1-Bit Direction of Arrival Estimation based on Compressed Sensing

Christoph Stöckle, Jawad Munir, Amine Mezghani and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Munich University of Technology, 80290 Munich, Germany
E-Mail: {christoph.stoeckle, jawad.munir, amine.mezghani, josef.a.nossek}@tum.de

***Abstract*—Massive MIMO plays an important role for future cellular networks since the large number of antenna elements is capable of increasing the spectral efficiency and the amount of usable spectrum. The 1-bit analog-to-digital converters can drastically reduce the resulting complexity and power consumption. Therefore, we investigate the Direction of Arrival (DoA) estimation using 1-bit measurements of many antenna elements in this paper. We extend Binary Iterative Hard Thresholding (BIHT), an efficient sparse recovery algorithm from the area of Compressed Sensing (CS) that takes the 1-bit quantization explicitly into account, to complex-valued signals and multiple measurement vectors such that it is applicable to 1-bit DoA estimation with multiple snapshots. The comparison of the resulting Complex-valued BIHT (CBIHT) algorithm to subspace- and CS-based methods in terms of both DoA estimation performance and computational complexity demonstrates that CBIHT is well suited for scenarios with many antenna elements and a few snapshots.**

***Index Terms*—Direction of Arrival Estimation, Compressed Sensing, Massive MIMO, 1-Bit Quantization.**

I. Introduction

The deployment of a very large number of antenna elements at the base station (BS) known as massive MIMO plays an important role in future cellular networks. This is due to the fact that the large antenna gains achievable by the large number of antenna elements can increase the spectral efficiency by beamforming, i.e., sending and receiving signals through narrow directed beams [1], and the amount of usable spectrum by making the communication in the underutilized mmWave frequency bands viable [2]. The main drawbacks of massive MIMO are the high complexity and power consumption resulting from the large number of antenna elements. The power consumption of the analog-to-digital converter (ADC), one of the most power-hungry devices, can be reduced exponentially by decreasing the resolution [3] and 1-bit quantization can drastically simplify other RF-components, e.g., amplifiers and mixers. Therefore, we investigate the Direction of Arrival (DoA) estimation from 1-bit measurements obtained by a large number of antenna elements in this paper. The estimated DoAs of the signals arriving at the BS can be used for beamforming.

DoA estimation from unquantized measurements based on Compressed Sensing (CS) has already been studied, e.g., in [4] and [5], by formulating it as a sparse recovery problem. Several powerful methods have been developed for solving sparse recovery problems, e.g., Basis Pursuit Denoise (BPDN) [6], [7]. Furthermore, there are also CS recovery algorithms like Binary Iterative Hard Thresholding (BIHT) that take the 1-bit quantization of the measurements explicitly into account [8]. However, BIHT was introduced for real-valued signals and a single measurement vector (SMV). In order to apply it to the 1-bit DoA estimation with multiple snapshots, we extend it to complex-valued signals and multiple measurement vectors (MMV). The resulting algorithm is named Complex-valued BIHT (CBIHT). It will be shown that the 1-bit DoA estimation performance of CBIHT is comparable to that of BPDN as well as the subspace-based DoA estimation methods MUSIC and ESPRIT or even better for a large number of antenna elements and a small number of snapshots while its computational complexity is relatively small.

The paper is organized as follows. In section II, a measurement model for the 1-bit DoA estimation with the help of a uniform linear array (ULA) is introduced and the 1-bit DoA estimation is formulated as a sparse recovery problem. In section III, we describe the sparse recovery methods BPDN and CBIHT for solving the sparse recovery problem. Section IV presents the results of the simulations we have conducted in order to examine the DoA estimation performance of BPDN, CBIHT, MUSIC and ESPRIT before the computational complexity of those methods is compared in section V. Finally, the paper is concluded in section VI.

II. Problem Formulation

A. Measurement Model

An M-element ULA with an inter-element spacing of half the wavelength is located in the far field of K sources emitting narrow-band signals. It receives the signal $s_k(t) \in \mathbb{C}$ of source $k \in \{1, 2, ..., K\}$ from DoA $\theta_k \in [-90°, 90°)$ corresponding to the spatial frequency $\mu_k = -\pi \sin(\theta_k)$. At time instant t, the 1-bit measurement acquired by the antenna element $m \in \{1, 2, ..., M\}$ of the ULA is given by

$$x_m(t) = \mathrm{Q}\left(\sum_{k=1}^{K} s_k(t)\, \mathrm{e}^{j(m-1)\mu_k} + n_m(t)\right), \qquad (1)$$

where $n_m(t) \in \mathbb{C}$ is an additive measurement noise and $\mathrm{Q}(.)$ is the 1-bit quantizer, which keeps only the sign of the real and imaginary part, i.e., $\mathrm{Q}(z) = \mathrm{sign}(\Re\{z\}) + j\,\mathrm{sign}(\Im\{z\})$ for $z \in \mathbb{C}$ with

$$\mathrm{sign}: \ \mathbb{R} \to \{-1, +1\}, x \mapsto \mathrm{sign}(x) = \begin{cases} -1, & x \leq 0 \\ +1, & x > 0 \end{cases}.$$

978-1-4799-1931-4/15/$31.00 ©2015 IEEE

Both $\mathrm{Q}(.)$ and $\mathrm{sign}(.)$ are applied element-wise to matrices.

