GRID-FREE DIRECTION-OF-ARRIVAL ESTIMATION WITH COMPRESSED SENSING AND ARBITRARY ANTENNA ARRAYS

Sebastian Semper[†], *Florian Roemer*[*], *Thomas Hotz*[†], *Giovanni Del Galdo*[*,+]

[†]Technische Universität Ilmenau, Institute for Mathematics, Ilmenau, Germany
[*]Technische Universität Ilmenau, Institute for Information Technology, Ilmenau, Germany
[+]Fraunhofer Institute for Integrated Circuits IIS, Ilmenau, Germany

ABSTRACT

We study the problem of direction of arrival estimation for arbitrary antenna arrays. We formulate it as a continuous line spectral estimation problem and solve it under a sparsity prior without any gridding assumptions. Moreover, we incorporate the array's beampattern in form of the Effective Aperture Distribution Function (EADF), which allows to use arbitrary (synthetic as well as measured) antenna arrays. This generalizes known atomic norm based grid-free DOA estimation methods (that have so far been limited to uniformly spaced arrays) to arbitrary antenna arrays. In addition, our formulation allows to incorporate compressed sensing in form of special linear combinations of the antennas' output ports. We provide conditions for the successful reconstruction of a certain number of targets depending on the amount of compression and the EADF of the antenna array. Our results are applicable to measurement matrices from any sub-Gaussian distribution.

***Index Terms*—** DOA estimation, Atomic Norm Minimization, Sparse Signal Recovery

1. INTRODUCTION

Direction of Arrival (DOA) estimation has been a field of active research for several decades [1] with a wide range of applications such as radar, sonar, communications, or channel sounding. Conventional techniques either exploit some algebraic structures of the underlying array manifolds or employ iterative solutions of the underlying non-convex maximum likelihood estimation problem [1].

Recently, connections between the DOA estimation problem and the emerging field of compressed sensing (CS) have been discovered [2]. In particular, since the observed signals are sparse in the angular domain, algorithms from the field of sparse signal recovery (SSR) can be applied for DOA estimation [3, 4]. Since the angle is a continuous parameter and its discretization introduces an unwanted model mismatch [5], grid-free SSR methods are of particular interest.

For the special case of uniform linear arrays (ULA) with isotropic antenna elements, the DOA estimation problem is equivalent to a line-spectral estimation problem and as such it can be cast into an atomic norm minimization (ANM) problem [6, 7]. However, this assumption is quite unrealistic in practice since isotropic antenna elements do not exist and ULAs are impractical for applications like direction finding and channel sounding where a uniform sensitivity of the array is desired.

This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG) project CoSMoS (grant GA 2062/2-1) and the Carl-Zeiss Foundation under the postdoctoral scholarship project "EMBiCoS".

The authors in [8] extend the grid-free approach from the ULA case to a randomly subsampled ULA arrays and demonstrate, that the grid-free sparse recovery approach to DOA generalizes to this setting. However, the array still has to fulfill the contraint, that it originates from an ULA and as such is not as flexible as the method proposed here.

This paper extends the prior work of grid-free sparsity-based DOA estimation to arbitrary antenna arrays. Since beampatterns of antenna arrays are periodic and rather smooth functions, they can typically be very well described by a truncated Fourier series [9]. This description is also known as the Effective Aperture Distribution Function (EADF). As we show, the EADF allows to rewrite the DOA estimation problem into a generalized line spectral estimation problem, for which an ANM-based approach is known [10]. Adopting [10] to the DOA setting allows us to develop grid-free sparsity-based DOA estimation algorithms that are applicable to arbitrary arrays. Moreover, we obtain performance guarantees that depend on the specific beam pattern of the array as well as the amount of compression applied to it. Our results are applicable to compression matrices drawn from an arbitrary sub-Gaussian distribution. In our numerical results, we demonstrate the performance in the noise-free as well as the noisy case for Gaussian as well as binary (± 1) compression matrices since the latter are very relevant in practice from a hardware realization point of view.

