MASSIVE DEVICE ACTIVITY DETECTION BY APPROXIMATE MESSAGE PASSING

Zhilin Chen and Wei Yu

Department of Electrical and Computer Engineering
University of Toronto, Toronto, ON, Canada M5S 3G4
Emails: {zchen, weiyu}@comm.utoronto.ca

ABSTRACT

User activity detection is a central problem in massive device communication scenarios in which an access point needs to detect active devices among large number of potential devices each transmitting sporadically. By exploiting sparsity in user activity, the detection problem can be formulated as a compressed sensing problem, thereby allowing the use of computationally efficient approximate message passing (AMP) algorithm for activity detection. This paper proposes an AMP-based user activity detector that accounts for the statistics of device geographic locations in a cellular network. The proposed scheme is based on a minimum mean squared error (MMSE) denoiser designed for specific wireless channel fading and path-loss distributions. This paper further provides an analytic characterization of the false alarm versus missed detection probabilities using state evolution for AMP. Simulation results show significantly improved detection threshold for the channel-aware denoiser as compared to standard soft threshold based AMP.

1. INTRODUCTION

Massive connectivity is a key requirement for future wireless cellular network in which a large number of devices (e.g., sensors) may communicate simultaneously to a central base-station (BS). A salient feature of these machine type communications is that the traffic pattern from each of these devices is typically sporadic, thus at any give time only a small fraction of potentially large number of devices are active. Accurate user activity detection is therefore a central component of the overall system design for massive connectivity.

This paper considers a pilot-based transmission system whereas active users synchronously transmit non-orthogonal signature sequences during a contention phase, and the BS detects the active users among a large pool of potential devices based on a linear combination of their signatures. Due to the *sparse* nature of the user transmission pattern, user activity detection can be formulated as a compressed sensing problem. The goal of this paper is to study the design and analysis of computationally efficient approximate message passing (AMP) algorithm for massive device detection.

Specifically, this paper proposes the design of a minimum mean squared error (MMSE) denoiser for AMP that exploits the distribution of user locations in a cellular network, as well as the fading statistics of wireless channels for the active users. Further, this paper shows that for the activity detection problem, the tradeoff between the false alarm probability and the missed detection probability can be analytically characterized via state evolution for AMP. Simulation results show that as compared to the conventional soft threshold denoiser based AMP, the proposed MMSE denoiser can significantly improve the detection threshold.

The user activity detection problem for massive connectivity has been studied in the literature from a compressed sensing perspective [1–7]. For example, [1] analyzes condition for reliable detection and [2] studies signature design, both without considering channel estimation. In [3–7], the joint user activity detection and channel estimation problem is studied. Specifically, [3] proposes a greedy algorithm based on orthogonal matching pursuit. By exploiting the statistics of channel path-loss, [6] proposes a modified Bayesian compressed sensing algorithm in a cloud radio-access network. In the context of cellular networks, [5] adopts the basis pursuit denoising method, and provides a detection error bound based on the restricted isometry property. The performance of such scheme in a practical setting is illustrated in [4, 5]. In contrast to these previous works, this paper adopts the computationally more efficient AMP algorithm for sparse recovery, which is more suitable for large-scale network with large number of devices. Although the use of AMP for activity detection has been previously proposed in [7], this paper makes further progress by exploiting channel statistics to design an MMSE-denoiser-based AMP that significantly outperforms conventional compressed sensing methods. Moreover, this paper provides an analytic detector performance characterization using state evolution.