Using the array steering matrix $\mathbf{A} = [\mathbf{a}(\mu_1), \ldots, \mathbf{a}(\mu_K)] \in \mathbb{C}^{M \times K}$ with $\mathbf{a}(\mu_k) = \left[1, \mathrm{e}^{j\mu_k}, \ldots, \mathrm{e}^{j(M-1)\mu_k}\right]^{\mathrm{T}} \in \mathbb{C}^M$, the signal vector $\mathbf{s}(t) = [s_1(t), \ldots, s_K(t)]^{\mathrm{T}} \in \mathbb{C}^K$ and the noise vector $\mathbf{n}(t) = [n_1(t), \ldots, n_M(t)]^{\mathrm{T}} \in \mathbb{C}^M$, the measurement model for one snapshot or an SMV $\mathbf{x}(t) = [x_1(t), \ldots, x_M(t)]^{\mathrm{T}} \in \mathbb{C}^M$ at time instant t can be formulated as

$$\mathbf{x}(t) = \mathrm{Q}(\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)). \tag{2}$$

It can be extended to multiple snapshots or MMV $\mathbf{x}(t_n)$ at N time instants t_n, $n = 1, 2, ..., N$:

$$\mathbf{X} = \mathrm{Q}(\mathbf{A}\mathbf{S} + \mathbf{N}) \tag{3}$$

with the matrix of measurements $\mathbf{X} = [\mathbf{x}(t_1), \ldots, \mathbf{x}(t_N)] \in \mathbb{C}^{M \times N}$, the signal matrix $\mathbf{S} = [\mathbf{s}(t_1), \ldots, \mathbf{s}(t_N)] \in \mathbb{C}^{K \times N}$ and the noise matrix $\mathbf{N} = [\mathbf{n}(t_1), \ldots, \mathbf{n}(t_N)] \in \mathbb{C}^{M \times N}$.

B. DoA Estimation as Sparse Recovery Problem

Compressed Sensing aims at solving sparse recovery problems [6]. In the SMV case, a signal vector $\mathbf{s} \in \mathbb{C}^P$ that is K-sparse, i.e., has at most $K \ll P$ non-zero entries, is to be recovered from $M < P$ linear measurements $\mathbf{x} = \mathbf{A}\mathbf{s} \in \mathbb{C}^M$ taken by the measurement matrix $\mathbf{A} \in \mathbb{C}^{M \times P}$. The support of the vector $\mathbf{s} = [s_1, \ldots, s_P]^{\mathrm{T}}$ is defined as the index set $\mathrm{supp}(\mathbf{s}) = \{i : s_i \neq 0\}$ of its non-zero entries. The so-called ℓ_0 "norm" $\|\mathbf{s}\|_0 = |\mathrm{supp}(\mathbf{s})|$ of the vector $\mathbf{s}$ counts its non-zero elements. In the MMV case, a signal matrix $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_N] \in \mathbb{C}^{P \times N}$ that consists of N jointly K-sparse signal vectors $\mathbf{s}_n$, $n = 1, 2, ..., N$, with the same support and is therefore row K-sparse, i.e., has at most $K \ll P$ non-zero rows, is to be recovered from the matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N] = \mathbf{A}\mathbf{S} \in \mathbb{C}^{M \times N}$ consisting of N measurement vectors $\mathbf{x}_n$. The number of non-zero rows of $\mathbf{S}$ can be expressed as $\|\mathbf{S}\|_{p,0}$. The mixed $\ell_{p,q}$ norm of $\mathbf{S}$ with rows $\mathbf{s}^i$, $i = 1, 2, ..., P$, is defined as

$$\|\mathbf{S}\|_{p,q} = \left\| \left[\|\mathbf{s}^1\|_p, \|\mathbf{s}^2\|_p, \ldots, \|\mathbf{s}^P\|_p \right] \right\|_q. \tag{4}$$

A further aspect of CS is the recovery of sparse signals from noisy and quantized measurements, which makes the CS framework applicable to the 1-bit DoA estimation.

