Blind Goal-Oriented Detection (BGOD) for Massive Access in Future Wireless Networks

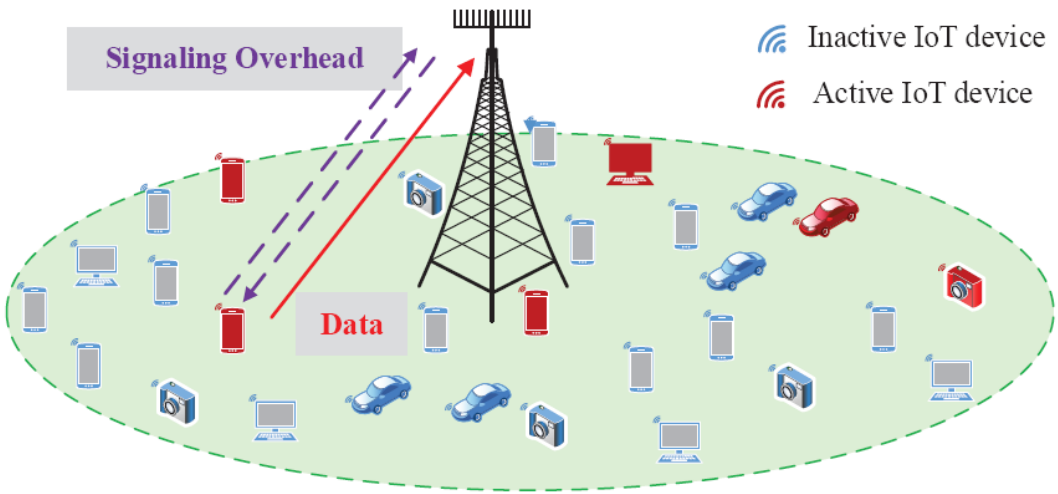
October 20, 2022

Outline

1. Introduction
2. System model
3. Prior works and their limitations
4. Proposed strategy
5. Simulations
6. Applications beyond wireless communications

Introduction

- Random Access (RA) is the key component of the communications between users and the BS in wireless networks.
- Various applications of IoT demand for massive machine machine-type communications (MTC).
- Uplink transmission of massive IoT devices with short packets
- Sporadic traffic of IoT devices



System Model

- Uniform linear array BS with N receive antennas serving $K = K_M + K_S$ single antenna users
- Only $K_a = K_{a,M} + K_{a,S} \ll K$ are simultaneously active, which use the same time-freq. resource.
- We assume a flat-fading model in which the channel is considered to be fixed during T signal dimensions in a channel of time-freq. coherence block:

$$\boldsymbol{h}_k = \rho_k \sum_{l=1}^{L_k} \alpha_l^k \boldsymbol{a}(\theta_l^k) \in \mathbb{C}^{N \times 1}$$

- The signal received at the BS can be given as:

$$\boldsymbol{Y}_\Omega = \mathcal{P}_\Omega(\boldsymbol{Y}) = \mathcal{P}_\Omega \left(\sum_{k=1}^{K} \boldsymbol{h}_k \boldsymbol{\phi}_k^H \right) + \boldsymbol{E} \in \mathbb{C}^{M \times T}$$

Prior RA works

Four-step RA

1. Each device picks a preamble from a predefined set at random.
2. The BS detects the preambles transmitted by active devices and sends responses to the active users to allocate a resource to each user.
3. The devices whose preambles are successfully detected transmit connection request messages including their device identities in order to request resources for subsequent data transmission.
4. The forth step is about contention resolution in which all colliding users need to re-transmit after a random time.

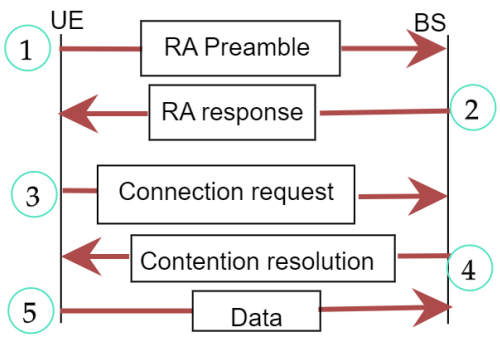


Figure 1: Four-step RA

Prior RA works

2-step RA

- Each active device transmits a preamble along with the data payload in a time division manner.
- At each time slot, the BS first detects the active devices by detecting which preamble sequences are used.
- Due to the limited channel coherence time and massive users, it is not possible to assign orthogonal pilots to all of the devices.

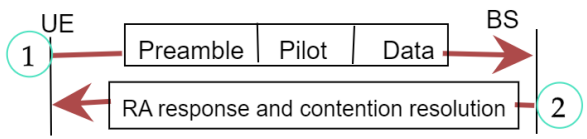


Figure 2: Two-step RA

RA frameworks

Sourced RA

- BS is interested in both the messages of the identities of the devices.
- The sourced random access is mainly applied to device-oriented applications such as status monitoring .
- Each device is preassigned to a unique signature sequence.

Unsourced RA

- BS is interested in the transmitted messages only and not the identity of the devices.
- Each user selects a B bit message from a shared codebook and transmit it.
- The unsourced random access is mainly applied to content-oriented applications such as quality inspections process of smart factories .

