Wireless Compressive Sensing for Energy Harvesting Sensor Nodes

Gang Yang, Vincent Y. F. Tan, Chin Keong Ho, See Ho Ting, and Yong Liang Guan

Abstract

We consider the scenario in which multiple sensors send spatially correlated data to a fusion center (FC) via independent Rayleigh-fading channels with additive noise. Assuming that the sensor data is sparse in some basis, we show that the recovery of this sparse signal can be formulated as a compressive sensing (CS) problem. To model the scenario in which the sensors operate with intermittently available energy that is harvested from the environment, we propose that each sensor transmits independently with some probability, and adapts the transmit power to its harvested energy. Due to the probabilistic transmissions, the elements of the equivalent sensing matrix are not Gaussian. Besides, since the sensors have different energy harvesting rates and different sensor-to-FC distances, the FC has different receive signal-to-noise ratios (SNRs) for each sensor. This is referred to as the *inhomogeneity* of SNRs. Thus, the elements of the sensing matrix are also not identically distributed. For this unconventional setting, we provide theoretical guarantees on the number of measurements for reliable and computationally efficient recovery, by showing that the sensing matrix satisfies the restricted isometry property (RIP), under reasonable conditions. We then compute an achievable system delay under an allowable mean-squared-error (MSE). Furthermore, using techniques from large deviations theory, we analyze the impact of inhomogeneity of SNRs on the so-called k-restricted eigenvalues, which governs the number of measurements required for the RIP to hold. We conclude that the number of measurements required for the RIP is not sensitive to the inhomogeneity of SNRs, when the number of sensors n is large and the sparsity of the sensor data (signal) k grows slower than the square root of n. Our analysis is corroborated by extensive numerical results.

Index Terms

Wireless compressive sensing, Energy harvesting, Restricted isometry property, Compressive sensing, Wireless sensor networks, Rayleigh-fading channels, Large deviations

 G. Yang, S. H. Ting and Y. L. Guan are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail:yang0305@e.ntu.edu.sg; {shting, eylguan}@ntu.edu.sg). G. Yang is supported in part by the Advanced Communications Research Program DSOCL06271, a research grant from the Defence Research and Technology Office (DRTech), Ministry of Defence, Singapore. This paper was presented in part at the International Conference on Communications (ICC), 2013.

V. Y. F. Tan and C. K. Ho are with the Institute for Infocomm Research, A*STAR, Singapore (e-mail: {tanyfv, hock}@i2r.a-star.edu.sg). V. Y. F. Tan is also with the Department of Electrical and Computer Engineering, National University of Singapore.

I. INTRODUCTION

The lifetimes of conventional wireless sensor networks (WSNs) are limited by the total energy available in the batteries. It is inconvenient to replace batteries periodically, or even impossible when sensors are deployed in harsh conditions, e.g., in toxic environments or inside human bodies. Energy harvesting of ambient energy such as solar, wind, thermal and piezoelectric energy, appears as a promising alternative to a fixed-energy battery, to prolong the lifetime and offer potentially maintenance-free operation for WSNs [1], [2]. Compared to limited but reliable power supply from conventional batteries, energy harvesters provide a virtually perpetual but unreliable energy source. Moreover, the sensors typically have different *energy harvesting rates*, due to varying harvesting conditions such as the spread of sunlight and difference in wind speeds.

This paper addresses the problem of data transmission in energy harvesting WSNs (EH-WSNs). We assume that energy harvesting sensors are deployed to monitor some physical phenomenon in space, e.g., temperature, toxicity of gas. Data collected from sensors are sent to the fusion center (FC). The data are typically correlated, and well approximated by a sparse vector in an appropriate transform (e.g., the Fourier transform). Recent developments in compressive sensing (CS) theory provide efficient methods to recover sparse signals from limited measurements [3]. CS theory states that if the sensing matrix satisfies the restricted isometry property (RIP), a small number of measurements (relative to the length of the data vector) is sufficient to accurately recover the sparse data. This advantage of CS potentially allows us to reduce the total number of transmissions, and this is particularly important for data transmission in bandwidth-limited wireless channels.

The accurate estimation of the sensor data by the FC has recently been addressed by using CS techniques in the literature. In [4], Haupt *et. al* presented a sensing scheme based on phase-coherent transmissions for all sensors. However, [4] made two practically limiting assumptions. First, it assumed that there was no channel fading, and path losses for all sensors were identical. Second, the transmissions from all sensors were synchronized such that signals arrived in phase at the FC. In [5], Aeron *et. al* derived information theoretic bounds on sensing capacity of sensor networks under a fixed signal-to-noise ratio (SNR) for all sensors. In contrast, [6] proposed a sparse approximation method in non-fading channels, which adapted a sensor's sensing activity according to its energy availability. In [7], Xue *et. al* successively applied CS in the spatial domain and the time domain, under a fixed SNR for all sensors. In [8], Fazel *et. al* proposed a random access scheme in underwater sensor networks. Each activated sensor picked a uniformly-distributed delay

to transmit. By simply discarding the colliding data packets from concurrent medium access, the FC used a CS decoder to recover the sensor data based only on the successfully received packets. Thus, the scheme did not exploit packet collisions for data recovery.

Since sensors are placed at different locations, it is commonly assumed that the sensors transmit data over independent but nonidentical channels with different fading conditions. Different energy harvesting rates also lead to different transmit powers and hence different (receive) SNRs. We refer to this generally as the *inhomogeneity* of SNRs. The proposed framework of *wireless compressive sensing* (WCS) that results in the scenario of inhomogeneous SNRs has, to the best of our knowledge, not been studied in the literature. We define the *system delay* as the number of concurrent sensor-to-FC transmissions (or channel uses) for estimating one data vector (among sensors). We aim to reduce the system delay, while ensuring a target estimation accuracy. Surprisingly, we observe that the required number of measurements for accurate recovery m is not overly sensitive to the inhomogeneity of SNRs provided that the number of sensors n is large and the sparsity of the data vector k grows slower than $\sqrt{n}$. This motivates us to further investigate the impact of inhomogeneity of SNRs, based on the recovery performance in terms of RIP.

The three main contributions are summarized as follows.

First, we propose *wireless compressive sensing*, which features probabilistic transmissions by sensor nodes and the over-the-air linear combination of the transmitted signals. In each time slot, sensor j decides to transmit with probability p_j, and adjusts the amplitude of the transmitted signal according to its energy availability. The FC thus receives a linear combination of wireless signals that are transmitted from a random subset of sensors. The transmissions over successive time slots result in a sensing matrix which is achieved through the mixing of signals in wireless channels.

Second, we prove that, for a given vector of transmit probabilities $\mathbf{p}$, the FC can recover the data accurately and efficiently via convex optimization, if the total number of transmissions (or measurements) m exceeds

$$O\left(\frac{k\rho_{\max}(k)}{p_{\min}^2\rho_{\min}(k)}\log\frac{n}{k}\right), \tag{1}$$

where n is the number of sensors, k is the sparsity of the sensor data, $p_{\min}$ is the minimum component of $\mathbf{p}$, and $\rho_{\max}(k)$ and $\rho_{\min}(k)$ are respectively the maximum and minimum k-restricted eigenvalues of a Gram matrix which depends on the inhomogeneity of SNRs. Different from previous works on CS, our bound depends on the ratio $\frac{\rho_{\max}(k)}{\rho_{\min}(k)}$. Based on this result, we also compute the achievable

system delay subject to a desired recovery accuracy.

Third, we analyze the impact of inhomogeneity of SNRs on the required number of measurements, in terms of the k-restricted eigenvalues. By using the theory of large deviations, we show that both the maximum and the minimum k-restricted eigenvalues concentrate around one (for all constant k) in the large n regime, and the rate of convergence to one depends on the inhomogeneity of SNRs. This allows us to explain the observation that the inhomogeneity of SNRs does not significantly affect the number of measurements required for the RIP to hold.

This remainder of this paper is organized as follows: Section II provides a description of the system model. Section III presents a new WCS scheme. Section IV details the main result on the RIP, the achievable system delay and investigates the impact of inhomogeneity of SNRs. Section V provides the simulation results. Section VI concludes this paper. The proofs for the RIP result and the result on the impact of inhomogeneity of SNRs are given in Appendix A and C.

II. System Model

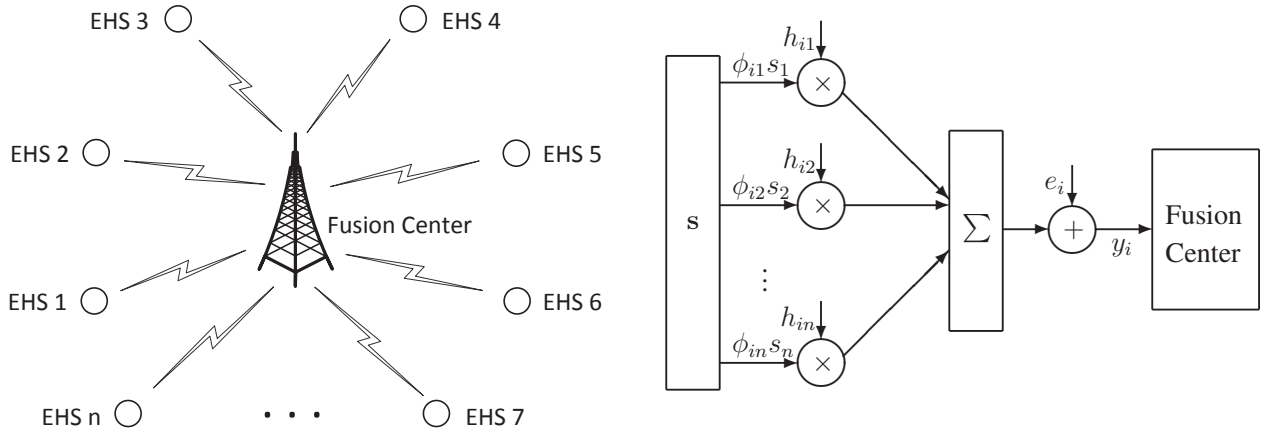
Consider a star-topology wireless sensor network (see Fig. 1a) that consists of n energy harvesting sensor nodes and a FC. Sensors transmit their data to the FC via a shared multiple-access channel (MAC). We consider slotted transmissions by first considering a single snapshot of the spatial-temporal field. Assuming the sensor data $\mathbf{s}$ is compressible, we can model it as being sparse with respect to (w.r.t.) a fixed orthonormal basis $\{\boldsymbol{\psi}_j \in \mathbb{C}^n : j = 1, \ldots, n\}$, i.e.,

$$\mathbf{s} = \boldsymbol{\Psi}\mathbf{x} = \sum_{j=1}^{n} \boldsymbol{\psi}_j x_j, \tag{2}$$

where $\mathbf{x} \in \mathbb{C}^n$ has at most $k < \lfloor n/2 \rfloor$ non-zero components and $\lfloor \cdot \rfloor$ is the floor operation.

We assume a flat-fading channel with complex-valued channel coefficients h_{ij}, where $1 \leq i \leq m$ denotes the slot index and $1 \leq j \leq n$ denotes the sensor index. The channel remains constant in each slot. We further assume a Rayleigh-fading channel, hence the channel coefficients for different slots are independent and identically distributed (i.i.d.) according to the complex Gaussian distribution.

We propose that sensors concurrently transmit to the FC in a probabilistic manner, such that the signals from sensors are linearly combined over the air. Sensor j multiplies its datum s_j by some



(a) A star-topology energy-harvesting WSN.

(b) The MAC structure for a WSN in the i-th time slot.

Fig. 1: System block diagram.

amplitude ϕ_{ij} (to be defined in (4)), then transmits in the i-th time slot[1]. The FC thus receives

$$y_i = \sum_{j=1}^{n} h_{ij}\phi_{ij}s_j + e_i,$$

where e_i is a noise term (not necessarily Gaussian). The signal model over one time slot is illustrated in Fig. 1b. After m time slots (i.e., one frame), the FC receives the measurement vector

$$\mathbf{y} = (\mathbf{H} \odot \mathbf{\Phi})\mathbf{s} + \mathbf{e} = \mathbf{Z}\mathbf{s} + \mathbf{e} = \mathbf{Z}\mathbf{\Psi}\mathbf{x} + \mathbf{e}, \tag{3}$$

where the matrix $\mathbf{Z} = \mathbf{H} \odot \mathbf{\Phi}$, and the operation $\odot$ is the element-wise product of two matrices. We assume all noise components are independent, zero mean and have variance σ^2.

From the perspective of signal recovery, we want to estimate $\mathbf{x}$ or equivalently $\mathbf{s}$, from $\mathbf{y}$, such that the *mean-squared-error* (MSE) $\mathbb{E}\|\widehat{\mathbf{x}} - \mathbf{x}\|_2^2$ does not exceed some threshold ϵ. Also, we would like to estimate the sparse vector using minimum network resources (i.e., channel uses), due to limited channel resources. Thus, given a fixed number of sensors n and an ϵ, our objective is to design a transmission scheme that minimizes the number of sensor-to-FC transmissions m.

Different from [6], [11], we consider Rayleigh-fading channels, and adopt concurrent transmissions in a probabilistic manner. Moreover, the SNRs of different sensors are considered to be different,

[1]Like in [4], [6], [9], we assume analog transmissions in WSNs and thus all quantities take continuous values. From an information-theoretic perspective [10], the analog/uncoded sensing or transmission scheme is better than the digital one, in terms of minimizing the distortion of reconstructed signal.

compared to the fixed SNR case in the literature [5], [7], [11].

III. Energy-Aware Wireless Compressive Sensing

In Section III-A, we first provide a CS perspective for the signal model in (3). Then in Section III-B, we present an *energy-aware wireless compressive sensing* (EAWCS) scheme. We also derive the probability distribution function (pdf) of elements in the random matrix $\mathbf{Z}$ in Section III-C, which will be used to show the RIP in Section IV-A.

