A Bayesian Approach for Nonlinear Equalization and Signal Detection in Millimeter-Wave Communications

Bin Li, Chenglin Zhao, Mengwei Sun, Haijun Zhang, Zheng Zhou, *Member, IEEE*, and Arumugam Nallanathan, *Senior Member, IEEE*

***Abstract*—For the emerging 5G millimeter-wave communications, the nonlinearity is inevitable due to RF power amplifiers of the enormous bandwidth operating in extremely high frequency, which, in collusion with frequency-selective propagations, may pose great challenges to signal detections. In contrast to classical schemes, which calibrate nonlinear distortions in transmitters, we suggest a nonlinear equalization algorithm, with which the multipath channel and unknown symbols contaminated by nonlinear distortions and multipath interferences are estimated in receiver-ends. Attributed to the nonlinearity and marginal integration, the involved posterior density is analytically intractable and, unfortunately, most existing linear equalization schemes may become invalid. To solve this problem, the Monte-Carlo sequential importance sampling based particle filtering is suggested, and the non-analytical distribution is approximated numerically by a group of random measures with the evolving probability-mass. By applying the Taylor's series expansion technique, a local-linearization observation model is further constructed to facilitate the practical design of a sequential detector. Thus, the unknown symbols are detected recursively as new observations arrive. Simulation results validate the proposed joint detection scheme. By excluding transmitting pre-distortion of high complexity, the presented algorithm is specially designed for the receiver-end, which provides a promising framework to nonlinear equalization and signal detection in millimeter-wave communications.**

***Index Terms*—5G millimeter-wave communications, nonlinear equalization, signal detection, nonlinear power amplifier, Bayesian recursive approach, particle filtering.**

Manuscript received January 9, 2014; revised July 10, 2014 and November 2, 2014; accepted February 21, 2015. Date of publication March 11, 2015; date of current version July 8, 2015. This work was supported by NSFC (61271180). This work also received support from The Ministry of Knowledge Economy (MIKE), Korea, under the Information Technology Research Center (ITRC) support program supervised by the National IT Industry Promotion Agency (NIPA) (NIPA-2011-C1090-1111-0007). The associate editor coordinating the review of this paper and approving it for publication was M. R. Bhatnagar.

B. Li, C. Zhao, M. Sun, and Z. Zhou are with the School of Information and Communication Engineering (SICE), Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China (e-mail: stonebupt@gmail.com).

H. Zhang is with the Department of Information Engineering, Beijing University of Chemical Technology (BUCT), Beijing 100029, China (e-mail: haijun.zhang@kcl.ac.uk).

A. Nallanathan is with the Department of Informatics, King's College London, London WC2R 2LS, U.K. (e-mail: nallanathan@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2412119

I. Introduction

ATTRIBUTED to the availability of abundant unauthorized spectrum resources in millimeter-wave (mm-Wave) frequency band, wireless personal communication networks (WPANs) and wireless local area networks (WLANs) operating in such frequencies (e.g., 60 GHz) have recently aroused general interests [1], [2]. The mainstream 60 GHz mm-Wave communication standards, i.e., IEEE 802.15.3c and 802.11ad [3], [4], are mainly oriented towards 5G communications aiming to support the high-speed data transmissions such as Uncompressed High Definition Television (UHDTV) and Gbps wireless accessing (e.g., Wi-Fi applications) [5], [6]. In all such scenarios, single carrier (SC) modulation has been recommended as one of the promising physical layer (PHY) techniques [7]. To achieve the ultra-high data rate up to 7 Gbps with a regulated transmission bandwidth of 2.16 GHz, high-order modulations such as M-order phase shift key (MPSK) and M-order quadrature amplitude modulation (M-QAM) have been commonly suggested to improve the spectrum efficiency furthermore [3], [4], [8].

As the absorption from oxygen to signals may reach a maximum in mm-Wave bands (about 15 dB/km), in practice, the propagation attenuation in such a high-frequency band is tremendous, correspondingly coming with a rather limited link budget [9]. To efficiently compensate the path-loss and, therefore, reinforce the signal-to-noise ratio (SNR) in receivers, usually high emission power is used in 60 GHz mm-Wave communications, except for the deployment of high resolution beam-forming techniques. Taking the high-order modulation signal with the high peak-to-average power ratio (PAPR) and the enormous bandwidth into accounts, mm-Wave communications may become extremely vulnerable to nonlinear distortions [10] and multipath propagations.

Unfortunately, the mm-Wave power amplifier (PA) may inevitably show nonlinear characteristics due to hardware imperfections [11], which may arouse serious nonlinear distortions in practice. The received signals, as a consequence, will be shifted sharply in the constellation-plane. Then, the symbols will be interpreted erroneously. In some bad cases, the bit error ratio (BER) will be significantly increased, leading to substantial performance deteriorations. In addition, further considering typical short-range indoor applications, the frequency-selective multipath propagations become very common to 60 GHz mm-Wave systems. The resulting multipath interference may also

remarkably degrade the detection performance. Thus, one of the major concerns in mm-Wave communications is to combat the linear and nonlinear distortions and, therefore, enhance the signal detection performance.

To overcome the destructive effect, many investigations have been dedicated to dealing with the PA nonlinearity in mm-Wave systems. A simple and direct approach is to reduce the radiation power [12], which makes the operational power away from the PA's saturation point and thereby alleviate nonlinear distortions to some extent. This output power back-off (OBO) method, nevertheless, may sacrifice the power efficiency and, as a consequence, the SNR in receiver may remarkably decline with a large OBO. Another approach to combat the PA nonlinearity is the widely recommended digital pre-distortion (DPD) techniques [13]–[17]. These methods may compensate or repair the operating condition of PAs [14], [15], which is practically implemented in the transmitter-end, by utilizing various complicated learning-based algorithms [15] or, alternatively, by a realization of the inverse nonlinear function deduced from the PA input-output response [16], [17]. Because such techniques are essentially designed for the baseband processing, the emitted radio-frequency (RF) signals are firstly required to be down-converted via an analog-to-digital converter (ADC). Subsequently, the baseband signal is calibrated via a feedback control loop (e.g., adaptive least mean squares scheme) or a look-up table (LUT). Given the additional *reception* sub-system in mm-Wave transmitters, the pre-distortion technique may be computationally intensive and impractical. More importantly, it is noted that such pre-distortion techniques can only compensate PA nonlinearity partly. Even utilizing such a complicated mechanism, serious nonlinear effects may still degrade the detection performance, especially when a small OBO is used and the frequency-selective multipath propagation of mm-Wave systems is further considered.

The first contribution of this investigation is that, by presenting a promising joint detection scheme, we directly address the nonlinear distortion and the frequency-selective multipath fading in the *receiver-end*. The original motivation of the new scheme is to alleviate the RF hardware requirements of mm-Wave devices by designing an effective baseband processing scheme. In sharp contrast to the widely used sophisticated pre-distorter in transmitter-ends, the received signal constellations, with serious distortions introduced by both PA contaminations and multipath interferences, are blindly calibrated via the Bayesian statistical inference. Since the formulated signal model involves both nonlinear distortions and linear multipath disruptions, the analytic form of the *a posteriori* probability of interest may be practically intractable. To cope with the difficulty, relying on the conception of Monte-Carlo random sampling [18], [19], a group of discrete measures with the evolving weights are simulated to approximate the realistic complex probability distribution function (PDF) numerically. Thus, the Bayesian inference is implemented by particle filtering (PF).

