrix can be done with a simple matrix-vector multiplication. This important feature is not lost with the enhancements proposed in this subsection. More precisely, denote by $\widetilde{R}_{\mathrm{d}}$ the estimate of the downlink covariance matrix R_{d}, and recall that $\widetilde{r}_{\mathrm{d},1} = \langle P_V(0), g_{\mathrm{d},1} \rangle, \ldots, \widetilde{r}_{\mathrm{d},|\mathcal{I}|} = \langle P_V(0), g_{\mathrm{d},|\mathcal{I}|} \rangle$ represent the real and imaginary parts of the components of $\widetilde{R}_{\mathrm{d}}$. Using (11) in these inner products, we verify that uplink-downlink channel covariance conversion can be performed with the following simple linear operation:

$$\mathrm{vec}[\mathrm{Real}(\widetilde{R}_{\mathrm{d}}) \ \mathrm{Imag}(\widetilde{R}_{\mathrm{d}})] = Q^T G^{\dagger} \underline{r} = Ar, \tag{16}$$

where $\underline{r} = [r^T, 0, \ldots, 0]^T \in \mathbb{R}^L$, $r := [r_{\mathrm{u},1}, \ldots, r_{\mathrm{u},|\mathcal{I}|}]^T \in \mathbb{R}^{|\mathcal{I}|}$, and $A \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{I}|}$ is the matrix obtained by keeping only the first $|\mathcal{I}|$ columns of the matrix $Q^T G^{\dagger}$. Note that the matrix A only depends on the support information, which is often assumed to be slowly time-varying information, and the array response, so this matrix is computed sporadically. If no support information is used, then A needs to be computed only once.

Before we finish this section, it is also worth noticing that, by increasing the subspace M with support information about ρ as described above, we also decrease the algorithm error term $\|\rho - P_M(\rho)\|$ in the bound (a) in (13), thus further improving the reliability of the conversion.

IV. EXAMPLE: UNIFORM LINEAR ARRAYS

We now further specialize the results in the previous section to uniform linear arrays. This particular choice enables us to relate the analysis in the previous sections to existing results in the literature that, unlike our approaches, do not seem easy to extend to schemes exploiting information about the structure of the angular power spectrum or to systems where the functions $(g_{\mathrm{u},k})_{k \in \mathcal{I}}$ and $(g_{\mathrm{d},k})_{k \in \mathcal{I}}$ are determined by measurements instead of models.[3]

---

[2] In these Hilbert spaces, which are used in Sect. IV, we typically work with classes of equivalent functions, with the equivalence relation between two functions f and g defined by $f \sim g \Leftrightarrow \|f - g\| = 0$. Equalities such as $f = g$ should be understood as equalities between the classes, not to the particular functions (in a pointwise sense) because f and g can differ, for example, in a countable set in their domains.

[3] If we normalize the angular power spectrum to satisfy $\|\rho\| = 1$, we can simply set the constant B in Corollary 1 to $B = 1$. In contrast, in addition to a similar normalization, the bound in [5, Theorem 1] requires knowledge of a constant that does not seem easy to determine. Without this constant, comparing directly the proposed bounds with the bound in [5, Theorem 1] seems difficult.

A. System model and bounds without support information

In a uniform linear array with N antennas, under very mild assumptions [6], [10], the uplink and downlink channel covariance matrices for typical frequency gaps are given by $\boldsymbol{R}_{\mathrm{u}} := \boldsymbol{R}(f_{\mathrm{u}}) \in \mathbb{C}^{N \times N}$ and $\boldsymbol{R}_{\mathrm{d}} := \boldsymbol{R}(f_{\mathrm{d}}) \in \mathbb{C}^{N \times N}$, where $f_{\mathrm{u}} \in \mathbb{R}_+$ and $f_{\mathrm{d}} \in \mathbb{R}_+$ are, respectively, the uplink and downlink frequencies;

$$\boldsymbol{R}(f) = \int_{-\pi/2}^{\pi/2} \rho(\theta) \boldsymbol{a}(\theta, f) \boldsymbol{a}(\theta, f)^H \mathrm{d}\theta$$

(the integral should be understood coordinate-wise) is the channel covariance matrix for a given frequency f; $\rho : [-\pi/2, \pi/2] \to \mathbb{R}_+$ is the angular power spectrum;

$$\begin{aligned}
\boldsymbol{a} : [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \times \mathbb{R}_+ \to \quad & \mathbb{C}^N \\
(\theta, f) \mapsto \quad & \left[1, e^{i2\pi \frac{f}{c} d \sin\theta}, \ldots, e^{i2\pi \frac{f}{c} d(N-1)\sin\theta}\right]
\end{aligned}$$
(17)

is the array response for a given angle θ and frequency f; c is the speed of the wave propagation; and d is the inter-antenna spacing.

