Sparse Bayesian Estimation of Millimeter-Wave Channel Correlation Matrix

Pajovic, M.; Wang, P.; Koike-Akino, T.; Orlik, P.V.

TR2019-042 July 03, 2019

Abstract

We propose an algorithm for estimating correlation matrix of mmWave channels. In addition to being sparse in angular/spatial domain, mmWave channel is commonly assumed to be time-invariant over a certain time period. However, more recent experiments indicate that while mmWave channel spatial directions can be assumed constant over some time period, their coefficients vary in time. Building upon this result, we probabilistically treat the correlation matrix estimation problem by associating sparse Bayesian learning prior to channel realizations, and performing statistical inference to recover channel directions and estimate their correlation matrix. The proposed algorithm is validated with simulations and shown to outperform benchmark methods based on greedy optimization-based sparse recovery.

IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)

Sparse Bayesian Estimation of Millimeter-Wave Channel Correlation Matrix

Milutin Pajovic, Pu Wang, Toshiaki Koike-Akino and Philip V. Orlik
Mitsubishi Electric Research Laboratories (MERL)
Cambridge, MA, 02139, USA
Email: pajovic@merl.com

***Abstract*—We propose an algorithm for estimating correlation matrix of mmWave channels. In addition to being sparse in angular/spatial domain, mmWave channel is commonly assumed to be time-invariant over a certain time period. However, more recent experiments indicate that while mmWave channel spatial directions can be assumed constant over some time period, their coefficients vary in time. Building upon this result, we probabilistically treat the correlation matrix estimation problem by associating sparse Bayesian learning prior to channel realizations, and performing statistical inference to recover channel directions and estimate their correlation matrix. The proposed algorithm is validated with simulations and shown to outperform benchmark methods based on greedy optimization-based sparse recovery.**

***Index Terms*—correlation matrix estimation, millimeter-wave channel, sparse Bayesian learning**

I. Introduction

Unprecedented data rates expected to be wirelessly delivered over millimeter-wave (mmWave) frequency ranges and different challenges associated with building actual mmWave systems have attracted considerable research interest in recent years [1]. Numerous measurements have shown that mmWave channel is sparse in spatial domain and, consequently, a variety of mmWave channel estimation algorithms exploit this inherent sparsity. One approach for mmWave channel estimation, employed by current mmWave communication standards [2], comprises of probing different spatial channel directions, one at a time, in consecutive time slots. This approach, however, leads to channel underutilization due to the need to exhaustively scan all possible spatial directions. A scheme trying to alleviate this issue is proposed in [3]. In an alternative approach, the channel is compressively sensed resulting in smaller number of required channel uses for training, as well as ability to train multiple channels in parallel in multi-user systems [4], [5]. With compressive sensing approach, the channel estimation is framed as Bayesian inference problem in [6], while [7] jointly estimates mmWave channel and compensates its frequency offset.

Commonly, the existing mmWave channel estimation algorithms assume the channel is time-invariant over a certain time period whose one part is used for channel estimation and the remaining one for information transmission. However, more recent experimental measurements have indicated that while spatial directions of mmWave channel may be unchanged during some time period, their coefficients are time-varying [8]. Therefore, channel correlation matrix, which contains information about channel spatial directions, i.e., angles of departures (AoD) and arrivals (AoA) of channel paths, as well as their powers, is necessary to establish the communication link between transmitter and receiver [9]. We propose in this paper an algorithm for estimation of mmWave channel correlation matrix. The algorithm comprises of two parts. The first part recovers channel spatial directions by modeling channel realizations with a hierarchal probabilistic generative model from sparse Bayesian learning (SBL) [10]. Although the SBL model assumes independent paths, the SBL prior promotes sparsity and admits tractable inference and is thus a handy approach for recovering non-zero channel directions even when they are correlated. Given the estimated AoDs and AoAs of the channel paths, we then develop a procedure to estimate correlation matrix of the recovered paths which then yields the overall channel correlation matrix. The proposed algorithm is tested with numerical simulations and shown to perform favorably in comparison to benchmarked optimization-based algorithms [9].

