

Article

Enhancing the Accuracy and Robustness of a Compressive Sensing Based Device-Free Localization by Exploiting Channel Diversity

Dongping Yu , Yan Guo *, Ning Li and Xiaoqin Yang

College of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China; yudongpingxm@126.com (D.Y.); lining_friend@sina.com (N.L.); 15261856573@139.com (X.Y.)

* Correspondence: guoyan_1029@sina.com; Tel.: +86-181-5103-2759

Received: 8 March 2019; Accepted: 15 April 2019; Published: 17 April 2019



Abstract: As an emerging and promising technique, device-free localization (DFL) estimates target positions by analyzing their shadowing effects. Most existing compressive sensing (CS)-based DFL methods use the changes of received signal strength (RSS) to approximate the shadowing effects. However, in changing environments, RSS readings are vulnerable to environmental dynamics. The deviation between runtime RSS variations and the data in a fixed dictionary can significantly deteriorate the performance of DFL. In this paper, we introduce ComDec, a novel CS-based DFL method using channel state information (CSI) to enhance localization accuracy and robustness. To exploit the channel diversity of CSI measurements, the DFL problem is formulated as a joint sparse recovery problem that recovers multiple sparse vectors with common support. To solve this problem, we develop a joint sparse recovery algorithm under the variational Bayesian inference framework. In this algorithm, dictionaries are parameterized based on the saddle surface model. To adapt to the environmental changes and different channel characteristics, dictionary parameters are modelled as tunable parameters. Simulation results verified the superior performance of ComDec as compared with other state-of-the-art CS-based DFL methods.

Keywords: changing environment; channel diversity; variational Bayesian inference; joint sparse recovery; device-free localization

1. Introduction

As a main piece of the context information, location information is essential for the location-based services (LBS) (all the used abbreviations are explained at the end of the article) in many context-aware applications. Nowadays, localization techniques have been an active field in pervasive and ubiquitous computing. Generally, the target localization methods can be classified as: device-free [1–3] and device-based [4–6]. Device-based methods need objects to attach assistant wireless devices for signal transmitting or receiving. However, in many applications, such as emergency rescue, intruder detection, and smart homes, it is difficult to attach objects with any transceivers. In this case, the device-based method will be infeasible.

Device-free localization (DFL) methods do not need to attach objects with any assistant devices. This approach has become a crucial component in many context-aware applications. As the transceiver-free objects cannot be directly perceived, DFL methods estimate their positions by analyzing their influences on surrounding radio environments. Among current DFL methods, some rely on specialized hardware, such as radar-based methods [7], camera-based methods [8], and infrared-based methods [9]. These methods need an extensive deployment of dedicated devices in the monitoring area and may involve privacy issues. In recent years, DFL methods based on existing infrastructures (e.g.,

WSNs/WiFi) have attracted a lot of research interests. They do not require dedicated hardware and only make use of the measurement information available from the already deployed wireless devices.

As a common link measurement, the received signal strength (RSS) is nearly ubiquitously available from the standard radio transceivers [10,11]. The RSS measurement information has been widely used in DFL. When targets located at different positions, the RSS readings will be different. To exploit the location dependence of RSS, RSS readings on multiple wireless links are recorded before and after targets entering into the monitoring area. To establish the relationship between RSS variations and target locations, the fingerprinting technique has been introduced [12]. Typically, the fingerprinting-based DFL consists of an online and offline phase. In the offline training phase, by gathering the RSS variations caused by a target at every possible location, a radio map can be built. In the online locating phase, target locations are estimated by matching current RSS variations with the fingerprints in radio map. However, as a major drawback of the fingerprinting-based DFL methods, gathering fingerprints to build a radio map is labor-intensive and time-consuming.

As an alternative to fingerprinting-based methods, model-based methods have been widely used in DFL. They map RSS variations to target positions by theoretical or empirical shadowing models. Based on these models, the dictionary (a.k.a. radio map) can be built without site survey. Unfortunately, in changing environments, RSS is extremely sensitive to environmental dynamics, such as temperature, humidity, electromagnetic characteristics, and pedestrians around [13]. In this case, shadowing effects cannot be well approximated by RSS variations. The spatial and temporal environmental dynamics may result in mismatches between runtime RSS measurements and the data in a fixed dictionary. In [14], a dictionary refinement algorithm is introduced, which alleviates the dictionary mismatches by iteratively refining the model-based dictionary. However, the real-time optimization of dictionary parameters leads to a high computational complexity. To avoid this, in ComDec, the dictionary parameters are modelled as tunable parameters to adapt to the changes of environment, and we do not need to explicitly estimate their values.

Recent years, in target localization, the channel state information (CSI) [15] has been exploited, which is a measurement information from PHY layer. As a fine-grained value, CSI depicts the channel quality on multiple orthogonal subcarriers. In wireless environments, due to the frequency-independent attenuation and frequency-selective fading, each channel will exhibit a unique amplitude and phase. Hence, the dictionaries corresponding to different channels are different. Over the past few years, in many device-based localization methods, CSI measurements has been leveraged [16]. Recently, more efforts have been paid in device-free technique. In [17], the fine-grained subcarrier information is exploited based on the multitask Bayesian compressive sensing (MBCS). However, the DFL method proposed in [17] do not provide a solution to the dictionary mismatch problem caused by environmental dynamics.

Compared with RSS measurements, CSI measurements are more robust and suitable for being utilized in DFL. In this paper, a novel ComDec method is proposed. It can enhance the accuracy and robustness of the CS-based DFL in changing environments. In ComDec, to enrich the measurement information, CSI measurements are collected from multiple frequency bands. Moreover, to reduce wireless sensor nodes, the compressive sensing (CS) theory [18] is applied in ComDec by taking advantage of the spatial sparsity of target localization. The CS-based DFL problem in multi-channel scenario is formulated as a joint sparse recovery problem. Furthermore, to bypass the dictionary training and retraining works, the dictionaries are built based on the saddle surface model [19]. We treat dictionary parameters as tunable parameters to adapt the changes of environment and different channel characteristics. Afterwards, to recover the sparse vectors of multiple channels, we develop a joint sparse recovery algorithm under the variational Bayesian inference framework [20]. The main contributions of this work are as follows:

- To enhance the localization accuracy and robustness of CS-based DFL, a novel ComDec method is proposed, which leverages the channel diversity of CSI measurements. In ComDec, the CS-based

DFL problem is extended to multi-channel scenario. It is formulated as a joint sparse recovery problem that recovers multiple jointly sparse vectors over two known dictionaries.

- To simultaneously recover the jointly sparse vectors, we develop a novel joint sparse recovery algorithm. The joint sparsity of the sparse vectors is induced by a novel two-layer hierarchical prior model. Then, the common support set of the sparse vectors is estimated by inferring the posteriors of the hidden variables that defined in the proposed prior model.
- To mitigate the influence of environmental dynamics in changing environments, the dictionary parameters with respect to multiple channels are modelled as tunable parameters to adapt the environmental changes and different channel characteristics. In this way, the dictionary mismatch problem can be solved without the need of explicitly estimating the dictionary parameters.
- To reduce the computational complexity, we introduce four methods in the proposed joint sparse recovery algorithm. Among them, the grid pruning method can improve the convergence speed of the proposed joint sparse recovery algorithm.

The remainder of this paper is organized as follows. Section 2 presents an overview of the related works on multi-target DFL. Section 3 gives the signal model and formulates the CS-based DFL problem as a joint sparse recovery problem. Section 4 proposes a novel joint sparse recovery algorithm to recover the jointly sparse vectors. Section 5 validates the proposed ComDec method with extensive numerical simulations. Finally, Section 6 concludes this paper.

2. Related Work

Beginning with the initial papers of Youssef et al. [1] and Zhang et al. [2], numerous research works on DFL have been carried out [17,21]. Typically, there are four types of DFL methods: (1) geometry-based methods, (2) fingerprinting-based methods, (3) radio tomographic imaging (RTI)-based methods, and (4) CS-based methods. The geometry-based methods estimate target locations based on the geometry information of wireless links. They have a restriction on target spacing and require to know the deployment information of wireless nodes. Fingerprinting-based methods estimate target locations by matching runtime measurements with fingerprints in a radio map. They can achieve enhanced accuracy, but a site survey is needed for radio map building and retraining. RTI-based methods do not require offline training effort. They treat target locations as the attenuation images of distorted links and can achieve an improved performance. However, they need a dense deployment of wireless links. It may lead to high hardware cost and great energy consumption. CS-based methods formulate the DFL problem as a sparse signal reconstruction problem. Compared with the aforementioned DFL methods, the CS-based DFL methods can ensure a high accuracy with fewer measurements. LCS [22] is a CS-based DFL method, in which the model-based dictionary satisfies the restricted isometry property (RIP) with high probability. E-HIPA [23] is a representative CS-based DFL method. It adopts an adaptive orthogonal matching pursuit (OMP) algorithm to estimate the target number and location vector. DR-DFL [14] introduced a real-time dictionary refinement algorithm for CS-based DFL. It can mitigate the influence of environmental dynamics in changing environments. The dictionary refinement is realized by optimizing the environment-related dictionary parameters. However, explicitly estimating the dictionary parameters may lead to a high computational complexity.

To leverage CSI for DFL, Pilot [24] regards the correlations of CSI as fingerprints and uses the maximum a posteriori probability (MAP) estimator to estimate target location. MonoPHY [25] is a fingerprinting-based DFL method that leverages CSI to locate a target with only one stream. Gao et al. [26] transformed CSI measurements into a radio image and adopted the machine learning method to estimate the position of a person. The aforementioned DFL method leverage CSI by fingerprinting technique. They seek appropriate location-dependent CSI features for DFL and trying to build a robust and precise relationship between CSI measurements and target position. However, to make full use of the channel diversity of CSI measurements is still a challenging issue in CS-based DFL.

3. Preliminaries and Problem Formulation

3.1. Overview of Multi-Target Device-Free Localization

The proposed ComDec method is a CS-based DFL method that aims to estimate the number and locations of multiple transceiver-free objects in changing environments. Figure 1 shows an example of the CS-based DFL. As can be seen, multiple transceiver-free objects are randomly distributed in an $l \times l$ two-dimensional (2D) monitoring area \mathcal{A} . Owing to the inherent spatial sparsity of the targets in \mathcal{A} , the position information of multiple targets can be considered as a sparse signal. To achieve this, we discretize \mathcal{A} into N equal-sized grids. When a target is located in grid n , we regard the grid center as its position. In this case, the target number and position information can be encoded in a sparse location vector, i.e.,

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_n, \dots, \theta_N]^T, \quad (1)$$

where $\boldsymbol{\theta} \in \mathbb{R}^{N \times 1}$ is a K -sparse vector. $\theta_n \in \{0, 1\}$ is its n -th component, which indicates whether a target is located in grid n . If there is a target, $\theta_n = 1$; otherwise, $\theta_n = 0$. Thus, the location information of K targets can be denoted as $\mathbb{L} = \{(x_n, y_n) | \theta_n = 1, n \in \{1, \dots, N\}\}$, where (x_n, y_n) is the coordinate of grid n . Moreover, the target number can be denoted as $K = \|\boldsymbol{\theta}\|_0$. When the target number K is far less than the grid number N ($K \ll N$), the sparsity of $\boldsymbol{\theta}$ can be ensured. As a matter of fact, the grid number N is usually a big number, while there is only a small number of targets in the monitoring area.

