Robust and sparse Dual Tree Complex Wavelet Transform-based Twin Support Vector regression for dense 5G InH Communications

anis charrada (anis.charrada@gmail.com)
University of Carthage: Universite de Carthage https://orcid.org/0000-0002-5966-6347
Abdelaziz Samet
INRS-EMT: Institut national de la recherche scientifique Centre energie Materiaux Telecommunications

Research Article

Keywords: mmWave, Dual Tree Complex Wavelet Transform, robust and sparse TSVR, Indoor Hotspot, path loss, Close-In model

Posted Date: September 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-787817/v1

Noname manuscript No.
(will be inserted by the editor)

Robust and sparse Dual Tree Complex Wavelet Transform-based Twin Support Vector regression for dense 5G InH Communications

Anis CHARRADA · Abdelaziz SAMET

the date of receipt and acceptance should be inserted later

Abstract A robust and sparse Twin Support Vector Regression based on Dual Tree Discrete Wavelet Transform algorithm is conceived in this paper and applied to 28, 38, 60 and 73-GHz LOS (Line-of-Sight) wireless multipath transmission system in 5G Indoor Hotspot (InH) settings (simple, semi-complex and complex conference rooms) under small receiver sensitivity threshold. The algorithm establishes a denoising process in the learning phase based on Dual Tree Discrete Wavelet Transform (DT-CWT) which is suitable for time-series data. Additionally, the Close-In (CI) free space reference distance path loss model is analyzed and the large-scale propagation and probability distribution functions are investigated by determining the PLE (Path Loss Exponent) and the standard deviation of Shadow Factor (SF) for each InH scenario under consideration. Performance are evaluated under twelve (12) configuration scenarios, according to three criteria: mobility (0/3mps), receiver sensitivity threshold (-80/-120 dBm) and type of the InH area (simple, semi-complex and complex conference room). Experimental results confirm the effectiveness of the proposed approach compared to other standard techniques.

Keywords: mmWave, Dual Tree Complex Wavelet Transform, robust and sparse TSVR, Indoor Hotspot, path loss, Close-In model.

A. Charrada
SERCOM, Tunisia Polytechnic School, University of Carthage
Box 743 - 2078 La Marsa Tunisia
Tel.: +216-52-075105
E-mail: anis.charrada@gmail.com

A. Samet
INRS, EMT Center 800, de la Gauchetiére W., Suite 6900 Montreal, QC, H5A 1K6, CANADA. E-mail: samet@emt.inrs.ca

1 Introduction

The challenge between bandwidth needs and spectrum constraints is becoming increasingly relevant with the exponential rise in wireless traffic demand. With the proliferation in mobile data usage, the mobile network of the fifth generation (5G) will leverage the vast amount of bandwidth in the millimeter wave (mmWave) bands to expand communication bandwidth considerably. In terms of high transmission path loss and blocking sensitivity, there are fundamental differences between mmWave communications and other current communications networks [1]. A specific insight of the free-licensed 60-GHz band was obtained to establish 5G wireless networks in Indoor Hotspot (InH) environments (office, shopping mall, conference rooms etc) and evolving mmWave networking systems for commercial demands [2] [3].

The high carrier frequency of the transmitter implies that mmWave communications have immense transmission losses, and the use of beamforming as an integral technology signify that mmWave communications are essentially directional. In particular, the 60-GHz Band is ideal for small-scale implementations and high data rate transmission applications including InH devices in future 5G networks [4]. Nevertheless, mmWave communications in indoor area are vulnerable to obstacles such as humans and furniture because of their weak diffraction ability. Consequently, mmWave connectivity is usually used in indoor settings, small cell connections and a cell backhaul of about 200 meters of length. Accordingly, the free space propagation loss at 60-GHz with a wavelength of around 5 mm is 28 decibels (dB) higher than at 2.4 GHz [5] [6].

Many indoor networks standards have been established by several international institutions, to use the free-licensed mmWave band for achieving high-throughput transmission data such that ECMA-387, IEEE 802.11ad and IEEE 802.15.3c [1].

Several publications have appeared in recent years documenting the Indoor mmWave communication [7] [8] [9]. In fact, path loss behaves at a significant number of obstacles and NLOS (Non Line-of-Sight) components experience major variations, which make transmission challenging at NLOS positions.

Alternatively, the Orthogonal Frequency Division Multiplexing (OFDM), which offers high-speed data rate, spectral efficiency and simple equalizer design has proven to be a good multicarrier transmission technique and an access to broadband digital communication networks [10]. The literature on deep fading channel identification shows a variety of approaches. Discrete Fourier transform [11], Linear minimum mean squares error (LMMSE) [12], Basis-expansion model [13], have been mostly employed. Study has been in-

creasingly popular in recent years on the adaptive weighted average estimator [14]. However, in the complex mmWave propagation setting, these approaches may not be feasible.

This means that a specific mmWave channel estimation approach based on Support Vector Regression (SVR) is helpful to develop the performance efficiency of the transmission system.

In the last decade, Twin SVR algorithm has attracted much attention from research teams. In [15], authors suggested a parametric Twin SVM for pattern classification. Authors in [16] proposed multi-target regression and defined some models for multiple outputs structures. A feature selection method based on TSVR is applied in [17]. Support vector machine classifier with truncated pinball loss is investigated by Shen et al. in [18]. Likewise, in [19], authors developed a K-nearest neighbor TSVR technique that was appropriate for clustering samples regression problems but not effective for time series data like the mmWave channel.

Recently, our paper [20] have suggested a TSVR procedure which is assessed for only one mmWave multipath environment (complex conference room) in small-scale propagation. However, there are still some interesting and relevant topics to be addressed such that large-scale propagation path loss in one hand, and the behavior of TSVR algorithm in various small-scale propagation InH scenarios in other hand at several mmWave frequency bands.

Here are synthesized the main contributions of this work.

1. The proposed Twin support vector regression based on Dual Tree Complex Discrete Wavelet Transform (DT-CWT) algorithm is conceived and applied to 28, 38, 60 and 73-GHz mmWave transmission bands in Indoor Hotspot settings (simple room, semi-complex room and complex room) under small receiver sensitivity threshold (-120 dBm).

2. The Close-In (CI) free space reference distance path loss model is analyzed and the large-scale propagation and probability distribution functions are investigated by determining the PLE (Path Loss Exponent) and the standard deviation of Shadow Factor (SF) for each InH scenario under consideration.

3. Bit Error Rate (BER) performance is evaluated under twelve (12) configuration scenarios, according to three criteria: mobility (0 / 3mps), receiver sensitivity threshold (-80 / -120 dBm) and type of the InH area (simple, semi-complex and complex conference room).

4. Normalized Mean Squares Error (NMSE) and Root Mean Squares Error Vector Magnitude (RMS EVM) are also evaluated.

The remainder of this paper is organized as follows. First, Section 2 outlines the system model. Twin SVR approach based on Dual Tree Discrete Wavelet Transform is illustrated in Section 3. In Section 4 simulation and experimental results are presented and discussed. Finally, Section 5 concludes this work.

Notations: $(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^{-1}$ mention the transpose conjugate, the transpose and the inverse, respectively.

2 System model

Take into consideration an OFDM system with N subcarriers. The time-domain transmitted sequence $x(n)$ can be explained as the following form:

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi nk/N}, \quad n = -M, ..., N-1, \tag{1}$$

where $X(k)$ stands for QAM constellation symbol and M symbolizes the cyclic prefix (CP) length which is equal to or larger than the channel order L. After eliminating CP at the receiver, the time-domain received signal $y(n)$ can be expressed as

$$y(n) = \sum_{l=0}^{L} h_l(n) x(n-l) + \kappa(n), \tag{2}$$

where $h_l(n)$ denotes the l^{th} baseband-correspondent channel impulse response at time n and $\kappa(n)$ symbolizes zero-mean additive white Gaussian noise (AWGN) with variance σ_n^2.

Otherwise, equation (2) can be represented in matrix form as following:

$$\mathbf{y} = \mathbf{h}\mathbf{F}^H\mathbf{X} + \kappa = \mathbf{h}\mathbf{x} + \kappa, \tag{3}$$

where $\mathbf{y} = [y(0), y(1), ..., y(N-1)]^T$, $\mathbf{x} = [x(0), x(1), \ldots, x(N-1)]^T$, $\mathbf{X} = [X(0), X(1), \ldots, X(N-1)]^T$ and $\kappa = [\kappa(0), \kappa(1), ..., \kappa(N-1)]^T$. Moreover, $\mathbf{F}^H$ involves N-point Inverse Discrete Fourier Transform (IDFT) matrix and $\mathbf{h} \in C^{N\times N}$ denotes the channel matrix. Once Fourier transform has been accomplished in both parts of (3), the following expressions are obtained:

$$\mathbf{F}\mathbf{y} = \mathbf{F}\mathbf{h}\mathbf{F}^H\mathbf{X} + \mathbf{F}\kappa \tag{4}$$

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{F}\kappa, \tag{5}$$

where $\mathbf{Y} = [Y(0), Y(1), \ldots, Y(N-1)]^T \in C^N$ stands for the frequency-domain received signal. Also, $\mathbf{H} = \mathbf{F}\mathbf{h}\mathbf{F}^H \in C^{N\times N}$ represents the channel

matrix counting Inter-Carrier-Interference (ICI) in frequency-domain caused by the time-varying channel, which elements can be described as follows:

$$\mathbf{H}(s,q) = \frac{1}{N}\sum_{l=0}^{L} e^{-j2\pi sl/N} \sum_{n=0}^{N-1} h_l(n) e^{-j2\pi(s-q)n/N}, \tag{6}$$

where $s, q = 0, 1, ..., N-1$.

Note that $\mathbf{H}$ can be separated into two parts, one of which represents the basic diagonal components $\mathbf{H}_{dg} \in C^{N\times N}$ and the other of the non-diagonal $\mathbf{H}_{ndg} \in C^{N\times N}$ consisting to ICI components.

