Joint Sparse Bayesian Learning for Channel Estimation in ISAC

Kangjian Chen, *Graduate Student Member, IEEE*, and Chenhao Qi, *Senior Member, IEEE*

***Abstract*— In this letter, we investigate channel estimation in integrated sensing and communication and propose a joint sparse Bayesian learning (JSBL) algorithm for the estimation of sensing and communication (S&C) channels. To capture the joint sparsity between S&C channels arising from the commonalities between the sensing targets and communication scatterers, we conceive an adaptive pattern-coupled prior for S&C channel coefficients, which allows us to formulate the joint S&C channel estimation as a maximum a posteriori (MAP) problem. Then, a tailored expectation-maximization (EM) method including an E-step and an M-step is developed to solve the MAP problem. Specifically, in the E-step, we derive the sparse Bayesian estimates for S&C channels, while in the M-step, we dynamically update the off-grid channel angles and prior parameters. Simulation results demonstrate that the proposed JSBL algorithm can achieve improved S&C channel estimation performance over existing methods due to the exploitation of the joint sparsity between S&C channels.**

***Index Terms*— Channel estimation, integrated sensing and communication (ISAC), joint sparsity, sparse Bayesian learning.**

I. Introduction

INTEGRATED sensing and communication (ISAC) is widely considered as a significant component for the future sixth-generation (6G) wireless systems [1], [2], [3]. For example, in the 6G visions released by the International Telecommunication Union in 2023, the ISAC is identified as one of the six major usage scenarios. By integrating sensing into communications, ISAC can potentially enhance the environmental awareness of the network, leading to improved system performance and enhanced ability to support various applications.

Integrated into one system, the sensing and communication (S&C) units can share various hardware and software resources, such as antenna arrays, radio frequency chains, and signal processing capabilities. As a result, ISAC can achieve desired S&C performance with fewer resources, leading to the integration gain [4]. Besides the integration gain, recent works have delved into exploring the coordination gain in ISAC, which aims at exploiting the mutual assistance between S&C to improve the performance of both functions [1]. One promising approach to achieving the coordination gain is exploiting the commonalities between S&C channels [5]. Specifically, certain sensing targets, such as environmental objects, may also serve as the scatterers of communication channels. These commonalities motivate us to perform joint S&C channel estimation (JSCCE). By utilizing independent observations of S&C signals from the same path, the JSCCE has the potential to improve the channel estimation performance of both S&C functions.

Channel estimation typically uses signal processing algorithms for channel parameter estimation. Among existing parameter estimation methods, Bayesian inference, which aims at maximizing the posterior probability of channel parameters, has the potential to achieve the optimal estimation [6]. Given sparse properties of S&C channels, sparse Bayesian learning (SBL), which incorporates a sparse prior into the Bayesian inference framework, has gained widespread adoption [7]. However, the sparse prior in conventional SBL usually assumes independent sparsity for S&C channels, and thus does not adapt to the JSCCE. In [8], a joint sparse prior is proposed to exploit joint sparsity between uplink and downlink channels and an off-grid SBL (OGSBL) is developed for channel estimation. However, the joint sparse prior in [8] is designed for scenarios where the support sets (SSs) of the channels are fully identical, and is not suitable for JSCCE where S&C channels share only partially identical SSs.

In this letter, we propose a joint sparse Bayesian learning (JSBL) algorithm for JSCCE. To capture the joint sparsity between S&C channels, we conceive an adaptive pattern-coupled prior for S&C channel coefficients, which allows us to formulate the JSCCE as a maximum a posteriori (MAP) problem. Then, a tailored expectation-maximization (EM) method including an E-step and an M-step is developed to solve the MAP problem. Specifically, in the E-step, we derive the sparse Bayesian estimates for S&C channels, while in the M-step, we dynamically update the off-grid channel angles and prior parameters.

Notations: Lowercase and uppercase bold symbols denote vectors and matrices, respectively. $[\boldsymbol{A}]_{:,m}$ denotes the mth column of a matrix $\boldsymbol{A}$. $[\boldsymbol{a}]_m$ denotes the mth entry of a vector $\boldsymbol{a}$. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, $\mathrm{diag}(\boldsymbol{a})$, and $\mathrm{tr}(\cdot)$ denote transpose, Hermitian transpose, a matrix with vector $\boldsymbol{a}$ as its diagonal entries, and the trace operator, respectively. $\propto$, $\mathbb{R}$, $\mathbb{C}$, and $\mathcal{CN}(\cdot)$ denote proportionality, real numbers, complex numbers, and complex Gaussian distribution, respectively.