Limitations of prior RA works

4-step RA

- Preamble collision happens when multiple active devices choose the same preamble.
- The probability of preamble collision increases exponentially with the number of active devices.
- Long access latency in massive access scenarios
- Prohibitive signalling overhead even when the devices have short packages to transmit

Limitations of prior RA works

2-step RA

- The number of active users that can access the BS is severely limited in massive access scenario ($K_a = \mathcal{O}(\frac{T^2}{\log^2(\frac{K}{T^2})})$).
- The detection accuracy directly depends on the total number of users.
- **Massive connectivity** requires massive **communication costs** at the user side

Limitations of prior RA works

Unsourced RA

- Users are forced to use the same and shared codebook
- Impossible to use individually different codebooks
- The length of the messages is considered to be the same.
- The users have to be synchronized for data transmission.
- For short payload length, the number of active devices that can be accommodated is limited by the detection capabilities of the inner decoder ($K_a = \mathcal{O}(\frac{T^2}{\log^2(\frac{2^B}{T^2})})$)

BGOD strategy

Expensive optimization problem

$$\min_{\substack{\boldsymbol{Z}_k \in \mathbb{C}^{N\times T} \\ k=1,\ldots,K}} \sum_{k=1}^{K} \|\boldsymbol{Z}_k\|_{\mathcal{A}_k} \; s.t. \|\boldsymbol{Y} - \sum_{k=1}^{K} \mathcal{P}_{\Omega}(\boldsymbol{Z}_k)\|_F \leq \eta \tag{1}$$

$$\mathcal{A}_k = \{\boldsymbol{a}(\theta)\boldsymbol{\phi}_k^H, \theta \in (0,\pi)\} \tag{2}$$

$$\|\boldsymbol{Z}_k\|_{\mathcal{A}_k} := \inf\{t > 0 : \boldsymbol{Z}_k \in t\mathrm{conv}(\mathcal{A}_k)\} =$$

$$\inf\{\sum_{l=1}^{L_k} |c_l^k| : \boldsymbol{Z}_k = \sum_{l=1}^{L_k} c_l^k \boldsymbol{a}(\theta_l^k)\boldsymbol{\phi}_k^H, \boldsymbol{a}(\theta_l^k)\boldsymbol{\phi}_k^H \in \mathcal{A}_k\} \tag{3}$$

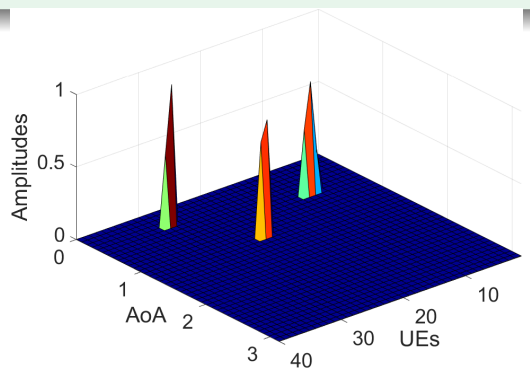


Figure 3: Angular sparse feature of channel.

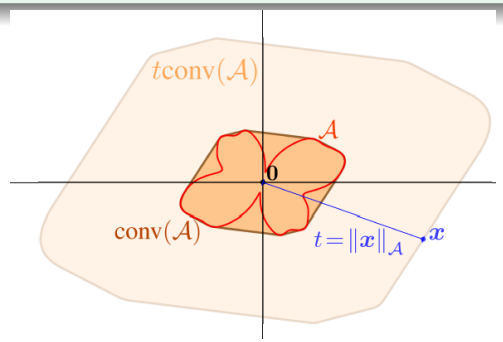


Figure 4: Atomic norm

BGOD strategy

Goal-oriented Optimization problem

$$\min_{\substack{\boldsymbol{v}\in\mathbb{C}^N,\boldsymbol{Z}\in\mathbb{C}^{N\times T}\\ \boldsymbol{Y}^\star\in\mathbb{C}^{M\times T},\boldsymbol{W}\in\mathbb{C}^{T\times T}}} \quad \mathrm{Re}(v_1)+\mathrm{Re}(\mathrm{tr}(\boldsymbol{W}))+\frac{\gamma}{2}\|\boldsymbol{Y}-\boldsymbol{Y}^\star\|_F^2$$

$$\text{s.t.} \quad \begin{bmatrix}\mathcal{T}(\boldsymbol{v}) & \boldsymbol{Z}\\ \boldsymbol{Z}^H & \boldsymbol{W}\end{bmatrix}\succeq \boldsymbol{0} \quad , \boldsymbol{Y}^\star=-2c_1\mathcal{P}_\Omega(\boldsymbol{Z}). \tag{4}$$

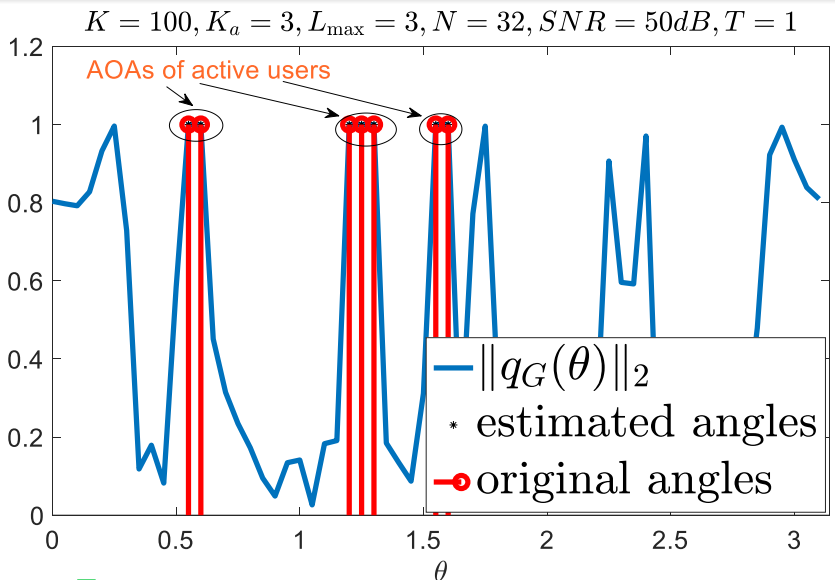


Figure 5: ℓ_2 norm of the goal-oriented dual polynomial function

BGOD strategy

- Use alternative minimization to recover the preamble and data of each active device

