

## RESEARCH

## Open Access



# Joint channel and phase noise estimation for mmWave full-duplex communication systems

Abbas Koohian<sup>1\*</sup> , Hani Mehrpouyan<sup>2</sup>, Ali A. Nasir<sup>3</sup> and Salman Durrani<sup>1</sup>

## Abstract

Full-duplex (FD) communication at millimeter-wave (mmWave) frequencies suffers from a strong self-interference (SI) signal, which can only be partially canceled using conventional RF cancellation techniques. This is because current digital SI cancellation techniques, designed for microwave frequencies, ignore the rapid phase noise (PN) variation at mmWave frequencies, which can lead to large estimation errors. In this work, we consider a multiple-input multiple-output mmWave FD communication system. We propose an extended Kalman filter-based estimation algorithm to track the rapid variation of PN at mmWave frequencies. We derive a lower bound for the estimation error of PN at mmWave and numerically show that the mean square error performance of the proposed estimator approaches the lower bound. We also simulate the bit error rate performance of the proposed system and show the effectiveness of a digital canceler, which uses the proposed estimator to estimate the SI channel. The results show that for a  $2 \times 2$  FD system with 64-QAM modulation and PN variance of  $10^{-4}$ , the residual SI power can be reduced to  $-25$  dB and  $-40$  dB, respectively, for signal-to-interference ratio of 0 and 15 dB.

**Keywords:** Full-duplex, Millimeter-wave, Joint channel and PN estimation, Residual self-interference power, Synchronization

## 1 Introduction

The next generation of wireless communication technologies, known as 5G, are expected to offer multi-gigabit data rates to mobile users [1, 2]. This has prompted wireless service providers to seek higher bandwidth at less crowded millimeter-wave (mmWave) frequencies. The short wave lengths of mmWave frequencies also allow for practical implementations of base stations with large number of antennas known as massive multiple-input multiple-output (MIMO) system, which is another promising technology for 5G networks [3]. Given these capabilities offered by mmWave communication, these systems have become increasingly popular in academia and industry. While MIMO systems can fully benefit from the capabilities offered by communication at mmWave frequencies [3], due to large peak-to-average power ratio,

orthogonal frequency division modulation (OFDM) is not popular for mmWave communication [4]. Since there is still an open debate about the modulation type at mmWave frequencies [5], we do not consider OFDM in this work.

Full-duplex (FD) communication has also emerged recently as a promising wireless technology, which allows for efficient use of bandwidth by enabling in-band transmission and reception [6–9]. The major obstacle in exploiting the full potential of FD communication is the self-interference (SI) signal, which is significantly stronger than the desired communication signal [10, 11]. The power of the SI signal can be reduced via two different suppression techniques: (i) *passive suppression*, where transmit and receive antennas are physically isolated to reduce the leakage of the transmit signal into the RF front end of the receiver chain, and (ii) *active suppression*, where the SI signal is suppressed via subtracting the analog replica of SI signal from the received signal [6]. The experimental results at microwave frequencies show that the successive combination of passive and active

\*Correspondence: [abbas.koohian@anu.edu.au](mailto:abbas.koohian@anu.edu.au)

<sup>1</sup>Research School of Electrical, Energy, and Materials Engineering, Australian National University, Canberra, Australia

Full list of author information is available at the end of the article

suppression can reduce the SI signal power to the receiver noise floor [12]. For this reason, in the majority of the radio architectures proposed for FD communication at microwave frequencies, the residual SI signal at baseband is treated as noise [6, 13–15]. The cancelation techniques that treat the SI signal as noise suffer from two fundamental problems: (i) they assume Gaussian distribution of the SI signal. However, as explained in [16, 17], the SI signal has a strong line of sight (LoS) component, and hence, it is not Gaussian, and (ii) treating SI as noise requires statistical knowledge of the SI channel, which might not be available.

Recently, SI channel measurements have been carried out for FD communication at mmWave frequencies [17, 18]. The measurements indicate that, as opposed to the microwave frequency band, the SI channel at mmWave has a non-line-of-sight (NLoS) component, which cannot be canceled using passive and active suppression techniques. This partial suppression of the SI signal results in a large residual SI signal at baseband, which is still significantly higher than the receiver noise floor [17]. Another challenge for mmWave FD communication systems is that the oscillator phase noise (PN) is large and rapidly changing [19]. Thus, the majority of the existing techniques for residual SI signal cancelation at baseband, which assume a very steady oscillator PN [20–23] cannot be used for FD mmWave communication. We note that the important aspect of mmWave communication considered in this paper is the estimation of fast varying PN. This fast variation of PN is in the order of symbol time at mmWave [24].

In this work, we consider the problem of joint channel and PN estimation for a mmWave FD MIMO system communication system. The main contributions of this work are as follows:

1. We construct a state vector for the joint estimation of the channel and PN and propose an algorithm based on extended Kalman filtering technique to track the fast PN variation at mmWave band.
2. We derive the lower bound on the estimation error of the proposed estimator and numerically show that the proposed estimator reaches the performance of the lower bound. We also show the effectiveness of a digital SI cancelation, which uses the proposed estimation technique to estimate the SI channel.
3. We present simulation results to show the mean square error (MSE) and bit error rate (BER) performance of a mmWave FD MIMO system with different PN variances and signal-to-interference ratios (SIR). The results show that for a  $2 \times 2$  FD system with 64-QAM modulation and PN variance of  $10^{-4}$ , the residual SI power can be reduced to  $-25$  dB and  $-40$  dB, respectively, for signal-to-interference ratio of 0 and 15 dB.

*Notation:* The following notation is used in this paper. Superscripts  $(\cdot)^\dagger$  and  $(\cdot)^T$  are the conjugate and the transpose operators, respectively. Bold face small letters, e.g.,  $\mathbf{x}$  are used for vectors, bold face capital letters, e.g.,  $\mathbf{X}$  are used for matrices.  $e^{j\theta}$  is the multivariate complex exponential function.  $|\cdot|$  is the absolute value operator, and  $\angle x$  is the phase of the complex variable  $x$ .  $\odot$  is the Hadamard (element-wise) product.  $\text{diag}(\mathbf{x})$  creates a matrix with elements of vector  $\mathbf{x}$  on the main diagonal.  $\text{trace}(\cdot)$  is the trace of a matrix, which sums up all the diagonal elements of a given matrix.  $x \in \mathcal{A}$ , means  $x$  is an element of set  $\mathcal{A}$ .  $\mathbf{0}_N$  is  $N$  by 1 vector of all zeros, and  $\mathbf{I}_N$  is  $N$  by  $N$  identity matrix.  $\mathbf{1}_{N \times N}$  is  $N \times N$  matrix of all 1.  $\mathbb{E}[\cdot]$  is the expectation operator.  $\Re\{\cdot\}$  returns the real part of a complex quantity. Finally, in Table 1, we present the important

**Table 1** Important symbols used in this paper

Symbol	Description
$\mathbf{y}(n)$	The $(N_r \times 1)$ vector of received symbols.
$\mathbf{x}(n)$	The $(N_t \times 1)$ vector of transmitted symbols.
$\mathbf{x}^{\text{SI}}(n)$	The $(N_t \times 1)$ vector of self-interfering (SI) symbols.
$\mathbf{w}(n)$	The $(N_r \times 1)$ Gaussian noise vector.
$\theta_t^{[m]}(n)$	The time-varying phase noise of $i$ th oscillator and $m \in \{t = \text{transmit}, r = \text{receive}, \text{SI}\}$ .
$\mathbf{H}(n)$	The $(N_r \times N_t)$ communication channel.
$\mathbf{H}^{\text{SI}}(n)$	The $(N_r \times N_t)$ self-interference (SI) channel.
$\bar{\mathbf{H}}(n)$	The $(N_r \times 2N_t)$ state transition matrix for joint PN and channel estimation.
$\boldsymbol{\beta}(n)$	The $(2N_t \times 1)$ state vector for joint channel and PN estimation.

symbols used in the mathematical representation of the system model. In general, if  $x$  is used in the mathematical representation of the system model, then  $\tilde{x}$  is used for the mathematical representation of the system model needed for joint channel and PN estimation, and  $\hat{x}$  is an estimate of  $x$ .