Up to now, the signal matrix $\mathbf{S}$ in the measurement model (3) is not row K-sparse. In order to make it row K-sparse, the angle $\theta \in [-90°, 90°)$ is discretized similarly to [4], which results in a sampling grid $\{\theta_1, \theta_2, ..., \theta_P\}$ of $P \gg K$ potential DoAs θ_i corresponding to the spatial frequencies $\mu_i = -\pi \sin(\theta_i)$, $i = 1, 2, ..., P$. Let $\mathcal{I} = \{i_1, i_2, ..., i_K\} \subset \{1, 2, ..., P\}$ be an ordered set of indices i_k indicating that the k^{th} source is located at the grid point i_k. The array steering matrix $\mathbf{A} = [\mathbf{a}(\mu_{i_1}), \ldots, \mathbf{a}(\mu_{i_K})] \in \mathbb{C}^{M \times K}$ of the measurement model (3), whose k^{th} column $\mathbf{a}(\mu_{i_k})$ corresponds to the k^{th} source, is extended to the array steering matrix $\mathbf{A} = [\mathbf{a}(\mu_1), \ldots, \mathbf{a}(\mu_P)] \in \mathbb{C}^{M \times P}$, whose i^{th} column $\mathbf{a}(\mu_i)$ corresponds to the direction θ_i of the sampling grid. The signal vectors $\mathbf{s}(t_n)$ in the signal matrix $\mathbf{S}$ of the measurement model (3) are extended from $\mathbf{s}(t) = [s_{i_1}(t), \ldots, s_{i_K}(t)]^{\mathrm{T}} \in \mathbb{C}^K$, whose k^{th} element $s_{i_k}(t)$ is the signal from source k, to $\mathbf{s}(t) = [s_1(t), \ldots, s_P(t)]^{\mathrm{T}} \in \mathbb{C}^P$, whose i^{th} element $s_i(t)$ is the signal from the direction θ_i of the sampling grid. Since

$$s_i(t) = \begin{cases} s_{i_k}(t), & i = i_k \in \mathcal{I} \\ 0, & \text{otherwise} \end{cases}, i = 1, 2, ..., P,$$

and thus $s_i(t)$ is only non-zero if one of the K sources is located at grid point i, the signal vectors $\mathbf{s}(t_n)$ are jointly K-sparse with $\mathrm{supp}(\mathbf{s}(t_n)) = \mathcal{I}$ and the signal matrix $\mathbf{S}$ is row K-sparse. The 1-bit DoA estimation can be considered as the sparse recovery problem of obtaining a row K-sparse estimate $\hat{\mathbf{S}} = [\hat{\mathbf{s}}(t_1), \ldots, \hat{\mathbf{s}}(t_N)]$ for the true row K-sparse signal matrix $\mathbf{S}$ from the 1-bit measurements $\mathbf{X}$ with the array steering matrix $\mathbf{A}$ as the measurement matrix. The indices of the non-zero rows of $\hat{\mathbf{S}}$, i.e., $\mathrm{supp}(\hat{\mathbf{s}}(t_n)) = \{l_1, l_2, ..., l_K\}$, finally determine the estimates θ_{l_k} for the true DoAs θ_{i_k} of the K sources.

III. Solution of Sparse Recovery Problem

A. Basis Pursuit Denoise (BPDN)

A natural approach is to treat the 1-bit measurements in $\mathbf{X}$ as signal values contaminated by errors resulting from the measurement noise and the 1-bit quantization. Then the row K-sparse signal matrix $\mathbf{S}$ can be recovered by using BPDN [7]:

$$\hat{\mathbf{S}} = \underset{\tilde{\mathbf{S}} \in \mathbb{C}^{P \times N}}{\mathrm{argmin}} \left\| \tilde{\mathbf{S}} \right\|_{2,1} \quad \text{s.t.} \quad \left\| \mathbf{A}\tilde{\mathbf{S}} - \mathbf{X} \right\|_F \leq \beta. \tag{5}$$

The objective of the optimization problem ensures that the estimate $\hat{\mathbf{S}}$ is row sparse while its constraint forces it to be consistent with the measurements $\mathbf{X}$. The regularization parameter β has to be chosen appropriately depending on the measurement errors caused by the noise and the 1-bit quantization, which is the main drawback of this approach.

B. Binary Iterative Hard Thresholding (BIHT)

In contrast to BPDN, the BIHT algorithm described in [8] takes the 1-bit quantization explicitly into account. It aims at recovering the K-sparse real-valued signal vector $\mathbf{s} \in \mathbb{R}^P$ from an SMV $\mathbf{x} = \mathrm{sign}(\mathbf{A}\mathbf{s}) \in \{-1, +1\}^M$ of measurements taken by the real-valued measurement matrix $\mathbf{A} \in \mathbb{R}^{M \times P}$ and quantized to 1 bit. As greedy methods, BIHT and its variant BIHT-ℓ_2 compute an estimate $\hat{\mathbf{s}}$ for $\mathbf{s}$ by trying to solve the optimization problem

$$\hat{\mathbf{s}} = \underset{\tilde{\mathbf{s}} \in \mathbb{R}^P}{\mathrm{argmin}} \left\| [\mathbf{x} \odot (\mathbf{A}\tilde{\mathbf{s}})]_- \right\|_p^p \text{ s.t. } \|\tilde{\mathbf{s}}\|_0 \leq K, \|\tilde{\mathbf{s}}\|_2 = 1 \tag{6}$$

iteratively for either $p = 1$ or $p = 2$. The first constraint guarantees that $\hat{\mathbf{s}}$ is K-sparse. Since the scaling of the true signal vector $\mathbf{s}$ is lost during the 1-bit quantization process, it can only be reconstructed up to a scaling factor. Therefore, the estimate $\hat{\mathbf{s}}$ is normalized to unit ℓ_2 norm as ensured by the second constraint. The objective forces the estimate $\hat{\mathbf{s}}$ to be consistent with the 1-bit measurements $\mathbf{x}$. Here, $\odot$ denotes the element-wise product and $[.]_-$ the negative function, which sets all non-negative entries to 0.