A. A Compressive Sensing Perspective

Since we assume the data vector $\mathbf{x}$ is sparse in some basis, it seems natural to adopt a CS method to recover $\mathbf{x}$. The over-the-air combination via the channel matrix $\mathbf{H}$ contributes to the effective equivalent sensing matrix $\mathbf{Z}$ in (3). However, there are two differences from the conventional CS setup that make the analysis more challenging. First, due to probabilistic transmissions, the elements of the sensing matrix $\mathbf{Z}$ are not Gaussian. Second, since sensors have different energy harvesting rates and different sensor-to-FC distances, the FC has different receive SNRs for all sensors. Thus, the elements of the sensing matrix $\mathbf{Z}$ are also not identically distributed.

The proposed transmission scheme calls for the analysis of non-Gaussian non-i.i.d. sensing matrices. Hence, we need to analyze the system performance in a more intricate way that differs from conventional CS problems. The key technique we employ is to show that the elements of the sensing matrix $\mathbf{Z}$ are sub-Gaussian (see Definition 6 in Appendix A), and make use of new results on sub-Gaussian random matrices.

B. Energy-Aware Wireless Compressive Sensing

We consider only the energy consumption for wireless transmissions, by assuming the energy consumption on sensing is negligible. The energy harvesting rate varies over sensors. For simplicity, we assume that each sensor allocates the same power for all slots. Let E_j be the accumulated harvested energy that is available for sensor j to transmit in each slot. We perform energy-aware wireless transmissions taking into account the different available energy. It is noted that a *causal energy constraint* that comes from energy harvesting should be satisfied, i.e., energy that is consumed for transmissions can not exceed the energy available in each slot.

For sensor j, set a transmit probability $p_j \in (0, 1]$ and a squared-amplitude $b_j > 0$. Let $\mathbf{\Phi}$ in (3) be a *selection-and-weight* (SW) matrix, whose elements are independently generated according to[2]

$$\phi_{ij} = \begin{cases} +\sqrt{b_j} & \text{w.p. } p_j/2 \\ 0 & \text{w.p. } 1-p_j, \\ -\sqrt{b_j} & \text{w.p. } p_j/2 \end{cases} \quad \forall\, i = 1, 2, \ldots, m. \tag{4}$$

Given available energy E_j, we choose b_j such that

$$p_j b_j \leq E_j, \quad \forall j = 1, 2, \ldots, n. \tag{5}$$

Clearly, each entry ϕ_{ij} is zero mean and has variance $p_j b_j$. Both p_j and b_j can be chosen such that the causal energy constraint is satisfied in expectation, i.e., $\mathbb{E}(\phi_{ij}^2) = p_j b_j \leq E_j$. This allows us to save energy to be used for future transmissions. The energy-saving feature can be crucial in the scenario where the energy harvesting rates are fluctuating over several snapshots of the spatial-temporal field. Further discussion on the choices of p_j's and b_j's will be made in Remark 3, Remark 4, and the the paragraphs following Remark 4.

In [6], all sensors consume the same amount of energy for transmissions. In contrast, each sensor here may adopt the transmit power to its available energy via the above-designed SW matrix. Furthermore, the SW matrix randomly selects the sensors to transmit, and weighs the data according to the sensors' harvested energy. In each time slot, a subset of sensors are selected at random to perform transmissions and over-the-air combination. The selections are performed in a distributed manner at each sensor node, since each node separately decides the slots that it transmits in. We couple random sensor selection and energy-aware transmission by the choice of the SW matrix.

Recall the signal model in (3), i.e., $\mathbf{y} = \mathbf{Z}\mathbf{\Psi}\mathbf{x} + \mathbf{e}$. With the knowledge[3] of the matrix $\mathbf{Z}$ and the sparsity-inducing basis $\mathbf{\Psi}$, the FC can implement CS decoding to obtain the recovered sparse coefficients $\widehat{\mathbf{x}}$ and further estimate the data vector as $\widehat{\mathbf{s}} = \mathbf{\Psi}\widehat{\mathbf{x}}$. Since the EAWCS scheme adopts probabilistic and concurrent transmissions presented in Section II, its diagram is similar to Fig. 1b in which the FC performs CS decoding by exploiting the sparsity of the spatial sensor data.

[2]The sensor j can also directly transmit the value $\sqrt{b_j}$ with probability p_j, and keep silent with probability $1 - p_j$. Due to the symmetry of the distribution of channel coefficients, the distribution in (8) is the same, and thus all results in the sequel still hold.

[3]The information can be obtained by the FC via channel estimation and feedback. Alternatively, the SW matrix can be generated by using predefined random seeds.

C. Probability Distribution Analysis and Equivalent Normalized Signal Model

Consider the signal model in (3). Denote each element in $\mathbf{Z}$ as $Z_{ij} = h_{ij}\phi_{ij} = Z_{ij}^{\mathrm{R}} + jZ_{ij}^{\mathrm{I}}$, where $Z_{ij}^{\mathrm{R}} \triangleq h_{ij}^{\mathrm{R}}\phi_{ij}$, and $Z_{ij}^{\mathrm{I}} \triangleq h_{ij}^{\mathrm{I}}\phi_{ij}$. Note that elements of the matrix $\mathbf{H}$ are independent, and each element h_{ij} has independent real and imaginary parts. Also $\mathbf{\Phi}$ consists of independent elements. All elements of matrix $\mathbf{Z}$ are thus independent, and have independent real and imaginary components. As such, it suffices to analyze the probability distribution of the real component, since the analysis is similar for the imaginary component. The marginal pdf of Z_{ij}^{R} can be shown to be

$$f_{Z_{ij}^{\mathrm{R}}}(z) = \frac{1}{\sqrt{b_j}} f_{H_j^{\mathrm{R}}}\left(\frac{z}{\sqrt{b_j}}\right) \cdot \frac{p_j}{2} + \frac{1}{\sqrt{b_j}} f_{H_j^{\mathrm{R}}}\left(-\frac{z}{\sqrt{b_j}}\right) \cdot \frac{p_j}{2} + (1-p_j)\cdot\delta(z), \tag{6}$$

where $f_{H_j^{\mathrm{R}}}(\cdot)$ is the pdf of channel coefficient of sensor j, and $\delta(\cdot)$ is the Dirac delta function. For the sake of brevity, we define a new pdf as follows.

Definition 1. *A random variable X follows a* mixed Gaussian *distribution, denoted as $X \sim \widetilde{\mathcal{N}}(\mu, \nu^2, p)$, if its pdf has the following form*

$$f_X(x) = p\frac{1}{\sqrt{2\pi\nu^2}} \exp\left(-\frac{(x-\mu)^2}{2\nu^2}\right) + (1-p)\delta(x), \tag{7}$$

where $p \in (0, 1]$ is the mixing parameter. The corresponding complex mixed Gaussian distribution, assuming the real and imaginary components are independent, is denoted as $\widetilde{\mathcal{N}}_c(\mu, \nu^2, p)$.

Assuming Rayleigh-fading channels, all elements in the channel matrix $\mathbf{H}$ are independent, zero mean and follow Gaussian distributions. Note that due to different fading channels for the sensors, the matrix $\mathbf{H}$ has column-dependent variances, where the j-th column follows a Gaussian distribution with variances ν_j^2. From (6) and (7), the marginal pdf of Z_{ij}^{R} can be expressed as

$$f_{Z^{\mathrm{R}}}(z) = p_j \frac{1}{\sqrt{\pi\nu_j^2 b_j}} \exp\left(-\frac{z^2}{\nu_j^2 b_j}\right) + (1-p_j)\delta(z). \tag{8}$$

Thus, we have $Z^{\mathrm{R}} \sim \widetilde{\mathcal{N}}\Big(0, \nu_j^2 b_j/2, p_j\Big)$.

Recall that $\mathbf{Z} = \mathbf{H} \odot \mathbf{\Phi}$. Due to independent sensor-to-FC channels, the matrix $\mathbf{H}$ has column-wise variances described in the diagonal matrix $\mathbf{\Gamma}_{\mathrm{H}} = \mathrm{diag}\{\nu_1, \nu_2, \ldots, \nu_n\}$. Similarly, because of different energy harvesting rates, the matrix $\mathbf{\Phi}$ has column-wise variances described in the diagonal matrix $\mathbf{\Gamma}_{\Phi} = \mathrm{diag}\{\sqrt{p_1 b_1}, \sqrt{p_2 b_2}, \ldots, \sqrt{p_n b_n}\}$. We then have the decompositions $\mathbf{H} = \widetilde{\mathbf{H}}\mathbf{\Gamma}_{\mathrm{H}}$ and

$\mathbf{\Phi} = \widetilde{\mathbf{\Phi}}\mathbf{\Gamma}_{\Phi}$, such that elements of $\widetilde{\mathbf{H}}$ (and $\widetilde{\mathbf{\Phi}}$) are i.i.d. and with unit variance. For convenience, we then decompose the matrix $\mathbf{Z}$ as follows

$$\mathbf{Z} = \sqrt{m}\widetilde{\mathbf{Z}}\mathbf{\Gamma}, \tag{9}$$

where we denote $\widetilde{\mathbf{Z}} = \left((1/m)\widetilde{\mathbf{H}}\right) \odot \widetilde{\mathbf{\Phi}}$ and $\mathbf{\Gamma} = \mathbf{\Gamma}_{\mathrm{H}}\mathbf{\Gamma}_{\Phi}$. Let $\mathbf{\Gamma} = \mathrm{diag}\{\sqrt{\gamma_1}, \sqrt{\gamma_2}, \ldots, \sqrt{\gamma_n}\}$, where the receive signal power of sensor j is[4] $\gamma_j = p_j b_j \nu_j^2$. We term the diagonal elements of $\mathbf{\Gamma}$ as the *signal power pattern*. The γ_j's are generally all different (i.e., inhomogeneous), and this directly leads to the inhomogeneous (receive) SNRs. We note that all elements of the matrix $\widetilde{\mathbf{Z}}$ are independent mixed Gaussian random variables, i.e., $\widetilde{Z}_{ij} \sim \widetilde{\mathcal{N}}_c\left(0, 1/(p_j m), p_j\right)$ and $\widetilde{Z}^{\mathrm{R}} \sim \widetilde{\mathcal{N}}\left(0, 1/(2p_j m), p_j\right)$.

Using the equivalent expression in (9), we rewrite the signal model in (3) as

$$\mathbf{y} = \sqrt{m}\widetilde{\mathbf{Z}}\mathbf{\Gamma}\mathbf{\Psi}\mathbf{x} + \mathbf{e}, \tag{10}$$

where the matrix $\mathbf{\Psi}$ is a unitary matrix. The distinct signal powers in $\mathbf{\Gamma}$ are spread along sparsity-inducing basis vectors (i.e., columns of $\mathbf{\Psi}$).

A matrix (or more correctly, a *sequence* of matrices) is *standard column regular* if all elements are uniformly bounded by some constant [12]. We normalize the matrix $\mathbf{\Gamma}\mathbf{\Psi}$ to be standard column regular. The normalization constant is $\|\mathbf{\Gamma}\mathbf{\Psi}\|_F/\sqrt{n} = \sqrt{P_{\mathrm{ave}}}$, where $P_{\mathrm{ave}} = \sum_{j=1}^{n} p_j b_j \nu_j^2/n$ denotes the average (receive) signal power in one time slot. Then the normalized matrix

$$\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Psi}/\sqrt{P_{\mathrm{ave}}} \tag{11}$$

has bounded spectral norm. Dividing both sides of (10) by $\sqrt{mP_{\mathrm{ave}}}$ yields the normalized signal model

$$\widetilde{\mathbf{y}} = \widetilde{\mathbf{Z}}\mathbf{\Sigma}\mathbf{x} + \widetilde{\mathbf{e}} = \mathbf{A}\mathbf{x} + \widetilde{\mathbf{e}}, \tag{12}$$

where all noise components are independent, zero mean and have normalized variance $\widetilde{\sigma}^2 \triangleq \sigma^2/(mP_{\mathrm{ave}})$. The average SNR is defined as

$$\mathrm{SNR}_{\mathrm{ave}} \triangleq \frac{P_{\mathrm{ave}}}{\sigma^2} = \frac{1}{n\sigma^2}\sum_{j=1}^{n} p_j b_j \nu_j^2. \tag{13}$$

[4]The receive signal power depends on both the channel condition (i.e., the variance of fading coefficients ν_j^2) and the average transmit power $p_j b_j$ that is governed by the accumulated harvested energy.

IV. Asymptotic Analysis of Energy-Aware Wireless Compressive Sensing

Having derived the probability distribution of elements of the matrix $\mathbf{Z}$ in Section III-C, we recall the definition of RIP [13] and state our main result, that is Theorem 1, in Section IV-A. Discussions and engineering implication of Theorem 1 are given in IV-B. We obtain an achievable system delay under an allowable MSE in Section IV-C. Finally we analyze the impact of inhomogeneity of SNRs on RIP and the required number of measurements in Section IV-D.

A. Restricted Isometry Property and Main Result

CS theory tells us that a sufficient condition for accurate and efficient recovery is that the sensing matrix satisfies the RIP. A matrix $\mathbf{A}$ satisfies the k-RIP, if there exists a $\delta_k \in (0,1)$ such that

$$(1-\delta_k)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\delta_k)\|\mathbf{x}\|_2^2 \tag{14}$$

holds for all k-sparse vectors $\mathbf{x}$. The smallest constant δ_k satisfying (14) is known as the *restricted isometry constant* (RIC) [13]. When the sensing matrix $\mathbf{A}$ is random, the inequality should hold with overwhelming probability that approaches one as n grows. Many families of random matrices, e.g., i.i.d. Gaussian random matrices and Bernoulli random matrices are known to satisfy the RIP [13], [14]. As a result, to evaluate the recovery performance, all we have to show is that the sensing matrix $\mathbf{A}$ in our scheme also obeys RIP with overwhelming probability.

