Power Allocation for Distributed Compressive Sensing with 1-Bit Quantization over Noisy Channels

Jiguang He†, Markus Leinonen†, Kien-Giang Nguyen†, Yong Li‡, Olli Silvén⋆, Markku Juntti†
†Centre for Wireless Communications, FI-90014, University of Oulu, Finland
‡College of Computer Science, Chongqing University, Chongqing 400044, China
⋆Center for Machine Vision and Signal Analysis (CMVS), FI-90014, University of Oulu, Finland
Email: {jiguang.he, markus.leinonen, giang.nguyen, olli.silven, markku.juntti}@oulu.fi and {yongli}@cqu.edu.cn

***Abstract*—Cost-efficient implementation with a low-complexity analog-to-digital converter is necessary for the sensor nodes in the internet of things. In the paper, we study the distributed compressive sensing (DCS) under the constraint of 1-bit quantization at each node. The entire transmission chain, composed of compressive sensing, 1-bit quantization, and joint source-channel coding (JSCC), is taken into consideration with joint signal reconstruction at the fusion center. A lower bound on the end-to-end mean square error distortion, which is a function of the measurement rate, distortion of 1-bit quantization, and that of JSCC, is derived under the assumption of the oracle reconstruction. The time-varying channel conditions have a major impact on the distortion of JSCC. Therefore, a suboptimal yet efficient power allocation scheme based on the successive convex approximation method is proposed to minimize the lower bound on the end-to-end distortion. Moreover, a practical coding and joint signal reconstruction scheme is provided to show its consistence with the derived theoretical limits.**

***Index Terms*—Distributed compressive sensing, Internet of things, joint sparsity model, rate distortion**

I. Introduction

The internet of things (IoT) is widely recognized as the next Industrial Revolution, aiming at connecting everything. In the IoT networks, a group of sensor nodes are deployed to sense and measure the same physical phenomenon. Due to the limited geometrical distance, spatial correlation usually exists between the sensed data from different sensor nodes. The data is also smooth or piece-wise smooth over time. Therefore, temporal correlation also exists within each time series. This motivates the research field of (distributed) compressive sensing (CS) [1], assuming that the sensor data has a sparse representation in a certain transform domain.

The research on distributed compressive sensing (DCS) mainly falls into two categories: 1) design of joint signal reconstruction algorithms and 2) information-theoretic analysis. Different joint signal reconstruction algorithms have their own advantages and drawbacks. For instance, the simultaneous orthogonal matching pursuit (SOMP) is simple to implement, but has unsatisfactory performance [2]. Bayesian algorithms are more complicated, but outperform SOMP in terms of reconstruction accuracy [3]. Approximate message passing (AMP) algorithms have a problem in finding the optimal soft-thresholding value during each iteration [4]. As for the information-theoretic analysis of the DCS, Baron et al. [5] studied a DCS setup with noiseless measurements and channel. Coluccia *et al.* [6] utilized source coding with side information, which is not exactly suitable for the DCS setup with recovering both users' data. Empirical and theoretical performance of quantized DCS were reported in, e.g., [7], [8].

In this paper, we focus on the framework of DCS with 1-bit scalar quantization, along with performing joint source-channel coding (JSCC) of the quantizer outputs to transmit them over noisy channels.[1] A lower bound on the end-to-end mean square error (MSE) distortion is derived under the assumption of the oracle reconstruction (i.e., the support of each signal is assumed known at the receiver), which is used, e.g., in [6], [7]. To the best of our knowledge, such a bound has not been derived earlier. In the proposed structure, each CS-based sensor can access the information source only via 1-bit quantized noisy compressive measurements, followed by joint source-channel coding (JSCC). The output of JSCC is transmitted over Rayleigh block fading channels to the fusion center. The Berger-Tung outer bound is used to characterize the distortion in the JSCC module. The DCS joint signal reconstruction is assumed to be an oracle estimator in the minimum mean square error (MMSE) sense. A suboptimal yet efficient power allocation scheme based on a successive convex approximation (SCA) approach is proposed with the objective of minimizing the lower bound on the end-to-end distortion. With the introduction of the power allocation, better performance in terms of reconstruction accuracy can be achieved compared to its equal power allocation counterpart. Practical results are produced by using a low-complexity channel code, e.g., accumulator (ACC)-aided turbo code, as JSCC, and SOMP[2] as the joint signal reconstruction algorithm.