Although PF has been applied widely to blind equalizations (see [20]–[27]), as far as we know, there are few works in the literature reported on the design of PF-based joint detector in the presence of both nonlinear distortions and linear multipath propagations. The second contribution, correspondingly, is that we designed a *nonlinear equalization* scheme for mm-Wave systems. A Taylor series expansion technique, as in [19], is applied to the nonlinear observation function. On this basis, a generalized PF framework with the local-linearization process is suggested for the mixed linear and nonlinear problem. Thus, the joint estimation of unknown symbols accompanying the linear multipath response can be derived recursively on reception of new observations. Experimental simulations further validate the proposed nonlinear equalization scheme. It is demonstrated that, even with realistic nonlinear PAs, the detection performance of the new scheme may remarkably surpass other existing techniques, e.g., a joint linear equalization scheme and the DPD methods.

Related Work: Attributed to the pioneering works of P. M. Djuric *et al.* [20], [21], PF has been applied widely to joint estimations in the signal processing context. The linearization method based on the Taylor series expansion was firstly introduced by Doucet *et al.* [19]. In this investigation, however, two differentiated points may be noted. 1) While a nonlinear estimation framework has been formulated, either the system model or application scenarios remain relatively different from this specific problem. The considered observation model, in contrast, involves two *heterogeneous* components, i.e., the nonlinear PA cascading a linear multipath channel. With the coupling of nonlinear effects from realistic PAs, the estimation of linear multipath becomes more challenging. The signal processing objective, i.e., linear de-convolution with nonlinear distortions, seems also to be different. 2) A particular significance of our nonlinear equalization scheme is that, by excluding complicated RF repairing operations (i.e., pre-distortion) in transmitter-ends, it may provide a more attractive approach to implement joint detection with nonlinear effects. To the best of our knowledge, the nonlinear equalization in mm-Wave systems, by fully taking both realistic nonlinear distortions and linearly frequency-selective fading, remain still as an unexploited area. This work extends the PF conception to such a new application scenario and presents a promising solution for 5G mm-Wave signal detections with nonlinear distortions.

The rest of this article is structured as follows. In Section II, the nonlinear PA model and the propagation model for mm-Wave communications are briefly described. Next, a new dynamic state-space model (DSM) is established to characterize the channel estimation and joint signal detections with nonlinear PAs. In Section III, the PF and Bayesian inference for traditional linear equalizations are introduced, and some difficulties in the considered nonlinear detection are further analyzed. In Section IV, a generalized PF is suggested by resorting to a local-linearization technique. Thus, a recurrence nonlinear equalization scheme is developed to cope with both linear multipath propagations and nonlinear distortions. Experimental simulations are provided in Section V. We finally conclude the investigation in Section VI.

II. System Description

A. Nonlinear Power Amplifier

For broadband mm-Wave communications, the nonlinear effect is always of crucial importance to the input signals with

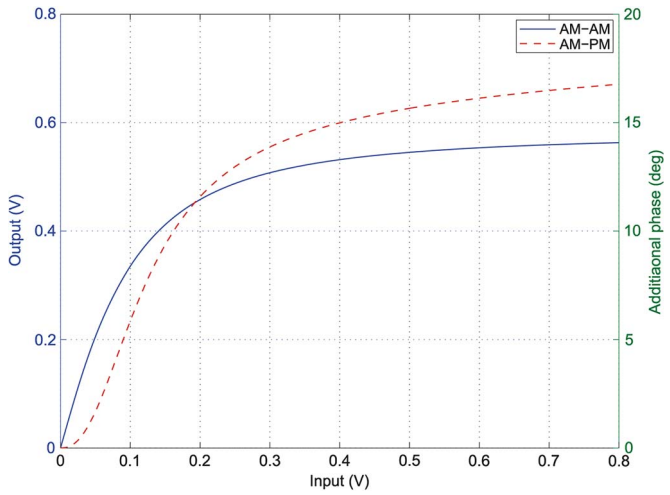


Fig. 1. Nonlinear PA model regulated by the IEEE 803.11ad TG.

a high operational power or PAPR. After propagated from a nonlinear PA, both the amplitude and phase of output signals will experience serious distortions, which can be commonly characterized by the amplitude modulation-amplitude modulation (AM-AM) model and the amplitude modulation-phase modulation (AM-PM) model [28], respectively. In this analysis, the nonlinear PA model regulated by the IEEE 802.11ad task group (TG) is considered [11].

Specifically, the AM-AM model is given by:

$$G(V_{in}) = \frac{gV_{in}}{[1 + (gV_{in}/V_{sat})]^{1/(2\sigma_p)}}, \tag{1}$$

Here, V_{in} and $G(V_{in})$ represent the input and output voltage amplitude in root mean square, respectively. g is the linear gain and we practically have $g = 4.65$; σ_p is the smoothness factor and $\sigma_p = 0.81$; V_{sat} denotes the saturation level and is typically configured to 0.58 V.

The AM-PM mode can be expressed as:

$$\Psi(V_{in}) = \frac{\alpha V_{in}^{q_1}}{1 + (V_{in}/\beta)^{q_2}}, \tag{2}$$

where $\Psi(V_{in})$ is the additional phase in degrees, and after multiplied by $\pi/180$ it can be transformed into radians form. The typical values of α, β, q_1, q_2 are respectively configured to 2560, 0.114, 2.4 and 2.3. For analysis clarity, we may denote the AM-AM model and AM-PM model together by the associative parameter set $(\mathbf{G}, \mathbf{\Psi}) \triangleq [g, \sigma_p, V_{sat}, \alpha, \beta, q_1, q_2]$.

Given the PA's parametric model, i.e., the input-output curves of AM-AM and AM-PM models shown in Fig. 1, the nonlinear distortion is mainly related with the input voltage V_{in}. To be specific, the *increment* on the output voltage $G(V_{in})$ will decrease with the increasing of the input voltage V_{in} and, therefore, the output-input voltage relations (i.e., the AM-AM curve) may become nonlinear. The additional phase of output signals, i.e., $\Psi(V_{in})$, on the other hand, will also increase with the increasing of the input voltage as shown by the AM-PM curve.

It is seen that if the input power p_{in} is relatively small, or the operational voltage keeps far away from the saturation point V_{sat}, then the PA will be treated as an ideal system with the linear gain and negligible phase shifts. In such a case, nonlinear distortions can be ignored without affecting the performance. In fact, this is also the principle of OBO techniques [12], in which the back-off value is measured by $-10log_{10}(p_{in}/p_{sat})$ dB. It follows that the nonlinear distortion may be alleviated partly by reducing the input power p_{in}. Note that, however, the efficiency of PA will decrease and the receiving SNR may not be guaranteed. With the further increasing of p_{in} (or the decreasing of the OBO value), the operational point may move towards to the saturation point and, correspondingly, the nonlinear distortion may be evident. I.e., the amplifier output may be increased slowly to a steady value by dramatically disrupting the linear relationship and, simultaneously, introducing remarkable phase shifts. As a result, signal constellations will be distorted and the detection performance will be degraded.