2. SYSTEM MODEL AND AMP ALGORITHM

2.1. System Model

Consider a cellular network with one BS and N potential devices, but in each coherence block only a subset of users are active. Let $a_n \in \{1, 0\}$ indicate whether or not user n is active. For the purpose of channel probing and user identification, each user is assigned to a unique pilot sequence $\mathbf{s}_n = [s_{1n}, s_{2n}, \cdots, s_{Ln}]^T \in \mathbb{C}^{L\times 1}$, where L is the length of the pilot sequence. We consider a block fading channel model where the channel is static in each block. Assuming that the BS and the users are all equipped with a single antenna each, the received signal at the BS can be modeled as

$$\mathbf{y} = \sum_{n=1}^{N} h_n a_n \mathbf{s}_n + \mathbf{w} \triangleq \mathbf{S}\mathbf{x} + \mathbf{w} \tag{1}$$

where $h_n \in \mathbb{C}$ is the channel coefficient between user n and the BS, $\mathbf{w} \in \mathbb{C}^{L\times 1}$ is the effective complex Gaussian noise whose variance σ_w^2 depends on the background noise power and the transmit power, $\mathbf{x} \triangleq [x_1, x_2, \cdots, x_N]^T \in \mathbb{C}^{N\times 1}$ where $x_n \triangleq h_n a_n$, and $\mathbf{S} \triangleq [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_N] \in \mathbb{C}^{L\times N}$.

The objective is to detect the active users, i.e., the nonzero entries of the vector $\mathbf{x}$ based on $\mathbf{y}$. We are interested in the regime where N is much larger than L, so that the user pilot sequences cannot be mutually orthogonal; but due to sporadic traffic, only a small number of devices transmit each block, resulting in a sparse $\mathbf{x}$. Therefore, the recovering of $\mathbf{x}$ falls into the compressed sensing paradigm. In this paper, we further assume that all pilot sequences are generated from i.i.d. complex Gaussian distribution with zero mean and variance $\frac{1}{L}$ such that each sequence has unit power.

A key observation of this paper is that the design of compressed sensing algorithm can be significantly enhanced by accounting for the knowledge of the statistics of $\mathbf{x}$. Toward this end, we assume that each user accesses the channel with probability λ in an i.i.d. fashion, i.e., $\Pr(a_n = 1) = \lambda, \forall n$, so that the entries of $\mathbf{x}$ follow an i.i.d. mixture distribution:

$$p_X = (1 - \lambda)\delta_0 + \lambda p_H \tag{2}$$

where δ_0 denotes the point mass measure at zero, and p_H denotes the distribution of the channel coefficient h_n. Further, we assume that the users are randomly and uniformly located in a circular coverage area of radius R around the BS, and the channel coefficients between the users and the BS follow an independent distribution that depends on the distance between the user and the BS. More specifically, p_H includes path-loss, shadowing and Rayleigh fading. The path-loss between each user and the BS is modeled (in dB) as $\alpha + \beta \log_{10}(d)$, where distance d is measured in meter, α is the fading coefficient at $d = 1$ and β is the pathloss exponent. The shadowing (in dB) follows a Gaussian distribution with zero mean and variance σ_{SF}^2, and the Rayleigh fading component is distributed as complex Gaussian random variable $\mathcal{CN}(0, 1)$.

2.2. AMP Algorithm

This paper adopts an iterative algorithm known as AMP, originally proposed in [8] and since extended along several directions [9–11], for detecting the non-zero entries of $\mathbf{x}$. Starting with $\mathbf{x}^0 = 0$ and $\mathbf{r}^0 = \mathbf{y}$, AMP proceeds at each iteration as

$$\mathbf{x}^{t+1} = \eta_t(\mathbf{S}^*\mathbf{r}^t + \mathbf{x}^t), \tag{3}$$

$$\mathbf{r}^{t+1} = \mathbf{y} - \mathbf{S}\mathbf{x}^{t+1} + \delta^{-1}\mathbf{r}^t\langle \eta_t'(\mathbf{S}^*\mathbf{r}^t + \mathbf{x}^t)\rangle \tag{4}$$

where $t = 0, 1, \cdots$ is the index of iteration, $\mathbf{x}^t$ is the estimate of $\mathbf{x}$ at iteration t, $\mathbf{r}^t$ is the residual, $\mathbf{S}^*$ is the conjugate transpose of $\mathbf{S}$, $\eta_t(\cdot)$ is a sequence of appropriately designed nonlinear functions known as *denoisers*, $\eta_t'(\cdot)$ is the first order derivative of $\eta_t(\cdot)$, $\delta \triangleq \frac{L}{N}$, and $\langle\cdot\rangle$ is the averaging operation over all entries of a vector.