In real physical systems, we can safely assume that ρ is an element of the Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ of Lebesgue (real) square-integrable functions $\mathcal{H} = L^2([-\pi/2, \pi/2])$ equipped with the inner product $(\forall \rho \in \mathcal{H})(\forall g \in \mathcal{H}) \langle \rho, g \rangle = \int_{-\pi/2}^{\pi/2} \rho(\theta) g(\theta) \mathrm{d}\theta$. As a result, by fixing f_{u}, in light of (17) the functions $(g_{\mathrm{u},k})_{k \in \{1,\ldots,2N^2\}}$ in (4) are obtained from the equality $(\forall \theta \in [-\pi/2, \pi/2])$

$$\begin{aligned}
&[g_{\mathrm{u},1}(\theta), \ldots, g_{\mathrm{u},2N^2}(\theta)]^T = \\
&\qquad \mathrm{vec}\left(\begin{bmatrix} \mathrm{Real}(\boldsymbol{a}(\theta, f_{\mathrm{u}})\boldsymbol{a}(\theta, f_{\mathrm{u}})^H) \\ \mathrm{Imag}(\boldsymbol{a}(\theta, f_{\mathrm{u}})\boldsymbol{a}(\theta, f_{\mathrm{u}})^H) \end{bmatrix}\right).
\end{aligned}$$
(18)

The downlink functions $(g_{\mathrm{d},k})_{k \in \{1,\ldots,2N^2\}}$ in (5) are obtained analogously by considering the downlink frequency f_{d} in (18).

In uniform linear arrays, the covariance matrices are Hermitian and Toeplitz [3], [5], [10], so, with the ordering in (18), we can consider only the functions $g_{\mathrm{u},1}, \ldots, g_{\mathrm{u},2N}$ responsible for the first column of $\boldsymbol{R}_{\mathrm{u}}$ because knowledge of this column is enough to reconstruct all elements of $\boldsymbol{R}_{\mathrm{u}}$. For the same reason, we use only the downlink functions $g_{\mathrm{d},1}, \ldots, g_{\mathrm{d},2N}$. By doing so, the set S' in (6) is given by $S' = \{g_{\mathrm{u},1}, \ldots, g_{\mathrm{u},2N}\} \subset \mathcal{H}$, and we can redefine the index set $\mathcal{I}$ accordingly; i.e., $\mathcal{I} := \{1, \ldots, 2N\}$.

Without any information about the support of ρ, we have $S = S'$, and we can use the results in [3, Sect. 4.1] to compute the algorithm independent term $\|g_{\mathrm{d},k} - P_M(g_{\mathrm{d},k})\|$ ($k \in \mathcal{I}$) of the bounds in Corollary 1 by using Bessel functions of the first kind, order zero, which we denote by $J_0 : \mathbb{R} \to \mathbb{R}$. In this case, the bound in the last inequality in Corollary 1(ii) reduces to

$$(\forall k \in \mathcal{I}) \; e_k \le B\sqrt{(\|g_{\mathrm{d},k}\|^2 - \boldsymbol{q}_k^T \boldsymbol{G}^\dagger \boldsymbol{q}_k)},$$
(19)

where

$$\|g_{\mathrm{d},k}\|^2 = \begin{cases} \frac{\pi}{2}\left(1 + J_0\left(4\pi \frac{f_{\mathrm{d}}}{c} d(k-1)\right)\right) & \text{if } 1 \le k \le N \\ \frac{\pi}{2}\left(1 - J_0\left(4\pi \frac{f_{\mathrm{d}}}{c} d(k-N-1)\right)\right) & \text{otherwise,} \end{cases}$$