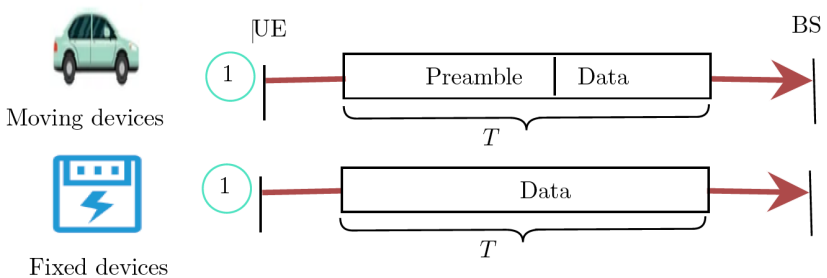
$$[\widehat{\boldsymbol{\alpha}}^1, ..., \widehat{\boldsymbol{\alpha}}^{\widehat{K}_a}] = \underset{\substack{\alpha_1^k, ..., \alpha_{\widehat{L}_k}^k, \\ k=1,...,\widehat{K}_a}}{\arg\min} \| \boldsymbol{Y} - \sum_{k \in \widehat{\mathcal{S}}_{\mathrm{AU}}} \sum_{l=1}^{\widehat{L}_k} \alpha_l^k \mathcal{P}_\Omega(\boldsymbol{a}(\widehat{\theta}_l^k) \widehat{\boldsymbol{\phi}}_k^H) \|_F \quad (5)$$

$$[\widehat{\boldsymbol{\phi}}_1, ..., \widehat{\boldsymbol{\phi}}_{\widehat{K}_a}] = \underset{\substack{\phi_1, ..., \phi_{\widehat{K}_a} \\ \phi_k \geq \mathbf{0}}}{\arg\min} \| \boldsymbol{Y} - \sum_{k \in \widehat{\mathcal{S}}_{\mathrm{AU}}} \sum_{l=1}^{\widehat{L}_k} \widehat{\alpha}_l^k \mathcal{P}_\Omega(\boldsymbol{a}(\widehat{\theta}_l^k) \boldsymbol{\phi}_k^H) \|_F \quad (6)$$

BGOD strategy

Motivations and Ideas

- The expensive optimization problem is computationally intractable in massive access scenarios
- Design a reconstruction-free optimization problem with the goal of active user detection in mind
- The information of channel gains and messages corresponding to all of the devices is completely lost and destroyed.
- **Big discovery:** The maximum of ℓ_2 norm of the dual polynomial functions occurs at the **same** locations for the expensive and goal-oriented problems.
- This shared information between the expensive and goal-oriented optimization problems is sufficient to find out the angles of all active devices.



BGOD features

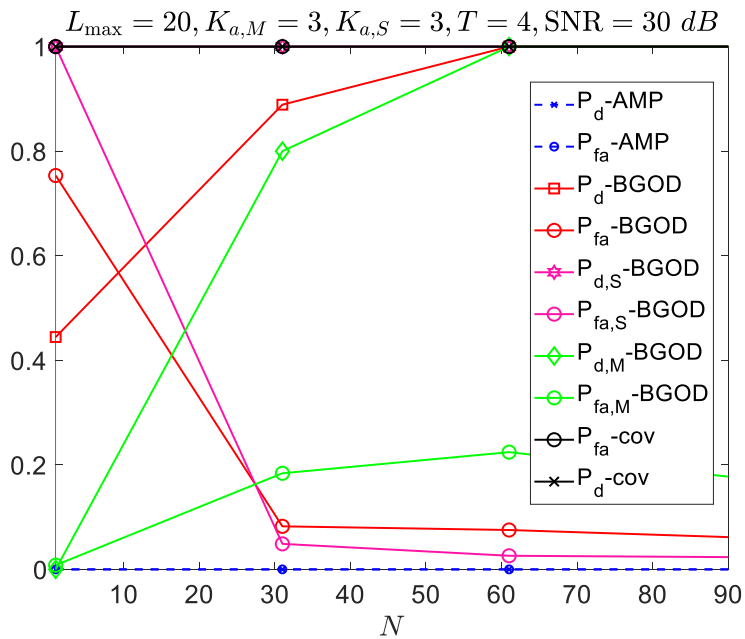
- Channel distribution is arbitrary
- Noise distribution is arbitrary
- The first Sourced RA strategy that is fully independent of the total number of devices.
- Communications costs at the users' side for RA are somehow transferred to the computational complexity at the BS.
- Separate the stationary active users by their LoS angles
- Separate the mobile active users by both their angles and preambles
- Stationary users do not have to spend any resources for RA
- Mobile active users have to include a preamble before their data payload