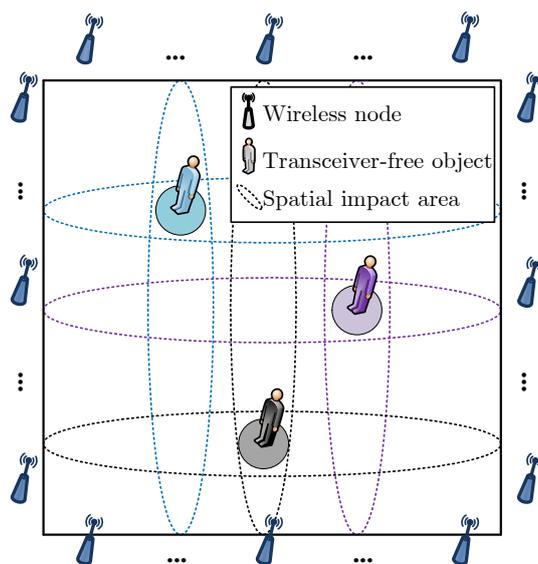


Figure 1. An example of the CS-based DFL.

As seen from Figure 1, \mathcal{A} is surrounded by several wireless nodes. Between these nodes, M bidirectional wireless links are established to cover \mathcal{A} . They travel through \mathcal{A} and sense the target-induced shadowing effects in the electromagnetic field. When there exists a target, some of the wireless links will be shadowed, and the signal power or other features of radio signals will be affected due to the scattering, reflection, refraction, and absorption of the signals. Fortunately, the changes of the signal features are closely related to target positions. Therefore, by analyzing the target-induced shadowing effects, the location information of multiple targets can be inferred.

3.2. CSI Collection and Feature Extraction

RSS is a coarse-grained measurement information from MAC layer. It represents the overall signal power across all subchannels. Due to the multipath fading, RSS is unreliable and varies with time even in a static environment. This may result in limited localization accuracy, especially in changing environments. To enhance the localization performance of CS-based DFL, ComDec adopts

CSI to characterize target-induced perturbations. CSI is a fine-grained measurement information from PHY layer. It can provide the phase and strength information of the signals on different subchannels. For link m , the CSI on each subchannel is a complex value which is defined as

$$H_m^f = |H_m^f| e^{j\sin\{\angle H_m^f\}}, \quad f = 1, 2, \dots, F, \quad (2)$$

where H_m^f denotes the CSI measurement corresponding to the f -th channel of link m . $|H_m^f|$ and $\angle H_m^f$ denote the amplitude and phase response, respectively. F is the number of channels. The amplitude response $|H_m^f|$ is the change of amplitude of link m on channel f . By converting it from linear space to logarithmic space, the corresponding power fading can be written as $\tilde{H}_m^f = 20 \log(|H_m^f|/10^3)$ (dBm) [27]. We collect the CSI measurements from F channels, and a set of power fading information with channel diversity can be obtained. They can provide redundancy information to alleviate the location ambiguity that is incurred by environmental dynamics.

3.3. Problem Formulation

To characterize shadowing effects, we measure the the change of power fading on each link. For each channel, the change of power fading on link m is

$$\Delta \tilde{H}_m^f = \tilde{H}_m^f - \tilde{H}_m^f(0) \approx S_m^f + \epsilon_m^f, \quad (3)$$

where \tilde{H}_m^f denotes the current power fading, $\tilde{H}_m^f(0)$ is the reference power fading recorded when \mathcal{A} is vacant, S_m^f represents the target-induced shadowing loss, and ϵ_m^f denotes the measurement noise.

The changes of power fading are measured on M links at runtime. For each link, we can obtain F measurements from F channels. Thus, the measurement information can be represented as

$$\mathbf{Y} = [\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^f, \dots, \mathbf{y}^F] = \begin{bmatrix} y_1^1 & y_1^2 & \cdots & y_1^F \\ y_2^1 & y_2^2 & \cdots & y_2^F \\ \vdots & \vdots & \ddots & \vdots \\ y_M^1 & y_M^2 & \cdots & y_M^F \end{bmatrix}_{M \times F}, \quad (4)$$

where $\mathbf{Y} \in \mathbb{R}^{M \times F}$ is the measurement matrix. $\mathbf{y}^f \in \mathbb{R}^{M \times 1}$ is the f -th column of \mathbf{Y} . Its m -th component $y_m^f = \Delta \tilde{H}_m^f$.

If any two targets are located sparsely [23] in the monitoring area, S_m^f can be represented as the summation of attenuations that occur in each cell [28]. Therefore, $\Delta \tilde{H}_m^f$ can be expressed as

$$\Delta \tilde{H}_m^f = \sum_{n=1}^N \theta_n \cdot \tilde{h}_{m,n}^f + \epsilon_m^f, \quad (5)$$

where $\tilde{h}_{m,n}^f$ is the shadowing loss that caused by a target located in grid n . For M links, the measurement vector \mathbf{y}^f can be written as

$$\mathbf{y}^f = \Phi^f \boldsymbol{\theta} + \boldsymbol{\epsilon}^f, \quad (6)$$

where $\Phi^f \in \mathbb{R}^{M \times N}$ is a dictionary. Its (m, n) -th element is $\tilde{h}_{m,n}^f$. $\boldsymbol{\epsilon}^f \in \mathbb{R}^{M \times 1}$ represents the noise vector of channel f , and its m -th component is ϵ_m^f . In ComDec, the DFL problem can be viewed as a problem of reconstructing the location vector $\boldsymbol{\theta}$ from the measurements \mathbf{Y} . Theoretically, this problem can be solved by existing joint sparse recovery methods. Nonetheless, as a common shortcoming, they require the true dictionaries $\{\Phi^f\}_{f=1}^F$ to be known in advance. However, in changing environments, it is impossible for us to accurately estimate these dictionaries.

Due to the difference in channel characteristics, the dictionaries with respect to different channels are different. Furthermore, in changing environments, the dictionary for an individual channel

may differ when observed at different times. In this case, the fingerprinting method will be infeasible because it may cumulate the effects of dictionary mismatches on multiple channels and deteriorate the localization performance significantly. As an alternative, we adopt the saddle surface model to characterize the shadowing effect and establish multiple model-based dictionaries for multiple channels. To adapt to the changes of environment and different channel characteristics, the environmental parameters in these dictionaries are considered as adjustable. For simplicity, we denote $\phi_{m,n}^f = \tilde{h}_{m,n}^f$, and thus Φ^f can be expressed as

$$\Phi^f = [\phi_1^f, \phi_2^f, \dots, \phi_m^f, \dots, \phi_M^f]^T = \begin{bmatrix} \phi_{1,1}^f & \phi_{1,2}^f & \cdots & \phi_{1,N}^f \\ \phi_{2,1}^f & \phi_{2,2}^f & \cdots & \phi_{2,N}^f \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M,1}^f & \phi_{M,2}^f & \cdots & \phi_{M,N}^f \end{bmatrix}_{M \times N}. \quad (7)$$

Figure 2 depicts the spatial impact area of a wireless link. As can be seen, the spatial impact area is an ellipse area, and we set up an U - V coordinate system in the area. Based on the saddle surface model, $\phi_{m,n}^f$ can be expressed as

$$\begin{aligned} \phi_{m,n}^f &= \gamma_m^f \cdot \left(\frac{1 - \rho_m^f}{\lambda_1^2} U_{m,n}^2 + \rho_m^f \cdot \left(1 - \frac{V_{m,n}^2}{\lambda_2^2} \right) \right), \\ \text{s.t. } & \frac{U_{m,n}^2}{\lambda_1^2} + \frac{V_{m,n}^2}{\lambda_2^2} \leq 1, \end{aligned} \quad (8)$$

where $(U_{m,n}, V_{m,n})$ denotes the coordinate of grid n . λ_1 and λ_2 denote the semi-major and semi-minor axes of the spatial impact area, respectively. In the saddle surface model, γ_m^f represents the maximum shadowing loss, and $\rho_m^f \in (0, 1]$ denotes the shadowing ratio, which represents the normalized shadowing loss at the midpoint of LOS path.

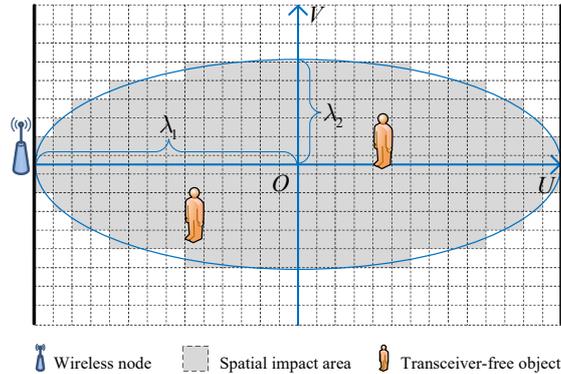


Figure 2. Spatial impact area of a wireless link.

Due to \mathcal{A} being an isotropic free space and all links sharing a common link length, we have $\gamma_m^f = \gamma^f$ and $\rho_m^f = \rho^f$. This means that the values of γ^f and ρ^f are only determined by the environmental characteristics. In this case, the shadowing loss $\phi_{m,n}^f$ can be rewritten as

$$\phi_{m,n}^f = \frac{U_{m,n}^2}{\lambda_1^2} \gamma^f + \gamma^f \rho^f \left(1 - \frac{U_{m,n}^2}{\lambda_1^2} - \frac{V_{m,n}^2}{\lambda_2^2} \right). \quad (9)$$

Let $\omega^f = \gamma^f \cdot \theta$ and $\mathbf{v}^f = \gamma^f \rho^f \cdot \theta$ be the unknown sparse vectors. Based on (6) and (9), \mathbf{y}^f can be rewritten as

$$\mathbf{y}^f = \mathbf{\Psi} \omega^f + \mathbf{\Psi}' \mathbf{v}^f + \epsilon^f, \quad (10)$$

where $\Psi \in \mathbb{R}^{M \times N}$ and $\Psi' \in \mathbb{R}^{M \times N}$ are known dictionaries. Their (m, n) -th elements are computed as $\psi_{m,n} = U_{m,n}^2/\lambda_1^2$ and $\psi'_{m,n} = (1 - U_{m,n}^2/\lambda_1^2 - V_{m,n}^2/\lambda_2^2)$, respectively. We can determine their values before the localization stage according to the deployment of wireless links. It is noteworthy that the location vector θ and sparse vectors $\mathbb{S} = \{\omega^f, v^f\}_{f=1}^F$ share a common sparse support set $\mathbb{T} \subseteq \{1, 2, \dots, N\}$. In other words, their nonzero entries are concentrated at some common locations.