This work has been performed in the framework of the IIoT Connectivity for Mechanical Systems (ICONICAL), funded by the Academy of Finland. This work is also partially supported by the Academy of Finland 6Genesis Flagship (grant 318927). The work of M. Leinonen has also been financially supported by Infotech Oulu, and the work of Y. Li is supported by China NSF under Grant No. 61771081.

[1] 1-bit CS is beneficial because 1) for a given total bit-budget, its performance is comparable to multi-bit alternatives, and 2) the quantizer, i.e., an analog-to-digital converter (ADC) becomes inexpensive [9].

[2] We choose the simplest methods in the literature to verify the consistence between the practical simulations and theoretical results.

II. Preliminaries

A. Compressive Sensing

CS is a promising technique for reconstructing sparse signals from limited number of measurements (far less than the dimension of the signal) [10]. The source signal $\mathbf{s} \in \mathbb{R}^{N\times 1}$ is compressively sampled by a measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ with $M < N$ and corrupted by a measurement noise $\mathbf{z} \in \mathbb{R}^{M\times 1}$. The measurement vector $\mathbf{y} \in \mathbb{R}^{M\times 1}$ can be expressed as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{s} + \mathbf{z}. \tag{1}$$

The source signal $\mathbf{s} = \mathbf{\Psi}\mathbf{u}$ is assumed to have a sparse representation in a transform domain, where $\mathbf{\Psi} \in \mathbb{R}^{N\times N}$ is usually an orthonormal matrix, i.e., $\mathbf{\Psi}^T\mathbf{\Psi} = \mathbf{I}_N$, with $()^T$ denoting the transpose and $\mathbf{I}_N$ being the identity matrix with dimension $N \times N$, and $\mathbf{u}$ is the sparse transform coefficient vector. The support of $\mathbf{u}$ is expressed as $\Omega_{\mathbf{u}} = \{i|u_i \neq 0\}$, and the cardinality of $\Omega_{\mathbf{u}}$ is K, i.e., $|\Omega_{\mathbf{u}}| = K$. One of the well-known properties of $\mathbf{\Phi}$ is mutual coherence [11], i.e.,

$$\mu \triangleq \max_{i\neq j} |\boldsymbol{\phi}_i^T \boldsymbol{\phi}_j|, \tag{2}$$

where $\boldsymbol{\phi}_i$ is the i-th column of $\mathbf{\Phi}$. Under the assumption that each column of $\mathbf{\Phi}$ is normalized to 1, we have $0 \leq \mu \leq 1$. The smaller the value of μ, the better the measurement matrix.

B. Distributed Compressive Sensing

1) Joint Sparsity Model: In the literature, there are three classic joint sparsity models (JSMs) [5] to characterize the relationship among the correlated sparse signals. JSM-1 supposes that the signals have a common sparse component but different sparse innovations. JSM-2 assumes that the sparse signals have the same support but different non-zero elements (correlation may also exist for the non-zero elements). JSM-3 supposes that the signals have a common non-sparse component but different sparse innovations.

2) Joint Signal Reconstruction: Unlike in reconstructing a single signal, the signals can be reconstructed efficiently in a joint fashion if there exists certain correlation among the signals, e.g., in the support and/or values. Namely, leveraging the inter-signal correlation is expected to improve the reconstruction performance. The joint signal reconstruction algorithms fall into four major categories, i.e., simultaneous OMP [2], Bayesian learning [3], AMP [4], and deep learning [12].

III. System Model

The IoT network is divided into different disjoint clusters with different tasks. Each cluster has an access point (AP), acting as a fusion center that collects the data from different sensor nodes, processes the data, and sends it to the remote center processing unit (CPU). Fig. 1 illustrates an IoT network with three clusters. The functionality of each sensor node is to sense, quantize, encode, and transmit the sensed data to the AP within the cluster via a wireless link.

To be specifical, we focus on a two-sensor DCS network, like Cluster 1 in Fig. 1, and assume JSM-2 signals. The block diagram of the entire DCS network is detailed in Fig. 2. We

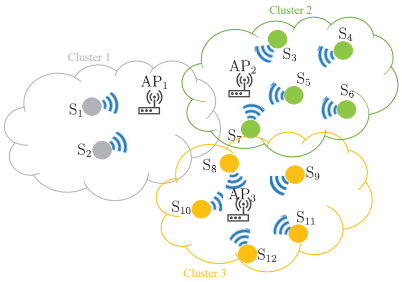

Fig. 1: An example of an IoT network.