BGOD features

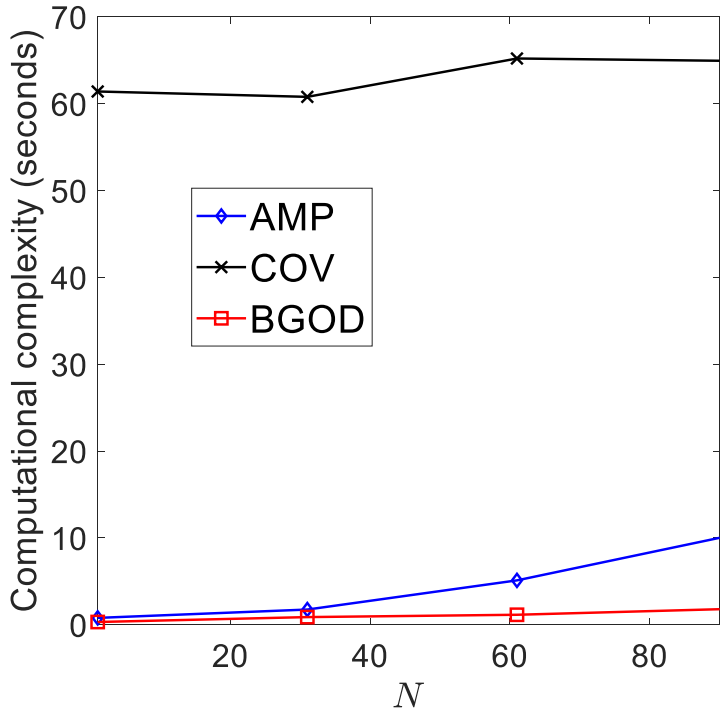
Unsourced

- The first blind unsourced RA
- The users are free to transmit any arbitrary messages from individually different codebooks
- Users do not need to be synchronized for data transmission.
- The messages can have different length

Simulation results



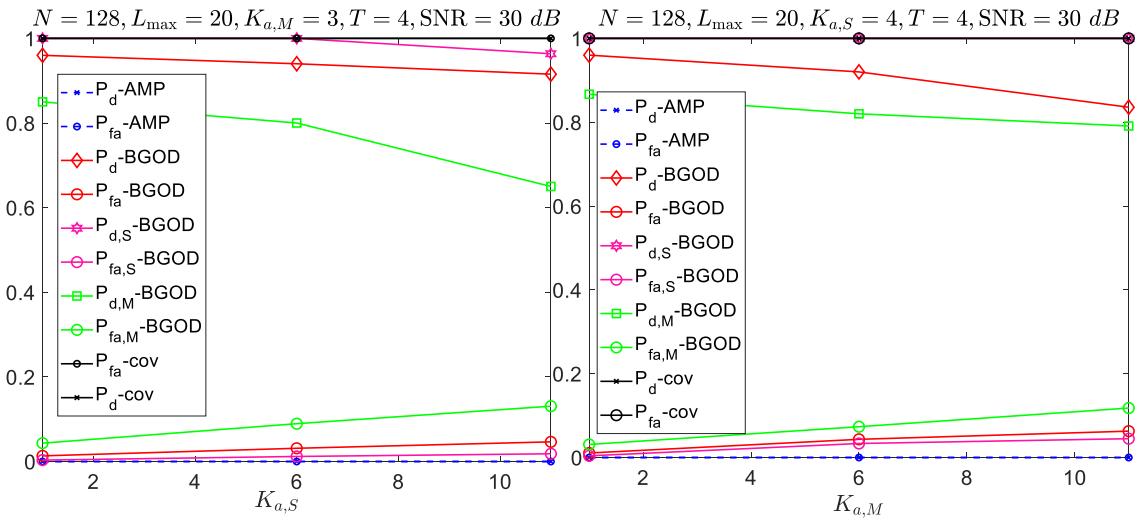
(a)



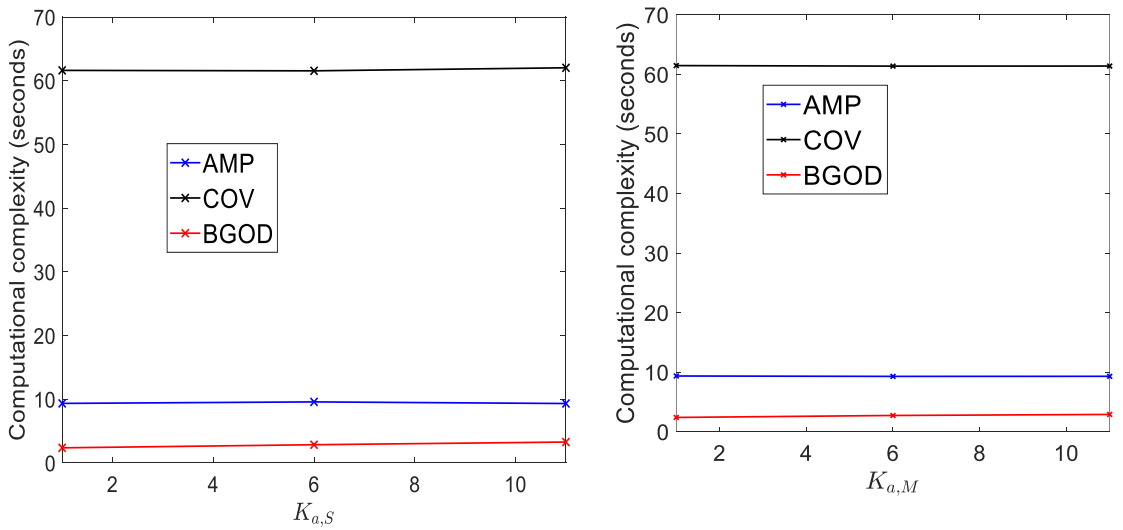
(b)

Figure 6: $K_S = 1500, K_M = 500, K_{a,\mathrm{M}} = K_{a,\mathrm{S}} = 3, L_{\max} = 20, T = 4$, and $\mathrm{SNR} = 30\mathrm{dB}$. (a) Detection accuracy (b) Computational complexity.