2. DATA MODEL

Let $\boldsymbol{a}(\theta) \,:\, [0, 2\pi) \to \mathbb{C}^M$ model the response of an array comprising of M antennas for a planar wave impinging from azimuth angle θ. Naturally, each element $a_m(\theta)$, $m = 1, 2, \dots, M$ is a periodic function in θ. Moreover, since beam patterns are typically quite smooth functions, they can be very well approximated by a truncated Fourier series [9] given by

$$a_m(\theta) \approx \sum_{\ell=-\frac{L-1}{2}}^{\frac{L-1}{2}} g_{m,\ell} \mathrm{e}^{\jmath\theta\ell}, \tag{1}$$

where we have considered an odd number of L terms. In matrix form, (1) can be written as

$$\boldsymbol{a}(\theta) \approx \boldsymbol{G} \cdot \boldsymbol{f}(\theta), \tag{2}$$

where $\boldsymbol{G} \in \mathbb{C}^{M\times L}$ contains the coefficients $g_{m,\ell}$ and $\mathrm{rk}(\boldsymbol{G}) = M$ and $\boldsymbol{f}(\theta) = [\mathrm{e}^{-\jmath\theta(L-1)/2}, \dots, \mathrm{e}^{\jmath\theta(L-1)/2}]^{\mathrm{T}}$ is the Fourier interpolation kernel. In practice $\boldsymbol{G}$ is obtained by measuring $\boldsymbol{a}(\theta)$ on a fine grid of angles, computing its Fourier coefficients, and truncating them at a reasonable, small threshold [9].

We assume that a superposition of S planar wavefronts from distinct directions θ_s is received by an antenna array with EADF $\boldsymbol{G} \in$

$\mathbb{C}^{M\times L}$ and a subsequent spatial compression step symbolized by a compression matrix $\mathbf{\Phi} \in \mathbb{C}^{K\times M}$ where $K < M$ and $\rho = K/M < 1$ is the compression factor. This means that $\mathbf{\Phi}$ achieves a reduction from M antenna ports to K channels that are actively downsampled and digitized. Such a reduction can be implemented with a network of power splitters, phase shifters, amplifiers, and combiners. Note that from the hardware complexity point of view it is attractive to limit $\mathbf{\Phi}$ to fewer degrees of freedom, e.g., adjusting only the phase or using only values of ± 1 (which only needs inverters that admit a broadband realization). For details on the design and the realization of a spatial compression, the reader is referred to [11]. The received signal can then be expressed as

$$\boldsymbol{z} = \mathbf{\Phi} \cdot \boldsymbol{G} \cdot \boldsymbol{x} + \boldsymbol{w} = \mathbf{\Gamma} \cdot \boldsymbol{x} + \boldsymbol{w}, \quad \text{where} \tag{3}$$

$$\boldsymbol{x} = \sum_{s=1}^{S} c_s \boldsymbol{f}(\theta_s), \tag{4}$$

the vector $\boldsymbol{w}$ represents the additive noise, and we have defined $\mathbf{\Gamma} = \mathbf{\Phi} \cdot \boldsymbol{G}$ for brevity.

3. RECONSTRUCTION

Our aim is now to recover $S \in \mathbb{N}$ as well as $c_s \in \mathbb{C}$ and $\theta_s \in [0, 2\pi)$ for $s = 1, 2, \ldots, S$ from the observed $\boldsymbol{z}$. Provided that $\boldsymbol{z}$ has a sparse representation shown in (4), this can be uniquely achieved given that S is small and a sufficient number of measurements is available.

To this end, we introduce the so called atomic norm $\|\cdot\|_{\mathcal{A}}$ for a given set $\mathcal{A} \subset \mathbb{C}^M$, which has to fulfill only very mild conditions in order for

$$\boldsymbol{x} \mapsto \|\boldsymbol{x}\|_{\mathcal{A}} = \inf \{ t > 0 \mid \boldsymbol{x} \in t \cdot \operatorname{conv}(\mathcal{A}) \} \tag{5}$$

to be a norm on $\mathbb{C}^M$, where $\operatorname{conv}(\mathcal{A})$ denoted the convex hull of the set $\mathcal{A}$. In our case, the sparse representation (4) is spanned by the Fourier vectors $\boldsymbol{f}(\theta)$ and therefore, we can think of the reconstruction as a line spectral estimation problem with an atomic set described as

$$\mathcal{A} = \left\{ \boldsymbol{f}(\theta) \in \mathbb{C}^M \;\middle|\; \theta \in [0, 2\pi) \right\}. \tag{6}$$

