


Super resolution passive radars based on 802.11ax Wi-Fi signals for human movement detection

Hasan Can Yildirim^{1,2}  | Jean-François Determe¹ | Laurent Storrer¹ |
François Rottenberg^{1,2} | Philippe De Doncker¹ | Jérôme Louveaux² | François Horlin¹

¹OPERA - WCG, Université Libre de Brussels, Brussels, Belgium

²ICTM - ELEN, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Correspondence

Hasan Can Yildirim, OPERA WCG UA3-215a, Ave Franklin Roosevelt 50, 1050, Brussels, Belgium.
Email: hayildir@ulb.ac.be

Abstract

Passive Radars based on Wi-Fi signals provide an excellent opportunity for human sensing without violating the privacy of individuals. Due to the limited integration time of Wi-Fi bursts and relatively low bandwidths, Fourier Transform-based methods do not provide the required accuracy. Herein, a Wi-Fi-based passive radar algorithm is proposed for indoor human movement detection with super resolution which relies on the ESPRIT algorithm to estimate range/speed parameters from limited number of measurements. To determine the number of targets in the environment, a new Model Order Selection (MOS) method is proposed which exploits the orthogonality between the basis vectors of signal and noise subspaces obtained from the sample covariance matrix of the measurements. The new MOS method along with the proposed algorithm are numerically analysed and compared with other existing methods. Finally, the performance of the algorithm is experimentally validated in indoor conditions.

1 | INTRODUCTION

Passive Radars (PR) are devices that make use of existing communication signals as Signal of Opportunity (SO) for detecting and tracking targets in an environment. Thereby, PRs do not emit any signals. It is well-known that radar estimation accuracy largely depends on three parameters [1]: the ambiguity function of the signal, which determines the suitability of the signal for radar processing; the bandwidth, which determines the range resolution; and the integration time, which determines the speed resolution. In recent years, the interest in PRs has increased [2] since communication technologies evolved and wider bandwidths became available to support higher throughput requirements. Wi-Fi 6 (802.11ax) signals [3] are ideal SOs for PRs since they are modulated with Orthogonal Frequency Division Multiplexing (OFDM) whose ambiguity function implies sufficient accuracy [4]. Moreover, the availability of the SOs is increased [5] since the Wi-Fi 6 standard is designed for densely deployed Access Points (AP). Furthermore, the widest

available bandwidth is 160 MHz, which provides roughly a 1-m range resolution.

However, 1-m range resolution is not enough to guarantee a robust range detection of human targets, especially when indoor conditions are considered. Moreover, the Wi-Fi signals consist of bursts of OFDM frames. Hence, the duration of the SO will not be long enough to achieve the desired speed resolution. This is one of the main differences between Wi-Fi and Digital Video Broadcasting (DVB) based PRs [6]. For the latter, the integration time can be adjusted as desired since the signal is always available due to its broadcasting nature. Moreover, approximately 0.5 m/s speed resolution is required to estimate the human main-body speed with an average velocity of 1.4 m/s. Thus, the integration time has to be at least 50 ms at 5.6 GHz with classical Discrete Fourier Transform (DFT)-based radar processing. This is an unrealistic assumption for Wi-Fi systems since a typical Wi-Fi burst is a few ms long. These two problems, regarding the limitations on the range and speed resolutions, are discussed within the recently formed Task

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *IET Radar, Sonar & Navigation* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

Group for Wi-Fi Sensing [7], that aims at indoor monitoring of human movements based on the Wi-Fi signals. In Ref. [8] different Wi-Fi protocols are considered as SO to estimate range/Doppler parameters as well as breathing response of humans with classical DFT-based processing.

In this work, the limitations of the classical processing methods on the accuracy of the range and speed estimations are addressed. To achieve higher accuracies than DFT-based methods, super-resolution algorithms constitute interesting alternatives. In Ref. [9], by exploiting the unique OFDM frame structure of Radar-Communications systems, the range and Doppler frequencies are estimated with ESPRIT and Least Square (LS) algorithms, respectively. However, the considered target velocities are 80 and 100 km/h, which yield high Doppler frequencies. Moreover, the ESPRIT algorithm requires certain characteristics from the sampled covariance matrix which are not discussed. In Ref. [10], 802.11p signals are used for radar processing, where the delay and Doppler are estimated with ESPRIT. However, the integration time is considered to be around 50 ms, which already provides high accuracy with DFT-based methods. Moreover, the methodology to estimate the number of targets is not mentioned. It is also discussed that, for targets with weak reflectivity (such as humans), it is difficult for ESPRIT to estimate the speeds between -2 and 2 m/s due to the presence of the static clutter. Therefore, a novel algorithm is proposed that estimates the range and speed of the targets with super-resolution. During the estimation stages, if a target with zero (or close-to-zero) speed is estimated, that target is assumed to be static clutter. Hence, its contribution is removed from the measured signal to guarantee a more robust detection.

In order to guarantee robust parameter estimation, subspace based super resolution algorithms require the size of the signal subspace, also known as Model Order Selection (MOS). Regarding the MOS methods that is, estimating the number of targets, the literature is extensive (see Refs. [11] and [12] and the references therein). Maximum a posteriori (MAP) method yields the optimal order selection rule [13]. Other approaches based on the Kullback-Leibler (KL) information criterion exists, which minimise the KL discrepancy between the probability density function (PDF) of the measured data and the PDF of the assumed model [14]. However, in real-life, neither MAP nor KL can be used since the associated PDFs are not available. Akaike Information Criterion (AIC) [15] is another MOS method which requires a set of candidate models. Once AIC values are obtained for each candidate model, the one with the smallest AIC value is selected as the best fit. AIC requires the computation of residual sum of squares (RSS) for all realistic and possible candidate models since the best model is determined among the candidate models. On the other hand, eigenvalue thresholding is proposed in [16], which computes a threshold by iterating through the eigenvalues of the sample covariance matrix by assuming that eigenvalues corresponding to noise subspace have much smaller amplitudes compared to ones corresponding to signal subspace. The contribution of our work is fourfold:

- A new algorithm inspired by ESPRIT is proposed, which provides range and speed estimations with sub-metre and sub-metres-per-second accuracies, respectively. The proposed algorithm requires relatively high signal-to-noise-power-ratio (SNR), which is a valid assumption when indoor scenarios are considered.
- The flexibility of the proposed algorithm allows static clutter removal, making the estimation of human movements more robust.
- A new metric is proposed as a MOS tool to select the correct size of the signal subspace, that is the number of targets in the environment.
- The proposed algorithm is experimentally validated in an indoor scenario by using Software Defined Radios (SDRs).

Moreover, the designed algorithm estimates the range and Doppler frequency of the targets in the environment.

This paper is structured as follows. In Section 2, the 802.11ax Wi-Fi frame structures and modulation parameters are briefly explained, along with the signal model. Then, the classical radar processing stages are reminded. In Section 3, the three stages of the proposed algorithm are explained in detail: the range estimation, MOS for range and the speed estimation. In Section 4, the extension of the algorithm is explained, as well as how our additions enhance its accuracy for more crowded areas. In Section 5, the impact of different parameter choices and the computational complexity of the extended algorithm are discussed. In Section 6, the numerical analysis for various scenarios are provided that demonstrates the accuracy of the algorithm. In Section 7, the proposed algorithm and MOS method are experimentally validated in indoor conditions by using SDRs both for APs and PRs. Finally, in Section 8, the conclusion is drawn.

Some notations

The letters n , m , q and v are used to indicate the fast-time (for range processing), slow-time (for Doppler processing), OFDM subcarrier and the Doppler spectrum bins, respectively. Functions with lowercase letters such as $f(t)$ and $f(m)$, indicate that f is a continuous or discrete function of t or m in the time-domain, respectively. Functions with two variables such as $f[n, m]$ mean that both variables are in the time-domain. Moreover, capital letters $F[q, m]$ are used to indicate that the corresponding equality is written as a function of the subcarrier indices and slow-time. Meanwhile, the calligraphic capital letters $\mathcal{F}[n, v]$ are used to indicate that the corresponding equality is a function of the fast-time and Doppler bins.

Moreover, \mathbf{u} and \mathbf{U} denote a vector and a matrix, respectively. u_i indicates the i -th entry of the vector \mathbf{u} , and U_{ij} indicates the (i, j) -th element of the matrix \mathbf{U} . \mathbf{U}_p denotes the matrix formed by the columns of \mathbf{U} indexed by the vector \mathbf{p} while \mathbf{u}_p denotes the p -th column vector of the matrix \mathbf{U} . Moreover, $[i: i + L]$ denotes a unit-spaced vector of size L , whose elements are the integers between i and $i + L$ where i and L are also integer numbers. Thus, $\mathbf{u}_{i:i+L}$ is defined as the

sub-vector of \mathbf{u} , such that $\mathbf{u}_{i:i+L} = [u_i \dots u_{i+L}]$. Similarly, the sub-matrix $\mathbf{U}_{i:i+L}$ is formed by the columns of \mathbf{U} , such that $\mathbf{U}_{i:i+L} = [\mathbf{u}_i \dots \mathbf{u}_{i+L}]$. When it is necessary, capital letters are used as indices to indicate the size of a matrix for example \mathbf{U}_Q is a square matrix of size Q -by- Q .

Moreover, the matrix \mathbf{I}_Q is defined as the identity matrix of size Q . \mathbf{X}^T , \mathbf{X}^H and \mathbf{X}^+ denote transpose, Hermitian transpose and Moore-Penrose inverse, respectively. The output of the operator $\text{diag}(\mathbf{X}) \in \mathbb{C}^Q$ is the vector whose entries are the diagonal elements of an arbitrary square matrix $\mathbf{X} \in \mathbb{C}^{Q \times Q}$. Similarly, the output of the operator $\text{Diag}(\mathbf{x}) \in \mathbb{C}^{Q \times Q}$ is the diagonal matrix, whose diagonal elements are the entries in the vector $\mathbf{x} \in \mathbb{C}^Q$. The hat symbol is used to denote the result of an estimation for example \hat{x} , $\hat{\mathbf{x}}$ or $\hat{\mathbf{X}}$, whereas tilde symbol is used to indicate that a decision is made on the estimated parameters for example \tilde{x} , $\tilde{\mathbf{x}}$ or $\tilde{\mathbf{X}}$. Moreover, the following pair of words are used interchangeably, since the link within the pairs are simple conversions: the propagation-delay/range and the Doppler-frequency/speed.

2 | SIGNAL MODEL

2.1 | 802.11ax Wi-Fi Frame Structure and Signal

The newest amendment of the Wi-Fi standard, 802.11ax, is designed to be used in dense deployment scenarios. A detailed overview of the 11ax standard can be found in Ref. [5]. The new OFDM parameters which are relevant for this work are summarised in Table 1 and the effects of these parameters on the PR performance are summarised here below:

- The highest carrier frequency (f_c) is 5 GHz, which yields higher Doppler shifts for the same target velocities.
- The maximum allowed bandwidth (B) is 160 MHz, increasing the range resolution.
- The number of subcarriers (Q) is increased by a factor of 4 for each bandwidth compared to the previous amendment of the standard. Therefore, the OFDM symbol duration is also increased by a factor of 4.

On the other hand, it is well known that the Wi-Fi APs work by transmitting bursts of OFDM frames, as plotted in Figure 1. Each frame consists of training symbols (known as the preamble) and data symbols. The training symbols are used to synchronize the devices, as well as to estimate the channel parameters. Therefore, these symbols are fixed by the standard, and are known a priori.