The RIP requires that the sensing matrix $\mathbf{A}$ approximately preserves the Euclidean norm of sparse vectors. For the signal model in (12), the entries in $\widetilde{\mathbf{Z}}$ can be shown to be independent *sub-Gaussian* random variables because each entry is a mixture of a Dirac delta and a Gaussian, i.e., a mixed Gaussian random variable. See details in the Definition 6 in Appendix A. It is known that large random matrices with independent sub-Gaussian entries approximately preserve the Euclidean norm of sparse vectors with high probability [15]. Since $\mathbf{A} = \widetilde{\mathbf{Z}}\mathbf{\Sigma}$, we need to analyze the norm-preserving property of $\mathbf{\Sigma}$ defined in (11). Note that $\mathbf{\Gamma}$ and thus $\mathbf{\Sigma}$ depend on the vector of probabilities $\mathbf{p} \triangleq [p_1, \ldots, p_n]$. We now define the *k-restricted extreme eigenvalues* of the Gram matrix $\mathbf{\Sigma}^*\mathbf{\Sigma}$ as

$$\begin{aligned} \rho_{\max}(k,\mathbf{p}) &\triangleq \max_{\mathbf{v}:\|\mathbf{v}\|_0 \leq k, \|\mathbf{v}\|_2=1} \|\mathbf{\Sigma}\mathbf{v}\|_2^2, \\ \rho_{\min}(k,\mathbf{p}) &\triangleq \min_{\mathbf{v}:\|\mathbf{v}\|_0 \leq k, \|\mathbf{v}\|_2=1} \|\mathbf{\Sigma}\mathbf{v}\|_2^2, \end{aligned} \tag{15}$$

where $\mathbf{v} \in \mathbb{C}^n$, and the "$l_0$-norm" $\|\mathbf{v}\|_0$ refers to the number of non-zero elements of $\mathbf{v}$. The extreme eigenvalues will be used to understand how the inhomogeneous SNRs affects the RIP.

Lemma 1. We have $1 \leq \rho_{\max}(k,\mathbf{p}) \leq k$ and $0 \leq \rho_{\min}(k,\mathbf{p}) \leq 1$.

Proof: Fix a vector $\mathbf{v} \in \mathbb{C}^n$ such that $\|\mathbf{v}\|_2 = 1$ and $\|\mathbf{v}\|_0 = k$. Let $\mathcal{T} \subset \{1,\ldots,n\}$ with $|\mathcal{T}| \leq k$ be the support of $\mathbf{v}$. Let $\mathbf{\Sigma}_{\mathcal{T}} \in \mathbb{C}^{n\times|\mathcal{T}|}$ be the submatrix of $\mathbf{\Sigma}$ with column indices $\mathcal{T}$. Denote the eigenvalues of the Gram matrix $\mathbf{\Sigma}_{\mathcal{T}}^*\mathbf{\Sigma}_{\mathcal{T}}$ by $\lambda_1 \geq \ldots \geq \lambda_k \geq 0$. Due to the normalization in (11), the trace of $\mathbf{\Sigma}_{\mathcal{T}}^*\mathbf{\Sigma}_{\mathcal{T}}$ is $\sum_{j=1}^k \lambda_j = k$. This implies that the λ_1 is at least one and at most k. Similarly, the λ_k is no larger than one. ■

We note that the sparsity level k is usually much smaller than the number of sensors n in large-scale WSNs. We assume $\rho_{\max}(k,\mathbf{p}) \in [1,2]$ and $\rho_{\min}(k,\mathbf{p}) \in [0,1]$. This simplifies some mathematical arguments. We further assume that both $\rho_{\max}(k,\mathbf{p})$ and $\rho_{\min}(k,\mathbf{p})$ concentrate around one as n becomes large. The validity of this assumption will be shown analytically in Section IV-D and numerically in Section V-B. To state our main result cleanly, we define two quantities that depend on $\mathbf{\Sigma}$ and k as follows

$$\begin{aligned}\xi_k(\mathbf{p}) = \xi_k(\mathbf{\Sigma}) &\triangleq \max\left\{1-\rho_{\min}(k,\mathbf{p}), \rho_{\max}(k,\mathbf{p})-1\right\},\\ \zeta_k(\mathbf{p}) = \zeta_k(\mathbf{\Sigma}) &\triangleq \max\left\{0, \frac{2-\rho_{\max}(k,\mathbf{p})-\rho_{\min}(k,\mathbf{p})}{\rho_{\max}(k,\mathbf{p})-\rho_{\min}(k,\mathbf{p})}\right\},\end{aligned} \tag{16}$$

where the argument $\mathbf{p}$ is introduced to make the dependence on $\mathbf{p}$ clear. The sparsity-inducing basis, channel conditions and amplitudes of transmit signals are kept fixed. Since $\rho_{\max}(k,\mathbf{p}) \in [1,2]$, we have $\xi_k(\mathbf{p}), \zeta_k(\mathbf{p}) \in [0,1]$. Let $\vartheta_k(\mathbf{p}) = (1+\zeta_k(\mathbf{p}))\rho_{\max}(k,(\mathbf{p})) - 1$. Given $\delta_k \in (\xi_k(\mathbf{p}),1)$, for convenience, we map δ_k to a "modified RIC" via a piecewise linear mapping as follows

$$\beta_k(\delta_k,\mathbf{p}) = \beta_k(\delta_k,\mathbf{\Sigma}) \triangleq \begin{cases} 1-(1-\delta_k)/\rho_{\min}(k,\mathbf{p}), & \delta_k \in (\xi_k(\mathbf{p}),\vartheta_k(\mathbf{p})),\\ (1+\delta_k)/\rho_{\max}(k,\mathbf{p})-1, & \delta_k \in (\vartheta_k(\mathbf{p}),1).\end{cases} \tag{17}$$

Let $\varsigma_k(\mathbf{p}) = 2/\rho_{\max}(k,\mathbf{p}) - 1$. For convenience, we express δ_k in terms of β_k as follows

$$\delta_k = \begin{cases} 1-(1-\beta_k)\,\rho_{\min}(k,\mathbf{p}), & \beta_k \in (0,\zeta_k(\mathbf{p}))\,,\\ (1+\beta_k)\,\rho_{\max}(k,\mathbf{p})-1, & \beta_k \in (\zeta_k(\mathbf{p}),\varsigma_k(\mathbf{p}))\,.\end{cases} \tag{18}$$

Based on the assumption on $\rho_{\max}(k,\mathbf{p})$ and $\rho_{\min}(k,\mathbf{p})$, we know that $\xi_k(\mathbf{p})$, which measures the inhomogeneity of the eigenvalues of $\mathbf{\Sigma}_{\mathcal{T}}^*\mathbf{\Sigma}_{\mathcal{T}}$ for $|\mathcal{T}| \leq k$, is a small positive number. This implies that $\zeta_k(\mathbf{p})$ is small, and the deviation between $\beta_k(\delta_k,\mathbf{p})$ and δ_k is also small. Define $p_{\min} \triangleq \inf_{j\in\mathbb{N}} p_j$ and assume that $p_{\min} > 0$.

Theorem 1. *Let $c_1, c_2 > 0$ be some universal constants. Given a sparsity level $k < \lfloor n/2 \rfloor$ and a number $\delta_k \in (\xi_k(\mathbf{p}), 1)$, if the number of measurements satisfies*

$$m > \frac{c_1 k \rho_{\max}(k, \mathbf{p})}{p_{\min}^2 \beta_k^2(\delta_k, \mathbf{p}) \rho_{\min}(k, \mathbf{p})} \log \frac{5en}{k}, \tag{19}$$

where $\beta_k(\delta_k, \mathbf{p})$ is defined in (17)*, then for any vector $\mathbf{x}$ with support of cardinality of at most k, we have that the RIP in* (14) *holds with probability at least*

$$1 - \exp\left(-c_2 m p_{\min}^2 \beta_k^2(\delta_k, \mathbf{p})/4\right). \tag{20}$$

Proof: (Sketch)When all quantities are real, we show in step 1 that $\widetilde{\mathbf{Z}}$ acts as isometry on the images of sparse vectors under matrix $\mathbf{\Sigma}$. Then by showing the rows of $\widetilde{\mathbf{Z}}$ are isotropic sub-Gaussian and by exploiting the so-called "restricted eigenvalue property" of $\mathbf{\Sigma}$, we derive an RIP for $\mathbf{A}$ in step 2. We finally extend the RIP to the complex case in step 3. See details in Appendix A. ■

B. Discussion

Note that the rows $\mathbf{a}_i$'s of the sub-Gaussian sensing matrix $\mathbf{A}$ are *non-isotropic*, because of the inhomogeneous signal power pattern. To the best of our knowledge, little is known about the RIP of non-isotropic sub-Gaussian random matrices. The only relevant result is [15, Remark 5.40] which gives a concentration inequality for such matrices in terms of an upper bound on the spectral norm. However, the authors did not study how the inhomogeneity affects the RIP, nor did they investigate the number of measurements required for the RIP. Theorem 1 fills this gap.

Remark 1 (Specialization to the homogeneous case). We say that the signal power pattern is *homogeneous in SNR* if $\mathbf{\Gamma}$ is a multiple of the identity matrix $\mathbf{I}_n$. Because the noise powers are identical, this means that the receive signal powers $b_j p_j \nu_j^2$'s do not depend on the sensor $1 \leq j \leq n$. From (19), we observe that the lower bound on the required number of measurements is $O(\frac{k\rho_{\max}(k,\mathbf{p})}{p_{\min}^2 \beta_k^2(\delta_k,\mathbf{p})\rho_{\min}(k,\mathbf{p})} \log \frac{n}{k})$. Specializing this to the homogeneous signal power pattern case, we have $O(\frac{k}{\delta_k^2} \log \frac{n}{k})$ because $\rho_{\max}(k, \mathbf{p}) = \rho_{\min}(k, \mathbf{p}) = 1$ and $\beta_k(\delta_k, \mathbf{p}) = \delta_k$. This specialization coincides with known results for i.i.d. sensing matrices. See [14, Theorem 5.2] and [15, Section 1.4.4].

Remark 2 (Asymptotic nature of Theorem 1). Because we assume that $\rho_{\min}(k, \mathbf{p})$ and $\rho_{\max}(k, \mathbf{p})$ are close to one (which happens when n is large), Theorem 1 is an asymptotic result. Data recovery in

the small-n regime will be discussed in Section V. In the large-n regime, let us compare $\rho_{\max}(k, \mathbf{p})$ to $\rho_{\max}(k, \bar{p}\mathbf{1}_{n\times 1})$, the maximum k-restricted eigenvalue when the p_j's are set to their empirical mean, i.e., $\bar{p} \triangleq \frac{1}{n}\sum_{j=1}^{n} p_j$. In the case where we treat the p_j's as i.i.d. with mean $\bar{p}$, then both ratios $\frac{\rho_{\max}(k,\mathbf{p})}{\rho_{\max}(k,\bar{p}\mathbf{1}_{n\times 1})}$ and $\frac{\rho_{\min}(k,\mathbf{p})}{\rho_{\min}(k,\bar{p}\mathbf{1}_{n\times 1})}$ converge to one in probability as n tends to infinity if $b_j\nu_j^2$ are uniformly bounded away from 0 and $+\infty$. See the justification in Appendix B.

Remark 3 (On the choice of transmit probability $\mathbf{p}$). Continuing from Remark 2, in the large-n regime and treating p_j's are i.i.d. random variables, we note that the lower bound in (19) is dominated by $p_{\min}$, since the quantities $\rho_{\max}(k, \mathbf{p})$, $\rho_{\min}(k, \mathbf{p})$ and $\beta_k(\delta_k, \mathbf{p})$ are shown to be asymptotically dependent *only* on $\bar{p}$. As such, for given $\bar{p}$, we should choose the p_j's to be identical in order to maximize $p_{\min}$ and thus minimize the lower bound on m.

Following Remark 3, henceforth, we will assume that all the transmit probabilities are identical and equal to some common $p \in (0, 1)$. Thus, the normalized matrix $\mathbf{\Sigma}$ defined in (11) is independent of p. The quantities $\rho_{\max}(k, \mathbf{p})$, $\rho_{\min}(k, \mathbf{p})$, $\xi_k(\mathbf{p})$, $\vartheta_k(\mathbf{p})$ and $\beta_k(\delta_k, \mathbf{p})$ will be denoted simply as $\rho_{\max}(k)$, $\rho_{\min}(k)$, ξ_k, ϑ_k and $\beta_k(\delta_k)$ respectively in the sequel. In addition, $p_{\min}^2$ in Theorem 1 is replaced by p^2 so the lower bound essentially scales as $1/p^2$.

In the proof of Lemma 6 in Appendix A, we show that for each element $\widetilde{Z}_{ij}$, its sub-Gaussian tail probability is bounded above by $pe^{-pt^2/2}$. Note that the sub-Gaussian norm of a random variable is the smallest constant $\varrho > 0$ for which the sub-Gaussian tail probability is $2e^{-t^2/(2\varrho^2)}$ (Definition 6). In view of the fact that the pre-factor in our bound is p (and not 2), there is some looseness with respect to p in the proof of Lemma 6 and thus Theorem 1. For larger p, the loss is reduced.

C. Achievable System Delay

The performance of the WCS scheme is characterized by two quantities, i.e., the MSE and the system delay. The MSE performance under bounded noise is studied in the CS literature [3], [16], [17]. Note that there is often a trade-off between the two quantities. Under an allowable MSE $\epsilon > 0$, we thus analyze the *achievable system delay* $D(\epsilon)$, which is defined as

$$D(\epsilon) \triangleq \min_{m} \; m \quad \text{subject to} \quad \mathbb{E}\|\widehat{\mathbf{x}} - \mathbf{x}\|_2^2 \leq \epsilon. \tag{21}$$

Corollary 1. *Let n, k, ξ_k, ϑ_k be as in Theorem 1. Let $\epsilon_{\rm th} \triangleq 1/(0.0942 \times \mathrm{SNR}_{\rm ave})$. Given an allowable MSE $\epsilon > \epsilon_{\rm th}$, with overwhelming probability (exceeding* (20)*), the achievable system delay is*

$$D(\epsilon) = \frac{c_1 k \rho_{\max}(k)}{p^2 \tilde{\beta}_k^2 \rho_{\min}(k)} \log \frac{5en}{k}, \tag{22}$$

where

$$\tilde{\beta}_k = \tilde{\beta}_k(\mathbf{\Sigma}, \epsilon, \mathrm{SNR}_{\rm ave}) \triangleq \begin{cases} 1 - \dfrac{0.693 + 1/\sqrt{\epsilon \mathrm{SNR}_{\rm ave}}}{\rho_{\min}(k)}, & \delta_k \in (\xi_k, \vartheta_k), \\ \dfrac{1.307 - 1/\sqrt{\epsilon \mathrm{SNR}_{\rm ave}}}{\rho_{\max}(k)} - 1, & \delta_k \in (\vartheta_k, 1). \end{cases} \tag{23}$$

Proof: We start the proof by leveraging on the following lemma.

