An Efficient Active Set Algorithm for Covariance Based Joint Data and Activity Detection for Massive Random Access with Massive MIMO

Ziyue Wang[⋆,§], Zhilin Chen[†], Ya-Feng Liu[§], Foad Sohrabi[†], and Wei Yu[†]

[⋆]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, China [†] Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada

[§]LSEC, ICMSEC, AMSS, Chinese Academy of Sciences, Beijing, China Email: wangziyue20@mails.ucas.ac.cn, {zchen, fsohrabi, weiyu}@ece.utoronto.ca, yafliu@lsec.cc.ac.cn

Motivation

- The uncoordinated random access is a challenging task in massive machine-type communication (mMTC).
- A large number of sporadically active devices wish to send small data to the base-station (BS) in the uplink.
- The BS acquires the active devices and their data by detecting the transmitted preassigned nonorthogonal signature sequences.
- Covariance based approach [1, 2, 3]: formulate the detection problem as a maximum likelihood estimation (MLE) problem.
- The state-of-the-art coordinate descent (CD) algorithm doesn't take advantage of the sparsity of the true solution.

Main Contribution

- Perform the covariance based approach for joint data and activity detection.
- Propose a computationally efficient active set algorithm with convergence guarantee.

System Model

- Single cell with one BS equipped with M antennas.
- N single-antenna devices, K of which are active at a time.
- Each active device wishes to transmit J bits of data to the BS.
- Each device n has a unique signature sequence set $\mathcal{S}_n = \{\mathbf{s}_{n,1}, \mathbf{s}_{n,2}, \ldots, \mathbf{s}_{n,Q}\}$, where $\mathbf{s}_{n,q} \in \mathbb{C}^{L\times 1}, 1 \le q \le Q \triangleq 2^J$, and L is the signature sequence length.
- Channel $\sqrt{g_n}\mathbf{h}_n \in \mathbb{C}^{M\times 1}$ of user n includes both
  - ◇ large-scale fading component $g_n \ge 0$;
  - ◇ Rayleigh fading component $\mathbf{h}_n \in \mathbb{C}^{M\times 1}$ following the i.i.d. complex Gaussian distribution.
- Whether or not $\mathbf{s}_{n,q}$ is transmitted is indicated as $\chi_{n,q} \in \{0, 1\}$, which satisfies $\sum_{q=1}^{Q} \chi_{n,q} \in \{0, 1\}$
  - ◇ $\sum_{q=1}^{Q} \chi_{n,q} = 1$ indicates that device n is active;
  - ◇ $\sum_{q=1}^{Q} \chi_{n,q} = 0$ indicates that device n is inactive.
- Define
  - ◇ $\mathbf{S} = [\mathbf{S}_1, \ldots, \mathbf{S}_N] \in \mathbb{C}^{L\times NQ}$, where $\mathbf{S}_n = [\mathbf{s}_{n,1}, \ldots, \mathbf{s}_{n,Q}]$.
  - ◇ $\mathbf{\Gamma}^{1/2} = \operatorname{diag}\{\mathbf{D}_1, \ldots, \mathbf{D}_N\} \in \mathbb{C}^{NQ\times NQ}$, where $\mathbf{D}_n = \sqrt{g_n}\operatorname{diag}\{\chi_{n,1}, \ldots, \chi_{n,Q}\}$.
  - ◇ $\mathbf{H} = [\mathbf{H}_1^T, \ldots, \mathbf{H}_N^T]^T \in \mathbb{C}^{NQ\times M}$, where $\mathbf{H}_n = [\mathbf{h}_n, \ldots, \mathbf{h}_n]^T$.

System Model (Cont.)

- The received signal $\mathbf{Y} \in \mathbb{C}^{L\times M}$ at the BS can be expressed as

$$\mathbf{Y} = \sum_{n=1}^{N}\sum_{q=1}^{Q} \chi_{n,q}\mathbf{s}_{n,q}\sqrt{g_n}\mathbf{h}_n^T + \mathbf{W} = \mathbf{S}\mathbf{\Gamma}^{1/2}\mathbf{H} + \mathbf{W}, \qquad (1)$$

where $\mathbf{W} \in \mathbb{C}^{L\times M}$ is the effective i.i.d. Gaussian noise with variance σ_w^2.

- For given $\boldsymbol{\gamma}$ (diagonal entries of $\mathbf{\Gamma}$), the m-th column of $\mathbf{Y}$ can be seen as independent samples from a complex Gaussian distribution as

$$\mathbf{y}_m \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}\mathbf{\Gamma}^{1/2}\mathbf{\Lambda}\mathbf{\Gamma}^{1/2}\mathbf{S}^H + \sigma_w^2\mathbf{I}\right), \qquad (2)$$

where $\mathbf{\Lambda}$ is a block diagonal matrix with each block being the all-one matrix $\mathbf{E} \in \mathbb{R}^{Q\times Q}$, and $\mathbf{I}$ is an identity matrix.