TABLE I: Parameters of the uniform linear array

Number of antennas (N)	30
Uplink frequency (f_{u})	1.8 MHz
Downlink frequency (f_{d})	1.9 MHz
Speed of wave propagation (c)	$3 \cdot 10^8$ m/s
Antenna spacing (d)	$1.05 \frac{c}{2f_{\mathrm{u}}}$

$$\boldsymbol{G} = \frac{\pi}{2}\begin{bmatrix} \boldsymbol{G}_{\mathrm{r}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_{\mathrm{j}} \end{bmatrix}, \boldsymbol{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{2N}] = \frac{\pi}{2}\begin{bmatrix} \boldsymbol{Q}_{\mathrm{r}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_{\mathrm{j}}, \end{bmatrix}$$

and the components of the nth row and mth column of the matrices $\boldsymbol{G}_{\mathrm{r}}, \boldsymbol{G}_{\mathrm{j}}, \boldsymbol{Q}_{\mathrm{r}}, \boldsymbol{Q}_{\mathrm{j}} \in \mathbb{R}^{N \times N}$ are given by $\boldsymbol{G}_{\mathrm{r},nm} = J_0(x_{nm}) + J_0(y_{nm})$, $\boldsymbol{G}_{\mathrm{j},nm} = J_0(x_{nm}) - J_0(y_{nm})$, $\boldsymbol{Q}_{\mathrm{r},nm} = J_0(p_{nm}) + J_0(q_{nm})$, $\boldsymbol{Q}_{\mathrm{j},nm} = J_0(p_{nm}) - J_0(q_{nm})$ with

$$x_{nm} = 2\pi \, d \, \frac{f_{\mathrm{u}}}{c}(n-m), \; y_{nm} = 2\pi \, d \, \frac{f_{\mathrm{u}}}{c}(n+m-2),$$

$$p_{nm} = 2\pi d\left(\frac{f_{\mathrm{u}}(n-1)}{c} - \frac{f_{\mathrm{d}}(m-1)}{c}\right),$$

and

$$q_{nm} = 2\pi d\left(\frac{f_{\mathrm{u}}(n-1)}{c} + \frac{f_{\mathrm{d}}(m-1)}{c}\right).$$

B. Numerical experiments

For a concrete example of the bounds, we use an antenna array with the configuration in Table I. The number of antennas in the array is relatively small to emphasize the fact that the proposed bounds do not appeal to asymptotic results. As discussed in [5], the configuration in Table I is challenging for uplink-downlink conversion for two main reasons: (i) the uplink frequency is lower than the downlink frequency, and (ii) the antenna spacing is larger than half of the wavelength $c/(2f_{\mathrm{d}})$ of the higher frequency f_{d}, so we have the undesirable phenomenon known as grating lobes [5], [11]. We show below that this challenging scenario for uplink-downlink conversion can be formally verified with the simple bounds in (19), and the problems for uplink-downlink conversion can be mitigated with information about the support of the angular power spectrum.

To illustrate the theoretical gains that can be achieved with the technique discussed in Sect. III-C, we assume that $\mathrm{Supp}(\rho) \subset C_{\mathrm{S}} = [0, \pi/2]$, and C_{S} is known. For all simulations in this section, we use only the scheme in Corollary 1(ii) for the estimation of ρ because of its low computational complexity, as discussed in Sect. III-C.

In Fig. 1, assuming $\|\rho\| = B = 1$, we show the bounds in the last inequality in (13) with and without support information (SI). For the computation of the former bound, we use the expressions in (19). For the latter bound, we construct the matrices $\boldsymbol{G}$ and $\boldsymbol{Q}$ in (9) and (12) by computing integrals numerically, unless the integral falls into one of the cases computed in (19). From Fig. 1, it is clear that, without any support information, the estimate $\widetilde{r}_{\mathrm{d},k}$ can be unreliable for many indices $k \in \mathcal{I}$, which is also in accordance with the results in [5]. In contrast, with support information, all estimates $(\widetilde{r}_{\mathrm{d},k})_{k \in \mathcal{I}}$ of the components of the downlink covariance matrix are reliable, even if the estimate $\widetilde{\rho} = P_V(0)$ of the angular power spectrum ρ is not necessarily accurate.

To illustrate the above fact, consider the following example for $\rho : [-\pi/2, \pi/2] \to \mathbb{R}_+$:
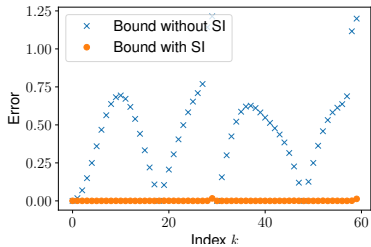
Fig. 1: Upper bound (13) on the error $(e_k)_{k \in \mathcal{I}}$ for the conversion performed with the algorithm in (16) with and without support information. $B = 1$.
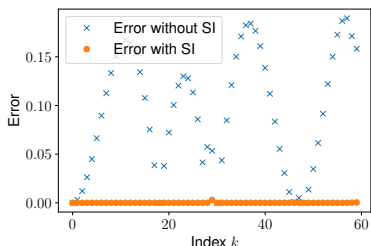