The AoA estimation problem, one application of which is mmWave channel estimation, has received significant attention in the literature. As such, a variety of basis pursuit (BP) and atomic norm denoising algorithms [14] have been proposed. Similarly, the SBL framework has been explored for mmWave channel estimation [11]–[13]. Like in those works, we assume the mmWave paths, i.e., their AoDs and AoAs, are unchanged over the observation period. However, contrary to those works, the path coefficients in our model are assumed to vary in time. Consequently, rather than estimating mmWave channel, which is commonly addressed, our challenge becomes the estimation of mmWave channel correlation matrix.

The closest work to this paper is [9], which frames estimation of mmWave channel correlation matrix as an optimization problem and proposes several greedy sparse recovery algorithms to solve it. In comparison to [9], we estimate channel correlation matrix through Bayesian inference procedure, where channel is modeled in a way to promote spatial sparsity and admit tractable inference. The simulations show that the proposed algorithm outperforms the methods from [9].

II. Signal Model

A. Channel Model

We assume transmitter and receiver are in communication over a narrowband mmWave channel. A baseband representation of the signal transmitted at discrete time n is $\mathbf{x}_n \in \mathbb{C}^{N_t \times 1}$, where N_t is the number of antennas on the transmitter side. The baseband representation of the corresponding signal, $\mathbf{y}_n \in \mathbb{C}^{N_r \times 1}$, received on N_r antennas is given by

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{v}'_n, \tag{1}$$

where $\mathbf{v}'_n$ is Gaussian noise and $\mathbf{H}_n \in \mathbb{C}^{N_r \times N_t}$ is time-varying narrowband mmWave channel. Numerous experimental measurements have shown that mmWave channel is sparse in angular domain, meaning that transmitted mmWave signal propagates over several distinct paths (line-of-sight, reflected and/or scattered path) before it reaches the receiver. To exploit channel sparsity in the spatial (i.e., angular) domain and facilitate its estimation, mmWave channel is commonly modeled as [4]

$$\mathbf{H}_n = \mathbf{A}_r \mathbf{G}_n \mathbf{A}_t^H, \tag{2}$$

where ^{H} denotes complex-conjugate transpose operator, $\mathbf{G}_n$ is channel representation in the virtual angular domain, $\mathbf{A}_r \in \mathbb{C}^{N_r \times G_r}$ and $\mathbf{A}_t \in \mathbb{C}^{N_t \times G_t}$ are manifold matrices of, respectively, receiver and transmitter arrays over corresponding angular domains discretized into G_r and G_t grid points. As a result of channel sparsity in the angular domain, $\mathbf{G}_n$ is a sparse matrix whose support directly indicates AoD and AoA of each non-zero channel path. Commonly, mmWave channel estimation algorithms from the literature assume the channel $\mathbf{G}_n$ is time-invariant over some time period. In comparison, we assume the AoD and AoA of channel paths remain constant during some time interval, while their coefficients vary in time. Formally, this means that the support of $\mathbf{G}_n$ is time-invariant while the values of its non-zero entries change with time n.

B. Sensing Model

A compressive sensing approach for training mmWave channel is to transmit/receive energy to/from random directions [4]. As such, the transmitter employs one radio-frequency (RF) chain to transmit a known symbol $s_n = 1$ modulated with a precoding vector $\mathbf{p}_n \in \mathbb{C}^{N_t \times 1}$. The precoding vector comprises of pseudo-random phasors such that the transmitted symbol is randomly phase shifted in each antenna yielding the transmitted signal $\mathbf{x}_n = \mathbf{p}_n$ that essentially insonifies random spatial channel directions.

The receiver applies a pseudo-random combining vector $\mathbf{q}_n \in \mathbb{C}^{N_r \times 1}$ in each of its M_r RF chains onto the received signal $\mathbf{y}_n$. The resulting signal $\mathbf{z}_n \in \mathbb{C}^{M_r \times 1}$ is given by

$$\mathbf{z}_n = \mathbf{Q}_n^H \mathbf{y}_n = \mathbf{Q}_n^H \mathbf{H}_n \mathbf{p}_n + \mathbf{v}_n, \tag{3}$$

where we approximately assume the noise is white, $\mathbf{v}_n \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{M_r})$, and $\mathbf{Q}_n \in \mathbb{C}^{N_r \times M_r}$ is the equivalent combining matrix containing individual combining vectors $\mathbf{q}_n$ in its columns. Vectorizing (3) and using the property that $\text{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right) \text{vec}(\mathbf{B})$ yields