B. Channel Modeling for 60 GHz Communications

Attributed to rich scatters in typical indoor environments and the excellent time resolution of transmitted signals, the short-range mm-Wave channel is known to be linearly dispersive with tens of resolved multipath components (MPCs) [29], [30]. Currently, the approved channel model is built on a modified Saleh-Valenzuela (S-V) indoor model [29]. The IEEE 802.15.3c TG has defined seven channel types for mm-Wave applications in dense multipath environments. The unified channel impulse response (CIR) is expressed as:

$$\begin{aligned} h(t,\theta,\phi) &= \alpha_{0,0}\delta(\theta - \alpha_{0,0})\delta(\phi - \phi_{0,0})\delta(t - \tau_{0,0}) \\ &\quad + \sum_{n=1}^{N}\sum_{l=1}^{L_n} \alpha_{n,l}\delta(\theta - \alpha_{n,l})\delta(\phi - \phi_{n,l})\delta(t - \tau_{n,l}), \end{aligned} \tag{3}$$

where N is the number of paths (or clusters) and L_n is number of sub-paths of the nth cluster; $\alpha_{n,l} = |\alpha_{n,l}|exp(-j\varphi_{n,l})$ corresponds to the complex coefficient of the lth sub-path of the nth cluster, with independent and identically distributed (i.i.d) random phase $\varphi_{n,l}$ distributed over $U[0, 2\pi)$. Here, j is the square root of -1. $\tau_{n,l}$ is the time delay; $\theta_{n,l}$ and $\phi_{n,l}$ are the angle of arrive (AoA) and angle of departure (AoD), respectively [30].

In the S-V channel model, the coefficients are grouped into N cluster. For the nth cluster, the arrival time is denoted by T_n, while the arrival time relative to T_n of the lth resolvable sub-path is $\tau_{n,l}$. The corresponding time delay of the (n,l)th MPC is $(T_n + \tau_{n,l}) \times T_s$, where T_s denotes the sampling interval. The mean power of MPCs, i.e., $|\bar{h}_{n,l}^2| \triangleq E(|\alpha_{n,l}|^2)$, is reorganized to:

$$|\bar{\mathbf{h}}|^2 = \left[\underbrace{\left|\bar{h}_{1,0}^2\right|\left|\bar{h}_{1,1}^2\right|\cdots\left|\bar{h}_{1,L_1-1}^2\right|}_{n=1}\cdots\underbrace{\left|\bar{h}_{N,0}^2\right|\cdots\left|\bar{h}_{N,L_N-1}^2\right|}_{n=N}\right]^T.$$

According to real measurements, the power delay profile (PDP) of MPCs will be modeled by a double exponential function [29], [30]. Thus, each component is given by:

$$\left|\bar{h}_{n,l}^2\right| = exp(-T_n/2\Gamma)exp(-\tau_{n,l}/2\gamma), \tag{4}$$

where Γ and γ denotes two arrival rate of clusters and sub-paths, respectively.

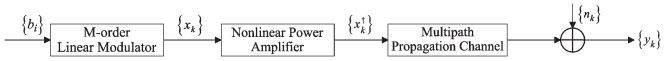


Fig. 2. Transmitter-end block diagram.

It is noteworthy that, owing to high-resolution spatial transmissions promised by beam-forming techniques [31]–[33], the line-of-sight (LOS) term in (3) is significantly larger than other non-line-of-sight (NLOS) multipath components. Following the experimental result, the energy of the 1st LOS path is about 20 dB higher than NLOS components [29], [30]. So, subpaths generated from the secondary reflecting may be ignored due to noticeable attenuations (e.g., about 10 dB for once reflection). Thus, the total number of multipaths L (i.e., $N = 1$) may become much smaller if the beam-forming technique is adopted.

C. Signal Model

Taking both the PA nonlinearity and the multipath fading into accounts, an uncoded mm-Wave communication system is considered and the block diagram of transmitter is shown in Fig. 2. The binary source sequence b_i $(i = 0, 1, 2, \cdots)$ is firstly fed to an M-order linear modulator (such as M-PSK or M-QAM modulator). Subsequently, each $m\,(m = log_2 M)$ bits are mapped to one output data symbol $x_k\,(k = 0, 1, 2, \cdots)$. Each symbol is then passed through a front-end nonlinear PA and, finally, the emitted symbols $x_k^\dagger$ are generated. The emission symbols are propagated through multipath fading channels and, consequently, the inter-symbols interference (ISI) will be inevitable in receivers.

Based on the descriptions above, a discrete-time baseband form of sampled receiving signal, denoted by $y_k(k = 0, 1, 2, \cdots)$, is given by:

$$\begin{aligned} y_k &= \sum_{l=0}^{L-1} h_{k,l}^* x_{k-l}^\dagger + n_k \\ &= \mathbf{h}_k^H \mathbf{x}_k^\dagger + n_k, \quad k = 0, 1, \cdots, K-1. \end{aligned} \tag{5}$$

where L specifies the memory order of multipath response; K is the size of a transmission burst. The baseband complex multipath response is $\mathbf{h}$, and $(.)^H$ denotes the conjugate transpose. For the convenience, the quasi-static channel is considered. That is, $\mathbf{h}_k$ remains invariant within each transmission burst but is allowed to be changed among different bursts. Thus, we may further simplify $\mathbf{h}_k$ to $\mathbf{h} = [h_0, h_1, h_2, \cdots, h_{L-1}]^T$.

It should be noted that the nonlinearly distorted symbol can be viewed as an unobserved intermediate state, which is denoted by $\mathbf{x}_k^\dagger = [x_k^\dagger, x_{k-1}^\dagger, \cdots, x_{k-L+1}^\dagger]^T$. Here, $x_k^\dagger$ is the distorted symbols of the kth discrete time index. Given the nonlinear mapping function $q(.)$, the PA outputs $x_k^\dagger$ are determined by $x_k^\dagger = q(x_k)$. Apart from contaminating the receiving constellations, this nonlinear transform will introduce a measure of memory to output symbols. As suggested by measurements [34], [35], the memory effect of nonlinear PAs may be aroused by *electrical* and *thermal* memory effects (especially for high power amplifier), which can be characterized by specific behavior models, e.g., a discrete time-domain *Volterra* model with the finite memory [35], [36]. The envelops of PA's output signal $\breve{x}_k^\dagger$, as demonstrated, may be then represented as a truncated Volterra series of previous input envelops, i.e., $\breve{x}_{k-i}(i = 0, 1, \cdots, l_m)$. Here, l_m accounts for the finite memory order of output symbols.

It is supposed that input complex symbols x_k keep independent of each other, which are drawn from a finite alphabet set $\mathcal{A} = \{a_1, a_2, \cdots, a_{|\mathcal{A}|}\}$. The *a priori* probabilities of each symbols x_k are specified by $p_\iota = P(x_k = a_\iota)(\iota = 1, 2, \cdots, |\mathcal{A}|)$. In practice, the transmitted symbols are equally likely, i.e., $p_\iota = 1/|\mathcal{A}|$. So, we have $\mathbf{x}_k \in \mathcal{A}^L$ and $p(\mathbf{x}_k) = \frac{1}{|\mathcal{A}|^L}$. It is further assumed that the additive channel noise n_k is a sequence of zero-mean i.i.d complex random variables, i.e., $n_k \sim \mathcal{N}(0, \sigma^2)$, which is independent of the symbols x_k.

After propagated from a nonlinear PA and multipath channels and corrupted by the ambient noise, the received signal $y_k(k = 0, 1, 2, \cdots)$ of mm-Wave systems will be distorted dramatically. In the receiver-end, then the nonlinearly contaminated symbols x_k would be estimated with the help of the *a priori* information and the estimated channel state information (CSI).