Let us define M_s and M_f as the number of OFDM symbols in one frame and the total number of transmitted frames by the AP in one burst, respectively. The transmitted signal, composed of $M = M_s M_f$ OFDM symbols, can be written as

$$s(t) = \sum_{m=0}^{M-1} \sum_{q=0}^{Q-1} X_{q,m} \exp\left(j2\pi \frac{q}{QT} (t - mT_s)\right) u(t - mT_s)$$

where T and T_s represent sample duration and OFDM symbol duration, respectively. Moreover, the element $X_{q,m}$ is picked from matrix $\mathbf{X} \in \mathbb{C}^{Q \times M}$, whose elements are the complex symbols of the m th OFDM symbol mapped on the subcarrier q . Therefore, the rows and columns of \mathbf{X} represent the different subcarrier indices and OFDM symbols, respectively. Moreover, the rectangle pulse-shaping function is defined as

$$u(t) = \begin{cases} 1, & 0 \leq t < T_s \\ 0, & \text{elsewhere} \end{cases}$$

Let P be the total number of propagation paths in the channel. The vectors $\boldsymbol{\alpha} \in \mathbb{C}^P$, $\boldsymbol{\tau} \in \mathbb{R}^P$ and $\mathbf{f} \in \mathbb{R}^P$ contain the path gains, propagation delays, and Doppler frequencies each associated with a path, such that

$$\begin{aligned} \boldsymbol{\alpha} &= [\alpha_1 \ \alpha_2 \ \dots \ \alpha_P] \\ \boldsymbol{\tau} &= [\tau_1 \ \tau_2 \ \dots \ \tau_P] \\ \mathbf{f} &= [f_1 \ f_2 \ \dots \ f_P] \end{aligned} \quad (1)$$

The time varying channel impulse response (CIR) can be written as

$$h(\tau, t) = \sum_{l=1}^P \alpha_p \delta(\tau - \tau_p) \exp(-j2\pi f_p t)$$

Here, the path gain and the propagation delay are considered as constants during the integration time, MT_s , which is in the order of a few milliseconds. This is a realistic assumption for the signals reflected from moving humans whose average speed is 1.5 m/s. Since the path gain and the propagation delay are functions of the path distance, they cannot significantly vary for the given integration time and speed. The received signal can be modelled as the convolution of the transmitted signal and the CIR, which can be written as

$$r(t) = \sum_{p=1}^P \alpha_p \left(\sum_{m=0}^{M-1} \sum_{q=0}^{Q-1} X_{q,m} \exp\left(j2\pi \frac{q}{QT} (t - mT_s - \tau_p)\right) \right) \exp(-j2\pi f_p t) u(t - mT_s - \tau_p) + z(t)$$

where $z(t)$ is the Additive White Gaussian Noise (AWGN) with zero-mean and variance σ^2 . Moreover, after the received signal is low pass filtered, it is sampled at instants $t = nT + mT_s$ where $n = -Q_{cp}, \dots, Q - 1$ and $m = 0, \dots, M - 1$. It is given by

$$r[n, m] := r(t = nT + mT_s)$$

The sampled windowing function, $u[n + (m' - m)] := u(nT + (m' - m)T_s - \tau_p) = \delta_{m,m'}$ for $nT + (m' - m)T_s < M * Q$. Moreover, if $Q_{cp}T > \tau_p, \forall p$ (the CP duration is longer than all propagation delays), the CP will contain the inter-block-interferences (IBI). Since the useful parts of

	Parameters	Values	Units
Carrier frequency	f_c	2.4 - 5	GHz
Bandwidth	B	20 - 40 - 80 - 160	MHz
Sampling duration	$T = 1/B$	50 - 25 - 12.5 - 6.25	Ns
Number of subcarriers	Q	256 - 512 - 1024 - 2048	
Number of active			
Subcarriers	Q_a	242 - 484 - 996 - 1992	
Subcarrier spacing	Δf	78,125	Hz
Cyclic prefix (CP) size	Q_{cp}	16 - 32 - 64	
OFDM symbol duration			
Without the CP	QT	12.8	M_s
OFDM symbol duration			
With the CP	$T_s = (Q + Q_{cp})T$	13 - 13.2 - 13.6	M_s
Number of OFDM			
Symbols in one frame	M_s	30 ...50	
Wi-Fi frame duration	$T_f = M_s T_s$	408 ...800	M_s
Number of Wi-Fi			
Frames in one burst	M_f	2 ...16	

TABLE 1 Key parameters of 802.11ax standard

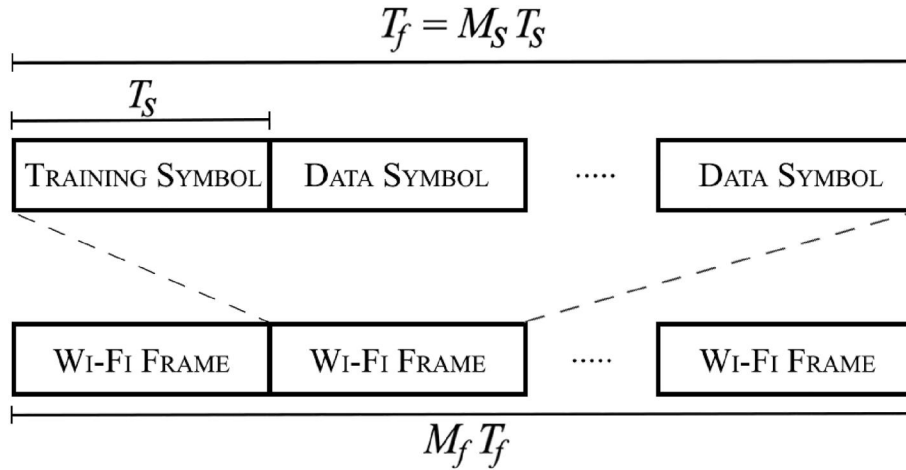


FIGURE 1 Wi-Fi frame structure

the OFDM symbols, obtained by cancelling the CP, are not corrupted by the IBI, the received sequence can be written as

$$r[n, m] = \sum_{p=1}^P \alpha_p \sum_{q=0}^{Q-1} \left(X_{q,m} \exp\left(j2\pi \frac{q}{QT} (nT - \tau_p)\right) \right. \\ \left. \exp(-j2\pi f_p nT) \exp(-j2\pi f_p mT_s) \right) + z[n, m] \quad (2)$$

where $n = 0, \dots, Q - 1$, and $z[n, m] = z(t = nT + mT_s)$ are the independently and identically distributed noise samples. It

is also assumed that the Carrier Frequency Offset (CFO) is already estimated and compensated. When the DFT is applied in order to obtain the OFDM spectrum, the second exponential causes a very small Inter-Carrier-Interference (ICI) and can be neglected.

In order to perform the radar processing, the PR has to reconstruct the transmitted data symbols, so that the channel transfer function (CTF) can be estimated. However, as mentioned earlier, Wi-Fi frames contain priorly known training symbols to estimate the channel parameters. Our recent analysis showed that using only the training field in the preamble is sufficient for radar processing [17], which is validated with experiments.

Let us assume that at the beginning of each frame, there is only one training symbol such that the CTF can be estimated by using only the received samples corresponding to the training symbols. In other words, only the columns of \mathbf{X} corresponding to the training symbols are used. Therefore, m , m in (2) is downsampled by M_s along its second dimension that is $r_d[n, m] = r[n, mM_s]$, leaving only M_f training symbols. For each training symbol, $r_d n, m$ contains the fast-time samples indexed by n and every sample is also identified by a slow-time index m . Instead of using the notation $r_d n, m$, we continue to use $r[n, m]$ for the sake of simplicity. Thus, for the rest of the paper $r[n, m]$ is the downsampled signal with $m = 0, \dots, M_f - 1$. Finally, this processing scheme allows us to model the Wi-Fi frames, transmitted by an AP and received by the PR, as pulse-compression radar signals. The pulse (which is the training symbol) is modulated with OFDM, and the time separation between the two pulses is given by T_f . Therefore, M_f can be considered as the number of pulses where T_f is the pulse repetition interval (PRI).

2.2 | Classical Wi-Fi-based Passive Radar processing

The classical Wi-Fi-based PR processing is based on non-parametric methods such as DFT which does not exploit the structure of the input signal but relies on the definition of the Power Spectral Density (PSD). Therefore, the range and Doppler accuracies are directly limited by the system bandwidth and integration time, respectively.

2.2.1 | Range processing

Once the received signal is filtered, sampled, synchronized, downsampled, moved to the frequency domain and the channel estimation is performed, the estimated CTF can be written as

$$\hat{H}q, m = \sum_{p=1}^P \alpha_p \exp(-j2\pi f_p m T_s) \exp\left(-j2\pi \frac{q\tau_p}{QT}\right) + W[q, m] \quad (3)$$

where $W[q, m]$ is defined as the noise samples divided by the transmitted symbols. The Inverse Discrete Fourier Transform (IDFT) of (3) along the subcarrier indices q , can be written as

$$\begin{aligned} \hat{h}[n, m] &= : \sum_{q=0}^{Q-1} \hat{H}[q, m] \exp\left(j2\pi \frac{qn}{Q}\right), \quad n = 0, \dots, Q-1 \\ &= \sum_{p=1}^P \alpha_p \exp(-j2\pi f_p m T_s) \\ &\quad \sum_{q=0}^{Q-1} \left(\exp\left(-j2\pi \frac{q}{Q} \left(n - \frac{\tau_p}{T}\right)\right) + W[q, m] \right) \end{aligned} \quad (4)$$

$$= \sum_{p=1}^P \alpha_p \exp(-j2\pi f_p m T_s) \Xi\left[n - \frac{\tau_p}{T}\right] + w[n, m] \quad (5)$$

where $\Xi[n]$ is the Dirichlet function and the fast-time index n corresponds to the delay bins, which forms M_f number of delay profiles. Moreover $w[n, m]$ is the IDFT of the noise spectrum, or equivalently, the noise added to the m -th delay profile. Notice that, the second sum in (4) takes its maximum value only if τ_p/T is equal to n . To satisfy this condition, the propagation delay has to be an integer multiple of the sampling interval. If the fractional part of τ_p/T is zero, the IDFT yields the Dirichlet function centred at the corresponding delay bin, while its nulls are exactly on the other delay bins. Therefore, the energy peak corresponding to a target appears only at the corresponding delay bin. Otherwise, the maximum and the nulls of the Dirichlet function lie between two delay bins. Thus, the energy of the target leaks to the other delay bins.