## 2 System model

We consider the MIMO communication system between two mmWave FD nodes  $a$  and  $b$ , each with  $N_t$  transmit and  $N_r$  receive antennas as illustrated in Fig. 1. The considered communication system can be a model for backhaul communication for cellular systems [24]. In this work, we make the following assumptions:

1. *The same number of transmit and receive antennas for both nodes:* We assume both nodes in the considered FD communication system have the same number of transmit and receive antennas.
2. *Modeling of RF impairments:* RF impairments due to imperfect transmitter and receiver chain electronics have been shown to significantly degrade the performance of the analog cancellation techniques [25, 26]. Since the focus of this work is residual SI cancellation, we only include PN in our model and assume that the other hardware impairments are dealt by a RF canceler. Such an assumption is also made in [20, 21, 27, 28].
3. *Assumptions on oscillators:* We make two assumptions about the oscillators. First, we assume that free-running oscillators are used. The assumption of using free-running oscillators for mmWave communications has also been made in

[24, 29]. Second, we assume each transmit and receive antenna is equipped with an independent oscillator.

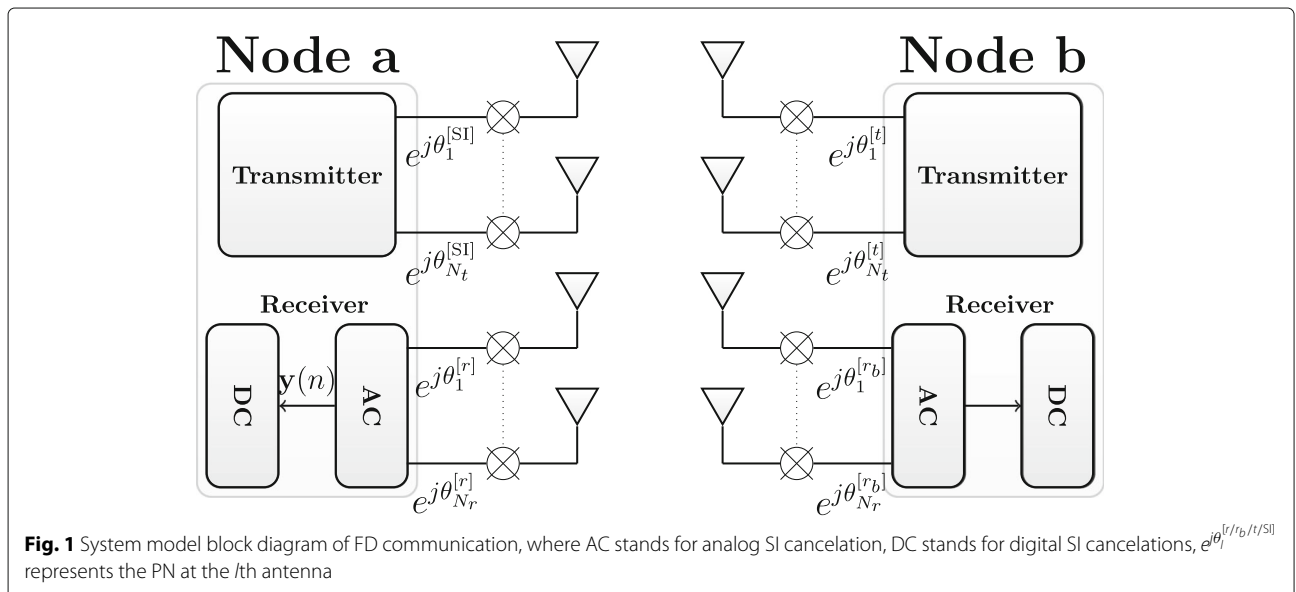
4. *Quasi-static flat-fading channel assumption:* The SI measurement results of [17] show that even with omnidirectional dipole antennas, the delay spread of the channel does not exceed 800 ns. This delay is significantly smaller than the proposed symbol durations for 5G communication [30, 31], which are in order of  $\mu s$ . Hence, not only can the channel be assumed flat but it can also be assumed to remain constant over transmission of one block of data (quasi-static). Similarly, measurement results of the desired communication channel show that the channel delays are relatively small compared to the symbol durations [32].
5. *Synchronized transmission and reception:* Although synchronizing transmission and reception of analog desired communication signal with the reception of analog SI signal is an important practical problem and requires its own detailed investigation, the synchronized FD communication assumption is widely used in the literature of channel and PN estimation for digital SI cancellation (DC) [20, 21, 33].

### 2.1 Mathematical representation of received vector

In this subsection, we present a mathematical model for the received vector of a FD MIMO communication system at mmWave frequencies. The received vector at node  $a$  and at time instant  $n$  is  $\mathbf{y}(n)$  and is given by

$$\mathbf{y}(n) = \mathbf{H}(n)\mathbf{x}(n) + \mathbf{H}^{SI}(n)\mathbf{x}^{SI}(n) + \mathbf{w}(n), \quad (1)$$

where  $\mathbf{y}(n) \triangleq [y_1(n), \dots, y_{N_r}(n)]^T$ , and  $y_i(n)$  is the received symbol at the  $i$ th antenna. For  $i \in \{1, \dots, N_r\}$



and  $k \in \{1, \dots, N_t\}$ , the element in the  $i$ th row and  $k$ th column of  $N_r \times N_t$  channel matrix  $\mathbf{H}(n)$  is given by  $h_{i,k} e^{j(\theta_i^{[r]}(n) + \theta_k^{[t]}(n))}$ , where  $h_{i,k}$  is the communication channel between the  $k$ th transmit antenna of node  $b$  to the  $i$ th receive antenna of node  $a$ , for  $m \in \{r, t, \text{SI}\}$ ,  $\theta_j^{[m]}(n)$  is the oscillator PN at the  $j$ th antenna and  $m$  determines the type of antenna such that  $m = r$  indicates a receive antenna,  $m = t$  means a transmit antenna, and  $m = \text{SI}$  indicates an interfering antenna. Furthermore, PN variation of a free-running oscillator follows a Wiener process [34], i.e.,

$$\theta_j^{[m]}(n) = \theta_j^{[m]}(n-1) + \delta(n), \tag{2}$$

where  $\delta(n)$  is Gaussian noise with mean 0 and variance  $\sigma_{[m]}^2$ , i.e.,  $\delta(n) \sim \mathcal{N}(0, \sigma_{[m]}^2)$ .

Similarly, the element in the  $i$ th row and  $k$ th column of  $N_r \times N_t$  SI channel matrix  $\mathbf{H}^{\text{SI}}(n)$  is given by  $h_{i,k}^{\text{SI}} e^{j(\theta_i^{[r]}(n) + \theta_k^{\text{SI}}(n))}$ , where  $h_{i,k}^{\text{SI}}$  is the interference channel between the  $k$ th transmit antenna and the  $i$ th receive antenna of node  $a$ .

In addition, the  $k$ th elements of  $N_t \times 1$  vectors  $\mathbf{x}(n)$  and  $\mathbf{x}^{\text{SI}}(n)$  are given by  $x_k(n)$  and  $x_k^{\text{SI}}(n)$ , respectively, which are the transmitted symbols from the  $k$ th transmit antenna of nodes  $b$  and  $a$ , respectively.

Finally,  $\mathbf{w}(n) \triangleq [w_1(n), \dots, w_{N_r}(n)]^T$ , where  $w_i(n)$  is the complex Gaussian noise, i.e.,  $w_i(n) \sim \mathcal{CN}(0, \sigma^2)$ .

### 2.2 Mathematical representation for joint channel and PN estimation

For received vector  $\mathbf{y}(n)$  and noise vector  $\mathbf{w}(n)$  in (1), a useful mathematical model for joint channel and PN estimation is of the form ([35], Ch. 13, pp.450, Eq. 13.66)

$$\mathbf{y}(n) = \bar{\mathbf{H}}(n)f(\boldsymbol{\beta}(n)) + \mathbf{w}(n), \tag{3}$$

where  $\bar{\mathbf{H}}(n)$  is the state transition model matrix,  $f$  is a nonlinear function, and  $\boldsymbol{\beta}(n)$  is the state vector to be estimated.

A fundamental step in the problem of joint channel and PN estimation is the construction of the state vector and the state transition matrix based on the system model given by (1). The state vector and the state transition matrix for the joint PN and channel estimation in the presence of SI signal are given by (4) and (5), respectively.