II. System Model

We consider an ISAC system including one ISAC base station (BS) and K single-antenna users. The ISAC BS employs a transmitter with N_{t} antennas and a collocated receiver with N_{r} antennas, where the antennas are arranged in uniform linear arrays with half-wavelength intervals. In the considered ISAC system, the transmitter emits sensing signals for environment sensing and the users transmit uplink pilots for channel estimation. Meanwhile, the receiver simultaneously collects the echo signals and uplink pilots. We assume orthogonal frequency

Manuscript received 14 June 2024; accepted 27 June 2024. Date of publication 3 July 2024; date of current version 14 August 2024. This work was supported in part by the National Natural Science Foundation of China under Grants U22B2007 and 62071116, and in part by the National Key Research and Development Program of China under Grant 2021YFB2900404. The associate editor coordinating the review of this letter and approving it for publication was M. Morales-Céspedes. *(Corresponding author: Chenhao Qi.)*

The authors are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: kjchen@seu.edu.cn; qch@seu.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2024.3422421

1558-2558 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Southeast University. Downloaded on August 16,2024 at 04:35:08 UTC from IEEE Xplore. Restrictions apply.

resources are allocated to the transmitter and K users so that their signals can be separated at the receiver. As a result, the receiver can collect $K+1$ independent signals. We specify that the first received signal corresponds to the echo signal and the kth, for $k \geq 2$, received signal corresponds to the uplink pilots of the $(k-1)$th user. The collected echo signals by the BS can be expressed as

$$\boldsymbol{y}_1 = \boldsymbol{H}_1 \boldsymbol{s} + \boldsymbol{n}_1, \tag{1}$$

where $\boldsymbol{s} \in \mathbb{C}^{N_\mathrm{t} \times 1}$ and $\boldsymbol{n}_1 \in \mathbb{C}^{N_\mathrm{r} \times 1}$ denote the transmit sensing signals and additive Gaussian noise, respectively. $\boldsymbol{H}_1 \in \mathbb{C}^{N_\mathrm{r} \times N_\mathrm{t}}$ denotes the sensing channels and can be modeled as

$$\boldsymbol{H}_1 = \sum_{l=1}^{L_1} \xi_1^{(l)} \boldsymbol{\alpha}\big(N_\mathrm{r}, \theta_1^{(l)}\big) \boldsymbol{\alpha}\big(N_\mathrm{t}, \theta_1^{(l)}\big)^\mathrm{H}, \tag{2}$$

where L_1, $\xi_1^{(l)}$, and $\theta_1^{(l)}$ denote the number of targets, the channel gain of the lth target, and the channel angle of the lth target, respectively. $\boldsymbol{\alpha}(\cdot)$ denotes the channel steering vector and is defined as

$$\boldsymbol{\alpha}(N, \theta) = \big[1, e^{j\pi\theta}, \ldots, e^{j\pi(N-1)\theta}\big]^\mathrm{T}. \tag{3}$$

Similarly, the collected uplink pilot of the kth user, for $k = 1, 2, \ldots, K$, can be expressed as

$$\boldsymbol{y}_{k+1} = \boldsymbol{h}_{k+1} x_k + \boldsymbol{n}_{k+1}, \tag{4}$$

where x_k and $\boldsymbol{n}_{k+1}$ denote the transmit pilot and additive Gaussian noise, respectively. $\boldsymbol{h}_{k+1} \in \mathbb{C}^{N_\mathrm{r} \times 1}$ denotes the communication channel and can be modeled as

$$\boldsymbol{h}_{k+1} = \sum_{l=1}^{L_{k+1}} \xi_{k+1}^{(l)} \boldsymbol{\alpha}(N_\mathrm{r}, \theta_{k+1}^{(l)}), \tag{5}$$

where L_{k+1}, $\xi_{k+1}^{(l)}$, and $\theta_{k+1}^{(l)}$ denote the number of channel paths, the channel gain of the lth path, and the channel angle of the lth path, respectively. We stack the collected signals from the transmitter and K users together as

$$\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{K+1}]. \tag{6}$$

Then, we focus on recovering S&C channels from $\boldsymbol{Y}$ by proposing a JSBL algorithm.

III. Joint Sparse Bayesian Learning

A. Off-Grid Sparse Bayesian Learning Framework

Define $\overline{\xi}_1^{(l)} \triangleq \xi_1^{(l)} \boldsymbol{\alpha}(N_\mathrm{t}, \theta_1^{(l)})^\mathrm{H} \boldsymbol{s}$ and $\overline{\xi}_{k+1}^{(l)} \triangleq \xi_{k+1}^{(l)} x_k$ for $k = 1, 2, \ldots, K$. Then, both (1) and (4) can be rewritten as

$$\boldsymbol{y}_k = \sum_{l=1}^{L_k} \overline{\xi}_k^{(l)} \boldsymbol{\alpha}\big(N_\mathrm{r}, \theta_k^{(l)}\big) + \boldsymbol{n}_k, \tag{7}$$

for $k = 1, 2, \cdots, K+1$. From (7), both echo signals and uplink pilots can be expressed as the combination of channel steering vectors. We quantize the channel angle into M samples with the mth sample expressed as $\varphi_m = -1 + (2m-1)/M$. Suppose the quantization is fine enough and the real channel angles are located in the quantized samples. Then, we can stack (7) in a more compact form as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{W} + \boldsymbol{N}, \tag{8}$$

where $[\boldsymbol{A}]_{:,m} \triangleq \boldsymbol{\alpha}(N_\mathrm{r}, \varphi_m)$, $\boldsymbol{N} \triangleq [\boldsymbol{n}_1, \boldsymbol{n}_2, \ldots, \boldsymbol{n}_{K+1}]$, and $\boldsymbol{W}$ denotes the corresponding channel coefficients. In practice, the real channel angles usually deviate from the quantized ones. To this end, a set of variables $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_M]^\mathrm{T}$ are introduced to deal with the deviation. Then, (8) can be rewritten as

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{\beta})\boldsymbol{W} + \boldsymbol{N}, \tag{9}$$

where $[\boldsymbol{A}(\boldsymbol{\beta})]_{:,m} \triangleq \boldsymbol{\alpha}(N_\mathrm{r}, \varphi_m + \beta_m)$.