Moreover, the delay profile can be converted to the range profile by simply multiplying each delay bin with the range resolution. Thus, (5) is known as the range/slow-time map, meaning that the range information is extracted from the CTF. Notice that for each object along the range profile that is $\forall n_p$, there is a phase that evolves over the slow time index m . These phase shifts contain the Doppler information for the given range bin, and they can be extracted by computing the DFT along the slow-time index m of (5)

$$\hat{\mathcal{H}}[n, v] = \sum_{m=0}^{M_f-1} \hat{h}[n, m] \exp\left(-j2\pi \frac{vm}{M_f}\right)$$

where $v = \frac{-M_f}{2}, \dots, \frac{M_f}{2} - 1$, yielding

$$\hat{\mathcal{H}}[n, v] = \sum_{p=1}^P \alpha_p \Xi\left[n - \frac{\tau_p}{T}\right] \Xi[v - f_p T_s] + \mathcal{W}[n, v] \quad (6)$$

where $\mathcal{W}n, v$ is the noise spectrum of the Doppler profiles. Here, a similar dilemma as before is faced. Depending on the system configuration and the speed of the object, the Dirichlet function will be sampled either at its maximum and nulls, or it will cause leakage. In either case, (6) is known as the range/Doppler Map (RDM). It can be seen that RDM is a function that associates each object in the environment with its relative distance and velocity. Notice that, after the range and Doppler processing, the AWGN at the receiver $z(t)$, acts like a noise floor on each cell of the RDM, denoted as $\mathcal{W}[n, v]$. In other words, if an object has lower amplitude α_p than the noise floor, the object will not appear on the RDM. Moreover, the noise floor corresponds to the mean of the noise, and the variance of the noise creates additional peaks on the RDM. Therefore, once the RDM is obtained, a detection algorithm is used to separate the real targets from the noise peaks [1].

3 | A NOVEL RANGE AND SPEED ESTIMATION ALGORITHM

The parametric, also known as model-based, methods assume that the signal satisfies a given model. Thus, the parameters of the model can be estimated from noisy measurements. This is essentially why parametric methods can perform better than DFT-based methods. However, the model order has to be carefully selected so that the parameter estimation is reliable. The novelty of the proposed algorithm is twofold. First, the number of targets in the environment is over-estimated and their corresponding range/speed parameters are estimated with ESPRIT. Second, the new MOS method is used to discard the non-real targets after the range estimation, such that the speed estimation is performed only for the real targets. For the sake of clarity, we start by focussing on a scenario where there is only one target, hence one Doppler frequency, at any given range. Then, the algorithm is extended for a more realistic scenario in Section 4.

Let us begin by defining the matrix \mathbf{D} whose columns are the cisoids in the frequency-domain, each associated with the propagation paths $p \in [1, P]$ between the AP and PR

$$\mathbf{D} = \begin{pmatrix} 1 & \dots & 1 \\ \exp\left(-j2\pi\frac{1}{QT}\tau_1\right) & \dots & \exp\left(-j2\pi\frac{1}{QT}\tau_p\right) \\ \vdots & \ddots & \vdots \\ \exp\left(-j2\pi\frac{Q-1}{QT}\tau_1\right) & \dots & \exp\left(-j2\pi\frac{Q-1}{QT}\tau_p\right) \end{pmatrix} \in \mathbb{C}^{Q \times P} \quad (7)$$

Similarly, the Doppler phases can be organised in matrix form

$$\mathbf{\Omega} = \begin{pmatrix} 1 & \dots & \exp(-j2\pi f_1(M_f - 1)T_f) \\ \vdots & \ddots & \vdots \\ 1 & \dots & \exp(-j2\pi f_p(M_f - 1)T_f) \end{pmatrix} \in \mathbb{C}^{P \times M_f} \quad (8)$$

where the rows of the matrix $\mathbf{\Omega}$ contain the Doppler phases which evolve with respect to the Doppler frequency f_p and the column (or slow-time) index m . For easy notation and future reference, let us define

$$\mathbf{B} = \text{Diag}(\boldsymbol{\alpha})\mathbf{\Omega}, \in \mathbb{C}^{P \times M_f} \quad (9)$$

whose rows are obtained after scaling the rows of $\mathbf{\Omega}$ by with respect to the path index p . Finally, the M_f estimated CTFs, given in (3), can be written in matrix form:

$$\widehat{\mathbf{H}} = \mathbf{D}\mathbf{B} + \mathbf{W}, \in \mathbb{C}^{Q \times M_f} \quad (10)$$

where $\widehat{\mathbf{H}}$ contains the channel estimates for M_f frames and Q subcarriers, and it is the input of the algorithm. The elements of \mathbf{W} are the noise samples in (3) that is $W_{q,m} = W[q, m]$.

Figure 2 shows the steps of the algorithm which is composed of three consecutive stages. The first group of blocks, named as Range Estimation, corresponds to the propagation delay estimation with ESPRIT and amplitude estimation with LS, both explained in Section 3.1. Then, the estimated ranges are refined with the new MOS method, described in Section 3.2. Finally, the Doppler frequencies are estimated as explained in Section 3.3.

3.1 | Range estimation

The range estimation stage is grouped in two subsequent stages. First, the partial covariance matrix and its eigenvalue decomposition are obtained for the estimation algorithms. Then, the target ranges, and their associated Doppler vectors, are estimated with ESPRIT and LS, respectively.

3.1.1 | Partial covariance matrix

First, we remind that ESPRIT is a subspace-based method that requires the eigenvalue decomposition of the sample covariance matrix. ESPRIT works with the estimates of basis vectors of the underlying signal [18]. Therefore, the rank of the sample covariance matrix has to be at least P , so that the signal basis vectors contain the necessary information on all propagation delays. Once this condition is satisfied, the signal basis vectors form the corresponding Vandermonde matrix, which can be derived through the Carathéodory parametrisation as shown in [19]. The K th-order sample covariance matrix is computed as follows

$$\mathbf{R}_K = \frac{1}{M_f} \sum_{m=1}^{M_f} \frac{1}{Q-K} \sum_{i=1}^{Q-K} \left[\widehat{\mathbf{H}}_{i:i+K,m} (\widehat{\mathbf{H}}_{i:i+K,m})^H \right] \quad (11)$$

where $\mathbf{R}_K \in \mathbb{C}^{K \times K}$ with $Q > K > P$, and the subscript K is used to indicate the size of the matrix. In words, M_f number of covariance matrices are obtained with a sliding window of size K . Then, all covariance matrices are averaged. This procedure allows the rank of the sample covariance matrix to be more than P , and possibly full-rank, regardless of the rank of the matrix $\widehat{\mathbf{H}}$. This rank requirement is crucial to provide robust parameter estimation with ESPRIT [19]. The eigenvalue decomposition of the matrix \mathbf{R}_K can be written as

$$\mathbf{R}_K = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H \quad (12)$$

where the matrix $\mathbf{V} \in \mathbb{C}^{K \times K}$ includes all the eigenvectors whose span is the signal subspace (with $K - P$ supernumerary noise eigenvectors), while the diagonal matrix $\boldsymbol{\Lambda} \in \mathbb{C}^{K \times K}$ contains the corresponding eigenvalues. The over-estimated number of targets is defined as P_o and it is set before the algorithm is executed. ESPRIT algorithm is applied on the P_o number of eigenvectors associated with the P_o largest eigenvalues where $K > P_o > P$.

3.1.2 | Parameter estimation

Now that the eigenvectors are obtained, the range of each possible target can be estimated with the ESPRIT algorithm, whose output is the vector defined as

$$\hat{\tau} = \text{ESPRIT}(\mathbf{V}_{1:P_o}) \in \mathbb{R}^{P_o} \quad (13)$$

whose elements are the estimated propagation delays. Let us construct the matrix whose elements are the approximated subcarrier-path pairs

$$\hat{\mathbf{D}} = \begin{pmatrix} 1 & \dots & 1 \\ \exp\left(-j2\pi \frac{1}{QT} \hat{\tau}_1\right) & \dots & \exp\left(-j2\pi \frac{1}{QT} \hat{\tau}_{P_o}\right) \\ \vdots & \ddots & \vdots \\ \exp\left(-j2\pi \frac{Q-1}{QT} \hat{\tau}_1\right) & \dots & \exp\left(-j2\pi \frac{Q-1}{QT} \hat{\tau}_{P_o}\right) \end{pmatrix} \in \mathbb{C}^{Q \times P_o} \quad (14)$$

In ideal conditions that is the propagation delays are estimated without any errors due to noise, the matrix \mathbf{D} is a submatrix of $\hat{\mathbf{D}}$. Thus, the goal of the proposed MOS stage is to find a set of column indices corresponding to the true paths. After the matrix $\hat{\mathbf{D}}$ is built, the complex amplitudes corresponding to its columns can be estimated with the LS method

$$\hat{\mathbf{B}} = \hat{\mathbf{D}}^+ \hat{\mathbf{H}}, \in \mathbb{C}^{P_o \times M_f} \quad (15)$$

where the entries in $\hat{\mathbf{B}}$ contain complex amplitude per slow-time index m as well as the Doppler phases that rotates over slow-time.

3.2 | Model Order Selection: range

The new MOS method relies on the orthogonality between basis vectors corresponding to signal and noise subspaces. In the MOS stage, the cisoids corresponding to the estimated propagation delays in (14) are projected to the noise subspace so that the angle between the constructed vector and the noise subspace can be obtained for each possible target delay. The proposed MOS metric corresponds to this angle, as illustrated in Figure 3. If the constructed vector does not belong to the noise subspace, the angle between its orthogonal projection and itself should be high that is it better fits to the signal subspace, than the noise subspace. On the other hand, if the constructed vector does not correspond to a real-path, the corresponding angle should be small. Our target recovery algorithm tends to identify the indices in which the upper-mentioned metric is sufficiently high.

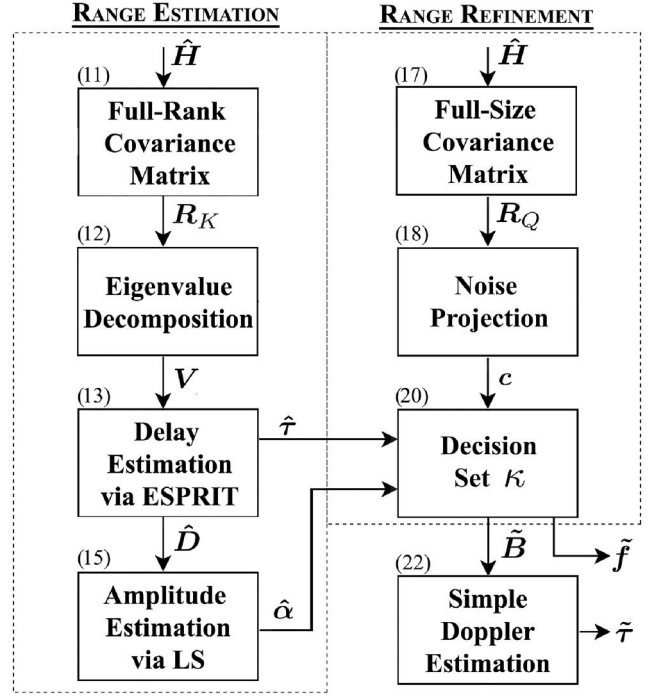


FIGURE 2 The block diagram of the proposed algorithm. The numbers located at the upper-left corner of each block, points to the corresponding equation number of the given block

Therefore, depending on this metric, a given path can be discarded, or kept for Doppler frequency estimation. In order to project the estimated vectors onto the noise subspace, the noise projection matrix is required.

3.2.1 | Full-size covariance matrix: let us define the following column vector, to obtain the noise projection matrix

$$\hat{\mathbf{h}} = \frac{1}{M_f} \sum_{m=1}^{M_f} \hat{\mathbf{h}}_m, \in \mathbb{C}^Q \quad (16)$$

In other words, $\hat{\mathbf{h}}$ is an estimate of the channel coefficients for all subcarriers averaged over M_f frames. Then, another sample covariance matrix can be defined

$$\mathbf{R}_Q = \hat{\mathbf{h}} \hat{\mathbf{h}}^H, \in \mathbb{C}^{Q \times Q} \quad (17)$$

Notice that the maximum possible rank of the sample covariance matrix \mathbf{R}_Q is 1, since only one vector is used to obtain the covariance matrix. This is one of the reasons for the averaging given in (16) (the other reason is to improve the SNR for a more clear separation between the signal and noise subspaces).