As pointed out earlier, the values of $\{\gamma^f, \rho^f\}_{f=1}^F$ are closely related to the environmental characteristics, which are different for different channels and times. Let $\gamma = [\gamma^1, \gamma^2, \dots, \gamma^F]$ and $\rho = [\rho^1, \rho^2, \dots, \rho^F]$. We define two matrices $\Theta, \Theta' \in \mathbb{R}^{N \times F}$, which are constructed as follows:

$$\Theta = \theta\gamma = [\omega^1, \omega^2, \dots, \omega^F], \tag{11}$$

$$\Theta' = \theta(\gamma \odot \rho) = [v^1, v^2, \dots, v^F]. \tag{12}$$

Both Θ and Θ' are K jointly sparse matrices. This means there are at most K rows in them that have nonzero elements. The column vectors of Θ and Θ' share the common support \mathbb{T} , which is the index set of the grids where a target exists. The cardinality of \mathbb{T} is $K := |\text{supp}\Theta| = |\text{supp}\Theta'|$, which also denotes the target number. Based on (10), (11) and (12), the measurement matrix \mathbf{Y} can be expressed as

$$\mathbf{Y} = \Psi\Theta + \Psi'\Theta' + \mathbf{E}, \tag{13}$$

where $\mathbf{E} \in \mathbb{R}^{M \times F}$ is the matrix of measurement noises, whose f -th column is ϵ^f . Θ and Θ' are the unknown matrices. By (13), the CS-based DFL problem is recast as a problem that needs to recover Θ and Θ' simultaneously over two known dictionaries. It should be pointed out that the problem is different from the conventional joint sparse recovery problem, which only needs to reconstruct a single sparse matrix over a single sparsifying dictionary. However, in this problem, there are two different sparse matrices to be estimated. Therefore, existing joint sparse recovery algorithms cannot be directly applied in ComDec. In this context, the key issue of ComDec is to design a joint sparse recovery algorithm to estimate Θ and Θ' simultaneously. It should be noted that, the channel diversity of CSI measurements can be exploited to enhance the accuracy and robustness of the joint sparse recovery.

Figure 3 shows an architectural overview of ComDec. The proposed ComDec consists of four main modules: dictionary construction, measurement information collection, joint sparse recovery, and location estimation. The process of target counting and localization is illustrated as follows: First, in the dictionary construction module, we establish two constant dictionaries Ψ and Ψ' according to the saddle surface model. Then, at runtime, CSI measurements are collected from M links. Each stream contains the CSI readings of F channels. Afterwards, in the joint sparse recovery module, the posteriors of all hidden variables are inferred by a joint sparse recovery algorithm. Finally, with the knowledge of the posteriors, in the location estimation module, the common support set \mathbb{T} can be estimated, and thus the estimated Cartesian coordinates of multiple targets can be obtained.

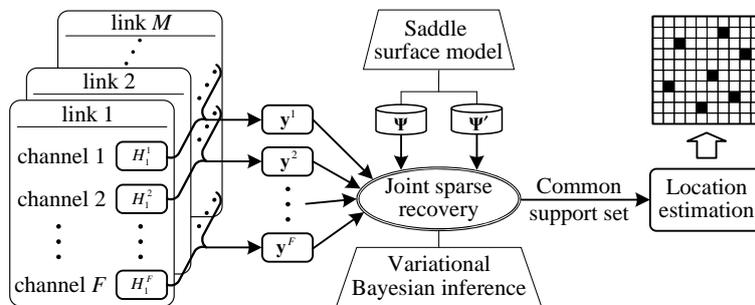


Figure 3. Architecture of the ComDec method.

4. Target Localization via Variational Bayesian Inference

In this section, we develop a novel joint sparse recovery algorithm. First, to induce the joint sparsity of the jointly sparse vectors \mathbb{S} , a two-layer hierarchical prior model is introduced. Then, by using the variational Bayesian inference technique, we infer the posteriors of the hidden variables that defined in the proposed hierarchical prior model. Finally, based on the posteriors of \mathbb{S} , target counting and localization are implemented.

4.1. Hierarchical Prior Model

The joint sparsity of \mathbb{S} is induced by a non-separable sparsity inducing prior model [29]. The graphical model for the joint sparse recovery is shown in Figure 4, which describes the dependencies between random variables. In the first layer, we regard $\{\omega^f, v^f\}_{f=1}^F$ as stochastic variables and define $\omega = \{\omega^f\}_{f=1}^F$ and $v = \{v^f\}_{f=1}^F$. Moreover, the Gaussian-inverse-Gamma prior is imposed on each sparse vector to encourage its sparsity. ω^f is treated as a Gaussian random variable, whose prior distribution is can be given as

$$p(\omega^f | \alpha) = \prod_{n=1}^N \mathcal{N}(\omega_n^f | 0, \alpha_n^{-1}) = (2\pi)^{-N/2} |\Lambda|^{1/2} \exp\left(-\frac{1}{2}(\omega^f)^T \Lambda \omega^f\right), \quad (14)$$

where $\mathcal{N}(\cdot | 0, \alpha_n^{-1})$ denotes the Gaussian distribution with zero mean and variance of α_n^{-1} . α_n is the inverse variance (precision) of $\{\omega_n^f\}_{f=1}^F$. We define $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$ and $\Lambda = \text{diag}(\alpha)$. As $v^f = \rho^f \omega^f$, where ρ^f is regarded as an unknown deterministic parameter, a multivariate Gaussian prior is also imposed on v^f . The variance of v_n^f can be given as $(\rho^f)^2 \alpha_n^{-1}$. We set $\rho^f = 1$ to accommodate the worst case of the variance. In this situation, the prior of v^f can be expressed as

$$p(v^f | \alpha) = \prod_{n=1}^N \mathcal{N}(v_n^f | 0, \alpha_n^{-1}). \quad (15)$$

The variances of v_n^f and ω_n^f are identical. Obviously, a large value of α_n will simultaneously drive v_n^f and ω_n^f to zero, which is in correspondence with the joint sparsity between v^f and ω^f . Intuitively, if most of the components of α have large values, the jointly sparse solutions will be obtained. Hence, we model the hyperparameters $\{\alpha_n\}_{n=1}^N$ as hidden variables to allow the flexibility to learn and adapt to the true situation.

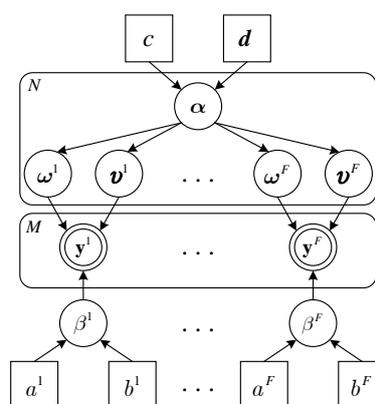


Figure 4. Graphical model for the joint sparse recovery.

To allow conjugate-exponential analysis [20], an independent identically distributed (i.i.d.) Gamma distribution is imposed on α , which can be expressed as

$$p(\alpha; c, \mathbf{d}) = \prod_{n=1}^N \text{Gamma}(\alpha_n | c, d_n) = \prod_{n=1}^N \frac{1}{\Gamma(c)} d_n^c \alpha_n^{c-1} \exp(-d_n \alpha_n), \quad (16)$$

where $\text{Gamma}(\cdot | c, d_n)$ denotes the Gamma distribution with parameters c and d_n , $\Gamma(c) = \int_0^\infty x^{c-1} e^{-x} dx$ is the Gamma function, $\mathbf{d} = [d_1, d_2, \dots, d_N]^T$. We set the hyperparameters c and d_n to very small values (e.g., 10^{-6}) to provide a non-informative hyperprior over α_n . The Gamma distribution is generally chosen as the prior for the inverse variance of a Gaussian distribution, because it is the conjugate prior of the Gaussian distribution. In this case, the associated Bayesian inference can be performed in closed form [4].

When a hierarchical Gaussian prior model imposed on ω^f , the true prior distribution of ω^f can be computed by marginalizing the parameter α , i.e.,

$$\begin{aligned} p(\omega^f; c, \mathbf{d}) &= \int p(\omega^f | \alpha) p(\alpha; c, \mathbf{d}) d\alpha \\ &= \int \prod_{n=1}^N \mathcal{N}(\omega_n^f | 0, \alpha_n^{-1}) \text{Gamma}(\alpha_n | c, d_n) d\alpha_n \\ &= \prod_{n=1}^N \text{St}(\omega_n^f | \lambda, v). \end{aligned} \quad (17)$$

In this case, the true distribution of ω^f is a Student-t pdf,

$$\text{St}(x | \kappa, \lambda, v) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)} \left(\frac{\lambda}{\pi v} \right)^{1/2} \left(1 + \frac{\lambda(x - \kappa)^2}{v} \right)^{-(v+1)/2}, \quad (18)$$

with mean $\kappa = 0$, parameter $\lambda = c/d$ and degrees of freedom $v = 2c$. According to the property of the Student-t distribution, when v is small, this distribution will exhibit very heavy tails. Thus, it favours sparse solutions, which include only few of the basis functions and prunes the remaining basis functions by setting the corresponding weights to very small values. In this case, a sparse vector ω^f can be induced with the hierarchical Gaussian prior model. As the case of ω^f , we can also induce the sparsity of v^f by the proposed hierarchical Gaussian prior model.

As illustrated earlier, ϵ^f follows an i.i.d. Gaussian distribution. The prior distribution of ϵ^f is defined as

$$p(\epsilon^f | \beta^f) = \prod_{n=1}^N \mathcal{N}(\epsilon_n^f | 0, (\beta^f)^{-1}) = \left(\frac{\beta^f}{2\pi} \right)^{\frac{M}{2}} \exp\left(-\frac{\beta^f}{2} \|\epsilon^f\|_2^2\right), \quad (19)$$

where β^f is the inverse variance of ϵ^f . We treat β^f as a hidden variable and define $\beta = \{\beta^f\}_{f=1}^F$. As ϵ^f follows a Gaussian distribution, a Gamma prior is also imposed on each β^f , i.e.,

$$p(\beta^f; a^f, b^f) = \text{Gamma}(\beta^f | a^f, b^f) = \frac{(b^f)^{a^f} (\beta^f)^{a^f-1}}{\Gamma(a^f)} \exp(-b^f \beta^f), \quad (20)$$

where a^f and b^f are deterministic parameters of the Gamma distribution. We denote $\mathbf{a} = \{a^f\}_{f=1}^F$ and $\mathbf{b} = \{b^f\}_{f=1}^F$. To assume uninformative priors for β , the hyperparameters \mathbf{a} and \mathbf{b} are also set to very small values (e.g., 10^{-6}).