Lemma 2 (Theorem 3.2 in [16]). Let $\widetilde{\mathbf{y}} = \mathbf{A}\mathbf{x} + \widetilde{\mathbf{e}}$, where $\mathbf{x}$ is a k-sparse vector in $\mathbb{C}^n$, $\widetilde{e} \in \mathbb{C}^m$ is a zero mean, white random vector whose entries have variance σ^2. If the $\mathbf{A}$ satisfies the RIP with RIC $\delta_k < 0.307$, then the solution $\widehat{\mathbf{x}}$ to the ℓ_1-minimization problem in CS decoder [3], [15] satisfies

$$\mathbb{E}\|\widehat{\mathbf{x}} - \mathbf{x}\|_2^2 \leq \frac{\sigma^2}{P_{\rm ave}(0.307 - \delta_k)^2}. \tag{24}$$

Recall the definition of $\mathrm{SNR}_{\rm ave}$ in (13). From Lemma 2, to achieve a MSE ϵ, it suffices to ensure the RIC satisfies $\delta_k^* = 0.307 - 1/\sqrt{\epsilon \mathrm{SNR}_{\rm ave}}$. From Theorem 1, the required minimum number of measurements such that the RIP holds with overwhelming probability is the right-hand-side expression of (22). The definition of the achievable system delay establishes Corollary 1. ■

Note that Corollary 1 applies only to the case where the MSE ϵ is greater than the threshold $\epsilon_{\rm th}$. If $\epsilon \leq \epsilon_{\rm th}$, then from (24), simple algebra reveals that $\delta_k = 0$, which implies that the sensing matrix $\mathbf{A}$ is a perfect isometry. Since $\mathbf{A}$ is random, and the entries are governed by a continuous density, this occurs with probability zero, implying that the constraint in (21) is almost surely not satisfied. Thus, we define the system delay to be ∞.

As either ϵ or $\mathrm{SNR}_{\rm ave}$ increases, $\tilde{\beta}_k$ increases, and thus the system delay $D(\epsilon)$ decreases. We will analyze the impact of inhomogeneity of SNRs on the deviation of $\rho_{\max}(k)$ and $\rho_{\min}(k)$ from unity and hence on the system delay, in Section IV-D.

Remark 4 (Total energy consumption). For transmissions in one frame, we note that the total energy consumption for all sensors is $D(\epsilon) p \sum_{j=1}^n b_j$. Then from (22) in Corollary 1, the total energy consumption is inversely proportional to $p\tilde{\beta}_k^2(\mathbf{\Sigma}, \epsilon, \mathrm{SNR}_{\rm ave})$. Thus, to minimize the total energy consumption, p should be chosen to be as large as possible, since $\tilde{\beta}_k$ increases with $\mathrm{SNR}_{\rm ave}$ and

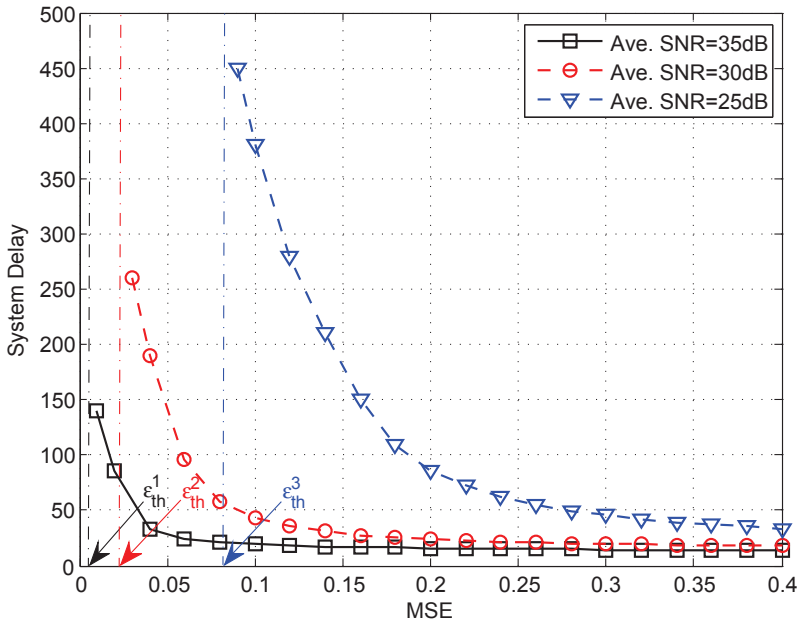


Fig. 2: Plot of achievable system delay v.s. allowable MSE.

the latter is proportional to p as shown in (13).

In practice, the squared amplitudes of transmitted signals from sensors are bounded, i.e., $b_{\min} \leq b_j \leq b_{\max}$ for all $1 \leq j \leq n$. Additionally, taking into account the maximum storage energy B, we have $pb_j \leq B$. Thus, all sensors should transmit with the probability $p^* = \min\{1, \frac{B}{b_{\min}}\}$. Clearly, some energy-rich sensors will not consume all its available energy in the current frame. The saved harvested energies during the current and previous frames can be used in subsequent frames.

Furthermore, one can address problems on power management over a long time period. First, to decrease the total system delay or minimize the total energy consumption, the transmit power can be allocated among successive frames. By optimizing p in each frame, one can perform power allocation among different frames subject to the causal availability of the harvested energy–a resource allocation problem with energy harvesting constraint as in [2]. Second, an energy threshold for transmission can be set. When the available energy exceed the threshold, it probabilistically transmit, otherwise it keeps silent and thus save energy for future transmissions. Analyses of such problems are beyond the scope of this paper.

Example 1. Let the number of sensors $n = 500$, the sparsity level $k = 5$ and the transmit probability $p = 0.8$. These parameters imply $\rho_{\max}(k) = 1.09, \rho_{\min}(k) = 0.88$ (see Section IV). We plot the

achievable system delay $D(\epsilon)$ against the allowable MSE ϵ, for different average SNRs in Fig. 2. We observe that beyond the MSE threshold (that depends on the average SNR), the system delay $D(\epsilon)$ decreases as either ϵ or $\mathrm{SNR}_{\mathrm{ave}}$ increases, which is expected.

D. Impact of Inhomogeneity of SNRs

We now study the impact of inhomogeneity of receive SNRs on the number of measurements needed to satisfy the RIP, in terms of the k-restricted eigenvalues. Without loss of generality, we assume all sensors have the same noise power, hence, it suffices to analyze the impact of inhomogeneity of receive signal powers.[5] We treat the overall receive signal powers as independent random variables, and thus denote the k-restricted eigenvalues as $\rho_{\max}(k)$ and $\rho_{\min}(k)$. We focus on the asymptotic scenario where the number of sensors n tends to infinity and, for the ease of analysis, k is kept constant. To make the dependence on n clear, we denote $\rho_{\max}(k)$ (resp. $\rho_{\min}(k)$) as $\rho_{\max}(k,n)$ (resp. $\rho_{\min}(k,n)$). It will be shown that both $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$ concentrate around one when n is large, and the rate of convergence to one depends on the inhomogeneity of SNRs. This implies that the recovery performance (required number of measurements and probability that the RIP holds in Theorem 1) is not sensitive to the inhomogeneity of SNRs when n is large.

Let $\mathbf{w} = \mathbf{\Sigma v}$, where the unit-norm, k-sparse vector $\mathbf{v}$ is supported on the set $\mathcal{T} \triangleq \{s_1, \ldots, s_k\} \subset \{1, \ldots, n\}$ and let $s_1 < \ldots < s_k$. To obtain further insights, we let $\mathbf{\Psi}$ be the n-point discrete Fourier transform (DFT) matrix. Then the squared ℓ_2-norm of $\mathbf{w}$ can be expressed as follows

$$\|\mathbf{w}\|_2^2 = \frac{1}{nP_{\mathrm{ave}}} \sum_{i=1}^{n} \gamma_i \left(1 + \sum_{q=1}^{k} \sum_{l=1, l<q}^{k} 2\mathrm{Re}\left\{ v_{s_q} v_{s_l}^* \exp\left(\frac{-j2\pi(i-1)(s_q - s_l)}{n} \right) \right\} \right). \tag{25}$$

Since $\|\mathbf{w}\|_2^2$ is strongly influenced by the inner summation terms, we analyze the behavior of these terms more carefully in the sequel. When the signal power pattern is homogeneous, i.e., $\mathbf{\Gamma} = \mathrm{diag}(\sqrt{\gamma}, \ldots, \sqrt{\gamma})$, we have $\|\mathbf{w}\|_2^2 = \|\mathbf{\Sigma v}\|_2^2 = 1$, hence $\rho_{\max}(k,n) = \rho_{\min}(k,n) = 1$ for all k, n.

We are interested to know how $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$ vary with different signal powers γ_i's. Thus, we consider a model in which the γ_i's are i.i.d. random variables following an approximate Gaussian distribution. By varying the variance of this distribution, we are in fact varying the

[5]The signal powers depend on the transmit probability p_j, the squared amplitude b_j and the variance of channel coefficients ν_j^2.

inhomogeneity of the signal powers. Specifically, to deal with the fact that the signal powers cannot be negative, we use the following truncated Gaussian distribution to model the signal powers.

Definition 2. *A random variable X is* truncated Gaussian*, denoted as $\mathcal{N}_{\rm tr}(\mu,\omega^2)$, if its pdf is*

$$g_X(x;\mu,\omega)=\frac{1}{\sqrt{2\pi}\omega(1-Q(\mu/\omega))}\exp\left(-\frac{(x-\mu)^2}{2\omega^2}\right) \tag{26}$$

for $x\geq 0$ and 0 else, where $Q(x)\triangleq\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}\,dt$ is the Q-function of a standard Gaussian pdf.

We assume that $\gamma_i\sim\mathcal{N}_{\rm tr}(\mu,\omega^2)$ for all $i=1,\ldots,n$ and they are mutually independent. Given μ, the "variance" ω^2 is a measure of the degree of inhomogeneity of the signal powers γ_i's. Also, the parameter $d\triangleq\mu/\omega$ is a measure of the homogeneity of the SNRs. If d is small (resp. large), the SNRs are less (resp. more) homogeneous. We use the exponential asymptotic notation $a_n\dot{\leq}\exp(-nE)$ to mean that $\limsup_{n\to\infty}\frac{1}{n}\log a_n\leq -E$. Under the above assumptions on the statistics of the signal powers, we have the following large deviations upper bound on $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$:

Theorem 2. *Let $d\triangleq\mu/\omega$. For any $t>0$, and any constant $1\leq k<\lfloor n/2\rfloor$,*

$$\begin{aligned}\mathbb{P}\left(\rho_{\max}(k,n)>1+t\right)&\dot{\leq}\exp\left[-nd^2E(k,t)^2\right],\\ \mathbb{P}\left(\rho_{\min}(k,n)<1-t\right)&\dot{\leq}\exp\left[-nd^2E(k,t)^2\right],\end{aligned} \tag{27}$$

where the exponent $E(k,t)$ is defined as $E(k,t)\triangleq t/(k-1+\sqrt{2}t)$.

Proof: (Sketch) We analyze $\|\mathbf{w}\|_2^2$ in (25) and show that it involves a ratio of sums of random variables. We then bound this ratio using the theory of large deviations. See Appendix C. ■

Recall that Theorem 1 says that both the required number of measurements and the probability that the RIP holds depends on the ratio $r(k,n)=\rho_{\max}(k,n)/\rho_{\min}(k,n)$. From Theorem 2, we note that both $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$ concentrate around one in the large n regime (for bounded k), and the rate of convergence to one depends on the inhomogeneity of SNRs. This allows us to conclude that for large-scale EHWSNs, the inhomogeneity of SNRs does not significantly affect the RIP and the system delay, which is a surprisingly positive observation.

We note that $E(k,t)$ is an increasing function of t and a decreasing function of the sparsity k which is expected. Also, the exponent $d^2E(k,t)^2$ increases with d, which means that the convergence of $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$ to unity is faster when d is large, or equivalently, when the signal powers are more homogeneous. It is observed that both $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$ are close to one in the

large n regime. This validates the assumption on $\rho_{\max}(k,n)$ and $\rho_{\min}(k,n)$ in Section IV-A.

Remark 5. One may wonder whether Theorem 2 depends strongly on $\mathbf{\Psi}$ being the DFT matrix. In fact, the only property of the DFT that we exploit in the proof of Theorem 2 is its circular symmetry, i.e., each basis vector of the DFT is uniformly distributed over the circle in the complex plane. Hence, certain Cesàro-sums converge to zero and the proof goes through. See (39) in Appendix C. Thus, Theorem 2 also applies for other sparsity-inducing bases whose basis vectors have the circular symmetric property, e.g., the discrete cosine transform (DCT) or the Hadamard transform.

V. Simulation Results

We now validate our results numerically. We use the truncated Gaussian with $\mu = 0.2$ to model the receive signal powers, and the basis pursuit de-noising (BPDN) algorithm [18] as the CS decoder.