– The state vector:

$$\boldsymbol{\beta}(n) \triangleq [\bar{\boldsymbol{\beta}}_1(n), \dots, \bar{\boldsymbol{\beta}}_{N_r}(n)]^T \tag{4}$$

– The state transition matrix:

$$\bar{\mathbf{H}}(n) \triangleq \begin{bmatrix} \bar{\mathbf{h}}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \bar{\mathbf{h}}_{N_r} \end{bmatrix} \tag{5}$$

where

$$\bar{\boldsymbol{\beta}}_i(n) \triangleq [\bar{\beta}_{i,1}, \dots, \bar{\beta}_{i,2N_t}], \tag{6a}$$

$$\bar{\beta}_{i,\bar{k}}(n) \triangleq \begin{cases} \theta_i^{[r]}(n) + \theta_{\bar{k}}^{[t]}(n), & \bar{k} \text{ is odd;} \\ \theta_i^{[r]}(n) + \theta_{\bar{k}}^{\text{SI}}(n), & \bar{k} \text{ is even} \end{cases}, \tag{6b}$$

$$\bar{\mathbf{h}}_i \triangleq [\bar{h}_{i,1}, \dots, \bar{h}_{i,2N_t}] \odot [\bar{x}_1(n), \dots, \bar{x}_{2N_t}(n)], \tag{6c}$$

$$\bar{h}_{i,\bar{k}} \triangleq \begin{cases} h_{i,k}, & \bar{k} \text{ is odd;} \\ h_{i,k}^{\text{SI}}, & \bar{k} \text{ is even} \end{cases}, \tag{6d}$$

$$\bar{x}_{\bar{k}}(n) \triangleq \begin{cases} x_k(n), & \bar{k} \text{ is odd;} \\ x_k^{\text{SI}}(n), & \bar{k} \text{ is even} \end{cases}, \tag{6e}$$

$$\bar{k} = \{1, \dots, 2N_t\} \tag{6f}$$

$$k = \begin{cases} \bar{k}, & \bar{k} < N_t; \\ \bar{k} - N_t, & \bar{k} > N_t \end{cases}. \tag{6g}$$

The principle idea behind the design of the state vector  $\boldsymbol{\beta}(n)$  and the state transition matrix  $\bar{\mathbf{H}}(n)$  as given by (4) and (5), respectively, is the fact that the PN noise is the only random variable that varies from one symbol to another and needs to be tracked. On the other hand, because of the quasi-static nature of the communication and SI channels, they remain constant over transmission of a single data packet. Therefore, these channels need to be estimated only once at the beginning of data transmission. This initial channel estimation for the constant channels can be done using pilot transmission.

Furthermore, we note that at each receive antenna, there are  $2N_t$  parameters that need to be estimated,  $N_t$  parameters for the communication channel, and  $N_t$  parameters for the SI channel. This explains the existence of index  $\bar{k} \in \{1, \dots, 2N_t\}$ .

Finally, with the state vector  $\boldsymbol{\beta}(n)$  and the state transition matrix  $\bar{\mathbf{H}}(n)$  given by (4) and (5), the discrete-time received vector at time instant  $n$  and at the baseband of node  $a$  is given by

$$\mathbf{y}(n) = \bar{\mathbf{H}}(n)e^{j\boldsymbol{\beta}(n)} + \mathbf{w}(n). \tag{7}$$

### 3 Joint channel and PN estimation

In this section, we use the state vector (4) and the state transition matrix (5) and present a joint channel and PN estimator based on the concept of extended Kalman filtering (EKF) [35]. The observation vector of EKF is given by  $\mathbf{y}(n)$  in (7), which is a nonlinear function of the states  $\boldsymbol{\beta}(n)$ . The EKF state equation is given by

$$\boldsymbol{\beta}(n) = \boldsymbol{\beta}(n-1) + \mathbf{u}(n), \tag{8}$$

where  $\mathbf{u}(n)$  is Gaussian with mean zero and covariance  $\mathbf{Q} \triangleq \mathbb{E}[\boldsymbol{\beta}(n)\boldsymbol{\beta}^T(n)]$ , i.e.,  $\mathbf{u}(n) \sim \mathcal{N}(\mathbf{0}_{2N_r N_t}, \mathbf{Q})$ . The  $2N_t N_r \times 2N_t N_r$  covariance matrix  $\mathbf{Q}$  is given by

$$\mathbf{Q} \triangleq \mathbb{E} \left[ \boldsymbol{\beta}(n) \boldsymbol{\beta}^T(n) \right] = \begin{bmatrix} \mathbf{R}_{1,1} & \cdots & \bar{\mathbf{R}}_{1,N_r} \\ \vdots & \vdots & \vdots \\ \mathbf{R}_{N_r,1} & \cdots & \bar{\mathbf{R}}_{N_r,N_r} \end{bmatrix}, \quad (9)$$

where, for  $m, n \in \{1, \dots, N_r\}$ ,  $\mathbf{R}_{m,n}$  is  $2N_t N_t$  matrix given by (10), where  $\sigma_r^2$ ,  $\sigma_t^2$ , and  $\sigma_{SI}^2$  are PN variances due to receive, transmit, and SI antennas, respectively.

$$\mathbf{R}_{m,n} = \begin{cases} \sigma_r^2 \mathbf{1}_{2N_t \times 2N_t} + \text{diag} \left( \underbrace{\sigma_t^2, \dots, \sigma_t^2}_{2N_t} \right), & m = n \text{ and is odd} \\ \sigma_r^2 \mathbf{1}_{2N_t \times 2N_t} + \text{diag} \left( \underbrace{\sigma_{SI}^2, \dots, \sigma_{SI}^2}_{2N_t} \right), & m = n \text{ and is even} \\ \text{diag} \left( \underbrace{\sigma_t^2, \dots, \sigma_t^2}_{2N_t} \right), & m \neq n \text{ and is odd} \\ \text{diag} \left( \underbrace{\sigma_{SI}^2, \dots, \sigma_{SI}^2}_{2N_t} \right), & m \neq n \text{ and is even} \end{cases} \quad (10)$$

The EKF state update equations are given by [35]

$$\hat{\boldsymbol{\beta}}(n|n) = \hat{\boldsymbol{\beta}}(n|n-1) + \Re \left\{ \mathbf{K}(n) \left( \mathbf{y}(n) - \bar{\mathbf{H}}(n) e^{j\hat{\boldsymbol{\beta}}(n|n-1)} \right) \right\}, \quad (11)$$

$$\hat{\boldsymbol{\beta}}(n|n-1) = \hat{\boldsymbol{\beta}}(n-1|n-1), \quad (12)$$

$$\mathbf{K}(n) = \mathbf{M}(n|n-1) \mathbf{D}^\dagger(n) \times \left( \sigma^2 \mathbf{I}_{N_r} + \mathbf{D}(n) \mathbf{M}(n|n-1) \mathbf{D}^\dagger(n) \right)^{-1}, \quad (13)$$

$$\mathbf{M}(n|n-1) = \mathbf{M}(n-1|n-1) + \mathbf{Q}, \quad (14)$$

$$\mathbf{M}(n|n) = \Re \left\{ \left( \mathbf{I}_{N_r} - \mathbf{K}(n) \mathbf{D}(n) \right) \mathbf{M}(n|n-1) \right\}, \quad (15)$$

where, for  $k \in \{1, \dots, N_t\}$ ,

$$\mathbf{D}(n) = \frac{\partial \bar{\mathbf{H}}(n) e^{j\boldsymbol{\beta}(n)}}{\partial \boldsymbol{\beta}^T(n)} = \begin{pmatrix} \mathbf{z}_1 & \mathbf{0}_{2N_t}^T & \mathbf{0}_{2N_t}^T \\ \mathbf{0}_{2N_t}^T & \ddots & \vdots \\ \mathbf{0}_{2N_t}^T & \mathbf{0}_{2N_t}^T & \mathbf{z}_{N_r} \end{pmatrix}, \quad (16)$$

$$\mathbf{z}_i = \begin{cases} h_{i,k} x_k(n) e^{j\hat{\beta}_{i,k}(n|n-1)} & k \text{ is even} \\ h_{i,k}^{SI} x_k^{SI}(n) e^{j\hat{\beta}_{i,k}(n|n-1)} & k \text{ is odd} \end{cases}, \quad (17)$$

and  $\hat{\beta}_{i,k}(n|n-1)$  is the  $2(i-1)N_t + k$  element of vector  $\hat{\boldsymbol{\beta}}(n|n-1)$ .

**Remark 1** We note that the state vector, as given by (8), is a real vector. This is because the state vector only contains the phases, which are real numbers. The complex channel coefficients are estimated using this estimated real vector and using the complex exponential function as given by (7). Since the states are all real, when updating the mean of the states in EKF, we can safely discard the imaginary part of the updated mean as in (11).

### 3.1 Symbol detection

The EKF Eq. (17) shows that  $\mathbf{z}_i$  requires the knowledge of the constant channels  $h_{i,k}$ ,  $h_{i,k}^{SI}$  and the transmitted symbols. Note that  $x_k^{SI}$ , the SI symbol is perfectly known at the receiver.