Usually, both S&C channels have limited paths. As a result, most entries in $\boldsymbol{W}$ would be zero, i.e., $\boldsymbol{W}$ is sparse. To characterize the sparsity of $\boldsymbol{W}$, the SBL framework provides a two-stage hierarchical prior as [7]

$$p(\boldsymbol{W}|\boldsymbol{G}) = \prod_{k=1}^{K+1} \mathcal{CN}(\boldsymbol{w}_k|\boldsymbol{0}, \mathrm{diag}(\boldsymbol{g}_k)^{-1}), \tag{10a}$$

$$p(\boldsymbol{g}_k) = \prod_{m=1}^{M} \Gamma(g_{m,k}; a+1, b). \tag{10b}$$

In (10a), $\boldsymbol{G} \in \mathbb{R}^{M \times (K+1)}$ denotes the stack of variances for entries in $\boldsymbol{W}$, $\boldsymbol{w}_k \triangleq [\boldsymbol{W}]_{:,k}$, and $\boldsymbol{g}_k \triangleq [\boldsymbol{G}]_{:,k}$. In (10b), $g_{m,k} \triangleq [\boldsymbol{g}_k]_m$ and $\Gamma(g_{m,k}; a+1, b)$ denotes the Gamma distribution.

Under the assumption of independently identical Gaussian noise, we have

$$p(\boldsymbol{N}|\zeta) = \prod_{k=1}^{K+1} \mathcal{CN}(\boldsymbol{n}_k|0, \zeta^{-1}\boldsymbol{I}), \tag{11}$$

where ζ^{-1} denotes the noise variance. Specifically, ζ is modeled as a Gamma distribution to provide a broad hyperprior, i.e., $p(\zeta) = \Gamma(\zeta; a+1, b)$.

B. Joint Sparsity Exploration

The communication channels are usually composed of the line-of-sight (LoS) path and non-LoS (NLoS) paths. The NLoS paths, which involve reflections from scatterers, contribute to a substantial portion of the communication channels. Therefore, accurately estimating the channel angles associated with these scatterers is crucial for reconstructing the communication channels. On the other hand, these scatterers could also be detected as sensing targets. Therefore, the collected echo signals may also contain information about the channel angles of the scatterers. This correlation indicates that the SSs of S&C channels have some overlaps because of the fact that the communication scatterers could also be the sensing targets. Besides, the sensing channels may have their individual SSs because part of targets may not serve as scatterers for any user channels. Moreover, the communication channels may also have their individual SSs because the communication channels may have LoS paths that do not pass through any scatterers. Based on the above discussions, we propose the following criterion for the joint sparsity.

Joint Sparsity Criterion*: If an entry in the sensing SS is zero or non-zero, the corresponding entry in the communication SS is more likely to be zero or non-zero than the case that the entry in the sensing SS is non-zero or zero, respectively.*

Remark 1: We provide an example to verify the ***Joint Sparsity Criterion***. Consider an ISAC system with $L_1 = 6$ targets, $L_k = 4$ communication channel paths, and $M = 32$ angle samples. Suppose $\overline{L}_k = 2$ targets serve as the communication scatterers. We define the events $\mathcal{S}_0$ and $\mathcal{S}_1$ as an entry in the sensing support set being zero and non-zero,

Authorized licensed use limited to: Southeast University. Downloaded on August 16,2024 at 04:35:08 UTC from IEEE Xplore. Restrictions apply.

respectively. Similarly, the events $\mathcal{C}_0$ and $\mathcal{C}_1$ represent an entry in the communication support set being zero and non-zero, respectively. Then, we have $p(\mathcal{C}_1|\mathcal{S}_0) = 1/13$, $p(\mathcal{C}_0|\mathcal{S}_0) = 12/13$, $p(\mathcal{C}_0|\mathcal{S}_1) = 2/3$, and $p(\mathcal{C}_1|\mathcal{S}_1) = 1/3$. We can find that $p(\mathcal{C}_0|\mathcal{S}_0) > p(\mathcal{C}_0|\mathcal{S}_1)$ and $p(\mathcal{C}_1|\mathcal{S}_1) > p(\mathcal{C}_1|\mathcal{S}_0)$, thus verifying the ***Joint Sparsity Criterion***.

Note that the two-stage hierarchical prior in (10) only considers the independent sparsity and neglects the joint sparsity between S&C channels. To exploit the joint sparsity, we replace the prior in (10a) with an adaptive pattern-coupled prior, expressed as

$$\begin{aligned} p(\boldsymbol{w}_1|\boldsymbol{g}_1) &= \mathcal{CN}(\boldsymbol{w}_1|\boldsymbol{0}, \mathrm{diag}(\boldsymbol{g}_1)^{-1}), \\ & p(w_{m,k}|g_{m,1}, g_{m,k}, \rho_{m,k}) \\ &\propto \mathcal{CN}(w_{m,k}|0, \rho_{m,k}^{-1} g_{m,1}^{-1}) \cdot \mathcal{CN}(w_{m,k}|0, g_{m,k}^{-1}), \\ & k = 2, 3, \cdots, K+1. \end{aligned} \tag{12}$$