In the proposed graphical model, the observed variables are $\mathbf{y} = \{\mathbf{y}^f\}_{f=1}^F$, and the hidden variables are $\mathbf{z} \triangleq \{\alpha, \omega, \mathbf{v}, \beta\}$. To estimate the jointly sparse vectors, we need to infer the posterior distributions of \mathbf{z} based on the predefined prior evidence and the measurement data. In addition,

the deterministic parameters of the prior model are $\Omega \triangleq \{c, d, a, b\}$, which are fixed at small values to allow uninformative hyperpriors for α and β .

4.2. Variational Bayesian Inference

In the joint sparse recovery module, the key task is to infer the posteriors of \mathbf{z} . Afterwards, based on these posteriors, the target number and locations can be estimated. For this objective, the variational Bayesian inference technique is adopted, which is applied due to it can deal with complicated Bayesian models [20]. Based on (10) and (19), the likelihood function of channel f can be written as

$$p(\mathbf{y}^f | \omega^f, \mathbf{v}^f, \beta^f) = \left(\frac{\beta^f}{2\pi} \right)^{\frac{M}{2}} \exp \left(-\frac{\beta^f}{2} \|\mathbf{y}^f - \Psi \omega^f - \Psi' \mathbf{v}^f\|_2^2 \right). \quad (21)$$

According to the chain rule of probability, the joint probability density function (PDF) of \mathbf{y} and \mathbf{z} can be written as

$$\begin{aligned} p(\mathbf{y}, \mathbf{z}) &= p(\mathbf{y} | \omega, \mathbf{v}, \beta) p(\beta) p(\omega | \alpha) p(\mathbf{v} | \alpha) p(\alpha) \\ &= \prod_{f=1}^F p(\mathbf{y}^f | \omega^f, \mathbf{v}^f, \beta^f) \prod_{f=1}^F p(\beta^f) \prod_{f=1}^F \prod_{n=1}^N p(\omega_n^f | \alpha_n) \prod_{f=1}^F \prod_{n=1}^N p(\mathbf{v}_n^f | \alpha_n) \prod_{n=1}^N p(\alpha_n). \end{aligned} \quad (22)$$

Figure 5 illustrates the factor graph representation of the joint PDF. In this figure, the circle nodes represent the random variables that the complicated global function relied on, while the square nodes represent the local functions.

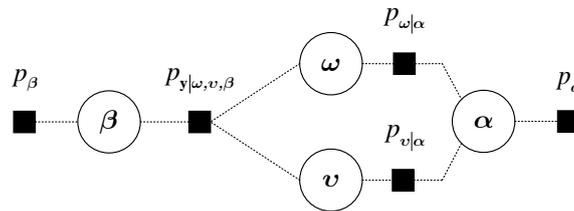


Figure 5. Factor graph representation of the joint PDF (22).

For an arbitrary density function $q(\mathbf{z})$, the evidence $p(\mathbf{y}) = \int p(\mathbf{y}, \mathbf{z}) d\mathbf{z}$ can be decomposed as

$$\ln p(\mathbf{y}) = \mathcal{F}(q; \Omega) + KL(q \| p), \quad (23)$$

where

$$\mathcal{F}(q; \Omega) = \int q(\mathbf{z}) \ln \left(\frac{p(\mathbf{y}, \mathbf{z}; \Omega)}{q(\mathbf{z})} \right) d\mathbf{z}, \quad (24)$$

$$KL(q \| p) = - \int q(\mathbf{z}) \ln \left(\frac{p(\mathbf{z} | \mathbf{y}; \Omega)}{q(\mathbf{z})} \right) d\mathbf{z}. \quad (25)$$

$\mathcal{F}(q; \Omega)$ is a lower bound of $\ln p(\mathbf{y})$, and $KL(q \| p)$ represents the Kullback-Leibler divergence (KLD) between the approximated posterior $q(\mathbf{z})$ and the exact posterior $p(\mathbf{z} | \mathbf{y}; \Omega)$. The proposed joint sparse recovery algorithm maximizes $\ln p(\mathbf{y})$ iteratively. At each iteration, we set $KL(q \| p) = 0$ to minimize the KLD and update $q(\mathbf{z})$ accordingly. This will lead to the lower bound $\mathcal{F}(q, \Omega)$ increasing to $\ln p(\mathbf{y})$. Meanwhile, the updating of $q(\mathbf{z})$ may enlarge $\ln p(\mathbf{y})$ and lead to a new non-negative KLD. We will minimize the new KLD and update $q(\mathbf{z})$ in the next iteration. By doing so, the log-likelihood $\ln p(\mathbf{y})$ will be maximized, and the approximated posterior $q(\mathbf{z})$ can be optimized iteratively.

However, $p(\mathbf{z}|\mathbf{y};\Omega)$ cannot be computed analytically. Thereby, directly updating $q(\mathbf{z})$ is intractable. To bypass this difficulty, we resort to the variational approximation method. It assumes that the posteriors of \mathbf{z} are independent, i.e.,

$$q(\mathbf{z}) = q(\boldsymbol{\alpha})q(\boldsymbol{\omega})q(\mathbf{v})q(\boldsymbol{\beta}) = q(\boldsymbol{\alpha}) \prod_{f=1}^F q(\boldsymbol{\omega}^f) \prod_{f=1}^F q(\mathbf{v}^f) \prod_{f=1}^F q(\boldsymbol{\beta}^f). \tag{26}$$

By applying the above assumption, in each iteration, the log-posterior of $z_i \in \mathbf{z}$ can be approximated as the expectation of the joint PDF with respect to other hidden variables $(\mathbf{z} - z_i)$. More specifically, the log-posteriors of \mathbf{z} are approximated as

$$\ln q(\boldsymbol{\omega}^f) = \langle \ln p(\mathbf{y}, \mathbf{z}) \rangle_{q(\boldsymbol{\alpha}) \prod_{i \neq f} q(\boldsymbol{\omega}^i) q(\mathbf{v}) q(\boldsymbol{\beta})} + \xi, \tag{27}$$

$$\ln q(\mathbf{v}^f) = \langle \ln p(\mathbf{y}, \mathbf{z}) \rangle_{q(\boldsymbol{\alpha}) q(\boldsymbol{\omega}) \prod_{i \neq f} q(\mathbf{v}^i) q(\boldsymbol{\beta})} + \xi, \tag{28}$$

$$\ln q(\boldsymbol{\beta}^f) = \langle \ln p(\mathbf{y}, \mathbf{z}) \rangle_{q(\boldsymbol{\alpha}) q(\boldsymbol{\omega}) q(\mathbf{v}) \prod_{i \neq f} q(\boldsymbol{\beta}^i)} + \xi, \tag{29}$$

$$\ln q(\boldsymbol{\alpha}) = \langle \ln p(\mathbf{y}, \mathbf{z}) \rangle_{q(\boldsymbol{\omega}) q(\mathbf{v}) q(\boldsymbol{\beta})} + \xi, \tag{30}$$

where ξ denotes a normalizing constant. It makes the corresponding $q(\cdot)$ a true PDF. The update rule for each hidden variable is derived below.

In (27), the terms independent of $\boldsymbol{\omega}^f$ can be treated as a constant value. Keeping only the terms that are related to $\boldsymbol{\omega}^f$, $\ln q(\boldsymbol{\omega}^f)$ can be given as

$$\ln q(\boldsymbol{\omega}^f) \propto \left\langle \ln p(\mathbf{y}^f | \boldsymbol{\omega}^f, \mathbf{v}^f, \boldsymbol{\beta}^f) + \ln p(\boldsymbol{\omega}^f | \boldsymbol{\alpha}) \right\rangle_{q(\boldsymbol{\alpha}) q(\mathbf{v}^f) q(\boldsymbol{\beta}^f)}. \tag{31}$$

Note that $q(\boldsymbol{\omega}^f)$ and $q(\mathbf{v}^f)$ follow the multivariate Gaussian distribution. We assume they have the following forms after the updating

$$q(\boldsymbol{\omega}^f) = \mathcal{N}(\boldsymbol{\omega}^f | \boldsymbol{\mu}_{\boldsymbol{\omega}}^f, \boldsymbol{\Sigma}_{\boldsymbol{\omega}}^f), \tag{32}$$

$$q(\mathbf{v}^f) = \mathcal{N}(\mathbf{v}^f | \boldsymbol{\mu}_{\mathbf{v}}^f, \boldsymbol{\Sigma}_{\mathbf{v}}^f), \tag{33}$$

where $\boldsymbol{\mu}_{\boldsymbol{\omega}}^f$ and $\boldsymbol{\mu}_{\mathbf{v}}^f$ are the mean vectors, $\boldsymbol{\Sigma}_{\boldsymbol{\omega}}^f$ and $\boldsymbol{\Sigma}_{\mathbf{v}}^f$ are the covariance matrices. The update of the posterior distribution is equivalent to seeking appropriate values for the parameters in the approximated posterior distribution. Our goal is to learn the values of the mean vectors and covariance matrices based on the prior distributions and likelihood function. Substituting (14) and (21) into (31), after some rearrangement, yields

$$\ln q(\boldsymbol{\omega}^f) \propto -\frac{1}{2}(\boldsymbol{\omega}^f)^T (\boldsymbol{\Sigma}_{\boldsymbol{\omega}}^f)^{-1} \boldsymbol{\omega}^f + (\boldsymbol{\omega}^f)^T (\boldsymbol{\Sigma}_{\boldsymbol{\omega}}^f)^{-1} \boldsymbol{\mu}_{\boldsymbol{\omega}}^f, \tag{34}$$

where

$$\boldsymbol{\Sigma}_{\boldsymbol{\omega}}^f = (\langle \boldsymbol{\beta}^f \rangle \boldsymbol{\Psi}^T \boldsymbol{\Psi} + \langle \boldsymbol{\Lambda} \rangle)^{-1}, \tag{35}$$

$$\boldsymbol{\mu}_{\boldsymbol{\omega}}^f = \langle \boldsymbol{\beta}^f \rangle \boldsymbol{\Sigma}_{\boldsymbol{\omega}}^f \boldsymbol{\Psi}^T (\mathbf{y}^f - \boldsymbol{\Psi}' \boldsymbol{\mu}_{\mathbf{v}}^f). \tag{36}$$

In the same manner, $\ln q(\mathbf{v}^f)$ can be given as

$$\ln q(\mathbf{v}^f) \propto \left\langle \ln p(\mathbf{y}^f | \boldsymbol{\omega}^f, \mathbf{v}^f, \boldsymbol{\beta}^f) + \ln p(\mathbf{v}^f | \boldsymbol{\alpha}) \right\rangle_{q(\boldsymbol{\alpha}) q(\boldsymbol{\omega}^f) q(\boldsymbol{\beta}^f)}. \tag{37}$$