Each iteration consists of two steps, a gradient descent step and a hard thresholding step. The gradient descent step

$$\check{\mathbf{s}}^{(i+1)} = \hat{\mathbf{s}}^{(i)} - \mu \mathbf{A}^{\mathrm{T}} \begin{cases} \left(\operatorname{sign}\left(\mathbf{A}\hat{\mathbf{s}}^{(i)}\right) - \mathbf{x}\right), & p = 1 \\ \left(\mathbf{x} \odot \left[\mathbf{x} \odot \left(\mathbf{A}\hat{\mathbf{s}}^{(i)}\right)\right]_-\right), & p = 2 \end{cases} \tag{7}$$

with step size $\mu \in \mathbb{R}^+$ starting at the current estimate $\hat{\mathbf{s}}^{(i)}$ reduces the consistency-enforcing objective [9]. The hard thresholding step ensures that the first constraint of (6) is fulfilled by applying the hard thresholding operator $H_K(.)$, which sets all but the K largest in magnitude elements to 0, to the resulting vector $\check{\mathbf{s}}^{(i+1)}$ to get a new K-sparse estimate $\hat{\mathbf{s}}^{(i+1)} = H_K\left(\check{\mathbf{s}}^{(i+1)}\right)$. The final estimate $\hat{\mathbf{s}}$ is obtained by normalizing the estimate from the last iteration to unit ℓ_2 norm such that the second constraint of (6) is fulfilled too.

C. Complex-valued BIHT (CBIHT)

Before BIHT can be used for the 1-bit DoA estimation, it has to be extended to complex-valued signals and MMV.

For the recovery of the K-sparse signal vector $\mathbf{s} \in \mathbb{C}^P$ from the SMV $\mathbf{x} \in \mathbb{C}^M$ according to the measurement model (2), we consider the consistency-enforcing objective

$$\mathcal{J}(\mathbf{s}) = \left\| \left[\begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} \odot \begin{bmatrix} \Re\{\mathbf{As}\} \\ \Im\{\mathbf{As}\} \end{bmatrix} \right]_- \right\|_p^p. \tag{8}$$

The objective $\mathcal{J}(\mathbf{s})$ can be expressed as

$$\mathcal{J}(\mathbf{s}) = \left\| [\mathbf{x}_r \odot (\mathbf{A}_r \mathbf{s}_r)]_- \right\|_p^p = \mathcal{J}_r(\mathbf{s}_r) \tag{9}$$

with the real-valued signal vector

$$\mathbf{s}_r = \begin{bmatrix} \Re\{\mathbf{s}\} \\ \Im\{\mathbf{s}\} \end{bmatrix} \in \mathbb{R}^{2P}, \tag{10}$$

the real-valued measurement matrix

$$\mathbf{A}_r = \begin{bmatrix} \Re\{\mathbf{A}\} & -\Im\{\mathbf{A}\} \\ \Im\{\mathbf{A}\} & \Re\{\mathbf{A}\} \end{bmatrix} \in \mathbb{R}^{(2M)\times(2P)} \tag{11}$$

and the real-valued 1-bit measurements

$$\mathbf{x}_r = \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} \in \{-1, +1\}^{2M} \tag{12}$$

since

$$\begin{bmatrix} \Re\{\mathbf{As}\} \\ \Im\{\mathbf{As}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{A}\} & -\Im\{\mathbf{A}\} \\ \Im\{\mathbf{A}\} & \Re\{\mathbf{A}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{s}\} \\ \Im\{\mathbf{s}\} \end{bmatrix} = \mathbf{A}_r \mathbf{s}_r. \tag{13}$$

According to (7), the standard BIHT algorithm reduces the objective $\mathcal{J}_r(\mathbf{s}_r)$ by computing the gradient descent step

$$\check{\mathbf{s}}_r = \mathbf{s}_r - \mu \mathbf{A}_r^{\mathrm{T}} \begin{cases} \left(\operatorname{sign}(\mathbf{A}_r \mathbf{s}_r) - \mathbf{x}_r\right), & p = 1 \\ \left(\mathbf{x}_r \odot [\mathbf{x}_r \odot (\mathbf{A}_r \mathbf{s}_r)]_-\right), & p = 2 \end{cases}. \tag{14}$$

Using (12), (13) and the 1-bit quantizer $\mathrm{Q}(.)$, $\operatorname{sign}(\mathbf{A}_r\mathbf{s}_r) - \mathbf{x}_r$ can be written as

$$\begin{aligned} &\operatorname{sign}(\mathbf{A}_r \mathbf{s}_r) - \mathbf{x}_r = \operatorname{sign}\left(\begin{bmatrix} \Re\{\mathbf{As}\} \\ \Im\{\mathbf{As}\} \end{bmatrix}\right) - \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} \\ &= \begin{bmatrix} \operatorname{sign}(\Re\{\mathbf{As}\}) - \Re\{\mathbf{x}\} \\ \operatorname{sign}(\Im\{\mathbf{As}\}) - \Im\{\mathbf{x}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathrm{Q}(\mathbf{As}) - \mathbf{x}\} \\ \Im\{\mathrm{Q}(\mathbf{As}) - \mathbf{x}\} \end{bmatrix}. \end{aligned} \tag{15}$$