To simplify the expression, we define a matrix $\boldsymbol{P} \in \mathbb{R}^{M\times(K+1)}$ to keep $\rho_{m,k}$, i.e., $[\boldsymbol{P}]_{m,k} \triangleq \rho_{m,k}$, where we specify $\rho_{m,1} = 0$ for consistency. In (12), the prior of $\boldsymbol{w}_1$ is the same as that in (10a) and only relies on $\boldsymbol{g}_1$, which implies that we provide an independent prior for the coefficients of sensing channels. However, the prior of $\boldsymbol{w}_k$, for $k \geq 2$, is coupled with the prior of $\boldsymbol{w}_1$ with $\rho_{m,k}$ adaptively adjusting the coupling effects, which implies that we provide a joint prior for the coefficients of communication channels. According to the Gaussian reproduction lemma [9], we have

$$\begin{aligned} & p(w_{m,k}|g_{m,1}, g_{m,k}, \rho_{m,k}) \\ & \propto \mathcal{CN}(w_{m,k}|0, (\rho_{m,k} g_{m,1} + g_{m,k})^{-1}), \end{aligned} \tag{13}$$

for $k = 2, 3, \cdots, K+1$. From (13), if $w_{m,1}$ equals zero, its variance $g_{m,1}^{-1}$ would be a small value that approaches zero. Then, $g_{m,1}$ would be a large value that approaches infinity. As a result, the variance of $w_{m,k}$ would also approach zero due to the small variance, which provides a more sparse prior for $w_{m,k}$ and means that $w_{m,k}$ would be more likely to be zero than independent prior in (10a). On the contrary, if $w_{m,1}$ is non-zero, its variance $g_{m,1}^{-1}$ would be a large value and $g_{m,1}$ would be a small value. As a result, the impact of $w_{m,1}$ on $w_{m,k}$ would be negligible, which indicates that $w_{m,k}$ would be more likely to be non-zero than the case that $w_{m,1}$ equals zero. Based on the above discussions, the prior of $\boldsymbol{W}$ in (12) aligns with ***Joint Sparsity Criterion***, and has the potential to exploit the joint sparsity between S&C channels.

Remark 2: In scenarios where the ***Joint Sparsity Criterion*** is not satisfied and the S&C channels have no correlations, the conventional SBL algorithm is sufficient to effectively solve the channel estimation problem, without the need to employ the proposed JSBL algorithm.

C. Detailed Implementation of the JSBL Algorithm

Following the procedures of conventional SBL [7], [8], the JSBL algorithm does not explicitly estimate $\boldsymbol{W}$, but treats $\boldsymbol{W}$ as hidden variables and infers ζ, $\boldsymbol{G}$, $\boldsymbol{P}$ and $\boldsymbol{\beta}$ by maximizing the posterior $p(\zeta, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\beta}|\boldsymbol{Y})$, i.e.,

$$(\zeta^\star, \boldsymbol{G}^\star, \boldsymbol{P}^\star, \boldsymbol{\beta}^\star) = \arg \max_{\zeta, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\beta}} p(\zeta, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\beta}|\boldsymbol{Y}). \tag{14}$$

According to the Bayes' Theorem, (14) is reformulated as

$$(\zeta^\star, \boldsymbol{G}^\star, \boldsymbol{P}^\star, \boldsymbol{\beta}^\star) = \arg \max_{\zeta, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\beta}} \ln p(\boldsymbol{Y}, \zeta, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\beta}). \tag{15}$$

In fact, directly solving (15) is difficult due to its integral form with respect to the hidden variables $\boldsymbol{W}$ [7], [8]. Alternatively, the EM algorithm, which iteratively calculates the posterior distribution of $\boldsymbol{W}$ in the E-step and updates the parameters in the M-step, is adopted to maximize the joint distribution with hidden variables and find solutions for (15) [7], [8].[1]

The EM algorithm begins by initializing ζ, $\boldsymbol{G}$, $\boldsymbol{P}$, and $\boldsymbol{\beta}$ as

$$\begin{aligned} & \widetilde{\zeta}^{(0)} = 1, \ \left[\widetilde{\boldsymbol{G}}^{(0)}\right]_{m,k} = \left|\left[\boldsymbol{A}(\widetilde{\boldsymbol{\beta}}^{(0)})^{\mathrm{H}}\boldsymbol{Y}\right]_{m,k}\right|^{-2}, \\ & [\widetilde{\boldsymbol{P}}^{(0)}]_{m,k} = 1, \ \text{for } k \geq 2, \ \text{and } \left[\widetilde{\boldsymbol{\beta}}^{(0)}\right]_m = 0, \end{aligned} \tag{16}$$

respectively. Then, we iteratively perform the E-step and M-step of the EM algorithm until the maximum number of iterations, Q, is reached.[2] In the qth iteration, for $q \geq 1$, we denote the updated ζ, $\boldsymbol{G}$, $\boldsymbol{P}$ and $\boldsymbol{\beta}$ in the last iteration as $\overline{\zeta} \triangleq \widetilde{\zeta}^{(q-1)}$, $\overline{\boldsymbol{G}} \triangleq \widetilde{\boldsymbol{G}}^{(q-1)}$, $\overline{\boldsymbol{P}} \triangleq \widetilde{\boldsymbol{P}}^{(q-1)}$, and $\overline{\boldsymbol{\beta}} \triangleq \widetilde{\boldsymbol{\beta}}^{(q-1)}$, respectively. Then, we elaborate on the details of the E-step and M-step as follows.