Intuitively, the AMP algorithm performs successive denoising on the matched filtered output $\tilde{\mathbf{x}}^t \triangleq \mathbf{S}^*\mathbf{r}^t + \mathbf{x}^t$, which can be modeled as signal $\mathbf{x}$ plus noise, i.e., $\tilde{\mathbf{x}}^t = \mathbf{x} + \mathbf{v}^t$. The denoiser is typically designed to minimize MSE at each iteration. In the compressed sensing literature, the prior distribution of $\mathbf{x}$ is usually assumed to be unknown. In this case, a minimax framework over the worse case $\mathbf{x}$ leads to a soft thresholding denoiser [8, 12]. When the prior distribution of $\mathbf{x}$ is known, the Bayesian framework then can be used to account for the prior information [13]. In this paper we adopt the Bayesian approach to design the MMSE denoiser.

The AMP algorithm can be analyzed in the asymptotic regime where $L, N \to \infty$ with fixed $\frac{L}{N}$ [8]. The state evolution predicts the per-coordinate performance of the AMP algorithm at each iteration as follows

$$\tau_{t+1}^2 = \sigma_w^2 + \delta^{-1}\mathbb{E}|\eta_t(X + \tau_t V) - X|^2 \tag{5}$$

where τ_t is referred to as the *state*, X and V are random variables with X following p_X and V following the standard complex Gaussian distribution, and the expectation is taken over both X and V. The random variables X, V and $\tilde{X}^t \triangleq X + \tau_t V$ capture the distributions of the entries of $\mathbf{x}$, $\mathbf{v}^t$ (up to a factor τ_t), and $\tilde{\mathbf{x}}^t$, respectively, with $\mathbb{E}|\eta_t(\tilde{X}^t) - X|^2$ characterizing the per-coordinate MSE of the estimate of $\mathbf{x}$ at iteration t. State evolution can be applied to a general family of AMP algorithms for any Lipschitz-continuous $\eta_t(\cdot)$ [14].

3. AMP BASED USER ACTIVITY DETECTION

3.1. MMSE Denoiser for AMP Algorithm

We now use the Baysian approach to design the denoiser $\eta_t(\cdot)$ for the user activity detection problem by minimizing the MSE. The MMSE denoiser is given by the conditional expectation $\mathbb{E}[X|\tilde{X}^t]$ at each iteration as function of τ_t. As X follows mixture distribution (2), we first characterize the channel fading component p_H accounting for the uniform distribution of user locations in the cell as well as path-loss, shadowing, and Rayleigh fading

$$p_H(h) = a\int_0^\infty z^{-\gamma-2} Q(b \ln z + c) \exp\left(\frac{-|h|^2}{z^2}\right) dz \tag{6}$$

where $Q(z) \triangleq \frac{2}{\sqrt{\pi}} \int_z^{\infty} \exp(-s^2) ds$, $\gamma \triangleq \frac{40}{\beta} + 1$, and a, b, and c are constants as shown below

$$a = \frac{20}{\pi R^2 \beta} \exp\left(\frac{2(\ln 10)^2 \sigma_{\rm SF}^2}{\beta^2} - \frac{2(\ln 10)\alpha}{\beta} \right),$$
$$b = \frac{-10\sqrt{2}}{(\ln 10)\sigma_{\rm SF}}, \quad c = -\frac{\alpha + \beta \log_{10} R}{\sqrt{2}\sigma_{\rm SF}} - \frac{20}{\beta b}.$$