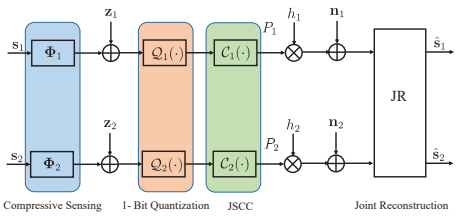

Fig. 2: The entire signaling chain at the sensor nodes and joint signal reconstruction at the fusion center.

assume that each non-zero element of the sparse coefficient vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ follows the independent and identically distributed (i.i.d.) Gaussian $\mathcal{N}(0,1)$. The relation between $\mathbf{u}_1$ and $\mathbf{u}_2$ is formulated by the first-order Gauss Markov process as

$$\mathbf{u}_{2,\Omega_{\mathbf{u}_2}} = \rho\mathbf{u}_{1,\Omega_{\mathbf{u}_1}} + \sqrt{1-\rho^2}\mathbf{i}, \tag{3}$$

$$\mathbf{u}_{2,\Omega\backslash\Omega_{\mathbf{u}_2}} = \mathbf{u}_{1,\Omega\backslash\Omega_{\mathbf{u}_1}} = \mathbf{0}, \tag{4}$$

where $\Omega = \{1, 2, \cdots, N\}$, $\Omega_{\mathbf{u}_1} = \Omega_{\mathbf{u}_2}$, $|\Omega_{\mathbf{u}_1}| = |\Omega_{\mathbf{u}_2}| = K$, $\mathbf{u}_{i,\Omega_{\mathbf{u}_i}}$ is a vector consisting of all the non-zero elements in $\mathbf{u}_i$, for $i = 1, 2$, $\Omega\backslash\Omega_{\mathbf{u}_i}$ is the complement of set $\Omega_{\mathbf{u}_i}$, $\mathbf{0}$ is an all-zero vector, ρ is the correlation coefficient determining the correlation level between $\mathbf{u}_{1,\Omega_{\mathbf{u}_1}}$ and $\mathbf{u}_{2,\Omega_{\mathbf{u}_2}}$, and $\mathbf{i} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_K)$. The source signals $\mathbf{s}_1$ and $\mathbf{s}_2$ can be expressed as

$$\mathbf{s}_i = \mathbf{\Psi}_i\mathbf{u}_i, \text{ for } i = 1, 2. \tag{5}$$

Without loss of generality, we assume that $\mathbf{\Psi}_1 = \mathbf{\Psi}_2 = \mathbf{\Psi}$.

The source signals are separately sensed through a measurement matrix as

$$\mathbf{y}_i = \mathbf{\Phi}_i\mathbf{s}_i + \mathbf{z}_i, \text{ for } i = 1, 2, \tag{6}$$

where each element of $\mathbf{\Phi}_i \in \mathbb{R}^{M_i\times N}$ follows $\mathcal{N}(0, \frac{1}{M_i})$ and $\mathbf{z}_i \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\mathbf{z}_i}\mathbf{I}_{M_i})$. We assume that $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ are fixed and known at the decoder.

The 1-bit scalar quantizer in Fig. 2 is described as

$$\mathcal{Q}_i : \mathbb{R} \to \mathbb{B}, \tag{7}$$

where $\mathbb{B} = \{0, 1\}$. For a realization of $y_{i,j}$ (with $y_{i,j}$ denoting the j-th element of $\mathbf{y}_i$), the quantizer operates as follows:

$$\mathcal{Q}_i(y_{i,j}) = \begin{cases} 1, & \text{if } \ y_{i,j} \geq 0, \\ 0, & \text{if } \ y_{i,j} < 0. \end{cases} \tag{8}$$

The code points of $\mathcal{Q}_i$ are set as the points that minimize the MSE distortion $E[|\tilde{y}_{i,j} - y_{i,j}|^2]$, where $\tilde{y}_{i,j}$ is given as

$\sqrt{2/\pi}\sigma_{\mathbf{y}_i}$ for region $y_{i,j} \geq 0$ and $-\sqrt{2/\pi}\sigma_{\mathbf{y}_i}$ for region $y_{i,j} < 0$ for $i = 1, 2$, $j = 1, \cdots, M_i$, and

$$\sigma_{\mathbf{y}_i} = \sqrt{K/M_i + \sigma^2_{\mathbf{z}_i}}, \tag{9}$$

which follows from the Central Limit Theorem (CLT) with $N \to \infty$ and $\lim_{N\to\infty} K/N$ being constant [6].