Simulations results

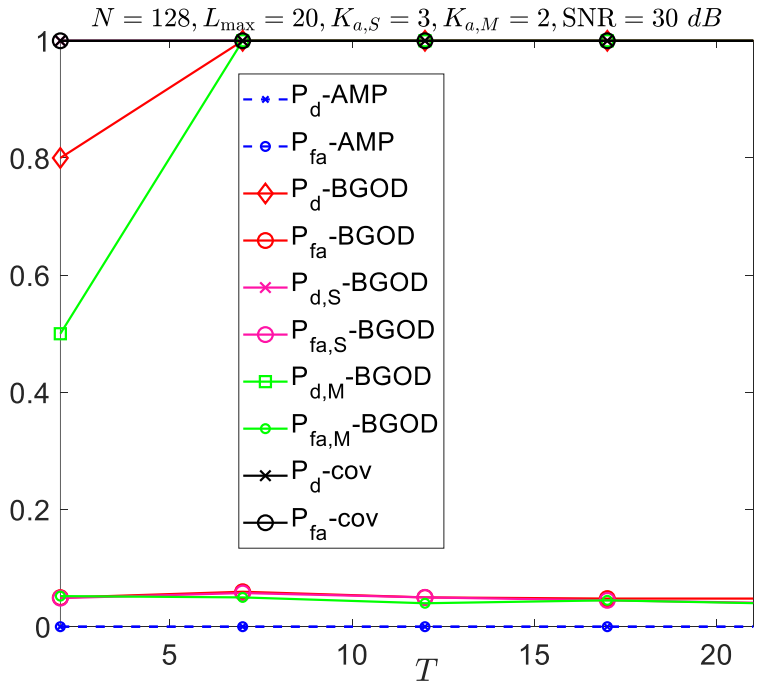


(a)

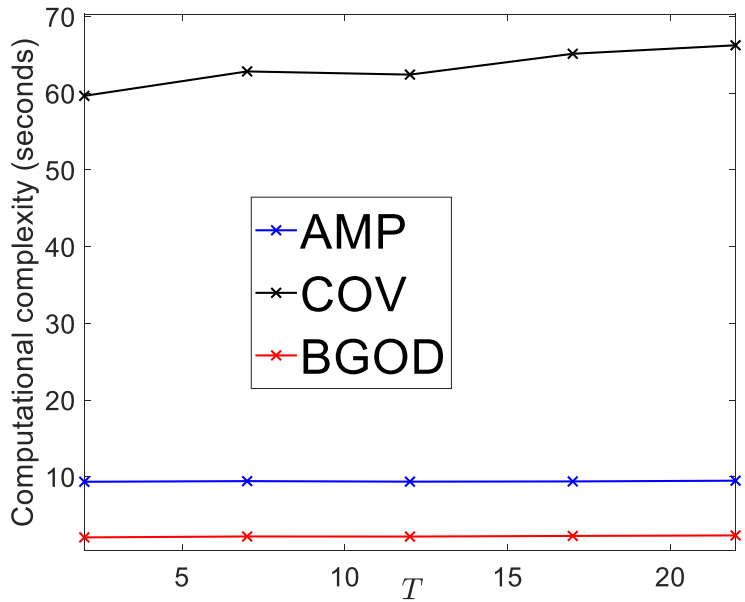


(b)

Simulation results



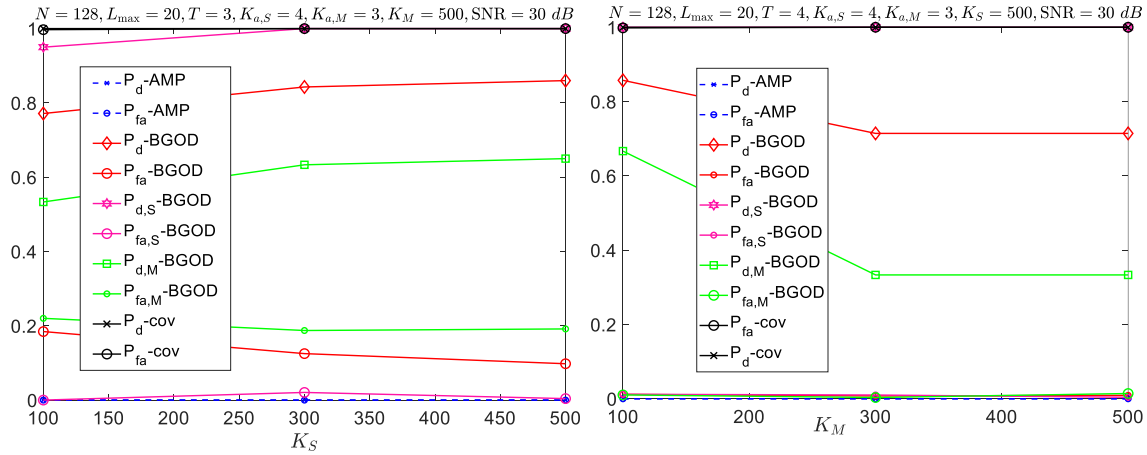
(a)