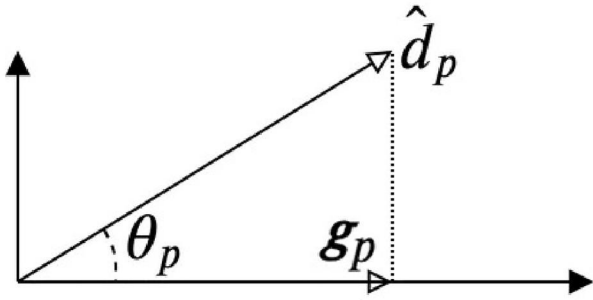


FIGURE 3 The x and y axes represent the noise and signal subspaces, respectively. The constructed vector $\hat{\mathbf{d}}_p$ is projected to the noise subspace, yielding the vector \mathbf{g}_p , and θ_p is used as the MOS metric

3.2.2 | Detection Refinement

Finally, the noise projection matrix can be directly obtained without any eigenvalue decomposition, since the rank of the sample covariance matrix is 1. The constructed cisoids can be projected to the noise subspace, as follows

$$\mathbf{G} = \left(\mathbf{I}_Q - \frac{\mathbf{R}_Q}{\|\hat{\mathbf{h}}\|^2} \right) \hat{\mathbf{D}}, \in \mathbb{C}^{Q \times P_o} \quad (18)$$

where \mathbf{I}_Q is the identity matrix of size Q . The matrix inside the first parenthesis corresponds to the noise projection matrix obtained without eigenvalue decomposition. Thus, the columns of the matrix \mathbf{G} corresponds to the noise projections of the cisoids in $\hat{\mathbf{D}}$. The angle between the columns of \mathbf{G} and $\hat{\mathbf{D}}$ can be obtained as follows

$$\theta_p = \arccos \left(\frac{\hat{\mathbf{d}}_p \mathbf{g}_p}{\|\hat{\mathbf{d}}_p\| \|\mathbf{g}_p\|} \right), \quad p = 1 \dots P_o \quad (19)$$

where θ_p is the p -th entry of the vector defined as $\boldsymbol{\theta} \in \mathbb{R}^{P_o}$. The entries in $\boldsymbol{\theta}$ are used as a metric to decide on the accuracy of an estimation. Let κ be the set defined as

$$\kappa := \{p, \theta_p > \mathcal{T}_d\} \quad (20)$$

where the cardinality of the set κ is \tilde{P} . Moreover, \mathcal{T}_d is a numerically obtained threshold which mainly depends on the SNR of the echo signals (see Section 6] for further discussion about the threshold).

In other words, all entries in the vector $\boldsymbol{\theta}$ are compared with the threshold. If the entry θ_p is above the threshold, its corresponding index p is placed in the decision set κ . The MOS stage is finalised by keeping the reliable estimations

$$\tilde{\boldsymbol{\tau}} = \hat{\boldsymbol{\tau}}_{\kappa}, \in \mathbb{R}^{\tilde{P}} \quad \tilde{\mathbf{D}} = \hat{\mathbf{D}}_{\kappa}, \in \mathbb{C}^{\tilde{P}} \quad \tilde{\mathbf{B}} = \hat{\mathbf{B}}_{\kappa}, \in \mathbb{R}^{\tilde{P}} \times M_f \quad (21)$$

where the tilde symbol is used to indicate that the corresponding vectors and matrices are obtained after the MOS stage.

3.3 | Speed estimation

As mentioned previously, the Doppler information is available in the rows of the matrix $\tilde{\mathbf{B}}$. Let us define the vector $\tilde{\mathbf{f}} \in \mathbb{R}^{\tilde{P}}$ whose entries are the estimated Doppler frequencies. Its entries can be obtained by the following equality

$$\tilde{f}_p = \frac{1}{M_f} \sum_{m=1}^{M_f} -\frac{\angle \tilde{\mathbf{B}}_{p,m}}{2\pi m T_f}, \quad p = 1, \dots, \tilde{P} \quad (22)$$

where \tilde{f}_p is the p -th entry of the vector $\tilde{\mathbf{f}}$. However, this method only works when there is only one Doppler frequency at the corresponding range. In other words, the matrix model given in (10), assumes that each row vector of \mathbf{B} has only one Doppler frequency for the corresponding propagation delay, or equivalently the corresponding column in \mathbf{D} . Clearly, this is not a realistic assumption for real channels where there can be multiple Doppler frequencies at each range. If there are multiple targets at a given range with different speeds, their contributions to the rows of \mathbf{B} , hence $\tilde{\mathbf{B}}$, would be combined. This additional and necessary condition requires an extension to the algorithm.

4 | EXTENSION OF THE ALGORITHM

In this section, the focus is on a more realistic scenario where multiple targets with different Doppler frequencies are present at any given range. Thus, the Simple Doppler Estimation block in Figure 2 is replaced by another layer of super-resolution parameter estimation. Before modelling the scenario and explaining the algorithm, let us make the link between the super-resolution range and speed estimations. For range estimation, the number of propagation delays are over-estimated. Then, their corresponding propagation delays are estimated from the columns of $\hat{\mathbf{H}}$. Afterwards, the matrix $\hat{\mathbf{D}}$ is constructed, with supernumerary propagation delays to select the real-paths with the new MOS method. Similarly, the number of Doppler frequencies can be over-estimated and their corresponding Doppler frequencies can be estimated from $\tilde{\mathbf{B}}$ with ESPRIT. The estimate of the matrix $\boldsymbol{\Omega}$, namely $\hat{\boldsymbol{\Omega}}$, can be constructed based on the estimated Doppler frequencies to be used in the MOS stage for speed. The corresponding block diagram is given in Figure 4.

Without losing any generality, let us assume that there are L number of objects in the environment, all at the same distance, causing the same propagation delays $\tau_o = \tau_l, \forall l$. Furthermore, let us assume that the matrix $\hat{\mathbf{D}}$ is successfully constructed, and

$\hat{\mathbf{B}}$ is estimated. After the range MOS stage, the decision set κ contains only one entry since all the objects cause the same propagation delay. That entry corresponds to the index of the column defined as $\tilde{\mathbf{d}} = \hat{\mathbf{D}}_{\kappa}$, such that

$$\tilde{\mathbf{d}} = \left[\exp\left(-j2\pi\frac{1}{QT}\tau_0\right) \dots \exp\left(-j2\pi\frac{Q-1}{QT}\tau_0\right) \right]^T \in \mathbb{C}^Q \quad (23)$$

Similarly, $\tilde{\mathbf{b}} = (\tilde{\mathbf{B}}_{\kappa,:})^T$ is the column vector obtained after the range MOS (21), and its entries are

$$\tilde{b}_m = \sum_{l=1}^L \alpha_l \exp(-j2\pi f_l T_f (m-1)), \quad m = 1 \dots M_f \quad (24)$$

In other words, the column vector $\tilde{\mathbf{b}}$ contains the sum of different Doppler cisoids, since all objects are located at the same range. Clearly, the equation given in (22) cannot be used to estimate the frequencies in (24).

As mentioned previously, in this section it is assumed that all the targets cause the same propagation delays. Therefore, the speed estimation and MOS stages will be explained only on the column vector $\tilde{\mathbf{b}}$. However, in reality, there can be multiple true ranges in the channel. In that case, the speed estimation and MOS stages will be repeated for each column of the matrix $(\tilde{\mathbf{B}})^T$ that is for each unique range, hence columns of $\tilde{\mathbf{D}}$.

4.1 | Speed estimation

Let us define the covariance matrix used for the speed estimation

$$\mathbf{R}_{L_o} = \frac{1}{M_f - L_o} \sum_{i=1}^{M_f - L_o} \left[\tilde{\mathbf{b}}_{i:i+L_o} (\tilde{\mathbf{b}}_{i:i+L_o})^H \right] \in \mathbb{C}^{L_o \times L_o} \quad (25)$$

where the value of L_o should satisfy the following condition to make sure that \mathbf{R}_{L_o} satisfies the rank condition: $L_o > M_f/2$. The eigenvalue decomposition of the matrix \mathbf{R}_{L_o} can be written as

$$\mathbf{R}_{L_o} = \mathcal{V} \Psi \mathcal{V}^H \quad (26)$$

where the matrix $\mathcal{V} \in \mathbb{C}^{L_o \times L_o}$ contains the eigenvectors associated with the eigenvalues in the diagonal matrix $\Psi \in \mathbb{C}^{L_o \times L_o}$. Notice that, the covariance matrix defined in (25) and its eigenvalue decomposition in (26) are the Doppler frequency counterparts of (11) and (12), respectively. Therefore, the columns of the matrix \mathcal{V} can be used to estimate the Doppler frequencies with ESPRIT. In contrast to the range estimation, all the columns of the matrix \mathcal{V} are used, since L_o is already small due to the restrictions on M_f . Therefore L_o is also the over-estimated number of Doppler frequencies obtained through ESPRIT

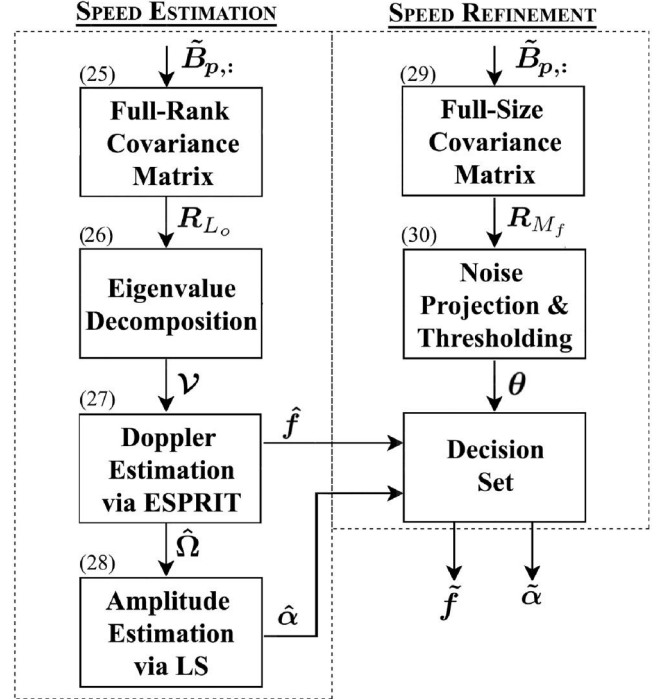


FIGURE 4 The block diagram of the proposed algorithm. The numbers located at the upper-left corner of each block, points to the corresponding equation number of the given block

$$\hat{\mathbf{f}} = \text{ESPRIT}(\mathcal{V}), \in \mathbb{R}^{L_o} \quad (27)$$

Moreover, the matrix $\hat{\mathbf{\Omega}} \in \mathbb{C}^{L_o \times M_f}$ can be built, which is the Doppler frequency counter-part of $\hat{\mathbf{D}}$, and its rows are defined as

$$\hat{\omega}_{l,:} = \left[1 \exp\left(-j2\pi\hat{f}_l T_f\right) \dots \exp\left(-j2\pi\hat{f}_l (M_f - 1) T_f\right) \right], \in \mathbb{C}^{M_f}$$

where $\hat{\omega}_{l,:}$ is the l -th row of $\hat{\mathbf{\Omega}}$. Finally, the amplitude of each Doppler cisoid can be estimated with LS

$$\hat{\alpha} = (\hat{\mathbf{\Omega}}^T)^+ \tilde{\mathbf{b}}, \in \mathbb{C}^{L_o} \quad (28)$$

where the entries in vector $\hat{\alpha}$ are the complex amplitudes per estimated Doppler frequency.