The problem we aim at solving is now

$$\min \|\boldsymbol{x}\|_{\mathcal{A}} \quad \text{subject to} \quad \|\boldsymbol{z} - \mathbf{\Phi}\boldsymbol{G}\boldsymbol{x}\|_2 \leq \epsilon. \tag{7}$$

where ϵ accounts for the additive noise. This type of optimization problem is called atomic norm minimization and it is a convex problem, which in the case of line spectral estimation can be reformulated as a semidefinite program (SDP) [6] and reads as

$$\begin{aligned} \min_{(\boldsymbol{x},\boldsymbol{u},t) \in \mathbb{C}^L \times \mathbb{C}^L \times \mathbb{R}} \quad & \frac{1}{2n} \operatorname{tr} \operatorname{Toep}(\boldsymbol{u}) + \frac{1}{2} t \\ \text{subject to} \quad & \begin{pmatrix} \operatorname{Toep}(\boldsymbol{u}) & \boldsymbol{x} \\ \boldsymbol{x}^{\mathrm{H}} & t \end{pmatrix} \succeq 0, \\ & \|\boldsymbol{z} - \mathbf{\Phi}\boldsymbol{G}\boldsymbol{x}\|_2 \leqslant \epsilon. \end{aligned} \tag{8}$$

Here, $\operatorname{Toep}(\boldsymbol{u})$ is a hermitian Toeplitz matrix with $\boldsymbol{u}$ as its first column, $\boldsymbol{A} \succeq 0$ denotes that the matrix $\boldsymbol{A}$ is positive semidefinite and $\operatorname{tr} \boldsymbol{A}$ is the trace of $\boldsymbol{A}$. Once the problem is solved, we can recover θ_s from $\operatorname{Toep}(\boldsymbol{u})$ by applying the Vandermonde decomposition via standard methods such as MUSIC or ESPRIT and the c_s via a least squares fit [6].

4. RECOVERY GUARANTEES

In this section we provide a sufficient condition such that the convex optimization problem shown in Section 3 has a unique solution. The derivation follows [10], adapting it to our setting where $\mathbf{\Gamma} = \mathbf{\Phi} \cdot \boldsymbol{G}$ with $\boldsymbol{G}$ being given by the array and $\mathbf{\Phi}$ representing the randomly chosen compression matrix. As in [10] we study the noise-free case only.

We begin with the following definition that describes a class of matrix-valued random distributions which have the property that an instance drawn from this ensemble yields successful recovery via (7) with high probability.

Definition 4.1 (sub-Gaussian matrices)**.** A matrix $\boldsymbol{A} \in \mathbb{C}^{K\times M}$ is b-sub-Gaussian with population covariance $\mathbf{\Sigma}$ if its rows are independent of each other and for all $k = 0, \ldots, K-1$ the k-th row $\boldsymbol{a}^k \in \mathbb{C}^{1\times M}$ of $\boldsymbol{A}$ satisfies

$$\mathbb{E}\{\boldsymbol{a}^k\} = 0, \quad \mathbb{E}\{\boldsymbol{a}^{k^{\mathrm{H}}} \boldsymbol{a}^k\} = \mathbf{\Sigma} \tag{9}$$

for an invertible covariance matrix $\mathbf{\Sigma} \in \mathbb{C}^{M\times M}$ and for any vector $\boldsymbol{x} \in \mathbb{C}^M$ it holds that

$$\mathbb{P}\left(\left| \boldsymbol{a}^k \boldsymbol{x} \right| \geqslant t \|\boldsymbol{x}\|_2 \right) \leqslant \mathrm{e}^{-\frac{t^2}{b^2}}. \tag{10}$$

In [10, Theorem 2] the authors state that these matrices are good compression matrices formalized in the following result.