Substituting (15) and (21) into (37), after some rearrangement, yields

$$\ln q(\mathbf{v}^f) \propto -\frac{1}{2}(\mathbf{v}^f)^T(\boldsymbol{\Sigma}_v^f)^{-1}\mathbf{v}^f + (\mathbf{v}^f)^T(\boldsymbol{\Sigma}_v^f)^{-1}\boldsymbol{\mu}_v^f, \quad (38)$$

where

$$\boldsymbol{\Sigma}_v^f = (\langle \beta^f \rangle \boldsymbol{\Psi}'^T \boldsymbol{\Psi}' + \langle \Lambda \rangle)^{-1}, \quad (39)$$

$$\boldsymbol{\mu}_v^f = \langle \beta^f \rangle \boldsymbol{\Sigma}_v^f \boldsymbol{\Psi}'^T (\mathbf{y}^f - \boldsymbol{\Psi} \boldsymbol{\mu}_\omega^f). \quad (40)$$

Keeping only the terms of (29) that are related to β^f , $\ln q(\beta^f)$ can be given as

$$\ln q(\beta^f) \propto \left\langle \ln p(\mathbf{y}^f | \boldsymbol{\omega}^f, \mathbf{v}^f, \beta^f) + \ln p(\beta^f) \right\rangle_{q(\boldsymbol{\omega}^f)q(\mathbf{v}^f)}. \quad (41)$$

We assume that the posterior of β^f follows a Gamma distribution, i.e.,

$$q(\beta^f) = \text{Gamma}(\beta^f | \tilde{a}^f, \tilde{b}^f), \quad (42)$$

where \tilde{a}^f and \tilde{b}^f denote the deterministic parameters of the updated posterior distribution. To infer them, we substitute (20) and (21) into (41). After some rearrangement, the posterior can be given as

$$\ln q(\beta^f) \propto (\tilde{a}^f - 1) \ln \beta^f - \tilde{b}^f \beta^f, \quad (43)$$

where

$$\tilde{a}^f = a^f + M/2, \quad (44)$$

$$\tilde{b}^f = b^f + \frac{1}{2} \left\| \mathbf{y}^f - \boldsymbol{\Psi} \boldsymbol{\mu}_\omega^f - \boldsymbol{\Psi}' \boldsymbol{\mu}_v^f \right\|_2^2 + \frac{1}{2} \left(\text{tr}(\boldsymbol{\Psi} \boldsymbol{\Sigma}_\omega^f \boldsymbol{\Psi}^T) + \text{tr}(\boldsymbol{\Psi}' \boldsymbol{\Sigma}_v^f \boldsymbol{\Psi}'^T) \right). \quad (45)$$

Similarly, we only keep the terms that are related to $\boldsymbol{\alpha}$ in (30), and thereby $\ln q(\boldsymbol{\alpha})$ can be given as

$$\ln q(\boldsymbol{\alpha}) \propto \left\langle \sum_{f=1}^F \ln p(\boldsymbol{\omega}^f | \boldsymbol{\alpha}) + \sum_{f=1}^F \ln p(\mathbf{v}^f | \boldsymbol{\alpha}) + \ln p(\boldsymbol{\alpha}) \right\rangle_{\prod_{f=1}^F q(\boldsymbol{\omega}^f) \prod_{f=1}^F q(\mathbf{v}^f)}. \quad (46)$$

The approximated posterior of $\boldsymbol{\alpha}$ is assumed to be a multivariate Gamma distribution, i.e.,

$$q(\boldsymbol{\alpha}) = \prod_{n=1}^N \text{Gamma}(\alpha_n | \tilde{c}, \tilde{d}_n), \quad (47)$$

where \tilde{c} and $\{\tilde{d}_n\}_{n=1}^N$ are deterministic parameters. Substituting (14), (15) and (16) into (46), after some rearrangement, yields

$$\ln q(\boldsymbol{\alpha}) \propto (\tilde{c} - 1) \sum_{n=1}^N \ln \alpha_n - \sum_{n=1}^N \tilde{d}_n \alpha_n, \quad (48)$$

where

$$\tilde{c} = c + F, \quad (49)$$

$$\tilde{d}_n = d_n + \frac{1}{2} \sum_{f=1}^F \left([\boldsymbol{\mu}_\omega^f]_n^2 + [\boldsymbol{\Sigma}_\omega^f]_{n,n} + [\boldsymbol{\mu}_v^f]_n^2 + [\boldsymbol{\Sigma}_v^f]_{n,n} \right). \quad (50)$$

The notation $[\cdot]_n$ denotes the n -th component of the input vector, and $[\cdot]_{n,n}$ denotes the (n, n) -th entry of the input matrix. Based on the results of posterior inference, the required expectations can be calculated as

$$\langle \beta^f \rangle = \tilde{a}^f / \tilde{b}^f, \quad f = 1, 2, \dots, F. \quad (51)$$

$$\langle \Lambda \rangle = \text{diag}(\langle \alpha_1 \rangle, \langle \alpha_2 \rangle, \dots, \langle \alpha_N \rangle), \quad (52)$$

where

$$\langle \alpha_n \rangle = \tilde{c} / \tilde{d}_n, \quad n = 1, 2, \dots, N. \quad (53)$$

4.3. Joint Sparse Reconstruction

According to the above update rules, we can successively update the posteriors of hidden variables \mathbf{z} . The jointly sparse vectors \mathbb{S} can be reconstructed according to these posteriors. The algorithm of reconstructing \mathbb{S} is summarized as follows:

1. For $f \in \{1, 2, \dots, F\}$, update $q(\omega^f)$ by using (35) and (36); update $q(v^f)$ by using (39) and (40). $\langle \beta^f \rangle$ and $\langle \Lambda \rangle$ are obtained based on the current posteriors of β^f and α .
2. For $f \in \{1, 2, \dots, F\}$, update $q(\beta^f)$ according to (44), (45) and the current posteriors of ω^f and v^f .
3. Update $q(\alpha)$ according to (49), (50) and the current posteriors of ω and v .
4. If a convergence criterion has been met, terminate and choose the posterior means of ω and v as the estimation of \mathbb{S} . Otherwise, go to step 1).

Based on the above joint sparse recovery algorithm, we can obtain a reconstructed sparse vector set $\hat{\mathbb{S}} = \{\mu_\omega^f, \mu_v^f\}_{f=1}^F$. The computational cost of the algorithm is dominated by the matrix inversion operations in (35) and (39), whose computational complexities are $\mathcal{O}(N^3F)$. Moreover, the matrix-vector multiplications in (36) and (40), as well as the matrix multiplication operations in (45) can also incur heavy computational burden. Their computational complexities are $\mathcal{O}(N^2MF)$. Thus, when applying ComDec in large-scale areas (where N is large), the joint sparse recovery algorithm will be computationally expensive. In this paper, we adopt the following four methods to alleviate the computational burden of the proposed algorithm.

In step 1, the covariance matrices Σ_ω^f and Σ_v^f are computed by (35) and (39), which contain matrix inversion operations. Using the matrix inversion lemma [30], the covariance matrices can be evaluated as

$$\Sigma_\omega^f = \langle \Lambda \rangle^{-1} - \langle \Lambda \rangle^{-1} \Psi^T (\Xi_\omega^f)^{-1} \Psi \langle \Lambda \rangle^{-1}, \quad (54)$$

$$\Sigma_v^f = \langle \Lambda \rangle^{-1} - \langle \Lambda \rangle^{-1} \Psi'^T (\Xi_v^f)^{-1} \Psi' \langle \Lambda \rangle^{-1}, \quad (55)$$

where

$$\Xi_\omega^f = \Psi \langle \Lambda \rangle^{-1} \Psi^T + \langle \beta^f \rangle^{-1} \mathbf{I}_M, \quad (56)$$

$$\Xi_v^f = \Psi' \langle \Lambda \rangle^{-1} \Psi'^T + \langle \beta^f \rangle^{-1} \mathbf{I}_M. \quad (57)$$

With these matrix transformations, we only need to compute the inversions of Ξ_ω^f and Ξ_v^f , whose computational complexities are $\mathcal{O}(M^3)$ ($M \ll N$). Note that, $\langle \Lambda \rangle$ is an $N \times N$ diagonal matrix. We can easily obtain its inversion.

As mentioned before, the expectation evaluations of μ_ω^f and μ_v^f can also lead to a high computational cost. To reduce their computational complexities, we reformulate (36) and (40) to cast them as a problem of solving the following linear systems of equations:

$$\left(\Psi^T \Psi + \frac{\langle \Lambda \rangle}{\langle \beta^f \rangle}\right) \mu_\omega^f = \Psi^T (\mathbf{y}^f - \Psi' \mu_v^f), \tag{58}$$

$$\left(\Psi'^T \Psi' + \frac{\langle \Lambda \rangle}{\langle \beta^f \rangle}\right) \mu_v^f = \Psi'^T (\mathbf{y}^f - \Psi \mu_\omega^f). \tag{59}$$

The equations can be solved by the conjugate gradient (CG) algorithm. Theoretically, we can reach the exact solution after N iterations. However, this may incur a considerable computational burden. Fortunately, in practice, a few iterations is sufficient to obtain a good solution. If W ($W \ll N$) iterations are required, the computational complexity of the CG algorithm will be $\mathcal{O}(WN \log N)$.

In step 2, the computational cost is mainly attributed to the matrix multiplications in the last term of (45). To mitigate this, (45) can be rewritten as

$$\tilde{b}^f = b^f + \frac{1}{2} \left\| \mathbf{y}^f - \Psi \mu_\omega^f - \Psi' \mu_v^f \right\|_2^2 + \frac{1}{\langle \beta^f \rangle} \sum_{n=1}^N \left(1 - \frac{\langle \alpha_n \rangle}{2} \left([\Sigma_\omega^f]_{n,n} + [\Sigma_v^f]_{n,n} \right) \right). \tag{60}$$

Here, matrix multiplications are replaced by simple element-wise operations. As a consequence, the computational complexity of step 2 can be reduced to $\mathcal{O}(NMF)$. By using the CG method and the above reformulations, the computational cost per iteration of the proposed joint sparse recovery algorithm can be reduced to $\mathcal{O}(NM^2F)$.

To further reduce the computational load and speed up the convergence, we conduct real-time grid pruning according to the posterior of α . In each iteration, when α_n is sufficiently large, $\{\omega_n^f, v_n^f\}_{f=1}^F$ will be driven to zero. This imply that the contribution of grid n to the signal power fading is negligible. In this case, we can remove grid n from the grid set Π , which is defined as the set of grids where a target possibly exists. With the reduction of grids in Π , the computational load of the next iteration will decrease.

We denote the initial grid set by $\Pi^{(0)} \triangleq \{1, 2, \dots, N\}$. Then, the grid set in each iteration can be updated as

$$\Pi^{(\tau)} = \Pi^{(\tau-1)} - \{n \mid \langle \alpha_n \rangle > \alpha_{th}\}, \tag{61}$$

where τ denotes the iteration number and α_{th} is the threshold of α_n . For the reconstruction of \mathcal{S} , the selection of α_{th} provides a trade-off between the localization accuracy and the convergence speed. After grid pruning, the dimension of θ can be reduced to $N = |\Pi^{(\tau)}|$. In addition, α , \mathcal{S} , Ψ , and Ψ' also should be pruned accordingly.