The knowledge of the constant channels can be obtained using pilot-based estimation during the initial half-duplex (HD) phase of the communication. In addition, the transmitted symbols at time  $n$  are detected using the initial channel estimates and the estimates of the state vector  $\boldsymbol{\beta}$  at time  $n-1$ . This is because at time  $n$  of the EKF algorithm,  $\boldsymbol{\beta}(n-1)$  has been successfully estimated. This procedure is shown in Fig. 2.

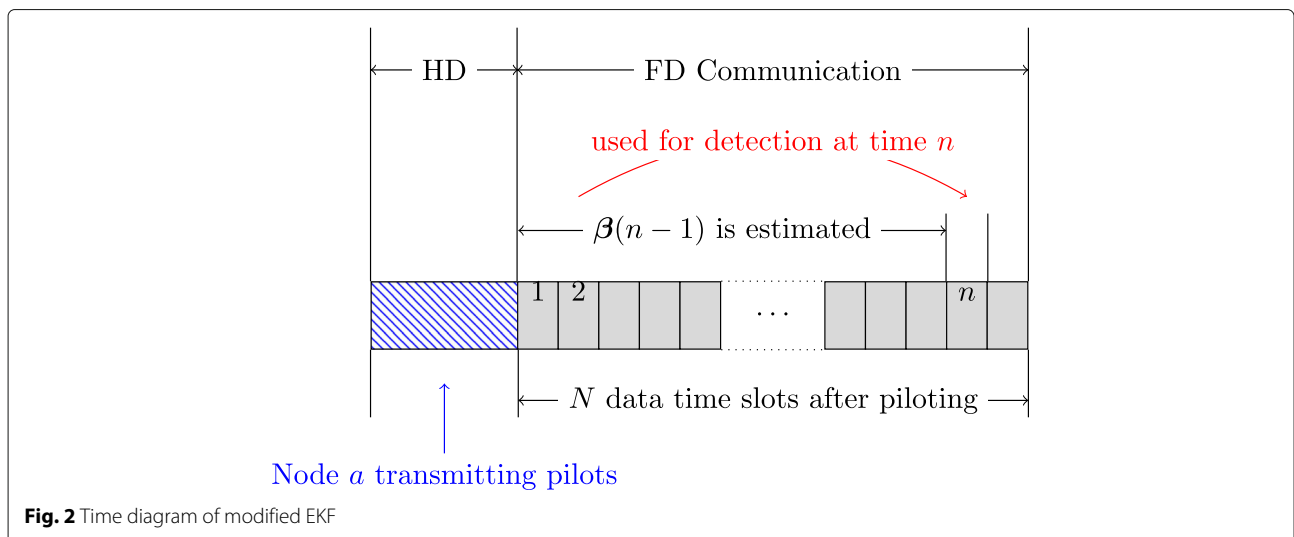


Fig. 2 Time diagram of modified EKF

### 3.2 Lower bound of estimation error

In this section, we derive a lower bound on the estimation error of the estimator proposed in the previous subsection. We first note that the mean square error (MSE) for estimating the state vector  $\beta(n)$  is given by

$$\text{MSE} = \text{trace} \left( \mathbb{E} \left[ \left( \beta(n) - \hat{\beta}(n) \right) \left( \beta_n - \hat{\beta}_n \right)^T \right] \right) \quad (18)$$

With the above definition of the MSE vector, we present the following proposition.

**proposition 1** *MSE of the EKF is lower bounded by trace (Q), i.e.,*

$$\text{MSE} \geq \text{trace} (\mathbf{Q}), \quad (19)$$

where  $\mathbf{Q}$  is the state covariance matrix given by (9).

*Proof* See Appendix 5. □

**Remark 2** *We note that (19) shows that the lower bound on the estimation error increases as the sum of diagonal elements of the covariance matrix of the states increases. Furthermore, (9) indicates that the diagonal elements of the state covariance matrix are the function of PN variance. Consequently, increasing the PN variance will result in worse estimation error. Since the residual SI cancelation is performed using the estimated SI channel, increasing the PN variance will result in worse SI cancelation performance. It is also worth to note that [34] shows that the PN variance is a monotonic increasing of function of carrier frequency. This means that the estimation error increases with increasing the carrier frequency and vice versa.*

### 3.3 Complexity analysis of EKF

For the complexity analysis of the proposed joint channel and PN estimation technique, we take the approach used by [36, 37] and count the number of multiplications and additions used in each step of EKF algorithm. Table 2 shows the complexity of each step of EKF algorithm using  $\mathcal{O}$ -notation. The corresponding complexity calculations for this table can be found in Appendix 5.

**Table 2** Complexity of each step of EKF algorithm

Equation no.	Complexity
(11)	$\mathcal{O}(N_t N_r)$
(13)	$\mathcal{O}(N_t^2 N_r^2)$
(14)	$\mathcal{O}(N_t^2 N_r^2)$
(15)	$\mathcal{O}(N_t^3 N_r^3)$

**Remark 3** *According to Table (2), the EKF has a polynomial complexity as a function of number of transmit  $N_t$  and receive  $N_r$  antennas. We can justify the increased complexity as follows. In [20], the authors propose an algorithm for channel estimation with linear complexity. However, the algorithm in [20] assumes a constant PN for a block of data. This could be an acceptable scenario in microwave communication but does not suit mmWave communication. Hence, the increased complexity of the proposed algorithm is justified because of fast variation of PN, i.e., PN variation over symbol time.*

## 4 Simulation results

In this section, we present simulation results for MIMO FD systems at 60 GHz frequency, which corresponds to mmWave frequency band [3]. For each simulation run, we assume a communication packet is 40 symbol long, i.e.,  $N = 40$ . This communication packet is transmitted after the training packet, which is  $2N_t$  symbols long, and is used for estimating the constant channels for EKF initialization as described in Section 3.1. We then use 10,000 simulation runs to obtain the desired simulation results.

Moreover, we use the assumptions presented in Section 2 to generate the random noise and PN. As summarized in [38], there are many mmWave channel models available for mmWave systems. In this work, similar to a large number of existing works in [24, 29, 39–41], we adopt a general Rician model. Note that the proposed estimator is independent of the adopted model. A performance comparison of the different mmWave channel models is outside the scope of this work.

We generate the random SI and communication channel ( $\mathbf{H}_{\text{SI/COM}}$ ) using Rician distribution as follows:

$$\mathbf{H}_{\text{SI/COM}} = \sqrt{\frac{K}{K+1}} \mathbf{H}_{\text{LoS}} + \sqrt{\frac{1}{K+1}} \mathbf{H}_{\text{NLoS}}, \quad (20)$$

where  $K$  is the Rician distribution  $K$ -factor;  $\mathbf{H}_{\text{LoS}}$  is the LoS component of the channel, and is generated assuming uniform distribution for angle of arrival, using the approach presented in [24];  $\mathbf{H}_{\text{NLoS}}$  is the NLoS component of the channel; and for both SI and communication channel is generated assuming Rayleigh fading. Furthermore, for both the SI and communication channel, we set the  $K$ -factor to 2 dB.

We note that the SI and communication channels have different power intensities, i.e.,  $\mathbb{E} \left[ \mathbf{H}_{\text{SI}} \mathbf{H}_{\text{SI}}^\dagger \right] \neq \mathbb{E} \left[ \mathbf{H}_{\text{COM}} \mathbf{H}_{\text{COM}}^\dagger \right]$ . Assuming that the LoS power of the residual SI (SI signal after the passive and analog cancelation) is the same as the LoS power of the communication signal, the signal to interference ratio (SIR) is given by:

$$\text{SIR} = \frac{\sigma_{\text{COM}}^2}{\sigma_{\text{SI}}^2}, \quad (21)$$

where  $\sigma_{\text{COM}}^2$  and  $\sigma_{\text{SI}}^2$  are the variances of NLoS components of the communication and SI channels, respectively.

In addition, SNR is defined as

$$\text{SNR} \triangleq \frac{\mathbb{E}[E_s]}{\sigma^2}, \tag{22}$$

where  $E_s$  is the symbol energy,  $\mathbb{E}[E_s] = 1$ , and  $\sigma^2$  is the noise variance.

Finally, we use the MSE for the state vector at time  $N = 40$ . This MSE is given by rewriting (18) in terms of the Euclidean norm of a vector, i.e.,  $\|\cdot\|_2$ ,

$$\mathbb{E} \left[ \left\| \boldsymbol{\beta}(N) - \hat{\boldsymbol{\beta}}(N) \right\|_2^2 \right]. \tag{23}$$

In what follows, we first present the MSE results for different FD MIMO communication systems. We then investigate the residual SI power after digital cancelation and the bit error rate (BER) performance of these systems with the proposed PN estimation technique.