$$\mathbf{z}_n = \left(\mathbf{p}_n^T \otimes \mathbf{Q}_n^H\right) \text{vec}(\mathbf{H}_n) + \mathbf{v}_n, \tag{4}$$

where $\otimes$ denotes Kronecker product. Substituting (2) into (4) and using the same property leads to

$$\begin{aligned} \mathbf{z}_n &= \left(\mathbf{p}_n^T \otimes \mathbf{Q}_n^H\right) \left(\mathbf{A}_t^T \otimes \mathbf{A}_r\right) \text{vec}(\mathbf{G}_n) + \mathbf{v}_n \\ &= \mathbf{B}_n \mathbf{g}_n + \mathbf{v}_n, \end{aligned} \tag{5}$$

where $\mathbf{g}_n \triangleq \text{vec}(\mathbf{G}_n)$ and $\mathbf{B}_n \triangleq \left(\mathbf{p}_n^T \otimes \mathbf{Q}_n^H\right) \left(\mathbf{A}_t^T \otimes \mathbf{A}_r\right)$.

We assume the channel is sensed over T time slots resulting in test statistics $\mathbf{z}_1, \dots, \mathbf{z}_T$ with known "sensing matrices" $\mathbf{B}_1, \dots, \mathbf{B}_T$ and unknown channel realizations $\mathbf{g}_1, \dots, \mathbf{g}_T$. As elaborated before, each channel realization $\mathbf{g}_n$ is a sparse vector with fixed support and varying values of non-zero entries over. Our goal is to estimate channel correlation matrix. Given that we represent mmWave channel in two domains, we correspondingly define channel correlation matrix in angular and array domain respectively as

$$\mathbf{\Sigma}_g = \mathbb{E}\left[\mathbf{g}_n \mathbf{g}_n^H\right] \tag{6}$$

$$\begin{aligned} \mathbf{\Sigma}_h &= \mathbb{E}\left[\text{vec}(\mathbf{H}_n)\text{vec}(\mathbf{H}_n)^H\right] \\ &= \left(\mathbf{A}_t^T \otimes \mathbf{A}_r\right) \mathbf{\Sigma}_g \left(\mathbf{A}_t^T \otimes \mathbf{A}_r\right)^H. \end{aligned} \tag{7}$$

As a side remark, we note that more than one RF chain on the transmitter side can be used for channel sensing. However, the resulting signal from each transmitter RF chain requires one channel use for transmission. Therefore, more transmitter RF chains do not bring more channel measurements within the fixed number of channel uses, alike with multiple receiver RF chains. Consequently, without loss of generality we assume a single transmitter RF chain.

III. Proposed Algorithm

The proposed algorithm for estimating correlation matrix of mmWave channel consists of two parts. The first part recovers the AoD and AoA of each channel path, while the second part estimates channel correlation matrix of the recovered paths.

A. Detection of Channel Paths

The channel support is recovered by describing channel vectors $\mathbf{g}_n$ with a specific hierarchical probabilistic generative model. The model is chosen so as to promote sparsity of channel vectors and, together with Gaussian likelihood for the observations $\mathbf{z}_n$, admit tractable Bayesian inference. The generative model assumes that channel vectors are conditionally independent samples from circularly symmetric Gaussian distribution

$$p(\mathbf{g}_n|\boldsymbol{\alpha}) = \mathcal{CN}(\mathbf{g}; \mathbf{0}, \mathbf{\Sigma}_\alpha), \tag{8}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{G_t G_r \times 1}$ is precision vector and $\mathbf{\Sigma}_\alpha = \text{diag}(\boldsymbol{\alpha})^{-1}$. The entries α_i, $i = 1, \dots, G_t G_r$, in the precision vector $\boldsymbol{\alpha}$ are further modeled as independent samples from Gamma distribution

$$p(\alpha_i) = \text{Gamma}(\alpha_i; c, d) = \frac{d^c}{\Gamma(d)} \alpha^c e^{-d\alpha_i}, \tag{9}$$

where c and d are hyper-parameters chosen such that the resulting prior distribution is uninformative. This is achieved for $c = \epsilon$ and $d = \epsilon$ with small $\epsilon \sim 10^{-5}$. The presented generative model has been used in various applications within the framework of sparse Bayesian learning (SBL) [10] with relatively recent theoretical validation [15]. The intuition behind this model stems from the fact that prior distribution $p(\mathbf{g}_n)$ (obtained by marginalizing out $\boldsymbol{\alpha}$) is peaked around zero and heavy tailed away from zero, meaning that the model promotes sparse solutions for $\mathbf{g}_n$.