By virtue of the local convergence property of variational Bayesian inference [20], the proposed joint sparse recovery algorithm is guaranteed to be convergent. The stop criterion is set as the residual error Res smaller than a pre-determined threshold r_{th} , where the Res is defined as

$$Res = \sum_{f=1}^F \left\| \mathbf{y}^f - \Psi \mu_\omega^f - \Psi' \mu_v^f \right\|_2^2. \tag{62}$$

Furthermore, to prevent the computational load from being excessively high, we set a maximum iteration number τ_{max} . In this situation, the iterative algorithm will stop when Res becomes smaller than r_{th} or τ reaches τ_{max} .

4.4. Target Counting and Localization

Based on $\hat{\mathbb{S}}$, the target number and locations can be estimated. However, these vectors are not strictly sparse. They may contain many negligible but nonzero coefficients. Thus, we estimate their common support by the following formulae:

$$\hat{\mathbb{T}} = \left\{ n \mid 20 \lg \left(\frac{[\boldsymbol{\mu}_{\omega}^{\hat{f}}]_n}{\max_{i \in \Pi} [\boldsymbol{\mu}_{\omega}^{\hat{f}}]_i} \right) \geq \mu_{\text{th}}, n \in \Pi \right\}, \quad (63)$$

where

$$\hat{f} = \arg \min_f \left\| \mathbf{y}^f - \boldsymbol{\Psi} \boldsymbol{\mu}_{\omega}^f - \boldsymbol{\Psi}' \boldsymbol{\mu}_v^f \right\|_2^2. \quad (64)$$

The common support of the jointly sparse vectors is estimated based on $\boldsymbol{\mu}_{\omega}^{\hat{f}}$, in which \hat{f} is the channel that has the minimum residual error. In (63), μ_{th} is the sparsity threshold. We use it to filter out the small coefficients in $\boldsymbol{\mu}_{\omega}^{\hat{f}}$. After that, we can estimate the target number and locations as $\hat{K} = |\hat{\mathbb{T}}|$ and $\hat{\mathbb{L}} = \{(x_n, y_n) | n \in \hat{\mathbb{T}}\}$, respectively.

5. Numerical Evaluation

5.1. Simulation Setup

The numerical simulations are carried out in MATLAB R2015b 64 bit version running on a PC with i7-8550U and 16 GB memory. In our simulations, the channel state information is assumed to be collected from the wireless devices that implement an orthogonal frequency division multiplexing (OFDM) system. The wireless devices are assumed to mode at 2.4 GHz with 20 MHz bandwidth. Table 1 summarizes the default values of simulation parameters. The numerical simulations are conducted in a complex radio transmission environment. In this transmission environment, as the impact of environmental dynamics, the environmental parameters are changing with time. Moreover, the measurement of each channel is corrupted by additive white Gaussian noise. The signal-to-noise ratio (SNR) of each channel is defined as

$$\text{SNR} = 10 \lg \frac{\|\boldsymbol{\Phi}^f \boldsymbol{\theta}\|_2^2}{M(\sigma^f)^2}, \quad f = 1, 2, \dots, F, \quad (65)$$

where σ^f is the standard deviation of the measurement noise vector.

In our simulations, all results are averaged over $T = 10^3$ Monte Carlo trials. For each trial, the localization error E_t is computed as

$$E_t = \frac{\sum_{k=1}^{\min\{K, \hat{K}\}} \|L_k - \hat{L}_k\|_2}{\min\{K, \hat{K}\}}, \quad (66)$$

where L_k and \hat{L}_k represent the true and estimated Cartesian coordinates of target k . We use the ‘‘average localization error’’ (*Avg.Error*) and ‘‘root-mean-square error’’ (*Rms.Error*) as the metrics to measure the localization accuracy. The *Avg.Error* is defined as

$$\text{Avg.Error} = \frac{\sum_{t=1}^T E_t}{T}. \quad (67)$$

The *Rms.Error* is defined as

$$\text{Rms.Error} = \sqrt{\frac{\sum_{t=1}^T E_t^2}{T}}. \quad (68)$$

Table 1. Default values of simulation parameters.

Parameters	Explains	Values
F	number of channels	15
K	number of targets	5
M	number of links	56
N	number of grids	784
l	link length	14 m
W	iteration number of CG algorithm	17
α_{th}	pruning threshold	10
τ_{max}	maximum number of iteration	600
r_{th}	residual error threshold	10^{-4}
μ_{th}	sparsity threshold	-20 dB

To evaluate the target counting performance, we introduce another metric *Prob.CoC*. It is the probability of correct counting (i.e., $\hat{K} = K$). We compare the performances of ComDec with other three CS-based DFL approaches: (i) LCS with the GMP algorithm (LCS-GMP) as the CS recovery algorithm [22], (ii) E-HIPA with the adaptive OMP algorithm (E-HIPA-OMP) as the CS recovery algorithm [23], and (iii) DR-DFL with the VEM algorithm for dictionary refinement and sparse recovery (DR-DFL-VEM) [14]. Table 2 reports the computational complexities and localization accuracies of multiple DFL methods. As can be seen from this table, the ComDec method can achieve the lowest average localization error (*Avg.Error*). It can enhance the accuracy and robustness of CS-based DFL by exploiting channel diversity. The computational complexity of the ComDec method is higher than the LCS and E-HIPA methods. The inferiority is mainly attributed to the estimation of the posterior distributions of sparse vectors. It noteworthy that the E-HIPA and LCS methods reconstruct the unknown sparse vector by using the OMP and GMP algorithms, respectively. They are greedy algorithms for sparse recovery, and can only provide a point estimation of the unknown sparse signal. In contrast, the proposed sparse recovery algorithm and VEM algorithm reconstruct sparse signal from a Bayesian perspective. They can provide a posterior belief (distribution function) for the values of the sparse signal, and therefore can achieve an improved accuracy. In real deployment, the sparse recovery algorithm is usually conducted on a server, where the computational resources are abundant. The server collects CSI measurements from all wireless nodes and estimates the target location accordingly. The simulation flowchart is shown in Figure 6. In the upper part of the flowchart, the measurement data is generated. In the lower part of the flowchart, the target locations can be estimated.

Table 2. Computational complexity and localization accuracy.

DFL Method	Sparse Recovery Algorithm	Computational Complexity	<i>Avg.Error</i>	<i>Rms.Error</i>
ComDec	VBI	$\mathcal{O}(NM^2F)$	0.296 m	0.612 m
DR-DFL	VEM	$\mathcal{O}(MN^2)$	0.981 m	1.420 m
E-HIPA	OMP	$\mathcal{O}(KMN + \log(2N))$	3.351 m	3.509 m
LCS	GMP	$\mathcal{O}(MN)$	2.188 m	2.462 m

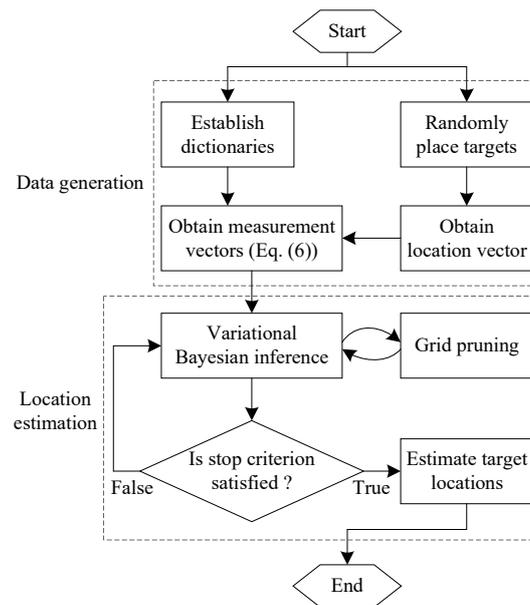


Figure 6. Simulation flowchart.

5.2. Impact of the Number of Channels

In this subsection, the impact of channel number on the performance of ComDec is evaluated. Theoretically, the accuracy of joint sparse recovery is closely related to the channel number. If we increase F , more useful information will be provided. Consequently, the target counting and localization performance will be improved.

In Figure 7, F varies from 1 to 25, and the *Avg.Error* of ComDec decreases as F increases. It is noteworthy that, the *Avg.Error* decreases dramatically when $F < 10$, while decreases slowly when $F > 10$. As in the case of *Avg.Error*, the *Prob.CoC* also increases as F increases, and its growth rate drops gradually. This is because the increase of F leads to a more complex hierarchical prior model, which contains $(3F + 1)$ hidden variables. When reconstructing \mathbb{S} , the posterior distributions of these hidden variables should be inferred. Thus, a large F may deteriorate the accuracy of the posterior inference. From this perspective, increasing the channel number has a negative effect on sparse recovery. In fact, when $F > 20$, the negative effect of increasing F will offset the positive effect contributed by channel diversity. Hence, we set $F = 15$ in our simulation to achieve a trade-off between model simplicity and the performance of target counting and localization.

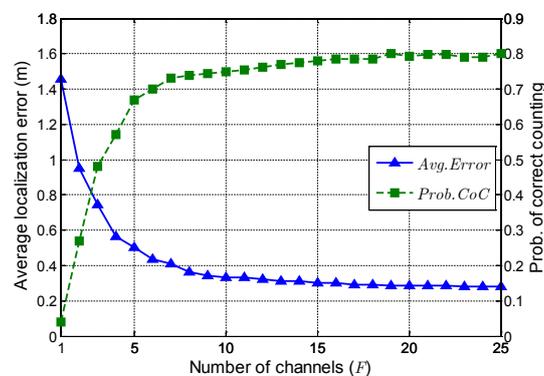


Figure 7. The performance of ComDec when channel number varies from 1 to 25.

5.3. Effectiveness of ComDec in Changing Environments

To mitigate the dictionary mismatches introduced by environmental dynamics, the proposed ComDec method combines the environmental parameters and the location vector to form a set of

jointly sparse vectors. We treat these vectors as random variables and learn their values by variational Bayesian inference. This makes the dictionary parameters can adapt to the environmental changes and different channel characteristics. In this subsection, we test the effectiveness and robustness of ComDec in changing environments.