With (12) and (13), it is also possible to write

$$\begin{aligned} &\mathbf{x}_r \odot [\mathbf{x}_r \odot (\mathbf{A}_r \mathbf{s}_r)]_- = \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} \odot \left[\begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} \right. \\ &\left. \odot \begin{bmatrix} \Re\{\mathbf{As}\} \\ \Im\{\mathbf{As}\} \end{bmatrix} \right]_- = \begin{bmatrix} \Re\{\mathbf{x}\} \odot [\Re\{\mathbf{x}\} \odot \Re\{\mathbf{As}\}]_- \\ \Im\{\mathbf{x}\} \odot [\Im\{\mathbf{x}\} \odot \Im\{\mathbf{As}\}]_- \end{bmatrix}. \end{aligned} \tag{16}$$

(15) and (16) can be combined to

$$\mathbf{y}_r = \begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} = \begin{cases} \operatorname{sign}(\mathbf{A}_r \mathbf{s}_r) - \mathbf{x}_r, & p = 1 \\ \mathbf{x}_r \odot [\mathbf{x}_r \odot (\mathbf{A}_r \mathbf{s}_r)]_-, & p = 2 \end{cases}, \tag{17}$$

where $\mathbf{y}$ is defined as

$$\mathbf{y} = \begin{cases} \mathrm{Q}(\mathbf{As}) - \mathbf{x}, & p = 1 \\ \Re\{\mathbf{x}\} \odot [\Re\{\mathbf{x}\} \odot \Re\{\mathbf{As}\}]_- + \\ +j\Im\{\mathbf{x}\} \odot [\Im\{\mathbf{x}\} \odot \Im\{\mathbf{As}\}]_-, & p = 2 \end{cases}. \tag{18}$$

Plugging (10), (17) and

$$\mathbf{A}_r^{\mathrm{T}} = \begin{bmatrix} \Re\{\mathbf{A}^{\mathrm{H}}\} & -\Im\{\mathbf{A}^{\mathrm{H}}\} \\ \Im\{\mathbf{A}^{\mathrm{H}}\} & \Re\{\mathbf{A}^{\mathrm{H}}\} \end{bmatrix} \tag{19}$$

into (14) results in

$$\check{\mathbf{s}}_r = \begin{bmatrix} \Re\{\mathbf{s}\} \\ \Im\{\mathbf{s}\} \end{bmatrix} - \mu \begin{bmatrix} \Re\{\mathbf{A}^{\mathrm{H}}\} & -\Im\{\mathbf{A}^{\mathrm{H}}\} \\ \Im\{\mathbf{A}^{\mathrm{H}}\} & \Re\{\mathbf{A}^{\mathrm{H}}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix}. \tag{20}$$

Applying (13) with $\mathbf{A}^{\mathrm{H}}$ instead of $\mathbf{A}$ and $\mathbf{y}$ instead of $\mathbf{s}$ yields

$$\check{\mathbf{s}}_r = \begin{bmatrix} \Re\{\mathbf{s}\} \\ \Im\{\mathbf{s}\} \end{bmatrix} - \mu \begin{bmatrix} \Re\{\mathbf{A}^{\mathrm{H}}\mathbf{y}\} \\ \Im\{\mathbf{A}^{\mathrm{H}}\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{s} - \mu\mathbf{A}^{\mathrm{H}}\mathbf{y}\} \\ \Im\{\mathbf{s} - \mu\mathbf{A}^{\mathrm{H}}\mathbf{y}\} \end{bmatrix}. \tag{21}$$

If $\check{\mathbf{s}}_r \in \mathbb{R}^{2P}$ is interpreted as $\check{\mathbf{s}}_r = \left[\Re\{\check{\mathbf{s}}\}^{\mathrm{T}}, \Im\{\check{\mathbf{s}}\}^{\mathrm{T}}\right]^{\mathrm{T}}$ with $\check{\mathbf{s}} \in \mathbb{C}^P$, from the last equation, the gradient descent step

$$\check{\mathbf{s}} = \mathbf{s} - \mu \mathbf{A}^{\mathrm{H}} \mathbf{y} \tag{22}$$

for decreasing the new objective $\mathcal{J}(\mathbf{s})$ in (8) follows.

In order to recover the row K-sparse signal matrix $\mathbf{S} \in \mathbb{C}^{P\times N}$ from the MMV $\mathbf{X} \in \mathbb{C}^{M\times N}$ according to the measurement model (3), we would like to solve the optimization problem

$$\begin{aligned} \hat{\mathbf{S}} = \underset{\tilde{\mathbf{S}}\in\mathbb{C}^{P\times N}}{\operatorname{argmin}} & \left\| \left[\begin{bmatrix} \Re\{\mathbf{X}\} \\ \Im\{\mathbf{X}\} \end{bmatrix} \odot \begin{bmatrix} \Re\left\{\mathbf{A}\tilde{\mathbf{S}}\right\} \\ \Im\left\{\mathbf{A}\tilde{\mathbf{S}}\right\} \end{bmatrix} \right]_- \right\|_{p,p}^p \\ \text{s.t.} \quad & \left\|\tilde{\mathbf{S}}\right\|_{q,0} \le K, \left\|\tilde{\mathbf{s}}_n\right\|_2 = 1, n = 1, 2, ..., N, \end{aligned} \tag{23}$$