From eq. (5), the resulting memory of multipath channels may facilitate the signal detection of mm-Wave systems. We will resort to the DSM to thoroughly characterize the nonlinear estimation process and, furthermore, exploit the underlying dynamics of the received signal. For the considered nonlinear estimation problem, a new DSM can be established as following

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{u}_k, \tag{6}$$

$$y_k = g(\mathbf{x}_k) + n_k. \tag{7}$$

In the *state equation* eq. (6), the states $\mathbf{x}_k$ can be modeled by an unobserved (or hidden) Markov process. For the adopted independent source, we may further employ a linear transform to fully describe the relationship between $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$ [20], [23], i.e., $f(\cdot) : \mathcal{A}^L \to \mathcal{A}^L$, which is specified by $f(\mathbf{x}_{k-1}) = \mathbf{F}\mathbf{x}_{k-1}$. That is, the evolution of unknown states would be characterized by a state transitional matrix (STM) $\mathbf{F}_{L\times L}$ and an $L \times 1$ driven vector $\mathbf{u}_k$. Here, the perturbation vector is specified by $\mathbf{u}_k = [x_k, 0, \cdots, 0]^T$, while the linear transitional matrix is specified by:

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

The received signals y_k is derived from the *measurement equation* involving a nonlinear observation function $g(.) : \mathcal{A}^L \to \mathbb{R}^1$, which is specified both by the nonlinear PA model $(\mathbf{G}, \mathbf{\Psi})$ and the unknown multipath response $\mathbf{h}$. Here, $\mathbb{R}^1$ gives the real-valued space of 1-dimension. With the established DSM, the main concern of the mm-Wave nonlinear equalization is to recover the unknown symbols $\mathbf{x}_k$ $(k = 0, 1, 2, \cdots)$ blindly, relying on the nonlinearly distorted and noisy-corrupted observations y_k $(k = 0, 1, 2, \cdots, K)$.

III. Bayesian Inference and Particle Filtering

A. Bayesian Inference

By excluding the pilot sequence, blind equalizations and signal detections, which may enhance the energy and spectrum efficiency, are of promise to mm-Wave applications. For the specific problem, the multipath channel $\mathbf{h}$ and transmitted symbols $\mathbf{x}_k$ are two unknown quantities, which are assumed to be independent of each other and have the *a priori* densities $p(\mathbf{h})$ and $p(\mathbf{x}_k)$, respectively. From a Bayesian point of view, the optimal signal detection should rely on the joint *a posteriori* probability $p(\mathbf{h}, \mathbf{x}_{0:k}|y_{0:k}, (\mathbf{G}, \mathbf{\Psi}))$, which is given by:

$$p\left(\mathbf{h}, \mathbf{x}_{0:k}|y_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) = \frac{p\left(y_{0:k}|\mathbf{h}, \mathbf{x}_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) p(\mathbf{h}, \mathbf{x}_{0:k})}{p(y_{0:k})}, \tag{8}$$

$$= \frac{p\left(y_{0:k}|\mathbf{h}, \mathbf{x}_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) p(\mathbf{x}_{0:k}) p(\mathbf{h})}{p(y_{0:k})}, \tag{9}$$

where the observations trajectory till the kth time index is denoted by $y_{0:k} \triangleq [y_0, y_1, \cdots, y_k]$, and the state trajectory is $\mathbf{x}_{0:k} \triangleq [\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_k]$. Given the nonlinear observation model in (7) and the white Gaussian noise n_k, the likelihood function $p(y_k|\mathbf{h}, \mathbf{x}_k, (\mathbf{G}, \mathbf{\Psi}))$ at time index k is expressed to eq. (10), where $\mathfrak{Re}(x)$ and $\mathfrak{Im}(x)$ represents the real part and image part of the complex variable x, respectively. It is seen from eq. (10), shown at the bottom of the page, that the likelihood distribution here is a non-Gaussian function of $\mathbf{x}_k$.

For simplicity, the multipath response $\mathbf{h}$ is assumed to follow *a priori* Gaussian distribution, with the mean vector of $\bar{\mathbf{h}}$ and the covariance matrix $\mathbf{\Sigma}$, i.e.,

$$p(\mathbf{h}) = \frac{1}{(2\pi)^{L/2} \left(\det(\mathbf{\Sigma})\right)^{\frac{1}{2}}} exp\left[-\frac{1}{2}(\mathbf{h} - \bar{\mathbf{h}})^H \mathbf{\Sigma}^{-1} (\mathbf{h} - \bar{\mathbf{h}})\right]. \tag{11}$$

The posterior probability of transmitted symbols $\mathbf{x}_k$ may be derived via the marginalization of the *a posteriori* density. As a consequence, signal detections will be realized by maximizing the posterior density $p(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi}))$, which is given by:

$$p\left(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) \propto \underbrace{\int_{\mathbf{h}} p(y_{0:k}|\mathbf{x}_{0:k}, \mathbf{h}, (\mathbf{G}, \mathbf{\Psi}))\, p(\mathbf{h}) d\mathbf{h}}_{p(y_{0:k}|\mathbf{x}_{0:k}, (\mathbf{G}, \mathbf{\Psi}))} \times p(\mathbf{x}_k). \tag{12}$$

For real-time transmissions of mm-Wave systems, it is significant to estimate the posterior probability $p(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi}))$ sequentially and accomplish signal detections on reception of new observations. By alleviating the computation burden repeated on each new observation, the recursive detection will be recommended which could adaptively incorporate the innovation information into the past inference. With the Chapman-Kolmogorov equation, $p(\mathbf{x}_{0:k}|y_{0:k})$ is estimated by:

$$p\left(\mathbf{x}_k, \mathbf{h}|y_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) = p\left(\mathbf{x}_{k-1}, \mathbf{h}|y_{0:k-1}, (\mathbf{G}, \mathbf{\Psi})\right) \times \frac{p\left(y_k|\mathbf{x}_k, \mathbf{h}, (\mathbf{G}, \mathbf{\Psi})\right) p(\mathbf{x}_k|\mathbf{x}_{k-1})}{\int_{\mathbf{x}} \int_{\mathbf{h}} p\left(y_k|\mathbf{x}_k, \mathbf{h}, (\mathbf{G}, \mathbf{\Psi})\right) p(\mathbf{h}) p(\mathbf{x}_k|y_{0:k-1}) d\mathbf{h} d\mathbf{x}_k}. \tag{13}$$

The above recurrence propagation of the posterior density forms a solid theoretical framework of Bayesian statistical inference. Note that, unfortunately, except for some special situations with a linear and Gaussian model that may be addressed elegantly without any approximations [20], e.g., the well-known Kalman filtering (KF), the above sequential estimation of the *a posteriori* probability can only be considered as a conceptual solution. More specifically, attributed to the encountered nonlinear and non-Gaussian process and the resulting intractable high-dimensional marginalization, the analytic form of the posterior distribution of interest could be hardly derived.

B. Sequential Importance Sampling

As a promising alternative that avoids the marginalization and implements the recursive estimation approximately, PF provides a great promise to the joint detection of mm-Wave communications. Relying on a simulated Monte-Carlo random sampling approach, PF obtains the consistent estimation of *a posteriori* (or target) probability via a group of discrete random measures (or particles) $\mathbf{x}^{(i)}$ with different probability masses (or weights) $w^{(i)} (i = 1, 2, \cdots, I)$, where I the size of discrete particles [18]. For clarity, here $\mathbf{x}$ denotes the hidden vector to be estimated. Essentially, the particles are simulated samples drawn from an unknown space associated with the target probability $p(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi}))$. Thus, the continuous target distribution can be approximated numerically by:

$$p\left(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) \simeq \sum_{i=1}^{I} w_k^{(i)} \delta\left(\mathbf{x} - \mathbf{x}_k^{(i)}\right), \tag{14}$$

where $\delta(\mathbf{x} - \mathbf{x}_k^{(i)})$ denotes the Dirac mass at the point $\mathbf{x}_k^{(i)}$. Based on eq. (14), the expectation of associated features of $\mathbf{x}$, i.e., $\mathsf{E}\{g(\mathbf{x}_k)\} = \int_{\mathbf{x}_k} g(\mathbf{x}_k) p(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi})) d\mathbf{x}_k$, may be evaluated by $\hat{\mathsf{E}}\{g(\mathbf{x}_k)\} \simeq \sum_{i=1}^{I} g(\mathbf{x}_k^{(i)}) w_k^{(i)}$ when the particle size I is sufficiently large [20], [24].