where $\alpha, \beta, \sigma_{\rm SF}, R$ are defined in system model. To simplify (6), we choose to ignore shadowing and approximate (6) as

$$p_H(h) \approx a' \int_{\epsilon}^{\infty} z^{-(\gamma+2)} \exp\left(-\frac{|h|^2}{z^2} \right) dz \tag{7}$$

where $a' \triangleq \frac{40}{\pi R^2 \beta} 10^{-2\alpha/\beta}$ and $\epsilon \triangleq 10^{-(\alpha+\beta \log_{10} R)/20}$ are constants. Based on (7), the MMSE estimator is derived as

$$\begin{aligned} \mathbb{E}[X|\tilde{X}^t] &= \int x p_{X|\tilde{X}^t}(x|\tilde{x}^t) dx \\ &= \frac{\tilde{x}^t}{\xi(\tilde{x}^t)} \int_{\epsilon}^{\infty} \frac{z^{2-\gamma}}{(z^2+\tau_t^2)^2} \exp\left(\frac{-|\tilde{x}^t|^2}{z^2+\tau_t^2} \right) dz \end{aligned} \tag{8}$$

where

$$\xi(\tilde{x}^t) = \frac{1-\lambda}{\lambda a' \pi \tau_t^2} \exp\left(\frac{-|\tilde{x}^t|^2}{\tau_t^2} \right) + \int_{\epsilon}^{\infty} \frac{z^{-\gamma} \exp\left(\frac{-|\tilde{x}^t|^2}{z^2+\tau_t^2} \right)}{z^2+\tau_t^2} dz. \tag{9}$$

Note that $\eta_t^{\rm mmse}(\tilde{x}^t)$ depends on τ_t. In practice, an empirical estimate $\hat{\tau}_t = \frac{1}{\sqrt{L}} \|\mathbf{r}^t\|_2$ can be used [15]. Although $\eta_t^{\rm mmse}(\tilde{x}^t)$ is in a complicated form, it can be pre-computed, so it does not add to run-time complexity. To gain some intuition, we illustrate the shape of $\eta_t^{\rm mmse}(\tilde{x}^t)$ as well as the soft thresholding in Fig. 1. Observe that the MMSE denoiser plays a role similar to the soft thresholding denoiser, shrinking the input towards the origin, especially when the input is small, thereby promoting sparsity.

3.2. Likelihood Ratio Test for Activity Detection

After recovering an estimate of $\mathbf{x}$ via the AMP, we employ the likelihood ratio test to perform user activity detection. For the hypothesis test problem

$$\begin{cases} H_0 : X = 0, \ \text{inactive user}, \\ H_1 : X \neq 0, \ \text{active user}; \end{cases} \tag{10}$$

the optimal decision rule is given by

$$LLR = \log \left(\frac{p_{\tilde{X}^t|X}(\tilde{x}^t|x \neq 0)}{p_{\tilde{X}^t|X}(\tilde{x}^t|x = 0)} \right) \underset{H_1}{\overset{H_0}{\lessgtr}} l_{th} \tag{11}$$

where LLR denotes the log-likelihood ratio, and l_{th} denotes the decision threshold typically determined by a cost function.

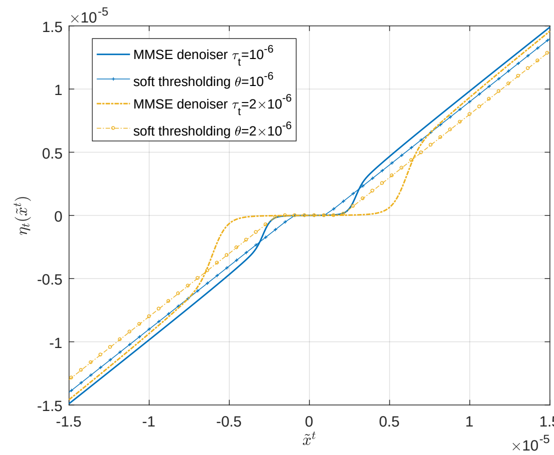

Fig. 1. MMSE vs. soft thresholding denoiser [10] $\eta_t^{\rm soft}(\tilde{x}^t) \triangleq (\tilde{x}^t - \frac{\theta \tilde{x}^t}{|\tilde{x}^t|}) \mathbb{I}(|\tilde{x}^t| > \theta)$, where $\mathbb{I}(\cdot)$ is the indicator function.

