Nested Sparse Approximation: Structured Estimation of V2V Channels Using Geometry-Based Stochastic Channel Model

Sajjad Beygi Student Member, IEEE, Urbashi Mitra Fellow, IEEE, and Erik G. Ström* Senior Member, IEEE

Abstract—Future intelligent transportation systems promise increased safety and energy efficiency. Realization of such systems will require vehicle-to-vehicle (V2V) communication technology. High fidelity V2V communication is, in turn, dependent on accurate V2V channel estimation. V2V channels have characteristics differing from classical cellular communication channels. Herein, geometry-based stochastic modeling is employed to develop a characterization of V2V channel channels. The resultant model exhibits significant structure; specifically, the V2V channel is characterized by three distinct regions within the delay-Doppler plane. Each region has a unique combination of specular reflections and diffuse components resulting in a particular element-wise and group-wise sparsity. This joint sparsity structure is exploited to develop a novel channel estimation algorithm. A general machinery is provided to solve the jointly element/group sparse channel (signal) estimation problem using proximity operators of a broad class of regularizers. The alternating direction method of multipliers using the proximity operator is adapted to optimize the mixed objective function. Key properties of the proposed objective functions are proven which ensure that the optimal solution is found by the new algorithm. The effects of pulse shape leakage are explicitly characterized and compensated, resulting in measurably improved performance. Numerical simulation and real V2V channel measurement data are used to evaluate the performance of the proposed method. Results show that the new method can achieve significant gains over previously proposed methods.

I. INTRODUCTION

VEHICLE-to-vehicle (V2V) communication is central to future intelligent transportation systems, which will enable efficient and safer transportation with reduced fuel consumption [1]. In general, V2V communication is anticipated to be short range with transmission ranges varying from a few meters to a few kilometers between two mobile vehicles on a road. A big challenge of realizing V2V communication is the inherent fast channel variations (faster than in cellular [2], [3]) due to the mobility of both the transmitter and receiver. Since channel state information can improve communication performance, a fast algorithm to accurately estimate V2V channels is of interest. Furthermore, V2V channels are highly dependent on the geometry of road and the local physical environment [1], [4].

A popular estimation strategy for fast time-varying channels is to apply Wiener filtering [5], [6]. Recently, [2], [6] present an adaptive Wiener filter to estimate V2V channels using subspace selection. The main drawback of Wiener-filtering is that the knowledge of the scattering function is required [5]; however, the scattering function is not typically known at the receiver. Often, a flat spectrum in the delay-Doppler domain is assumed, which introduces performance degradation due to the mismatch with respect to the true scattering function [6].

In this work, we adopt a V2V channel model in the delay-Doppler domain, using the geometry-based stochastic channel model proposed in [7]. Our characterization of this model reveals the special structure of the V2V channel components in the delay-Doppler domain. We show that the delay-Doppler representation of the channel exhibits three key regions; within these regions, the channel is a mixture of specular reflections and diffuse components. While the specular contributions appear all over the delay-Doppler plane sparsely, the diffuse contributions are concentrated in specific regions of the delay-Doppler plane. We see that the channel measurements from a real data experiment also confirm to our analysis of the V2V channel structures in the delay-Doppler domain.

In our prior work [8], [9], a Hybrid Sparse/Diffuse (HSD) model was presented for a mixture of sparse and Gaussian diffuse components for a static channel, which we have adapted to estimate a V2V channel [10]. This approach requires information about the V2V channel such as the statistics, and power delay profile (PDP) of the diffuse and sparse components [8]. Another approach for time-varying frequency-selective channel estimation is via compressed sensing (CS) or sparse approximation based on an $l_1$-norm regularization [11–13]. These algorithms perform well for channels with a small number of scatterers or, clusters of scatterers. For V2V channels, diffuse contributions from reflections along the roadside will degrade the performance of CS methods that only consider element-wise sparsity [6], [12].

Herein, we exploit the particular structure of the V2V channel, inspired by recent work in 2D sparse signal estimation [14], [15], to design a novel joint element- and group-wise sparsity estimator to estimate the 2D time-varying V2V channel using received data. Our proposed method provides a general machinery to solve the joint sparse structured estimation problem with a broad class of regularizers that promote sparsity. We show that our proposed algorithm covers both well-known convex and non-convex regularizers such as smoothly clipped absolute deviation (SCAD) regularizers [16], and the minimax concave penalty (MCP) [17] that were proposed for element-wise sparsity estimation. We also present a general way to design a proper regularizer function for joint sparsity problem in our previous work [18]. Our method can be applied to scenarios beyond V2V channels.

Recent algorithms for hierarchical sparsity (sparse groups with sparsity within the groups) [19–21] also consider a mixture of penalty functions (group-wise and element-wise).
Of particular note is [20] where a similar nested solution is determined, also in combination with the alternating direction method of multipliers (ADMM) as we do herein. Our modeling assumptions can be viewed as a generalization of their assumptions which results in the need for different methods for proving the optimization of the nested structures. In particular, [20] examines a particular proximity operator for which the original regularizing function is never specified.1 This proximity operator is built by generalizing the structure of the proximity operator for the $l_p$ norm. In contrast, we begin with a general class of regularizing functions, we specify the properties needed for such functions (allowing for both convex and non-convex functions). Thus, our proof methods rely only on the properties induced by these assumptions. Furthermore, our results are also applicable to the problem of hierarchical sparsity [19–21]. We observe that the results in [19–21] cannot be applied to non-convex regularizer functions such as SCAD and MCP due to their concavity and a non-linear dependence on the regularization parameter.

To find the optimal solution of our joint sparsity objective function, we take advantage of ADMM [22], which is a very flexible and efficient tool for optimization problems whose objective functions are the combination of multiple terms. Furthermore, we use the proximity operator [23] to show that the estimation can be done by using simple thresholding operations in each iteration, resulting in low complexity.

We also address the channel leakage effect due to finite block length and bandwidth for channel estimation in the delay-Doppler plane. In [12], the basis expansion for the scattering function is optimized to compensate for the leakage which can degrade performance. The resulting expansion in [12] is computationally expensive. Herein, we take an alternative view and show that with the proper sampling resolution in time and frequency, we can explicitly derive the leakage pattern and robustify the channel estimator with this knowledge at the receiver to improve the sparsity, compensate for leakage, and maintain modest algorithm complexity. The main contributions of this work are as follows:

1) A general framework for joint sparsity estimation problem is proposed, which covers a broad class of regularizers including convex and non-convex functions. Furthermore, we show that the solution for joint sparse estimation problem is computed by applying the element-wise and group-wise structure in a nested fashion using simple thresholding operations.

2) We provide a simple model for the V2V channel in the delay-Doppler plane, using the geometry-based stochastic channel modeling proposed in [7]. We characterize the three key regions in the delay-Doppler domain with respect to the presence of sparse specular and diffuse components. This structure is verified by experimental channel measurement data, as presented in Section VIII.

3) The leakage pattern is explicitly computed and a compensation procedure proposed.

4) A low complexity joint element- and group-wise sparsity V2V channel estimation algorithm is proposed exploiting the aforementioned channel model and optimization result.

5) We use extensive numerical simulation and experimental channel measurement data to investigate the performance of the proposed joint sparse channel estimators and show that our method outperforms classical and compressed sensing methods [6], [10–12]

The rest of this paper is organized as follows. In Section II, we review some definitions from variational analysis and present our key optimization result for a joint sparse and group sparse signal estimation. In Section III, the system model for V2V communications is presented. In Section IV, the geometry-based V2V channel model is developed. The observation model and leakage effect are computed in Section V. In Section VI, the channel estimation algorithm for the time-varying V2V channel model using joint sparsity structure is presented. In Section VII, we provide simulation results and compare the performance of the estimators. In Section VIII, the real channel measurements are provided to confirm the validity of the channel model and numerical simulation. Finally, Section IX concludes the paper. We present our proofs, region specifier algorithm, and analysis of our proximal ADMM in the Appendices.

**Notation:** We denote a scalar by $x$, a column vector by $x$, and its $i$-th element with $x[i]$. Similarly, we denote a matrix by $X$ and its $(i,j)$-th element by $X[i,j]$. The transpose of $X$ is given by $X^T$ and its conjugate transpose by $X^H$. A diagonal matrix with elements $x$ is written as $\text{diag}(x)$ and the identity matrix as $I$. The set of real numbers by $\mathbb{R}$, and the set of complex numbers by $\mathbb{C}$. The element-wise (Schur) product is denoted by $\circ$.

## II. JOINTLY SPARSE SIGNAL ESTIMATION: OPTIMIZATION RESULT

In this section, we propose a unified framework using proximity operators [24] to solve the optimization problem imposed by a structured sparse signal estimation problem. Then, we apply this machinery to estimate the V2V channel, exploiting the group- and element-wise sparsity structure discovered in Section IV.

Proximal methods have drawn increasing attention in the signal processing (e.g., [23], and numerous references therein) and the machine learning communities (e.g., [25], [26], and references therein), due to their convergence rates (optimal for the class of first-order techniques) and their ability to accommodate large, non-smooth, convex (and non-convex) problems. In proximal algorithms, the base operation is evaluating the proximal operator of a function, which involves solving a small optimization problem. These sub-problems can be solved with standard methods, but they often admit closed-form solutions or can be solved efficiently with specialized numerical methods. Our main theoretical contribution is stated in Theorem 1 in Section II.B. In this theorem, we show that, for our proposed class of regularization functions, the nested joint sparse structure can be recovered by applying the element-wise sparsity and group-wise sparsity structures in a nested fashion, see Eq. (10).

---

1Note given a particular regularization function, there is a unique proximity operator, but for a given proximity operator there may exist more than one regularizer function.

2A jointly sparse signal in this paper, is a signal that has both element-wise and group-wise sparsity.
A. Proximity Operator

We start with the definition of a proximity operator from variational analysis [24].

**Definition 1.** Let $\phi(\mathbf{a}; \lambda)$ be a continuous real-valued function of $\mathbf{a} \in \mathbb{R}^N$, the proximity operator $P_{\lambda, \phi}(\mathbf{b})$ is defined as

$$
P_{\lambda, \phi}(\mathbf{b}) := \text{argmin}_{\mathbf{a} \in \mathbb{R}^N} \left\{ \frac{1}{2} \| \mathbf{b} - \mathbf{a} \|_2^2 + \phi(\mathbf{a}; \lambda) \right\},$$

where $\mathbf{b} \in \mathbb{R}^N$ and $\lambda > 0$.