Although PF may provide a feasible approach to estimate various complicated distributions, it has to be emphasized that, it is usually infeasible to sample directly from the target posterior distribution $p(\mathbf{x}_k|y_{0:k}, (\mathbf{G}, \mathbf{\Psi}))$. This is because, in practice, an analytical form of posterior PDFs will be unavailable. A proposal distribution or an importance function $\pi(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi}))$, therefore, is designed from which discrete particles can be simulated, i.e., $\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi}))$. Accordingly, the associated

$$p\left(y_k|\mathbf{h}, \mathbf{x}_k, (\mathbf{G}, \mathbf{\Psi})\right) = \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{1}{\sigma^2}\left(\left|\mathfrak{Re}(y_k) - \mathfrak{Re}\left(g(\mathbf{x}_k)\right)\right|^2 + \left|\mathfrak{Im}(y_k) - \mathfrak{Im}\left(g(\mathbf{x}_k)\right)\right|^2\right)\right\} \tag{10}$$

probability masses (or importance weigh) $w^{(i)}$ may be determined by:

$$w_k^{(i)} = \frac{p\left(y_{0:k}|\mathbf{x}_{0:k}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right) p\left(\mathbf{x}_{0:k}^{(i)}\right)}{\pi\left(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right)}. \tag{15}$$

As an approximation of realistic distributions, the importance weights should be normalized.

$$w_k^{*(i)} = w_k^{(i)} \Big/ \sum_{i=1}^{I} w_k^{(i)}. \tag{16}$$

Note that, usually an importance function will be factored as:

$$\pi(\mathbf{x}_k|y_{0:k}) = \pi(\mathbf{x}_k|\mathbf{x}_{0:k-1}, y_{0:k}) \times \pi(\mathbf{x}_{0:k-1}|y_{0:k-1}). \tag{17}$$

With the particles $\mathbf{x}_k^{(i)}(i = 1, 2, \cdots, I)$ sequentially sampled from $\pi(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi}))$ as each new observation arrives, the associated importance weight $w_k^{(i)}$ can be updated recursively by:

$$w_k^{(i)} \propto \frac{p\left(y_k|\mathbf{x}_k, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right) p\left(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}\right)}{\pi\left(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0,k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right)} \times w_{k-1}^{(i)}. \tag{18}$$

It should be noted that the importance distribution, which should be carefully designed in accordance with different realistic situations, will have a significant impact on the estimation performance of PF [25], [26]. In practice, two popular importance functions can be recommended, i.e., the prior and the optimal importance function. The prior importance function is chosen to $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}\right)$, which is easy to implement but rarely exploits the innovation information of observations. The optimal importance function, which would minimize the one-step variance $var(w^{(i)})$, is given by $p\left(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, y_k\right)$. In our analysis, the optimal importance function will be adopted and, accordingly, the importance weight $w_k^{(i)}$ will be propagated via:

$$w_k^{(i)} = p\left(y_k|\mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right) \times w_{k-1}^{(i)}. \tag{19}$$

To sum up, two steps are involved in the implementation of PF: (1) draw the random particles by sampling from the importance distribution, i.e., $\mathbf{x}_k^{(i)} \sim \pi\left(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right)$; and (2) update the associated weight $w_k^{(i)}$ by using (19). To deal with the degeneration of particle weights, a *resample* process is usually necessary to eliminate the negligible particles and further improve the estimation performance, one can refer to [18], [19], [23] for details.

C. Application Considerations

With the linear multipath and Gaussian noise (i.e., in absence of nonlinear effects), PF will be implemented conveniently. For the ideal linear and Gaussian model, the posterior distribution of $\mathbf{h}$, after observing y_k, is given by:

$$p(\mathbf{h}|\mathbf{x}_{0:k}, y_{0:k}) = p(\mathbf{h}, \mathbf{x}_{0:k}|y_{0:k})/p(\mathbf{x}_{0:k}|y_{0:k}) \tag{20}$$

$$\propto p(y_{0:k}|\mathbf{h}, \mathbf{x}_{0:k})p(\mathbf{h}). \tag{21}$$

Note that, for the posterior distribution of $\mathbf{h}$, here the term $p(\mathbf{x}_{0:k}|y_{0:k})$ may be treated as a constant which is independent of $\mathbf{h}$. Thus, $p(\mathbf{h}|\mathbf{x}_{0:k}, y_k)$ also follows a Gaussian posterior distribution and, accordingly, its mean and covariance may be updated by (22) and (23).

$$\mathbf{\Sigma}_k^{-1} = \mathbf{x}_k\mathbf{x}_k^H/\sigma^2 + \mathbf{\Sigma}_{k-1}^{-1}, \tag{22}$$

$$\bar{\mathbf{h}}_k = \mathbf{\Sigma}_k\left(\mathbf{x}_k y_k/\sigma^2 + \mathbf{\Sigma}_{k-1}^{-1}\bar{\mathbf{h}}_{k-1}\right). \tag{23}$$

The closed analytical form of the importance distribution, which is also related to the likelihood function $p\left(y_k|\mathbf{x}_k^{(i)}\right)$, may be derived [20], [21], [23]. Thus, the random sampling may be realized effectively and discrete measures $\left\{\mathbf{x}_k^{(i)}, w_k^{(i)}\right\}$ can be drawn. As a result, the posterior distribution $p\left(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}\right)$ will be calculated approximately via (24).

$$p\left(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right) \simeq \frac{\sum_{i\in\mathcal{X}_j} w_k^{(i)}}{\sum_{i=1}^{I} w_k^{(i)}},$$

$$\mathcal{X}_j = \left\{i|\mathbf{x}_k^{(i)} \to \mathcal{A}^L(j)\right\}. \tag{24}$$

Moving on, the transmitted symbols $\mathbf{x}_k$ will be estimated by maximizing the approximated *a posteriori* distribution $p\left(\mathbf{x}_k|y_{0:k}, \mathbf{x}_{0:k-1}^{(i)}, (\mathbf{G}, \mathbf{\Psi})\right)$. If the particle size I is sufficiently large, then the estimation obtained from PF will be considered as the maximum *a posteriori* (MAP) estimation.

Based on the following two considerations, traditional PF-based linear equalization schemes, unfortunately, may be invalid in mm-Wave systems with non-ideal PAs. For one thing, the Bayesian estimation of the multipath response $\mathbf{h}$, from (20)–(23), relies primarily on a *linear* map between $\mathbf{x}_k$ and y_k, which is derived via a linear minimum mean square error (LMMSE) criterion. The nonlinear observation function (i.e., $g(\mathbf{x})$), however, has disorganized the linear relationship. Accordingly, the recursive estimation (i.e., $\hat{\mathbf{h}}$), premised on the traditional MAP scheme, may tend to be erroneous. For another, attributed to the involved nonlinear coupling relationship, a direct sequential importance sampling (SIS) process involved by recursive estimations will be infeasible.