In the E-step of the qth iteration, we compute the posterior distribution of $\boldsymbol{W}$ as

$$\begin{aligned} p(\boldsymbol{W}|\boldsymbol{Y}, \overline{\zeta}, \overline{\boldsymbol{G}}, \overline{\boldsymbol{P}}, \overline{\boldsymbol{\beta}}) &\propto \prod_{k=1}^{K+1} p(\boldsymbol{w}_k|\boldsymbol{y}_k, \overline{\zeta}, \overline{\boldsymbol{g}}_k, \overline{\boldsymbol{p}}_k, \overline{\boldsymbol{\beta}}) \\ &\propto \prod_{k=1}^{K+1} \mathcal{CN}(\boldsymbol{w}_k|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \end{aligned} \tag{17}$$

where

$$\begin{aligned} & \boldsymbol{\mu}_k = \overline{\zeta}\boldsymbol{\Sigma}_k \boldsymbol{A}(\overline{\boldsymbol{\beta}})^{\mathrm{H}}\boldsymbol{y}_k, \ \overline{\boldsymbol{g}}_k \triangleq [\overline{\boldsymbol{G}}]_{:,k}, \ \overline{\boldsymbol{p}}_k \triangleq [\overline{\boldsymbol{P}}]_{:,k}, \\ & \boldsymbol{\Sigma}_k = \left(\mathrm{diag}(\overline{\boldsymbol{g}}_k) + \mathrm{diag}(\overline{\boldsymbol{g}}_1)\overline{\boldsymbol{p}}_k + \overline{\zeta}\boldsymbol{A}(\overline{\boldsymbol{\beta}})^{\mathrm{H}}\boldsymbol{A}(\overline{\boldsymbol{\beta}})\right)^{-1}. \end{aligned} \tag{18}$$

Then, we compute the expectation of $\ln p(\boldsymbol{Y}, \boldsymbol{W}, \zeta, \boldsymbol{G}, \boldsymbol{\beta}, \boldsymbol{P})$ over $\boldsymbol{W}$, where

$$\begin{aligned} & p(\boldsymbol{Y}, \boldsymbol{W}, \zeta, \boldsymbol{G}, \boldsymbol{\beta}, \boldsymbol{P}) \\ & \quad \propto p(\boldsymbol{Y}|\boldsymbol{W}, \zeta, \boldsymbol{\beta}) p(\boldsymbol{W}|\boldsymbol{G}, \boldsymbol{P}) p(\boldsymbol{G}) p(\zeta). \end{aligned} \tag{19}$$

In the M-step of the qth iteration, we update ζ, $\boldsymbol{G}$, $\boldsymbol{P}$, and $\boldsymbol{\beta}$ by maximizing $\mathbb{E}\{\ln p(\boldsymbol{Y}, \boldsymbol{W}, \zeta, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\beta})\}$, i.e.,

$$\widetilde{\zeta}^{(q)} = \max_{\zeta} \mathbb{E}\{\ln p(\boldsymbol{Y}, \boldsymbol{W}, \zeta, \overline{\boldsymbol{G}}, \overline{\boldsymbol{P}}, \overline{\boldsymbol{\beta}})\}, \tag{20a}$$

$$\widetilde{\boldsymbol{G}}^{(q)} = \max_{\boldsymbol{G}} \mathbb{E}\{\ln p(\boldsymbol{Y}, \boldsymbol{W}, \widetilde{\zeta}^{(q)}, \boldsymbol{G}, \overline{\boldsymbol{P}}, \overline{\boldsymbol{\beta}})\}, \tag{20b}$$

$$\widetilde{\boldsymbol{P}}^{(q)} = \max_{\boldsymbol{P}} \mathbb{E}\{\ln p(\boldsymbol{Y}, \boldsymbol{W}, \widetilde{\zeta}^{(q)}, \widetilde{\boldsymbol{G}}^{(q)}, \boldsymbol{P}, \overline{\boldsymbol{\beta}})\}, \tag{20c}$$

$$\widetilde{\boldsymbol{\beta}}^{(q)} = \max_{\boldsymbol{\beta}} \mathbb{E}\{\ln p(\boldsymbol{Y}, \boldsymbol{W}, \widetilde{\zeta}^{(q)}, \widetilde{\boldsymbol{G}}^{(q)}, \widetilde{\boldsymbol{P}}^{(q)}, \boldsymbol{\beta})\}. \tag{20d}$$

Subsequently, we provide how we update ζ, $\boldsymbol{G}$, $\boldsymbol{P}$, and $\boldsymbol{\beta}$ based on (20).

1) Updating ζ: From (19) and (20a), we have

$$\begin{aligned} & \ln p(\boldsymbol{Y}, \boldsymbol{W}, \zeta, \overline{\boldsymbol{G}}, \overline{\boldsymbol{P}}, \overline{\boldsymbol{\beta}}) \\ & \propto (M(K+1) + a) \ln \zeta \\ & \quad - \zeta\left(\sum_{k=1}^{K+1} \|\boldsymbol{y}_k - \boldsymbol{A}(\overline{\boldsymbol{\beta}})\boldsymbol{w}_k\|_2^2 + b\right). \end{aligned} \tag{21}$$

[1]The detailed implementation of the JSBL algorithm follows the standard procedures of the SBL developed in [7]. Interested readers can refer to [7] and [8] for further information.