Fig. 2: Simulated conversion error $(e_k)_{k \in \mathcal{I}}$ of the algorithm in (16) with and without support information. Angular power spectrum in (20).

$$\rho(\theta) = n e^{-|\theta - .5|/.05} + 4n \ e^{-|\theta - 1.4|/.05}, \qquad (20)$$

where $n \in \mathbb{R}_+$ is a normalizing constant chosen to guarantee that $\|\rho\| = 1$. This exemplary ρ can be interpreted as coming from a user with two multipath components at angles 0.5 rad and 1.4 rad. Note that we have violated the assumption of the support of ρ, but the signal energy outside C_S is small compared to the energy in C_S, so we can expect the bounds shown in Fig. 1 to be accurate. This fact is illustrated in Fig. 2, which shows the absolute error $(e_k = |r_{d,k} - \widetilde{r}_{d,k}|)_{k \in \mathcal{I}}$ [with $\tilde{r}_{u,k}$ computed by using (16)] of the estimates. Note that, by including support information, uplink-downlink conversion has been performed reliably for all components of the downlink covariance matrix, even though the estimate of angular power spectrum (APS) is not necessarily accurate, as depicted in Fig. 3. We verify, for example, that the schemes used to estimate the function in (20) produce functions taking negative values for some angles.

V. SUMMARY AND CONCLUSIONS

Recent work has proved that, without side information about the angular power spectrum, existing algorithms in the literature may not be able to estimate reliably all components of the downlink covariance matrix. In this study we have introduced alternative reliability bounds that are based on elementary arguments in infinite dimensional Hilbert spaces. The main advantages of the proposed analysis are the simplicity and the generality. Unlike previous results, the bounds shown here can be straightforwardly used to analyze the performance of algorithms that exploit information about
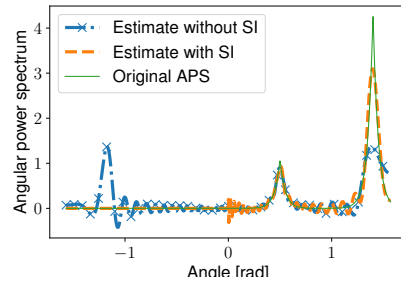

Fig. 3: Estimate $\widetilde{\rho} = P_V(0)$ [see (11)] of the angular power spectrum (APS) ρ in (20) with and without support information for the construction of the linear variety V.

the support of the angular power spectrum in challenging scenarios that take into account the polarization of antennas and physical impairments of real antenna arrays. To illustrate a possible application of the bounds, we have improved a simple set-theoretic algorithm that does not require any parameter tuning. We have shown that, with coarse information about the angular power spectrum, all components of the downlink covariance matrix can be reliably estimated from the uplink covariance matrix with a simple linear operation. This result suggests that, in some scenarios, the main challenge may be the estimation of the uplink covariance matrix, not necessarily the uplink-downlink conversion problem.

REFERENCES

[1] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing–the large-scale array regime," *IEEE transactions on information theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[2] A. Decuninge, M. Guillaud, and D. T. Slock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *IEEE Globecom*. IEEE, 2015, pp. 1–6.

[3] L. Miretti, R. L. G. Cavalcante, and S. Stanczak, "FDD massive MIMO channel spatial covariance conversion using projection methods," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018.

[4] ——, "Downlink channel spatial covariance estimation in realistic FDD massive MIMO systems," 2018, (submitted) *arXiv preprint arXiv:1804.04892*.

[5] S. Haghighatshoar, M. B. Khalilsarai, and G. Caire, "Multi-band covariance interpolation with applications in massive MIMO," in *IEEE International Conference on Information Theory (ISIT)*, June 2018, pp. 386–390, extended version available as arXiv preprint arXiv:1801.03714.

[6] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, 2016.

[7] J. Dai, A. Liu, and V. K. N. Lau, "FDD massive MIMO channel estimation with arbitrary 2D-array geometry," *IEEE Trans. Signal Processing*, vol. 66, no. 10, pp. 2584–2599, May 2018.

[8] M. Yukawa, "Adaptive filtering based on projection method," Lecture Notes for Block Seminar at University of Erlangen Nuremberg and Technical University of Munich, Germany, 2010.

[9] D. G. Luenberger, *Optimization by vector space methods*. USA: Wiley, 1969.

[10] S. Haghighatshoar and G. Caire, "Massive MIMO channel subspace estimation from low-dimensional projections," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 303–318, 2017.

[11] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.