In the case the AWGN variance σ^2 is unknown, it is modeled in a similar vein by assuming for precision $\alpha_0 \triangleq \sigma^{-2}$, $p(\alpha_0) = \text{Gamma}(\alpha_0; c_0, d_0)$ with hyper-parameters c_0 and d_0 chosen so as make this prior uninformative. As a final piece in the probabilistic formulation for the problem at hand, the likelihood model for observations $\mathbf{z}_n$ directly follows from (5) and Gaussian noise statistics such that

$$p(\mathbf{z}_n|\mathbf{B}_n, \mathbf{g}_n, \alpha_0) = \mathcal{CN}(\mathbf{g}; \mathbf{B}_n\mathbf{g}_n, \alpha_0^{-1}\mathbf{I}_{M_r}). \quad (10)$$

Having specified the probabilistic model, we now infer unknown parameters from the observations. To facilitate this process, we first note that

$$p(\mathbf{g}_n|\mathbf{z}_n, \boldsymbol{\Sigma}_\alpha, \alpha_0) = \mathcal{CN}(\mathbf{g}; \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n), \quad (11)$$

because likelihood (10) and conditional prior (8) are Gaussian distributions. The covariance matrix and mean vector of the Gaussian conditional posterior (11) are respectively given by

$$\boldsymbol{\Sigma}_n = \left(\alpha_0\mathbf{B}_n^H\mathbf{B}_n + \text{diag}(\boldsymbol{\alpha})\right)^{-1} \quad (12)$$

$$\boldsymbol{\mu}_n = \alpha_0\boldsymbol{\Sigma}_n\mathbf{B}_n^H\mathbf{z}_n. \quad (13)$$

The point estimates of precision vector $\boldsymbol{\alpha}$ and inverse variance α_0 can be obtained using the expectation-maximization (EM) algorithm [16] such that

$$\hat{\boldsymbol{\alpha}}, \hat{\alpha_0} = \arg\max_{\boldsymbol{\alpha},\alpha_0} \mathbb{E}\left[\log p(\{\mathbf{z}_n\}_{n=1}^T, \{\mathbf{g}_n\}_{n=1}^T, \boldsymbol{\alpha}, \alpha_0)\right], \quad (14)$$

where the expectation is taken with respect to

$$p(\{\mathbf{g}_n\}_{n=1}^T|\{\mathbf{z}_n\}_{n=1}^T, \boldsymbol{\alpha}, \alpha_0) = \prod_{n=1}^T p(\mathbf{g}_n|\mathbf{z}_n, \boldsymbol{\alpha}, \alpha_0), \quad (15)$$

which is further evaluated using (11). The log-likelihood of complete data in (14) is expanded into a sum of decoupled terms using conditional independence of $\mathbf{z}_n$ given $\mathbf{g}_n$, and $\mathbf{g}_n$ given $\boldsymbol{\alpha}$, as well as independence of entries α_i in $\boldsymbol{\alpha}$. Taking the first derivatives of the resulting expression with respect to α_i's and α_0, equating them to zero and solving for α_i's and α_0 yields

$$\hat{\alpha}_i = \frac{T + c - 1}{d + \sum_{n=1}^T (|[\boldsymbol{\mu}_n]_i|^2 + [\boldsymbol{\Sigma}_n]_{ii})}, i = 1, \ldots, G_tG_r \quad (16)$$

$$\hat{\alpha}_0 = \frac{M_rT + c_0 - 1}{d_0 + \sum_{n=1}^T (\|\mathbf{z}_n - \mathbf{B}_n\boldsymbol{\mu}_n\|_2^2 + \text{tr}\{\mathbf{B}_n\boldsymbol{\Sigma}_n\mathbf{B}_n^H\})}, \quad (17)$$

where $[\boldsymbol{\mu}_n]_i$ is the ith entry in $\boldsymbol{\mu}_n$, $[\boldsymbol{\Sigma}_n]_{ii}$ denotes the ith diagonal entry of $\boldsymbol{\Sigma}_n$, $\|\mathbf{x}\|_2$ is the $l2$ norm of $\mathbf{x}$, and $\text{tr}\{\mathbf{A}\}$ is the trace of a matrix $\mathbf{A}$.