[2]Our simulation results show that the EM algorithm converges within 100 iterations. Therefore, we recommend setting Q as 100 in the implementation.

Authorized licensed use limited to: Southeast University. Downloaded on August 16,2024 at 04:35:08 UTC from IEEE Xplore. Restrictions apply.

Then, we have

$$\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\zeta,\overline{\boldsymbol{G}},\overline{\boldsymbol{P}},\overline{\boldsymbol{\beta}})\}\propto(M(K+1)+a)\ln\zeta-\zeta R, \tag{22}$$

where

$$R\triangleq b+\sum_{k=1}^{K+1}(\|\boldsymbol{y}_k-\boldsymbol{A}(\overline{\boldsymbol{\beta}})\boldsymbol{\mu}_k\|_2^2+\mathrm{Tr}(\boldsymbol{A}(\overline{\boldsymbol{\beta}})^{\mathrm{H}}\boldsymbol{A}(\overline{\boldsymbol{\beta}})\boldsymbol{\Sigma}_k)). \tag{23}$$

Letting $\partial\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\zeta,\overline{\boldsymbol{G}},\overline{\boldsymbol{P}},\overline{\boldsymbol{\beta}})\}/\partial\zeta=0$, we have

$$\widetilde{\zeta}^{(q)}=R/(M(K+1)+a). \tag{24}$$

2) Updating $\boldsymbol{G}$: From (19) and (20b), we have

$$\begin{aligned}&\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\boldsymbol{G},\overline{\boldsymbol{P}},\overline{\boldsymbol{\beta}})\\&\propto\sum_{m=1}^{M}\left(K\ln g_{m,1}-\left(\sum_{k=2}^{K+1}\rho_{m,k}|w_{m,k}|^2+b\right)g_{m,1}\right)\\&\quad+\sum_{m=1}^{M}\sum_{k=1}^{K+1}\left((a+1)\ln g_{m,k}-(b+|w_{m,k}|^2)g_{m,k}\right).\end{aligned} \tag{25}$$

Then, we have

$$\begin{aligned}&\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\boldsymbol{G},\overline{\boldsymbol{P}},\overline{\boldsymbol{\beta}})\}\\&\propto\sum_{m=1}^{M}\left(K\ln g_{m,1}-g_{m,1}\left(\sum_{k=2}^{K+1}\rho_{m,k}\eta_{m,k}+b\right)\right)\\&\quad+\sum_{m=1}^{M}\sum_{k=1}^{K+1}\left((a+1)\ln g_{m,k}-(b+\eta_{m,k})g_{m,k}\right),\end{aligned} \tag{26}$$

where

$$\eta_{m,k}=|[\boldsymbol{\mu}_k]_m|^2+[\boldsymbol{\Sigma}_k]_{m,m}. \tag{27}$$

Letting $\partial\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\boldsymbol{G},\overline{\boldsymbol{P}},\overline{\boldsymbol{\beta}})\}/\partial g_{m,k}=0$, we have

$$\begin{aligned}&g_{m,1}=(K+1+a)\Big/\left(\sum_{k=2}^{K+1}\rho_{m,k}\eta_{m,k}+\eta_{m,1}+b\right),\\&g_{m,k}=(a+1)/\left(\eta_{m,k}+b\right)\ \text{for}\ k\geq 2.\end{aligned} \tag{28}$$

3) Updating $\boldsymbol{P}$: From (19) and (20c), we have

$$\begin{aligned}&\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\widetilde{\boldsymbol{G}}^{(q)},\boldsymbol{P},\overline{\boldsymbol{\beta}})\\&\propto\sum_{m=1}^{M}\sum_{k=2}^{K+1}\ln\rho_{m,k}-\rho_{m,k}g_{m,1}|w_{m,k}|^2.\end{aligned} \tag{29}$$

Then, we have

$$\begin{aligned}&\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\widetilde{\boldsymbol{G}}^{(q)},\boldsymbol{P},\overline{\boldsymbol{\beta}})\}\\&\propto\sum_{m=1}^{M}\sum_{k=2}^{K+1}\ln\rho_{m,k}-\rho_{m,k}g_{m,1}\eta_{m,k}.\end{aligned} \tag{30}$$

Letting $\partial\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\widetilde{\boldsymbol{G}}^{(q)},\boldsymbol{P},\overline{\boldsymbol{\beta}})\}/\partial\rho_{m,k}=0$, we have

$$\rho_{m,k}=1/(g_{m,1}\eta_{m,k}),\text{for}\ k\geq 2. \tag{31}$$

4) Updating $\boldsymbol{\beta}$: From (19) and (20d), we have

$$\begin{aligned}&\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\widetilde{\boldsymbol{G}}^{(q)},\widetilde{\boldsymbol{P}}^{(q)},\boldsymbol{\beta})\\&\propto-\sum_{k=1}^{K+1}\left\|\boldsymbol{y}_k-\boldsymbol{A}(\boldsymbol{\beta})\boldsymbol{w}_k\right\|_2^2.\end{aligned} \tag{32}$$