IV. Nonlinear Equalization and Signal Detection

In this section, the realistic challenge of Bayesian estimation aroused by the nonlinear effects will be addressed via suggesting a nonlinear equalization and signal detection scheme, which is embedded with a local-linearization process. As in [19], with the linearly approximated observations, the difficulty of the SIS and marginalization process will be avoided by the new paradigm. As a consequence, the posterior density of interest, even given the nonlinear coupling between unknown symbols and observations, can be derived numerically. With the generalized PF detection framework, the nonlinearly contaminated symbols will be recursively estimated by receivers, accompanying the unknown multipath CIR.

A. Local Linearization of Observations

Nonlinearity has long been remained as one major difficulty in Bayesian inference, which will make the analytic PDF of

interests unavailable. To deal with the encountered PA distortion, from a suboptimal filtering point of view [19], [37], a local linearization model is introduced before the SIS procedure.

For convenience, the nonlinear function in (7) is reformatted as:

$$g(\mathbf{x}_k) = \mathbf{h}^T \left\{ G(\mathbf{x}_k) \circ exp\left[-j\boldsymbol{\theta}_k - j\psi(\mathbf{x}_k) \right] \right\}, \tag{25}$$

Here, the notation $\circ$ represents the Hadamard product between two vectors with the same length L; $\boldsymbol{\theta}_k$ denotes the phase vector of the complex state $\mathbf{x}_k$. Based on eqs. (1), (2), and (5), it is visually shown that the composite function $g(\mathbf{x}_k)$ remains continuous and derivable in the domain of definition. Thus, the high-order derivatives $\frac{\partial^L \mathbf{G}(\mathbf{x})}{\partial^L \mathbf{x}}$ and $\frac{\partial^L \boldsymbol{\Psi}(\mathbf{x})}{\partial^L \mathbf{x}}$ exist for $L \geq 1$. Therefore, it may be represented by the L-term Taylor series expansion (TSE).

By utilizing the first-order TSE at the point $\mathbf{x}_k = f(\mathbf{x}_{k-1})$, then a linear approximation of $g(\mathbf{x}_k)$ will be obtained, i.e.,

$$y_k = g(f(\mathbf{x}_{k-1})) + \left. \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = f(\mathbf{x}_{k-1})} [\mathbf{x}_k - f(\mathbf{x}_{k-1})] + n_k. \tag{26}$$

Two remarks should be made concerning the local-linearization model in (26). First, as a linear approximation to the original nonlinear function, the TSE-based simplified model may lead to the loss of information to some extent. Thus, only a suboptimal estimation performance will be obtained. Fortunately, from Fig. 1, the *piece-wise* linear property of AM-AM and AM-PM curves indicates that such a linearly approximated model will provide the competitive detection performance, which will be also demonstrated by subsequent experimental results. Second, the expansion point should be carefully configured. As a compromise between the implementation complexity and the estimation performance, in practice the expanding point $f(\mathbf{x}_{k-1})$ is suggested.

B. Sequential Nonlinear Equalization

With the local linearization model, the proposed nonlinear equalization scheme, which is more flexible to address the mm-Wave nonlinear signal detection problem, will involve three recursive steps, i.e., (1) draw discrete particles, (2) update of multipath statistics, and (3) estimate the partial derivative.

1) Draw Discrete Particles: Relying on the locally-linearized model at the expanding point $f(\mathbf{x}_{k-1}) \triangleq \mathbf{T}\mathbf{x}_{k-1}$ and taking the *a priori* Gaussian distribution of n_k, the analytical importance distribution in the presence of both realistic 60 GHz PA nonlinear effect and multipath propagations is derived analytically.

Proposition: For the formulated nonlinear estimation of 60 GHz mm-Wave systems, the optimal importance function $\pi\left(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, y_{0:k}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right)$ follows a Gaussian distribution (see Appendix A), i.e.,

$$\pi\left(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, y_{0:k}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right) \triangleq p\left(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, y_{0:k}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right) \sim \mathcal{N}(m_*, \sigma_*), \tag{27}$$

with the variance σ_* determined by:

$$\sigma_*^{-1} = \frac{1}{\sigma} \times \left[\left. \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = f(\mathbf{x}_{k-1})} \right]^H \left. \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = f(\mathbf{x}_{k-1})}, \tag{28}$$

and the mean vector $\mathbf{m}_k$ calculated from:

$$m_* = \frac{\sigma_*}{\sigma} \times \left[\left. \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = f(\mathbf{x}_{k-1})} \right]^H \times \left[y_k - g\left(f(\mathbf{x}_{k-1})\right) + \left. \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = f(\mathbf{x}_{k-1})} \right]. \tag{29}$$

The basic idea of approximating the optimal proposal distribution is that, with the local linearization model in eq. (26), the AM-AM and AM-PM model are treated as two linear curves locally. Despite the simplicity, part of underlying properties aroused by nonlinear PAs may be lost as mentioned and, therefore, the estimation performance is compromised. Note that, the particle vector at time index k, i.e., $\left\{\mathbf{x}_k^{(i)}\right\}$, will be decomposed into the latest particle component $\left\{x_k^{(i)}\right\}$ and the former estimated symbols $\{\hat{x}_{k-1}, \cdots, \hat{x}_{k-L+1}\}$. For simplicity, only the latest particle needs to be derived from the importance function $p\left(\mathbf{x}_k | \mathbf{x}_{0:k-1}^{(i)}, y_{0:k}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right)$, which is then appended to $\hat{x}_{k-L+1:k-1}$.

When updating the weights of simulated particles, it is relatively difficult to obtain $p\left(y_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right) = p(y_k | \mathbf{x}_k, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})) \times p\left(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}\right)$, owing to the involved nonlinear observation. For the considered DSM, we may further have $p\left(y_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right) \propto p\left(y_k | \mathbf{x}_k^{(i)}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right)$, as $p\left(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}\right) \propto p(\mathbf{x}_k)$ (constant) is independent of k. To have an effective recurrence update of particle weights, we may propagate the derived particles $\mathbf{x}_k^{(i)}$ through a nonlinear mapping modular and obtain a group of auxiliary states $\mathbf{x}_k^{\dagger(i)}$. It is seen that the likelihood function, premised on new auxiliary particles $\mathbf{x}_k^{\dagger(i)} (i = 0, 1, \cdots, I)$, now follows a Gaussian distribution, i.e., $p\left(y_k | \mathbf{x}_k^{(i)}, \mathbf{h}, (\mathbf{G}, \boldsymbol{\Psi})\right) \sim \mathcal{N}\left(m_k^{(i)}, \sigma_k^{(i)}\right)$, with its mean and variance calculated from:

$$m_k^{(i)} = \mathbf{x}_k^{\dagger(i)T} \hat{\mathbf{h}}_{k-1}, \tag{30}$$

and

$$\sigma_k^{(i)^2} = \mathbf{x}_k^{\dagger(i)T} \hat{\boldsymbol{\Sigma}}_{k-1} \mathbf{x}_k^{\dagger(i)} + \sigma^2. \tag{31}$$

Here, $\hat{\mathbf{h}}_k$ and $\hat{\boldsymbol{\Sigma}}_k$ denote the channel statistics to be estimated, i.e., the mean vector and covariance matrix. Note that, for the assumed quasi-static channel with the invariant mean and variance, it is practically feasible to employ the one-step *delay* of the latest estimation (i.e., $\hat{\mathbf{h}}_{k-1}$ and $\hat{\boldsymbol{\Sigma}}_{k-1}$). So, in eqs. (30) and (31) the previously estimated statistics will be used to evaluate $m_k^{(i)}$ and $\sigma_k^{(i)}$ at time index k.