**Remark 1.** If $\phi(\cdot)$ is a separable function, i.e., $\phi(\mathbf{a}; \lambda) = \sum_{i=1}^N \phi_i(\mathbf{a}[i]; \lambda)$. Then, $[P_{\lambda, \phi}(\mathbf{a})]_i = \text{argmin}_{x \in \mathbb{R}^N} \left\{ \frac{1}{2} \| \mathbf{b} - \mathbf{a} \|_2^2 + \phi(\mathbf{a}[i]; \lambda) \right\}$.

**Remark 2.** If the objective function $J(\mathbf{a}) = \frac{1}{2} \| \mathbf{b} - \mathbf{a} \|_2^2 + \phi(\mathbf{a}; \lambda)$ is a strictly convex function, the proximity operator of $\phi(\mathbf{a}; \lambda)$ admits a unique solution.

**Remark 3.** Furthermore, $P_{\lambda, \phi}(\mathbf{b})$ is characterized by the inclusion

$$
\forall (\mathbf{a}^*, \mathbf{b}), \quad \mathbf{a}^* = P_{\lambda, \phi}(\mathbf{b}) \iff \mathbf{a}^* - \mathbf{b} \in \partial \phi(\mathbf{a}^*; \lambda), \quad (2)
$$

where $\partial \phi(\cdot)$ is the sub-gradient of the function $\phi$ [24].

Note that $\phi$ does not need to be a convex or differentiable function to satisfy the conditions noted in Remarks 2 and 3. Proximity operators have a very natural interpretation in terms of denoising [23], [24]. Consider the problem of estimating a vector $\mathbf{a} \in \mathbb{R}^N$ from an observation $\mathbf{b} \in \mathbb{R}^N$, namely $\mathbf{b} = \mathbf{a} + \mathbf{n}$ where $\mathbf{n}$ is additive white Gaussian noise. If we consider the regularization function $\phi(\mathbf{a}; \lambda)$ as the prior information about the vector $\mathbf{a}$, then $P_{\lambda, \phi}(\mathbf{b})$ can be interpreted as a maximum a posteriori (MAP) estimate of the vector $\mathbf{a}$ [27]. Two well-known types of prior information about the structure of a vector are element-wise sparsity and group-wise sparsity structures. A $N$-vector $\mathbf{a}$ is element-wise sparse if, the number of non-zero (or larger than some threshold) entries in the vector is small compared to the length of the vector.

To define the group-wise sparsity, let us consider $(I_i)_{i=1}^{N_g}$ be a partition of the index set $\{1, 2, \ldots, N\}$ into $N_g$ groups, namely $\cup_{i=1}^{N_g} I_i = \{1, 2, \ldots, N\}$ and $I_i \cap I_j = \emptyset$ for all $i \neq j$. We define group vectors as follows

$$
\mathbf{a}_i[k] = \begin{cases} 
\mathbf{a}[k] & k \in I_i \\
0 & k \notin I_i
\end{cases}, \quad (3)
$$

for $i = 1, \ldots, N_g$, where $N_g$ is the total number of group vectors. Based on above definition, we have $\mathbf{a} = \sum_{i=1}^{N_g} \mathbf{a}_i$, and the nonzero elements of $\mathbf{a}_i$ and $\mathbf{a}_j$ are non-overlapping for $i \neq j$. The vector $\mathbf{a}$ is called a group-wise sparse vector, if the number of group vectors $\mathbf{a}_i$ for $i = 1, \ldots, N_g$ with non-zero $l_2$-norm (or $l_2$-norm larger than some threshold) is small compared to the total number of group vectors, $N_g$.

B. Optimality of the Nesting of Proximity Operators

We consider the estimation of the vector $\mathbf{a}$ from vector $\mathbf{b}$ as noted above. Furthermore, suppose that the vector $\mathbf{a}$ is a jointly sparse vector. The desired optimization problem is as follows

$$
\hat{\mathbf{a}} = \text{argmin}_{\mathbf{a} \in \mathbb{R}^N} \left\{ \frac{1}{2} \| \mathbf{b} - \mathbf{a} \|_2^2 + \phi_y(\mathbf{a}; \lambda_y) + \phi_c(\mathbf{a}; \lambda_c) \right\}, \quad (4)
$$

where $\phi_y(\mathbf{a}; \lambda_y)$ is a regularization term to induce group sparsity and $\phi_c(\mathbf{a}; \lambda_c)$ is a term to induce the element-wise sparsity. In general, the weighting parameters, $\lambda_y > 0$, $\lambda_c > 0$, can be selected from a given range via cross-validation, by varying one of the parameters and keeping the other fixed [21]. We further consider penalty functions $\phi_y(\mathbf{a}; \lambda_y)$ and $\phi_c(\mathbf{a}; \lambda_c)$ of the form:

$$
\phi_y(\mathbf{a}; \lambda_y) = \sum_{j=1}^{N_g} g_j \| \mathbf{a}_j \|_2; \lambda_y), \quad (5)
$$

$$
\phi_c(\mathbf{a}; \lambda_c) = \sum_{i=1}^{N} f_i (\mathbf{a}[i]; \lambda_c), \quad (6)
$$

where $g_j : \mathbb{R} \rightarrow \mathbb{R}$ and $f_i : \mathbb{R} \rightarrow \mathbb{R}$ are continuous functions to promote sparsity on groups and elements, respectively, $N$ is the length of vector $\mathbf{a}$, and $N_g$ is the number of groups in vector $\mathbf{a}$. Our goal here is to derive the solution of the optimization problem in (4) using the proximity operators of the functions $f_i$ and $g_j$. We state the conditions imposed on the regularizers, in terms of the univariate functions $g_j(x; \lambda)$ and $f_i(x; \lambda)$ to promote sparsity and also to control the stability of the solution of the optimization problem in (4).

**Assumption 1:** For $k \in \{e, g\}$

i. $f_k$ is a non-decreasing function of $x$ for $x \geq 0$; $f_k(0; \lambda) = 0$, and $f_k(x; 0) = 0$.

ii. $f_k$ is differentiable except at $x = 0$.

iii. For $\forall z \in \partial f_k(0; \lambda)$, then $|z| \leq \lambda$, where $\partial f_k(0; \lambda)$ is the subgradient of $f_k$ at zero.

iv. There exists a $\mu \leq \frac{1}{2}$ such that the function $f_k(x; \lambda) + \mu x^2$, is convex.

v. $g_j$ is a homogeneous function, i.e., $g_j(\alpha x; \alpha \lambda) = \alpha^2 g_j(x; \lambda)$ for all $\alpha > 0$.

vi. $f_i$ is a scale invariant function, i.e., $f_i(\alpha x; \alpha \lambda) = f_i(x; \alpha \lambda) = \alpha f_i(x; \lambda)$ for all $\alpha > 0$.

It can be observed that conditions (i), (iv), (v), and (vi) ensure the existence of the minimizer of the optimization problem in Eq. (4), and they induce norm properties on the regularizer function. Assumption (ii) promotes sparsity, (iii) controls the stability of the solution in Eq. (4) and guarantees the optimality of solution of optimization problem in Eq. (4) or ($P_0$) in Section VI. Finally Assumption (iv) enables the inclusion of many non-convex functions in the optimization problem. Note that the scale invariant property of $f_i$ implies that $f_i$ also satisfies (v) (is a homogeneous function).

Many pairs of regularizer functions satisfy Assumption I. For instance, the $l_1$-norm, namely $f_{l_1}(x; \lambda) = \lambda |x|$ and $f_i(x; \lambda) = \lambda |x|$, satisfy Assumption I (see Appendix A). We note that two recently popularized non-convex functions, SCAD and MCP regularizers, also satisfy Assumption I. It is worth pointing out that SCAD and MCP are more effective in promoting sparsity than the $l_p$ norms. The SCAD regularizer [16], is given by

$$
f_{\mu}(x; \lambda) = \begin{cases} 
\lambda |x| & \text{for } |x| \leq \lambda \\
-\frac{x^2 - 2\mu_0 \lambda |x| + \lambda^2}{2(\mu_0 - 1)} & \text{for } |x| > \mu_0 \lambda
\end{cases}, \quad (7)
$$

where $\mu_0 > 1$. A precise definition of subgradients of functions is given in [24], page 301.
where $\mu_S > 2$ is a fixed parameter, and the MCP regularizer [17], is
\[
f_g(x; \lambda) = \text{sign}(x) \int_0^{\|x\|} \left( \lambda - \frac{z}{\mu_M} \right) dz,
\]
and finally the proximity operator for the MCP regularizer [17] is
\[
P_{\lambda,f_g}(x) = \begin{cases} 0, & |x| \leq \lambda \\ \frac{x - \text{sign}(x) \lambda}{1 - \frac{\lambda}{\mu_M}}, & |x| > \lambda \mu_M \\ \text{sign}(x), & |x| < \lambda \mu_M \end{cases}
\]
In Fig. 1, MCP and SCAD proximity operators are depicted for $\lambda = 1$ and three different values of the parameters $\mu_S$ and $\mu_M$. It is clear that when $\mu_S$ and $\mu_M$ are large, both SCAD and MCP operators behave like the soft-thresholding operator (for $x$ smaller than $\mu_S \lambda$ and $\mu_M \lambda$, respectively).

In the sequel, first we model a V2V communication system. Then, we show that V2V channel representation in the delay-Doppler domain has both element-wise and group-wise sparsity structures. Finally, we apply our key optimization result derived in this section to estimate the V2V channel using an ADMM algorithm.