In these simpler forms, equations (4) can be expressed as:

$$\mathbf{Y} = \mathbf{H}_{dg}\mathbf{X} + \mathbf{H}_{ndg}\mathbf{X} + \mathbf{F}\kappa \tag{7}$$

$$= diag(\mathbf{X})\mathbf{H}_{dg}^{'} + \mathbf{H}_{ndg}\mathbf{X} + \mathbf{F}\kappa, \tag{8}$$

with $diag(\cdot)$ denotes the diagonal operator. Additionally, the column vector $\mathbf{H}_{dg}^{'} \in C^N$ symbolizes the elements taken from the diagonal matrix $\mathbf{H}_{dg}$ and the amount $(\mathbf{H}_{ndg}\mathbf{X})$ stands for the ICI elements.

3 DTCWT-Twin SVR approach

Having regard to the training set $\mathbf{S} = \{(\mathbf{t}_1, o_1), (\mathbf{t}_2, o_2), ..., (\mathbf{t}_m, o_m)\}$, with $\mathbf{t}_i \in R^2$ and $o_i \in R$, $i = 1, 2, ..., m$. The training data input matrix can then be defined as $\mathbf{T} = (\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_m)^T \in R^{m\times 2}$ and the vector output can be described as $\mathbf{O} = (o_1, o_2, ..., o_m)^T \in R^m$. The time and frequency domain coordinates of the training samples can be represented in the first and second columns of $\mathbf{T}$, respectively.

In our suggested methodology, the perceived OFDM symbols comprise reference and data symbols in both times and frequencies, where reference symbols are inserted equidistantly. The reference symbols' positions are defined as $(n\Delta_t)$, $n = 0, 1, ..., N_t - 1$, with N_t and Δ_t represent the number of reference symbols and the time-domain reference symbol interval, respectively. In addition, in each OFDM symbol, the transferring position of reference subcarrier is referenced as $(m\Delta_f)$, $m = 0, 1, ..., N_f - 1$, with Δ_f displays the frequency-domain reference interval and N_f symbolizing the number of reference symbols per OFDM symbol.

Consider the reference matrix expressed in this sense as $\mathbf{X}_r = diag(X(n\Delta_t, m\Delta_f)) \in C^{N_tN_f\times N_tN_f}$. According to (7), it is possible to obtain the estimated channel frequency responses at reference subscarriers using the following Least-Squares (LS) criterion:

$$\hat{\mathbf{H}}_r = (\mathbf{X}_r)^{-1}\mathbf{Y}_r, \tag{9}$$

with $\hat{\mathbf{H}}_r = \hat{\mathbf{H}}(n\Delta_t, m\Delta_f) \in C^{N_t \times N_f}$ displays the estimated channel frequency responses at reference positions $(n\Delta_t, m\Delta_f)$ and $\mathbf{Y}_r = \mathbf{Y}(n\Delta_t, m\Delta_f) \in C^{N_t \times N_f}$ stands for the received reference vector.

In fact, the DT-CWT (Dual-Tree Complex Wavelet Transform) improves the conventional DWT (Discrete Wavelet Transform) efficiently, solving low performance in complex signals (e.g. radar, voice, etc). As demonstrated in [21], two real DWTs form a DT-CWT with real and imaginary components provided by the first real DWT and the second real DWT. The forward DT-CWT process is shown in Fig. 1. Providing two real DWTs expressed as real $(\mathbf{\Psi}_h(t))$ and imaginary $(\mathbf{\Psi}_g(t))$, the standard formulation of the DT-CWT can be given by [21]

$$\mathbf{\Psi}(t) = \mathbf{\Psi}_h(t) + j\mathbf{\Psi}_g(t). \tag{10}$$

Because there is no interrelationship between $\mathbf{\Psi}_h(t)$ and $\mathbf{\Psi}_g(t)$, they may be computed independently.

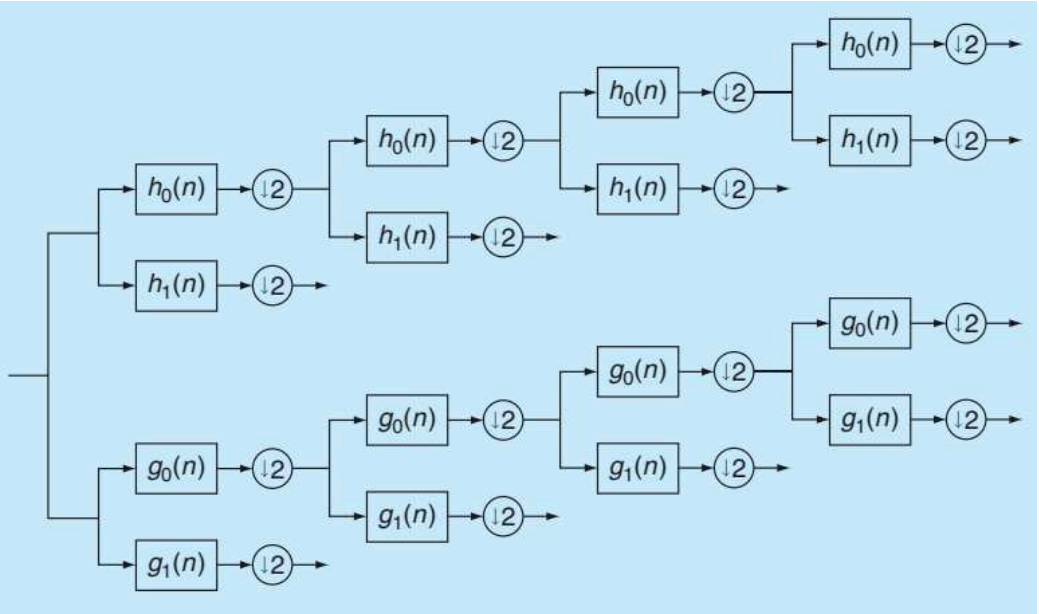


Fig. 1: Analysis filter bank for the DT-CWT [21] .

The DT-CWT overcomes the typical problems of limited directional selectivity and shift variance in two and more dimensional issues with the usually-used of DWT. In order to accomplish the denoising process, the performance of the dual-tree complex wavelet transform is evaluated in this work.

The estimated channel frequency responses at reference locations can therefore be articulated as follows after the denoising process:

$$\hat{\mathbf{H}}_{dn} = \mathbf{\Psi}(\hat{\mathbf{H}}_r). \tag{11}$$

For all OFDM data position subcarriers, the predicted channel frequency response can therefore be calculated by the following equation:

$$\tilde{\mathbf{H}} = f(\hat{\mathbf{H}}_{dn}), \tag{12}$$

with $f(\cdot)$ denotes the Twin SVR formulation.

In this way, two bound functions construct Twin SVR: the time-domain down-bound function $f_1(\cdot)$ and the frequency-domain up-bound function $f_2(\cdot)$ consisting to two nonparallel hyperplanes, where each hyperplane creates its own ϵ-insensitive bound regressor. Afterward, the overall TSVR regressor corresponds to the mean of the sum of $f_1(\cdot)$ and $f_2(\cdot)$ as

$$f(\cdot) = \frac{1}{2}\left((f_1(\cdot) + f_2(\cdot)\right). \tag{13}$$

It is interesting to note that we have used the Gaussian kernel to resolve the fast fading of the mmwave channel thanks to its potentiel to map the input space through an implicit transformation $\boldsymbol{\varphi}$ into an infinite-dimensional output space. Thus, we only have to determine the Gram matrix $\boldsymbol{G}$ by accomplishing the Gaussian kernel inner product $\boldsymbol{K}$ expressed as follows:

$$\boldsymbol{G}(1,2) = \boldsymbol{K}(\boldsymbol{t}_1, \boldsymbol{t}_2) = < \boldsymbol{\varphi}(\boldsymbol{t}_1), \boldsymbol{\varphi}(\boldsymbol{t}_2) > . \tag{14}$$

In this sense, the down and up-bound functions generated by the kernel can be expressed in these simpler forms:

$$f_1(\mathbf{t}) = \boldsymbol{K}(\mathbf{t}, \mathbf{T}^T)\boldsymbol{w}_1 + b_1 + e_1, \tag{15}$$

$$f_2(\mathbf{t}) = \boldsymbol{K}(\mathbf{t}, \mathbf{T}^T)\boldsymbol{w}_2 + b_2 + e_2, \tag{16}$$

with $\boldsymbol{w}_1, \boldsymbol{w}_2 \in C^m$, $b_1, b_2, e_1, e_2 \in C$.

To minimize approximation errors e_1 and e_2, we exploited the ε-Huber loss function presented in [22]. For purposes of convenience, we symbolize by $(\mathbf{t}_n^1)$ the $\mathbf{T}(n,1)$ value corresponding to the first reference symbol location at the n^{th} line. In the same manner, $(\mathbf{t}_m^2)$ symbolizes the $\mathbf{T}(m,2)$ value corresponding to the second reference symbol location at the m^{th} line.