Given the likelihood distribution $p\left(y_k|\mathbf{x}_{k-1}^{(i)},\mathbf{h},(\mathbf{G},\mathbf{\Psi})\right)$, the associated weights of new particles $\left\{\mathbf{x}_k^{(i)}\right\}$ is updated via:

$$w_k^{(i)} \propto \frac{1}{\sqrt{2\pi}\sigma_k^{(i)}} exp\left[-\left(y_k - m_k^{(i)}\right)^2 \Big/ 2\sigma_k^{(i)2}\right] \times w_{k-1}^{(i)}. \quad (32)$$

It is observed from eqs. (28) and (29) that the analytic form of the first-order derivative (i.e., $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$) of the nonlinear function is further required during each SIS iteration. For the considered blind scenario, however the analytic expression of nonlinear function $g(\mathbf{x})$, which is associated with unknown multipath CIR, is unavailable. Therefore, to make the nonlinear signal detector revolve, the first-order derivative needs to be estimated in practice. To accomplish this, we firstly investigate the recursive estimation of channel statistics.

2) Update of Multipath Statistics: If we treat the nonlinear PA output $\mathbf{x}_k^\dagger$ to be one auxiliary state (i.e., the intermediate input of multipath channel), we may then obtain a quasi-linear map, i.e., $\mathbf{y}_k = \mathbf{X}_k^H\mathbf{h} + \mathbf{n}_k$, where $\mathbf{X}_k \triangleq \left(\mathbf{x}_0^\dagger, \mathbf{x}_1^\dagger, \cdots, \mathbf{x}_k^\dagger\right)^T$ and $\mathbf{y}_k \triangleq [y_0\ y_1\ \cdots\ y_k]^T$. Based on the kth observation y_k and the estimated symbols $\hat{\mathbf{x}}_k = arg\max_{\mathbf{x}} \sum_{i=1}^I w_k^{(i)}\delta\left(\mathbf{x} - \mathbf{x}_k^{(i)}\right)$, then the posterior probability of $\mathbf{h}$ follows a Gaussian distribution (see Appendix B), i.e.,

$$p(\mathbf{h}|\mathbf{X}_k, y_{0:k}) \sim \mathcal{N}\left(\hat{\mathbf{h}}_k, \hat{\mathbf{\Sigma}}_k\right), \quad (33)$$

where the channel statistics $(\hat{\mathbf{h}}_k, \hat{\mathbf{\Sigma}}_k)$ will be refined via:

$$\hat{\mathbf{\Sigma}}_k = \hat{\mathbf{\Sigma}}_{k-1} - \frac{\hat{\mathbf{\Sigma}}_{k-1}\hat{\mathbf{x}}_k^\dagger\hat{\mathbf{x}}_k^{\dagger H}\hat{\mathbf{\Sigma}}_{k-1}}{\sigma^2 + \hat{\mathbf{x}}_k^{\dagger H}\hat{\mathbf{\Sigma}}_{k-1}\hat{\mathbf{x}}_k^\dagger}, \quad (34)$$

$$\hat{\mathbf{h}}_k = \hat{\mathbf{h}}_{k-1} + \frac{y_k - \hat{\mathbf{x}}_k^{\dagger H}\hat{\mathbf{h}}_{k-1}}{\sigma^2 + \hat{\mathbf{x}}_k^{\dagger H}\hat{\mathbf{\Sigma}}_{k-1}\hat{\mathbf{x}}_k^\dagger}\hat{\mathbf{\Sigma}}_k\hat{\mathbf{x}}_k^\dagger. \quad (35)$$

3) Estimate the Derivative: From eq. (25), although we have coordinated realistic effects from both the linear multipath blurring and the nonlinear distortion together into the DSM, the channel response $\mathbf{h}$ may still be separated from $g(\mathbf{x})$. That is, the nonlinear transform function $g(\mathbf{x})$ could be resolved to a linear part $\mathbf{h}$ and the nonlinear part $q(\mathbf{x}_k) = G(\mathbf{x}_k) \circ exp[-j\boldsymbol{\theta}_k - j\psi(\mathbf{x}_k)]$. Thus, the first-order derivative of the nonlinear function $g(\mathbf{x})$ will be calculated by:

$$\begin{aligned}\left.\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} &= \hat{\mathbf{h}}^H \left.\frac{\partial q(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \\ = \hat{\mathbf{h}}^H \Bigg\{ &\left.\frac{\partial \boldsymbol{G}(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \circ exp\left[-j\boldsymbol{\theta}_k - j\mathbf{\Psi}(\mathbf{x}_k)\right] \\ &-j\left.\frac{\partial \mathbf{\Psi}(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \boldsymbol{G}(\mathbf{x}_k) \circ exp\left[-j\boldsymbol{\theta}_k - j\mathbf{\Psi}(\mathbf{x}_k)\right] \Bigg\},\end{aligned} \quad (36)$$

where $\frac{\partial \boldsymbol{G}(\mathbf{x}_k)}{\partial \mathbf{x}_k}|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})}$ and $\frac{\partial \mathbf{\Psi}(\mathbf{x}_k)}{\partial \mathbf{x}_k}|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})}$ represent two $L \times 1$ partial derivative vectors, respectively. The MMSE estimator of the unknown frequency-selective multipath response, i.e., $\hat{\mathbf{h}} = \mathsf{E}\{\mathbf{h}|\mathbf{y}\}$, is obtained from:

$$\hat{\mathbf{h}} = \mathbf{h}_0 + \mathbf{\Sigma}\mathbf{X}_k^H\left(\mathbf{X}_k^H\mathbf{\Sigma}\mathbf{X}_k^H + \mathbf{\Sigma}_\sigma\right)^{-1}\left(\mathbf{y}_k - \mathbf{X}_k^H\mathbf{h}_0\right),$$

where the diagonal matrix $\mathbf{\Sigma}_\sigma = diag(\sigma^2)$ is the covariance matrix of noise vector $\mathbf{n}_k$. The MSE of the above channel estimation is given by:

$$\begin{aligned}MSE &= tr\left\{\mathsf{E}\left\{\left[\mathbf{h} - E\left(\hat{\mathbf{h}}|\mathbf{y}\right)\right]^H\right\}\right\} \\ &= tr\left[\hat{\mathbf{\Sigma}}_{k-1} - \hat{\mathbf{\Sigma}}_{k-1}\mathbf{X}_k^H\left(\mathbf{X}_k\hat{\mathbf{\Sigma}}_{k-1}\mathbf{X}_k^H + \mathbf{\Sigma}_\sigma\right)^{-1}\mathbf{X}_k\hat{\mathbf{\Sigma}}_{k-1}\right]\end{aligned} \quad (37)$$

With the help of Duncan-Guttman inverse formula, we may further have

$$\begin{aligned}MSE &= \mathrm{tr}\left[\left(\mathbf{X}_k\mathbf{\Sigma}_\sigma^{-1}\mathbf{X}_k^H + \hat{\mathbf{\Sigma}}_{k-1}^{-1}\right)^{-1}\right] \\ &\simeq \frac{\sigma^2}{kE(\mathbf{x}^{\dagger 2}) + \sigma^2\sum_{l=1}^L 1/\delta_l^2} \overset{\delta_l=\delta}{=} \frac{\sigma^2\delta^2}{kE(\mathbf{x}^{\dagger 2})\delta^2 + L\sigma^2}.\end{aligned} \quad (38)$$

where δ_l is the lth diagonal element of the covariance matrix $\mathbf{\Sigma}$. From eq. (38), it is seen that the MSE will decrease as the increasing of $E(\mathbf{x}^{\dagger 2})/\sigma^2$. It is also supposed that, with the recurrence estimation process, the updated statistics of multipath response may gradually become more accurate.