Overall, the EM algorithm alternates between estimating α_i's and α_0 using (16) and (17), and inferring posterior distribution (11). The procedure is initialized with some large values for α_0 and $\boldsymbol{\alpha}$, and executed a certain number of iterations or until convergence is established. The EM algorithm yields posterior mean and covariance of each channel realization $\mathbf{g}_n$, as well as point estimates of precisions α_0 and $\boldsymbol{\alpha}$. Since the power in the spatial direction i is $1/\alpha_i$, the channel support is directly recovered from vector $\boldsymbol{\alpha}' = 1/\{\alpha_i\}_{i=1}^{G_tG_r}$ by pruning small values in $\boldsymbol{\alpha}'$, or as indices of S largest entries in $\boldsymbol{\alpha}'$ in the case the number of non-zero paths S is known and/or estimated using some other method. Simulations indicate that non-zero channel paths give rise to prominent peaks in $\boldsymbol{\alpha}'$.

B. Correlation Matrix of Channel Paths

In general, the channel correlation matrix $\boldsymbol{\Sigma}_g$ and AWGN variance σ^2 can be estimated, respectively, as $\boldsymbol{\Sigma}_\alpha$ and $1/\hat{\alpha}_0$. However, as the SBL generative model assumes independent α_i's, it is not suitable for correlation matrix estimation of channels with correlated paths. On the other hand, given the SBL prior promotes sparsity and admits tractable inference, it is a handy approach for recovering non-zero channel directions even when they are correlated. Therefore, the method from the previous part is used to recover channel support and here we present an algorithm for estimating correlation matrix of the detected channel directions (i.e, paths) and AWGN variance.

Given the channel support, the observations $\mathbf{z}_n$ are expressed as

$$\mathbf{z}_n = \tilde{\mathbf{B}}_n\tilde{\mathbf{g}}_n + \mathbf{v}_n, \quad (18)$$

where $\tilde{\mathbf{g}}_n$ is non-zero portion of $\mathbf{g}_n$ and $\tilde{\mathbf{B}}_n$ contains columns from $\mathbf{B}_n$ corresponding to the detected channel support. Similarly, the non-zero portion of the channel correlation matrix $\boldsymbol{\Sigma}_g$ is the correlation matrix of channel paths, denoted $\tilde{\boldsymbol{\Sigma}}_g$. Assuming the channel path coefficients $\tilde{\mathbf{g}}_n$ are samples from zero-mean complex Gaussian distribution of unknown correlation matrix $\tilde{\boldsymbol{\Sigma}}_g$, we employ the EM algorithm to estimate $\tilde{\boldsymbol{\Sigma}}_g$ as well as the AWGN variance $\hat{\sigma}^2$ by optimization

$$\sigma^2, \tilde{\boldsymbol{\Sigma}}_g = \arg\max_{\sigma^2, \tilde{\boldsymbol{\Sigma}}_g} \mathbb{E}\left[p(\{\mathbf{z}_n\}_{n=1}^K, \{\tilde{\mathbf{g}}_n\}_{n=1}^K; \sigma^2, \tilde{\boldsymbol{\Sigma}}_g)\right], \quad (19)$$

where the expectation is taken over $\prod_{n=1}^K p(\tilde{\mathbf{g}}_n|\mathbf{z}_n; \sigma^2, \tilde{\boldsymbol{\Sigma}}_g)$. Using the Gaussian likelihood model for observations $\mathbf{z}_n$ and Gaussian prior on channel path coefficients $\tilde{\mathbf{g}}_n$, we express the log-likelihood of complete data in (19) as the sum of decoupled terms, and solve the nonlinear optimization problem using first derivative method, yielding

$$\tilde{\boldsymbol{\Sigma}}_g = T^{-1}\sum_{n=1}^T \left(\tilde{\boldsymbol{\Sigma}}_n + \tilde{\boldsymbol{\mu}}_n\tilde{\boldsymbol{\mu}}_n^H\right) \quad (20)$$

and

$$\hat{\sigma}^{-2} = \frac{M_rT}{\sum_{n=1}^T \left(\|\mathbf{z}_n - \tilde{\mathbf{B}}_n\tilde{\boldsymbol{\mu}}_n\|_2^2 + \text{tr}\{\tilde{\mathbf{B}}_n\tilde{\boldsymbol{\Sigma}}_n\tilde{\mathbf{B}}_n^H\}\right)}, \quad (21)$$