Theorem 4.1. *Let $\boldsymbol{A} \in \mathbb{C}^{K\times M}$ be a b-sub-Gaussian matrix with population covariance $\mathbf{\Sigma}$ and measurements given by $\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x}$ for $\boldsymbol{x}$ as in* (4)*. Assume furthermore that*

$$\min_{i \neq j \in [S]} |\theta_i - \theta_j| \geqslant \frac{4}{L}.$$

Then as long as

$$K \geqslant cS \log(L) b^{-2} \sigma(\mathbf{\Sigma})$$

for a fixed constant c, $\boldsymbol{x}$ is the unique minimizer of (7) *with probability at least* $1 - \exp(-(K-2)/8)$*. Here $\sigma(\mathbf{\Sigma})$ is the condition number of $\mathbf{\Sigma}$.*

In our setting, since $\mathbf{\Gamma} = \mathbf{\Phi} \cdot \boldsymbol{G}$ and the matrix $\boldsymbol{G}$ is completely determined by the antenna array, we can only choose $\mathbf{\Phi}$ freely. We draw the rows $\boldsymbol{\varphi}^k$ of $\mathbf{\Phi}$ for $k = 1, \ldots, K$ i.i.d. from a random distribution F with values in $\mathbb{C}^M$. Then clearly $\mathbf{\Gamma} \sim \hat{F}$ for some distribution $\hat{F}$ with values in $\mathbb{C}^{K\times M}$. The next definition describes the type of distribution we use for the entries of $\mathbf{\Phi}$.

Definition 4.2 (sub-Gaussian random variable)**.** A real valued random variable X is called sub-Gaussian with variance factor c, if

$$\mathbb{E} \exp(\lambda X) \leqslant \lambda^2 \frac{c}{2}$$

for all $\lambda \in \mathbb{R}$.

Let $\varphi_{k,m}$ be independently and identically distributed for each $k \in 1, \ldots, K, m \in 1, \ldots, M$ as centered sub-Gaussian distributions with variance 1. With these, we define

$$\mathbf{\Phi} = (\varphi_{k,m})_{k,m} \quad \text{and as above} \quad \mathbf{\Gamma} = \mathbf{\Phi} \cdot \boldsymbol{G}. \tag{11}$$

Now, we only need to verify that the distribution of $\mathbf{\Gamma}$ is b-sub-Gaussian (in the matrix sense) and calculate the parameters b and $\mathbf{\Sigma}$ to specify the number of measurements needed.

To this end, we need two standard results from probability theory, which we summarize from results in [12, Chapters 2.2, 2.3]:

Lemma 4.1. *Let $\xi_1, \ldots, \xi_n$ for $n \in \mathbb{N}$ be independent real valued random variables. Then*

1. $\mathbb{E}\left\{ \exp\left(t \cdot \sum_{i=1}^{n} \xi_i \right) \right\} = \prod_{i=1}^{n} \mathbb{E}\left\{ \exp(t \cdot \xi_i) \right\}$
2. *Let X be a real valued sub-Gaussian random variable X with variance factor c, then*

$$\mathbf{P}(X > t), \mathbf{P}(X < -t) \leqslant exp(-\frac{t^2}{2c}) \tag{12}$$

With this Lemma at hand, we calculate the moment generating function of $\mathbf{\Gamma}_{k,\ell}$ for $k = 1, \ldots, K$ and $\ell = 1, 2, \ldots, L$ via

Proposition 4.1. *Using the measurement setup from* (11)*, the matrix* $\mathbf{\Phi} \cdot \boldsymbol{G}$ *is* b*-sub-Gaussian with* $b^{-1} = \max_{1 \leqslant \ell \leqslant L} \|\boldsymbol{g}_\ell\|_2^2$ *and* $\mathbf{\Sigma} = \boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}$*.*