### 4.1 MSE performance

In this section, we investigate the MSE performance of the proposed PN estimation technique for a  $2 \times 2$  FD MIMO system and assume that SIR = 0 dB, i.e., the SI signal is as strong as the desired communication signal.

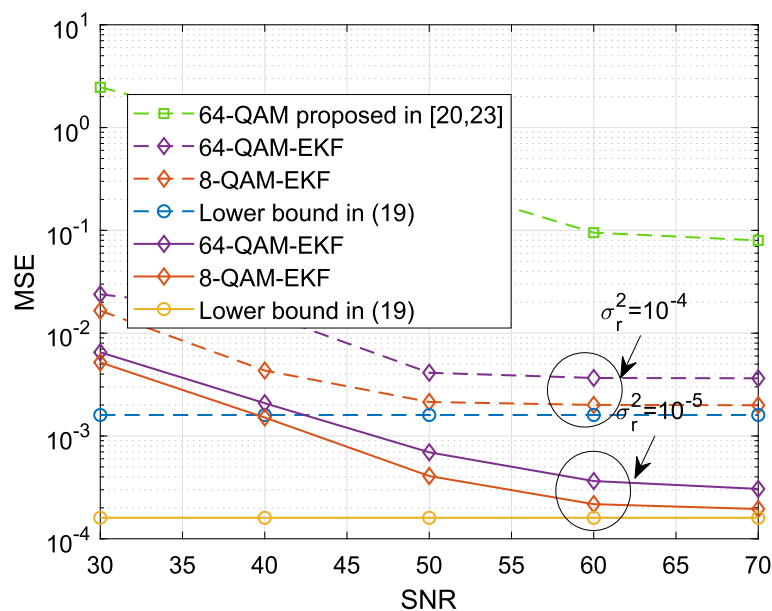
Figure 3 shows the MSE performance of the proposed system against the derived theoretical bound in Section 3.2 for different quadrature amplitude modulations (QAM) and different PN variances. Firstly, as discussed in Remark 2, with increasing PN variance, the estimation performance degrades. Secondly, it can be observed from this figure that lower order modulations

have better performance compared to the higher order modulations. This is because as shown in Section 3.1, the EKF algorithms require to detect the transmitted symbols. Hence, the MSE of EKF is affected by the detection error. Finally, Fig. 3 shows that at high SNRs, the MSE performance of the proposed estimator approaches the lower bound.

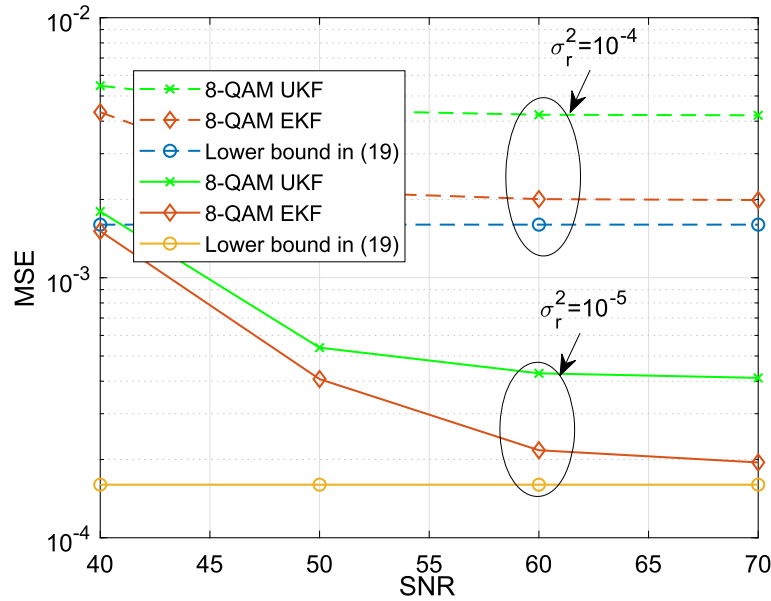
In Fig. 3, we also plot the MSE result of the state-of-the-art pilot-based phase noise estimator in [20, 23] for microwave frequency. As expected, this estimator does not perform well compared to our proposed estimator. This is because it assumes that the PN variations are small, which is not applicable for the case for mmWave frequency. Note that we only show the MSE result of the estimator in [20, 23] for 64-QAM modulation since the MSE performance is invariant with respect to the modulation order (the estimator uses pilots and does not require detection).

### 4.2 Comparison with unscented Kalman filter

We compare the performance of the proposed EKF estimator with unscented Kalman filter (UKF). UKF provides an alternative for linearizing the observations. The detailed implementation of the UKF is provided in Appendix 5. Figure 4 shows the performance of the EKF and UKF estimators for 8-QAM modulation, SIR = 0 dB, and different PN variances. We can see that the MSE performance of the proposed EKF estimator is better than the UKF estimator. This is because (i) UKF estimator works with the sigma point approximation of the mean of the state process, while EKF tracks the PN based on the true



**Fig. 3** MSE performance for PN variances  $\sigma_r^2 = \sigma_t^2 = \sigma_{\text{SI}}^2 = 10^{-4}, 10^{-5}$  and different QAM modulations for a  $2 \times 2$  FD MIMO system with SIR = 0 dB



**Fig. 4** MSE performance of the UKF and proposed EKF for PN variances  $\sigma_r^2 = \sigma_t^2 = \sigma_{SI}^2 = 10^{-4}, 10^{-5}$  and 8-QAM modulation for a  $2 \times 2$  FD MIMO system with SIR= 0 dB

mean of the linear state vector; (ii) while the MSE performances of both EKF and UKF are degraded because of the detection error, this error affects UKF algorithm more than EKF. This is because the sigma points calculation are affected more by the error due to the symbol detection (Section 3.1); and (iii) UKF is inherently more suitable for the systems which experience high nonlinearities, i.e., both the state and process models are nonlinear and noise is nonlinear too. In our case, only the process model in (7) is nonlinear.

### 4.3 Residual SI power

In this section, we numerically investigate the remaining SI power after digital cancellation for a  $2 \times 2$  MIMO FD system with 64-QAM modulation, assuming the PN variance for all the oscillators is  $10^{-4}$ . This residual power is given by

$$P_{SI} = \left\| \left( \mathbf{H}^{SI}(n) - \tilde{\mathbf{H}}^{SI}(n) \right) \mathbf{x}^{SI}(n) \right\|_2, \quad (24)$$

where  $\|\cdot\|_2$  is the Euclidean norm of a vector, and  $\tilde{\mathbf{H}}^{SI}(n)$  is an estimate of the SI channel using the proposed EKF estimator. Figure 5 shows the residual SI power for different SIR values, where a SIR value of 0 dB indicates that passive and analog cancellation stages have managed to reduce the SI power to the same level as the desired signal power.

The numerical result of Fig. 5 shows that the performance of digital canceler depends on the residual SI power after passive and analog cancellation stages. As the residual SI power after passive and analog cancellation decreases,

so does the residual SI power after the digital cancellation. The results show that the residual SI power can be reduced to  $-25$  and  $-40$  dB for SIR of 0 and 15 dB, respectively. This is important as it shows the effectiveness of digital SI cancellation after passive and analog cancellation.

### 4.4 BER performance

Finally, in this section, we present the BER results of a  $2 \times 2$  FD MIMO system with different QAM modulations, assuming that PN variance for all oscillators is  $10^{-4}$ . Figure 6 shows the BER performance of the system for different values of SIR. The results are consistent with the results of the residual SI power in Fig. 5, i.e., the higher the SIR, the better the BER results. Furthermore, 8-QAM system performs better than the 64-QAM system, which is consistent with the results of Fig. 6.

## 5 Conclusion

In this paper, we considered a MIMO FD system for mmWave communication and proposed a joint channel and PN estimation algorithm<sup>1</sup>. We also derived a lower bound on the estimation error and numerically showed that the MSE of the proposed estimator approaches the error bound. Furthermore, we investigated the residual SI power after the digital cancellation and showed that the digital canceler, which uses the estimated SI channel can reduce the SI power to  $-25$  to  $-40$  dB. These results indicate the effectiveness of digital cancellation after passive and analog cancellation stages.