Then, we have

$$\begin{aligned}&\mathbb{E}\{\ln p(\boldsymbol{Y},\boldsymbol{W},\widetilde{\zeta}^{(q)},\widetilde{\boldsymbol{G}}^{(q)},\widetilde{\boldsymbol{P}}^{(q)},\boldsymbol{\beta}))\}\\&\propto-\sum_{k=1}^{K+1}\left(\left\|\boldsymbol{y}_k-\boldsymbol{A}(\boldsymbol{\beta})\boldsymbol{\mu}_k\right\|_2^2+\mathrm{Tr}(\boldsymbol{A}(\boldsymbol{\beta})^{\mathrm{H}}\boldsymbol{A}(\boldsymbol{\beta})\boldsymbol{\Sigma}_k)\right).\end{aligned} \tag{33}$$

Algorithm 1 Joint Sparse Bayesian Learning Algorithm

Input: N_{r}, K, Q, $\boldsymbol{A}$ and $\boldsymbol{Y}$.
Initialize ζ, $\boldsymbol{G}$, $\boldsymbol{P}$, and $\boldsymbol{\beta}$ via (16). $q\leftarrow 0$.
while $q\leq Q$ **do**
  Update $\widetilde{\zeta}^{(q)}$, $\widetilde{\boldsymbol{G}}^{(q)}$, $\widetilde{\boldsymbol{P}}^{(q)}$, and $\widetilde{\boldsymbol{\beta}}^{(q)}$ via (24), (28), (31) and (33), respectively.
  $q\leftarrow q+1$.
end while
Obtain $\widehat{\boldsymbol{h}}_k$, for $k=1,2,\ldots,K+1$, via (34).
Obtain $\widehat{\theta}_1^{(l)}$, for $l=1,2,\ldots,\widehat{L}_1$ via (35).
Output: $\widehat{\boldsymbol{h}}_k$ and $\widehat{\theta}_1^{(l)}$.

Due to the complicated expressions in (33), it is difficult to obtain a closed-form solution for (20d). Following [8], we find a local solution for $\boldsymbol{\beta}$ via the gradient descent.

D. Channel Reconstruction

After Q iterations, we denote the estimated $\boldsymbol{\beta}$ and $\boldsymbol{\mu}_k$ as $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\mu}}_k$, respectively. Note that $\widehat{\boldsymbol{\mu}}_k$ is the posterior mean of the channel coefficient $\boldsymbol{w}_k$. Therefore, we reconstruct communication channels as

$$\widehat{\boldsymbol{h}}_k=\boldsymbol{A}(\widehat{\boldsymbol{\beta}})\widehat{\boldsymbol{\mu}}_k,\ k=2,3,\cdots,K+1. \tag{34}$$

Suppose $\widehat{L}_1$ targets are detected based on $\widehat{\boldsymbol{\mu}}_1$. We define a set $\boldsymbol{\chi}$ to keep the indices of the $\widehat{L}_1$ largest absolute values in $\widehat{\boldsymbol{\mu}}_1$. Then the estimated target angles can be expressed as

$$\widehat{\theta}_1^{(l)}=-1+(2[\boldsymbol{\chi}]_l-1)/M+[\widehat{\boldsymbol{\beta}}]_{[\boldsymbol{\chi}]_l},\ l=1,\ldots,\widehat{L}_1. \tag{35}$$

Now, we evaluate the computational complexity of the proposed JSBL algorithm. The JSBL algorithm undergoes Q iterations, where each iteration contains the E-step and M-step. The computational complexity of the E-step mainly comes from updating $\boldsymbol{\Sigma}_k$ in (18), which involves the matrix inversion and has a computational complexity of $\mathcal{O}(M^3)$. Since we provide closed-form expressions for updating ζ, $\boldsymbol{G}$ and $\boldsymbol{P}$ in (24), (28), and (31), respectively, the computational complexity of the M-step mainly comes from updating $\boldsymbol{\beta}$ in (33), which employs the gradient descent. Note that the gradient descent is a widely adopted method. We omit the detailed analysis and denote its computational complexity as $\mathcal{O}(D)$. In total, the computational complexity of the proposed JSBL algorithm is $\mathcal{O}(Q((K+1)M^3+D))$.

Finally, we summarize the proposed JSBL algorithm in **Algorithm 1**.

IV. Simulation Results

Now, we evaluate the performance of the proposed JSBL algorithm. We consider an ISAC system that contains one BS and four users. The ISAC BS employs a transmitter and a receiver both with 32 antennas to sense six targets. The communication channels between the users and the ISAC BS are supposed to contain four paths. In each implementation of Monte Carlo simulations, we randomly select part of the targets to serve as the communication scatterers, which indicates that the S&C channels have their individual SSs in addition to the common SSs. The channel angles distribute

Authorized licensed use limited to: Southeast University. Downloaded on August 16,2024 at 04:35:08 UTC from IEEE Xplore. Restrictions apply.

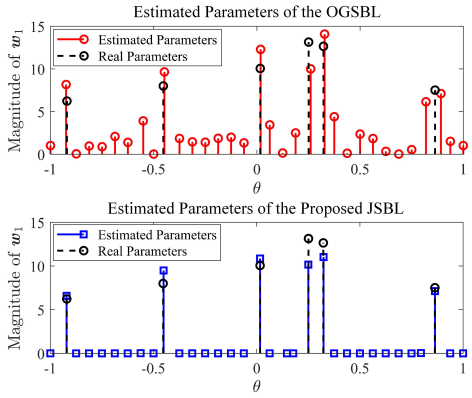

Fig. 1. Comparisons of the parameter estimation performance for the OGSBL and proposed JSBL.