where $\tilde{\boldsymbol{\Sigma}}_n$ and $\tilde{\boldsymbol{\mu}}_n$ are covariance matrix and mean vector of Gaussian posterior $p(\tilde{\mathbf{g}}_n|\mathbf{z}_n)$, respectively given by

$$\tilde{\boldsymbol{\Sigma}}_n = \left(\sigma^{-2}\tilde{\mathbf{B}}_n^H\tilde{\mathbf{B}}_n + \tilde{\boldsymbol{\Sigma}}_g^{-1}\right)^{-1} \tilde{\boldsymbol{\mu}}_n = \sigma^{-2}\tilde{\boldsymbol{\Sigma}}_n\tilde{\mathbf{B}}_n^H\mathbf{z}_n. \quad (22)$$

The EM procedure initializes $\tilde{\mathbf{\Sigma}}_g$ with a diagonal matrix whose diagonal elements are inverses of α_i's corresponding to the non-zero support detected in the previous part. Similarly, $\hat{\sigma}^2$ is initialized with α_0^{-1}. The alternating procedure is run a certain number of iterations or until convergence is established, and finally outputs $\hat{\sigma}^2$ and $\tilde{\mathbf{\Sigma}}_g$. The estimate of the channel correlation matrix in the angular domain, $\hat{\mathbf{\Sigma}}_g$, directly follows from $\tilde{\mathbf{\Sigma}}_g$, while that in the array domain, $\hat{\mathbf{\Sigma}}_h$, is obtained from $\hat{\mathbf{\Sigma}}_g$ using (7).

IV. Simulation Study

In all simulation tests, we assume uniform linear arrays (ULA), precoders and combiners introduce four possible phase shits (0^o, 90^o, 180^o and 270^o), and the array manifold matrices are discretised into 256 points.

As an illustrative example, we first consider a scenario with $N_t = N_r = 8$ antennas on both ends, $M_r = 4$ RF chains in the receiver, $S = 5$ independent channel paths, $T = 50$ pilots and SNR $= 0$ dB. The locations of true and estimated channel paths in the AoD-AoA domain and their powers are shown in Fig. 1. Notably, the proposed algorithm accurately recovers channel paths in the angular domain.

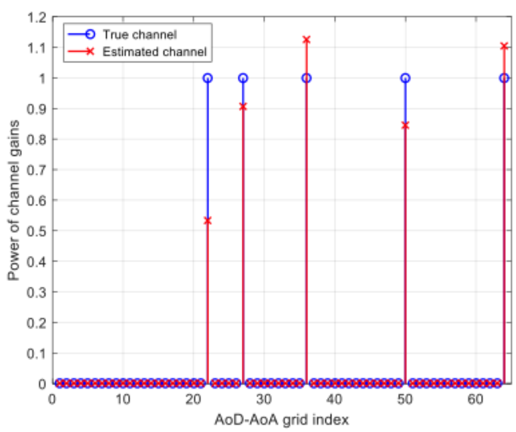

Fig. 1: True and estimated channel paths for $N_t = N_r = 8$, $M_r = 4$, $S = 5$, $T = 50$ and SNR $= 0$ dB.

The performance of the proposed algorithm is analyzed with simulations using the same scenario as in [9] so as to benchmark it against the performance of greedy sparse recovery algorithms reported therein. As such, the simulation scenario assumes channel with $S = 8$ independent mmWave paths, where transmitter and receiver respectively employ $N_t = 1$ and $N_r = 64$ antennas. The performance is measured with efficiency defined as

$$\eta_h = \frac{\mathrm{tr}\{\mathbf{U}^H_{\hat{\mathbf{\Sigma}}_h}\mathbf{\Sigma}_h\mathbf{U}_{\hat{\mathbf{\Sigma}}_h}\}}{\mathrm{tr}\{\mathbf{U}^H_{\mathbf{\Sigma}_h}\mathbf{\Sigma}_h\mathbf{U}_{\mathbf{\Sigma}_h}\}}, \tag{23}$$

where $\mathbf{\Sigma}_h$ and $\hat{\mathbf{\Sigma}}_h$ are, respectively, true and estimated channel correlation matrices in the array domain, while $\mathbf{U}_\mathbf{A}$ denotes eigenvector matrix of matrix $\mathbf{A}$. Due to lack of space, we skip further elaboration of the selected metric and remark that the same metric is used in [9]. The performance of greedy-based algorithms from [9] under the considered scenario are shown in Fig. 2.