The output of the quantizer is further encoded by a JSCC, defined as

$$\mathcal{C}_i : \mathbb{B}^{M_i} \to \mathbb{B}^{J_i}, \tag{10}$$

with the joint transmission rate $R_{\mathrm{J},i} = \frac{M_i}{J_i}$. Each encoder output has a transmit power constraint P_i. We assume that the encoded data sequences are transmitted to the fusion center over orthogonal i.i.d. Rayleigh block fading channels. The channel coefficients are denoted as h_1 and h_2, which follow $\mathcal{CN}(0, 1)$.

At the fusion center, joint signal reconstruction is operated after receiving the encoded signals from the sensor nodes. Joint source-channel decoding (JSCD) is carried out by utilizing the correlation information of quantizer outputs,[3] followed by dequantization and joint signal reconstruction of the DCS part with final outputs $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$. The end-to-end distortion is defined as

$$D_{\mathrm{End}} = \sum_{i=1}^{2} D_{\mathrm{End},i} = \sum_{i=1}^{2} E\{||\mathbf{s}_i - \hat{\mathbf{s}}_i||_2^2\}. \tag{11}$$

IV. End-to-End Distortion Analysis and Power Allocation

A. End-to-End Distortion Analysis

The end-to-end distortion in (11) is a function of the measurement rate ($R_{\mathrm{M},i} = \frac{M_i}{N}$), quantization rate ($R_{\mathrm{Q},i} = 1$), joint transmission rate ($R_{\mathrm{J},i} = M_i/J_i$), transmit powers ($P_i$), and instantaneous channel coefficients (h_i), i.e.,

$$\begin{aligned} D_{\mathrm{End},i} = f_i(&R_{\mathrm{M},1}, R_{\mathrm{M},2}, R_{\mathrm{Q},1}, R_{\mathrm{Q},2}, ... \\ &R_{\mathrm{J},1}, R_{\mathrm{J},2}, P_1, P_2, h_1, h_2), \text{ for } i = 1, 2. \end{aligned} \tag{12}$$

Since a closed-form expression for $D_{\mathrm{End},i}$ seems technically infeasible, we next derive a lower bound on the end-to-end distortion under the assumption of the oracle DCS reconstruction.

The output sequences of the 1-bit quantizers, denoted by $\mathbf{q}_1$ and $\mathbf{q}_2$, are correlated (even though each of them is uniformly distributed), and the correlation depends on the value of ρ in (3). $\mathbf{q}_1$ and $\mathbf{q}_2$ can be modeled by the bit-flipping model as

$$q_{1,j} = q_{2,j} \oplus e_j, \text{ for } j = 1, 2, \cdots, \min\{M_1, M_2\}, \tag{13}$$

where e_j is an i.i.d. binary variable with $\Pr(e_j = 1) = p = f(\rho)$. The larger the value of ρ, the smaller the value of p. The calculation of p is based on the joint probability of $y_{1,j}$ and $y_{2,j}$, i.e.,

$$p = 2\Pr(y_{1,j} > 0, y_{2,j} < 0) = \int_0^{+\infty}\int_{-\infty}^{0} f(y_{1,j}, y_{2,j}) dy_{1,j} dy_{2,j}, \tag{14}$$

where

$$f(y_{1,j}, y_{2,j}) = \frac{\exp\{-1/2[y_{1,j}\ y_{2,j}]\mathbf{\Sigma}^{-1}[y_{1,j}\ y_{2,j}]^T\}}{\sqrt{(2\pi)^2|\mathbf{\Sigma}|}}, \tag{15}$$

and $\mathbf{\Sigma}$ is the covariance matrix for $[y_{1,j}\ y_{2,j}]$, in the form of

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2_{\mathbf{y}_1} & \rho\sigma_{\mathbf{y}_1}\sigma_{\mathbf{y}_2} \\ \rho\sigma_{\mathbf{y}_1}\sigma_{\mathbf{y}_2} & \sigma^2_{\mathbf{y}_2} \end{pmatrix}. \tag{16}$$

Due to the assumption of orthogonal transmissions, distortions $D_{\mathrm{J},1}$ and $D_{\mathrm{J},2}$ in the JSCC are achievable if the following inequalities hold:

$$R(D_{\mathrm{J},1})R_{\mathrm{J},1} \leq C(\gamma_1), \tag{17}$$

$$R(D_{\mathrm{J},2})R_{\mathrm{J},2} \leq C(\gamma_2), \tag{18}$$

where $R(\cdot)$ is the rate distortion function for the binary input $\mathbf{q}_1$ and $\mathbf{q}_2$ with Hamming distortion, i.e., $R(D) = 1 - H_b(D)$ with $H_b(x) = -x\log(x) - (1-x)\log(1-x)$, $0 \leq x \leq 1$, $C(\gamma_i) = \log(1+\gamma_i)$ is the Gaussian channel capacity by assuming the additive noise in the received signal has unit variance with $\gamma_i = P_i|h_i|^2$.