Since $\tilde{X}^t \triangleq X + \tau_t V$ in AMP, where V follows a complex Gaussian distribution asymptotically, we can write down the likelihood of $\tilde{X}^t$ given $X = 0$ or $X \neq 0$ in the asymptotic regime, while accounting for path-loss, shadowing, and fading components, as

$$p_{\tilde{X}^t|X}(\tilde{x}^t|x=0) = \frac{1}{\pi \tau_t^2} \exp\left(-\frac{|\tilde{x}^t|^2}{\tau_t^2} \right) \tag{12}$$

$$p_{\tilde{X}^t|X}(\tilde{x}^t|x \neq 0) = a \int_0^{\infty} \frac{Q(b \ln z + c)}{z^{\gamma}(z^2+\tau_t^2)} \exp\left(\frac{-|\tilde{x}^t|^2}{z^2+\tau_t^2} \right) dz \tag{13}$$

The log-likelihood ratio is then given as

$$LLR = \log \int_0^{\infty} \frac{a \pi \tau_t^2 z^{-\gamma}}{z^2+\tau_t^2} Q(b \ln z + c) \exp(|\tilde{x}^t|^2 \Delta) dz \tag{14}$$

where $\Delta \triangleq \frac{1}{\tau_t^2} - \frac{1}{z^2+\tau_t^2}$. By observing that LLR is monotonic in $|\tilde{x}^t|$, we can simplify the decision rule in (11) as $|\tilde{x}^t| \underset{H_1}{\overset{H_0}{\lessgtr}} l'_{th}$, i.e., user activity detection can be performed based on the magnitude of $\tilde{x}^t$ only.

3.3. Performance Analysis

The fact that the estimation error in AMP can be asymptotically modeled as Gaussian allows analytic characterization of false alarm and the missed detection probabilities for user activity detection as follows:

$$P_F = \int_{|\tilde{x}^t| > l'_{th}} p_{\tilde{X}^t|X}(\tilde{x}^t|x=0) d\tilde{x}^t \tag{15}$$

$$P_M = \int_{|\tilde{x}^t| < l'_{th}} p_{\tilde{X}^t|X}(\tilde{x}^t|x \neq 0) d\tilde{x}^t \tag{16}$$