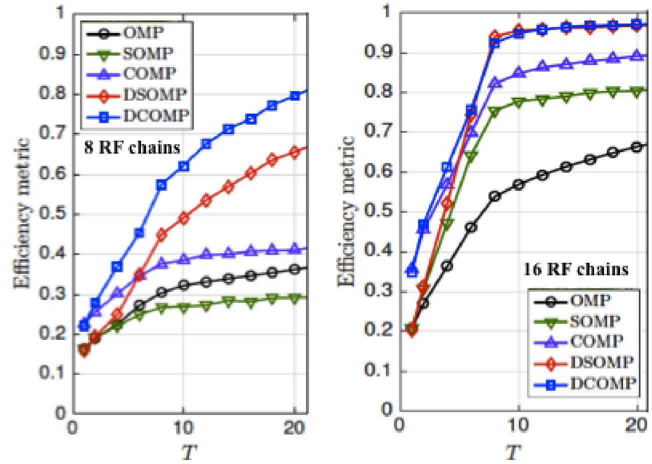

Fig. 2: Performance of greedy sparse recovery algorithms (Orthogonal Matching Pursuit (OMP), simultaneous OMP (SOMP), covariance OMP (COMP), dynamic SOMP (DSOMP) and dynamic COMP (DCOMP)) from [9] for $M_r = 8$ (left) and $M_r = 16$ (right).

The efficiencies of the proposed algorithm, simulated over 100 Monte-Carlo runs for each considered number of pilots T and $M_r = 8$ or $M_r = 16$ RF chains are shown in Fig. 3, where different cases as to which information is available on the receiver side are considered. More specifically, we consider the cases where information that the channel paths are independent, i.e., channel correlation matrix is of diagonal structure, as well as the noise variance are available/unavailable. In the case the receiver is aware the channel correlation matrix has diagonal structure, the estimation is carried out by updating only the diagonal elements in (20). As can be noted from the performance plots, the performance does not deteriorate when noise variance is unknown and estimated together with the channel correlation matrix. Furthermore, the lack of knowledge about diagonal structure of the channel correlation matrix causes insignificant performance deterioration only for a small number of pilots. Overall, for $M_r = 8$ RF chains and $T = 20$ pilots, our algorithm achieves nearly perfect efficiency, while the efficiency of the best greedy algorithm from Fig. 2 is 0.8. Similarly, for $M_r = 16$ RF chains, our algorithm achieves efficiency close to 1 and outperforms the best greedy algorithm from Fig. 2.

Finally, we consider the same communication system with 8 channel paths that exhibit different correlation patterns and measure efficiency of the proposed algorithm when the structure of the channel correlation matrix and noise variance are unknown. The channel paths are grouped into G mutually exclusive groups such that gains of the paths within the same group have unit correlation coefficient (i.e., are "fully correlated"), while those from different groups are uncorrelated. In other words, the underlying channel correlation matrix is block diagonal, where size of each block is equal to the

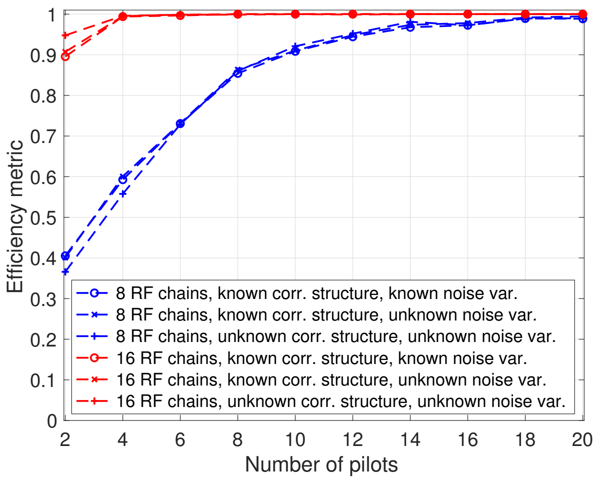

Fig. 3: Performance of the proposed algorithm with known/unknown path correlation matrix structure and noise variance.