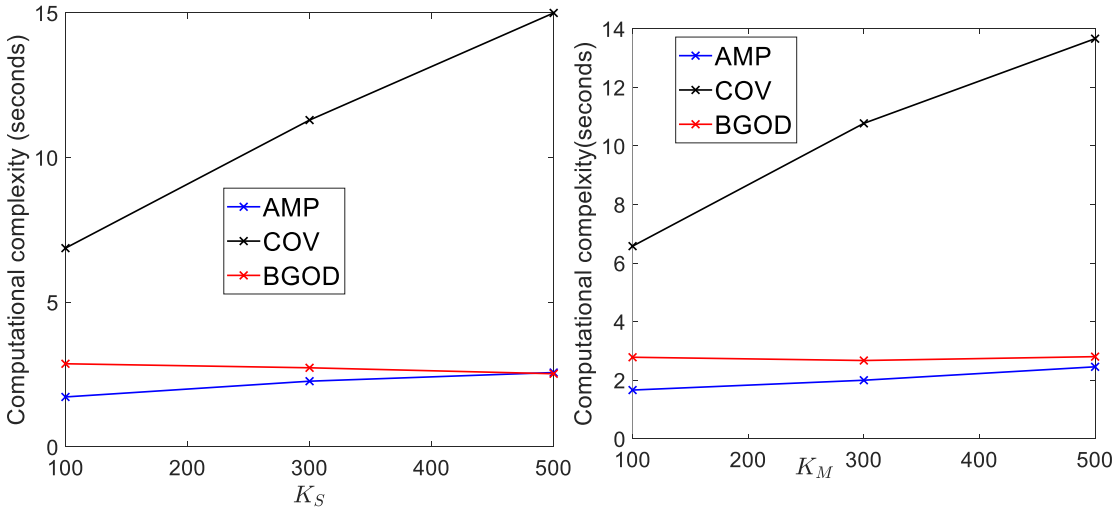
(b)

Figure 8: $N = 128$, $K_S = 1500$, $K_M = 500$, $K_{a,S} = 3$, $K_{a,M} = 2$, $L_{\max} = 20$, and $\mathrm{SNR} = 30\mathrm{dB}$. (a) Detection accuracy (b) Computational complexity.

Simulation results



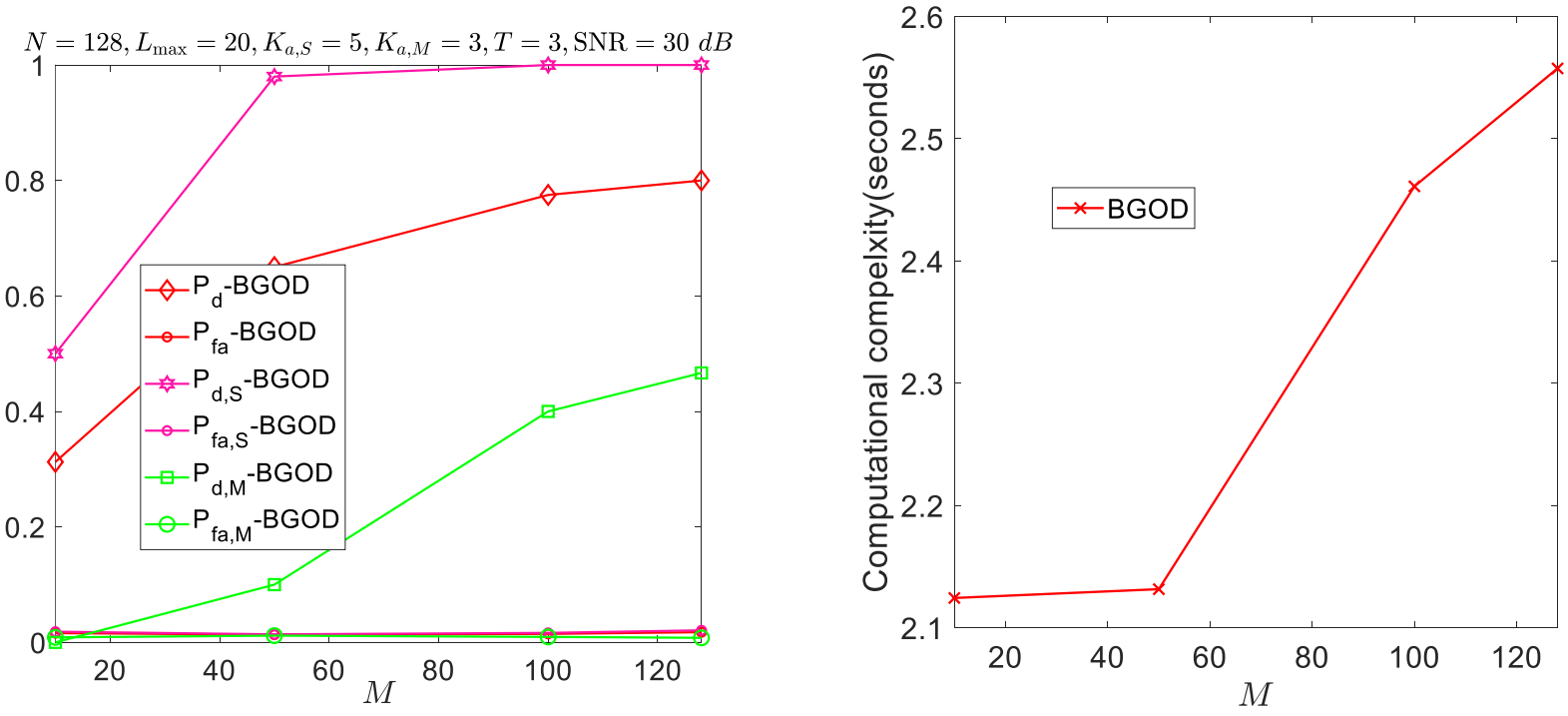
(a)



(b)

Figure 9: $N = 128$, $K_{a,M} = 3$, $K_{a,S} = 4$, $L_{\max} = 20$, and SNR $= 30$dB