where $\tilde{\mathbf{s}}_n$ is the n^{th} column of $\tilde{\mathbf{S}}$. The first constraint guarantees that $\hat{\mathbf{S}}$ is row K-sparse. The further constraints ensure that the columns of the estimate $\hat{\mathbf{S}}$ are normalized to unit ℓ_2 norm since the true signal vectors, i.e., the columns of the true signal matrix $\mathbf{S}$, can only be reconstructed up to a scaling factor due to the 1-bit quantization. The objective again forces the estimate $\hat{\mathbf{S}}$ to be consistent with the 1-bit measurements $\mathbf{X}$. In each iteration, the consistency-enforcing objective is reduced by computing the gradient descent step

$$\check{\mathbf{S}}^{(i+1)} = \hat{\mathbf{S}}^{(i)} - \mu \mathbf{A}^{\mathrm{H}} \mathbf{Y}^{(i)} \tag{24}$$

starting at the current estimate $\hat{\mathbf{S}}^{(i)}$ simultaneously for all MMV according to (22) with step size μ and

$$\mathbf{Y}^{(i)} = \begin{cases} \mathrm{Q}\left(\mathbf{A}\hat{\mathbf{S}}^{(i)}\right) - \mathbf{X}, & p = 1 \\ \Re\{\mathbf{X}\} \odot \left[\Re\{\mathbf{X}\} \odot \Re\left\{\mathbf{A}\hat{\mathbf{S}}^{(i)}\right\}\right]_- + \\ +j\Im\{\mathbf{X}\} \odot \left[\Im\{\mathbf{X}\} \odot \Im\left\{\mathbf{A}\hat{\mathbf{S}}^{(i)}\right\}\right]_-, & p = 2 \end{cases}. \tag{25}$$

Applying the hard thresholding operator $\bar{H}_K(.)$ to the resulting matrix $\check{\mathbf{S}}^{(i+1)}$ sets all but the K rows with the largest ℓ_2 norm to $\mathbf{0}^{\mathrm{T}}$ to get a new row K-sparse estimate $\hat{\mathbf{S}}^{(i+1)} = \bar{H}_K\left(\check{\mathbf{S}}^{(i+1)}\right)$, which ensures that the first constraint of (23) is fulfilled. Normalizing the columns of the estimate from the last iteration to unit ℓ_2 norm by the operator $\bar{U}(.)$ yields the final estimate $\hat{\mathbf{S}}$ such that the further constraints of (23) are fulfilled too. We call the resulting algorithms for the cases $p = 1$ and $p = 2$ Complex-valued BIHT (CBIHT) and CBIHT-ℓ_2, respectively. They are summarized in Algorithm 1.

Algorithm 1 Complex-valued BIHT(-ℓ_2) (CBIHT(-ℓ_2))

Input: $\mathbf{X}$, $\mathbf{A}$, K, μ
  Initialize: $\hat{\mathbf{S}}^{(0)} = \mathbf{A}^{\mathrm{H}}\mathbf{X}$, $i = 0$
  while stopping criterion not met **do**
    $\hat{\mathbf{S}}^{(i+1)} = \bar{H}_K\left(\hat{\mathbf{S}}^{(i)} - \mu\mathbf{A}^{\mathrm{H}}\mathbf{Y}^{(i)}\right)$
    $i := i + 1$
  end while
  $\hat{\mathbf{S}} = \bar{U}\left(\hat{\mathbf{S}}^{(i)}\right)$
Output: $\hat{\mathbf{S}}$

IV. DoA Estimation Performance

We have conducted simulations in order to evaluate the performance of the 1-bit DoA estimation. The results were obtained in $R = 100$ Monte Carlo runs. The 1-bit DoA estimation performance is measured in terms of the root mean square error (RMSE)

$$\mathrm{RMSE}_\theta = \sqrt{\frac{1}{RK}\sum_{r=1}^{R}\sum_{k=1}^{K}\left|\hat{\theta}_{k,r} - \theta_k\right|^2} \tag{26}$$

between the DoAs θ_k of the sources $k = 1, 2, ..., K$ and their estimates $\hat{\theta}_{k,r}$ in the runs $r = 1, 2, ..., R$. All sources are uncorrelated and emit equipower QPSK-symbols with zero mean and variance $\sigma_s^2 = 1$, whereas the measurement noise samples are drawn from an i.i.d. circularly symmetric complex Gaussian random process with zero mean and variance σ_n^2. The SNR in dB is defined as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{K\sigma_s^2}{\sigma_n^2}\right)\ \mathrm{dB}. \tag{27}$$

In order to reduce the computational complexity of BPDN, the matrix of measurements $\mathbf{X} \in \mathbb{C}^{M\times N}$ is compressed to the matrix $\mathbf{X}_{\mathrm{SVD}} \in \mathbb{C}^{M\times K}$ used in BPDN instead of $\mathbf{X}$ by the ℓ_1-SVD described in [4] if the number of snapshots is larger than the number of sources, i.e., $N > K$. The sampling grid