### III. Communication System Model

We will consider communication between two vehicles as shown in Fig. 2. The transmitted signal $s(t)$ is generated by the modulation of the transmitted pilot sequence $s[n]$ onto the transmit pulse $p_t(t)$ as,
\[
s(t) = \sum_{n=-\infty}^{+\infty} s[n] p_t(t - nT_s),
\]
where $T_s$ is the sampling period. Note that this signal model is quite general, and encompasses OFDM signals as well as single-carrier signals. The signal $s(t)$ is transmitted over a linear, time-varying, V2V channel. The received signal $y(t)$ can be written as,
\[
y(t) = \int_{-\infty}^{+\infty} h(t, \tau) s(t - \tau) d\tau + z(t).
\]
Here, $h(t, \tau)$ is the channel’s time-varying impulse response, and $z(t)$ is a complex white Gaussian noise. At the receiver, $y(t)$ is converted into a discrete-time signal using an anti-aliasing filter $p_r(t)$. That is,
\[
y[n] = \int_{-\infty}^{+\infty} y(t) p_r(t N_s - t) dt.
\]
The relationship between the discrete-time signal $s[n]$ and received signal $y[n]$, using Eqs. (16)–(18), can be written as,
\[
y[n] = \sum_{m=-\infty}^{+\infty} h_t[n, m] s[n - m] + z[n],
\]
where \( h_t[n, m] \) is the discrete time-delay representation of the observed channel, which is related to the continuous-time channel impulse response \( h(t, \tau) \) as follows,

\[
h_t[n, m] = \int_{-\infty}^{+\infty} h(t + nT_s, \tau) p_i(t - \tau + mT_s)p_r(-t) \, dt \, d\tau.
\]

With some loss of generality, we assume that \( p_r(t) \) has a root-Nyquist spectrum with respect to the sample duration \( T_s \), which implies that \( z[n] \) is a sequence of i.i.d. circularly symmetric complex Gaussian random variables with variance \( \sigma^2_z \), and that \( h_t[n, m] \) is causal with maximum delay \( M - 1 \), i.e., \( h_t[n, m] = 0 \) for \( m \geq M \) and \( m < 0 \). We can then write

\[
y[n] = \sum_{m=0}^{M-1} \left( \sum_{k=-K}^{K} H_t[k, m] e^{j \frac{2\pi nk}{\lambda T}} \right) s[n - m] + z[n], \quad \text{for } n = 0, 1, \ldots, N_r - 1,
\]

where \( 2K + 1 \geq N_r \), and

\[
H_t[k, m] = \frac{1}{2K + 1} \sum_{n=0}^{N_r-1} h_t[n, m] e^{-j \frac{2\pi nk}{\lambda T}}, \quad \text{for } |k| \leq K,
\]

is the discrete delay-Doppler spreading function of the channel. Here, \( N_r \) denotes the total number of received signal samples used for the channel estimation.

### IV. Joint Sparsity Structure of V2V Channels

In this section, we adopt the V2V geometry-based stochastic channel model from [7] and analyze the structure such a model imposes on the delay-Doppler spreading function. The V2V channel model considers four types of multipath components (MPCs): (i) the effective line-of-sight (LOS) component, which may contain the ground reflections, (ii) discrete components generated from reflections of discrete mobile scatterers (MD), e.g., other vehicles, (iii) discrete components reflected from discrete static scatterers (SD) such as bridges, large traffic signs, etc., and (iv) diffuse components (DI). Thus, the V2V channel impulse response can be written as

\[
h(t, \tau) = h_{\text{LOS}}(t, \tau) + \sum_{i=1}^{N_{MD}} h_{\text{MD},i}(t, \tau)
\]

\[
+ \sum_{i=1}^{N_{SD}} h_{\text{SD},i}(t, \tau) + \sum_{i=1}^{N_{DI}} h_{\text{DI},i}(t, \tau),
\]

where \( N_{MD} \) denotes the number of discrete mobile scatterers, \( N_{SD} \) is the number of discrete static scatterers and \( N_{DI} \) is the number of diffuse scatterers, respectively. Typically, \( N_{DI} \) is much larger than \( N_{SD} \) and \( N_{MD} \) [7]. In the above representation, the multipath components can be modeled as

\[
h_i(t, \tau) = \eta_i \delta(\tau - \tau_i)e^{-j2\pi\nu_i t},
\]

where \( \eta_i \) is the complex channel gain, \( \tau_i \) is the delay, and \( \nu_i \) is the Doppler shift associated with path \( i \) and \( \delta(t) \) is the Dirac delta function. The channel description in (22) and (23)

\[ \text{iv} \]

\[ \text{V} \]

The subscript \( \text{iv} \) hereafter denotes the channel with leakage. We discuss the channel leakage effect with more details in Section V.

Fig. 2: Geometric representation of the V2V channel. The shaded areas on each side of the road contain static discrete (SD) and diffuse (DI) scatters, while the road area contains both SD and moving discrete (MD) scatters.

Explicitly models distance-dependent pathloss and scatterer parameters [7]. We assume that these parameters can be approximated as time-invariant over the pilot signal duration. This is a reasonable assumption in practical systems as will be illustrated in Figures 12 and 13 for the experimental channel measurement data. The spatial distribution of the scatterers and the statistical properties of the complex channel gains are specified in [7] for rural and highway environments. It is shown that the channel power delay profile is exponential. Further details about the spatial evolution of the gains can be found in [7], [28]. In geometry-based stochastic channel modeling, point scatterers are randomly distributed in a geometry according to a specified distribution. The position and speed of the scatterers, transmitter, and receiver determine the delay-Doppler parameters for each MPC, which in turn, together with the transmitter and receive pulse shapes, determine \( H_t[k, m] \).

We next determine the delay and Doppler contributions of an ensemble of point scatterers of type (i)-(iv) for the road geometry depicted in Fig. 2. If vehicles are assumed to travel parallel with the \( x \)-axis, the overall Doppler shift for the path from the transmitter (at position TX) via the point scatterer (at position P) to the receiver (at position RX) can be written as [7]

\[
\nu(\theta_t, \theta_r) = \frac{1}{\lambda_\nu} \left[ (v_T - v_P) \cos \theta_t + (v_R - v_P) \cos \theta_r \right],
\]

where \( \lambda_\nu \) is the wavelength, \( v_T, v_P, \) and \( v_R \) are the speed of the transmitter, scatterer, and receiver, respectively, and \( \theta_t \) and \( \theta_r \) is the angle of departure and arrival, respectively. The path delay is

\[
\tau = \frac{d_1 + d_2}{c_0},
\]

where \( c_0 \) is the propagation speed, \( d_1 \) is the distance from TX to P, and \( d_2 \) is the distance from P to RX. The path parameters \( \theta_t, \theta_r, d_1, \) and \( d_2 \) are easily computed from TX, P, and RX. The delay and Doppler parameters of each component (i)-(iv) can now be specified by Eqs. (24) and (25).

**LOS**: If it exists, the most significant component of the V2V channel is the line of sight (LOS) component, which will have delay and Doppler parameters \( \tau_0 = \frac{d_0}{c_0} \) and \( \nu_0 = \frac{1}{\lambda_\nu} (v_T - v_R) \cos(\theta) \), where \( d_0 \) is the distance from TX and
Fig. 3: Delay-Doppler contribution of line-of-sight (LOS) and from static scatterers (SD/DI) placed on a parallel line beside the road.

RX and \( \theta \) is the angle between the x-axis (i.e., the moving direction) and the line passing through TX and RX.

**Diffuse Scatterers:** The diffuse (DI) scatterers are static (\( v_P = 0 \)) and uniformly distributed in the shadowed regions in Fig. 2. Suppose we place a static scatterer at the coordinates \((x, y)\). The delay-Doppler pair, \((\tau(x, y), \nu(x, y))\), for the corresponding MPC can be calculated from Eqs. (24) and (25). If we fix \( y = y_0 \) and vary \( x \) from \(-\infty\) to \(+\infty\), the delay-Doppler pair will trace out a U-shaped curve in the delay-Doppler plane, as depicted in Fig. 3.

Repeating this procedure for the permissible \( y \)-coordinates for the DI scatterers, \([y_0] \in [D/2, d + D/2]\), will result in a family of curves that are confined to a U-shaped region, see Fig. 4. Hence, the DI scatterers will result in multipath components with delay-Doppler pairs inside this region. The maximum and minimum Doppler values of the region is easily found from Eq. (24). In fact, it follows from Eq. (24) that the Doppler parameter of an MPC due to a static scatterer will be confined to the symmetric interval \([-\nu_S, \nu_S]\), where \( \nu_S = \frac{1}{2\nu_T} (\nu_T + \nu_R) \).

**Static Discrete Scatterers:** The static discrete (SD) scatterers can appear outside the shadowed regions in Fig. 2. In fact, the \( y \)-coordinates of the SD scatterers are drawn from a Gaussian mixture consisting of two Gaussian pdfs with the same standard deviation \( \sigma_{y,SD} \) and means \( y_1,SD \) and \( y_2,SD \) [7]. The delay-Doppler pair for an MPC due to an SD scatterer can therefore appear also outside the U-shaped region in Fig. 4. However, since the SD scatterers are static, the Doppler parameter is in the interval \([-\nu_S, \nu_S]\), i.e., the same interval as for the diffuse scatterers.

**Mobile Discrete Scatterers:** We assume that no vehicle travels with an absolute speed exceeding \( v_{\text{max}} \). It then follows from Eq. (24) that the Doppler due to a mobile discrete (MD) scatterer is in the interval \([-v_{\text{max}}, v_{\text{max}}]\), where \( v_{\text{max}} = \frac{4\nu_T}{\nu_S} \). For example, in Fig. 4, the Doppler shift \( \nu_{P} \) is due to an MD scatterer (vehicle) that travels in the oncoming lane (\( v_P < 0 \)).

Based on the analysis above, we can conclude that the delay-Doppler parameters for the multipath components can be divided into three regions:

\[
R_1 \triangleq \{ (\tau, \nu) \in \mathbb{R}^2 : \tau \in (\tau_0, \tau_0 + \Delta \tau), \nu \in (-\nu_S, \nu_S) \} \\
R_2 \triangleq \{ (\tau, \nu) \in \mathbb{R}^2 : \tau \in (\tau_0 + \Delta \tau, \tau_{\text{max}}), |\nu| \in [\nu_S - \Delta \nu, \nu_S) \} \\
R_3 \triangleq \{ (\tau, \nu) \in \mathbb{R}^2 : \tau \in (\tau_0, \tau_{\text{max}}), |\nu| \leq \nu_{\text{max}} \} \setminus (R_1 \cup R_2),
\]

where \( \tau_{\text{max}} - \tau_0 \) is the maximum significant excess delay for the V2V channel. Here, \( \Delta \tau \) and \( \Delta \nu \) are chosen such that the contributions from all diffuse scatterers are confined to \( R_1 \cup R_2 \). The exact choice of \( \Delta \tau \) is somewhat arbitrary. In this paper, we consider a thresholding rule to compare the noise level and channel components, which results in a particular choice of \( \Delta \tau \); the method is specified in Appendix C. However, regardless of method, once \( \Delta \tau \) is chosen, we can compute the height \( \Delta \nu \) of the two strips that make up \( R_2 \). This can be done by placing an ellipsoid with its foci at the transmitter and receiver such that the path from the transmitter to the receiver via any point on the ellipsoid has propagation delay \( \tau_0 + \Delta \tau \). By computing the associated Doppler along the part of ellipsoid that is in the diffuse region (i.e., in the strips just outside the highway, see Fig. 2), we can determine the smallest absolute Doppler value among them as \( \nu' \) and calculate \( \Delta \nu \) as \( \Delta \nu = \nu_S - \nu' \). In Appendix C, we present a data driven approach to approximate \( \Delta \nu \). Note that the regions gather channel components with similar behavior. Region \( R_1 \), contains the LOS, ground reflections, and (strong) discrete and diffuse components due scatterers near the transmitter and receiver. In Region \( R_2 \), the delay-Doppler contribution of static discrete and diffuse scatterers from farther locations appear.