Proof. First of all note that since each element of $\mathbf{\Phi}$ has zero mean with independently drawn rows, we trivially have $\mathbb{E}(\boldsymbol{\gamma}^k) = \mathbf{0}$ as well as $\mathbb{E}(\boldsymbol{\gamma}^k \boldsymbol{\gamma}^{\ell^{\mathrm{H}}}) = \delta_{k\ell}$ where $\boldsymbol{\gamma}^k$ is the k-th row of $\mathbf{\Gamma}$. Also, we have $\mathbb{E}(\boldsymbol{\gamma}^{k^{\mathrm{H}}} \boldsymbol{\gamma}^k) = \boldsymbol{G}^{\mathrm{H}}\boldsymbol{G} = \mathbf{\Sigma}$ for all k which is invertible since $\boldsymbol{G}$ was assumed to have full row-rank. What remains to be shown is condition (10) for an appropriate value of b. To this end, we first calculate

$$\mathbb{E}\left\{\exp(t\mathbf{\Gamma}_{r,s})\right\} = \prod_{m=1}^{M} \mathbb{E}\exp(t\boldsymbol{\varphi}_{r,m} g_{m,s})) \leqslant \frac{1}{2}t^2 \sum_{m=1}^{M} g_{m,s}^2$$

where we made use of both statements of Lemma 4.1. For some given $\boldsymbol{v} \in \mathbb{C}^L$ and $\boldsymbol{\gamma}^r$ being the r-th row of $\mathbf{\Gamma}$ this can also be used to calculate

$$\mathbb{E}\left\{\exp(t\boldsymbol{\gamma}^r \boldsymbol{v})\right\} = \prod_{\ell=1}^{L}\prod_{k=1}^{K} \mathbb{E}\left\{\exp(v_\ell g_{k,\ell} t)\right\} \leqslant \frac{1}{2}\|\boldsymbol{v}\|_2^2 t^2 \max_{1 \leqslant \ell \leqslant L} \|\boldsymbol{g}_\ell\|_2^2,$$

where $\boldsymbol{g}_\ell$ denotes the ℓ-th column of $\boldsymbol{G}$. With this we have shown that $\mathbf{\Gamma}$ obeys (10) and hence it is b-sub-Gaussian for $b^{-1} = \max_{1 \leqslant \ell \leqslant L} \|\boldsymbol{g}_\ell\|_2^2$. □

This results in the validity of the measurement scheme described above if the number of measurements obeys

$$K \geqslant \hat{c} S \log(M) \cdot \max_{1 \leqslant \ell \leqslant L} \|\boldsymbol{g}_\ell\|_2^2 \cdot \sigma(\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}), \tag{13}$$

where $\hat{c}$ is a fixed constant. It is worth noting that the above result is optimal up to the logarithmic factor, as already stated in [10]. This illuminates that fact that sub-Gaussian measurements with similar constants variance factors will qualitatively perform similarly at atomic norm minimization. As explained in Section 2 an example for compression matrices that are easy to realize in hardware is given by binary ± 1 matrices. These correspond to the i.i.d. Rademacher distribution, which yields $c = 1$ as used in above Proposition.

5. NUMERICAL RESULTS

To demonstrate the performance of the proposed estimator, we implemented it using CVX and the SDPT3 solver [13, 14]. For the first experiment, we use a stacked polarimetric uniform circular patch array (SPUCPA) which is depicted in Figure 1. It consists of two stacked 12-element uniform circular patch arrays and an additional cube of five patch elements on top. Each element has two ports for vertical and horizontal polarization. For the simulations, we use only the two rings of 12 elements and only the vertical port so that $M = 24$ ports are available. Moreover, the EADF for the beam pattern (which was measured in an anechoic chamber) contains $L = 25$ coefficients per antenna element.

To quantify the performance of our estimator, we compare it to the deterministic Cramér-Rao Bound (CRB). Without spatial compression, the CRB can be computed via

$$C(\boldsymbol{\theta}) = \frac{\sigma^2}{2} \operatorname{tr}\left(\left[\Re(\boldsymbol{D}^{\mathrm{H}} \mathbf{\Pi}_{\boldsymbol{A}}^{\perp} \boldsymbol{D} \odot (\boldsymbol{c}\boldsymbol{c}^{\mathrm{H}})^{\mathrm{T}})\right]^{-1}\right), \tag{14}$$

Fig. 1. Stacked polarimetric uniform circular patch array with 58 ports. We have used only the ports corresponding to the two stacked circular arrays with 12 elements per ring.