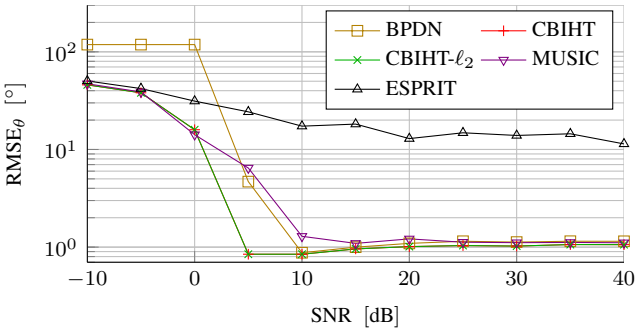

Fig. 1. RMSE_θ vs. SNR for $M = 20$, $N = 2$, $P = 128$, $K = 2$, $\theta_1 = 20.11°$, $\theta_2 = 44.68°$.

used for BPDN, CBIHT, CBIHT-ℓ_2 and MUSIC is constructed from the angles $\theta_i = \arcsin\left(\frac{2}{P}(i-1) - 1\right)$ corresponding to the equally spaced spatial frequencies $\mu_i = \pi - \frac{2\pi}{P}(i-1)$, $i = 1, 2, ..., P$, such that the maximum absolute inner product of two distinct columns of the array steering matrix $\mathbf{A}$ is minimum. In this case, the squared ℓ_2 norm of the rows of the initial guess $\hat{\mathbf{S}}^{(0)} = \mathbf{A}^{\mathrm{H}}\mathbf{X}$ for the greedy methods CBIHT and CBIHT-ℓ_2 is the Fourier spectrum. In contrast to DoA estimation using the Fourier spectrum, DoA estimation using CBIHT or CBIHT-ℓ_2, however, is not restricted to ULAs and takes the 1-bit quantization into account, which is important. Since the absolute inner product of neighboring columns of $\mathbf{A}$ is still pretty high for a large P, we decided to let the hard thresholding operator $\bar{H}_K(.)$ in CBIHT and CBIHT-ℓ_2 set all but the K rows at which the ℓ_2 norm of the rows has its largest peaks to $\mathbf{0}^{\mathrm{T}}$. $\bar{H}_K(.)$ is also applied to the estimate of BPDN to make it row K-sparse if it has more than K non-zero rows. In all simulations, 10 iterations of CBIHT and CBIHT-ℓ_2 are performed while the step size is chosen to be $\mu = \|\mathbf{A}\|_2^{-2}$.

In Fig. 1, the RMSE is plotted over the SNR for $M = 20$ antenna elements, $N = 2$ snapshots, $P = 128$ grid points, and $K = 2$ sources at $\theta_1 = 20.11°$ and $\theta_2 = 44.68°$. ESPRIT is not able to estimate the DoAs, because the number of snapshots is too small to estimate the subspaces. Since MUSIC has the same problem, its performance is worse than the one of the CS-based methods for medium SNR. The greedy methods perform equally well and better than BPDN, which exhibits a performance similar to the one of MUSIC and the greedy methods if the SNR is high.

The RMSE is plotted over the number of snapshots N in Fig. 2 where the SNR is 10 dB, $M = 100$, $P = 1024$, and $K = 2$ sources are located at $\theta_1 = 4.93°$ and $\theta_2 = 10.01°$. The performance of all methods becomes better with increasing number of snapshots. While MUSIC and the CS-based methods perform equally well, they clearly outperform ESPRIT, whose RMSE is quite high for a small number of snapshots because the subspace estimation is not good enough. To achieve an RMSE of at most $0.5°$, for all methods except ESPRIT, only one snapshot is sufficient. In comparison, ESPRIT requires more than 8 snapshots for that. Moreover, MUSIC and the CS-based methods, which choose their estimate from the grid of potential DoAs, do not make any mistake in estimating the

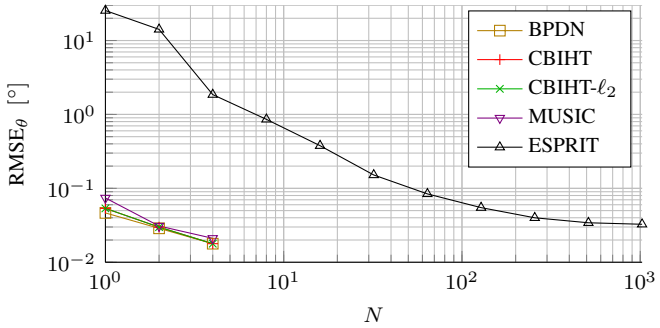

Fig. 2. RMSE$_\theta$ vs. N for SNR $= 10$ dB, $M = 100$, $P = 1024$, $K = 2$, $\theta_1 = 4.93°$, $\theta_2 = 10.01°$.