Firstly, we set up a simulated changing environment. To simulate the environmental dynamics in changing environments, we add Gaussian noises to the environmental parameters (γ^f and ρ^f). Based on these noisy environmental parameters and the saddle surface model, the dictionaries $\{\Phi^f\}_{f=1}^F$ can be built. After that, we can obtain the measurement vectors $\{y^f\}_{f=1}^F$ according to (6). In our simulation, the values of γ^f and ρ^f are calculated by

$$\gamma_t^f = \gamma_0^f + \sum_{i=1}^t \omega_{\gamma}^i, \quad f = 1, 2, \dots, F, \quad (69)$$

$$\rho_t^f = \rho_0^f + \sum_{i=1}^t \omega_{\rho}^i, \quad f = 1, 2, \dots, F, \quad (70)$$

where ω_{γ}^i and ω_{ρ}^i are additive white Gaussian noises whose variances are 0.6 and 0.06, respectively. The setting of the noise variances is according to the real data in changing environments. t represents the times of environmental changes and also denotes the number of additive noises added to the original environmental parameters (γ_0^f and ρ_0^f). In our simulations, we set $\gamma_0^f = 17.0831$ and $\rho_0^f = 0.35122$. In addition, uncorrelated Gaussian noises are added to the dictionaries $\{\Phi^f\}_{f=1}^F$ to produce an SNR of 40 dB. Figure 8 depicts the variations of two dictionary atoms in the simulated changing environment. As expected, the accumulation of additive noises can lead to a deviation of dictionary atoms, which is a simulation of the effect of environmental dynamics in real-world environments.

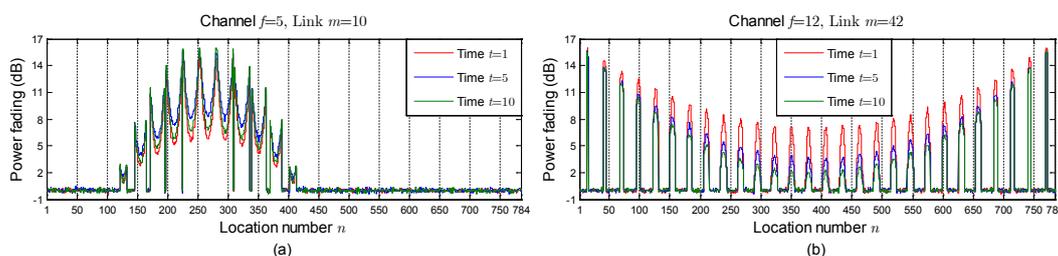


Figure 8. The values of the dictionary atoms: (a) ϕ_{10}^5 ; (b) ϕ_{42}^{12} , when $t = 1, 5$, and 10.

Secondly, in the simulated changing environment, the localization performances of multiple CS-based multi-target DFL methods are evaluated. Figure 9 compares the localization performances when t increases from 1 to 10. The proposed ComDec method that using the variational Bayesian inference technique for joint sparse recovery (ComDec-VBI) can achieve the minimum *Avg.Error*. It should be noted that, its localization performance is robust to the environmental changes. In contrast, with the increase of t , the localization accuracies of other three DFL methods deteriorate gradually. This reveals that these methods are sensitive to environmental dynamics. The simulation results demonstrate the impact of environmental dynamics on *Avg.Error* and verify the accuracy and robustness of the ComDec method.

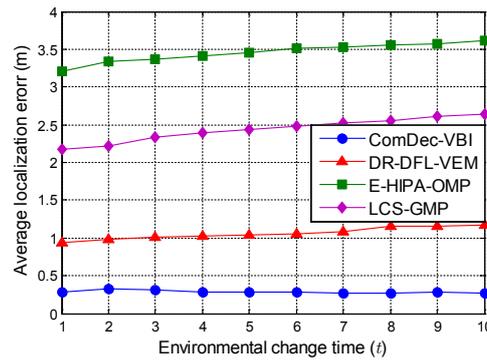


Figure 9. Effect of environmental changes on the average localization error (*Avg.Error*) of multiple DFL approaches.

Finally, we investigate the sparse reconstruction performances of the CS recovery algorithms that adopted in the CS-based DFL methods. We run E-HIPA, LCS, DR-DFL, and the ComDec method with $F = 5$ and $F = 15$ to reconstruct the target location vector. Figure 10 shows an example of the recovered sparse vectors of different DFL approaches. As can be seen, the recovered sparse vectors corresponding to E-HIPA and DR-DFL have many negligible but nonzero coefficients. Although the reconstructed sparse vector in Figure 10c does not have small coefficients, the indices of the nonzero components are different from those of the original location vector. Fortunately, the recovered sparse vector corresponding to our proposed ComDec method has a few small coefficients, and the indices of their significant coefficients are the same as those in the original location vector. Based on them, the target number and target locations can be estimated correctly.

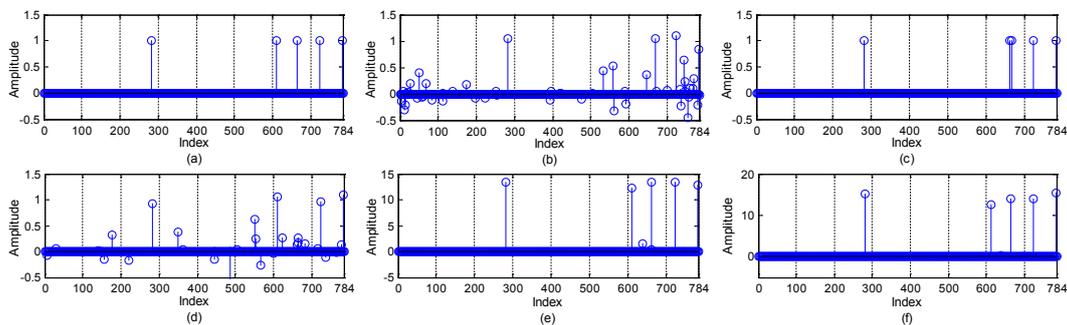


Figure 10. Comparison of the original location vector and the reconstructed sparse vectors of different DFL approaches. (a) the original location vector; The reconstructed sparse vector corresponding to (b) E-HIPA; (c) LCS; (d) DR-DFL; (e) ComDec ($F = 5$); (f) ComDec ($F = 15$).

Figure 11 shows the target positions estimated by multiple CS-based DFL methods. As can be seen, the ComDec method can accurately estimate target positions, whereas other DFL methods have 1 or 2 incorrect estimated positions. This is because the proposed ComDec method can eliminate the position ambiguities introduced by environmental dynamics by taking advantage of the channel diversity.

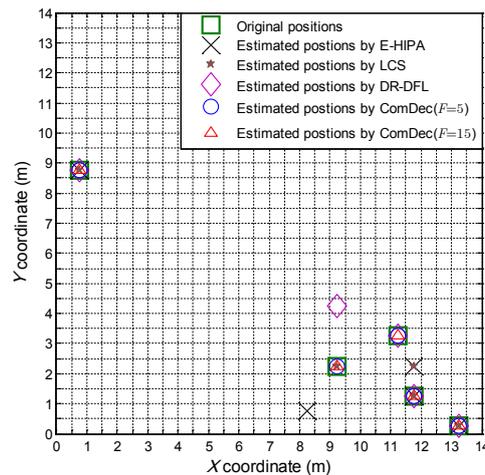


Figure 11. Comparison of the original target positions and the estimated target positions of different DFL approaches.

5.4. Localization Error vs. SNR

In this subsection, the localization performances of different DFL methods under different SNRs are evaluated. In cellular communication (e.g., LTE), the typical SNR value is below 0 dB. However, in short range communication protocols such as 802.11 b/g, the typical SNR value is from 10–15 dB (low signal) to 40 dB (excellent signal). In fact, the scenario that is chosen for our simulations is the short-range communication scenario. Thus, the SNR in this simulation is varied from 0 dB to 40 dB. Figure 12 shows the *Avg.Error* of the ComDec method ($F = 10, 15,$ and 20) and other DFL methods. As can be seen, the *Avg.Error* of all DFL methods decrease with the increase of SNR. Generally, the localization error is mainly caused by the measurement noise and the environmental dynamics in changing environments. As the ComDec and DR-DFL have countermeasures to mitigate the dictionary mismatches introduced by environmental dynamics, their *Avg.Error* decrease dramatically when SNR increases from 0 dB to 25 dB.

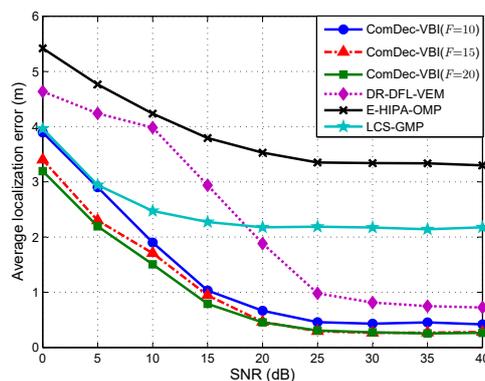


Figure 12. Average localization error with different SNR.

Furthermore, as can be seen from Figure 12, by increasing the channel number, the accuracy improvement of ComDec in low SNR case ($\text{SNR} < 20$ dB) is much larger than the improvement in high SNR case ($\text{SNR} > 20$ dB). This demonstrates that the channel diversity can help to reduce the location uncertainty that introduced by measurement noises especially in the low SNR case. Moreover, the DR-DFL performs better than LCS in the high SNR case, while performs worse in low SNR cases. This is because, in the low SNR case, the measurement noise can significantly affect the dictionary refinement process. It may lead to a wrong estimation of the dictionary parameters, and the dictionary mismatches will deteriorate the localization performance seriously. Fortunately, in the high SNR case, the DR-DFL can correctly estimate the dictionary parameters and performs better than LCS.

5.5. Localization Error vs. Number of Targets

In this subsection, we examine the performances of multiple CS-based DFL methods when K varies from 2 to 20 at a step of 2. In this simulation, we set the environmental change time $t = 1$ to simulate the environmental dynamics in changing environments. With the increase of K , the joint sparsity level of the jointly sparse vectors \mathbb{S} changes accordingly. This will make the localization accuracies of all CS-based DFL decrease. Figure 13 plots the *Avg.Error* as a function of K . As we can see, the *Avg.Error* of all DFL methods increase as K increases, and the ComDec is more accurate than other DFL methods. Furthermore, we can also observe that the DR-DFL and ComDec perform better than the E-HIPA and LCS. This verifies that the Bayesian recovery algorithm can provide a more accurate sparse reconstruction than the greedy algorithm for CS recovery.

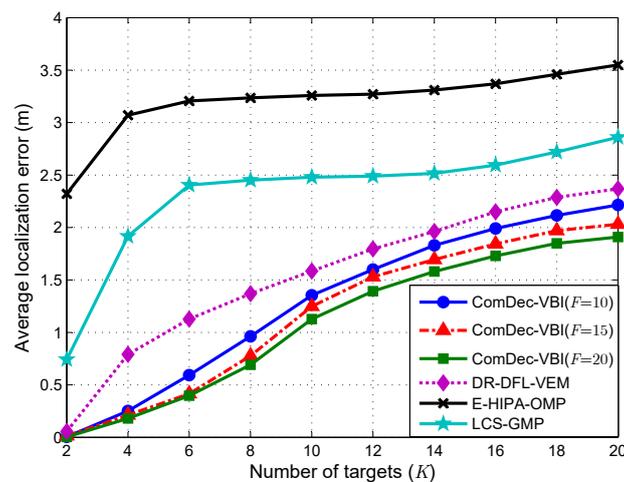


Figure 13. Average localization error with different K .