Region \( R_3 \) contains contributions from moving discrete and static discrete scatterers only.

**Remark 4.** In Fig. 4, we see that there are sparse contributions from the SD and MD scatterers in all regions \( R_1, R_2, \text{ and } R_3 \). However, clusters of DI components are confined to \( R_1 \cup R_2 \). Therefore, V2V channel components can be divided into two main clusters. One is the element-wise sparse components (mobile and static discrete scatterers) that are distributed in all three regions \( R_1, R_2, \text{ and } R_3 \), and the other one is the group-wise sparse components (diffuse components) that are located in regions \( R_1 \) and \( R_2 \). We note that, in V2V channels due to the geometry of the channel and antenna heights, there exists more diffuse components compared to the cellular communication. Therefore, a proper V2V channel estimation algorithm should consider the estimation of diffuse components with higher quality compared to the cellular communication channel. Since the diffuse components are located in a specific part of delay-Doppler domain and the rest of the channel support in the delay-Doppler is essentially zero, we partition the channel into group vectors and take advantage of group-wise sparsity of the channel to enhance the accuracy of the estimate of the diffuse components. In Sec. VI, we propose a method based on joint element-wise and group-wise sparsity and Theorem 1 of Section II-B to estimate the discrete delay-Doppler representation of V2V channel exploiting this structure.
V. Observation Model and Leakage Effect

In this section, we show how pulse shaping and a finite-length training sequence can be taken into account when formulating a linear observation model of the V2V channel. We assume that \( p_t(t) \) and \( p_r(t) \) are causal with support \([0, T_{\sup}]. \) The contribution to the received signal from one of the \( 1 + N_{MD} + N_{SD} + N_{DJ} \) terms in (22) is of the form
\[
\begin{align*}
    s(t) * h_i(t, \tau) &= \sum_{l=-\infty}^{\infty} s[l] \eta e^{j 2 \pi \nu_i (lT_s - \tau_i)} p_i(t - lT_s - \tau_i) \\
    &\approx \sum_{l=-\infty}^{\infty} s[l] \eta e^{j 2 \pi \nu_i (lT_s + \tau_i)} p_i(t - lT_s - \tau_i),
\end{align*}
\]
where the approximation is valid if we make the (reasonable) assumption that \( \nu_i T_{\sup} \ll 1, \) and * denotes convolution. If we let \( p(t) = p_t(t) * p_r(t), \) we can write the contribution after filtering and sampling as
\[
    y_i[n] = s(t) * h_i(t, \tau) * p_r(t) \big|_{t = nT} = \sum_{l=-\infty}^{\infty} s[l] \eta e^{j 2 \pi \nu_i (lT_s + \tau_i)} p((n - l)T_s - \tau_i) = \sum_{m=-\infty}^{\infty} s[n - m] \eta e^{j 2 \pi \nu_i ((n - m)T_s + \tau_i)} p(mT_s - \tau_i),
\]
and identify \( h_i[n, m] \) as the DFT of a discrete-time complex exponential with frequency \( \nu_i \). The leakage with respect to delay decreases with frequency \( \nu_i \).

The channel without leakage is given by
\[
    H[k, m] = \sum_i \eta_i \delta[k - k_i] \delta[m - m_i],
\]
where \( \delta[n] \) is the Kronecker delta function. The channels in (32) and (33) can be represented for \( k \in K \) and \( m \in M \) by the vectors \( \mathbf{x}_i \in \mathbb{C}^N \) and \( \mathbf{x} \in \mathbb{C}^N, \) respectively, where \( N = |K||M| = (2K + 1)M, \)
\[
    \mathbf{x}_l = \text{vec} \begin{bmatrix} H_l[-K, 0] & \cdots & H_l[-K, M - 1] \\ \vdots & \ddots & \vdots \\ H_l[K, 0] & \cdots & H_l[K, M - 1] \end{bmatrix},
\]
\[
    \mathbf{x} = \text{vec} \begin{bmatrix} H[0, 0] & \cdots & H[K, M - 1] \\ \vdots & \ddots & \vdots \\ H[K, 0] & \cdots & H[K, M - 1] \end{bmatrix}.
\]
where \( \text{vec}(\mathbf{H}) \) is the vector formed by stacking the columns of \( \mathbf{H} \) on the top of each other. The relationship between \( \mathbf{x}_l \) and \( \mathbf{x} \) can be written as
\[
    \mathbf{x}_l = \mathbf{G} \mathbf{x},
\]
where \( \mathbf{G} \in \mathbb{C}^{N \times N} \) is defined as
\[
    \mathbf{G} = \begin{bmatrix} \text{vec}(G_0) & \text{vec}(G_1) & \cdots & \text{vec}(G_{N-1}) \end{bmatrix},
\]
\[
    G_j = \begin{bmatrix} g[-K, 0, k', m'] & \cdots & g[-K, M - 1, k', m'] \\ \vdots & \ddots & \vdots \\ g[K, 0, k', m'] & \cdots & g[K, M - 1, k', m'] \end{bmatrix},
\]
where the one-to-one correspondence between \( j = 0, 1, \ldots, (2K + 1)M - 1 \) and \( (k', m') \in K \times M \) is given by \( j = m' + (2K + 1)k'. \)

The structure of \( \mathbf{G} \) is a direct consequence of how we vectorize \( H_l[k, m] \) in (35). If we consider an alternative way of vectorizing \( H_l[k, m] \), then the leakage matrix \( \mathbf{G} \) needs to be recomputed accordingly, by appropriate permutation of the columns and rows of leakage matrix defined in (37). As \( K, M, \) and the pulse shape are known, \( \mathbf{G} \) is completely determined in (36). Thus, we can utilize the relationship in (36) to compensate for leakage.

Consider that the source vehicle transmits a sequence of \( N_r + M - 1 \) pilots, \( s[n], \) for \( n = -(M - 1), -(M - 2), \ldots, N_r - 1, \) over the channel. We collect the \( N_r \) received samples in a column vector
\[
    \mathbf{y} = [y[0], y[1], \ldots, y[N_r - 1]]^\top.
\]
Using (20), we have the following matrix representation:
\[
    \mathbf{y} = \mathbf{S} \mathbf{x}_l + \mathbf{z},
\]
where $z \sim \mathcal{CN}(0, \sigma_z^2 I_{N_r})$ is a Gaussian noise vector, and $S$ is a $N_x \times N$ block data matrix of the form

$$S = [S_0, \ldots, S_{M-1}],$$

where each block $S_m \in \mathbb{C}^{N_x \times (2K+1)}$ is of the form

$$S_m = \text{diag} \{ s[-m], \ldots, s[N_r - m - 1] \} \Omega,$$

for $m = 0, 1, \ldots, M-1$, and $\Omega \in \mathbb{C}^{N_x \times (2K+1)}$ is a Vandermonde matrix, $\Omega[i, j] = \omega_{2K+1}^{(i-j)K}$, where $i = 0, 1, \ldots, N_r - 1$ and $j = 0, 1, \ldots, 2K$. Finally, by combining (39) and (36) we have

$$y = Sx + z = Ax + z,$$

where $A = SG$ and $A \in \mathbb{C}^{N_r \times N}$.

VI. CHANNEL ESTIMATION

Based on our analysis in Section IV, the components in the vector $x$ have both element- and group-wise sparsity. Given estimates of the parameters that define regions $R_1$, $R_2$, and $R_3$ (see Section IV and Appendix C), we illustrate how to partition the elements in the channel vector $x$ to enforce the group-wise sparsity structure. Our partitioning is based on the sparsity structure of regions $R_1$, $R_2$, and $R_3$, and Eq. (35), which maps the entries of $x$ into channel components, $H[k, m]$ for $k = -K, \ldots, K$ and $m = 0, \ldots, M-1$. To partition the elements in regions $R_1$ and $R_2$, we collect channel components with a common Doppler value into a single group, e.g., $I_1, I_2, \ldots, I_{N_1}$ (blue segments), and $I_{N_1+1}, \ldots, I_{N_1+N_2}$ (red segments) as depicted in Fig. 5. For $R_3$, we know that this region contains only the element-wise sparse components, thus we consider each element of $x$ in this region as a single partition, e.g., $I_{N_1+1}, I_{N_1+N_2+1}, \ldots, I_{N_r}$, where $N_g = N_{N_1,1} + N_{N_1,2} + N_{N_1,3}$ as depicted in Fig. 5. Now that we have a partition of all the elements in $x$, we can easily determine the group vectors $x_i$ as follows:

$$x_i[k] = \begin{cases} x[k] & k \in I_i \\ 0 & k \notin I_i \end{cases} \text{ for } i = 1, \ldots, N_g$$

We know that most of components in $R_3$ are zero (or close to zero) and there is a significant numbers of non-zero diffuse components in region $R_1$ and $R_2$. Therefore, if we enforce the group sparsity regularizer over this partitioning of elements in $x$, it will improve the quality of estimation of diffuse components.