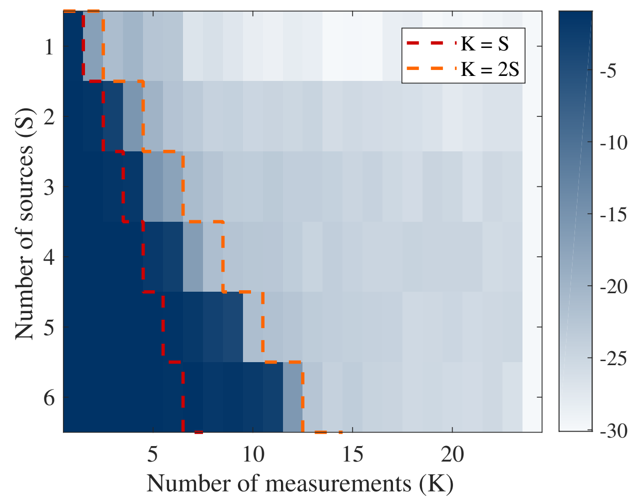

Fig. 2. Estimation error (logarithmic scale) vs. K, S for the noise-free case and Rademacher distributed compression matrices.

where $\boldsymbol{A} = \boldsymbol{G}\boldsymbol{F}(\boldsymbol{\theta})$, $\boldsymbol{D} = \jmath \boldsymbol{G} \operatorname{diag}(\boldsymbol{\mu})\boldsymbol{F}(\boldsymbol{\theta})$ and $\mathbf{\Pi}_{\boldsymbol{A}}^{\perp} = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{H}}$. Here $\boldsymbol{F}(\boldsymbol{\theta}) = [\boldsymbol{f}(\theta_1), \ldots, \boldsymbol{f}(\theta_S)]$, the vector $\boldsymbol{\mu} = -(L-1)/2, \ldots, +(L-1)/2$ and $\Re$ denotes the real part of a complex number. If we incorporate compression, (14) changes to [11]

$$C(\boldsymbol{\theta}) = \frac{\sigma^2}{2} \operatorname{tr}\left(\left[\Re(\bar{\boldsymbol{D}}^{\mathrm{H}} \mathbf{\Pi}_{\bar{\boldsymbol{A}}}^{\perp} \bar{\boldsymbol{D}} \odot (\boldsymbol{c}\boldsymbol{c}^{\mathrm{H}})^{\mathrm{T}})\right]^{-1}\right), \tag{15}$$

where $\bar{\boldsymbol{A}} = \mathbf{\Phi}\boldsymbol{A} = \mathbf{\Gamma}\boldsymbol{F}(\boldsymbol{\theta})$ and $\bar{\boldsymbol{D}} = \mathbf{\Phi}\boldsymbol{D} = \jmath\mathbf{\Gamma} \operatorname{diag}(\boldsymbol{\mu})\boldsymbol{F}(\boldsymbol{\theta})$.

In Figure 2 we show the empirical phase transition for the noiseless case $\boldsymbol{w} = \mathbf{0}$. We vary the number of sources S and the number of measurements K and draw the source positions randomly such that the separation condition from Theorem 4.1 is always satisfied. Moreover, the amplitudes $\boldsymbol{c}$ are drawn randomly on the complex unit circle. We depict the empirical estimation error defined as $\sum_{s=1}^{S}(\theta_s - \hat{\theta}_s)^2$ on a logarithmic scale, i.e., values below -10 correspond to an estimation error below 10^{-10} which can be considered to be rounding errors of the floating point representation. Figure 2 considers a ± 1 binary Rademacher distribution, showing the best realization among 100 trials. We observe a quite sharp phase transition that occurs between $K = S$ and $K = 2S$ illustrated by the fact that left of the line corresponding to $K = S$ the reconstruction error is constantly high, whereas right of the line indicating $K = 2S$ we only observe rounding errors. This confirms the theoretical prediction from Theorem 4.1 that suggests that