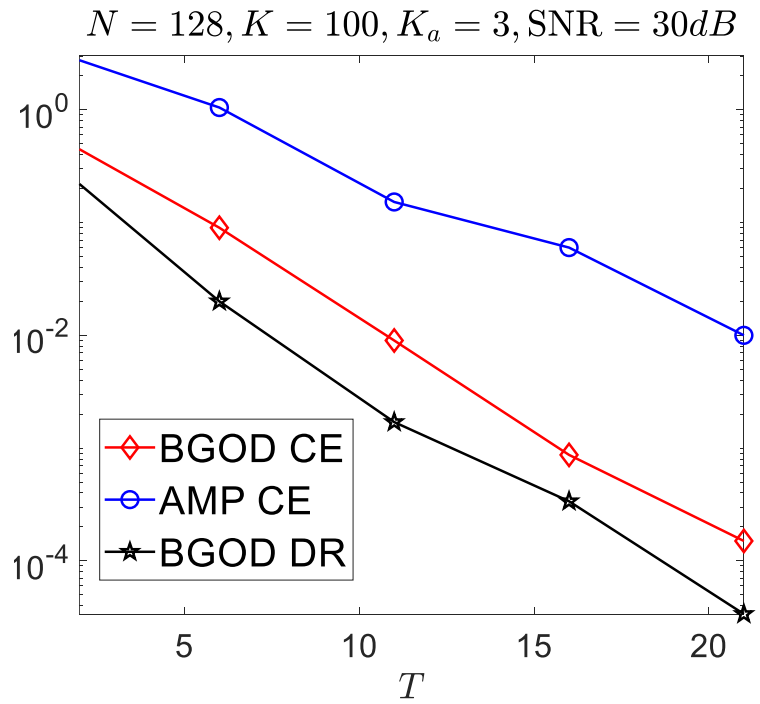
Simulation results



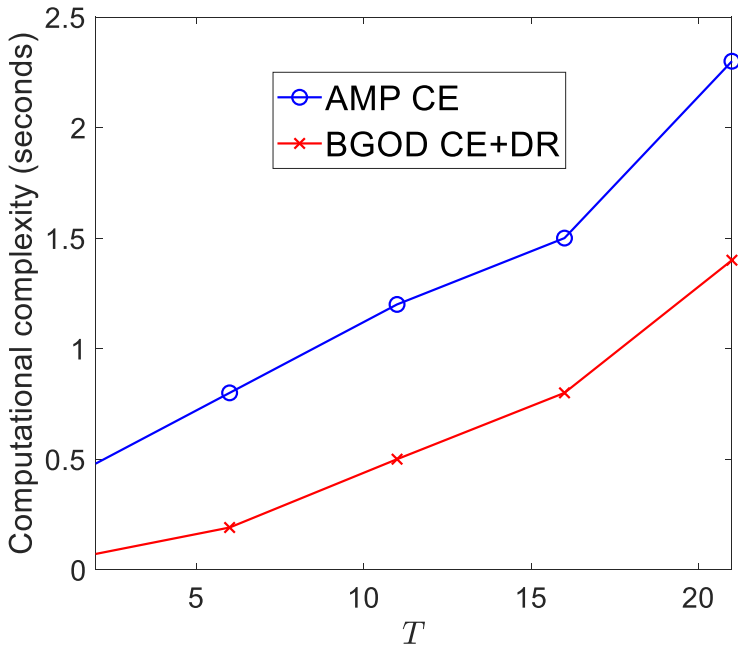
(a) (b)

Figure 10: Random access performance of the proposed method versus the number M of observed arrays at the BS. Figures (a) and (b) show detection accuracy and computational complexity, respectively.

Simulation results



(a)



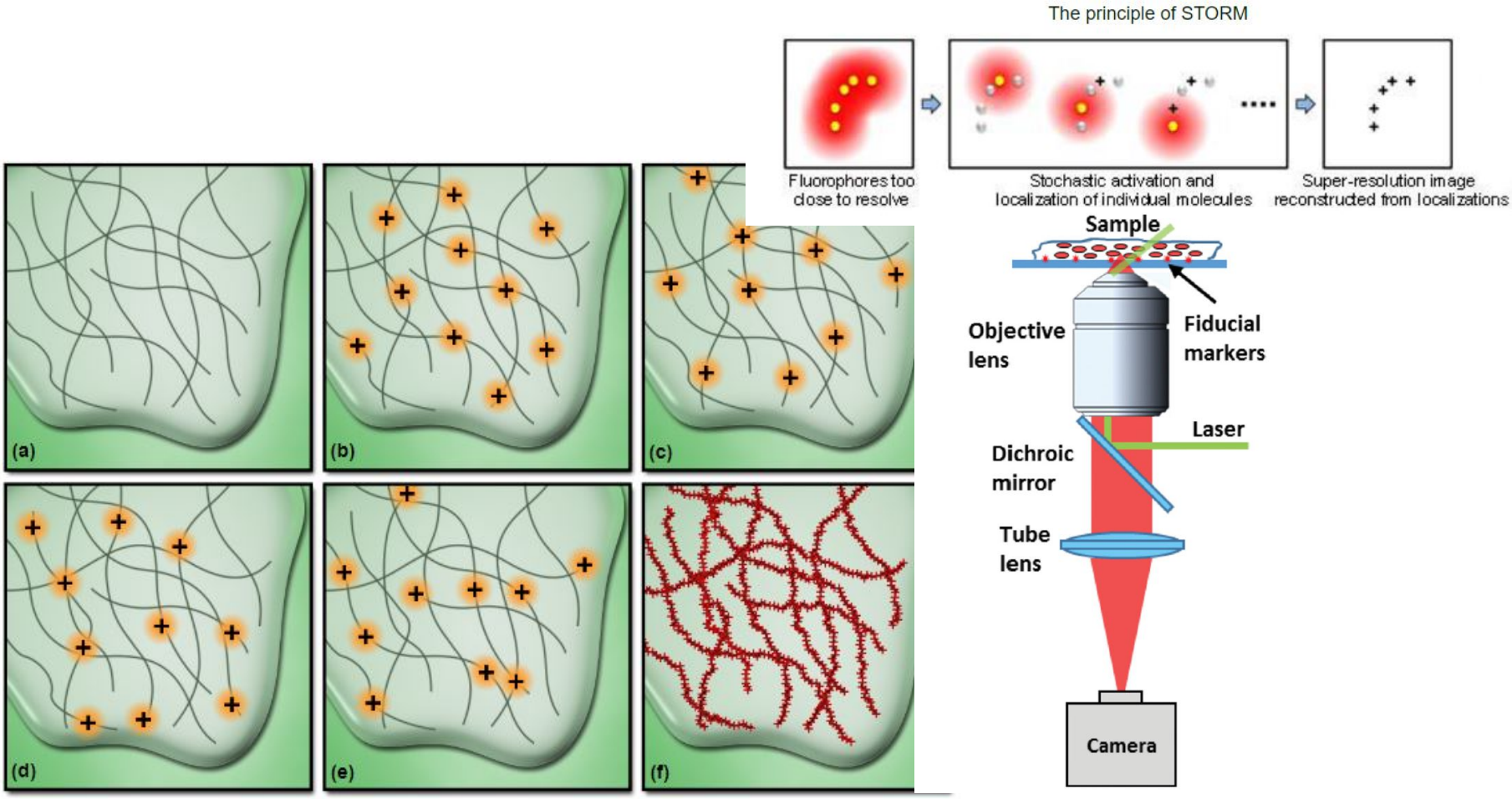
(b)