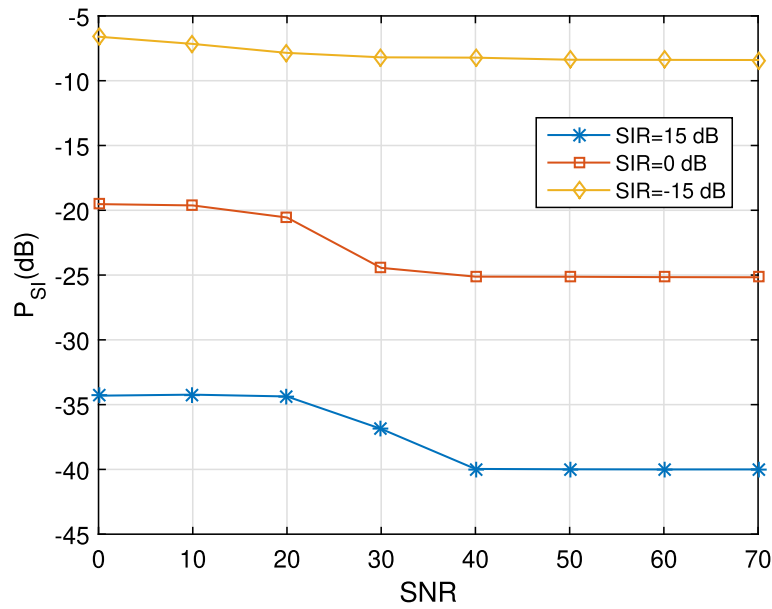


Fig. 5 The residual SI power  $P_{S_i}$  after digital cancellation

**Endnote**

<sup>1</sup>Indeed, the main focus of this work is to correctly estimate the channel and PN for effective SI cancellation. In case of inter-node interference [45], the proposed estimator would need to be modified. However, in the special case, if the inter-node interference can be treated as Gaussian, then the system model given by (1) can

capture the effect of the inter-node interference by including an additional Gaussian noise term due to inter-node interference.

**A lower bound of the estimation error**

In this section, we derive the lower bound of the estimation error. We start the proof by expanding

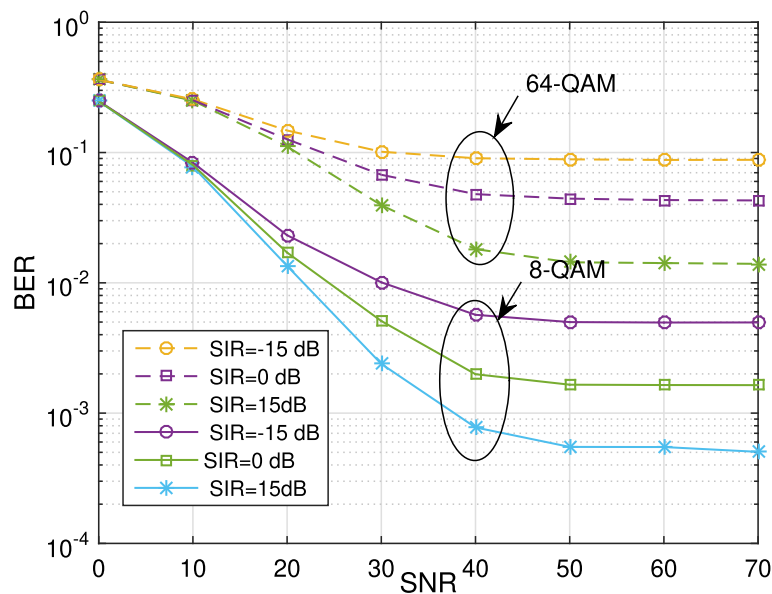


Fig. 6 BER performance of the proposed system for a 2 × 2 MIMO FD system with different QAM modulations

$$\begin{aligned} & \mathbb{E} \left[ \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right) \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right)^T \right]. \\ & \mathbb{E} \left[ \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right) \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right)^T \right] = \mathbb{E} \left[ \boldsymbol{\beta}(n) \boldsymbol{\beta}^T(n) \right] \\ & \quad + \mathbb{E} \left[ \hat{\boldsymbol{\beta}}(n) \hat{\boldsymbol{\beta}}^T(n) \right] - \mathbb{E} \left[ \boldsymbol{\beta}(n) \hat{\boldsymbol{\beta}}^T(n) \right] - \mathbb{E} \left[ \hat{\boldsymbol{\beta}}(n) \boldsymbol{\beta}^T(n) \right] \end{aligned} \quad (25)$$

Next, we show that the last two terms of (25) are zero. We do this by showing only  $\mathbb{E} \left[ \boldsymbol{\beta}(n) \hat{\boldsymbol{\beta}}^T(n) \right] = 0$  as a similar approach can be used to show that  $\mathbb{E} \left[ \hat{\boldsymbol{\beta}}(n) \boldsymbol{\beta}^T(n) \right] = 0$ .

We first note that  $\boldsymbol{\beta}(n)$  given by (8) is a Gaussian autoregressive model (AR) with mean zero, i.e.,  $\mathbb{E} [\boldsymbol{\beta}(n)] = 0$ . Hence,

$$\begin{aligned} \mathbb{E} \left[ \boldsymbol{\beta}(n) \hat{\boldsymbol{\beta}}^T(n) \right] &= \int \int \boldsymbol{\beta}(n) \hat{\boldsymbol{\beta}}(n) p(\boldsymbol{\beta}(n), \boldsymbol{y}(n)) d\boldsymbol{\beta}(n) d\boldsymbol{y}(n) \\ &= \int \hat{\boldsymbol{\beta}}(n) \int \boldsymbol{\beta}(n) p(\boldsymbol{\beta}(n)) d\boldsymbol{\beta}(n) p(\boldsymbol{y}(n) | \boldsymbol{\beta}(n)) d\boldsymbol{y}(n) \\ &= \int \hat{\boldsymbol{\beta}}(n) \mathbb{E} [\boldsymbol{\beta}(n)] p(\boldsymbol{y}(n) | \boldsymbol{\beta}(n)) d\boldsymbol{y}(n) = 0. \end{aligned} \quad (26)$$

Consequently, we can rewrite (25) as follows:

$$\begin{aligned} & \mathbb{E} \left[ \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right) \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right)^T \right] \\ &= \mathbb{E} \left[ \boldsymbol{\beta}(n) \boldsymbol{\beta}^T(n) \right] + \mathbb{E} \left[ \hat{\boldsymbol{\beta}}(n) \hat{\boldsymbol{\beta}}^T(n) \right] \end{aligned} \quad (27)$$

It is easy to show that  $\mathbb{E} \left[ \hat{\boldsymbol{\beta}}(n) \hat{\boldsymbol{\beta}}^T(n) \right]$  is a positive semi-definite matrix and hence

$$\mathbb{E} \left[ \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right) \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right)^T \right] \geq \mathbb{E} \left[ \boldsymbol{\beta}(n) \boldsymbol{\beta}^T(n) \right] \quad (28)$$

Furthermore, the properties of *trace* allows us to write

$$\text{trace} \left( \mathbb{E} \left[ \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right) \left( \boldsymbol{\beta}(n) - \hat{\boldsymbol{\beta}}(n) \right)^T \right] \right) \geq \text{trace} \left( \mathbb{E} \left[ \boldsymbol{\beta}(n) \boldsymbol{\beta}^T(n) \right] \right) \quad (29)$$

Finally, using (29) and the definitions of  $\mathbf{Q}$  and MSE in (9) and (18), we can establish the proof of the proposition.

### B Complexity analysis of EKF

In this section, we provide the complexity analysis of the EKF algorithm by counting the number of multiplications and additions. However, before we proceed, it can easily be shown that every entry of product of a  $K \times L$

matrix by a  $L \times M$  matrix requires  $L$  multiplications and  $L-1$  additions, and hence, the whole matrix requires  $KML$  multiplications and  $KM(L-1)$  additions, where  $KM$  is the size of the resulting matrix. Furthermore, it is known that matrix inversion has the same complexity in terms of additions and multiplication as the matrix multiplication, up to a multiplicative constant  $\gamma$  [42]. We can now proceed with calculating the complexity of EKF algorithm in (30) to (32).

$$\hat{\boldsymbol{\beta}}(n|n) = \hat{\boldsymbol{\beta}}(n|n-1) \Re \left\{ \mathbf{K}(n) \left( \mathbf{y}(n) - \underbrace{\bar{\mathbf{H}}(n) e^{j\hat{\boldsymbol{\beta}}(n|n-1)}}_{N_r(2N_tN_r) + N_r(2N_tN_r - 1)} \right) \right\} \quad (30)$$