3.1 Primal problem

The time and frequency domains primal optimization formulation can be written as:

$$\begin{aligned} \textit{minimize} \quad & \frac{1}{2}\left\|\boldsymbol{w}_{(1,2)}\right\|^2 + \frac{1}{2\gamma_{(1,2)}} \sum_{n\in A, m\in A'} (\xi_{(n,m)} + {\xi_{(n,m)}}^*)^2 \\ & + C_{(1,2)} \sum_{n\in B, m\in B'} (\xi_{(n,m)} + {\xi_{(n,m)}}^*) + \frac{1}{2\gamma_{(1,2)}} \sum_{n\in C, m\in C'} (\zeta_{(n,m)} + {\zeta_{(n,m)}}^*)^2 \\ & + C_{(1,2)} \sum_{n\in D, m\in D'} (\zeta_{(n,m)} + {\zeta_{(n,m)}}^*) - \frac{1}{2} \sum_{n\in\{B,D\}; m\in\{B',D'\}} \gamma_{(1,2)} (C_{(1,2)})^2 \end{aligned} \tag{17}$$

subject to

$$\begin{aligned} \Re\left(\hat{H}(\mathbf{t}^{(1,2)}_{(n,m)}) - \boldsymbol{w}^H_{(1,2)}\boldsymbol{\varphi}_{(1,2)}(\mathbf{t}^{(1,2)}_{n,m}) - b_{(1,2)}\right) &\leq \varepsilon_{(1,2)} + \xi_{(n,m)} \\ \Im\left(\hat{H}(\mathbf{t}^{(1,2)}_{(n,m)}) - \boldsymbol{w}^H_{(1,2)}\boldsymbol{\varphi}_{(1,2)}(\mathbf{t}^{(1,2)}_{n,m}) - b_{(1,2)}\right) &\leq \varepsilon_{(1,2)} + \zeta_{(n,m)} \\ \Re\left(-\hat{H}(\mathbf{t}^{(1,2)}_{(n,m)}) + \boldsymbol{w}^H_{(1,2)}\boldsymbol{\varphi}_{(1,2)}(\mathbf{t}^{(1,2)}_{n,m}) + b_{(1,2)}\right) &\leq \varepsilon_{(1,2)} + \xi^*_{(n,m)} \\ \Im\left(-\hat{H}(\mathbf{t}^{(1,2)}_{(n,m)}) + \boldsymbol{w}^H_{(1,2)}\boldsymbol{\varphi}_{(1,2)}(\mathbf{t}^{(1,2)}_{n,m}) + b_{(1,2)}\right) &\leq \varepsilon_{(1,2)} + \zeta^*_{(n,m)} \\ {\xi_{n,m}}^{(*)}, {\zeta_{n,m}}^{(*)} &\geq 0 \ . \end{aligned} \tag{18}$$

with $n = 1, \cdots, N_t$; $m = 1, \cdots, N_f$, and $\xi^{(*)}_{(n,m)}$ denote slack variables for both positive and negative real values in time and frequency-domains, respectively. Furthermore, $\zeta^{(*)}_{(n,m)}$ represent slack variables for both positive and negative imaginary values in time and frequency-domains, respectively. Variables $C_{(1,2)}$ and $\gamma_{(1,2)}$ essentially monitor the equilibrium in time and frequency domains between losses and regularization.

$\{A, B, C, D\}$ and $\{A^{'}, B^{'}, C^{'}, D^{'}\}$ describe the sequence of points represented in time and frequency-domains by:
$\{A, A^{'}\}$: quadratic field that contains real error values ;
$\{B, B^{'}\}$: linear field that contains real error values;
$\{C, C^{'}\}$: quadratic field that contains imaginary error values;
$\{D, D^{'}\}$: linear field that contains imaginary error values.

3.2 Dual problem

To transform the minimization primal problem (17) subject to (18) into a dual optimization problem, constraints should first be incorporated into the primal problem with a view to the accomplishment of the Lagrangian correspondent.

After deriving this Lagrangian with respect to $\boldsymbol{w}_i$ and derivatives canceling, we obtain the optimal weights which can be represented in the following expressions:

$$\boldsymbol{w}_{(1,2)} = \sum_{n=1,m=1}^{N_t,N_f} \psi_{(n,m)}^{(1,2)} \boldsymbol{\varphi}_{(1,2)}(\mathbf{t}_{(n,m)}^{(1,2)}), \tag{19}$$

with parameters $\psi_{(n,m)}^{(1,2)}$ can be written as

$$\psi_{(n,m)}^{(1,2)} = (\alpha_{R,(n,m)}^{(1,2)} - \alpha_{R,(n,m)}^{(1,2)\ *}) + j(\alpha_{I,(n,m)}^{(1,2)} - \alpha_{I,(n,m)}^{(1,2)\ *}), \tag{20}$$

with $(\alpha_{R,n}^{1}, \alpha_{R,n}^{1\ *}, \alpha_{I,n}^{1}, \alpha_{I,n}^{1\ *})$ and $(\alpha_{R,m}^{2}, \alpha_{R,m}^{2\ *}, \alpha_{I,m}^{2}, \alpha_{I,m}^{2\ *})$ correspond to multipliers in time and frequency-domains for real and imaginary residual elements. It is more important that our considered nonlinear SVR structure should be viewed as a compact matrix format. The following dual formulations can therefore be expressed in time and frequency-domains as:

$$\begin{aligned} maximize \quad & -\frac{1}{2}\boldsymbol{\psi}^{(1,2)H}\left(\boldsymbol{G}_{(1,2)} + \gamma_{(1,2)}\mathbf{I}\right)\boldsymbol{\psi}^{(1,2)} + \Re\left(\boldsymbol{\psi}^{(1,2)H}\mathbf{Y}_r\right) \\ & -\left(\boldsymbol{\alpha}_{\boldsymbol{R}}^{(\mathbf{1,2})} + \boldsymbol{\alpha}_{\boldsymbol{R}}^{(\mathbf{1,2})*} + \boldsymbol{\alpha}_{\boldsymbol{I}}^{(\mathbf{1,2})} + \boldsymbol{\alpha}_{\boldsymbol{I}}^{(\mathbf{1,2})*}\right)\boldsymbol{1}\varepsilon_{(1,2)}, \end{aligned} \tag{21}$$

subject to

$$0 \leq \ \alpha_{R,(n,m)}^{(1,2)}, \alpha_{R,(n,m)}^{(1,2)\ *}, \alpha_{I,(n,m)}^{(1,2)}, \alpha_{I,(n,m)}^{(1,2)\ *} \ \leq C_{(1,2)} \tag{22}$$

where $\boldsymbol{\psi}^{(1,2)} = [\psi_{(1,1)}^{(1,2)}, \cdots, \psi_{N_{(t,f)}}^{(1,2)}]^T$. Notice that $\mathbf{I}$ and $\mathbf{1}$ symbolize the identity matrix and all ones vector, respectively. Moreover, $\boldsymbol{\alpha}_{\boldsymbol{R}}^{(\mathbf{1,2})}$ and $\boldsymbol{\alpha}_{\boldsymbol{I}}^{(\mathbf{1,2})}$ represent real and imaginary vectors of dual variables in time and frequency-domains. It can be easily seen that we can determine the weights solution by optimizing the formulation (21) in relation to $\alpha_{R,(n,m)}^{(1,2)}$, $\alpha_{R,(n,m)}^{(1,2)\ *}$, $\alpha_{I,(n,m)}^{(1,2)}$, $\alpha_{I,(n,m)}^{(1,2)\ *}$ and next replacing into (19).

Thus, the functions down-bound $f_1(\cdot)$ and up-bound $f_2(\cdot)$ according to equations (15) and (16) may be defined. Correspondingly, by averaging the sum of these two functions as mentioned above, the Twin SVR regressor can be determined.

3.3 Algorithm summary

In this subsection, we illustrate the dual tree complex wavelet transform-based Twin SVR algorithm.

Algorithm: DTCWT-Twin SVR algorithm

Input:

- Reference symbols positions at the transmitter.
- Reference symbols matrix at the transmitter.
- Received reference symbols vector.
- Appropriate TSVR parameters.

Output:

The optimal solution of the predicted frequency response at all subcarriers.

Process:

1. Calculate channel frequency response at reference subcarriers according to (Eq.(9)).

2. Perform Dual Tree Complex Wavelet Transform (DWT) in denoising operation at training phase by (Eq.(11)).

3. Solve the dual problem according to (Eq.(21)).

4. Substitute the obtained solution into (Eq.(19)) to obtain the optimal weights in time-domain.

5. Resolve the time-domain down-bound function $f_1(\cdot)$ by (Eq.(15)).

6. Repeat **1-5** once to resolve the frequency-domain up-bound function $f_2(\cdot)$ by (Eq.(16)).

7. Get the DTCWT-TSVR solution by averaging the sum of $f_1(\cdot)$ and $f_2(\cdot)$

8. Determine the frequency responses of all subcarriers in all OFDM symbols by (Eq.(12)).

4 Simulation and experimental results

4.1 Large-scale propagation path loss analysis

In order to provide the large-scale channel characterization (path loss) for mmWave transmissions in Indoor hotspot environment, we perform simulation for customizable (simple and complex) conference room in LOS situations. Thus, we simulate the input data files containing the measured path loss values provided by the "Wireless InSite" ray tracer by Remcom [23] at several frequencies 28, 38, 60 and 73 GHz. Hence, we analyze Close-In (CI) free space reference distance path loss model which corresponds to a multi-frequency stochastic model that explains large-scale propagation path loss over distance at all appropriate frequencies for a specific scenario [24]. The

CI model expression can be stated as follows [25]:

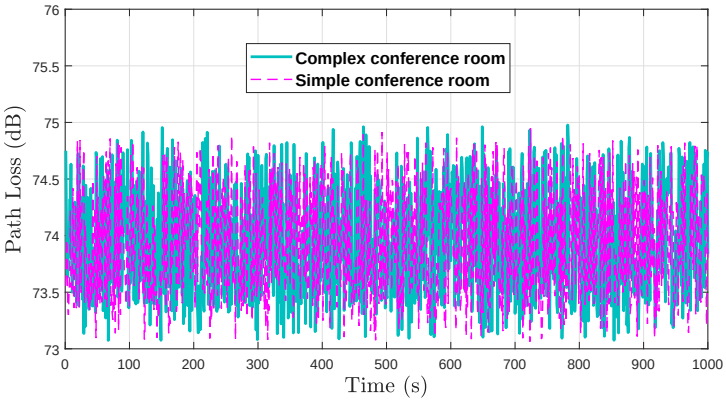
$$PL(f,d)[dB] = FSPL(f,d_0)[dB] + 10\ n\ log_{10}\ (\frac{d}{d_0}) + \chi_\sigma, \quad where\ d > d_0, \tag{23}$$

where n characterizes the Path Loss Exponent (PLE), d_0 stands for the close-in free space reference distance, and χ_σ symbolizes a random Gaussian zero-mean variable with a standard deviation σ in dB. Only one parameter, the PLE, is necessary for the CI model to calculate the mean path loss with frequency and distance. A valuable property of (23) is that 10 n represents the path loss in dB for decades of distance starting at d_0. In (23), d denotes the 3D Tx-Rx distance of separation and $FSPL(f,d_0)$ refers to the free space path loss in dB at the carrier frequency f and at a Tx-Rx distance of separation d_0:

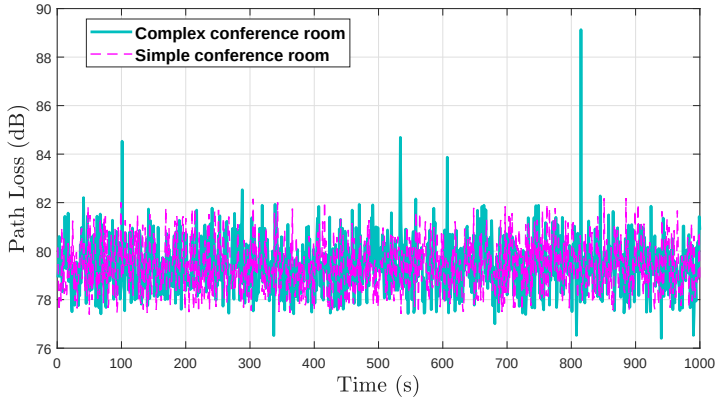
$$PSPL(f,d_0)[dB] = 20\ log_{10}(\frac{4\pi f d_0}{c}). \tag{24}$$

The selection of $d_0 = 1m$ as the reference distance for close-in free space demonstrates superior stability of parameters and model precision for indoor channels across a wide variety of microwave and mmwave frequencies, and build a standardized modeling methodology [25].
Consider Fig. 2, which plots path loss versus time for simple and complex conference rooms at 60 GHz with (a) d=2.5 m and (b) d=3.5m, respectively. As may be seen, path loss is more serious when 3D Tx-Rx distance is longer. For simple conference room, the mean path loss value at d=2.5m is equivalent to 73.9558 dB and 73.9388 dB for complex conference room. In the other hand, at d=3.5m, the mean path loss value is 79.3829 dB for simple conference room and 79.4095 dB for complex conference room. We can observe that the path loss differential of around 5 dB is accomplished by going from 2.5m to 3.5m for simple and complex conference room scenarios. Additionally, the difference in path loss of about .2 dB is achieved between transmission in a simple room and a complex conference room. This meaning is negligible as the transmission is realized under LOS conditions for both indoor settings, and we can see from Fig. 4 the path losses are the same for both cases ($FSPL(d_0) = 68.0048\ dB$) in free space.

In the CI model, the PLEs are 2.0481 and 2.0468 at 60 GHz for both (a) simple conference room and (b) complex conference room scenarios. Both values are very similar, which means that path losses in large scale are practically the same in both indoor rooms. This deduction is confirmed for all other frequencies 28, 38 and 73 GHz, as depicted in Figs. 3, 5 and 6.
Fig. 4 also displays the probability distribution function with $d_0 = 1m$. The standard deviations σ of Shadow Factor (SF) are .67069 and .71215 for (a)



(a)



(b)

Fig. 2: Variations of path loss for simple and complex conference room with distances equal to (a) 2.5 m and (b) 3.5 m, respectively.

simple conference room and (b) complex conference room scenarios, respectively. The SF σ of the CI model for customizable complex room is larger than that of a simple room due to the influence of electronic equipment and materials. The SF σ represents the standard deviation of the large-scale fluctuations of signal power about the mean path loss over distance. These findings explain the large-scale InH path losses in the 60 GHz LOS transmission, where the distance has the same implications for path loss in reasonably equal proportion. The standard deviations σ of Shadow Factor (SF) are 1.3989, .95134 and .81583 for simple conference room at 28, 38 and 73 GHz, respectively. It can be seen that σ progresses inversely with frequencies.

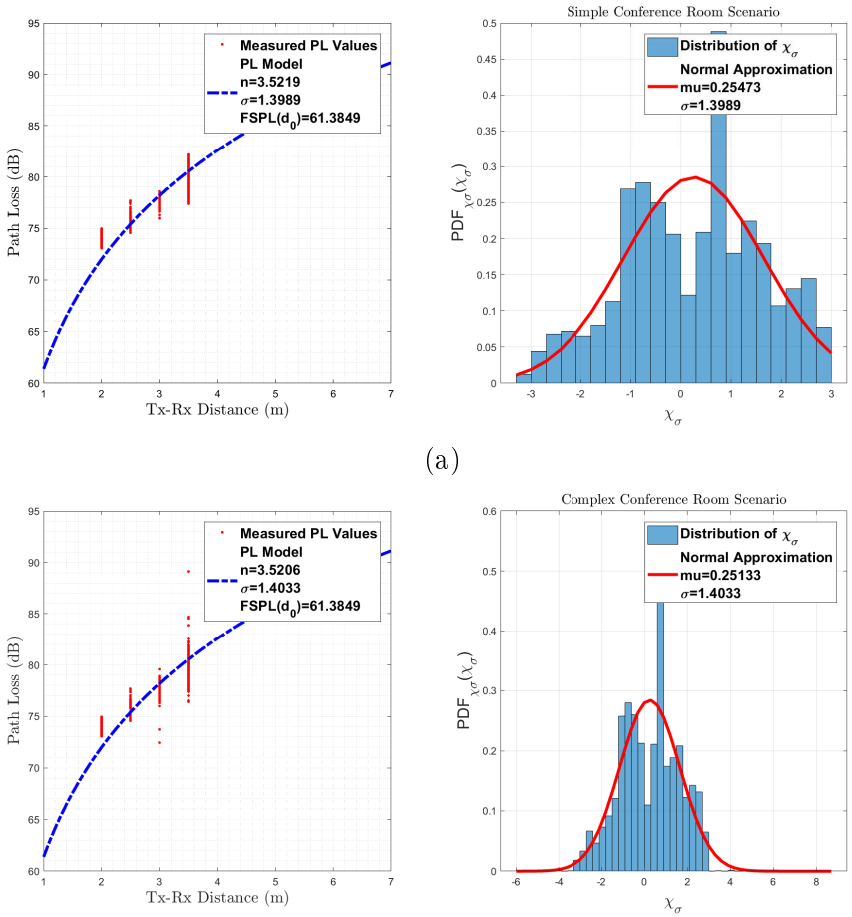
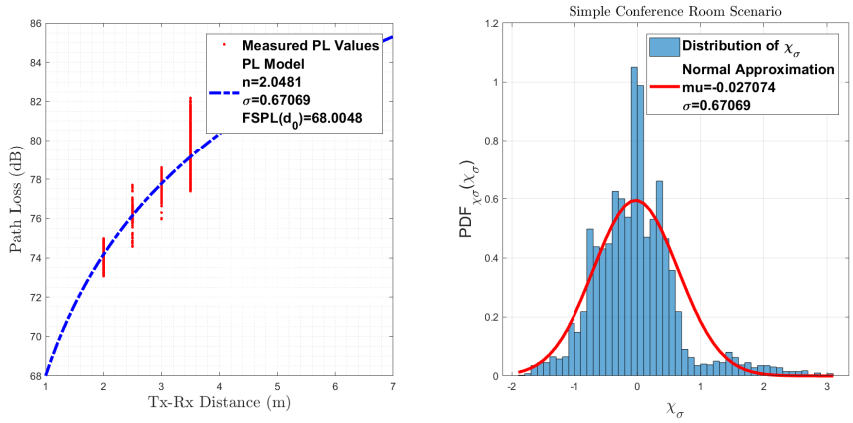


Fig. 3: Snapshot of Path Loss values and probability distribution function with $d_0 = 1m$ at 28 GHz for (a) simple conference room and (b) complex conference room scenarios, respectively.

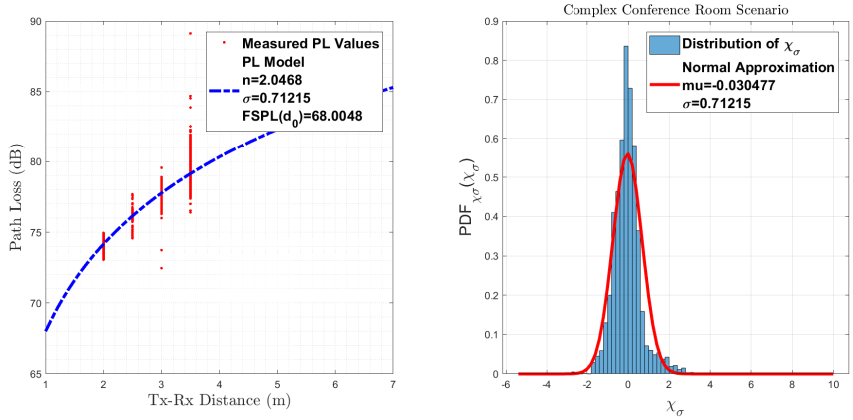
Next, particular attention will be paid to the small-scale multipath propagation analysis of the 28, 38, 60 and 73 GHz mmWave wireless systems for several indoor environments.

4.2 Multipath propagation analysis

The obtained power levels and the phase values of the multipath components entered at the receiver shall be assisted with the output of the ray tracer in the function of time for a consistent threshold.



(a)



(b)

Fig. 4: Snapshot of Path Loss values and probability distribution function with $d_0 = 1m$ at 60 GHz for (a) simple conference room and (b) complex conference room scenarios, respectively.

Fig. 7 presents the received power and phase vs. time of arrival with receiver sensitivity threshold S= -80 dBm for (a) simple conference room (20 paths), (b) semi complex conference room (22 paths) and (c) complex conference room (41 paths) scenarios, respectively.