We next specify the regularizations to exploit the jointly sparse structure of the V2V channel as follows

$$\hat{x} = \arg\min_{x \in \mathbb{C}^N} \left\{ \frac{1}{2} \| y - Ax \|_2^2 + \phi_g(|x|; \lambda_g) + \phi_c(|x|; \lambda_c) \right\}, \quad (P_0)$$

where $|x| = \|x[1], \ldots, x[N]\|^T$ with $N = M(2K + 1)$ and $x[i] = \sqrt{\text{Re}(x[i])^2 + \text{Im}(x[i])^2}$. Here the regularization functions are

$$\phi_g(|x|; \lambda_g) = \sum_{j=1}^{N} f_g(\|x_j\|; \lambda_g),$$

$$\phi_c(|x|; \lambda_c) = \sum_{i=1}^{N} f_c(|x[i]|; \lambda_c).$$

We develop a proximal ADMM algorithm to solve problem $P_0$. Problem $P_0$ can be rewritten using an auxiliary variable $w$ as follows,

$$\min_{x, w \in \mathbb{C}^N} \frac{1}{2} \| y - Ax \|_2^2 + \phi_g(|w|; \lambda_g) + \phi_c(|w|; \lambda_c)$$

s.t. $w = x$ \hspace{1cm} (46)

For the optimization problem in (46), ADMM consists of the following iterations,

- **Initialize:** $\rho \neq 0, \lambda_{pg} = \frac{\lambda_x}{\rho}, \lambda_{pe} = \frac{\lambda_x}{\rho}, \theta^0 = w^0 = 0, A_0 = (\rho^2 I + A^H A)^{-1}, \text{ and } x_0 = A_0 A^H y$.
- **Update-$x$:**

$$x^{n+1} = \rho^2 A_0 \left( w^n - \theta^n \right) + x_0,$$

where $\theta^n = \frac{\theta^n}{\rho^2}$.
- **Update-$w$:** for $i = 1, 2, \ldots, N_g$,

$$w_i^{n+1} = \arg\min_{w_i} \frac{1}{2} \|x_i^{n+1} + \theta_{pi} - w_i\|_2^2 + g(|w_i|; \lambda_{pg}) + E(|w_i|; \lambda_{pe}),$$

where the index $i$ denotes the group number, and

$$E(|w_i|; \lambda_{pe}) = \sum_f f_c \left( \left| \frac{|w_i[1]|}{\rho} \right|; \lambda_{pe} \right),$$

$$g(|w_i|; \lambda_{pg}) = f_g \left( \left| \frac{|w_i[1]|}{\rho} \right|; \lambda_{pg} \right).$$

- **Update-dual variable-$\theta$:**

$$\theta^{n+1} = \theta^n + (x^{n+1} - w^n)^+.$$}

Details of this derivation are provided in Appendix D. Note that both $A_0$ and $x_0$ are known and can be computed in advance. In the update-$w$ step, the index $i$ denotes the group number, thus, this step can be done in parallel for all groups, simultaneously.

Since the vectors in optimization problem for updating $w$ in (48) are complex vectors, Theorem 1 in Section II-B, cannot directly be applied to find a closed-form solution for this optimization problem. However, in order to apply Theorem 1, we introduce the following notation and lemma. The vector $w_i \in \mathbb{C}^n$ can be written as $w = |w| \oplus \text{Phase}(w)$, where the $n$th element of Phase($w$) is $\exp(j \text{Ang}(w[n]))$, and $\text{Ang}(w[n])$ is the angle of $w[n]$ in polar form, i.e., $w[n] = |w[n]| \exp(j \text{Ang}(w[n]))$, and $\oplus$ is component-wise multiplication (Schur product).
TABLE I: Proposed V2V Channel Estimation Method

| Input: y, A, \(\lambda_g\), \(\lambda_e\), \(\rho\), \(n_{\text{max}}\), \(\epsilon\). |
| Initialize: \(w^0 = \theta^0_p = 0\). |
| Pre-computation: \(A_0 = (\rho^2 I + A^H A)^{-1}\), \(x^0 = A_0 A y\). |
| For \(n = 0\) to \(n_{\text{max}} - 1\) |
| \(w^{n+1}_i = P_{\lambda_{pg}} \left( P_{\lambda_e, E} \left( |x^{n+1}_i + \theta^{n}_{\rho_i}| \right) \right) \odot \text{Phase} \left( x^{n+1}_i + \theta^{n}_{\rho_i} \right) \) |
| for \(i = 1, 2, \ldots, N_g\) |
| \(\theta^{n+1}_i = \theta^n_i + (x^{n+1}_i - w^{n+1}_i)\) if \(\|x^{n+1}_i - w^{n+1}_i\|_2 < \epsilon\) then break |
| End |
| Output: Vector \(x\). |

Lemma 1. For any \(c \in \mathbb{C}^N\)

\[
\arg\min_{w \in \mathbb{C}^N} \|c - w\|_2^2 = \text{Phase}(c) \odot \arg\min_{|w| \in \mathbb{R}^N} \|c - |w|\|_2^2. \tag{52}
\]

The proof is provided in Appendix E. Since the two last terms in (48) are independent of the phase of \(w_i\), we use Lemma 1 to write the \(i\)th group problem in (48) as

\[
w^{n+1}_i = \text{Phase} \left( x^{n+1}_i + \theta^n_{\rho_i} \right) \odot \left( \arg\min_{|w_i|} \frac{1}{2} \|x^{n+1}_i + \theta^n_{\rho_i} - |w_i|\|_2^2 + g(|w_i|; \lambda_{pg}) + E(|w_i|; \lambda_{pe}) \right) \tag{53}
\]

Now, the vectors in the optimization problem in (53) involve only real vectors, therefore, the solution for this optimization problem can be directly computed using Theorem 1. We determine a closed form solution for the update-\(w\) step in Corollary 1, below, using the proximity operators of the univariate functions \(f_e\) and \(f_g\). This update rule is a direct consequence of Theorem 1.

Corollary 1. The second step, update-\(w\), can be performed as follows

\[
w^{n+1}_i = P_{\lambda_{pg}} \left( P_{\lambda_e, E} \left( |x^{n+1}_i + \theta^n_{\rho_i}| \right) \right) \odot \text{Phase} \left( x^{n+1}_i + \theta^n_{\rho_i} \right), \tag{54}
\]

where \(E\) and \(g\) are defined in Equations (49) and (50).

Based on Corollary 1, the update-\(w\) step only depends on the proximity operators of the regularizer functions \(f_e\) and \(f_g\). The proposed algorithm to estimate the channel vector \(x\) from the received data vector \(y\) is summarized in Table I. The complexity of our joint sparse signal estimation algorithm is as follows:

Step 1) update \(x\), incurs a computational complexity of \(O(N^2)\), where \(N = M(2K + 1)\), due to a matrix/vector multiplication. Step 2) updating \(w\), requires \(N_g\) group-wise threshold comparisons and \(N_{\text{element-wise threshold comparisons}}\). Thus, it has complexity of \(O(N + N_g) \approx O(N)\). Updating the dual variable in Step 3) has \(O(N)\) complexity.

Therefore, the overall complexity of our proposed algorithm is \(O(N^2)\). In our algorithm (similar to the other methods), we compute a Least-Squares (LS) solution for the initialization. Computation of a least squares (LS) solution can be implemented with complexity \(O(N^3)\). Algorithms such as the Wiener filter, the Hybrid Sparse/Diffuse (HSD) estimator that are designed based on statistical knowledge of channel parameters require a large number of samples to estimate the required covariance matrices. If the correlation matrices are known, then the Wiener filter and the Hybrid Sparse/Diffuse (HSD) estimator have computational complexity in order of \(O(N^3)\).

VII. NUMERICAL RESULTS

In this section, we demonstrate the performance gains that can be achieved with our proposed, structured, estimation using both convex and non-convex sparsity-inducing regularizers, in comparison to prior methods such as Wiener filtering [6], the Hybrid Sparse/Diffuse (HSD) estimator [8–10], and the compressed sensing (CS) method [11, 12].

To simulate the channel, we consider a geometry with length of 1 km around the transmitter-receiver pair, road width \(D = 50\) m, and the width of the diffuse strip around the road \(d = 25\) m, as Fig. 2. The locations of the transmitter and receiver are chosen in this geometry with distance 100 m to 200 m from each other. The speeds of the transmit and receive vehicles are chosen randomly from the interval \([60, 160]\) (km/h), the speed limits for a highway. It is assumed that the number of MD scatterers \(N_{MD} = 10\), and their speeds are also chosen randomly from the interval \([60, 160]\) (km/h); we have \(N_{SD} = 10\) and \(N_{DI} = 400\) SD and DI scatterers, respectively [7]. Using these parameters, we compute the delay and Doppler values for each scatterer. The statistical parameter values for different scatterers are selected as in Table I of [7], which are determined from experimental measurements. The scatterer amplitudes were randomly drawn from zero mean, complex Gaussian distributions with three different power delay profiles for the LOS and mobile discrete (MD) scatterers, static discrete scatterers, and diffuse scatterers. We assume that the mean power of the static scatterers is 10 dB less than the mean power of the LOS and MD scatterers, and the mean power of the diffuse scatterers also is 20 dB less than the mean power of the LOS and MD scatterers [7]. Furthermore, we consider \(f_c = 5.8\) GHz, \(T_s = 10\) ns, \(N_r = 1024\), \(K = 512\), and \(M = 256\). The pilot samples are drawn from a zero-mean, unit variance Gaussian distribution. The interpolation/anti-aliasing filters \(p_r(t) = p_c(t)\) are root-raised-cosine filters with roll-off factor 0.25 and \(T_{\text{supp}} = 1\) ms. The required regularization parameters were found by trial and error using a cross-validation algorithm on the data [29]. Performance is measured by the normalized mean square error (NMSE), normalized by the mean energy of the channel coefficients. The NMSE is defined as

\[
\text{SNR} = \frac{E(\|\mathbf{x} - \hat{x}\|_2^2)}{E(\|\mathbf{x}\|_2^2)}. \tag{50}
\]

Fig. 6 depicts the MSE using our proposed estimator, our previously proposed hybrid sparse/ diffuse (HSD) estimator as adapted to V2V channels [10], a compressed sensing (CS) method [11, 12], and the Wiener based estimator of [6]. We have also included a curve (known support) corresponding to the case that the support (location of non-zero components) of vector \(x\) is known when we apply our joint sparse esti-
Our algorithm is the case considered in [19] and [20] for estimator with a soft-thresholding regularizer, i.e., the algorithm [8].