6. Conclusions and Future Work

In this paper, a novel multi-channel CS-based DFL method, ComDec is proposed. It can solve the dictionary mismatch problem caused by environmental dynamics in changing environments. The key novelty of ComDec is the making use of the channel diversity of CSI measurements in CS-based DFL. Moreover, the dictionary parameters of multiple channels are assumed to be adjustable. They can adapt to the changes of environment and different channel characteristics. In this manner, we can avoid the site survey process that is typically adopted in fingerprinting-based DFL and improve the robustness of DFL against environmental dynamics. We formulate the CS-based DFL problem as a joint sparse recovery problem, and develop a novel joint sparse recovery algorithm under the variational Bayesian inference framework. Furthermore, four methods are presented to reduce its computational complexity. Finally, numerical simulation results validate the effectiveness of the ComDec method.

As future work, we will further investigate the dictionary mismatch problem in CS-based multi-target DFL caused by more realistic and complex environmental dynamics. Additionally, we intend to design and implement an efficient and accurate DFL framework that can utilize the phase information of CSI measurements.

Author Contributions: All authors contributed extensively to the work in this paper. D.Y., Y.G., and N.L. conceived the proposed scheme, derived inference procedure and conducted numerical simulation. D.Y. analyzed the data and wrote the manuscript. Y.G., N.L. and X.Y. provided valuable comments and contributed to the revision of the manuscript.

Funding: This work was supported in part by the Natural Science Foundation of Jiangsu Province under grant BK20171401; the National Natural Science Foundation of China under grant 61871400 and 61571463.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DFL	Device-free localization
VBI	Variational Bayesian inference
LBS	Location-based service
OFDM	Orthogonal Frequency Division Multiplexing
LTE	Long Term Evolution
SNR	Signal-to-noise ratio
CS	Compressive sensing
RIP	Restricted isometry property
RTI	Radio tomographic imaging
GMP	Greedy matching pursuit
LCS	Device-free localization based on compressive sensing
PDF	Probability density function
CG	Conjugate gradient
OMP	Orthogonal matching pursuit
E-HIPA	The energy-efficient framework for high-precision multi-target-adaptive device-free localization
VEM	Variational expectation-maximization
DR-DFL	Dictionary refinement based device-free localization
ComDec	Compressive sensing-based multi-target device-free localization
M-OMP	Multiple measurement vector orthogonal matching pursuit
MMV	Multiple measurement vector
COTS	Commercial off-the-shelf
RSS	Received signal strength
WSN	Wireless sensor networks
CSI	Channel state information
LASSO	Least absolute shrinkage and selection operator
LOS	Line-of-sight
M-SBL	Multiple sparse Bayesian learning
KLD	Kullback-Leibler divergence
M-BMP	Multiple measurement vector basic matching pursuit

References

1. Youssef, M.; Mah, M.; Agrawala, A. Challenges: Device-Free Passive Localization for Wireless Environments. In Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking (MobiCom'07), Montreal, QC, Canada, 9–14 September 2007; pp. 222–229. [\[CrossRef\]](#)
2. Zhang, D.; Ma, J.; Chen, Q.; Ni, L. An RF-Based System for Tracking Transceiver-Free Objects. In Proceedings of the Fifth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom'07), White Plains, NY, USA, 19–23 March 2007; pp. 135–144. [\[CrossRef\]](#)
3. Lei, Q.; Zhang, H.; Sun, H.; Tang, L. Fingerprint-Based Device-Free Localization in Changing Environments using Enhanced Channel Selection and Logistic Regression. *IEEE Access* **2018**, *6*, 2569–2577. [\[CrossRef\]](#)
4. Sun, B.; Guo, Y.; Li, N.; Fang, D. Multiple Target Counting and Localization Using Variational Bayesian EM Algorithm in Wireless Sensor Networks. *IEEE Trans. Commun.* **2017**, *65*, 2985–2998. [\[CrossRef\]](#)
5. Qian, P.; Guo, Y.; Li, N. Multitarget Localization With Inaccurate Sensor Locations via Variational EM Algorithm. *IEEE Sens. Lett.* **2019**, *3*, 1–4. [\[CrossRef\]](#)
6. Guo, Y.; Sun, B.; Li, N.; Fang, D. Variational Bayesian Inference-based Counting and Localization for Off-Grid Targets with Faulty Prior Information in Wireless Sensor Networks. *IEEE Trans. Commun.* **2018**, *66*, 1273–1283. [\[CrossRef\]](#)
7. Lin, A.; Ling, H. Doppler and Direction-Of-Arrival (DDOA) Radar for Multiple-Mover Sensing. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 1496–1509. [\[CrossRef\]](#)

8. Krumm, J.; Harris, S.; Meyers, B.; Brumitt, B.; Hale, M.; Shafer, S. Multi-Camera Multi-Person Tracking for EasyLiving. In Proceedings of the IEEE 3rd International Workshop on Visual Surveillance, Dublin, Ireland, 1 July 2000; pp. 3–10. [\[CrossRef\]](#)
9. Want, R.; Hopper, A.; Falco, V.; Gibbons, J. The Active Badge Location System. *ACM Trans. Inf. Syst.* **1992**, *10*, 91–102. [\[CrossRef\]](#)
10. Botta, M.; Simek, M. Adaptive Distance Estimation Based on RSSI in 802.15.4 Network. *Radioengineering* **2013**, *22*, 1162–1168.
11. Neburka, J.; Tlamse, Z.; Benes, V.; Polak, L.; Kaller, O.; Bolecek, L.; Sebesta, J.; Kratochvil, T. Study of the Performance of RSSI Based Bluetooth Smart Indoor Positioning. In Proceedings of the IEEE 26th International Conference Radioelektronika (RADIOELEKTRONIKA), Kosice, Slovakia, 19–20 April 2016; pp. 121–125. [\[CrossRef\]](#)
12. Zhang, D.; Liu, Y.; Guo, X.; Ni, L.M. RASS: A Real-Time, Accurate, and Scalable System for Tracking Transceiver-Free Objects. *IEEE Trans. Parallel Distrib. Syst.* **2013**, *24*, 996–1008. [\[CrossRef\]](#)
13. Mager, B.; Lundrigan, P.; Patwari, N. Fingerprint-Based Device-Free Localization Performance in Changing Environments. *IEEE J. Sel. Areas Commun.* **2015**, *33*, 2429–2438. [\[CrossRef\]](#)
14. Yu, D.; Guo, Y.; Li, N.; Fang, D. Dictionary Refinement for Compressive Sensing Based Device-Free Localization via the Variational EM Algorithm. *IEEE Access* **2016**, *4*, 9743–9757. [\[CrossRef\]](#)
15. Yang, Z.; Zhou, Z.; Liu, Y. From RSSI to CSI: Indoor Localization via Channel Response. *ACM Comput. Surv.* **2013**, *46*, 1–32. [\[CrossRef\]](#)
16. Wang, X.; Gao, L.; Mao, S.; Pandey, S. CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach. *IEEE Trans. Veh. Technol.* **2017**, *66*, 763–776. [\[CrossRef\]](#)
17. Guo, Y.; Yu, D.; Li, N. Exploiting Fine-Grained Subcarrier Information for Device-Free Localization in Wireless Sensor Networks. *Sensors* **2018**, *18*, 3110. [\[CrossRef\]](#)
18. Candes, E.; Wakin, M. An Introduction to Compressive Sampling. *IEEE Signal Process. Mag.* **2008**, *25*, 21–30. [\[CrossRef\]](#)
19. Wang, J.; Gao, Q.; Pan, M.; Zhang, X.; Yu, Y.; Wang, H. Towards Accurate Device-Free Wireless Localization with a Saddle Surface Model. *IEEE Trans. Veh. Technol.* **2016**, *65*, 6665–6677. [\[CrossRef\]](#)
20. Tzikas, D.G.; Likas, A.C.; Galatsanos, N.P. The Variational Approximation for Bayesian Inference. *IEEE Signal Process. Mag.* **2008**, *25*, 131–146. [\[CrossRef\]](#)
21. Wang, J.; Zhang, X.; Gao, Q.; Ma, X.; Feng, X.; Wang, H. Device-Free Simultaneous Wireless Localization and Activity Recognition with Wavelet Feature. *IEEE Trans. Veh. Technol.* **2017**, *66*, 1659–1669. [\[CrossRef\]](#)
22. Wang, J.; Fang, D.; Chen, X.; Yang, Z.; Xing, T.; Cai, L. LCS: Compressive Sensing Based Device-Free Localization for Multiple Targets in Sensor Networks. In Proceedings of the IEEE INFOCOM 2013, Turin, Italy, 14–19 April 2013; pp. 145–149. [\[CrossRef\]](#)
23. Wang, J.; Fang, D.; Yang, Z.; Jiang, H.; Chen, X.; Xing, T.; Cai, L. E-HIPA: An Energy-Efficient Framework for High-Precision Multi-Target-Adaptive Device-Free Localization. *IEEE Trans. Mob. Comput.* **2017**, *16*, 716–729. [\[CrossRef\]](#)
24. Xiao, J.; Wu, K.; Yi, Y.; Wang, L.; Ni, L. Pilot: Passive Device-Free Indoor Localization Using Channel State Information. In Proceedings of the IEEE 33rd International Conference on Distributed Computing Systems, Philadelphia, PA, USA, 8–11 July 2013; pp. 236–245. [\[CrossRef\]](#)
25. Abdel-Nasser, H.; Samir, R.; Sabek, I.; Youssef, M. Monophy: Mono-Stream-Based Device-Free WLAN Localization via Physical Layer Information. In Proceedings of the IEEE Wireless Communication Network Conference (WCNC), Shanghai, China, 7–10 July 2013; pp. 4546–4551. [\[CrossRef\]](#)
26. Gao, Q.; Wang, J.; Ma, X.; Feng, X.; Wang, H. CSI-Based Device-Free Wireless Localization and Activity Recognition Using Radio Image Features. *IEEE Trans. Veh. Technol.* **2017**, *66*, 10346–10356. [\[CrossRef\]](#)
27. Wang, J.; Xiong, J.; Jiang, H.; Jamieson, K.; Chen, X.; Fang, D.; Wang, C. Low Human-Effort, Device-Free Localization with Fine-Grained Subcarrier Information. *IEEE Trans. Mob. Comput.* **2018**, *17*, 2550–2563. [\[CrossRef\]](#)
28. Wilson, J.; Patwari, N. Radio Tomographic Imaging with Wireless Networks. *IEEE Trans. Mob. Comput.* **2010**, *9*, 621–632. [\[CrossRef\]](#)

29. Wipf, D.; Rao, B.; Nagarajan, S. Latent Variable Bayesian Models for Promoting Sparsity. *IEEE Trans. Inf. Theory* **2011**, *57*, 6236–6255. [[CrossRef](#)]
30. Wipf, D.; Rao, B. Sparse Bayesian Learning for Basis Selection. *IEEE Trans. Signal Process.* **2004**, *52*, 2153–2164. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).