A. Recovery Performance

First, we simulate the recovery performance for the scenario where different sensors transmit with different probabilities p_j. We fix the numbers of sensors $n = 500$ and the sparsity level $k = 20$. For Case 1, we let the p_j's take values from $\{0.6, 0.8, 1\}$ with equal probability. For comparison, in Case 2 and Case 3, we set $p_j = 0.6$ and $p_j = 0.8$ respectively. Thus for these two cases, the p_j's are identical. We define an outage event if the normalized squared error is greater than 0.005. The outage probability is calculated from 10^5 independent simulations and plotted in Fig. 3.

From Fig. 3, we observe that, for sufficiently large m, the outage probability in log scale for Case 1 and Case 2 have approximately the same rates of decrease, while Case 1 and Case 3 have different rates of decreases even though both have the same average transmit probabilities. This observation is consistent with Theorem 1. It is because from (20), the outage probability scales as $\exp(-cmp_{\min}^2)$ where c is some constant. Since Case 1 and Case 2 have the same $p_{\min} = 0.6$, while Case 3 has $p_{\min} = 0.8$, the outage probabilities decrease exponentially at the same rate for Cases 1 and 2 but at a higher rate for Case 3. Thus, even though $\rho_{\max}(k, \mathbf{p})$ (resp. $\rho_{\min}(k, \mathbf{p})$) is close to $\rho_{\max}(k, \bar{p}\mathbf{1}_{n\times 1})$ (resp. $\rho_{\min}(k, \mathbf{p})$) when n is large (cf. Remark 2), $\bar{p}$ alone is *not* a good prediction of the performance for the case of different p_j's; rather $p_{\min}$ governs the lower bound on m.

Second, we simulate the recovery performance in terms of the MSE when the transmit probabilities are all fixed to be $p = 0.8$ and the parameter $d = 2$. We use numerical simulations to obtain insights when n is small. Before we do so, we make an observation from Theorem 1. We see that given k,

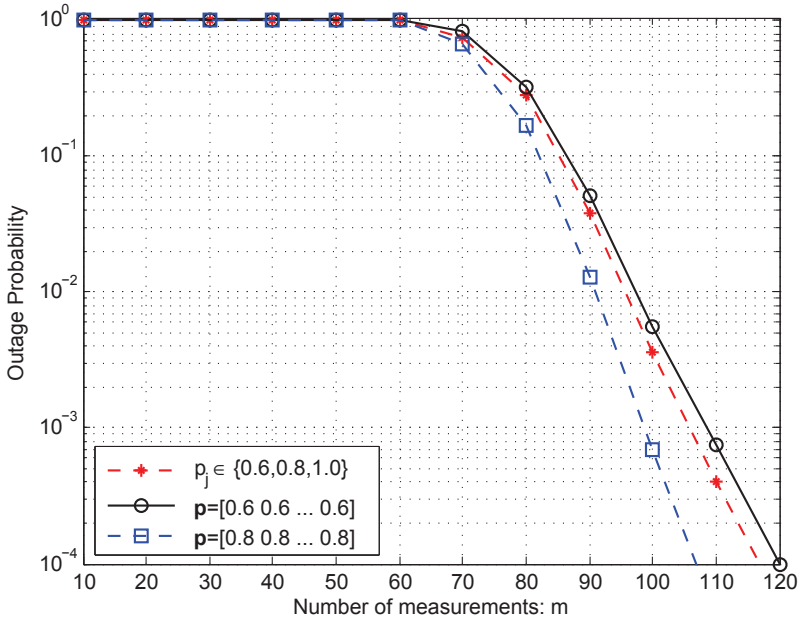


Fig. 3: Plot of MSE against number of measurements.

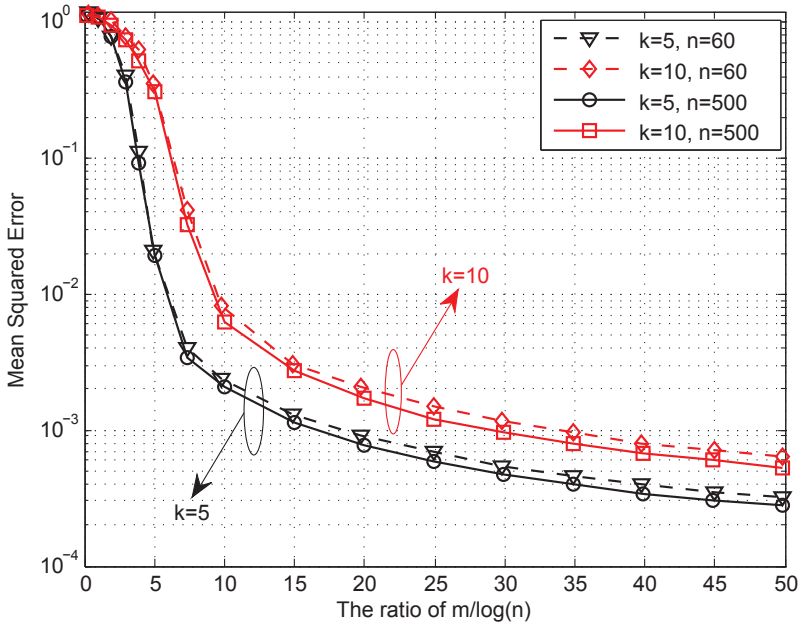
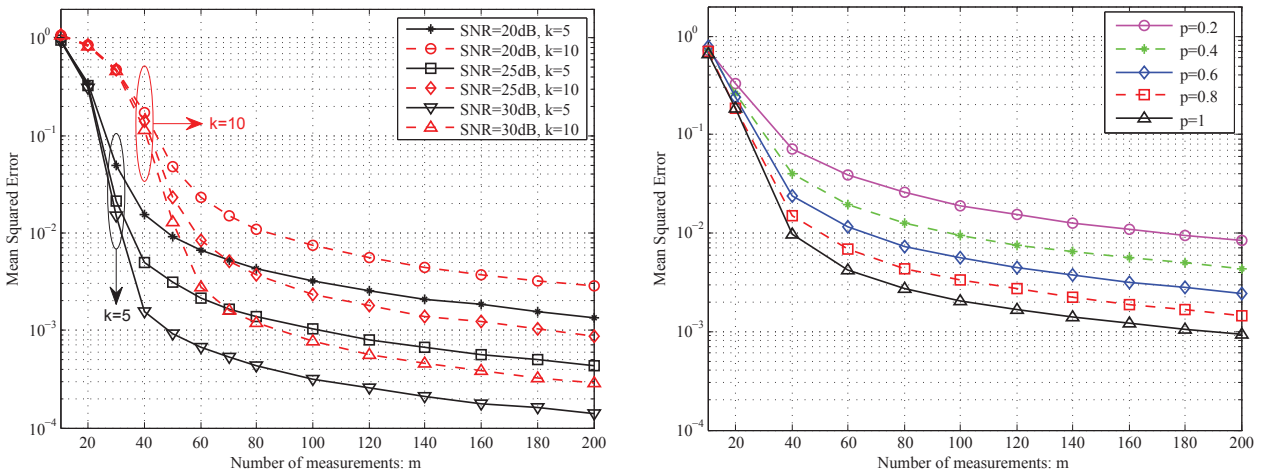


Fig. 4: Plot of MSE against the ratio m/n.

the lower bound on the ratio $m/\log(n)$ is given by $c_4/\log(n) + c_3$, where c_3 and c_4 are constants that are dependent only on k. Therefore, the lower bound asymptotically approaches a constant for large n. Fig. 4 compares the MSE performances for $n = 60, 500$. We observe from Fig. 4 that for a



(a) Plot for different k and average SNRs.

(b) Plot for different p (identical for all sensors).

Fig. 5: Plot of MSE against the number of measurements.

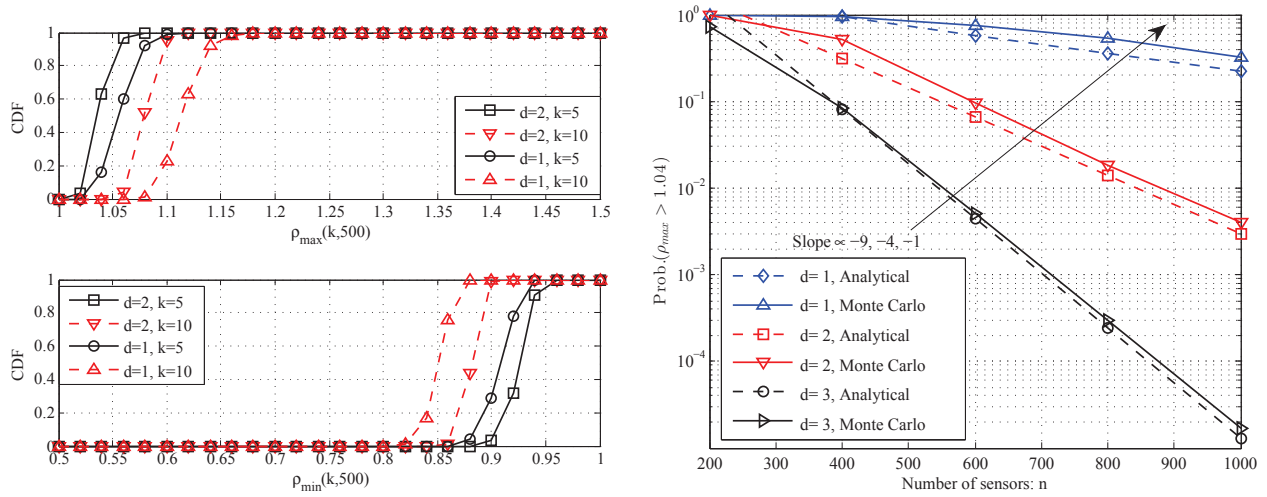
given sparsity k, the MSE curves for different n's are almost the same. Hence, even in the small-n scenario, the WCS scheme can also recover the sensor data reliably. Thus, the numerical results for fairly small n corroborate very well with analysis which applies for large n. Hence, we conclude that our asymptotic analysis can also be applied to the practical case of moderate n.

Fig. 5a plots the MSE against m for different sparsities k and different average SNRs, given $n = 500$. As expected, the MSE decreases as either k decreases or the average SNR increases. Consider the MSE level 2×10^{-3}. When the average SNR is 25 dB, the WCS scheme achieves a smaller system delay of $D = 68$ for $k = 5$ compared to $D = 115$ for $k = 10$.

Fig. 5b compares the MSE performances for different transmit probabilities $p = 0.2, 0.4, 0.6, 0.8, 1$, respectively, given $n = 500$. It is observed that for a desired MSE level, the required number of measurements decreases as p increases. In practice, we can choose p that leads to reduced system delay, for a given MSE requirement. To achieve a system delay $D \leq 120$, we choose $p = 0.8$ for the high-accuracy ($\text{MSE} \leq 0.003$) recovery, or choose $p = 0.4$ for the low-accuracy ($\text{MSE} \leq 0.01$) recovery. We observe that as p decreases from one, this lower bound on m increases at a rate slower than p^{-2}. This observation coincides with the discussion (after Remark 3) for Theorem 1.

B. Impact of Inhomogeneity of SNRs

By modeling SNRs as independent truncated Gaussian random variables, Fig. 6a shows the cumulative distribution function (CDF) of $\rho_{\max}(k, 500)$ and $\rho_{\min}(k, 500)$. We note that both $\rho_{\max}(k, 500)$



(a) CDF of $\rho_{\max}(k, 500)$ and $\rho_{\min}(k, 500)$.

(b) Plot of the probability of $\rho_{\max}(k, n) > 1.04$ v.s. n

Fig. 6: Plot of impact of inhomogeneity of SNR on $\rho_{\max}(k)$ and $\rho_{\min}(k)$.

and $\rho_{\min}(k, 500)$ converge to one faster for larger d, or equivalently, for more homogeneous SNRs. Also, under the same inhomogeneous SNRs, both converge to one faster for smaller k.

We also numerically validate the asymptotic behavior of $\rho_{\max}(k, n)$ as n grows. Fig. 6b shows the probability that $\rho_{\max}(k, n) > 1.04$ for different n. It is observed that the logarithm of the probability decreases linearly as n grows (when n/k is large) and furthermore, the slope varies quadratically w.r.t. d, i.e., the slope is proportional to $-1, -4, -9$ for $d = 1, 2, 3$, respectively. This observation corroborates the exponentially decreasing nature of the tail probability in (27) in Theorem 2.

Finally, Fig. 7 compares the MSEs of the inhomogeneous SNR and the homogeneous SNR scenarios, for the sparsity levels $k = 5, 10, 20$, given the average SNR to be 25dB. It is observed that in the inhomogeneous scenario, the MSE performance is slightly worse than that of the homogeneous-SNR scenario. Note that the degradation becomes larger as the sparsity k increases. This is because the convergence rate for $\rho_{\max}(k)$ and $\rho_{\min}(k)$ to one is faster if k is small relative to n. This corroborates the observation in Section IV-D.