The Berger-Tung outer bound is further utilized to characterize the rate distortion region for the distributed multiterminal lossy source coding (binary source case), defined as [14]

$$R(D_{\mathrm{J},1}) \geq H_b[2p * H_b^{-1}(1 - R(D_{\mathrm{J},2}))] - H_b(D_{\mathrm{J},1}), \tag{19}$$

$$R(D_{\mathrm{J},2}) \geq H_b[2p * H_b^{-1}(1 - R(D_{\mathrm{J},1}))] - H_b(D_{\mathrm{J},2}), \tag{20}$$

$$R(D_{\mathrm{J},1}) + R(D_{\mathrm{J},2}) \geq 1 + H_b(2p) - H_b(D_{\mathrm{J},1}) - H_b(D_{\mathrm{J},2}), \tag{21}$$

where $a * b = a(1-b) + (1-a)b$ and $H_b^{-1}(\cdot)$ is the inverse function of $H_b(\cdot)$.

The aggregate distortion of 1-bit quantization and JSCC can be written as

$$\begin{aligned} D_{\mathrm{Q},i} = &D_{\mathrm{J},i}\int_{-\infty}^{0}(y_i - \sqrt{\frac{2}{\pi}}\sigma_{\mathbf{y}_i})^2 f(y_i)dy_i \\ &+ D_{\mathrm{J},i}\int_{0}^{\infty}(y_i + \sqrt{\frac{2}{\pi}}\sigma_{\mathbf{y}_i})^2 f(y_i)dy_i \\ &+ (1 - D_{\mathrm{J},i})\int_{-\infty}^{0}(y_i + \sqrt{\frac{2}{\pi}}\sigma_{\mathbf{y}_i})^2 f(y_i)dy_i \\ &+ (1 - D_{\mathrm{J},i})\int_{0}^{\infty}(y_i - \sqrt{\frac{2}{\pi}}\sigma_{\mathbf{y}_i})^2 f(y_i)dy_i \\ = &(1 + \frac{2}{\pi})\sigma^2_{\mathbf{y}_i} + (4D_{\mathrm{J},i} - 2)\alpha_i, \end{aligned} \tag{22}$$

where $\alpha_i = \int_0^{\infty} 2\sqrt{\frac{2}{\pi}}\sigma_{\mathbf{y}_i} y_i f(y_i)dy_i$ and $f(y_i)$ is the probability density function for each element in $\mathbf{y}_i$.

We set the aggregate MSE distortion in (22) as the variance of an additive Gaussian random variable. Accordingly, the relationship between the dequantizer output $\hat{y}_{i,j}$ and the quantizer input $y_{i,j}$ is expressed by $\hat{y}_{i,j} = y_{i,j} + \hat{n}_{i,j}$, where $\hat{n}_{i,j} \sim \mathcal{N}(0, D_{\mathrm{Q},i})$.

For the joint signal reconstruction, we assume that the support information is perfectly known at the fusion center. Therefore, the end-to-end distortion is lower bounded as

$$D_{\mathrm{End},i} \geq \sum_{k=1}^{K} \frac{\lambda_k^2(\sigma^2_{\mathbf{z}_i} + D_{\mathrm{Q},i})}{(\lambda_k^2 + \sigma^2_{\mathbf{z}_i} + D_{\mathrm{Q},i})^2}, \text{ for } i = 1, 2. \tag{23}$$

[3]The correlation information can be estimated during the decoding process [13].

under the MMSE estimation, where λ_k^2, for $k = 1, \cdots, K$, are the eigenvalues of the matrix $\mathbf{\Phi}_{\Omega_{\mathbf{u}_i}}^T \mathbf{\Phi}_{\Omega_{\mathbf{u}_i}}$.