- Since there is at most one non-zero entry in each diagonal block $\mathbf{D}_n$ in $\mathbf{\Gamma}^{1/2}$, the covariance matrix in (2) can be simplified as

$$\mathbf{S}\mathbf{\Gamma}^{1/2}\mathbf{\Lambda}\mathbf{\Gamma}^{1/2}\mathbf{S}^H + \sigma_w^2\mathbf{I} = \mathbf{S}\mathbf{\Gamma}\mathbf{S}^H + \sigma_w^2\mathbf{I}.$$

- The MLE problem can be formulated as

$$\min_{\boldsymbol{\gamma}} \quad \log\left|\mathbf{S}\mathbf{\Gamma}\mathbf{S}^H + \sigma_w^2\mathbf{I}\right| + \operatorname{Tr}\left(\left(\mathbf{S}\mathbf{\Gamma}\mathbf{S}^H + \sigma_w^2\mathbf{I}\right)^{-1}\hat{\mathbf{\Sigma}}\right) \qquad (3a)$$

$$\text{s.t.} \quad \boldsymbol{\gamma} \ge \mathbf{0}. \qquad (3b)$$

- The sample covariance matrix $\hat{\mathbf{\Sigma}} = \mathbf{Y}\mathbf{Y}^H/M$ is computed by averaging over different antennas.
- The constraint $\boldsymbol{\gamma} \ge \mathbf{0}$ is due to the fact that $\gamma_{n,q} = g_n\chi_{n,q} \ge 0$ for all n and q.

Problem Formulation and Analysis

- Let $f(\boldsymbol{\gamma})$ denote the objective function of problem (3). The gradient of $f(\boldsymbol{\gamma})$ with respect to $\gamma_{n,q}$ is

$$\left[\nabla f(\boldsymbol{\gamma})\right]_{n,q} = \mathbf{s}_{n,q}^H\mathbf{\Sigma}^{-1}\mathbf{s}_{n,q} - \mathbf{s}_{n,q}^H\mathbf{\Sigma}^{-1}\hat{\mathbf{\Sigma}}\mathbf{\Sigma}^{-1}\mathbf{s}_{n,q}.$$

- The first-order (necessary) optimality condition of problem (3) is

$$\left[\nabla f(\boldsymbol{\gamma})\right]_{n,q} \begin{cases} = 0, & \text{if } \gamma_{n,q} > 0; \\ \ge 0, & \text{if } \gamma_{n,q} = 0, \end{cases} \quad \forall\, q, n, \qquad (4)$$

- Let $[\cdot]_+$ denote the projection operator onto the nonnegative orthant. Then (4) is equivalent to

$$[\boldsymbol{\gamma} - \nabla f(\boldsymbol{\gamma})]_+ - \boldsymbol{\gamma} = \mathbf{0}.$$

Active Set Algorithm

- To fully exploit the sparsity of the true solution of (3), the active set should
  - – contain the indices of active sequences;
  - – have the smallest possible cardinality.
- At the k-th iteration, the proposed selection strategy of the active set $\mathcal{A}^k$ is

$$\mathcal{A}^k = \left\{(i, q) \mid \gamma_{i,q}^k > \omega_k \text{ or } [\nabla f(\boldsymbol{\gamma}^k)]_{i,q} < -\nu_k\right\}, \qquad (5)$$

where $\omega_k, \nu_k > 0$ and $\omega_k \downarrow 0$ and $\nu_k \downarrow 0$ (monotonically decrease and converge to zero).

- Once the active set $\mathcal{A}^k$ is selected, we solve the following subproblem

$$\min \quad \hat{f}(\boldsymbol{\gamma}_{\mathcal{A}^k}) \qquad (6a)$$

$$\text{s.t.} \quad \boldsymbol{\gamma}_{\mathcal{A}^k} \ge \mathbf{0}, \qquad (6b)$$

where $\boldsymbol{\gamma}_{\mathcal{A}^k}$ is the subvector of $\boldsymbol{\gamma}$ indexed by $\mathcal{A}^k$ and $\hat{f}(\boldsymbol{\gamma}_{\mathcal{A}^k})$ is $f(\boldsymbol{\gamma})$ defined over $\boldsymbol{\gamma}_{\mathcal{A}^k}$ with all the other variables fixed being zero.

- If the set $\mathcal{A}^k$ in (6) is properly chosen, the dimension of problem (6) is potentially much smaller than that of problem (3).
- We apply the spectral PG algorithm [4] to solve the subproblem in (6) until $\boldsymbol{\gamma}_{\mathcal{A}^k}^{k+1}$ satisfying

$$\left\|\left[\boldsymbol{\gamma}_{\mathcal{A}^k}^{k+1} - \nabla\hat{f}(\boldsymbol{\gamma}_{\mathcal{A}^k}^{k+1})\right]_+ - \boldsymbol{\gamma}_{\mathcal{A}^k}^{k+1}\right\| < \varepsilon_k, \qquad (7)$$

where $\varepsilon_k > 0$ is the solution tolerance at the k-th iteration.