$\underbrace{\hspace{10em}}_{2N_tN_r^2 + 2N_tN_r(N_r - 1) + N_r + N_r(2N_tN_r) + N_r(2N_tN_r - 1)}$

$$\begin{aligned} \mathbf{K}(n) &= \mathbf{M}(n|n-1) \mathbf{D}^\dagger(n) \\ & \quad \left( \sigma^2 \mathbf{I}_{N_r} + \mathbf{D}(n) \mathbf{M}(n|n-1) \mathbf{D}^\dagger(n) \right)^{-1}, \\ & 4N_t^2N_r^3 + 2N_tN_r^2(2N_tN_r - 1) + 2N_tN_r^3 + 2N_tN_r^2(N_r - 1) \\ & + \gamma N_r^3 + \gamma N_r^2(N_r - 1) + N_r^2 + 2N_tN_r^2 + N_r(2N_tN_r - 1) \\ & + 4N_t^2N_r^3 + 2N_tN_r^2(2N_tN_r - 1) \end{aligned}$$

$$\mathbf{M}(n|n-1) = \underbrace{\mathbf{M}(n-1|n-1) + \mathbf{Q}}_{4N_t^2N_r^2} \quad (31)$$

$$\mathbf{M}(n|n) = \Re \left\{ \mathbf{M}(n|n-1) \left( \mathbf{I}_{2N_tN_r} - \underbrace{\mathbf{K}(n) \mathbf{D}(n)}_{2N_tN_r^3 + 2N_tN_r^2(N_r - 1)} \right) \right\} \quad (32)$$

$\underbrace{\hspace{10em}}_{8N_t^3N_r^3 + 4N_t^2N_r^2(2N_tN_r - 1) + 2N_tN_r^3 + 2N_tN_r^2(N_r - 1)}$

### C Unscented Kalman filter (UKF)

Unscented Kalman filter (UKF) provides an alternative to EKF for nonlinear state vector estimation. In UKF instead of linearizing the observation vector, the probability distributions of states and observations are approximated using *sigma points* [43]. UKF can solve a very general class of problems, where both state process and observations are nonlinear. However, the joint channel and PN estimation problem, as given by the observation vector (7) and

the state vector (8), has a linear state process and additive noise. This allows for the use of non-augmented state vectors for UKF [44]. For the state vector  $\beta(n)$  in (8), the sigma points  $\mathcal{B}(i, n)$  are given by

$$\mathcal{B}(0, n|n-1) = \mathcal{B}(0, n-1), \tag{33a}$$

$$\mathcal{B}(i, n|n-1) = \mathcal{B}(i, n-1) + \left(\sqrt{\{(L+\lambda)\mathbf{Q}\}}\right)_i, \tag{33b}$$

$$i = 1, \dots, L,$$

$$\mathcal{B}(i, n|n-1) = \mathcal{B}(i, n-1) - \left(\sqrt{\{(L+\lambda)\mathbf{Q}\}}\right)_i, \tag{33c}$$

$$i = L+1, \dots, 2L$$

where,  $\sqrt{\{\cdot\}}$  is the matrix square root,  $(\cdot)_i$  is the  $i$ th column of the matrix,  $L = 4N_tN_r$ ,  $\lambda = \alpha^2L - L$ , where  $\alpha = 10^{-3}$  [43], and  $\mathbf{Q}$  is the state covariance matrix given by (9). Subsequently, the mean of the sigma points, which is used as an approximate to the true mean of the probability distribution of states, is given by

$$\bar{\beta}(n) = \sum_{i=0}^{2L} \mathcal{W}_i^m \mathcal{B}(i, n|n-1), \tag{34}$$

where,

$$\mathcal{W}_0^m = \frac{\lambda}{L+\lambda} \tag{35}$$

$$\mathcal{W}_i^m = \frac{1}{2(L+\lambda)}, \quad i = 1, \dots, 2L. \tag{36}$$

Similarly, the covariance of the state vector based on the sigma points approximation is given by

$$\bar{\mathbf{P}}_n = \sum_{i=0}^{2L} \mathcal{W}_i^m [\mathcal{B}(i, n|n-1) - \bar{\beta}(n)] [\mathcal{B}(i, n|n-1) - \bar{\beta}(n)]^* \tag{37}$$

Moreover, the sigma points for the observations, and the corresponding approximate mean of probability distribution of observations are given by

$$\mathcal{Y}(n|n-1) = \bar{\mathbf{H}}(n)e^{j\mathcal{B}(i,n|n-1)}, \tag{38a}$$

$$\bar{\mathbf{y}}(n) = \sum_{i=0}^{2L} \mathcal{W}_i^c \mathcal{Y}(n|n-1), \tag{38b}$$

where  $\mathcal{W}_0^c = \mathcal{W}_0^m + (1 - \alpha^2 + \beta)$ ,  $\beta = 2$ , and  $\mathcal{W}_i^c = \mathcal{W}_i^m$  for  $i = 1, \dots, 2L$ . Once the state and the process models are approximated by the sigma points using (33a)–(33c), and (38a), respectively, the updated mean  $\hat{\beta}(n)$  and variance  $\hat{\mathbf{P}}_n$  can be calculated as follows:

$$\hat{\beta}(n) = \bar{\beta}(n) + \mathcal{K}(\mathbf{y}(n) - \bar{\mathbf{y}}(n)), \tag{39}$$

$$\hat{\mathbf{P}}_n = \bar{\mathbf{P}}_n - \mathcal{K}\mathbf{P}_{y,y}\mathcal{K}^T, \tag{40}$$

---

**Algorithm 1** UKF for joint channel and PN estimation

---

- 1: **Initialize:**  
 $\hat{\beta}(0) = \mathbf{0}_{2N_tN_r \times 1}$   
 $\mathbf{P}_0 = \mathbf{Q}$
  - 2: **for**  $n = 0$  to  $N$  **do**
  - 3:   Calculate sigma points using (33a) to (33c).
  - 4:   Calculate mean of state  $\bar{\beta}(n)$  using (34).
  - 5:   Calculate the covariance matrix of the state vector using (37).
  - 6:   Calculate the sigma points for observations using (38a).
  - 7:   Calculate the mean of observation using (38b).
  - 8:   Update the mean  $\hat{\beta}(n)$  using (39).
  - 9:   Update the variance  $\hat{\mathbf{P}}_n$  using (40).
- 

where

$$\mathcal{K} = \mathbf{P}_{x,y}\mathbf{P}_{y,y}^{-1}, \tag{41}$$

$$\mathbf{P}_{x,y} = \sum_{i=0}^{2L} \mathcal{W}_i^c [\mathcal{B}(i, n|n-1) - \bar{\beta}(n)] [\mathcal{Y}(n|n-1) - \bar{\mathbf{y}}(n)]^*, \tag{42}$$

$$\mathbf{P}_{y,y} = \sum_{i=0}^{2L} \mathcal{W}_i^c [\mathcal{Y}(n|n-1) - \bar{\mathbf{y}}(n)] [\mathcal{Y}(n|n-1) - \bar{\mathbf{y}}(n)]^*. \tag{43}$$

Algorithm 1 summarizes the UKF joint channel and PN estimation algorithm.

**Acknowledgements**

The authors would like to thank Professor Taneli Rihhonen for his insightful comments on this work.

**Funding**

The work of Abbas Koohian was supported by an Australian Government Research Training Program (RTP) Scholarship. The work of Hani Mehrpouyan was partially funded by the NSF ERAS grant award number 1642865.

**Authors' contributions**

All authors contributed equally to this work. The final manuscript has been read and approved by all authors for submission.

**Competing interests**

The authors declare that they have no competing interests.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Author details**

<sup>1</sup>Research School of Electrical, Energy, and Materials Engineering, Australian National University, Canberra, Australia. <sup>2</sup>Department of Electrical and Computer Engineering, Boise State University, Boise, ID, USA. <sup>3</sup>Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