As follows from Fig. 8, received power and phase vs. time of arrival with receiver sensitivity threshold S= -120 dBm are shown for (a) simple conference room (250 paths), (b) semi complex conference room (250 paths) and (c) complex conference room (250 paths) scenarios, respectively.

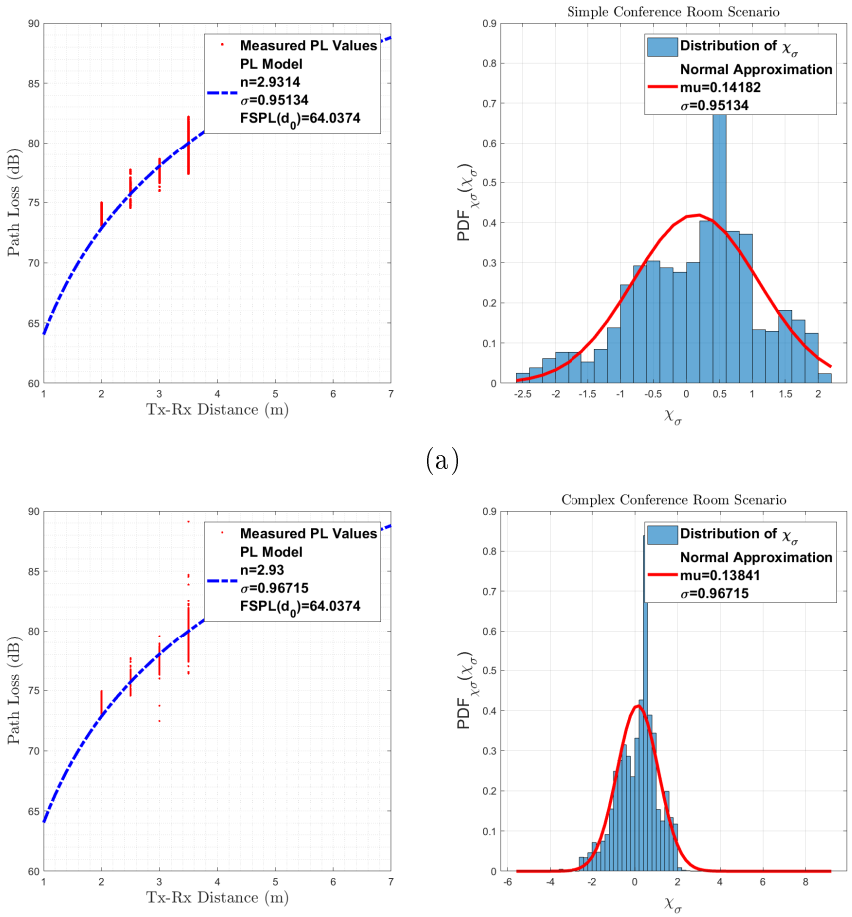


Fig. 5: Snapshot of Path Loss values and probability distribution function with $d_0 = 1m$ at 38 GHz for (a) simple conference room and (b) complex conference room scenarios, respectively.

As can be seen from Figs. 7 and 8, the number of paths to receiver sensitivity threshold -80 dBm is less than that of -120 dBm for all types of conference rooms.

In this regard, the OFDM structure with $N = 1024$ subcarriers and 128-QAM modulation scheme are taken into account. Three use-case scenario simulations are represented to determine the impacts of small-scale multipath propagation for simple, semi complex and complex conference rooms. It is worth noticing that, we appropriately choose TSVR parameters by the grid search approach from the value collection $\{10^i | i = -5, -4, \cdots, 5\}$.

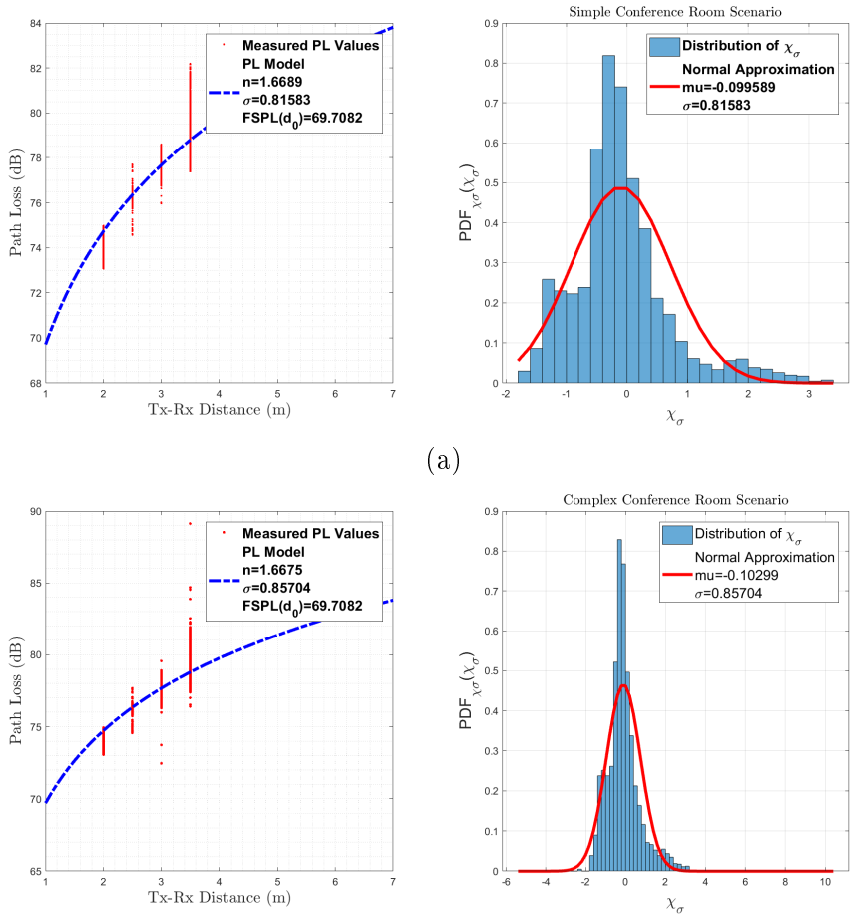
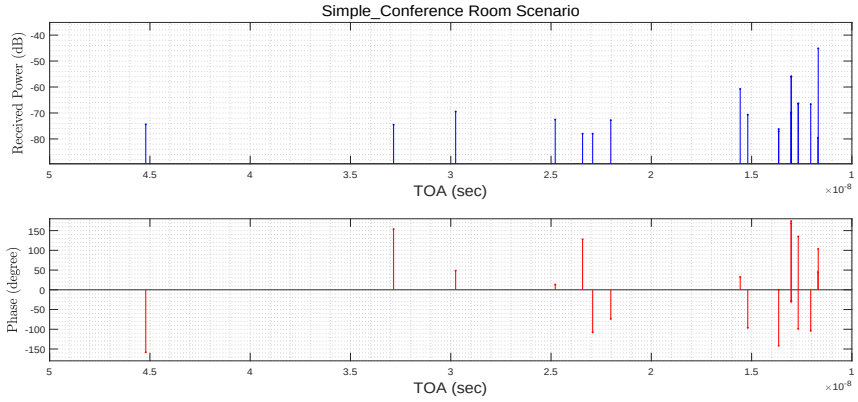


Fig. 6: Snapshot of Path Loss values and probability distribution function with $d_0 = 1m$ at 73 GHz for (a) simple conference room and (b) complex conference room scenarios, respectively.

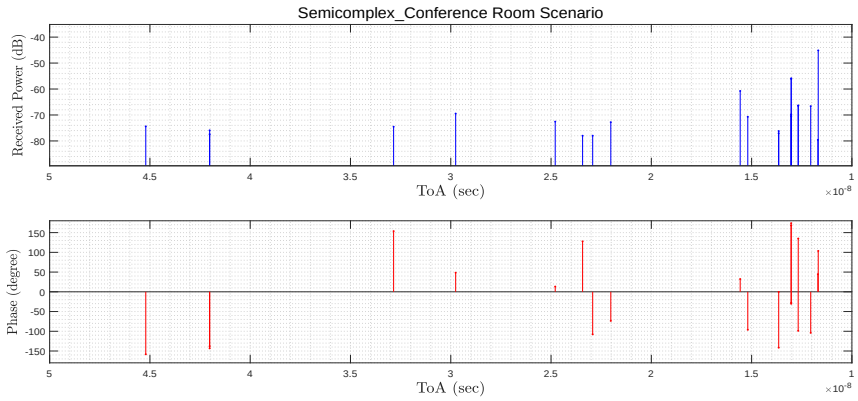
Figs. 9 and 10 show an example of the obtained coefficients with real and imaginary components from our experimental setup parameters at 28 and 60 Ghz, respectively. These dual coefficients, as mentioned in the third section, constitute the TSVR solution.

Additionally, the efficiency of the channel estimation based on Dual Tree Complex Wavelet Transform Twin Support Vector Regression and other methods is determined by the BER (Bit Error Rate) criteria.