- The pseudocodes of the proposed algorithm are given in Algorithm 1.

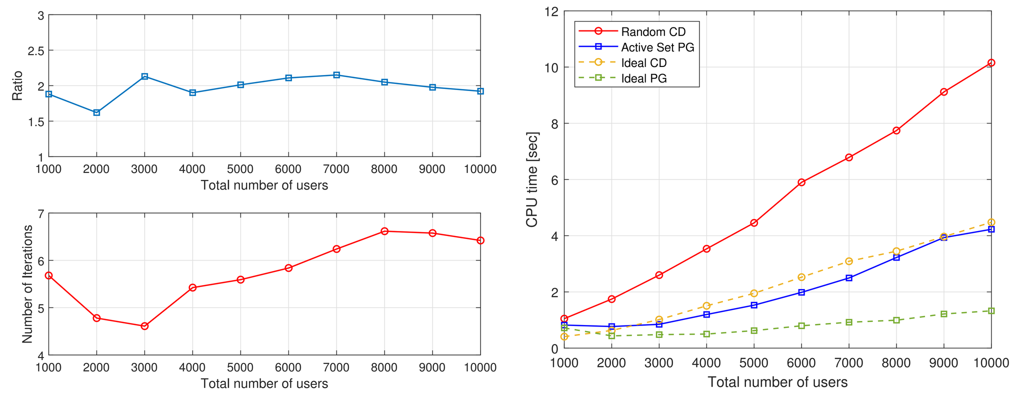
Algorithm 1 Proposed active set PG algorithm for solving problem (3)

Initialize: $\boldsymbol{\gamma}^0 = \mathbf{0}$, $k = 0$, $\{\omega_k, \nu_k, \varepsilon_k\}_{k\ge 0}$, and $\varepsilon > 0$;
repeat
  Select the active set $\mathcal{A}^k$ according to (5);
  Apply the spectral PG algorithm [4] to solve the subproblem (6) until (7) is satisfied;
  Set $k \leftarrow k + 1$;
until $\|[\boldsymbol{\gamma}^k - \nabla f(\boldsymbol{\gamma}^k)]_+ - \boldsymbol{\gamma}^k\| < \varepsilon$
Output: $\boldsymbol{\gamma}^k$

- Convergence property: For any given tolerance $\varepsilon > 0$, suppose that the parameters ω_k and ν_k in (5) satisfy $\omega_k \downarrow 0$ and $\nu_k \downarrow 0$ and the parameter ε_k in (7) satisfy $\lim_{k\to\infty} \varepsilon_k < \varepsilon$, then the active set PG Algorithm 1 will terminate within a finite number of iterations.

Simulation Results

- The power spectrum density of the background noise is -169dBm/Hz over 10 MHz and the transmit power of each device is 25dBm;
- A single cell of radius 1000m, all devices are located in the cell edge, g_n's are the same for all devices;
- All signature sequences from i.i.d. complex Gaussian distribution with zero mean and unit variance
- Parameters setting: $M = 256$, $L = 150$, and $J = 1$ (and thus $Q = 2$), $K/N = 0.1$ (10% of the total devices are active).
- Compare the proposed Algorithm 1 with
  - – random CD algorithm in [1];
  - – Ideal CD/PG algorithm: apply the CD/PG algorithm to solve problem (3) defined over the indices of active sequences;
- Parameters setting: $\omega_k = 10^{-6-k}$, $\varepsilon_k = \max\left\{10^{-k}, 0.8 * 10^{-3}\right\}$, $\nu_k = \min\left\{10^{4-k}, 0.5\left|\min_{n,q}\left\{\left[\nabla f(\boldsymbol{\gamma}^k)\right]_{n,q}\right\}\right|\right\}$, $\varepsilon = 10^{-3}$.
- Average over 500 Monte-Carlo runs.



Left: Average ratio $|\mathcal{A}^k|/K$; Average number of iterations to terminate; Right: Average CPU time comparison.

- The ratio is in the interval $[1.5, 2.5]$, and Algorithm 1 will generally terminate within 4–7 iterations.
- The proposed active set selection strategy (5) is very efficient.
- In CPU time, the proposed Algorithm 1 significantly outperforms the random CD algorithm, and even achieves slightly better efficiency than the ideal CD algorithm.

References

[1] S. Haghighatshoar, P. Jung, and G. Caire, "Improved scaling law for activity detection in massive MIMO systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT), Vail, CO, USA*, June 2018, pp. 381–385.

[2] Z. Chen, F. Sohrabi, Y.-F. Liu and W. Yu, "Covariance based joint activity and data detection for massive random access with massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC), Shanghai, China*, May 2019, pp. 1–6.

[3] Z. Chen, F. Sohrabi, Y.-F. Liu, and W. Yu, "Phase transition analysis for covariance based massive random access with massive MIMO," 2020. [Online]. Available: https://arxiv.org/abs/2003.04175

[4] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM J. Optim.*, vol. 10, no. 4, pp. 1196–1211, 2000.