B. Minimization of Lower Bound on End-to-End Distortion via Power Allocation

The objective is to minimize the lower bound on the end-to-end distortion in (23) subject to the constrains in (17), (18), (19)–(21) and a total power constraint $P_1 + P_2 \leq P$. Thus, the optimization problem is formulated as follows:

$$\min_{\{P_i, D_{\mathrm{J},i}\}} \quad \sum_{i=1}^{2}\sum_{k=1}^{K} \frac{\lambda_k^2(\sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i})}{\left(\lambda_k^2 + \sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i}\right)^2} \tag{24a}$$

subject to

$$H_b[2p * H_b^{-1}(1 - \frac{C(\gamma_{\bar{i}})}{R_{\mathrm{J},\bar{i}}})] - H_b(D_{\mathrm{J},i}) \leq \frac{C(\gamma_i)}{R_{\mathrm{J},i}}, \tag{24b}$$

$$1 + H_b(2p) - \sum_{i=1}^{2} H_b(D_{\mathrm{J},i}) \leq \sum_{i=1}^{2} \frac{C(\gamma_i)}{R_{\mathrm{J},i}}, \tag{24c}$$

$$D_{\mathrm{J},i} \in [0, 0.5], \sum_{i=1}^{2} P_i \leq P, i \in \{1, 2\}, \bar{i} \in \{1, 2\}\backslash\{i\}. \tag{24d}$$

Our first observation is that problem (24) is intractable to solve. To be specific, $\frac{\lambda_k^2(\sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i})}{\left(\lambda_k^2 + \sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i}\right)^2}$ is a quasi-concave function w.r.t. $D_{\mathrm{J},i} \geq 0$ which can easily be checked by second-order conditions [15, Sect. 3.4]. Thus, the convexity of the objective is difficult to prove [16]. In addition, function $H_b[2p * H_b^{-1}(1 - \frac{C(\gamma_{\bar{i}})}{R_{\mathrm{J},\bar{i}}})]$ in constraint (24b) is not straightforward to handle. We propose a suboptimal but efficient algorithm to solve (24). To this end, we define a new set of variables, i.e., $v_i \triangleq 1 - \frac{C(\gamma_i)}{R_{\mathrm{J},i}}$ and $w_i \triangleq H_b^{-1}(v_i)$. Then, by noting that $P_i = \frac{2^{2R_{\mathrm{J},i}(1-v_i)} - 1}{|h_i|^2}$, $v_i = H_b(w_i)$, and following the epigraph transformation [15, Sect. 3.1.7], we can equivalently rewrite problem (25) in a more tractable form as

$$\min_{\{t_{i,k}, w_i, r_i, D_{\mathrm{J},i}\}} \quad \sum_{i=1}^{2}\sum_{k=1}^{K} t_{i,k} \tag{25a}$$

subject to

$$2 + \frac{\lambda_k^2}{\sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i}} + \frac{\sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i}}{\lambda_k^2} \geq \frac{1}{t_{i,k}} \tag{25b}$$

$$H_b(2p - 4pw_{\bar{i}} + w_{\bar{i}}) + H_b(w_i) \leq 1 + H_b(D_{\mathrm{J},i}), \tag{25c}$$

$$1 + H_b(2p) + \sum_{i=1}^{2} H_b(w_i) \leq 2 + \sum_{i=1}^{2} H_b(D_{\mathrm{J},i}) \tag{25d}$$

$$\sum_{i=1}^{2} \frac{2^{r_i} - 1}{|h_i|^2} \leq P, \ \frac{r_i}{2R_{\mathrm{J},i}} + H_b(w_i) \geq 1 \tag{25e}$$

$$D_{\mathrm{J},i} \in [0, 0.5], \ i \in \{1, 2\}, \bar{i} \in \{1, 2\}\backslash\{i\}, \tag{25f}$$

where $\{t_{i,k}\}$ and $\{r_i\}$ are newly introduced slack variables. Note that $H_b(x)$ is a concave function. Thus, the nonconvexity of problem (25) can easily be verified from the left-side of constraints (25b)–(25d). Interestingly, problem (25) is an SCA-applicable formulation [17]. In the light of the SCA principle, we can convexify (25) as

$$\min_{\{t_{i,k}, w_i, r_i, D_{\mathrm{J},i}\}} \quad \sum_{i=1}^{2}\sum_{k=1}^{K} t_{i,k} \tag{26a}$$

subject to

$$2 + \tilde{f}^{(n)}(D_{\mathrm{J},i}) + \frac{\sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i}}{\lambda_k^2} \geq \frac{1}{t_{i,k}} \tag{26b}$$

$$\tilde{H}_b^{(n)}(2p - 4pw_{\bar{i}} + w_{\bar{i}}) + \tilde{H}_b^{(n)}(w_i) \leq 1 + H_b(D_{\mathrm{J},i}), \tag{26c}$$

$$H_b(2p) + \sum_{i=1}^{2} \tilde{H}_b^{(n)}(w_i) \leq 1 + \sum_{i=1}^{2} H_b(D_{\mathrm{J},i}) \tag{26d}$$