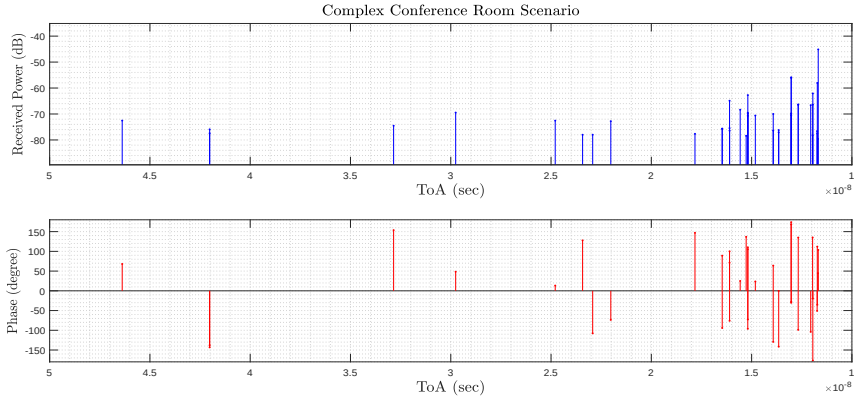
Fig. 11 illustrates the BER performance versus SNR (ranging from 0 to 25 dB) under 250 paths in semi-complex conference room for 128-QAM modulation scheme at 60 GHz under -120 dBm receiver sensitivity threshold



(a)

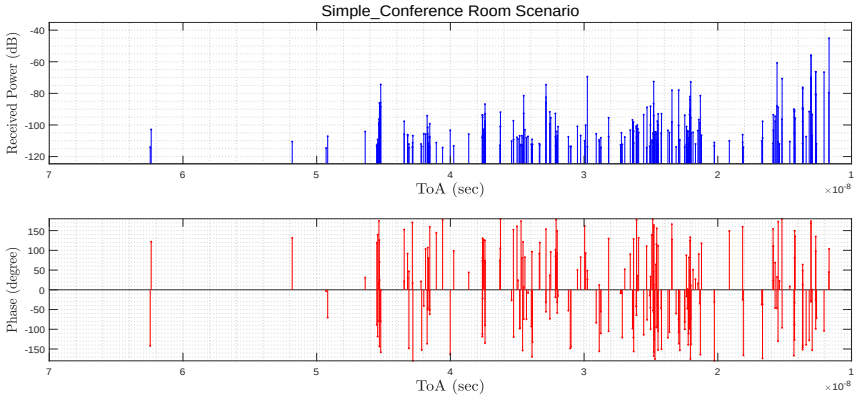


(b)

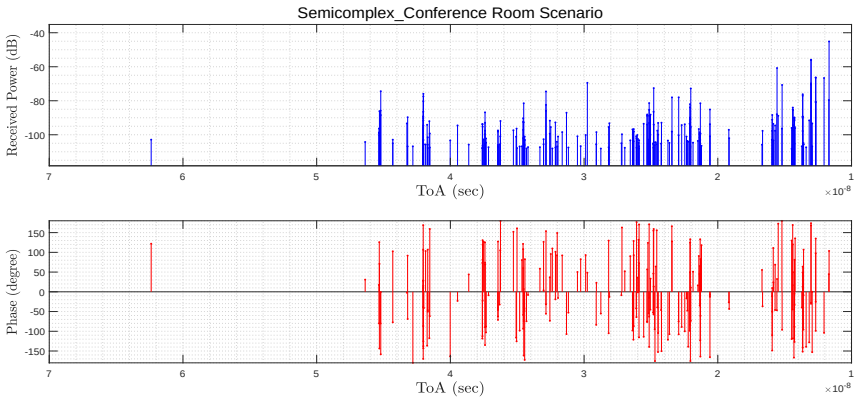


(c)

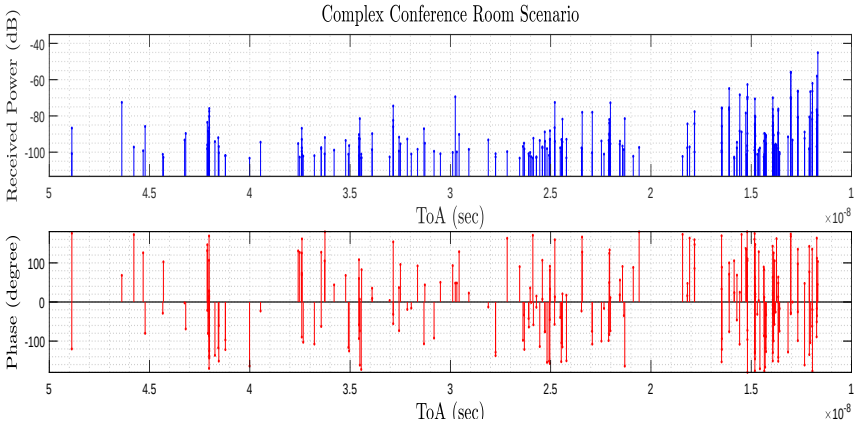
Fig. 7: Rx power and phase vs. time with receiver sensitivity threshold = -80 dBm for (a) simple conference room (20 paths), (b) semi complex conference room (22 paths) and (c) complex conference room (41 paths) scenarios, respectively.



(a)

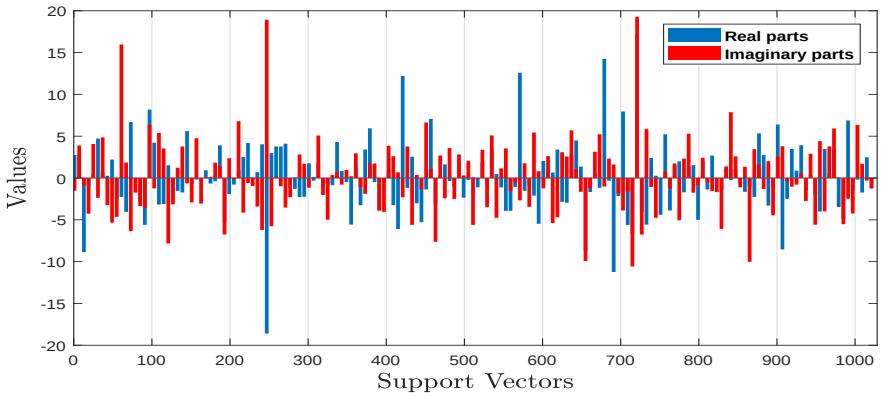


(b)

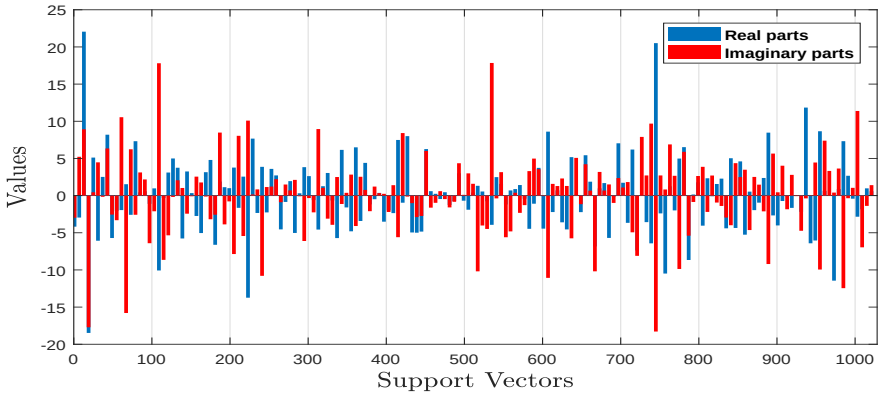


(c)

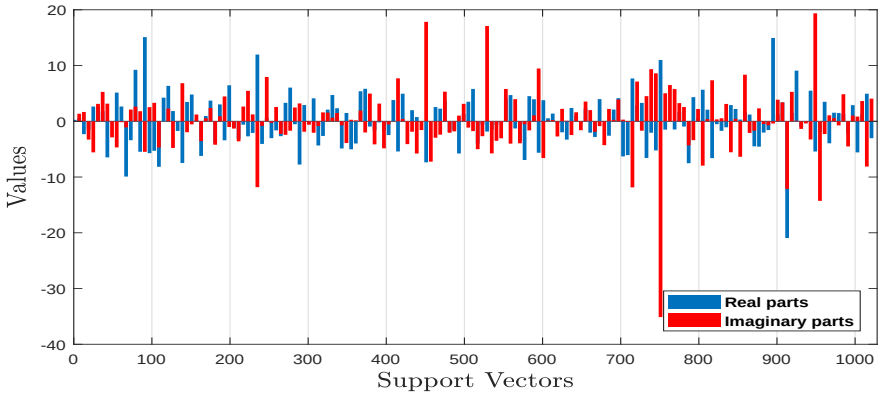
Fig. 8: Rx power and phase vs. time with receiver sensitivity threshold = -120 dBm for (a) simple conference room (250 paths), (b) semi complex conference room (250 paths) and (c) complex conference room (250 paths) scenarios, respectively.



(a)

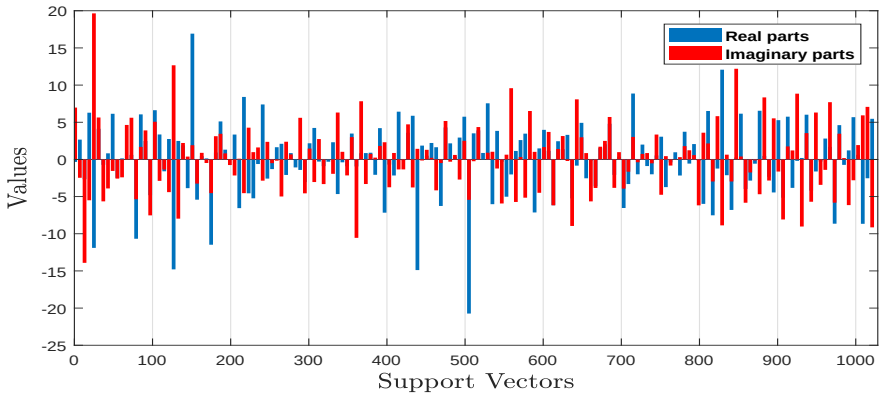


(b)

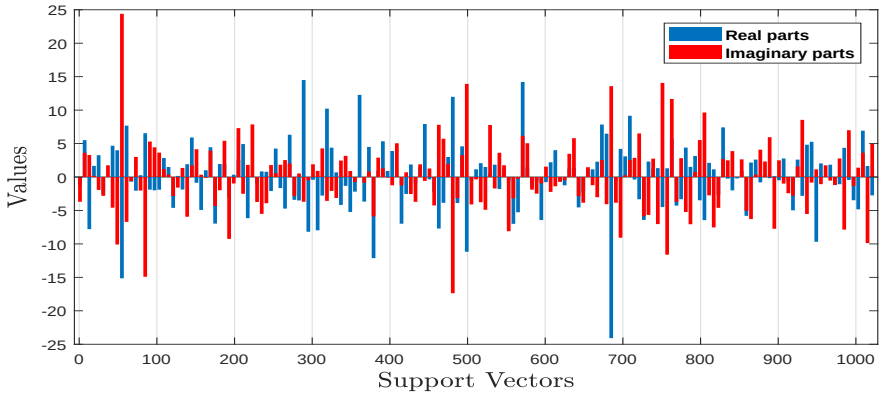


(c)