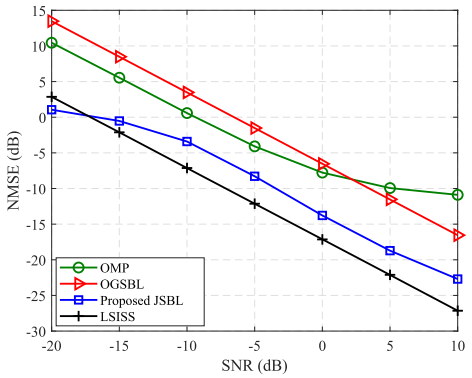

Fig. 2. Comparisons of the communication channel estimation performance for different methods.

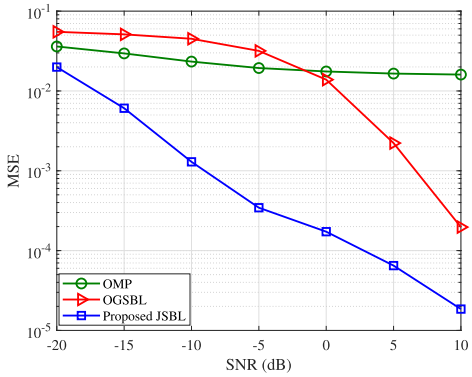

Fig. 3. Comparisons of the target angle estimation performance for different methods.

randomly within $[-1, 1]$ and the channel gains obey the complex Gaussian distribution with a mean of zero and a variance of one. The classic orthogonal matching pursuit (OMP) and the off-grid SBL (OGSBL) [8], neither exploiting the joint sparsity between S&C channels, are adopted as benchmarks for comparisons.

In Fig. 1, we compare the parameter estimation performance of the OGSBL and proposed JSBL, taking the sensing channel as an example. The x-axis and the y-axis denote the estimated channel angles and the corresponding estimated channel coefficients, respectively. We set the signal-to-noise ratio (SNR) as -5 dB. From the figure, the proposed JSBL can achieve a more sparse channel representation and a better parameter estimation performance than the OGSBL owing to the conceived pattern-coupled prior.

In Fig. 2, we compare the communication channel estimation performance for different methods in terms of the normalized mean squared error (NMSE). The least squares based on ideal SS (LSISS), which assumes the channel angles are known and estimates the channel coefficients with the LS, is adopted as a benchmark. From the figure, the JSBL significantly outperforms the OMP and OGSBL and can approach the LSISS for all SNRs, which verifies that exploiting the joint sparsity between S&C channels can effectively improve the channel estimation performance of communications. Notably, the JSBL slightly outperforms the LSISS when the SNR is less than -17.5 dB. This can be attributed to the fact that the JSBL only infers part of channel paths and treats the remaining paths as noise due to the large noise powers, resulting in an NMSE close to 0 dB. However, the large noise power leads to significant channel gain estimation errors for LSISS, resulting in a larger NMSE than JSBL.

In Fig. 3, we compare the target angle estimation performance for different methods in terms of the mean squared error (MSE). From the figure, the proposed JSBL outperforms OMP and OGSBL due to the exploitation of the partially joint sparsity between S&C channels. Specifically, the JSBL can achieve an MSE performance of nearly 10^{-5} when the SNR is 10 dB, which demonstrates its superior angle estimation capability.

V. Conclusion

In this letter, channel estimation in ISAC systems has been investigated and a JSBL algorithm has been proposed. To capture the joint sparsity between S&C channels, an adaptive pattern-coupled prior for the channel coefficients has been conceived. Simulation results have demonstrated the effectiveness of the proposed JSBL algorithm and shown that exploiting the joint sparsity can improve channel estimation performance for both S&C systems.

References

[1] S. Lu et al., "Integrated sensing and communications: Recent advances and ten open challenges," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19094–19120, Jun. 2024.

[2] K. Chen, C. Qi, O. A. Dobre, and G. Y. Li, "Simultaneous beam training and target sensing in ISAC systems with RIS," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 2696–2710, Apr. 2024.

[3] W. Xu, Y. Xiao, A. Liu, M. Lei, and M.-J. Zhao, "Joint scattering environment sensing and channel estimation based on non-stationary Markov random field," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 3903–3917, May 2024.

[4] R. Liu, M. Li, H. Luo, Q. Liu, and A. L. Swindlehurst, "Integrated sensing and communication with reconfigurable intelligent surfaces: Opportunities, applications, and future directions," *IEEE Wireless Commun.*, vol. 30, no. 1, pp. 50–57, Feb. 2023.

[5] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, Jun. 2020.

[6] H. L. Harney, *Bayesian Inference: Parameter Estimation and Decisions*. Heidelberg, Germany: Springer, May 2003.

[7] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Jun. 2001.

[8] J. Dai, A. Liu, and V. K. N. Lau, "FDD massive MIMO channel estimation with arbitrary 2D-array geometry," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2584–2599, May 2018.

[9] C. E. Rasmussen and C. K. Williams, *Gaussian Process for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

Authorized licensed use limited to: Southeast University. Downloaded on August 16,2024 at 04:35:08 UTC from IEEE Xplore. Restrictions apply.