Figure 11: CE and DR performance of our proposed method versus the time resources T for $N = 128, K = 100, K_a = 5, L_{\max} = 3$, and $\mathrm{SNR} = 30\mathrm{dB}$.

BGOD goes beyond wireless communications

Three-dimensional super-resolution single-molecule imaging

- Fluorescence microscopy is an essential tool for biologists to visualize molecular structures and their interactions.
- STORM is a type of super-resolution optical microscopy technique based on stochastic switching of single-molecule fluorescence signal.
- By employing photoswitchable fluorescent molecules, the imaging process of single-molecule microscopies is divided into many frames , where in each frame, a sparse number of fluorophores (point sources) are randomly activated and then deactivated.
- Reconstruct a 3-D object from 2-D image frames, by introducing a cylindrical lens to modulate the ellipticity of the point spread function based on the depth of the fluorescent object.
- The acquired image in each frame can be regarded as a superposition of returns from multiple depth layers, where the return from each layer corresponds to the convolution outcome of the fluorophores in that depth layer with the depth-dependent point spread function.


The principle of STORM
Fluorophores too close to resolve
Stochastic activation and localization of individual molecules
Super-resolution image reconstructed from localizations
(a)
(b)
(c)
(d)
(e)
(f)
Sample
Objective lens
Fiducial markers
Laser
Dichroic mirror
Tube lens
Camera

BGOD goes beyond wireless communications

Spike sorting for neural recording in neuroscience

- Neurons in the brain communicate by firing action potentials, i.e. spikes, and it is possible to capture their communications through the use of a microelectrode, which records simultaneous activities of multiple neurons within a local neighborhood.
- The neural recording can be modeled as a superposition of returns from multiple neurons, where the return from each neuron corresponds to the convolution of its characteristic spike shape with the sequence of its firing times.

DNA sequencing in biology

- recover the underlying DNA sequence from the vector time series.


THANKS

BGOD strategy: Details

Dual formulation of the expensive problem

$$\max_{\boldsymbol{V}\in\mathbb{C}^{N\times T}} \mathrm{Re}\langle \boldsymbol{V}, \boldsymbol{Y}\rangle_F - \frac{1}{2\gamma}\|\boldsymbol{V}\|_F^2$$
$$\text{s.t.} \quad \|(\mathcal{P}_\Omega^{\mathrm{Adj}}(\boldsymbol{V}))^H \boldsymbol{a}(\theta_k)\|_2 \|\boldsymbol{\phi}_k\|_2 \leq 1$$
$$\forall \theta_k \in (0,\pi), k = 1, ..., K \qquad (7)$$

Goal oriented dual problem

$$\max_{\substack{\boldsymbol{V}\in\mathbb{C}^{M\times T}\\ \boldsymbol{Q}\in\mathbb{C}^{N\times N}}} \mathrm{Re}\langle \boldsymbol{V}, \boldsymbol{Y}\rangle_F - \frac{1}{2\gamma}\|\boldsymbol{V}\|_F^2$$
$$\begin{bmatrix} \boldsymbol{Q} & \mathcal{P}_\Omega^{\mathrm{Adj}}(\boldsymbol{V})c_1 \\ (\mathcal{P}_\Omega^{\mathrm{Adj}}(\boldsymbol{V}))^H c_1 & \boldsymbol{I}_T \end{bmatrix} \succeq \boldsymbol{0},$$
$$\langle \boldsymbol{\Theta}_q, \boldsymbol{Q}\rangle = 1_{q=0}, \quad q = -N+1, ..., N-1. \qquad (8)$$

Theorem

The maximum of the ℓ_2 norm of the dual polynomial function $\boldsymbol{q}_G(\theta) = (\mathcal{P}_\Omega^{\mathrm{Adj}}(\boldsymbol{V}))^H \boldsymbol{a}(\theta)$ occurs at the same locations for the expensive and the goal-oriented dual problems