$$(25e), (25f), \tag{26e}$$

where $\tilde{f}^{(n)}(D_{\mathrm{J},i})$ and $\tilde{H}_b^{(n)}(.)$ are the first-order approximations of $\frac{\lambda_k^2}{\sigma_{\mathbf{z}_i}^2 + D_{\mathrm{Q},i}}$ and $H_b(.)$, respectively, at some point $\{t_{i,k}^{(n)}, w_i^{(n)}, r_i^{(n)}, D_{\mathrm{J},i}^{(n)}\}$ in the feasible set. We have

$$\tilde{f}^{(n)}(x) \triangleq \frac{\lambda_k^2}{\sigma_{\mathbf{z}_i}^2 + x^{(n)}} - \frac{\lambda_k^2}{(\sigma_{\mathbf{z}_i}^2 + x^{(n)})^2}(x - x^{(n)}),$$

$$\tilde{H}_b^{(n)}(x) \triangleq -\frac{\log(1 - x^{(n)}) + x(\log(x^{(n)}) + \log(1 - x^{(n)}))}{\ln 2}.$$

Here, the superscript n denotes the n-th iteration of the SCA algorithm. Problem (26) is solved at step 3 of the iterative method outlined in Algorithm 1. The optimal solution for (26) is used to form the problem to be solved at the next iteration. The iterative procedure of Algorithm 1 is guaranteed to converge by following the same arguments as those in [18].

Algorithm 1 SCA for solving (25)

1: **Initialization:** Set $n \coloneqq 0$, find an initial feasible point $\{w_i^{(0)}, D_{\mathrm{J},i}^{(0)}\}$
2: **repeat**
3: Solve (26) and obtain optimal value $\{w_i^*, D_{\mathrm{J},i}^*\}$
4: Update $\{w_i^{(n+1)}, D_{\mathrm{J},i}^{(n+1)}\} \coloneqq \{w_i^*, D_{\mathrm{J},i}^*\}$
5: Update $n \coloneqq n + 1$
6: **until** Convergence
7: **Output** $\{t_{i,k}^{(n)}, w_i^{(n)}, r_i^{(n)}, D_{\mathrm{J},i}^{(n)}\}$

V. Numerical Results

In this section, we evaluate the theoretical end-to-end distortion and practical implementation for the DCS under 1-bit quantization at each sensor node. The ACC-aided turbo code with polynomial $G = ([3, 2])_8$ is taken as a JSCC [13] and SOMP [2] is applied as the DCS joint signal reconstruction scheme for practical implementation. In all the experiments, we set $M_1 = M_2 = 50$, $N = 60$, $R_{\mathrm{J},1} = R_{\mathrm{J},2} = 0.5$, $\rho = 0.95$, and $K = 4$. In the derivation of the theoretical lower bound of the end-to-end distortion, we assume that $\boldsymbol{\lambda}^2 = [1.3 \ \ 1.0 \ \ 0.8 \ \ 0.6]$ is perfectly known for a given measurement matrix.[4]

[4]In the practical implementation, we assume that the non-zero elements are in the first K positions of $\mathbf{u}_1$ and $\mathbf{u}_2$. The measurement matrices $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ are designed to be the same.

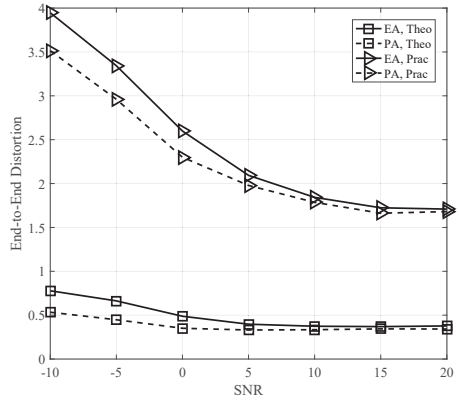

Fig. 3: End-to-end distortion for DCS with noiseless measurements, theoretical analysis vs. practical scheme.

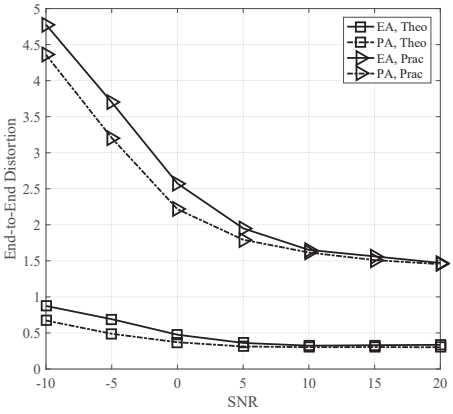

Fig. 4: End-to-end distortion for DCS with noisy measurements, theoretical analysis vs. practical scheme.