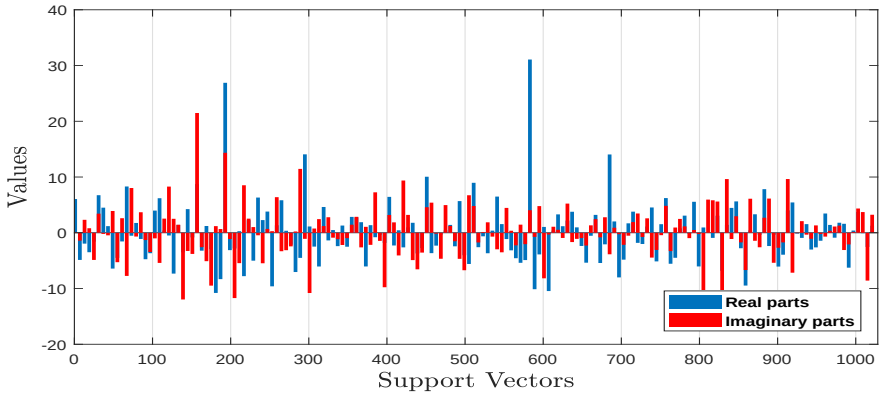
Fig. 9: Real and imaginary parts of support vectors with receiver sensitivity threshold = -120 dBm and 128 QAM modulation scheme at 28 GHz for (a) simple, (b) semi complex and (c) complex conference rooms scenarios, respectively.



(a)



(b)



(c)

Fig. 10: Real and imaginary parts of support vectors with receiver sensitivity threshold = -120 dBm and 128 QAM modulation scheme at 60 GHz for (a) simple, (b) semi complex and (c) complex conference rooms scenarios, respectively.

without mobility. With linear interpolation, Basis Expansion Model-based estimation, Least Squares-SVR estimation, K-nearest Twin SVR and perfect channel estimation, we evaluate the BER efficiency of the DWT-TSVR approach.

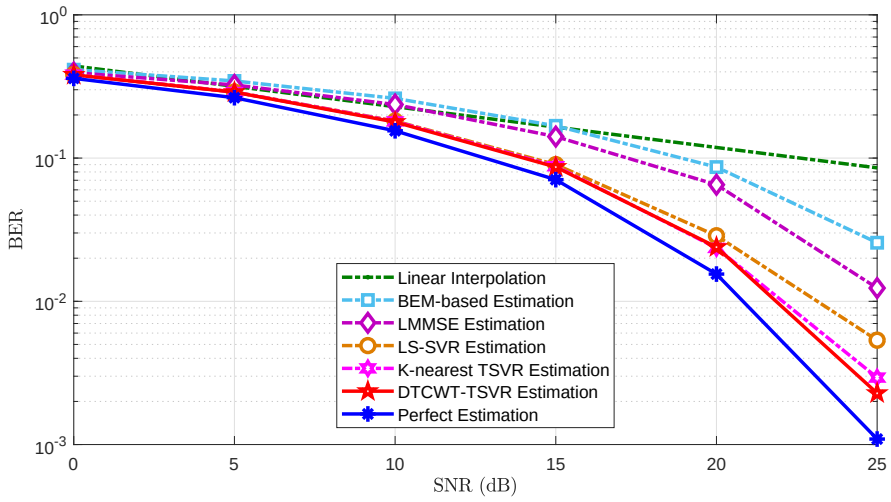


Fig. 11: BER vs. SNR for 128-QAM modulation scheme in semi-complex conference room with S=- 120 dBm.

It can clearly be shown that BER performance of the proposed method beats other techniques of estimation by reaching the minimum value (2.10^{-3}) at SNR = 25 dB and thus having the best frequency response estimate for the channel. Therefore, due to its generalization ability and the efficiency of the denoising process based on Dual Tree Discrete Wavelet Transform, the BER evaluation of the proposed algorithm is good for the adjustment of the noise in training samples, particularly in non-linear regression situations, and then to follow quick variations in the channels.

Fig. 12 shows BER performance versus SNR for several frequencies 28, 38 60 and 73 GHz with 128-QAM modulation scheme in semi-complex conference room with S=- 120 dBm with and without mobility. It can be seen that best performance is obtained for 28 GHz with BER equal to (2.10^{-4}). It should be noted here that, BER performance decreases for high frequencies, especially for 60 GHz band due to its attenuation properties.

In order to evaluate the effect of receiver sensitivity threshold S, number of paths, movement speed v and type of the conference room on the BER performance of the mmWave wireless system, Fig. 13 displays BER vs. SNR for 128-QAM modulation scheme at 60 GHz band under twelve (12) configu-

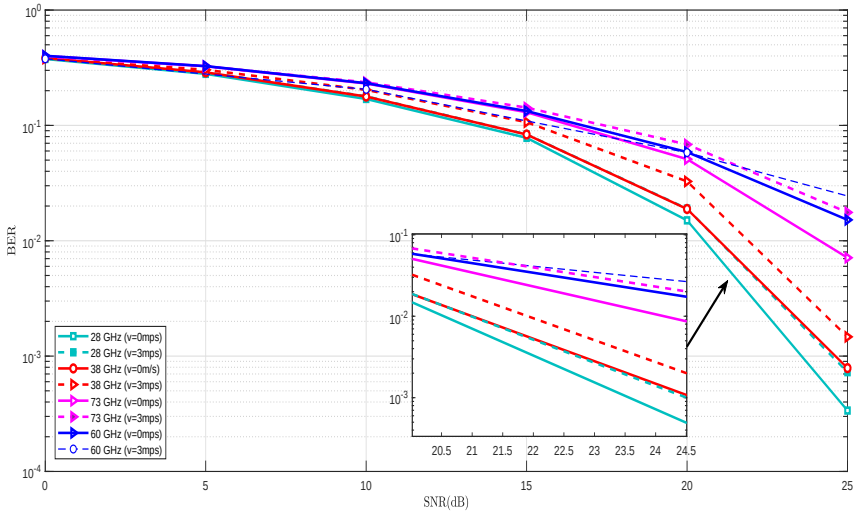


Fig. 12: BER vs. SNR for several frequencies with 128-QAM modulation scheme in semi-complex conference room with S=- 120 dBm

ration scenarios: four configurations for each type of conference room (simple, semi-complex and complex) with (S = -80 / -120 dBm) and (v = 0 / 3mps).

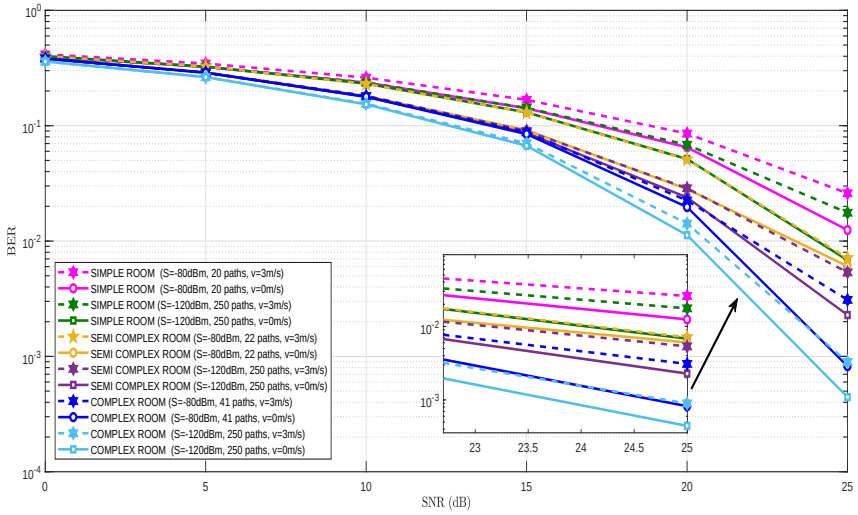


Fig. 13: BER vs. SNR for 128-QAM modulation scheme under several configuration scenarios.

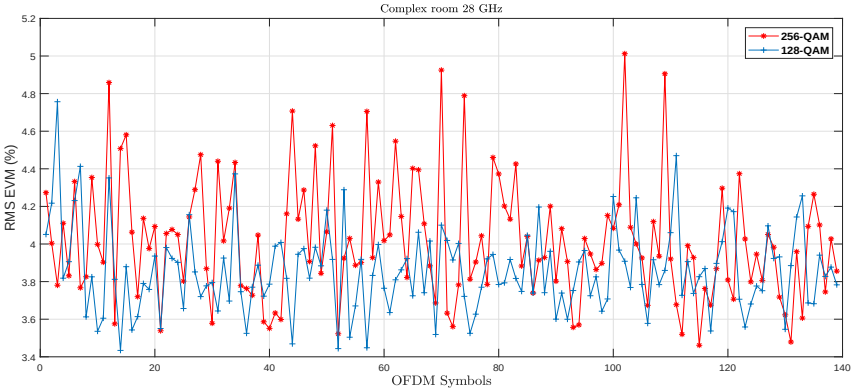
As shown in Fig. 13, the number of paths, movement speed and sensitivity threshold influence the efficiency of the wireless system. Indeed, slow velocity allows for improved BER performance for both receiver sensitivity thresholds -80 dBm and -120 dBm. From this figure it can be seen that worst performance is obtained in simple conference room with receiver sensitivity threshold -80 dBm (20 paths) under (3 m/s) movement speed. On the other hand, best performance is achieved in complex conference room for low receiver sensitivity threshold (-120 dBm) under low movement speed (0 m/s) and high number of paths (250 paths). Interestingly, multipaths can be perceived as a diversity source for the receiver, helping to maximize the average signal strength ratio.