In Fig. 6, we can conclude that since the channel components in V2V channels (sparse and diffuse components) have different levels of energy, proximity operators such as SCAD and MCP with a multi-threshold nature of their proximal operators (see Fig. 1 and Eqs. (14) and (15)) are (more) suited to the channel structure.

In Fig. 7, we consider the effect of leakage compensation (Section V) on the performance of sparse estimator of V2V channels, such as the HSD estimator and our proposed joint sparsity estimator using SCAD regularizer function for group sparsity. From the results in Fig. 7, we observe that the uncompensated leakage effect reduces performance severely, more than 7 dB, particularly at higher SNR, due to the channel mismatch introduced by the channel leakage.

Next, we assess the performance of our proposed V2V channel estimation algorithm for different values of $N_{SD}$, $N_{DI}$, and $N_{MD}$ in the channel. In Fig. 8(a), the performance of our algorithm for different choices of $N_{SD} + N_{MD}$ is depicted. Note that the static (SD) and mobile discrete (MD) components have similar effects on the channel sparsity pattern. Therefore, we consider the value of $N_{SD} + N_{MD}$ for our simulations. In Fig. 8(b), the performance of our algorithm for different choices of $N_{DI}$ is depicted.
Fig. 8: Performance of algorithm for different values of $N_{DI}$ and $N_{SD} + N_{MD}$.

(a) NMSE for varying $N_{SD}$ and $N_{MD}$ and fixed $N_{DI} = 500$

(b) NMSE for varying $N_{DI}$ and fixed $N_{SD} + N_{MD} = 50$

Fig. 9: Performance of algorithm for different values of $\Delta \tau$ and $\Delta \nu$.

(a) Different choice of $\Delta \tau$

(b) Different choice of $\Delta \nu$

Fig. 10: Performance of the proposed algorithm under different values of $K = \gamma \lfloor (N_r - 1)/2 \rfloor$.

(a) Different choice of $N_r$ and $M = 256$

(b) Different choice of $M$ and $N_r = 256$

Fig. 11: Performance of the proposed algorithm under different values of $N_r$ and $M$.

In numerical simulation, we consider $K = 1024$, $M = 512$, and $N_r = 2048$, and SNR=10 dB. It can be seen that by decreasing the number of channel components, the performance of our method improves.

Furthermore, we consider the performance of our proposed method under different values of $\Delta \tau$, and $\Delta \nu$. Our numerical results in Figures 9.(a) and 9.(b) show that by increasing both $\Delta \tau$ and $\Delta \nu$ the performance of our algorithm degrades. This is due to the fact that increasing $\Delta \tau$ and $\Delta \nu$ reduces the sparsity of the channel, which results in less noise reduction. Here $N_r = 512$, $M = 256$, $N_{SD} + N_{MD} = 20$ and $N_{DI} = 200$ are considered.

The value of parameter $K$ is lower bounded by $N_r$, namely the total number of measurements, in the sense that $2K + 1 \geq N_r$, as given in Eq. (21). To decrease the computational complexity, we consider $K = \lfloor (N_r - 1)/2 \rfloor$, which provides a good performance results as seen in Fig. 10.

However, we can consider $K \geq \lfloor (N_r - 1)/2 \rfloor$ by zero padding the DFT transform in Eq. (21). The larger the $K$, the better the leakage is compensated, but the higher the signal dimension of $x$. The first will improve the signal recovery, but the latter degrades the signal reconstruction as more unknown variables are introduced. To understand the effect of increasing $K$, we consider $K = \gamma \lfloor (N_r - 1)/2 \rfloor$, where $1 \leq \gamma$. Here for computational efficiency, we consider $\gamma = 2^n$, where $n$ is a nonnegative integer. Furthermore, $N_r = 512$, $M = 256$, $N_{SD} + N_{MD} = 20$ and $N_{DI} = 200$ are considered. Results in Fig.10 show that by increasing $\gamma$, there is an improvement in the performance of the algorithm for $n \leq 4$. But after $n \geq 5$ (increasing the signal dimension) increasing $K$ increases the NMSE.

In Figures 11.(a) and 11.(b), we have considered the effect of different choices of $M$ and $N_r$ on the NMSE of the proposed V2V channel estimation algorithm. Note that the values of $N_r$ and $M$ are dictated by the sampling time, training signal length and channel delay spread. Therefore, we cannot increase the value of these parameters arbitrarily. Figures 11.(a) and 11.(b) indicate that increasing the number of measurements, $N_r$, and $M$, reduces the NMSE, as expected.

VIII. REAL CHANNEL MEASUREMENTS

In this section, we use experimental channel data, recorded by WCAE measurement systems in a highway environment, to model the V2V channel and also assess the performance of our proposed channel estimation algorithm. The channel measurement data was collected using the RUSK-Lund channel sounder in Lund and Malmö city in Sweden. The complex $4 \times 4$ multi-input-muti-output (MIMO) channel transfer functions were recorded at a 5.6 GHz carrier frequency over a bandwidth of 240 MHz in several different propagation environments with two standard Volvo V70 cars used as the TX and RX cars during the measurements [30].

A. Measurement setup

V2V channel measurements were performed with the RUSK-LUND sounder using a multi-carrier signal with carrier frequency 5.6 GHz to sound the channel and record the time-variant complex channel transfer function $H(t, f)$. The measurement bandwidth was 240 MHz, and a test signal length of $3.2 \mu s$ was used. The time-varying channel was sampled every 0.307 ms, corresponding to a sampling frequency of 3255 Hz during a time window of roughly 31.2 ms. The sampling frequency implies a maximum resolvable Doppler shift of 1.5 kHz, which corresponds to a relative speed of about 350 km/h at 5.6 GHz.

\textsuperscript{3}Wireless Communication in Automotive Environment
By Inverse Discrete Fourier Transforming (IDFT) the recorded frequency responses \( H(t, f) \), with a Hanning window to suppress side lobes, the complex channel impulse responses \( h(t, \tau) \) are obtained. Finally, by taking Discrete Fourier Transform (DFT) of \( h(t, \tau) \) with respect to \( t \), the channel scattering function in the delay-Doppler domain, \( H[k, m] \), is computed.

In this experiment, \( K = 116 \) and \( M = 256 \) is considered. Three recorded channel scattering functions, \( H[k, m] \), are plotted in Figures 12 and 13. Fig. 12 presents a V2V channel in the delay-Doppler domain. A discrete component is visible at approximately 0.65 \( \mu \)s propagation delay. Also plotted in the figure is the Doppler shift vs. distance as produced by Equation (24) and (25), i.e., for scatterers located on a line parallel to (and a distance \( 5 \) m away from) the TX/RX direction of motion. We notice that V2V channel scattering (in the delay-Doppler domain) in Figures 12 and 13 (a) and (b) are highly structured as predicted in Section IV. As seen in these figures, the diffuse components are confined in a U-shaped area that was also predicted by our analysis in Section IV. Furthermore, we can observe the sparse structure of the discrete components in all the figures.

In Fig. 14, we compare the performance of our proposed nested estimators with the CS method [11], [12] to estimate the channel given in Fig. 13 (a). The training pilot samples are generated as discussed in Section VII. To vary the SNR, we add additive white Gaussian noise to the signal at the output of channel. Note that bandwidth and observation time for the channel measurements are large enough to allow us to ignore the leakage effect. The results in Fig. 14 confirm our numerical analysis in Section VII and show that our proposed joint sparse estimation algorithm has a better performance compared to the (only) element-wise sparse estimator methods [11], [12].

In Fig. 15, we investigate the effect of specifying regions on the channel estimation algorithm. We consider the channel given in Fig. 13 (b) for this experiment. We have considered three different region scenarios. In the first scenario, we determine the regions as computed by our heuristic method given in Appendix C. In the second scenario, we keep \( \Delta \tau \) the same as the first scenario, but we set \( \Delta \nu = \nu_{\text{max}} / 2 \), which means that we neglect the structure of the diffuse components in regions \( R_2 \) and we assume that the diffuse components can occur in the entire delay-Doppler domain. Finally, in the third scenario, we set \( \Delta \tau = \tau_{\text{max}}, \text{i.e.,} R_1 \) extends to cover the entire delay-Doppler domain. Results in Fig. 15 indicate that considering the structural information of the V2V channel in the delay-Doppler (three regions) significantly improves the performance of our joint sparse estimation algorithm.

**IX. Conclusions**

We provide a comprehensive analysis of V2V channels in the delay-Doppler domain using the well-known geometry-based stochastic channel modeling. Our characterization reveals that the V2V channel model has three key regions, and these regions exhibit different sparse/hybrid structures which can be exploited to improve channel estimation. Using this structure, we have proposed a joint element- and group-wise sparse approximation method using general regularization functions. We prove that for the needed optimization, the optimal solution results in a nested estimation of the channel vector based on the group and element wise penalty functions.

Our proposed method exploits proximity operators and the alternating direction method of multipliers, resulting in a modest complexity approach with excellent performance. We characterized the leakage effect on the sparsity of the channel and robustified the channel estimator by explicitly compensating for pulse shape leakage at the receiver using the leakage matrix. Simulation results reveal that exploiting the joint
Lemma 2. Consider that \( g(x; \lambda) = f \left( \frac{|x|}{\lambda} \right); \) where \( f(x; \lambda) \) is a non-decreasing function of \( x \). Furthermore, \( f(x; \lambda) \) is a homogeneous function, i.e., \( f(\alpha x; \alpha \lambda) = \alpha^2 f(x; \lambda) \). Then,

\[ \begin{align*}
& i) \ P_{\lambda,g}(a) = \gamma a, \text{ where } \gamma \in [0,1]. \\
& ii) \ \gamma = \begin{cases} \\
\frac{1}{n \|a\|_2} \sum_{i \in S} f \left( \frac{|a_i|}{\|a\|_2} \right) & \text{if } \|a\|_2 > 0 \\
0 & \text{if } \|a\|_2 = 0 \\
\end{cases}
\end{align*} \]

**Proof:** \( i) \) For every \( z \), we can write \( a = a^+ + \gamma a \), where \( a^+ \perp a \) and \( \gamma \in \mathbb{R} \). Therefore, based on the proximity operator definition, we have,

\[ P_{\lambda,g}(a) = \arg \min_{\gamma} \left\{ \frac{1}{2} \|a - z\|_2^2 + g(\gamma a; \lambda) \right\} \]

\[ = \arg \min_{\gamma} \left\{ \frac{1}{2} \|a - a^+ - \gamma a\|_2^2 + g(\gamma a + a^+; \lambda) \right\} \]

Since \( \|a + a^+\|_2 \geq \max\{\|a\|_2, \|a^+\|_2\} \) and \( f \) is a non-decreasing function, we have \( g(\gamma a + a^+; \lambda) \geq g(\gamma a; \lambda) \). Therefore, \( a^+ = 0 \) in the optimization problem (59) and we can rewrite it as,

\[ P_{\lambda,g}(a) = \arg \min_{\gamma} \left( \frac{1}{2} (\gamma - 1)^2 \|a\|_2^2 + g(\gamma a; \lambda) \right). \]

The two terms in the objective function in (60) are increasing when \( \gamma \) increases from \( \gamma = 1 \) or decreases from \( \gamma = 0 \). Hence, the minimizer lies in the interval \([0, 1]\). Therefore, we have \( P_{\lambda,g}(a) = \gamma a \), where \( \gamma \in [0, 1] \).