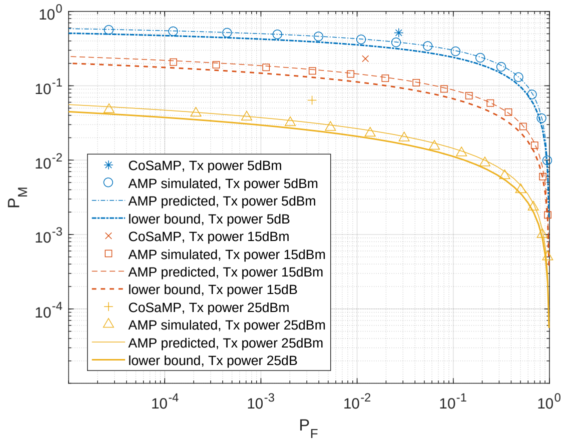

Fig. 2. Performance of AMP algorithm with MMSE denoiser

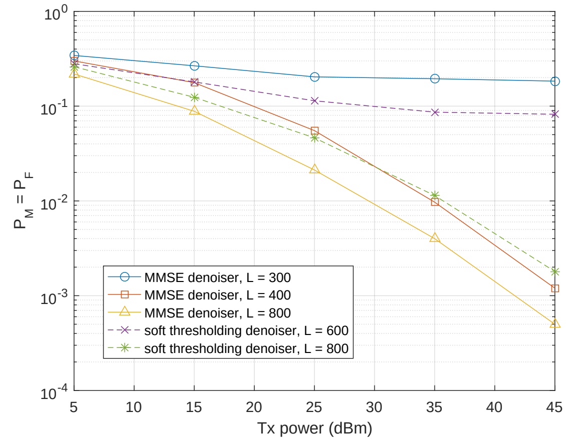

Fig. 3. Impact of transmit power and pilot length

Algorithm 1 AMP Based User Activity Detection

Input and initialization: $\mathbf{y}$, $\mathbf{S}$, l'_{th}, $\mathbf{x}^0 = \mathbf{0}$, and $\mathbf{r}^0 = \mathbf{y}$.
Repeat
1) Update the estimate of $\mathbf{x}$ as in (3) with $\eta_t^{\text{mmse}}(\cdot)$.
2) Update the residual as in (4) with $\eta_t^{\text{mmse}}(\cdot)$.
Until convergence
3) Compute $\tilde{\mathbf{x}}^t = \mathbf{S}^*\mathbf{r}^t + \mathbf{x}^t$.
4) For each entry of $\tilde{\mathbf{x}}^t$, perform LLR test based on l'_{th}.

where parameter τ_t converges to τ_∞ as the algorithm converges. To compute τ_∞, we use the state evolution (5), where $\mathbb{E}|\eta_t(X + \tau_t V) - X|^2$ in (5) can be interpreted as the MSE of the denoiser, given as

$$MSE^{\text{mmse}}(\tau_t) = \int var(X|\tilde{X}^t)p_{\tilde{X}^t}(\tilde{x}^t)d\tilde{x}^t \tag{17}$$

where

$$\begin{aligned} var(X|\tilde{X}^t) =& \mathbb{E}[|X - \mathbb{E}[X|\tilde{X}^t]|^2 \mid \tilde{X}^t] \\ =& \frac{|\tilde{x}^t|^2\mu(\tilde{x}^t) + \tau_t^2\nu(\tilde{x}^t)}{\xi(\tilde{x}^t)} - \frac{|\tilde{x}^t|^2\nu^2(\tilde{x}^t)}{\xi^2(\tilde{x}^t)} \end{aligned} \tag{18}$$

and

$$\mu(\tilde{x}^t) = \int_\epsilon^\infty \frac{z^{4-\gamma}}{(z^2+\tau_t^2)^3} \exp\left(-\frac{|\tilde{x}^t|^2}{z^2+\tau_t^2}\right) \tag{19}$$

$$\nu(\tilde{x}^t) = \int_\epsilon^\infty \frac{z^{2-\gamma}}{(z^2+\tau_t^2)^2} \exp\left(-\frac{|\tilde{x}^t|^2}{z^2+\tau_t^2}\right) \tag{20}$$

Then, based on (5), we have $\tau_\infty^2 = \sigma_w^2 + \frac{1}{\delta}MSE^{\text{mmse}}(\tau_\infty)$.

4. SIMULATION RESULTS

We evaluate the performance of the proposed user activity detection method in a cell of radius $R = 1000$m with potential $N = 4000$ users among which 200 are active, i.e., $\lambda = 0.05$. The channel fading parameters are $\alpha = 15.3, \beta = 37.5$ and $\sigma_{\text{SF}} = 8$, and the background noise is -169dBm/Hz over 10MHz. Fig. 2 shows the tradeoff between the missed detection and the false alarm probabilities of AMP with MMSE denoiser with pilot sequence length $L = 800$. We see that the predicted P_M and P_F match the analysis very well, validating the approximation in (7) for ignoring shadowing. We also plot a lower bound using $\tau_\infty = \sigma_w$. The lower bounds are very close to the actual performance, indicating that after convergence AMP is able to almost completely eliminate multiuser interference; the remaining error is dominated by the background noise. For comparison, we also plot the performance of the widely used CoSaMP algorithm [16]. We observe that AMP outperforms CoSaMP due partly to the fact that CoSaMP does not exploit channel distribution information. Note that performing user activity detection by solving the LASSO problem via direct convex optimization would have been too computationally complex.