4.2 | Model Order Selection: speed

In order to refine the estimated Doppler frequencies, one last covariance matrix is defined as follows

$$\mathbf{R}_{M_f} = \tilde{\mathbf{b}} \tilde{\mathbf{b}}^H, \in \mathbb{C}^{M_f \times M_f} \quad (29)$$

TABLE 2 Complexity of different algorithms

	Equation	Time Complexity	Number of Multiplications
\mathbf{R}_K	(11)	$\mathcal{O}((Q-K)K^2M_f)$	$10^{9.06}$
\mathbf{R}_Q	(17)	$\mathcal{O}(Q^2)$	$10^{6.02}$
\mathbf{V}	(12)	$\mathcal{O}(K^3)$	$10^{7.52}$
$\hat{\tau}$	(13)	$\mathcal{O}(P_o^3 + P_o^2K)$	$10^{5.93}$
$\hat{\mathbf{B}}$	(15)	$\mathcal{O}(P_oQ(P_o + M_f))$	$10^{6.5}$
\mathbf{G}	(18)	$\mathcal{O}(Q^2P_o)$	$10^{7.7}$
\mathbf{R}_{L_o}	(25)	$\mathcal{O}(\tilde{P}(M_f - L_o)L_o^2)$	$10^{4.31}$
\mathbf{R}_{M_f}	(29)	$\mathcal{O}(\tilde{P}M_f^2)$	$10^{4.01}$
ν	(26)	$\mathcal{O}(\tilde{P}L_o^3)$	$10^{4.31}$
$\hat{\mathbf{f}}$	(27)	$\mathcal{O}(\tilde{P}L_o^3)$	$10^{4.31}$
$\hat{\alpha}$	(28)	$\mathcal{O}(\tilde{P}(L_o^3 + L_o^2M_f))$	$10^{4.79}$
\mathcal{G}	(30)	$\mathcal{O}(\tilde{P}M_f^2L_o)$	$10^{4.91}$

where \mathbf{R}_{M_f} is rank 1. Hence, similar to (17) and (18), the Doppler cisoids can be projected to the noise subspace as follows

$$\mathcal{S} = \left(\mathbf{I}_{M_f} - \frac{\mathbf{R}_{M_f}}{\|\hat{\mathbf{b}}\|^2} \right) \left(\text{Diag}(|\hat{\alpha}|) \hat{\boldsymbol{\Omega}} \right)^T, \in \mathbb{C}^{M_f \times L_o} \quad (30)$$

where \mathbf{I}_{M_f} is the identity matrix of size M_f and $|\hat{\alpha}|$ is the vector whose entries are the absolute values of the entries in $\hat{\alpha}$. Similarly to the range MOS stage, the MOS metric can be obtained through the definition given in (19) that is, by computing the angle between columns of \mathcal{S} and $\hat{\boldsymbol{\Omega}}^T$. In order to decide whether the path l is reliable or not, each angle is compared against a numerically obtained threshold as shown in (20). Once the decision set with cardinality \tilde{L} is obtained, the speed MOS stage yields the vectors $\hat{\mathbf{f}}$ and $\hat{\alpha}$.

Moreover, obtaining the amplitude and Doppler frequency pairs for a given range allows to filter out the static clutter. For human movement detection, the static clutter creates a crucial problem since the energy of the clutter echoes is usually higher than the echoes from humans [20]. Thus, human echoes can be concealed by the static clutter echoes. Once the amplitude of the clutter is estimated in (29), its contribution can be removed from the signal given in (24), and the estimation stages, described between (25) and (29), can be repeated to increase the estimation accuracy of the remaining, non-clutter objects. Finally, the dynamic targets in most indoor scenarios can be characterised by their micro-Doppler signatures on the range-Doppler maps [21]. However, the considered integration time in this work roughly corresponds to 6 – 8 ms, which is quite low for extracting the micro-Doppler information.

5 | COMPUTATIONAL COMPLEXITY ANALYSIS AND PARAMETER DEFINITIONS

In this section, the computational complexity of the proposed algorithm and MOS method is analysed by mainly focussing on the most resource demanding computations, such as matrix multiplications, eigenvalue decompositions, and pseudo-inverses. Moreover, comments on how to fix the parameters of the algorithm are provided. Table 2 shows the computational complexity for the extended version of the algorithm. The first and second columns of the table correspond to the output of the computation and the equation where the output is defined, respectively. The third column indicates the time complexities with Big-O notation [22,23]. The last column shows the estimated number of operations for fixed values of the parameters (e.g., P_o , K , L_o). On the other hand, the processing time in seconds can also be obtained to assess the complexity of the algorithm. However, such a study would be a very unique case since the time complexity depends on the implementation level (MATLAB, C/C++, HDL, etc.), type of the processing unit, and many variables. Therefore, the complexity analysis is only focused on the Big-O notation. In the remainder of this section, the range-group refers to the computations that appear in the range estimation and MOS stages, which are the first seven rows of Table 2. Similarly, the speed-group refers to the last seven rows, hence the speed estimation and MOS stages.

First of all, the condition $Q - K > K$ has to be satisfied so that the rank requirement of \mathbf{R}_K in (11) is also satisfied. However, if $P_o > K$ then the eigenvectors in (12) will not have the Vandermonde structure, and the propagation delays will not be estimated accurately [19]. The time complexity of \mathbf{R}_Q is determined only by the number of subcarriers Q . The eigenvectors in \mathbf{V} are obtained from the eigenvalue decomposition of the covariance matrix \mathbf{R}_K . Hence, the corresponding complexity is proportional only to the parameter K . The entries in the vector $\hat{\tau}$ are estimated by the ESPRIT algorithm which involves an LS estimation and an eigenvalue decomposition. Hence, the complexity of $\hat{\tau}$ depends on K and the over-estimated number of propagation delays P_o . Notice that, setting $P_o < P$ entails miss detections. On the other hand, setting P_o to a high value for example, $P_o = 200$, is not a realistic assumption since it implicitly means that 200 unique ranges are expected in the environment. This also increases the time consumption in MOS step since 200 range estimations have to be tested. Moreover, the matrix $\hat{\mathbf{B}}$ is obtained through the LS algorithm, and the parameters that determines its complexity are P_o and M_f . Finally, the complexity of the matrix multiplication for the noise projection is mainly determined by Q and P_o .

The processes that appear in the range-group are equivalent to the processes that appear in the speed-group with only two differences. First, the size of the matrices and vectors in the speed-group are much smaller than their counterparts in the range-group. In the speed-group, the number of available measurements M_f determines the size of the matrices and vectors, and M_f is assumed to be a

relatively small number for example $M_f \leq 32$. Second, all the computations in the speed-group are repeated for each uniquely estimated propagation delay that is for each column of \mathbf{D} . Therefore, their complexities scale with the parameter \tilde{P} .

In order to quantitatively compare the complexities, the following values are set based on Table 1 for all the upper-mentioned parameters: $Q = 1024$, $K = 320$, $M_f = 16$, $L_o = 8$ and $P_o = 48$. Finally, the number of unique ranges is set to $\tilde{P} = 40$ for demonstrative purposes. There are a few observations that can be made regarding the complexities.

- The three most demanding computations are in the range-group. Since $Q \approx 10^3$, it is not a surprise that the three most demanding computations are functions of Q as well.
- Since Q is not controlled by the PR designer, and K is set automatically as a function of Q , the only solution to reduce the time consumption is to optimise the corresponding computations.
- If the signals transmitted by the AP has lower bandwidth, Q will be a smaller value, as shown in Table 1. Hence, the complexity of the corresponding computations will be lowered at the cost of losing estimation accuracy.
- Even though the computations in the speed-group are repeated \tilde{P} times, the range-group still requires most of the resources, which means that the main optimization is required in the range-group.
- Since M_f and hence L_o are limited by the nature of Wi-Fi signals. Therefore, the only parameter that can significantly vary in the speed-group is \tilde{P} .

6 | NUMERICAL ANALYSIS

This section is divided into two subsections. First subsection aims at showing the statistical consistency of the proposed MOS metric. Moreover, the threshold defined in (20) and its relation to SNR conditions are also discussed. The second subsection starts by defining the various scenarios that will be used to compare the proposed MOS method with other MOS methods.

6.1 | MOS metric and threshold

In order to demonstrate the statistical relevance of the proposed metric and to find the related threshold, a target is simulated assuming a varying distance and the system parameters are summarised in Table 3. For each target distance 1000 realisations are executed. After each realization, the entries in $\boldsymbol{\theta} \in \mathbb{R}^{P_o}$ (19) are sorted in decreasing order. Since there is only one target, θ_1 and θ_2 are the metrics corresponding to the real and non-real paths. In other words, θ_1 and θ_2 are the angles between the constructed vectors corresponding to the real and non-real paths respectively, and their orthogonal projections on the noise subspace. More importantly, θ_2 corresponds to the spurious noise

target whose constructed vector is closest to the signal subspace since the entries in $\boldsymbol{\theta}$ are sorted.

In Figure 5, the probability density functions corresponding to θ_1 and θ_2 parameters are shown for a fixed distance and different M_f values. When the number of available measurements M_f increases, the sample covariance matrix in (11) will be closer to its noise-free counter-part due to the averaging. This has two outcomes: i) the estimation accuracy is improved; ii) the proposed MOS method becomes more reliable. The standard deviation of the metrics corresponding to the non-real paths is very small that is the metric is well concentrated around its mean for spurious noise targets. On the other hand, the underlying distribution of the metric corresponding to the real-path is highly affected by the variation of M_f : when M_f increases, standard deviation decreases and mean increases, meaning that the reliability of the metric increases along with the estimation accuracy.

In Figure 6, 95% confidence regions of the underlying normal distributions are shown for the θ_1 values. On the other hand, only the means are shown for the non-real paths, since their variations are very small. It is clear that the threshold for the metric can be set to different values as a function of the detected target distance for example, if the estimated target distance is below 15 m, the threshold can be set to 20° , else it can be set to 10° . Moreover, regardless of the number of measurements and the target distance, θ_2 does not vary significantly since the thermal noise power is constant for a given signal bandwidth. Therefore, the metric corresponding to the non-real paths are close to zero. However, once the target distance is increased, the standard deviation corresponding to the target's underlying normal distribution increases. Since the estimation accuracy is reduced, the constructed vectors corresponding to the real-paths start to be misaligned with the signal subspace.

Therefore, based on our analysis, three conclusions can be made regarding the proposed metric: i) the threshold based decision making gets more reliable as SNR increases; ii) for low but realistic values of M_f , the proposed metric is reliable as long as the target distance is below 25 m; iii) the gap between the metrics of real and non-real paths widens as SNR increases, which provides a sufficient margin for decision errors and shows the statistical consistency of the proposed metric.