\( ii) \) Let \( \tau = \gamma \|a\|_2 \), then the optimization problem in (60) for \( \|a\|_2 > 0 \), can be written as,

\[ P_{\lambda,g}(a) = \frac{a}{\|a\|_2} \arg \min_{\tau \in \mathbb{R}} \left\{ \frac{1}{2} \left( \frac{\|a\|_2 - \tau}{\|a\|_2} \right)^2 + f \left( \frac{\tau}{\rho} \right) \right\} \]

\[ = \frac{a}{\|a\|_2} \arg \min_{\tau \in \mathbb{R}} \left\{ \frac{1}{2} \left( \frac{\|a\|_2 - \tau}{\|a\|_2} \right)^2 + \frac{1}{\rho^2} f(\tau; \rho) \right\} \]

\[ = \frac{\rho}{\|a\|_2} P_{\lambda,g}(a). \]

Equality (a) is due to the homogeneity of function \( f \), i.e., \( f \left( \frac{\tau}{\rho} \right) = f \left( \frac{\tau \lambda}{\rho} \right) = \frac{1}{\rho} f(\tau; \rho) \). Equality (b) is due to the definition of the proximal operator. Thus, \( \gamma = \frac{1}{\|a\|_2} \sum_{i \in S} f \left( \frac{|a_i|}{\|a\|_2} \right) + \rho \), and the proof of the Lemma is completed.

**Lemma 3.** If the function \( f(x; \lambda) \) is homogenous i.e., \( f(\alpha x; \alpha \lambda) = \alpha^2 f(x; \lambda) \) for all \( \alpha > 0 \), then \( P_{\alpha f}(ab) = \alpha P_{f}(b) \) for all \( b \in \mathbb{R} \) and \( \lambda > 0 \).

**Proof:** By definition of the proximity operator, we have \( P_{\alpha f}(ab) = \arg \min_{\gamma} \left\{ \frac{1}{2} \|ab - x\|_2^2 + f(x; \lambda) \right\} \). Consider \( x = \alpha z \), and using the homogenous properties of \( f \), we have \( P_{\alpha f}(ab) = \alpha \arg \min_{\zeta} \left\{ \frac{a^2}{2} (b - z)^2 + \alpha^2 f(z; \lambda) \right\} = \alpha \arg \min_{\zeta} \left\{ \frac{1}{2} (b - z)^2 + f(z; \lambda) \right\} = \alpha P_{f}(b) \).
where \( g(a; \lambda_{pg}) = f_g \left( \frac{\|a\|_2}{\rho}; \lambda_{pg} \right) \) and \( E(a; \lambda_{pe}) = \sum_j f_e \left( \frac{u_j}{\rho}; \lambda_{pe} \right) \). Here \( \lambda_{pg} = \frac{\alpha}{\rho} \) and \( \lambda_{pe} = \frac{\alpha}{\rho} \). Note that for \( p = 1 \), we have the claim in Theorem 1.

Assume \( v = P_{\lambda_{pe},E}(b) \) and \( u = P_{\lambda_{pg},g}(v) \).

Based on above definitions, to prove the claim of Theorem 1, we need to show that \( a = u \) is the minimizer of \( J(a) = \frac{1}{2} \|b-a\|_2^2 + g(a; \lambda_{pg}) + E(a; \lambda_{pe}) \). To prove this claim, we consider two cases: I: \( u \neq 0 \), and II: \( u = 0 \).

Case (I): \( u \neq 0 \). Since \( u = P_{\lambda_{pg},g}(v) \) and \( g(a; \lambda_{pg}) = f_g \left( \frac{\|a\|_2}{\rho}; \lambda_{pg} \right) \), and \( f_g \) is a homogenous non-decreasing function, by Lemma 2 we have \( u = \gamma v \), for some \( \gamma \in (0,1) \). Furthermore, \( u \) should satisfy the first order optimality condition for the objective function in \( u = \arg \min_a \left\{ \frac{1}{2} \|v-a\|_2^2 + g(a; \lambda_{pg}) \right\} \), namely

\[
0 \in u - v + \partial g(u; \lambda_{pg}).
\]

Using the definition of the proximity operator and Remark 1, we have \( \|P_{\lambda_{pe},E}(b)\|_i = P_{\lambda_{pe},E}(\|b_i\|/\rho) \). Since \( f_e \) is a homogenous function, using Lemma 3, we have \( P_{\lambda_{pe},f_e}(\|b_i\|/\rho) = \gamma P_{\lambda_{pe},f_e}(\|b_i\|/\rho) \) or equivalently,

\[
P_{\lambda_{pe},E}(\gamma b) = \arg \min_a \left\{ \frac{1}{2} \|g - a\|_2^2 + E(a; \lambda_{pe}) \right\} \quad (63)
\]

and by the first order optimality condition (of \( u \)) for the objective function in (63), we have \( 0 \in v - b + \partial E(u; \gamma \lambda_{pe}) \). Since \( \gamma \neq 0 \), Equation can be rewritten as

\[
0 \in v - b - \frac{1}{\gamma} \partial E(u; \gamma \lambda_{pe}).
\]

Since \( E(u; \lambda_{pe}) = \sum_j f_e \left( \frac{u_j}{\rho}; \lambda_{pe} \right) \), applying scale invariant property of function \( f_e \), i.e., \( f_e \left( \frac{u_j}{\rho}; \lambda_{pe} \right) = \gamma f_e \left( \frac{u_j}{\rho}; \lambda_{pe} \right) \), we have \( \frac{1}{\gamma} \partial E(u; \gamma \lambda_{pe}) = \partial E(u; \lambda_{pe}) \).

Therefore, we can rewrite (64) as

\[
0 \in v - b + \partial E(u; \lambda_{pe}).
\]

Summing Equations (62) and (65), we have \( 0 \in u - b + \partial g(u; \lambda_{pg}) + \partial E(u; \lambda_{pe}) \), which is the first order optimality condition for \( u \) for the objective function \( J(a) \). Case (II): \( u = 0 \). Here, we need to show the first-order optimality conditions for \( u = 0 \) for the objective function \( J(a) \), i.e., \( 0 \in \{ u - b + \partial g(u; \lambda_{pg}) + \partial E(u; \lambda_{pe}) \} |_{u=0} = 0 - b + \partial g(0; \lambda_{pg}) + \partial E(0; \lambda_{pe}) - b \). This is equivalent to showing the existence of 1. \( \chi_1 \in [-1,1]^N \), (equivalent to the term \( \partial E(0; \lambda_{pe}) \)), where \( \chi_2 \) with \( \|\chi_2\|_2 \leq \frac{1}{\gamma} \) (equivalent to the term \( \partial g(0; \lambda_{pg}) \)) such that \( b = \lambda_{pe} \chi_1 + \lambda_{pg} \chi_2 \), due to property (iii) in Assumption 1.

By definition of the proximity operator we have

\[
u = P_{\lambda_{pg},g}(v) = \arg \min_z \left\{ \frac{1}{2} \|v - z\|_2^2 + g(z; \lambda_{pg}) \right\} \quad (66)
\]

\[
u = \arg \min_z \left\{ \frac{1}{2} \|v - z\|_2^2 + f_g \left( \frac{\|z\|_2}{\rho}; \lambda_{pg} \right) \right\}.
\]

Fig. 16: Discrete representation of the regions \( R_1, R_2 \), and \( R_3 \). Using the first optimality condition of \( u = 0 \) for the objective function in (66), we have

\[
0 \in -v + \partial f_g(0; \lambda_{pg}) \frac{\partial (\|0\|)}{\rho}.
\]

Since \( \partial (\|0\|) = \{x \in \mathbb{R}^M, \|x\|_2 \leq 1 \} \) and \( \|z\|_2 \leq \lambda_{pg} \) for all \( z \in \partial f_g(0; \lambda_{pg}) \) (using property (ii) in Assumption 1), using Equation (67) we have \( \|v\|_2 \leq \frac{2 \alpha}{\gamma} \). Furthermore, since \( v = P_{\lambda_{pe},E}(b) \), the first-order optimality condition implies that \( 0 \in \partial E(v; \lambda_{pe}) + v - b \). Thus for \( \chi_1 \in \partial E(v; \lambda_{pe}) \) and \( \chi_2 = \lambda_{pg} \), we have \( b = \lambda_{pe} \chi_1 + \lambda_{pg} \chi_2 \) and proof is completed.

Appendix C

Region and Group Specifications

Here, we propose a heuristic method to find the regions \( R_1, R_2 \), and \( R_3 \), depicted in Fig. 16, and introduced in Section IV. To describe the regions, we need to compute the value of \( k_s, \Delta k, \) and \( \Delta m \), i.e., the discrete Doppler and delay parameters that corresponds to \( v_S, \Delta v, \) and \( \Delta T \), respectively. To estimate the \( \Delta m \) and \( \Delta k \), we use a regularized least-squares estimate of \( x \) given by \( x_{LS} = A_0 y = A_0 (A x + z) \approx x + e \) where \( e = A_0 z \) and \( A_0 = (A H A^T + \rho^2 I)^{-1} A H^T \) and \( \rho \) is a small real value. Based on relationship in the Eq. (35), we can write \( x_{LS} = \sum_{i} \{H_i \} \), where \( H_i \) is an estimate of the discrete delay-Doppler spreading function. Let us define the function \( E_d(m) = \sum_{i=1}^{\Delta m} \sum_{j=1}^{\Delta k} [H_{LS}[i,j]]^2 \) for \( 1 \leq m \leq M \). This function represent the energy profile of channel components in delay direction. Then

\[
\Delta m = \min_m \{ m | E_d(m) \leq T_1 \},
\]

where \( T_1 = \alpha_d E_d(0) \). Here \( 0 < \alpha_d < 1 \) is a tuning parameter. For a highway environment [7], based on our numerical analysis, \( \alpha_d \approx 0.4 \) is a reasonable choice.