Received: 13 November 2018 Accepted: 21 February 2019

Published online: 15 March 2019

## References

1. S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, J. Rodriguez, Millimeter-wave massive MIMO communication for future wireless systems: a survey. *IEEE Commun. Surveys Tuts.* **20**(2), 836–869 (2018)
2. R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, A. M. Sayeed, An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE J. Sel. Topics Signal Process.* **10**(3), 436–453 (2016)
3. T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, F. Gutierrez, Millimeter wave mobile communications for 5G cellular: it will work! *IEEE Access.* **1**, 335–349 (2013)
4. S. Rajagopal, S. Abu-Surra, J. Zhang, in *Proc. IEEE SPAWC*. Spectral mask filling for PAPR reduction in large bandwidth mmWave systems, (2015), pp. 131–135
5. S. Buzzi, C. D'Andrea, T. Foggi, A. Ugolini, G. Colavolpe, Single-carrier modulation versus OFDM for millimeter-wave wireless MIMO. *IEEE Trans. Commun.* **66**(3), 1335–1348 (2018)
6. M. Duarte, C. Dick, A. Sabharwal, Experiment-driven characterization of full-duplex wireless systems. *IEEE Trans. Wireless Commun.* **11**(12), 4296–4307 (2012)
7. Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, L. Hanzo, Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection. *IEEE Commun. Mag.* **53**(5), 128–137 (2015)
8. H. Mehrpouyan, M. R. Khanzadi, M. Matthaiou, A. M. Sayeed, R. Schober, Y. Hua, Improving bandwidth efficiency in E-band communication systems. *IEEE Commun. Mag.* **52**(3), 121–128 (2014)
9. V. Syrjala, M. Valkama, L. Anttila, T. Riihonen, D. Korpi, Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers. *IEEE Trans. Wirel. Commun.* **13**(6), 2977–2990 (2014)
10. S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, P. Levis, Applications of self-interference cancellation in 5G and beyond. *IEEE Commun. Mag.* **52**(2), 114–121 (2014)
11. A. Koohian, H. Mehrpouyan, M. Ahmadian, M. Azarbad, in *Proc. IEEE ICC*. Bandwidth efficient channel estimation for full duplex communication systems, (2015), pp. 4710–4714
12. M. Duarte, A. Sabharwal, in *Proc. Asilomar Conf. on Signals, Syst. and Computers*. Full-duplex wireless communications using off-the-shelf radios: feasibility and first results, (2010), pp. 1558–1562
13. M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. K. Ramakrishnan, C. W. Rice, N. K. Shankaranarayanan, Design and characterization of a full-duplex multi-antenna system for WiFi networks. *IEEE Trans. Veh. Technol.* **63**(3), 1160–1177 (2014)
14. E. Everett, A. Sahai, A. Sabharwal, Passive self-interference suppression for full-duplex infrastructure nodes. *IEEE Trans. Wirel. Commun.* **13**(2), 680–694 (2014)
15. A. A. Nasir, S. Durrani, H. Mehrpouyan, S. D. Blostein, R. A. Kennedy, Timing and carrier synchronization in wireless communication systems: a survey and classification of research in the last 5 years. *EURASIP J. Wirel. Commun. Netw.* **180**(1) (2016). Available: <https://doi.org/10.1186/s13638-016-0670-9>
16. E. Ahmed, A. M. Eltawil, On phase noise suppression in full-duplex systems. *IEEE Trans. Wirel. Commun.* **14**(3), 1237–1251 (2015)
17. B. Lee, J. B. Lim, C. Lim, B. Kim, J. Y. Seol, in *Proc. IEEE Globecom Workshops*. Reflected self-interference channel measurement for mmWave beamformed full-duplex system, (2015), pp. 1–6
18. A. Demir, T. Haque, E. Bala, P. Cabrol, in *Proc. WAMICON*. Exploring the possibility of full-duplex operations in mmWave 5G systems, (2016), pp. 1–5
19. T. A. Thomas, M. Cudak, T. Kovarik, in *Proc. IEEE ICC*. Blind phase noise mitigation for a 72 GHz millimeter wave system, (2015), pp. 1352–1357
20. A. Masmoudi, T. Le-Ngoc, A maximum-likelihood channel estimator for self-interference cancellation in full-duplex systems. *IEEE Trans. Veh. Technol.* **65**(7), 5122–5132 (2016)
21. A. Masmoudi, T. Le-Ngoc, Channel estimation and self-interference cancellation in full-duplex communication systems. *IEEE Trans. Veh. Technol.* **66**(1), 321–334 (2017)
22. X. Xiong, X. Wang, T. Riihonen, X. You, Channel estimation for full-duplex relay systems with large-scale antenna arrays. *IEEE Trans. Wireless Commun.* **15**(10), 6925–6938 (2016)
23. R. Li, A. Masmoudi, T. Le-Ngoc, Self-interference cancellation with nonlinearity and phase-noise suppression in full-duplex systems. *IEEE Trans. Veh. Technol.* **67**(3), 2118–2129 (2018)
24. H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. K. Karagiannidis, T. Svensson, Joint estimation of channel and oscillator phase noise in MIMO systems. *IEEE Trans. Signal Process.* **60**(9), 4790–4807 (2012)
25. D. Korpi, L. Anttila, V. Syrjala, M. Valkama, Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver. *IEEE J. Sel. Areas Commun.* **32**(9), 1674–1687 (2014)
26. D. Korpi, T. Riihonen, V. Syrjala, L. Anttila, M. Valkama, R. Wichman, Full-duplex transceiver system calculations: analysis of ADC and linearity challenges. *IEEE Trans. Wirel. Commun.* **13**(7), 3821–3836 (2014)
27. L. Samara, M. Mokhtar, O. Ozdemir, R. Hamila, T. Khattab, Residual self-interference analysis for full-duplex OFDM transceivers under phase noise and I/Q imbalance. *IEEE Commun. Lett.* **21**(2), 314–317 (2017)
28. E. Ahmed, A. M. Eltawil, A. Sabharwal, Rate gain region and design tradeoffs for full-duplex wireless communications. *IEEE Trans. Wirel. Commun.* **12**(7), 3556–3565 (2013)
29. A. A. Nasir, H. Mehrpouyan, R. Schober, Y. Hua, Phase noise in MIMO systems: Bayesian Cramer Rao bounds and soft-input estimation. *IEEE Trans. Signal Process.* **61**(10), 2675–2692 (2013)
30. K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, A. Szufarska, A flexible 5G frame structure design for frequency-division duplex cases. *IEEE Commun. Mag.* **54**(3), 53–59 (2016)
31. S. Dutta, M. Mezzavilla, R. Ford, M. Zhang, S. Rangan, M. Zorzi, Frame structure design and analysis for millimeter wave cellular systems. *IEEE Trans. Wirel. Commun.* **16**(3), 1508–1522 (2017)
32. C. Gustafson, K. Haneda, S. Wyne, F. Tufvesson, On mm-wave multipath clustering and channel modeling. *IEEE Trans. Antennas Propag.* **62**(3), 1445–1455 (2014)
33. D. Kim, H. Ju, S. Park, D. Hong, Effects of channel estimation error on full-duplex two-way networks. *IEEE Trans. Veh. Technol.* **62**(9), 4666–4672 (2013)
34. M. R. Khanzadi, R. Krishnan, D. Kuylentierna, T. Eriksson, in *Proc. IEEE Globecom Workshops*. Oscillator phase noise and small-scale channel fading in higher frequency bands, (2014), pp. 410–415
35. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. (Prentice-Hall, Inc., Upper Saddle River, 1993)
36. A. A. Nasir, H. Mehrpouyan, S. D. Blostein, S. Durrani, R. A. Kennedy, Timing and carrier synchronization with channel estimation in multi-relay cooperative networks. *IEEE Trans. Signal Process.* **60**(2), 793–811 (2012)
37. A. Koohian, H. Mehrpouyan, A. A. Nasir, S. Durrani, S. D. Blostein, Superimposed signaling inspired channel estimation in full-duplex systems. *EURASIP J. Adv. Signal Process.* **8**(2018). Available: <https://doi.org/10.1186/s13634-018-0529-9>
38. I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, L. Hanzo, Millimeter-wave communications: physical channel models, design considerations, antenna constructions, and link-budget. *IEEE Commun. Surveys Tuts.* **20**(2), 870–913 (2018)
39. A. G. Siamarou, Digital transmission over millimeter-wave radio channels: a review [wireless corner]. *IEEE Antennas Propag. Mag.* **51**(6), 196–203 (2009)
40. J. Zhang, L. Dai, X. Zhang, E. Bjornson, Z. Wang, Achievable rate of rician large-scale MIMO channels with transceiver hardware impairments. *IEEE Trans. Veh. Technol.* **65**(10), 8800–8806 (2016)
41. T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, S. Deng, Small-scale, local area, and transitional millimeter wave propagation for 5G communications. *IEEE Trans. Antennas Propag.* **65**(12), 6474–6490 (2017)
42. S. Arora, B. Barak, *Computational Complexity: A Modern Approach*, 1st edn. (Cambridge University Press, New York, 2009)
43. E. A. Wan, R. V. D. Merwe, in *Proc. IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*. The unscented Kalman filter for nonlinear estimation, (2000), pp. 153–158
44. Y. Wu, D. Hu, M. Wu, X. Hu, Unscented Kalman filtering for additive noise case: augmented versus nonaugmented. *IEEE Signal Process. Lett.* **12**(5), 357–360 (2005)
45. W. Feng, Y. Wang, D. Lin, N. Ge, J. Lu, S. Li, When mmWave communications meet network densification: a scalable interference coordination perspective. *IEEE J. Sel. Areas Commun.* **35**(7), 1459–1471 (2017)