The accuracy of the proposed algorithm is compared with the DFT-based processing, and the results are shown in Figure 7. In this scenario, three targets are simulated: two of them 5 m away with 1 and 2 m/s speeds, the third target 6 m away with 2 m/s speed. For the DFT-based radar processing, 2000 OFDM symbols are used. Hence, the integration time is roughly 28 ms, corresponding to 35 Hz Doppler frequency resolution, or equivalently, 1 m/s speed resolution. Moreover, Blackman window is applied on Doppler profiles in order to suppress the sidelobes, while also increasing the width of the mainlobe. First of all, the accuracy of the DFT-based processing is poor since range and speed accuracies are limited to 1.5 m and 1 m/s, respectively. Moreover, since the Blackman window is applied, the increased mainlobe width of the targets

TABLE 3 Parameters used in the numerical analysis

	Parameter	Value	Unit
Carrier frequency	f_c	5.6	GHz
Bandwidth	B	80	MHz
Number of subcarriers	Q	1024	
CP Length	Q_{cp}	64	
OFDM symbol duration	T	13.6	μ s
Number of OFDM symbols in one frame	M_s	40	
Number of frames	M_f	16	
Transmit power	P_{tx}	20	dBm
Target RCS	σ	-4	dBsm
Covariance matrix size	K	320	
Over-estimated number of unique ranges	P_o	32	
Over-estimated number of speeds	L_o	8	

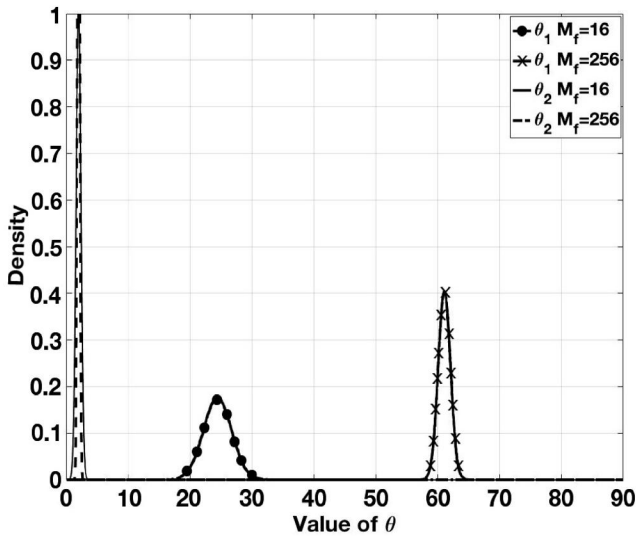


FIGURE 5 The PDFs of θ_1 and θ_2 are shown when $M_f = 16$ and $M_f = 256$ values, even though the latter one is an unrealistic value when Wi-Fi standard is considered. The target distance is 15 m

cause the detection algorithm to detect additional targets. On the other hand, the proposed algorithm shows small errors on the range and speed estimations that is, sub-metres and sub-m/s accuracies are reached.

The scenarios illustrated in Figure 8 and summarised in Table 4 are considered. In Figure 8, ΔR and ΔV represent the relative range and speed difference between the targets, respectively. The PR/AP pair is located at the origin, hence a quasi-monostatic radar geometry is considered. Moreover, the reference range of the first target is denoted by R_0 . The first two scenarios aim to identify the minimum required ΔR and ΔV between two targets such that they can be resolved at the fixed reference range R_0 . The third and fourth scenarios aim to show the impact of SNR on the estimation accuracy by only varying the reference range R_0 . The accuracy of the algorithm

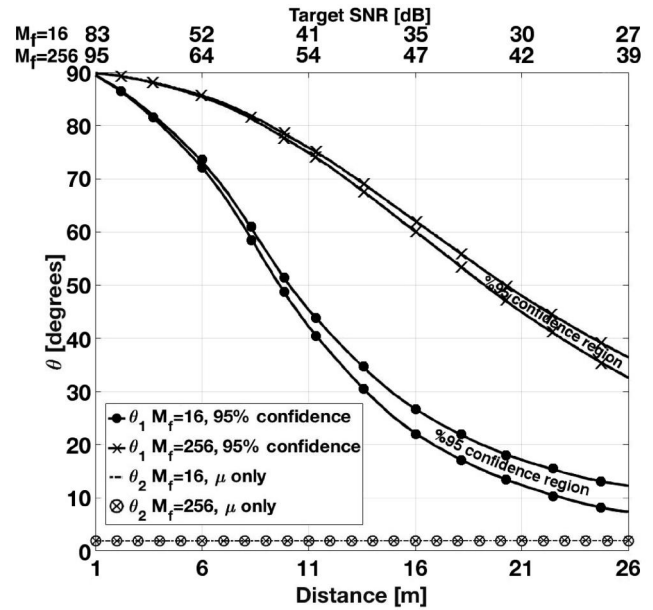


FIGURE 6 The horizontal axis above the graph corresponds to the theoretical SNR for the given distance and M_f values

is measured by computing the Root Mean Square Error (RMSE) of the range and speed estimations. The estimation errors are averaged over 1000 realisations of the parameters. The simulation parameters are selected to comply with the 802.11ax standard, and they are summarised in Table 3. The radar cross section (RCS) of the targets is set to -4 dBsm to emulate the reflectivity of humans. Finally, notation A1(1,10), mainly used in the captions, indicates that the target A1 is moving at 1 m/s speed and located at 10 m range.

In Figure 9, the proposed metric is compared with the AIC [15] and eigenvalue thresholding [16] as a function of the number of targets. The distance between any two targets is 2 m, roughly corresponding to the DFT-based range resolution. First of all, as new targets are included in the simulation, the overall estimation accuracy is reduced since the targets at further distances have relatively lower SNR. In other words, if the SNR of a target echo is weak, its propagation delay will be estimated with considerably higher errors and the corresponding vector in (14) will be poorly constructed. Since our target recovery algorithm relies on the angle between the constructed vector and its orthogonal projection on the noise subspace, a poorly constructed vector yields a miss detection that is the corresponding target will be considered as noise. However, on average 9 out of 12 targets are still detected in the worst case. On the other hand, the eigenvalues of the covariance matrix R_K , given in (11), do not hold the required information about the signal and noise subspaces explaining why the eigen-thresholding method shows such poor statistical performance.

In Figure 10, the performance of the three MOS methods are illustrated as a function of ΔR and ΔV . When the first scenario is considered, the AIC and the proposed metric follow the same trend. The targets cannot be resolved due to the poor

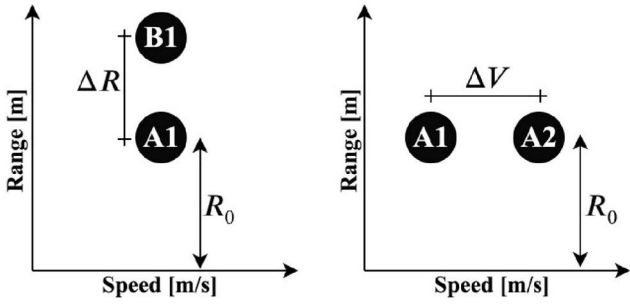


FIGURE 7 The target detection output of the DFT-based processing is compared to the proposed algorithm. Blackman window is applied on Doppler dimension to suppress target leakage

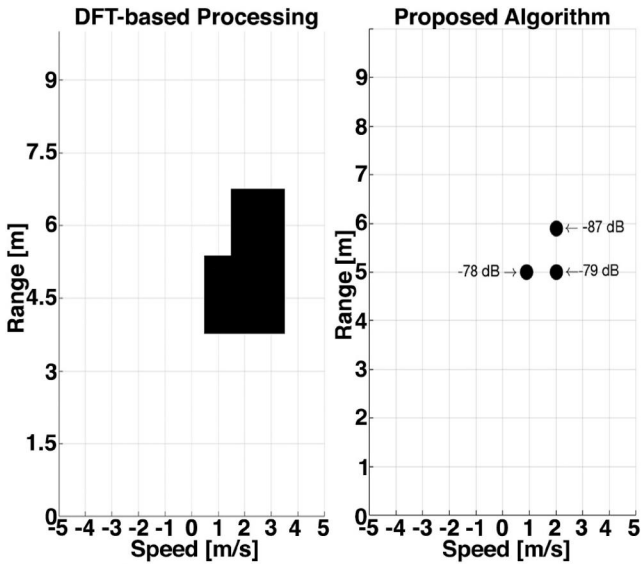


FIGURE 8 Pseudo range-speed maps for the simulation scenarios

TABLE 4 Summary of the scenarios

Scenario	Objects	R_0	ΔR	ΔV
1	A1 & B1	10 m	Varies	0
2	A1 & A2	10 m	0	Varies
3	A1 & B1	Varies	Fixed	0
4	A1 & A2	Varies	0	Fixed

SNR at the reference distance when $\Delta R < 0.5$ m. Once $\Delta R \geq 0.5$ m, both the AIC and the proposed MOS method correctly estimate the number of real-paths. The eigen-thresholding performs very poorly because the eigen-structure of the covariance matrix R_K is altered and the estimated noise threshold is too large. In the second scenario, the eigen-thresholding requires $\Delta V \geq 1.2$ m/s to find the correct model order due to the same reasoning as before. However, both AIC and the proposed metric can find the correct model order as long as $\Delta V \geq 0.5$ m/s at the given reference distance. We remind that AIC estimates the quality of each candidate model, relative to the other candidate models to select the best fit. To

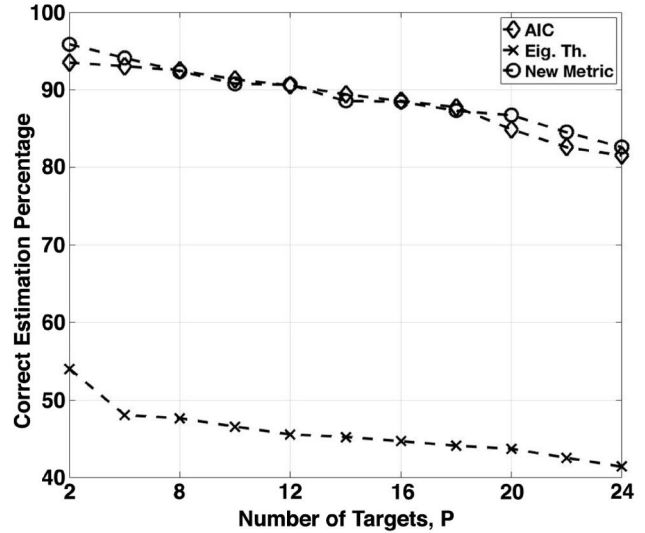


FIGURE 9 Each target is separated by 2 m distance, and all targets are stationary

do so, RSS is computed and an AIC value is obtained for each candidate model. Once all AIC values are obtained, the one with the smallest value is selected as the best fit. On the other hand, the proposed MOS method evaluates each possible path independent from other paths by executing the proposed algorithm only once. Finally, if the proposed MOS method fails to detect the true number of targets during the range stage, this error can have two outcomes. If it is a false detection, the MOS during the speed stage can still remove that particular target. Else, it is a miss detection, and the target will not be detected by the algorithm. However, if the proposed MOS method fails during the speed stage, there is no chance for recovering/removing the targets.