group size. The considered group sizes are 1 (i.e., all paths are independent), 2, 4 and 8 (i.e., gains of all paths vary in the same manner) and the obtained performance is shown in Fig. 4. As can be seen, the efficiency gap between the most and least favorable cases of, respectively, uncorrelated and fully correlated paths is no more than 0.1, occurring when the number of pilots is 8. However, the efficiency gap closes as the number of pilots increases, leading to relatively insignificant degradation in the most extreme case of fully correlated paths with respect to other more practical cases. Finally, comparing the results in Fig. 4 with those in Fig 2 for 8 RF chains yields that the proposed algorithm in the most extreme case of full correlation among channel paths outperforms the best greedy optimization-based algorithm in the most favorable case of uncorrelated paths.

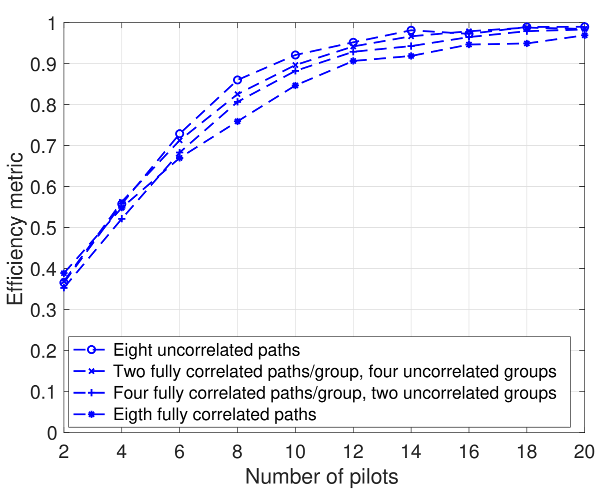

Fig. 4: Performance of the proposed algorithm for 8 RF chains and different correlation matrix structures of 8 mmWave paths.

V. Conclusion

We presented in this paper an algorithm for mmWave channel correlation matrix estimation where mmWave channel is assumed to have sparse and time-invariant directions with correlated time-varying coefficients. The algorithm builds upon sparse Bayesian learning framework leveraged to detect mmWave channel directions and then estimates channel correlation matrix using separate procedure. The algorithm is tested with simulations and shown to outperform the benchmark state-of-the-art methods based on greedy optimization-based sparse recovery.

References

[1] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K.Wang, G. N.Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

[2] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*, Pearson Education, 2014.

[3] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, "Millimeter wave MIMO channel estimation using overlapped beam patterns and rate adaptation," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 601–616, February 2017.

[4] A.Alkhateeb, O.El Ayach, G.Leus, and R.Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Jour. Select. Topics in Sig. Proc.*, vol. 8, no. 5, pp. 831–846, October 2014.

[5] A. Alkhateeby, G. Leus, and R.W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2909–2913, April 2015.

[6] R. T. Suryaprakash, M. Pajovic, K. J. Kim, and P. Orlik, "Millimeter wave communications channel estimation via Bayesian group sparse recovery," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3406–3410, March 2016.

[7] M. Pajovic, P. Wang, T. Koike-Akino, and P. Orlik, "Estimation of frequency unsynchronized millimeter-wave channels," *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1205–1209, November 2017.

[8] P. A. Eliasi, S. Rangan, and T. S. Rappaport, "Low-rank spatial channel estimation for millimeter wave cellular systems," *arXiv: 1410.4831v1*, October 2017.

[9] S. Park and R. W. Heath, "Spatial channel covariance estimation for the hybrid MIMO architecture: A compressive sensing based approach," *arXiv:1711.04207v1*, November 2017.

[10] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, September 2001.

[11] A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Sparse Bayesian learning-based channel estimation in millimeter wave hybrid mimo systems," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2017, pp. 1–5.

[12] H. Tang, J. Wang, and L. He, "Off-grid sparse Bayesian learning-based channel estimation for mmwave massive mimo uplink," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 45–48, Feb 2019.

[13] M. Umar Aminu, M. Codreanu, and M. Juntti, "Bayesian learning based millimeter-wave sparse channel estimation with hybrid antenna array," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 1–5.

[14] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, Dec 2013.

[15] M. Pajovic, "Misspecified Bayesian Cramer-Rao bound for sparse bayesian learning," *IEEE Statistical Signal Processing (SSP) Workshop*, July 2018.

[16] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, 2006.