VI. Conclusion

In this paper, we considered the scenario in which each sensor independently decides whether or not to transmit with some probability, and the overall transmission power depends on its available energy. Hence, only a subset of sensors transmits concurrently to the FC. Moreover, we exploit the spatial combination inherent in wireless channels. We use techniques from CS theory to prove

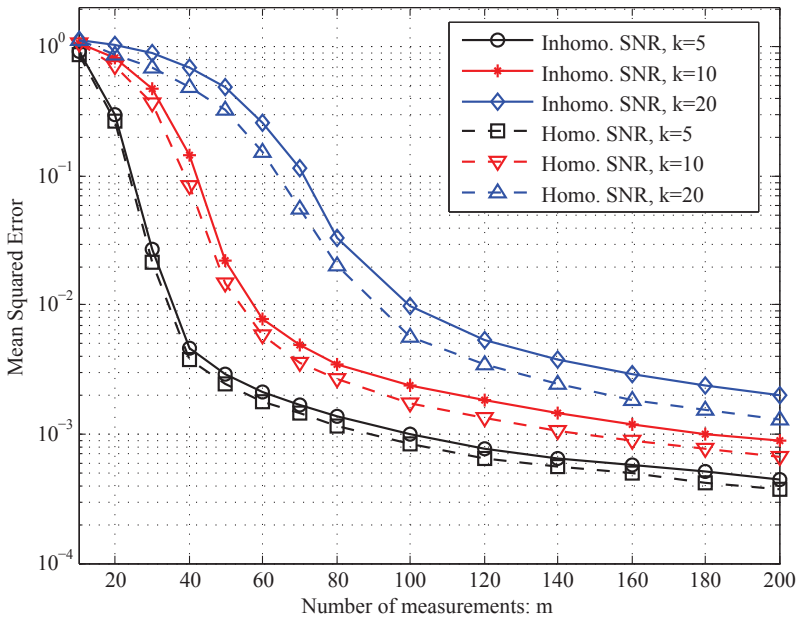


Fig. 7: Plot of MSE against the number of measurements.

a lower bound on the required number of measurements to satisfy the RIP and hence to ensure that the data recovery is both computationally efficient (and amenable to convex optimization) and accurate. The lower bound is used to obtain insights on system design, such as the choice of the transmit probabilities, the number of sensors, etc., to achieve a minimum system delay with high probability. Finally, we analyze the impact of inhomogeneity on the k-restricted extreme eigenvalues. These eigenvalues govern the number of measurements required for the RIP to hold. In large-scale EH-WSNs, we showed using large deviation techniques that the recovery accuracy and the system delay are not sensitive to the inhomogeneity of SNRs.

APPENDIX A
PROOF OF THEOREM 1

Proof: The sketch of the three-step proof is given below Theorem 1 in Section IV-A. Before step 1, we start with the following preliminaries. Let $d(\mathbf{u}, \mathbf{v})$ be the Euclidean distance in $\mathbb{R}^n$.

Definition 3 (Nets, covering numbers [15])**.** *Let$(\mathcal{U}, d)$ be a metric space and fix $\epsilon > 0$. A subset $\mathcal{N}_\epsilon \subset \mathcal{U}$ is called an* ϵ-net *if every $\mathbf{u} \in \mathcal{U}$ can be approximated to within ϵ by some $\mathbf{v} \in \mathcal{N}_\epsilon$, i.e., $d(\mathbf{u}, \mathbf{v}) \leq \epsilon$. The* covering number *$\mathcal{N}(\mathcal{U}, \epsilon)$ is the cardinality of the smallest ϵ-net of $\mathcal{U}$.*

Definition 4 (Set of sparse vectors)**.** *Let $\mathcal{S}^{n-1}$ be the unit sphere in $\mathbb{R}^n$ and $1 \leq k \leq n$. Define $\mathcal{U}_k \triangleq \{\mathbf{u} \in \mathcal{S}^{n-1} : \ \|\mathbf{u}\|_0 \leq k\}$ and also the subset of the Euclidean unit ball $\mathcal{B}_2^n$ with (at most) k-sparse vectors as $\widetilde{\mathcal{U}}_k \triangleq \{\mathbf{u} \in \mathcal{B}_2^{n-1} : \ \|\mathbf{u}\|_0 \leq k\}$.*

Lemma 3 (Upper bound on covering numbers, Lemma 2.3 in [19]). Let $0 < \epsilon < 1$ and $1 \leq k \leq n$. There exists an ϵ-net of $\widetilde{\mathcal{U}}_k$, namely $\mathcal{N}_\epsilon$, whose cardinality can be upper bounded as $|\mathcal{N}_\epsilon| \leq \left(\frac{5}{2\epsilon}\right)^k \binom{n}{k}$.

Definition 5 (Complexity measure [19])**.** *The* complexity *of a set $\mathcal{V} \subset \mathbb{R}^n$ is defined as $\ell_*(\mathcal{V}) \triangleq \mathbb{E}\left[\sup_{\mathbf{v}\in\mathcal{V}} |\langle \mathbf{v}, \mathbf{u}\rangle|\right]$, where $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and the supremum is over all vectors $\mathbf{v} \in \mathcal{V}$.*

Recall the signal model in (12), i.e., $\widetilde{\mathbf{y}} = \widetilde{\mathbf{Z}}\mathbf{\Sigma}\mathbf{x} + \widetilde{\mathbf{e}} = \mathbf{A}\mathbf{x} + \widetilde{\mathbf{e}}$. Given a subset $\mathcal{V} \subset \mathbb{R}^n$, we aim to measure the complexity of $\mathcal{W}(\mathcal{V})$, which is the image set of the set $\mathcal{V}$ under a fixed linear mapping $\mathbf{\Sigma}$. More precisely, we define $\mathcal{W}(\mathcal{V}) \triangleq \{\mathbf{w} \in \mathbb{R}^n : \ \mathbf{w} = \mathbf{\Sigma}\mathbf{v}, \text{ for some } \mathbf{v} \in \mathcal{V}\}$. Define the complexity of $\mathcal{W}(\mathcal{V})$ as $\ell_*\left(\mathcal{W}(\mathcal{V})\right) \triangleq \mathbb{E}\left[\sup_{\mathbf{v}\in\mathcal{V}} |\langle \mathbf{v}, \mathbf{\Sigma}\mathbf{u}\rangle|\right]$.

Lemma 4 (Upper bound on complexity measure, Lemma B.6 in [20]). Let $\mathcal{N}_{\frac{1}{2},k}$ be a $\frac{1}{2}$-net of $\widetilde{\mathcal{U}}_k$ provided by Lemma 3. Then for all $1 \leq k \leq n$, it holds that

$$\ell_*\left(\mathcal{W}(\mathcal{N}_{\frac{1}{2},k})\right) \leq 3\sqrt{k\rho_{\max}(k,\mathbf{p})\log\frac{5en}{k}}, \qquad \text{and} \qquad \ell_*\left(\mathcal{W}(\mathcal{U}_k)\right) \leq 2\ell_*\left(\mathcal{W}(\mathcal{N}_{\frac{1}{2},k})\right).$$

Define the set $\mathcal{E}_k \triangleq \{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{\Sigma}\mathbf{v}\|_2 = 1, \|\mathbf{v}\|_0 = k\}$, then for $\mathcal{V} = \mathcal{E}_k$, the complexity measure of the set $\mathcal{W}(\mathcal{E}_k)$ is bounded in the following Lemma.

Lemma 5. The complexity measure of the set $\mathcal{W}(\mathcal{E}_k)$ is upper bounded as

$$\ell_*\left(\mathcal{W}(\mathcal{E}_k)\right) \leq 6\sqrt{k\frac{\rho_{\max}(k,\mathbf{p})}{\rho_{\min}(k)}\log\frac{5en}{k}}. \tag{28}$$

Proof: For any vector $\mathbf{v} \in \mathcal{E}_k$ and any random vector $\mathbf{u} \in \mathbb{R}^n$, with probability one,

$$|\langle \mathbf{u}, \mathbf{\Sigma}\mathbf{v}\rangle| = |\langle \mathbf{v}, \mathbf{\Sigma}\mathbf{u}\rangle| = \|\mathbf{v}\|_2 \left|\left\langle \frac{\mathbf{v}}{\|\mathbf{v}\|_2}, \mathbf{\Sigma}\mathbf{u}\right\rangle\right| \leq \|\mathbf{v}\|_2 \sup_{\mathbf{r}\in\mathcal{U}_k} |\langle \mathbf{r}, \mathbf{\Sigma}\mathbf{u}\rangle|, \tag{29}$$

where the inequality follows from the definition of the set $\{\frac{\mathbf{v}}{\|\mathbf{v}\|_2} : \mathbf{v} \in \mathcal{E}_k\} \subset \mathcal{U}_k$. From Lemma 4,

$$\mathbb{E}\left[\sup_{\mathbf{v}\in\mathcal{E}_k} |\langle \mathbf{u}, \mathbf{\Sigma}\mathbf{v}\rangle|\right] \overset{(a)}{\leq} \sup_{\mathbf{v}\in\mathcal{E}_k} \|\mathbf{v}\|_2\, \mathbb{E}\left[\sup_{\mathbf{r}\in\mathcal{U}_k} |\langle \mathbf{r}, \mathbf{\Sigma}\mathbf{u}\rangle|\right] \overset{(b)}{\leq} 6\sqrt{k\frac{\rho_{\max}(k,\mathbf{p})}{\rho_{\min}(k,\mathbf{p})}\log\frac{5en}{k}}, \tag{30}$$

where (a) comes from (29) and (b) follows from Lemma 4 and the definitions in (15). ■

Step 1: Isometry on the images of sparse vectors. We first assme that the sensor data and

all matrices are real. We show that all row vectors $\widetilde{\mathbf{Z}}$ are isotropic sub-Gaussian (see Definition 7 below) in Lemma 6. Then we use Lemma 5 to obtain an isometry on the images of sparse vectors.

Definition 6 (sub-Gaussian random variables [15])**.** *Let X be a zero mean random variable that has unit variance. It is* sub-Gaussian *if for any $t \geq 0$, there exist a positive number ϱ such that*

$$\mathbb{P}\left(|X| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2\varrho^2}\right).$$

The sub-Gaussian norm $\|X\|_{\psi_2}$ *is the smallest number ϱ for which the above inequality holds.*

Definition 7 (Isotropic sub-Gaussian random vectors [15])**.** *Let $\mathbf{u}$ be a random vector in $\mathbb{R}^n$. If $\mathbb{E}[\mathbf{u}\mathbf{u}^T] = \mathbf{I}_n$, then $\mathbf{u}$ is called* isotropic*. The random vector $\mathbf{u}$ is* sub-Gaussian *with constant α if*

$$\sup_{\mathbf{r}\in\mathbb{R}^n:\|\mathbf{r}\|_2=1} \|\langle \mathbf{u}, \mathbf{r}\rangle\|_{\psi_2} < \alpha.$$

Lemma 6. Let $\mathbf{u} \in \mathbb{R}^n$ be a random vector with independent elements, each distributed as $\widetilde{\mathcal{N}}(0, 1/p_j, p_j)$. Then $\mathbf{u}$ is isotropic sub-Gaussian with constant $\alpha = c_0/\sqrt{p_{\min}}$ where c_0 is an absolute constant.

Proof: Since all elements in $\mathbf{u}$ are independent zero mean random variables, and has unit variance, we have $\mathbb{E}[\mathbf{u}\mathbf{u}^T] = \mathbf{I}_n$. Let $X_j \sim \widetilde{\mathcal{N}}(0, 1/p_j, p_j)$ be a mixed Gaussian random variable with pdf defined in (7). Then, we have for every $t \geq 0$ that

$$\mathbb{P}(|X_j| > t) = 2\int_{\sqrt{p_j}t}^{\infty} p_j \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp(-\frac{x^2}{2})dx \overset{(a)}{\leq} p_j e^{-pt^2/2} \overset{(b)}{\leq} 2e^{-p_j t^2/2},$$

where (a) follows from the Chernoff bound on Gaussian Q-function, and (b) from $p_j \in (0, 1]$. Hence, the sub-Gaussian norm of X is bounded above by $1/\sqrt{p_j}$. From Lemma 5.24 in [15], we have that the vector $\mathbf{u}$ is sub-Gaussian with constant $\alpha = c_0/\sqrt{p_{\min}}$, where c_0 is an absolute constant. ■

Recall that the signal model is $\widetilde{\mathbf{y}} = \widetilde{\mathbf{Z}}\mathbf{\Sigma}\mathbf{x}+\widetilde{\mathbf{e}}$. We note that all elements in matrix $\widetilde{\mathbf{Z}}$ are independent, and the element $\widetilde{Z}_{ij} \sim \widetilde{\mathcal{N}}(0, 1/(mp_j), p_j)$. Then Lemma 6 implies that all row vectors of scaled matrix $\sqrt{m}\widetilde{\mathbf{Z}}$ are independent, and isotropic sub-Gaussian with constant $\alpha = c_0/\sqrt{p_{\min}}$. The key idea to prove Theorem 1 is to apply one result in [19], which is given without proof as follows.

Lemma 7 (Theorem 2.1 in [19]). Set $1 \leq m \leq n$ and $0 < \beta < 1$. Let $\mathbf{b}$ be an isotropic sub-Gaussian random vector on $\mathbb{R}^n$ with constant $\alpha \geq 1$. Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be independent copies of $\mathbf{b}$. Let the random matrix $\mathbf{B}$ have rows $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Let $\mathcal{V} \subset \mathcal{S}^{n-1}$. If $m > \frac{c'\alpha^4}{\beta^2}\ell_*(\mathcal{V})^2$, then with probability at least $1 - \exp\left(-\bar{c}\beta^2 m/\alpha^4\right)$, for all $\mathbf{v} \in \mathcal{V}$, we have $1 - \beta \leq \frac{\|\mathbf{B}\mathbf{v}\|_2^2}{m} \leq 1 + \beta$, where

$c', \bar{c}$ are positive absolute constants.

Put $\mathcal{V} = \mathcal{W}(\mathcal{E}_k)$. Then from Lemmas 5, 6 and 7, we obtain the following result: If

$$m > \frac{c_1 k \rho_{\max}(k, \mathbf{p})}{p_{\min}^2 \beta^2 \rho_{\min}(k, \mathbf{p})} \log \frac{5en}{k}, \tag{31}$$

then with probability at least $1 - \exp\left(-c_2 \beta^2 p_{\min}^2 m/4\right)$, for all $\mathbf{v} \in \mathcal{E}_k$, we have $1 - \beta \leq \|\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{v}\|_2^2 \leq 1 + \beta$, where $c_1 \triangleq 36c'c_0^4$ and $c_2 \triangleq \bar{c}/c_0^4$ are positive absolute constants. Furthermore, by replacing $\mathbf{v}$ with the $\boldsymbol{\Sigma}$-normalized vector $\mathbf{v}/\|\boldsymbol{\Sigma}\mathbf{v}\|_2$, with overwhelming probability, we have

$$(1 - \beta)\|\boldsymbol{\Sigma}\mathbf{v}\|_2^2 \leq \|\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{v}\|_2^2 \leq (1 + \beta)\|\boldsymbol{\Sigma}\mathbf{v}\|_2^2. \tag{32}$$

Step 2: Restricted Isometry Property. From (32) and the definitions of the k-restricted extreme eigenvalues in (15), for any k-sparse vector $\mathbf{x}$, we obtain that the following inequality

$$(1 - \beta)\rho_{\min}(k, \mathbf{p})\|\mathbf{x}\|_2^2 \leq \|\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{x}\|_2^2 \leq (1 + \beta)\rho_{\max}(k, \mathbf{p})\|\mathbf{x}\|_2^2 \tag{33}$$

holds with probability at least $1 - \exp\left(-c_2 m p_{\min}^2 \beta^2/4\right)$.