The range and speed estimation accuracy of the algorithm are provided for the first scenario in Figure 11. Whether the MOS stage is used or not, once $\Delta R > 60$ cm, the two targets are resolvable with high accuracy. If the MOS stage is not used, the number of targets to be estimated is directly provided to the estimator that is, P_o is set to two in (13). Therefore, when $\Delta R < 60$ cm, the algorithm estimates two different ranges regardless of the resolvability of the targets. However, there is only one range that can be reliably estimated due to the presence of the noise: the combined range of the two targets. Even though the estimation of the combined range is accurate, the second estimated range is incorrect due to the SNR conditions, which increases the overall estimation error. On the other hand, the proposed MOS method identifies only one resolvable target, and only its range is estimated. Moreover, since the Doppler estimation is performed for the targets identified during the range estimation, the quality of the Doppler estimation directly depends on the quality of the range estimation. Therefore the Doppler estimation RMSE follows the same trends as the range estimation RMSE. Furthermore, the range estimation RMSE can be compared to the Cramer Rao Lower Bound (CRLB). It should be reminded that the CRLB shows the accuracy for separately estimating the

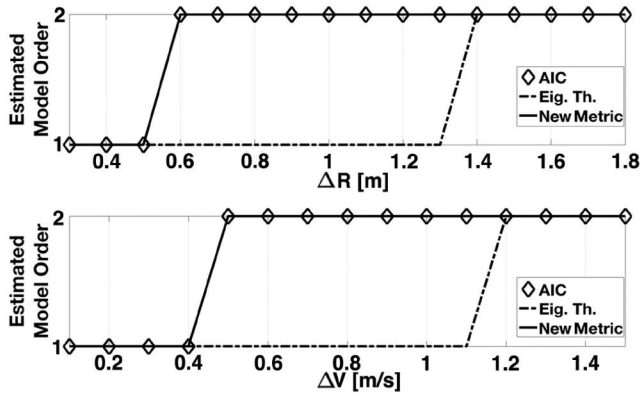


FIGURE 10 Upper graph, scenario 1: $A1(1,10)$ and $B1(1,10+\Delta R)$. Lower graph, scenario 2: $A1(0.5, 10)$ and $A2(0.5+\Delta V, 10)$

two ranges, while the proposed algorithm estimates only one range when $\Delta R < 60$ cm.

In Figure 12, the numerical results are shown for the second scenario. The range error is constant since there is only one fixed range, R_0 , to be estimated. However, even though the range estimation is highly accurate, the speed estimation can still perform poorly due to the small number of available measurements M_f . Furthermore, once $\Delta V > 0.5$ m/s the algorithm correctly estimates the speed of the two targets, independent from the MOS stage for speed. When $\Delta V \leq 0.5$ m/s and if the MOS stage is not used, the speed estimation RMSE is high since only the combined speed of the two targets can be reliably estimated at low ΔV . On the other hand, when the estimated speeds are refined, the RMSE shows similar performance to scenario 1. In other words, when the two targets are moving at similar speeds at the same range, the MOS stage identifies only one target speed which can be reliably estimated.

In Figure 13, the numerical results for the scenario three are plotted. The range estimation errors are very small due to the reduced path loss and increased SNR at closer distances. However, beyond 24 m, range and speed estimations are poor due to the decreased SNR as well as the reduced reliability of the metric. In Figure 14, the numerical results for the fourth scenario are plotted. The range estimation does not fail for the simulated object distances since there is only one range to estimate, namely R_0 . However, due to the reduced SNR and low number of available measurements, the speed estimation accuracy is poor beyond 23-24 m.

7 | EXPERIMENTAL SETUP AND RESULTS

The Software Defined Radios (SDRs) are widely used in the academy and industry to validate signal processing algorithms without any dedicated hardware. Essentially, an SDR hands over the received baseband samples to a host machine for offline processing. One of the most widely used SDR is the Universal Software Radio Peripheral (USRP) [24]. In order to reach high sampling rates, such as 80-160 Msps, overflows at

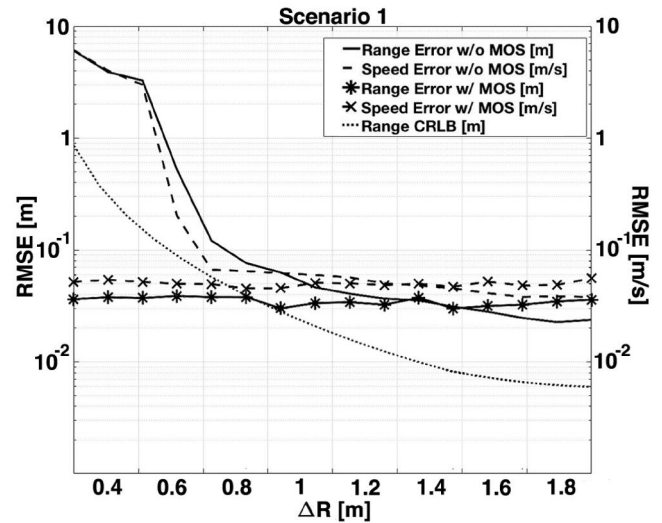


FIGURE 11 $R_0 = 10m$, $A1(1,R_0)$ and $B1(1,R_0+\Delta R)$

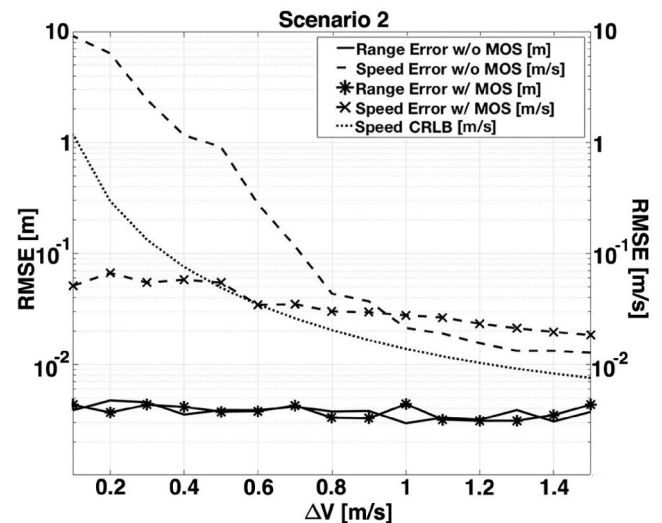


FIGURE 12 $R_0 = 10m$, $A1(0.5, R_0)$ and $A2(0.5+\Delta V, R_0)$

the receiver and underflows at the transmitter are the main bottlenecks that brings latency. The main cause of these issues is the so-called context switching. It allows a single Central Processing Unit (CPU), possibly with multiple cores, to be shared by multiple processes which brings additional latencies. The Data Plane Development Kit (DPDK) [25] is specifically designed to accelerate data processing on the CPUs. In short, DPDK allows us to dedicate multiple cores to specific processes which enables us to stream at higher sampling rates. In our recent work [17,26], we have shown that up-to six USRPs can be connected to one host machine equipped with a 12-core CPU, 64 GB RAM, and 6×10 Gbe Ethernet connections, and the experiments can be done without any interruptions at 100 Msps sampling rate.

In Figure 15, the experimental setup is shown, which is composed of two USRPs: one for the transmitter (TX) and

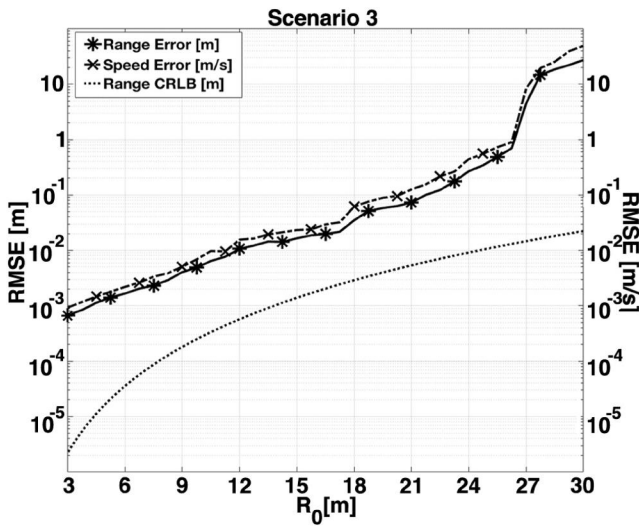


FIGURE 13 $\Delta R = 1\text{ m}$, R_0 varies, $A1(1, R_0)$ and $B1(1, R_0 + \Delta R)$

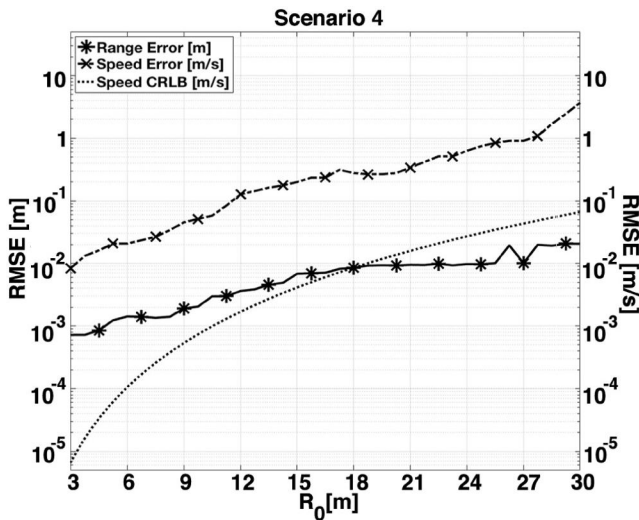


FIGURE 14 $\Delta V = 0.6\text{ m/s}$, R_0 varies, $A1(1, R_0)$ and $A2(1 + \Delta V, R_0)$

the other for the receiver (RX). The same parameters given in Table 3 are used in the experiments as well. The measurements are obtained in a $4 \times 8\text{ m}$ rectangular room when two individuals are walking, emulating a typical household scenario where SNR conditions are satisfied since the TX/RX are inside the same room, and thanks to the short ranges. In order to identify the number of targets with the proposed MOS method, the thresholds given in Figure 5 are used with respect to the estimated target ranges. Finally, a target on the RDM is denoted by its speed/range vector: (speed [m/s], range[m]).

The goal of the first experiment is to illustrate the detrimental effect caused by the static clutter when two individuals are walking on a straight line in opposite directions. Figure 16 illustrates the range-speed map obtained at the output of the radar processing. The estimated amplitudes in (28),

corresponding to the detected targets are also shown. The ground truth of the two targets are plotted with plus signs while the dots correspond to the detected targets. The static clutter is visible on the vertical axis around 0 m/s . Due to the estimation errors and small vibrations, some of the clutter components appear with non-zero velocities. The person who is moving away that is the plus sign at $(-0.75, 3.15)$, is detected with speed and range errors of 0.1 m/s and 0.8 m , respectively. The target detected at $(-0.9, 5.4)$ corresponds to a multipath reflection of the target at $(-0.75, 3.15)$. However, the person who is approaching towards the radar setup that is, the plus sign at $(1.1, 2)$, is not detected at all, while its multipath component is detected at $(1.2, 5.95)$. Since the signals reflected from the static clutter have significantly higher magnitude than the signals reflected from the individuals, some of the targets are missed or removed during the MOS stage. In this particular case, the target at $(1.1, 2)$ is closer to the radar system. However, the static clutter around the same range of this target that is, the target at $(0.2, 2.1)$, has a stronger magnitude as well which yields a miss detection. Therefore, the estimation accuracy is significantly reduced due to the presence of the strong static clutter.