Fig. 14 presents the Root Mean Squares Error Vector Magnitude (RMS EVM) per OFDM symbol for SNR=30 dB with 128 and 256-QAM modulation schemes in complex conference room at (a) 28 GHz and (b) 60 GHz, respectively. In fact, the EVM measurement computes the vector difference between an ideal reference signal and an impaired signal. We can normalize measurements according to average constellation power or peak constellation power [26]. The mean values of the RMS EVM are 3.854, 3.975, 3.866 and 3.992 for 128-QAM (28GHz), 256-QAM (28GHz), 128-QAM (60GHz) and 256-QAM (60GHz), respectively. We can observe that RMS EVM values for 60 GHz are higher than those of 28 GHz for both constellation schemes 128 and 256-QAM. In addition, the maximum values of the RMS EVM confirms this results. We can conclude that transmission with 28 GHz presents less error vector magnitude than 60 GHz.

The simulation experiments listed in table 1 show the performance of our proposed method for several frequencies (28, 38, 60 and 73 GHz) in a range of SNR varying from 0 dB to 20 dB with a step of 5 dB in terms of NMSE (Normalized Mean Squares Error) for 128-QAM modulation scheme under receiver sensitivity threshold = -120 dBm with empty, simple, semi-complex and complex conference rooms scenarios. It can be easily seen that, for each frequency, the best NMSE value is obtained for high SNR (SNR=20 dB). For complex conference rooms, with SNR=20 dB, the worst NMSE value (6.010e-06) is achieved at 60 GHz due to the propagation attenuation properties at this frequency band. We conclude that NMSE performance confirms the obtained results in Fig. 12.

5 Conclusions

This article proposes an innovative sparse channel estimator for mmWave transmission based on Dual Tree Discrete Wavelet Transform Twin Support



(a)

(b)

Fig. 14: RMS EVM per OFDM symbol for SNR=30 dB with 128 and 256-QAM modulation schemes in complex conference room at (a) 28 GHz and (b) 60 GHz, respectively.

Vector Regression adapted to OFDM system structure. In the training phase, the proposed method is based on Dual Tree Discrete Wavelet Transform in order to denoise training signals in time and frequency domains simultaneously, then perform Twin SVR procedure during the estimation phase. The real channel impulse response and path loss values are obtained by the "Wireless InSite" ray tracer by Remcom. The proposed approach showed large and small scales path loss analysis for dense 5G Indoor Hotspot environments with simple, semi-complex and complex conference room scenarios at several mmWave frequencies 28, 38, 60 and 73 GHz. We contrasted our algorithm to popular literature techniques and showed the effectiveness of our method in terms of BER, NMSE and RMS EVM metrics. Thus, several configuration experiments were investigated and discussed.

Table 1: Normalized MSE values for 128-QAM modulation scheme under receiver sensitivity threshold = -120 dBm.

SNR (dB)	0	5	10	15	20
			f=28 GHz		
Empty	0.537e-03	0.171e-03	5.539e-05	1.712e-05	5.323e-06
Simple	0.559e-03	0.175e-03	5.608e-05	1.731e-05	5.356e-06
Semi-complex	0.571e-03	0.177e-03	5.728e-05	1.763e-05	5.602e-06
Complex	0.572e-03	0.181e-03	5.693e-05	1.786e-05	5.726e-06
			f=38 GHz		
Empty	0.560e-03	0.175e-03	5.559e-05	1.729e-05	5.661e-06
Simple	0.566e-03	0.176e-03	5.594e-05	1.766e-05	5.707e-06
Semi-complex	0.568e-03	0.178e-03	5.671e-05	1.784e-05	5.756e-06
Complex	0.573e-03	0.182e-03	5.760e-05	1.790e-05	5.777e-06
			f=60 GHz		
Empty	0.555e-03	0.174e-03	5.436e-05	1.786e-05	5.423e-06
Simple	0.557e-03	0.180e-03	5.648e-05	1.788e-05	5.545e-06
Semi-complex	0.566e-03	0.181e-03	5.770e-05	1.793e-05	5.567e-06
Complex	0.583e-03	0.191e-03	5.817e-05	1.837e-05	6.010e-06
			f=73 GHz		
Empty	0.549e-03	0.173e-03	5.433e-05	1.705e-05	5.543e-06
Simple	0.555e-03	0.179e-03	5.637e-05	1.749e-05	5.610e-06
Semi-complex	0.562e-03	0.180e-03	5.663e-05	1.805e-05	5.710e-06
Complex	0.582e-03	0.187e-03	5.762e-05	1.821e-05	5.816e-06

6 Declarations

On behalf of all authors, the corresponding author states that there is no conflict of interest.

7 Data availability statement

My manuscript has no associated data.

References

1. Niu, Y., Li, Y., Jin, D., Su, L., Vasilakos, A.: A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges. Wireless Networks **21**, 2657–2676 (2015)

2. T.S. Rappaport, al.: Millimeter-wave mobile communications for 5G cellular: It will work! IEEE Access **2013**, 335–349 (2013)

3. Fan, Y., Zhang, Z., Li, H.: Message passing based distributed learning for joint resource allocation in millimeter wave heterogeneous networks. IEEE Transactions on Wireless Communications **18**(15), 2872–2885 (2019)

4. Liu, P., Di Renzo, M., Springer, A.: Variable-N_u generalized spatial modulation for indoor los mmwave communication: performance optimization and novel switching structure. IEEE Transactions on Communications **65**(6), 2625–2640 (2017)

5. M. Mezzavilla, et al.: End-to-End Simulation of 5G mmWave Networks. IEEE Communications Surveys Tutorials **20**(3), 2237–2263 (2018)

6. Rath, H., Timmadasari, S., Panigrahi, B., Simha, A.: Realistic indoor path loss modeling for regular wifi operations in india. In: 2017 Twenty-third National Conference on Communications (NCC), Chennai, pp. 1–6 (2017)

7. Liu, P., Di Renzo, M., Springer, A.: Line-of-Sight spatial modulation for Indoor mmwave communication at 60 GHz. IEEE Transactions on Wireless Communications **15**(11), 7373–7389 (2016)

8. Liu, P., Blumenstein, J., Perovic, N., Di Renzo, M., Springer, A.: Performance of generalized spatial modulation MIMO over measured 60 GHz Indoor channels. IEEE Transactions on Communications **66**(1), 133–148 (2018)

9. F. Fuschini, et al.: Analysis of In-Room mm-Wave propagation: directional channel measurements and ray tracing Simulations. J. Infrared Milli Terahz Waves **38**, 727–744 (2017)

10. Sengupta, A., Alvarino, A., Catovic, A., Casaccia, L.: Cellular terrestrial Broadcast-Physical layer evolution from 3GPP Release 9 to Release 16. IEEE Transactions on Broadcasting **66**(2), 459–470 (2020)

11. F. Gu, et al.: A universal channel estimation algorithm based on DFT smoothing filtering. IEEE Access **7**, 33–39 (2019)

12. Neumann, D., Wiese, T., Utschick, W.: Learning the MMSE channel estimator. IEEE Transactions on Signal Processing **66**(11), 2601–2613 (2018)

13. X Liu, et al.: BEM-PSP for single-carrier and SC-FDMA communication over a doubly selective fading Channel. IEEE Transactions on Wireless Communications **19**(6), 3924–3937 (2020)

14. Zhang, M., Zhou, X., Wang, C.: A novel noise suppression channel estimation method based on adaptive weighted averaging for OFDM systems. Symmetry **11**(8), 33–42 (2019)

15. Rastogi, R., Sharma, S., Chandra, S.: Robust parametric Twin Support Vector Machine for pattern classification. Neural Process Lett. **41**(1), 293–323 (2017)

16. Melki, G., Cano, A., Kecman, V., Ventura, S.: Multi-target support vector regression via correlation regressor chains. Information Sciences **415-416**, 53–69 (2017)

17. Wu, Q., Zhang, H., Jing, R., Li, Y.: Feature selection based on twin support vector regression. In: 2019 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 2903–2907 (2019)

18. Shen, X., Niu, L., Qi, Z., Tian, Y.: Support vector machine classifier with truncated pinball loss. Pattern Recognition **68**, 199–210 (2017)

19. Balasundaram, S., Meena, Y.: K-nearest neighbor-based weighted twin support vector regression. Appl. Intell. **41**(1), 299–309 (2014)

20. Charrada, A., samet, A.: Application of TSVR algorithm in 5G mmWave indoor networks. Wireless Networks **27**(2), 1491–1502 (2021)

21. I. W. Selesnick, R.G.B., Kingsbury, N.C.: The dual-tree complex wavelet transform. IEEE signal processing magazine **22**(6), 123–151 (2005)

22. Charrada, A., Samet, A.: Fast-Fading channel environment estimation using Linear Minimum Mean Squares Error-Support Vector Regression. Wireless Personal Communications **106**, 1897–1913 (2019)

23. Tehrani Moayyed, M.: Channel impulse response for mmwave communication. https://www.github.com/NEU-MathWorks-mmWaveProject/Channel-Impulse-Response, GitHub. (Retrieved May 12, 2020)

24. MacCartney, G., Rappaport, T., Sun, S., Deng, S.: Indoor office wideband millimeter-wave propagation measurements and channel models at 28 GHz and 73 GHz for ultra-dense 5G wireless networks. IEEE Access **3**, 2388–2424 (2015)

25. S. Sun, et al.: Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5G wireless communications. IEEE Transactions on Vehicular Technology **65**(5), 2843–2860 (2016)

26. D. Carrera, et al.: Comparative study of channel estimators for massive MIMO 5G NR systems. IET Communication **14**(7), 1175–1184 (2020)