After computing \( \Delta m \), to compute the values of \( \Delta k \) and \( k_s \) as labeled in Fig. 8, due to symmetry of channel diffuse components around zero Doppler value, we define functions \( E_{\nu}(k) = \sum_{i=1}^{\Delta m} |H_{LS}[i,k]|^2 + |H_{LS}[-k,i]|^2 \) for \( 0 \leq k \leq K \). This function represents the energy profile of channel components in Doppler direction. Let us define \( k_0 = \max_k E_{\nu}(k) \). Then, we can estimate \( \Delta k \) and \( k_s \) using following equations,

\[
k_s = \Delta k = \max_k \{ k | E_{\nu}(k) < T_2, 0 \leq k < k_0 \}, \quad (69)
\]

\[
k_s = \min_k \{ k | E_{\nu}(k) < T_2, k_0 < k \leq K \}, \quad (70)
\]

where \( T_\nu = \alpha_\nu E_{\nu}(k_0) \) with \( 0 < \alpha_\nu < 1 \). For highway environment, based on our numerical analysis, \( \alpha_\nu \approx 0.6 \) is a good choice.
APPENDIX D

PROXIMAL ADMM ITERATION DEVELOPMENT

For the optimization problem given in (46), we form the augmented Lagrangian
\[
L_\rho(x, w, \theta) = \frac{1}{2} \|y - Ax\|^2 + \phi_g(|w|; \lambda_g) + \phi_e(|w|; \lambda_e) + \langle \theta, x - w \rangle + \frac{\rho^2}{2} \|x - w\|^2,
\]
where \(\theta\) is the dual variable, \(\rho \neq 0\) is the augmented Lagrangian parameter, and \((a, b) = \text{Re}(b^H a)\). Thus, ADMM consists of the iterations:

- **update-x:** \(x^{n+1} = \text{argmin}_x \frac{1}{2} \|y - Ax\|^2\)
- **update-w:** \(w^{n+1} = \text{argmin}_w \phi_g(|w|; \lambda_g) + \phi_e(|w|; \lambda_e) + \langle \theta^n, x^n - w^n \rangle + \frac{\rho^2}{2} \|x^n - w^n\|^2\)
- **update-dual variable:** \(\theta^{n+1} = \theta^n + \rho \langle x^{n+1} - w^{n+1}, w^n \rangle\)

Deriving a closed form expressions for update-x is straightforward, \(x^{n+1} = \rho^2 A_0 (w^n - \theta^n) + x_0\), where \(\theta^n = \frac{\theta^n}{\rho^2}\), \(A_0 = (\rho^2 I + A^H A)^{-1}\) and \(x_0 = A_0 A^H y\). If we pull the linear terms into the quadratic ones in the objective function of update-w and ignoring additive terms, independent of \(w\), then we can express this step as

\[
w^{n+1} = \text{argmin}_w \left\{ \frac{1}{2} \|x^n + \theta^n - w^n\|^2 + \frac{1}{\rho^2} \phi_g(|w|; \lambda_g) + \phi_e(|w|; \lambda_e) \right\}
\]
\[
= \text{argmin}_w \sum_{i=1}^{N_i} \left\{ \frac{1}{2} \|x_i^n + \theta_i^n - w_i^n\|^2 + f_g \left( \frac{|w_i|_2}{\rho} ; \frac{\rho}{\lambda_g} \right) + \sum_j f_e \left( \frac{|w_i[j]|}{\rho} ; \frac{\rho}{\lambda_e} \right) \right\}
\]

where \(x_i, w_i\), and \(\theta_i^n\) are computed using the partitions introduced for the channel vector in Section VI. Thus, we can perform the update-w step in parallel for all groups,

\[
w_i^{n+1} = \text{argmin}_{w_i} \left\{ \frac{1}{2} \|x_i^n + \theta_i^n - w_i^n\|^2 + f_g \left( \frac{|w_i|_2}{\rho} ; \frac{\rho}{\lambda_g} \right) + \sum_j f_e \left( \frac{|w_i[j]|}{\rho} ; \frac{\rho}{\lambda_e} \right) \right\}
\]

Here, for simplicity in representation, we define \(\lambda_{pg} = \lambda_g / \rho\) and \(\lambda_{pe} = \lambda_e / \rho\). In addition, we define \(E(|w_i|; \lambda_{pg}) = \sum_j f_e \left( \frac{|w_i[j]|}{\rho} ; \frac{\rho}{\lambda_{pe}} \right)\), \(g(|w_i|; \lambda_{pg}) = f_g \left( \frac{|w_i|_2}{\rho} ; \frac{\rho}{\lambda_{pg}} \right)\). Thus, we have

\[
w_i^{n+1} = \text{argmin}_{w_i} \left\{ \frac{1}{2} \|x_i^n + \theta_i^n - w_i^n\|^2 + g(|w_i|; \lambda_{pg}) + E(|w_i|; \lambda_{pe}) \right\}
\]

To guarantee convergence to the optimal solution in (P0), the overall objective function, i.e., \(\frac{1}{2} \|y - Ax\|^2 + \phi_g(|x|; \lambda_g) + \phi_e(|x|; \lambda_e)\), should be a convex function [22]. Note that since the first term in the objective function, i.e., the quadratic penalty function, is convex, for any functions \(f_e\) and \(g\) that satisfies condition (iv) in Assumption I, the overall objective function is convex as well. Thus, ADMM yields convergence for all choices of the convex and non-convex functions given in Section II-B.

APPENDIX E

PROOF OF LEMMA 1

The function \(|c - w|_2^2 = ||c||^2_2 + ||w||_2^2 - 2 \text{Re}(c^H w)\) is minimized, with respect to phase of \(w\), when \(\text{Re}(c^H w)\) is maximized. Now,

\[
\text{Re}(c^H w) = \sum_{n=1}^{N} |c[n]| |w[n]| \cos(\text{Ang}(c[n]) - \text{Ang}(w[n]))
\]
\[
\leq \sum_{n=1}^{N} |c[n]| |w[n]| = |c| \text{Tr} |w|
\]

with equality if and only if \(\text{Phase}(w) = \text{Phase}(c)\), which in turn implies that \(|c - w|_2^2 = ||c||^2 - ||w||_2^2\). Hence,

\[
|c - w|_2^2 = \text{argmin}_{|w| \in \text{Phase}(w) \in C^N} \|c| - |w||_2^2
\]

and the lemma follows.

ACKNOWLEDGMENTS

We would like to thank Fredrik Tufvesson and Taimoor Abbas of Lund University for providing channel measurement data. The experimental work was funded by the Swedish Governmental Agency for Innovation Systems - VINNOVA, through the project Wireless Communication in Automotive Environment.

REFERENCES


Sajjad Beygi is currently working toward the Ph.D. degree in electrical engineering at the University of Southern California. His current research interests are statistical signal processing, sparse signal processing, and underwater wireless communications.

Since 2008, he has been a member of the Institute of Electrical and Electronics Engineers (IEEE) and has been serving as a reviewer of various journals in this society, including IEEE Transaction on Signal Processing, IEEE Signal Processing Letters, IEEE Transactions on Communication. He received his B.Sc. and M.Sc. degrees from Amirkabir University of Technology and Sharif University of Technology, Iran, in 2008 and 2010, respectively.

Urbashi Mitra received the B.S. and the M.S. degrees from the University of California at Berkeley and her Ph.D. from Princeton University. After a six year stint at the Ohio State University, she joined the Department of Electrical Engineering at the University of Southern California, Los Angeles, where she is currently a Professor. Dr. Mitra is a member of the IEEE Information Theory Society’s Board of Governors (2002-2007, 2012-2017) and the IEEE Signal Processing Society’s Technical Committee on Signal Processing for Communications and Networks (2012-2016). Dr. Mitra is a Fellow of the IEEE. She is the recipient of: 2015 Inspiring Women in STEM Award from Insight into Diversity magazine, 2012 Globecom Signal Processing for Communications Symposium Best Paper Award, 2012 NAE Lilian Gilbreth Lectureship, USC Center for Excellence in Research Fellowship (2010-2013), the 2009 DC0SS Applications & Systems Best Paper Award, Texas Instruments Visiting Professor (Fall 2002, Rice University), 2001 Okawa Foundation Award, 2000 OSU College of Engineering Lumley Award for Research, 1997 OSU College of Engineering MacQuigg Award for Teaching, and a 1996 National Science Foundation (NSF) CAREER Award. Dr. Mitra currently serves on the IEEE Fourier Award for Signal Processing committee, IEEE James H. Mulligan, Jr. Education Medal committee and the IEEE Paper Prize Committee. She is the inaugural Editor-in-Chief for the IEEE Transactions on Molecular, Biological and Multi-scale Communications. Dr. Mitra has been/is an Associate Editor for the following IEEE publications: Transactions on Signal Processing (2012–), Transactions on Information Theory (2007–2011), Journal of Oceanic Engineering (2006–2011), and Transactions on Communications (1996–2001). She has co-chaired: (technical program) 2014 IEEE International Symposium on Information Theory in Honolulu, HI, 2014 IEEE Information Theory Workshop in Hobart, Tasmania, IEEE 2012 International Conference on Signal Processing and Communications, Bangalore India, and the IEEE Communication Theory Symposium at ICC 2003 in Anchorage, AK; and general co-chair for the first ACM Workshop on Underwater Networks at Mobicom 2006, Los Angeles, CA. Dr. Mitra was the Tutorials Chair for IEEE ISIT 2007 in Nice, France and the Finance Chair for IEEE ICASSP 2008 in Las Vegas, NV. Dr. Mitra has held visiting appointments at: the Delft University of Technology, Stanford University, Rice University, and the Eurecom Institute. She served as co-Director of the Communication Sciences Institute at the University of Southern California from 2004–2007. Her research interests are in: wireless communications, communication and sensor networks, detection and estimation and the interface of communication, sensing and control.