As explained in Section 4, once a target with zero (or close-to-zero) speed is detected during the speed estimation stages, its contribution can be removed from the measured signal, and the estimation stages can be repeated to increase the robustness of the estimations. The second experiment, shown in Figure 17, is identical to the first experiment, except that upper-mentioned clutter removal is used to improve the estimation accuracy. After the disturbance caused by the static clutter is removed, the range and speed errors are now below 0.5 m and 0.5 m/s for both targets, respectively. Still the experimental results do not achieve the performance expected from the numerical results because i) the numerical analysis are provided after multiple realisations; ii) when the numerical analysis are performed, the non-idealities caused by the hardware are not considered. However, it is well-known that non-idealities such as phase noise, have significant impact on OFDM modulated signals and subspace-based estimation methods [27]. Moreover, since the static clutter is removed, additional multipath components appear along the range profile of the two targets, causing the so-called ghost targets. We remind that the ghost targets are high-order reflections caused by the real targets in the environment, which is common when indoor scenarios are considered. However, the ghost targets can be eliminated by using a tracking algorithm [28] or a multi-static radar configuration [29].

8 | CONCLUSION

This work aims at designing a new Passive Radar based on 802.11ax compliant preambles for indoor human movement detection. A new method is proposed to determine the number of resolvable targets based on which the robust range and speed estimations can be performed with sub-metre and sub-m/s accuracy, respectively.

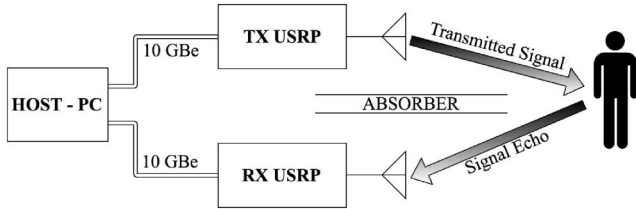


FIGURE 15 The experimental setup from top-view. The distance between the transmitter and receiver is 1.5 m, and the absorber is used to remove the direct signal between antennas

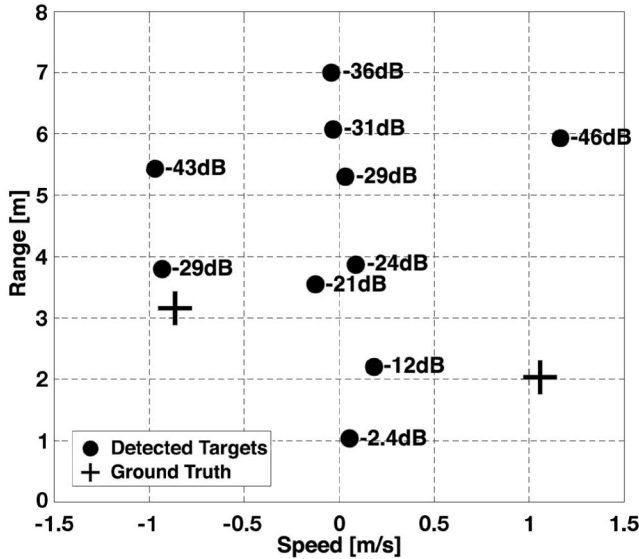


FIGURE 16 Two targets walking towards opposite directions

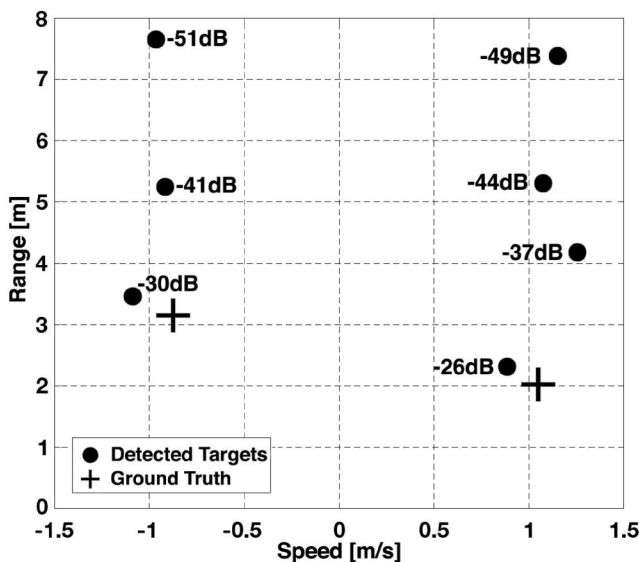


FIGURE 17 Two targets walking towards opposite directions. The range and speed of the targets are estimated after the clutter removal

The statistical consistency of the proposed MOS method is shown to be reliable for correctly estimating the number of targets. Moreover, when closely spaced targets are simulated in typical indoor SNR conditions, the proposed algorithm is shown to be accurate enough to detect human movement.

Highest computational resources are required by the range-group due to large matrices involved in the processing. However, parallel computation can reduce the consumption time of the related computations.

Finally, the algorithm is validated with a state-of-the-art SDR test-bench. There are still open challenges that require future work: the hardware non-idealities that significantly reduce the estimation accuracy and the multipath components which cause ghost targets on a range-speed map.

ORCID

Hasan Can Yildirim  <https://orcid.org/0000-0002-9905-6615>

REFERENCES

- Richards, M.A., Scheer, J., Holm, W.A.: Principles of modern radar: Basic principles. Institution of Engineering and Technology, Chennai (2010)
- Falcone, P., Colone, F., Lombardo, P.: Potentialities and challenges of WiFi-based passive radar. *IEEE Aero. Electron. Syst. Mag.* 27(11), 15–26 (2012). <https://doi.org/10.1109/MAES.2012.6380822>
- IEEE Standards Association.: P802.11ax: “IEEE Draft standard for information Technology – Telecommunications and information Exchange between systems local and Metropolitan Area Networks – specific requirements Part 11: Wireless LAN Medium access Control (MAC) and Physical layer (PHY) Specifications amendment Enhancements for high Efficiency WLAN, pp. 1–746. IEEE, New York City (2019). <https://ieeexplore.ieee.org/document/867264>
- Colone Fabiola, Woodbridge Karl, Guo Hui, Mason David, Baker Chris J. Ambiguity Function Analysis of Wireless LAN Transmissions for Passive Radar. *IEEE Transactions on Aerospace and Electronic Systems* 47, (1), 240–264 (2011). <http://dx.doi.org/10.1109/taes.2011.5705673>
- Aruba Networks 802.11ax white paper, Aruba Networks, Santa Clara, CA (2019). https://www.arubanetworks.com/assets/wp/WP_802.11AX.pdf. Accessed 19 Jun. 2020
- Palmer James E., Harms H. Andrew, Searle Stephen J., Davis LindaM. DVB-T Passive Radar Signal Processing. *IEEE Transactions on Signal Processing* 61, (8), 2116–2126 (2013). <http://dx.doi.org/10.1109/tsp.2012.2236324>
- IEEE, IEEE 802.11, “Wi-Fi sensing”, IEEE, New York City (2019). https://mentor.ieee.org/802.11/documents?is_dcn=DCN%2C%20Title%2C%20Author%20or%20Affiliation&is_group=SENS. Accessed 27 Dec 2019
- Li Wenda, Piechocki Robert J., Woodbridge Karl, Tang Chong, Chetty Kevin Passive WiFi Radar for Human Sensing Using a Stand-Alone Access Point. *IEEE Transactions on Geoscience and Remote Sensing* 1–13 (2020). <http://dx.doi.org/10.1109/tgrs.2020.3006387>
- Gu, J.F., Moghaddasi, J., Wu, K., Delay and Doppler shift estimation for OFDM-based radar-radio (RadCom) system., pp. 1–4. *IWS 2015, Shenzhen* (2015). <https://doi.org/10.1109/IEEE-IWS.2015.7164629>
- Nguyen, D.H.N., Heath, R.W.: Delay and Doppler processing for multi-target detection with IEEE 802.11 OFDM signaling. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3414–3418. New Orleans (2017). <https://doi.org/10.1109/ICASSP.2017.7952790>
- Stoica P., Selen Y. Model-order selection. *IEEE Signal Processing Magazine* 21, (4), 36–47 (2004). <http://dx.doi.org/10.1109/msp.2004.1311138>

12. Carotenuto Vincenzo, De Maio Antonio, Orlando Danilo, Stoica Petre Model Order Selection Rules for Covariance Structure Classification in Radar. *IEEE Transactions on Signal Processing* 65, (20), 5305–5317 (2017). <http://dx.doi.org/10.1109/tsp.2017.2728523>
13. Louis, L., Scharf : Statistical signal processing. Pearson, New York City (/1997 1991/1997)
14. Kullback S., Leibler R.A.: On information and sufficiency, *Ann. Math. Stat.* 22(1), 79–86 (1951)
15. Akaike, H.: Selected Papers of Hirotugu Akaike, Springer Series in Statistics (Perspectives in Statistics). Selected Papers of Hirotugu Akaike, Springer, New York. https://doi.org/10.1007/978-1-4612-1694-0_29
16. Chen W, Wong K.M., Reilly J.P. Detection of the number of signals: a predicted eigen-threshold approach. *IEEE Transactions on Signal Processing* 39(5), 1088–1098 (1991). <http://dx.doi.org/10.1109/78.80959>
17. Yildirim, H.C., et al.: Passive Radar based on 802.11ac Signals for Indoor Object Detection. 2019 16th European Radar Conference (EuRAD), pp. 153–156, Paris (2019)
18. Roy R., Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37(7), 984–995 (1989). <http://dx.doi.org/10.1109/29.32276>
19. Stoica, P., Moses, R.L.: Spectral analysis of signals. Pearson Prentice Hall, New Jersey, NJ (2005)
20. Storrer, L., et al.: Clutter removal for Wi-Fi-based passive radar. In: Proceedings of the IEEE 91st Vehicular Technology Conference, vol. 25-28. VTC-Spring, Antwerp (2020)
21. Chen, V. C., et al.: Micro-Doppler effect in radar: phenomenon, model, and simulation study. *IEEE Trans. Aerosp. Electron. Syst.* 42(1), 2–21 (2006). <https://doi.org/10.1109/TAES.2006.1603402>
22. Big-O notation Tutorial, MIT, Massachusetts (2017). http://web.mit.edu/16.070/www/lecture/big_o.pdf. Accessed 3 July 2020
23. Bürgisser, P., Clausen, M., Amin, S.: Algebraic complexity theory. Springer, New York City (1997)
24. Universal software radio peripheral, Ettus, Austin, TX (2017). https://files.ettus.com/manual/page_usrp_x3x0.html. Accessed 3 July 2020
25. Data plane development kit. <https://www.dpdk.org/>. Accessed 3 July 2020
26. Storrer, L., et al., Experimental implementation of a multi-antenna 802.11 ax-based passive bistatic radar for human target detection, In: Proceedings of the IEEE Radar Conference Radar Conference, 21–25, Florence (2020)
27. Horlin, F., Bourdoux, A.: Digital compensation for analog front-ends: A new approach to wireless transceiver design. John Wiley and Sons Ltd., New Jersey (2008)
28. Ryu, I.-H., Won, I., Kwon, J.: Detecting host targets using multilayer perceptron in multiple-target tracking. *Symmetry*. 10, 16 (2018)
29. Zhang, Y.D., Himed, B.: Moving target parameter estimation and SFN ghost rejection in multistatic passive radar. In: Proceedings of 2013 IEEE Radar Conference (RadarCon13), pp. 1–5, Ottawa (2013)

How to cite this article: Yildirim HC, Determe JF, Storrer L, et al. Super resolution passive radars based on 802.11ax Wi-Fi signals for human movement detection. *IET Radar Sonar Navig.* 2021;15:323–339. <https://doi.org/10.1049/rsn2.12038>