Fig. 3 compares AMP with MMSE denoiser with soft thresholding denoiser as function of transmit power and pilot length. For convenience, we set $P_F = P_M$ by properly choosing l'_{th}. Observe first that the MMSE denoiser outperforms soft thresholding significantly, but more importantly, the phase transition (i.e., the minimum L needed to drive P_F and P_M to zero as transmit power increases) is about 300 for the MMSE denoiser and 600 for the soft threshloding denoiser, indicating the clear advantage of accounting for channel statistics in user activity detector design.

5. CONCLUSIONS

This paper proposes to use the AMP algorithm with MMSE denoiser for user activity detection for massive connectivity. We provide a performance analysis of the false alarm and the missed detection probabilities and illustrate the advantage of exploiting the statistics of wireless channel in detector design.

6. REFERENCES

[1] A. K. Fletcher, S. Rangan, and V. K. Goyal, "A sparsity detection framework for on-off random access channels," in *IEEE Inter. Symp. Inf. Theory (ISIT)*, June 2009, pp. 169–173.

[2] L. Zhang, J. Luo, and D. Guo, "Neighbor discovery for wireless networks via compressed sensing," *Performance Evaluation*, vol. 70, no. 7, pp. 457–471, 2013.

[3] H. F. Schepker, C. Bockelmann, and A. Dekorsy, "Exploiting sparsity in channel and data estimation for sporadic multi-user communication," in *Inter. Symp. Wireless Commun. Sys. (ISWCS)*, Aug. 2013, pp. 1–5.

[4] G. Wunder, H. Boche, T. Strohmer, and P. Jung, "Sparse signal processing concepts for efficient 5G system design," *IEEE Access*, vol. 3, pp. 195–208, 2015.

[5] G. Wunder, P. Jung, and M. Ramadan, "Compressive random access using a common overloaded control channel," in *IEEE Globecom (GC) Workshops*, Dec. 2015, pp. 1–6.

[6] X. Xu, X. Rao, and V. K. N. Lau, "Active user detection and channel estimation in uplink CRAN systems," in *IEEE Inter. Conf. Commun. (ICC)*, June 2015, pp. 2727–2732.

[7] G. Hannak, M. Mayer, A. Jung, G. Matz, and N. Goertz, "Joint channel estimation and activity detection for multiuser communication systems," in *IEEE Inter. Conf. Commun. (ICC) Workshop*, June 2015, pp. 2086–2091.

[8] D. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

[9] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *IEEE Inter. Symp. Inf. Theory (ISIT)*, July 2011, pp. 2168–2172.

[10] A. Maleki, L. Anitori, Z. Yang, and R. G. Baraniuk, "Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP)," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4290–4308, July 2013.

[11] J. Ziniel and P. Schniter, "Efficient high-dimensional inference in the multiple measurement vector problem," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 340–354, Jan. 2013.

[12] D. L. Donoho, I. Johnstone, and A. Montanari, "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3396–3433, June 2013.

[13] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *IEEE Inf. Theory Workshop (ITW)*, Jan. 2010, pp. 1–5.

[14] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.

[15] A. Montanari, "Graphical models concepts in compressed sensing," in *Compressed Sensing: Theory and Applications*, Y. C. Eldar and G. Kutyniok, Eds., chapter 9, pp. 394–438. Cambridge University Press, New York, 2012.

[16] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comp. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, 2009.