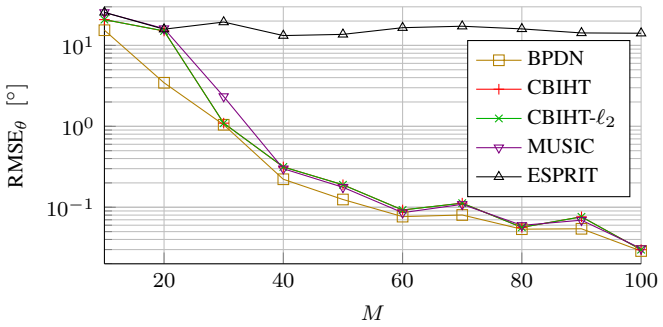

Fig. 3. RMSE$_\theta$ vs. M for SNR $= 10$ dB, $N = 2$, $P = 1024$, $K = 2$, $\theta_1 = 4.93°$, $\theta_2 = 10.01°$.

DoAs for 8 and more snapshots.

The relationship between the RMSE and the number of antenna elements M is illustrated in Fig. 3 where the SNR is 10 dB, $N = 2$, $P = 1024$, and $K = 2$ sources are located at $\theta_1 = 4.93°$ and $\theta_2 = 10.01°$. Since the number of snapshots is too small to get a good estimate of the subspaces, ESPRIT is not able to estimate the DoAs. All other methods exhibit a similar performance, which is better than the one of ESPRIT, and achieve an RMSE of at most $0.5°$ if the number of antenna elements is 40 or larger.

V. Computational Complexity

If the matrix of measurements is compressed by the ℓ_1-SVD, the DoAs can be estimated using BPDN with a computational complexity of $\mathcal{O}\left(K^3P^3\right)$ [4]. This is larger than the computational costs $\mathcal{O}\left(M^2P + M^2N\right)$ and $\mathcal{O}\left(M^3 + M^2N\right)$ of the subspace-based methods MUSIC and ESPRIT, respectively, whose subspace estimation using the eigenvalue decomposition of the covariance matrix has a computational complexity of $\mathcal{O}\left(M^3\right)$. We found out that a few iterations of CBIHT and CBIHT-ℓ_2 are sufficient for DoA estimation and the number of iterations is empirically almost independent of the problem size, which results in a computational complexity of $\mathcal{O}\left(MNP\right)$. In addition, many multiplications to be performed in CBIHT and CBIHT-ℓ_2 can be implemented as simple sign flips and bit shifts due to the fact that the 1-bit quantization of the measurements is explicitly taken into account. Table I lists the computational complexity of all considered methods. If the number of snapshots N is small and the number of antenna elements M becomes very large, the subspace estimation becomes computationally intractable such that the computational complexity of CBIHT and CBIHT-ℓ_2 is smaller than the one of the subspace-based methods.

TABLE I
Computational Complexity.

BPDN	$\mathcal{O}\left(K^3P^3\right)$
CBIHT(-ℓ_2)	$\mathcal{O}\left(MNP\right)$
MUSIC	$\mathcal{O}\left(M^2P + M^2N\right)$
ESPRIT	$\mathcal{O}\left(M^3 + M^2N\right)$

VI. Conclusion

In this paper, we examined the DoA estimation from 1-bit measurements of many antenna elements. After introducing a measurement model and formulating the 1-bit DoA estimation as a sparse recovery problem, the methods BPDN and BIHT from the area of CS for solving sparse recovery problems were described. Since BIHT was introduced for real-valued signals and an SMV [8], we extended it to complex-valued signals and MMV such that the resulting CBIHT algorithm is applicable to 1-bit DoA estimation with multiple snapshots. The presented simulation results demonstrate that the 1-bit DoA estimation performance of CBIHT is comparable to the one of BPDN as well as the subspace-based methods MUSIC and ESPRIT or even better in a scenario with a large number of antenna elements and a small number of snapshots. Through the comparison of the computational complexity, we have shown that CBIHT has a smaller computational complexity than the other methods. Therefore, CBIHT is well suited for massive MIMO with many antenna elements if only a few snapshots are available.

References

[1] Nokia Solutions and Networks, "Technology Vision 2020," White Paper, June 2013.

[2] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[3] C. Svensson, S. Andersson, and P. Bogner, "On the power consumption of analog to digital converters," in *Norchip Conference, 2006. 24th*, Nov 2006, pp. 49–52.

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

[5] L. Carin, D. Liu, and B. Guo, "Coherence, compressive sensing, and random sensor arrays," *Antennas and Propagation Magazine, IEEE*, vol. 53, no. 4, pp. 28–39, Aug 2011.

[6] Y. C. Eldar and G. Kutyniok, *Compressed Sensing Theory and Applications*. Cambridge Univ. Press, 2012.

[7] E. van den Berg and M. P. Friedlander, "Sparse optimization with least-squares constraints," *SIAM J. Optimization*, vol. 21, no. 4, pp. 1201–1229, 2011.

[8] L. Jacques, J. Laska, P. Boufounos, and R. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *Information Theory, IEEE Transactions on*, vol. 59, no. 4, pp. 2082–2102, April 2013.

[9] X. Zeng and M. Figueiredo, "Robust binary fused compressive sensing using adaptive outlier pursuit," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 7674–7678.