Recall the definitions of the parameters $\xi_k(\mathbf{p})$, $\zeta_k(\mathbf{p})$, $\vartheta_k(\mathbf{p})$, δ_k, and $\beta_k(\delta_k, \mathbf{p})$ defined prior to Theorem 1. As in (33), the LHS and the RHS may have different deviations from one. Hence, the maximum operation and piecewise linear mappings are used in those definitions, such that after some simple substitutions and algebraic manipulations, the following inequality

$$(1 - \delta_k)\|\mathbf{x}\|_2^2 \leq \|\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{x}\|_2^2 \leq (1 + \delta_k)\|\mathbf{x}\|_2^2 \tag{34}$$

holds with probability at least $1 - \exp\left(-c_2 m p_{\min}^2 \beta_k^2(\delta_k, \mathbf{p})/4\right)$. Collecting the results in (31) and (34), we obtain Theorem 1 for the real case.

Step 3: Generalization to the complex case. We generalize the above RIP result to the complex case. First, we show that the matrix $\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}$ satisfies the RIP for the complex data $\mathbf{x} = \mathbf{x}^{\mathrm{R}} + j\mathbf{x}^{\mathrm{I}}$. With probability at least $1 - \exp\left(-c_2 m p_{\min}^2 \beta_k^2(\delta_k, \mathbf{p})/4\right)$, we have $(1 - \delta_k)\|\mathbf{x}^{\mathrm{R}}\|_2^2 \leq \|\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{x}^{\mathrm{R}}\|_2^2 \leq (1 + \delta_k)\|\mathbf{x}^{\mathrm{R}}\|_2^2$, and similarly for the imaginary part. This yields $(1-\delta_k)\|\mathbf{x}\|_2^2 \leq \|\widetilde{\mathbf{Z}}\boldsymbol{\Sigma}\mathbf{x}\|_2^2 \leq (1+\delta_k)\|\mathbf{x}\|_2^2$. Second, we show that when the sensing matrix $\mathbf{A}$ in our scheme is complex random matrix, it still satisfies the RIP. Let $\mathbf{A} = \mathbf{A}^{\mathrm{R}} + j\mathbf{A}^{\mathrm{I}}$. It is assumed that the real part $\mathbf{A}^{\mathrm{R}}$ and the imaginary part $\mathbf{A}^{\mathrm{I}}$ are independent, and have the same probability distribution. Recall that the sensing matrix $\mathbf{A} = \widetilde{\mathbf{Z}}\boldsymbol{\Sigma}$. For any k-sparse complex vector $\mathbf{x}$, we have $\frac{1}{2}(1 - \delta_k)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}^{\mathrm{R}}\mathbf{x}\|_2^2 \leq \frac{1}{2}(1 + \delta_k)\|\mathbf{x}\|_2^2$,

and similarly for the imaginary part. Combining the two parts yields the RIP in (14). ■

APPENDIX B
PROOF OF CLAIM IN REMARK 3

Proof: If $\mathbf{p} = \bar{p}\,\mathbf{1}_{n\times 1}$, we denote the corresponding matrix as $\bar{\mathbf{\Sigma}} \triangleq \mathbf{\Gamma\Phi}/\sqrt{P_{\text{ave}}}$. Set $\mathbf{u} = \mathbf{\Phi v}$ for some $\mathbf{v} \in \mathcal{V} \triangleq \{\mathbf{v} : \|\mathbf{v}\|_0 = k,\ \|\mathbf{v}\|_2 = 1\}$. We consider the ratio

$$R_n \triangleq \frac{\|\mathbf{\Sigma v}\|_2^2}{\|\bar{\mathbf{\Sigma}}\mathbf{v}\|_2^2} = \left(\frac{\bar{p}\sum_{j=1}^n b_j\nu_j^2}{\sum_{j=1}^n p_j b_j \nu_j^2}\right)\left(\frac{\sum_{j=1}^n p_j b_j \nu_j^2 u_j^2}{\bar{p}\sum_{j=1}^n b_j\nu_j^2 u_j^2}\right). \tag{35}$$

Denote the first and second ratios on the right of (35) as R_{1n} and R_{2n} respectively. For convenience, we also define $a_j \triangleq b_j\nu_j^2$, $a_{\max} \triangleq \sup_{j\in\mathbb{N}} a_j < \infty$ and $a_{\min} \triangleq \inf_{j\in\mathbb{N}} a_j > 0$. We use upper case $P_j \triangleq p_j$ to emphasize that the transmit powers are i.i.d. with mean $\bar{p}$ and variance σ_{p}^2. We now bound the probability that R_{1n} exceeds $1+t$ for some $t > 0$ by considering the following inequalities

$$\begin{aligned}\mathbb{P}(R_{1n} > 1+t) &= \mathbb{P}\left(\frac{1}{n}\sum_{j=1}^n (P_j - \bar{p})a_j \geq \frac{t\bar{p}}{n}\sum_{j=1}^n a_j\right) \overset{(a)}{\leq} \mathbb{P}\left(\frac{1}{n}\sum_{j=1}^n (P_j-\bar{p})a_j \geq \tau_{\min}\right)\\ &\overset{(b)}{\leq} \frac{\text{Var}\left(\sum_{j=1}^n a_j P_j\right)}{\tau_{\min}^2 n^2} \leq \frac{\sigma_{\text{p}}^2 a_{\max}^2}{\tau_{\min}^2 n},\end{aligned} \tag{36}$$

where (a) comes from the assumption $\frac{t\bar{p}}{n}\sum_{j=1}^n a_j \geq t\,\bar{p}\,a_{\min} \triangleq \tau_{\min}$ and (b) is from Chebyshev's inequality. Clearly $\mathbb{P}(R_{1n} > 1+t) \to 0$ and so does $\mathbb{P}(R_{1n} < 1-t)$. Thus $R_{1n} \to 1$ in probability. The same statement holds for R_{2n} (due to the analogous structure as R_{1n}) and thus for R_n. Recalling the definitions of $\rho_{\max}(k,\mathbf{p})$ and $\rho_{\min}(k,\mathbf{p})$ in (15), it is clear that the desired claim holds, because the two objective functions approximate each other uniformly on the compact set $\{\mathbf{\Phi v} : \mathbf{v} \in \mathcal{V}\}$. ■

APPENDIX C
PROOF OF THEOREM 2

Proof: Clearly, we have $\rho_{\max}(1,n) = \rho_{\min}(1,n) = 1$ so the bounds are satisfied for $k = 1$. We will first prove Theorem 2 for the case $k = 2$. Subsequently, we generalize the result to arbitrary $2 \leq k < \lfloor n/2 \rfloor$. Let the two non-zero elements be $v_{s_1} = A_1 e^{j\theta_1}$ and $v_{s_2} = A_2 e^{j\theta_2}$, where $A_1^2 + A_2^2 = 1$ (because $\|\mathbf{v}\|_2 = 1$). Then from (25), and the fact that $P_{\text{ave}} = \sum_{i=1}^n \gamma_i/n$, we obtain

$$\|\mathbf{w}\|_2^2 = \frac{1}{nP_{\text{ave}}}\sum_{i=1}^n \gamma_i + \frac{2A_1A_2}{nP_{\text{ave}}}\sum_{i=1}^n \gamma_i \cos\left(\theta + \frac{2\pi(i-1)\Delta}{n}\right) = 1 + 2A_1A_2\frac{\sum_{i=1}^n a_i\gamma_i}{\sum_{i=1}^n \gamma_i}, \tag{37}$$

where $\theta \triangleq \theta_1 - \theta_2 \in (0, 2\pi]$, $\Delta \triangleq s_2 - s_1 \in \{1, \ldots, n-1\}$, and $a_i \triangleq \cos(\theta + 2\pi(i-1)\Delta/n)$. We now set $X_i = \gamma_i$ to emphasize that the signal powers are random variables. Recall that the distributions of X_i's are truncated Gaussian, denoted by $\mathcal{N}_{\rm tr}(\mu, \omega)$. We consider the random variable

$$S_n \triangleq \frac{\sum_{i=1}^n a_i X_i}{\sum_{i=1}^n X_i}. \tag{38}$$

We define the Cesàro-sum of the a_i's as

$$\bar{a}_n \triangleq \frac{1}{n}\sum_{i=1}^n a_i = \frac{1}{n}\sum_{i=1}^n \cos\left(\theta + \frac{2\pi(i-1)\Delta}{n}\right), \tag{39}$$

and note that as $n \to \infty$, the Cesàro-sum converges. Indeed, $\bar{a}_n \to \bar{a} = \frac{1}{2\pi\Delta}\int_0^{2\pi\Delta} \cos(\theta + \Delta t)dt = 0$. We now bound the probability that S_n exceeds some $t > 0$ by considering the inequalities

$$\begin{aligned}
\mathbb{P}\left(S_n > t\right) &\overset{(a)}{\leq} \mathbb{P}\left(\left\{\sum_{i=1}^n a_i X_i > t\sum_{i=1}^n X_i\right\} \cap \left\{\frac{1}{n}\sum_{i=1}^n X_i > \tau\mu\right\}\right) + \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \leq \tau\mu\right) \\
&\overset{(b)}{\leq} \mathbb{P}\left(\left\{\frac{1}{n}\sum_{i=1}^n a_i X_i > t\tau\mu\right\} \cap \left\{\frac{1}{n}\sum_{i=1}^n X_i > \tau\mu\right\}\right) + \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \leq \tau\mu\right) \\
&\leq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n a_i X_i > t\tau\mu\right) + \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \leq \tau\mu\right),
\end{aligned} \tag{40}$$

where (a) is from the fact $\mathbb{P}(\mathcal{A}) \leq \mathbb{P}(\mathcal{A} \cap \mathcal{B}) + \mathbb{P}(\mathcal{B}^c)$ and (b) comes from monotonicity of measure.

Define $t' \triangleq t\tau\mu$ and let s be an arbitrary non-negative number. Then from Markov's inequality, the first term in (40) can be upper bounded as follows

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n a_i X_i > t'\right) \leq \exp(-nst')\mathbb{E}\left[\exp\left(\sum_{i=1}^n s a_i X_i\right)\right], \tag{41}$$

which implies by the independence of the X_i's that

$$\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n a_i X_i > t'\right) \leq -st' + \frac{1}{n}\sum_{i=1}^n \log\mathbb{E}[\exp(s a_i X_i)]. \tag{42}$$

To bound the sum in (42), we find the cumulant-generating function (CGF) of $X \sim \mathcal{N}_{\rm tr}(\mu, \omega^2)$ in terms of a Gaussian with mean μ and variance ω^2. By simple algebraic manipulations, we have

$$\log\mathbb{E}[\exp(sX)] = \mu s + \frac{1}{2}\omega^2 s^2 + \varphi(\mu, \omega, s), \tag{43}$$

where $\varphi(\mu, \omega, s) \triangleq \log\left(1 - Q\left(\mu/\omega + \omega s\right)\right) - \log\left(1 - Q\left(\mu/\omega\right)\right)$. Note that given the pair of positive

numbers (μ,ω), $s\mapsto\varphi(\mu,\omega,s)$ for $s\geq 0$ is concave, because $s\mapsto -Q(\mu/\omega+\omega s)$ (for $\mu/\omega>0$) and $t\mapsto\log(1+t)$ are both concave and the latter is non-decreasing. Moreover, $s\mapsto\varphi(\mu,\omega,s)$ is continuous for each positive (μ,ω) pair, because every concave function on an open set is continuous.

Substituting the CGF of the truncated Gaussian distribution in (43) into (42) yields

$$\begin{aligned}
&\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}a_iX_i>t'\right)\\
&\leq -st'+\mu s\bar{a}_n+\frac{\omega^2s^2}{2n}\sum_{i=1}^{n}a_i^2+\frac{1}{n}\sum_{i=1}^{n}\varphi(\mu,\omega,a_is)\\
&\overset{(a)}{=} -st'+\mu s\bar{a}_n+\frac{\omega^2s^2}{4}+\frac{\omega^2s^2}{4n}\sum_{i=1}^{n}\cos\left(2\theta+\frac{4\pi\Delta(i-1)}{n}\right)+\frac{1}{n}\sum_{i=1}^{n}\varphi(\mu,\omega,a_is)\\
&\overset{(b)}{\leq} -st'+\mu s\bar{a}_n+\frac{\omega^2s^2}{4}+\frac{\omega^2s^2}{4n}\sum_{i=1}^{n}\cos\left(2\theta+\frac{4\pi\Delta(i-1)}{n}\right)+\varphi\left(\mu,\omega,\frac{s}{n}\sum_{i=1}^{n}a_i\right),
\end{aligned}\tag{44}$$

where (a) comes from the definition of a_i and the double-angle formula for the cosine, and (b) follows the fact $\varphi(\mu,\omega,s)$ is concave in s for any positive (μ,ω) pair.