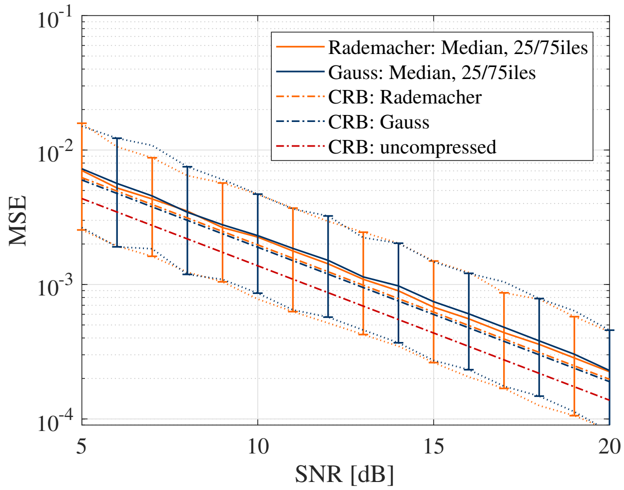

Fig. 3. Estimation error vs. SNR for $S = 2$ sources and $K = 12$

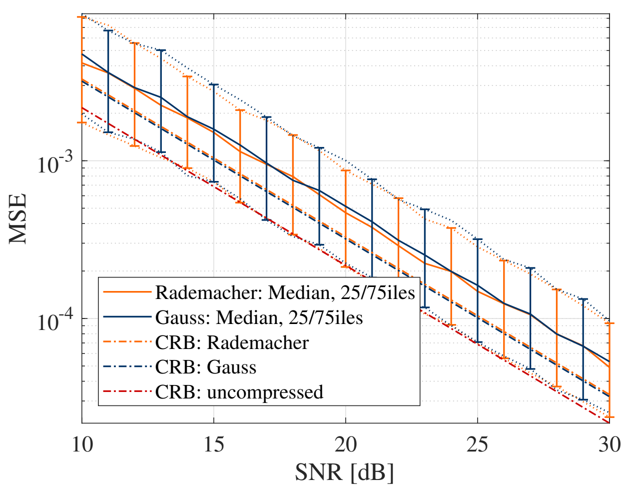

Fig. 4. Estimation error vs. SNR for $S = 3$ sources and $K = 15$

the required number of measurements K scales linearly with S.

To investigate the performance in the presence of noise with variance σ^2, the results shown in Figures 3 and 4 show the empirical estimation error of the proposed method vs. the CRB for the compressed and the uncompressed cases, where the optimization in (7) was run with $\varepsilon = \sigma/K$. We compare the effect of using Gaussian and Rademacher distributed compression matrices. For both we show the median (solid lines) and the 25/75 percentiles (error bars and dotted lines). Moreover, the corresponding CRBs are shown in dash-dotted lines. In Figure 3 we consider $S = 2$ sources and $K = 12$ (i.e., a compression rate of 50 %) whereas in Figure 4 we choose $S = 3$ sources and $K = 15$ (i.e., a compression rate of 62.5 %). As before, the source positions are drawn randomly such that they always obey the separation condition from Theorem 4.1. The results show the statistics over 2500 trials and confirm that both distributions behave very similarly and provide estimation errors that are close to the CRB.

In the last experiment we consider an array of $M = 29$ isotropic antenna elements in a randomly generated array geometry, which is depicted in Figure 5. We generate $S = 5$ sources at random positions in a manner similar to the previous experiment. Figure 6 shows

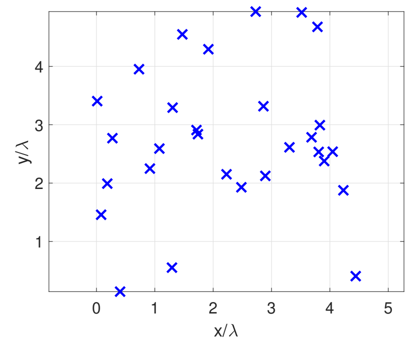

Fig. 5. Randomly drawn array geometry ($M = 29$) for the experiment shown in Figure 6.