First, we consider the noiseless measurements, i.e., $\sigma^2_{\mathbf{z}_1} = \sigma^2_{\mathbf{z}_2} = 0$, resulting in $p = 0.101$ and $\alpha_1 = \alpha_2 = 0.0424$. The simulation results are shown in Fig. 3. The terms "PA" and "EA" in the legend denote power allocation and equal power allocation, respectively. The terms "Theo" and "Prac" in the legend denote theoretical results and practical results, respectively. The signal-to-noise ratio (SNR) is defined as $10\log_{10}(P_1 + P_2)$. Second, we set the measurement noise variances as $\sigma^2_{\mathbf{z}_1} = \sigma^2_{\mathbf{z}_2} = 0.01$, resulting in $p = 0.101$ and $\alpha_1 = \alpha_2 = 0.0446$. The simulation results are shown in Fig. 4. As shown in Figs. 3 and 4, better performance can be achieved for the proposed power allocation scheme, especially in the low SNR regime. The gap between the practical schemes and derived theoretical limits is reasonable due to three reasons: 1) The JSCC code with low complexity is not capacity-approaching, even in an additive white Gaussian noise (AWGN) channel. 2) Errors in the support detection further enlarge the gap. 3) The DCS joint signal reconstruction algorithm itself can not achieve the best performance in the MMSE sense.

VI. Conclusion

Theoretical end-to-end distortion has been analyzed for the distributed compressive sensing under the assumptions of 1-bit quantization and the oracle estimator. The correlation information of the quantizer outputs has been utilized in the JSCC. A suboptimal yet efficient power allocation scheme has been operated with the aim of minimizing the lower bound on the end-to-end distortion via the SCA approach. Practical coding and joint signal reconstruction has been performed to show its consistence with the derived theoretical distortion bounds. With the consideration of the low-resolution 1-bit quantization, the sampling method may violate the conventional concept of compressive sensing with $M_i > N$. Thus, a more complicated high-resolution quantization scheme could be considered for better complexity-performance trade-offs by carefully allocating the quantization rate, measurement rate, JSCC rate, etc.

References

[1] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb 2006.

[2] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572 – 588, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168405002227

[3] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun 2008.

[4] G. Hannak, A. Perelli, N. Goertz, G. Matz, and M. E. Davies, "Performance analysis of approximate message passing for distributed compressed sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 5, pp. 857–870, Oct 2018.

[5] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," Jan 2009. [Online]. Available: http://arxiv.org/abs/0901.3403

[6] G. Coluccia, A. Roumy, and E. Magli, "Operational rate-distortion performance of single-source and distributed compressed sensing," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 2022–2033, Jun 2014.

[7] A. Shirazinia, S. Chatterjee, and M. Skoglund, "Distributed quantization for measurement of correlated sparse sources over noisy channels," Jul 2015. [Online]. Available: http://arxiv.org/abs/1404.7640

[8] M. Leinonen, M. Codreanu, and M. Juntti, "Distributed distortion-rate optimized compressed sensing in wireless sensor networks," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1609–1623, Apr 2018.

[9] J. N. Laska and R. G. Baraniuk, "Regime change: Bit-depth versus measurement-rate in compressive sensing," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3496–3505, July 2012.

[10] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[11] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5030–5043, Oct 2010.

[12] H. Palangi, R. Ward, and L. Deng, "Distributed compressive sensing: A deep learning approach," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4504–4518, Sep. 2016.

[13] J. He, V. Tervo, S. Qian, Q. Xue, M. Juntti, and T. Matsumoto, "Performance analysis of lossy decode-and-forward for non-orthogonal MARCs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1545–1558, Mar. 2018.

[14] X. He, X. Zhou, P. Komulainen, M. Juntti, and T. Matsumoto, "A lower bound analysis of hamming distortion for a binary CEO problem with joint source-channel coding," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 343–353, Jan 2016.

[15] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[16] H. J. Greenberg and W. P. Pierskalla, "A review of quasi-convex functions," *Operations research*, vol. 19, no. 7, pp. 1553–1570, 1971.

[17] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, no. 1, pp. 29–51, 2010.

[18] K. Nguyen, L. Tran, O. Tervo, Q. Vu, and M. Juntti, "Achieving energy efficiency fairness in multicell MISO downlink," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1426–1429, Aug 2015.