Taking the limsup on both sides of (44) and using the definition of $\bar{a}_n$ yields

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}a_iX_i>t'\right)\\
&\overset{(a)}{\leq} -st'+\frac{\omega^2s^2}{4}+\frac{\omega^2s^2}{16\pi\Delta}\int_0^{4\pi\Delta}\cos\left(2\theta+t\right)dt+\limsup_{n\to\infty}\varphi\left(\mu,\omega,\bar{a}_ns\right)\\
&\overset{(b)}{=} -st'+\frac{\omega^2s^2}{4}+\limsup_{n\to\infty}\varphi\left(\mu,\omega,\bar{a}_ns\right)\overset{(c)}{=}-st'+\frac{\omega^2s^2}{4}\triangleq f(s),
\end{aligned}\tag{45}$$

where (a) follows from Riemann sums, (b) comes from the fact cosine has zero mean over an integer number of periods (note $\Delta\in\mathbb{Z}$) and (c) follows from the continuity of $\varphi(\mu,\omega,s)$ and $\bar{a}_n\to 0$. Note that the minimum $f(s)$ in (45) is $f(s^*)=-\tau^2d^2t^2$ (attained at $s^*=2t'/\omega^2$). Hence,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}a_iX_i>t'\right)\dot{\leq}\exp\left[-n\tau^2d^2t^2\right].\tag{46}$$

The second term in (40) can be bounded using standard techniques from the large deviations theory [21] (Cramér's theorem) and along the same lines as the derivation above. As such we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\leq\tau\mu\right)\overset{(a)}{\leq}\exp\left[-n\left(s\mu(1-\tau)-\omega^2s^2/2-\varphi(\mu,\omega,-s)\right)\right]$$

$$\overset{(b)}{\leq} \exp\left[-n\left(s\mu(1-\tau) - \omega^2 s^2/2\right)\right],$$

where (a) follows from using the CGF of X_i in (43), and (b) follows from the fact that $\varphi(\mu, \omega, -s) \leq 0$ for all $s \geq 0$. Hence, setting $s \triangleq \mu(1-\tau)/\omega^2$, we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i \leq \tau\mu\right) \dot{\leq} \exp\left[-n(1-\tau)^2 d^2/2\right]. \tag{47}$$

Using the largest-exponent-dominates principle, we have from (40), (46) and (47) that

$$\mathbb{P}(S_n > t) \dot{\leq} \exp\left[-n \min\left\{\tau^2 d^2 t^2, (1-\tau)^2 d^2/2\right\}\right]. \tag{48}$$

Since $\tau > 0$ is a free parameter, we can set it to be $\tau^* \triangleq \frac{1}{1+\sqrt{2}t}$. Substituting τ^* into (48) yields

$$\mathbb{P}\left(S_n > t\right) \dot{\leq} -nd^2\tilde{t}^2, \tag{49}$$

where $\tilde{t} \triangleq t/(1+\sqrt{2}t)$ and similarly for $\mathbb{P}\left(S_n < -t\right)$ by symmetry.

Recall that $\rho_{\max}(k,n)$ is the maximum value of $\|\mathbf{w}\|_2^2 = \|\mathbf{\Sigma v}\|_2^2$ over all unit-norm k-sparse vectors $\mathbf{v}$. From (37), $\|\mathbf{w}\|_2^2$ depends only on A_1A_2. Note that $0 < A_1A_2 \leq 1/2$ because $\sqrt{A_1A_2} \leq (A_1+A_2)/2$. We set $A_1A_2 = 1/2$, whence $\|\mathbf{w}\|_2$ attains its maximum value. From (37),

$$\mathbb{P}\left(\rho_{\max}(2,n) > 1+t\right) \dot{\leq} \exp\left[-nd^2\tilde{t}^2\right], \tag{50}$$

and similarly for the probability $\mathbb{P}\left(\rho_{\min}(2,n) < 1-t\right)$.

We now generalize it to the case where $k > 2$. Set the non-zero elements of the vector $\mathbf{v}$ to be $v_{s_q} = A_q e^{j\theta_q},\ q = 1, \ldots, k$, where $\sum_{q=1}^{k} A_q^2 = 1$. Equation (25) can be written as

$$\begin{aligned}
\|\mathbf{w}\|_2^2 &= \frac{1}{nP_{\text{ave}}}\sum_{i=1}^{n}\gamma_i\left(1 + \sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k} A_qA_l\cos\left(\theta_{q,l} + \frac{j2\pi(i-1)\Delta_{q,l}}{n}\right)\right) \\
&= 1 + \sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k} A_qA_l\sum_{i=1}^{k}\cos\left(\theta_{q,l} + \frac{2\pi(i-1)\Delta_{q,l}}{n}\right), \\
&= 1 + \sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k} A_qA_lS_n^{q,l} \triangleq 1 + B_n,
\end{aligned} \tag{51}$$

where $S_n^{q,l}$ is defined as in (38) but involving the q-th and the l-th nonzero elements of $\mathbf{v}$, i.e.,

$\theta_{q,l} = \theta_q - \theta_l$, and $\Delta_{q,l} = s_l - s_q$. On the other hand, we can bound B_n^2 as follows

$$B_n^2 \overset{(a)}{\leq} \left(\sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k} A_q^2 A_l^2\right)\left(\sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}(S_n^{q,l})^2\right) = \left(\left(\sum_{q=1}^{k} A_q^2\right)^2 - \sum_{q=1}^{k} A_q^4\right)\left(\sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}(S_n^{q,l})^2\right)$$
$$\overset{(b)}{\leq} \left(1 - \frac{1}{k}\left(\sum_{q=1}^{k} A_q^2\right)^2\right)\left(\sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}(S_n^{q,l})^2\right) = \frac{k-1}{k}\sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}(S_n^{q,l})^2, \tag{52}$$

where (a) comes from the Cauchy-Schwartz inequality and (b) comes from the basic inequality relating the arithmetic and quadratic means, namely $1/M\sum_{j=1}^{M}\alpha_j \leq (1/M\sum_{j=1}^{M}\alpha_j^2)^{1/2}$. Now, given any $t > 0$, we can bound the probability that $|B_n|$ exceeds t as follows

$$\mathbb{P}\left(B_n^2 > t^2\right) \overset{(a)}{\leq} \mathbb{P}\left(\sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}(S_n^{q,l})^2 > \frac{kt^2}{k-1}\right) \leq \mathbb{P}\left(\max_{l\neq q}(S_n^{q,l})^2 > \frac{t^2}{(k-1)^2}\right)$$
$$\overset{(b)}{\leq} \sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}\mathbb{P}\left(\left(S_n^{q,l}\right)^2 > \frac{t^2}{(k-1)^2}\right) = \sum_{q=1}^{k}\sum_{l=1,l\neq q}^{k}\mathbb{P}\left(S_n^{q,l} > \frac{t}{k-1}\right), \tag{53}$$

where (a) comes from (52) and monotonicity of measure and (b) comes from the union bound. Set $E(k,t) \triangleq t/(k-1+\sqrt{2}t)$. Applying the result for $k = 2$ in (50) to (53), we have

$$\mathbb{P}\left(|B_n| > t\right) \dot{\leq} k(k-1)\exp\left[-nd^2E(k,t)^2\right]. \tag{54}$$

Recall the definition of $\rho_{\max}(k,n)$ in (15). From (51) and (54), we conclude that

$$\mathbb{P}\left(\rho_{\max}(k,n) > 1+t\right) \dot{\leq} \exp\left[-nd^2E(k,t)^2\right]. \tag{55}$$

The analysis of $\mathbb{P}\left(\rho_{\min}(k,n) < 1-t\right)$ proceeds *mutatis mutandis*. This completes the proof. ■

REFERENCES

[1] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 6, Sept. 2007.

[2] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal Process.*, vol. 60, pp. 4808–4818, Sept. 2012.

[3] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, Mar. 2008.

[4] J. D. Haupt and R. D. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inf. Theory*, vol. 52, pp. 4036–4048, Sept. 2006.

[5] S. Aeron, M. Zhao, and V. Saligrama, "Information theoretic bounds to sensing capacity of sensor networks under fixed SNR," in *IEEE Inf. Th. Workshop*, (Lake Tahoe, CA, USA), pp. 84–89, Sept. 2007.

[6] R. Rana, W. Hu, and C. T. Chou, "Energy-aware sparse approximation technique (EAST) for rechargeable wireless sensor networks," in *European Conf. on Wireless Sensor Networks*, (Coimbra, Portugal), pp. 306–321, Feb. 2010.

[7] T. Xue, X. Dong, and Y. Shi, "A multiple access scheme based on multi-dimensional compressed sensing," in *IEEE Int. Conf. on Commun. (ICC)*, (Ottawa, Canada), pp. 3832–3836, Jun. 2012.

[8] F. Fazel, M. Fazel, and M. Stojanovic, "Random access compressed sensing for energy-efficient underwater sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 29, pp. 1660–1670, Sept. 2011.

[9] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 745–754, April 2005.

[10] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1147–1158, May 2003.

[11] W. Bajwa, *New information processing theory and methods for exploiting sparsity in wireless systems*. PhD thesis, University of Wisconsin-Madison, 2009.

[12] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Hanover, MA, USA: now Publishers Inc. Press, 2004.

[13] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, pp. 4203–4215, Dec. 2005.

[14] R. Baraniuk, M. Davenport, R. D. Vore, and M. Wakin, "A simple proof of the restricted isometry property," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.

[15] Y. C. Eldar and G. Kutyniok, *Compressed sensing: Theory and applications*. Cambridge Univ. Press, 2012.

[16] T. T. Cai, M. Wang, and G. Xu, "New bounds for restricted isometry constants," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4388–4394, Sept. 2010.

[17] M. A. Davenport, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range," *IEEE Trans. Signal Process.*, vol. 60, pp. 4628–4642, Sept. 2012.

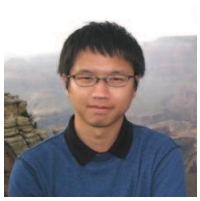
[18] E. V. D. Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *Proc. of Soc. Ind. Appl. Math.*, vol. 31, no. 2, pp. 890–912, 2008.

[19] S. Mendelson, A. Pajor, and N. T. Jaegermann, "Uniform uncertainty principle for Bernoulli and sub-Gaussian ensembles," *Constr. Approx.*, vol. 28, pp. 277–289, 2008.

[20] S. Zhou, "Restricted eigenvalue conditions on sub-Gaussian random matrices." Website, 2009. http://arxiv.org/abs/0912.4045v2.

[21] A. S. Dembo and O. Zeitouni, *Large deviation techniques and applications*. Springer Press, 1998.

Gang Yang (S'13) received the B.Eng. and M.Eng. degrees (First-Class Honors) in Communication Engineering, Communication and Information Systems in 2008, 2011 respectively from University of Electronic Science and Technology of China (UESTC), Chengdu, China. He is currently pursuing a Ph.D. degree in Infocomm Centre of Excellence of the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, since 2011. His current research interests include green wireless communications with energy harvesting constraints, wireless information and power transfer, compressive sensing and statistical signal processing. He is a student member of IEEE and a member of IEEE Communications Society.



Vincent Y. F. Tan (S'07-M'11) received the B.A. and M.Eng. degrees in Electrical and Information Engineering from Sidney Sussex College, Cambridge University, in 2005. He received the Ph.D. degree in Electrical Engineering and Computer Science (EECS) from the Massachusetts Institute of Technology (MIT) in 2011. From 2011 to 2012, he was a postdoctoral researcher in the Electrical and Computer Engineering Department at the University of Wisconsin-Madison. He is now a scientist at the Institute for Infocomm Research (I^2R), Singapore, and an adjunct assistant professor in the Department of Electrical and Computer Engineering at the National University of Singapore. He has held summer research internships at Microsoft Research. His research interests include network information theory, detection and estimation, and learning and inference of graphical models.

Dr. Tan is a recipient of the 2005 Charles Lamb Prize, a Cambridge University Engineering Department prize awarded annually to the top candidate in Electrical and Information Engineering. He also received the 2011 MIT EECS Jin-Au Kong outstanding doctoral thesis prize. He is a member of the IEEE and of the IEEE Machine Learning for Signal Processing Technical Committee. He has served as a reviewer for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON INFORMATION THEORY, and the Journal of Machine Learning Research.

Chin Keong Ho (S'05-M'07) received the B.Eng. (First-Class Hons., Minor in Business Admin.), and M. Eng degrees from the Department of Electrical Engineering, National University of Singapore in 1999 and 2001, respectively. He obtained his Ph.D. degree at the Eindhoven University of Technology, The Netherlands, where he concurrently conducted research work in Philips Research. Since August 2000, he has been with Institute for Infocomm Research (I^2R), A*STAR, Singapore. He is currently Lab Head of Energy-Aware Communications Lab, Department of Modulation and Coding, in I^2R. His research interest includes green wireless communications with focus on energy-efficient solutions and with energy harvesting constraints; cooperative and adaptive wireless communications; and implementation aspects of multi-carrier and multi-antenna communications.

See Ho Ting (S'02-M'07) received his B.Eng., M.Eng. and Ph.D. degrees in electrical and electronic engineering in 2002, 2004 and 2006 respectively from Tokyo Institute of Technology, Japan. From April 2006 to August 2012, he was an assistant professor in Nanyang Technological University (NTU), Singapore. He is currently an associate professor in NTU.

Dr. Ting received the Young Researcher Encouragement Award from IEEE Vehicular Technology Society (Japan Chapter) in 2002, IEICE Outstanding Paper Award and Ericsson Young Scientist Award in 2005. In 2007, and again in 2010, his team won 1st Prize in the ASEAN Virtual Instrumentation Applications Contest. In 2010, Dr. Ting was awarded the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award. He is also a certified IEEE Wireless Communication Professional. His current research interests include cognitive radios, cooperative communications, wireless network coding and MIMO-OFDM systems.

Yong Liang Guan (M'99) received his Ph.D. degree from the Imperial College of Science, Technology and Medicine, University of London, in 1997, and B.Eng. degree with first class honors from the National University of Singapore in 1991. He is now an associate professor at the School of Electrical and Electronic Engineering, and the Director of Infocomm Centre of Excellence (http://www.infinitus.eee.ntu.edu.sg), Nanyang Technological University. His research interests include modulation, coding and communication signal processing, digital watermarking for information security, and channel modeling. [http://www3.ntu.edu.sg/home/eylguan]