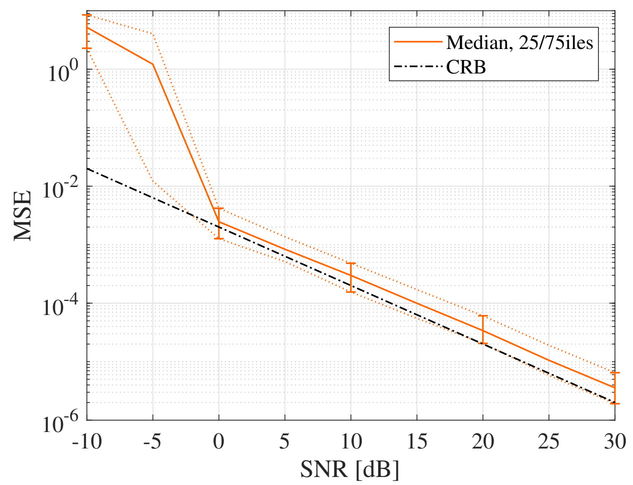

Fig. 6. Estimation error vs. SNR for $S = 5$ sources the $M = 29$ element array shown in Figure 5 (no compression).

the empirical estimation error vs. the CRB for the case where no compression is applied ($\boldsymbol{\Phi} = \boldsymbol{I}_M$). The result demonstrates that the proposed method enables grid-free sparsity-based DOA estimation with arbitrary array geometries and that it achieves the Cramér-Rao Bound.

6. CONCLUSION

In this paper we investigate the problem of grid-free sparsity-based direction of arrival (DOA) estimation using arbitrary antenna arrays and compressed sensing. Applying a Fourier-based description of the antenna array response we show that the DOA problem can be cast as a generalized line spectral estimation problem that has recently been studied in this context. In particular, we show that an atomic norm minimization based framework can be applied that provides the DOA estimates based on solving a convex semidefinite program. Our description allows to incorporate spatial compressive sensing via a randomly chosen compression matrix. Tuning existing recovery guarantees to our setting allows to derive a condition on the required number of measurements that depends on the array's beam pattern as well as the distribution of the compression matrix. Our approach is applicable to arbitrary sub-Gaussian distributions, including the binary Rademacher distribution that can be realized in hardware very efficiently and shows a performance very close to the Gaussian distribution.

7. REFERENCES

[1] H. Karim and M. Viberg, “Two decades of array signal processing research: The parametric approach,” *IEEE Signal Processing Magazine*, vol. 13, no. 4, Jul. 1996.

[2] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[3] D. Malioutov, M. Cetin, and A. S. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[4] M. Hyder and K. Mahata, “Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation,” *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4646–4655, Sep. 2010.

[5] H. Zhu, G. Leus, and G. B. Giannakis, “Sparsity-cognizant total least-squares for perturbed compressive sampling,” *IEEE Transactions on Signal Processing*, vol. 59, no. 5, May 2011.

[6] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, “Compressed sensing off the grid,” *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[7] E. J. Candès and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[8] A. Xenaki and P. Gerstoft, “Grid-free compressive beamforming,” *The Journal of the Acoustical Society of America*, vol. 137, no. 4, pp. 1923–1935, 2015. [Online]. Available: https://doi.org/10.1121/1.4916269

[9] M. Landmann, A. Richter, and R. S. Thoma, “DoA resolution limits in MIMO channel sounding,” in *IEEE Antennas and Propagation Society Symposium*, vol. 2, Jun. 2004, pp. 1708–1711.

[10] R. Heckel and M. Soltanolkotabi, “Generalized line spectral estimation via convex optimization,” *arXiv preprint arXiv:1609.08198*, 2016.

[11] M. Ibrahim, V. Ramireddy, A. Lavrenko, J. König, F. Römer, M. Landmann, M. Grossmann, G. D. Galdo, and R. S. Thomä, “Design and analysis of compressive antenna arrays for direction of arrival estimation,” *Elsevier Signal Processing*, vol. 138, pp. 35 – 47, Sep. 2017.

[12] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013.

[13] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” Mar. 2014.

[14] ——, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.