Inspecting the estimated channel $\hat{\mathbf{h}}$ and eq. (35), we easily find $\hat{\mathbf{h}} = \hat{\mathbf{h}}_k$. By using the one-step delay of $\hat{\mathbf{h}}_k$, the channel estimation can be then simplified to $\hat{\mathbf{h}} \simeq \hat{\mathbf{h}}_{k-1}$. With the estimated channel and the approximated 1st-order derivative, new particles $\mathbf{x}_k^{(i)}$ will be generated from the importance distribution $p\left(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, y_{0:k}, \mathbf{h}, (\mathbf{G},\mathbf{\Psi})\right)$. Thus, the sequentially nonlinear equalization is realized relying on the new extended PF-based paradigm.

Algorithm 1 Nonlinear Equalization and Signal Detection

Input: Observation $y_k, k = 0, 1, \cdots, K-1$
Nonlinear PA Model $(\mathbf{G}, \mathbf{\Psi})$
Output: Estimated symbols $\{x_k\}$.
Initialize the particles $\left\{\mathbf{x}_0^{(i)}, w_0^{(i)}\right\}$ and the channel statistics (i.e., the mean $\hat{\mathbf{h}}_0$ and variance $\hat{\mathbf{\Sigma}}_0$).
for $k \to 0$ to $K-1$ **do**
  for $i \to 1$ to I **do**
    Based on eq. (36), calculate the first-order derivation $\partial g(\mathbf{x}_k)/\partial \mathbf{x}_k$ at the point $\mathbf{x}_k = f(\hat{\mathbf{x}}_{k-1})$.
    Derive the importance density by eqs. (27)–(29).
    Draw $x_k^{(i)}$ from the updated importance density.
    Let the particles pass a nonlinear mapping module and obtain the auxiliary particles $x_k^{(i)\dagger}$.
    Derive the likelihood function based on eqs. (30), (31).
    Update the particle weights by eq. (32).
  end for
  Normalize the particle weights $w_k^{*(i)} = w_k^{(i)} / \sum_i w_k^{(i)}$.
  Let the effective size $N_{eff} = 1/\sum_i \left(w_k^{(i)}\right)^2$, and choose a random term $\varepsilon \in (0, 1)$.

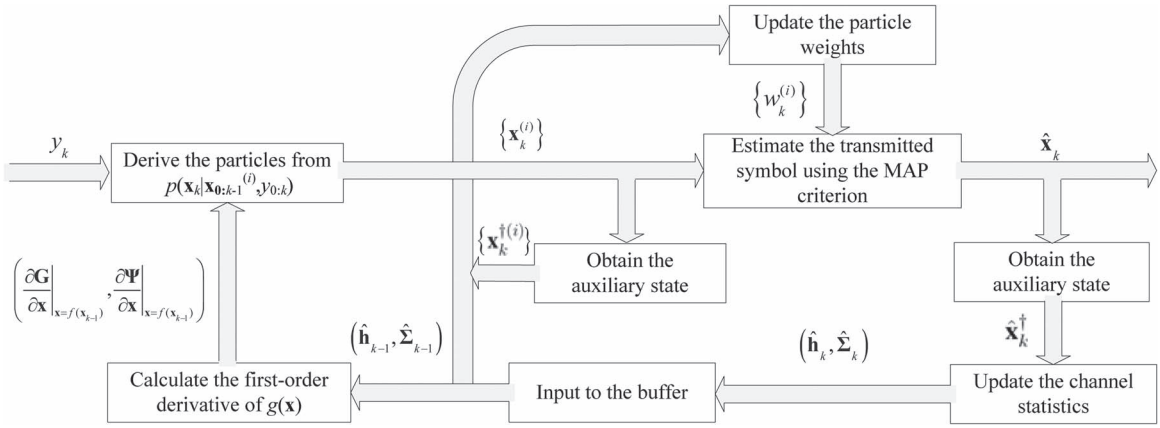


Fig. 3. Schematic implementation of the proposed PF-based nonlinear detection.

if $N_{eff} \leq \varepsilon$ and $k < K$ **then**
  for $i \rightarrow 1$ to I **do**
    $\mathbf{x}_{0:k}^{(i)} = \mathbf{x}_{0:k}^{(i)}$ with the probability of $w_k^{(i)}$.
  end for
end if
Perform MAP estimation by following eq. (24).
Update the channel statistics $(\hat{\mathbf{h}}_k, \hat{\mathbf{\Sigma}}_k)$ by following eqs. (34), (35).
end for

C. Implementations

The schematic implementation of the presented recursive estimation algorithm is illustrated by Fig. 3. Based on the one-step delay of the estimated channel statistics, the derivative of nonlinear observation function is evaluated. Subsequently, new particles are generated relying on the new observations, which is further fed into a baseband nonlinear transform modular (maybe DSP-based). The resulting auxiliary states are then employed to update the likelihood function and the associated weights. Finally, the estimation of transmitted symbols will be derived premised on the approximated posterior density and an MAP criterion. Meantime, the multipath channel statistics $(\hat{\mathbf{h}}_k, \hat{\mathbf{\Sigma}}_k)$ will be refined sequentially by utilizing the estimated auxiliary state (i.e., $\hat{\mathbf{x}}_k^{\dagger}$). The flow of the suggested nonlinear equalization and signal detection scheme is given by **Algorithm 1**.

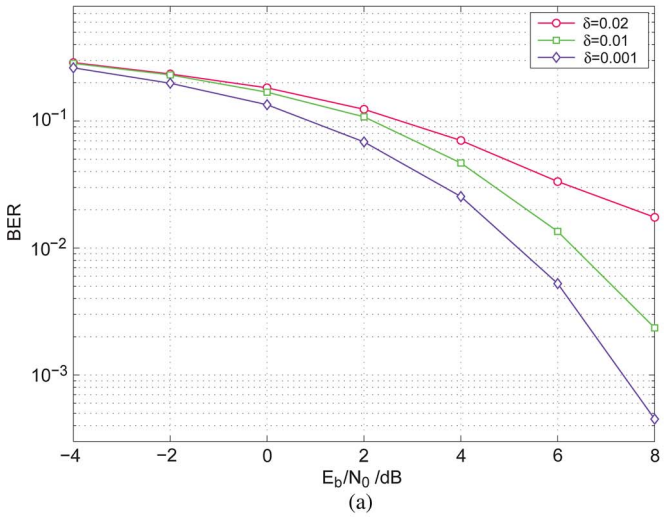
As another approach to construct the importance density, we may alternatively adopt an M-length discrete probability mass function (PMF), which shares a much similar conception with [21]. The basic idea is that the new particles can take values only from an M-elements discrete space $\mathcal{A}$ (e.g., for M-QAM/QPSK), thus the optimal importance density $p(x_k|x_{0:k-1}; y_{0:k}) \propto p(y_k|x_k; x_{0:k-1})$ may be obtained by constructing an M-length PMF after translating total M states through a nonlinear PA [21]. Then, the SIS process will be performed. This PMF-based scheme, i.e., directly deriving its importance density by feeding possible symbols into the nonlinear PA, is theoretically optimal for unknown *independent* symbols x_k. It is noteworthy that, as mentioned, nevertheless the intermediate or received symbols (i.e., $x_k^{\dagger}$ or y_k) would present some level of *memory* after propagated from a nonlinear PA [34]–[36]. The linearization approach, which fully takes the memory among output symbols, may be more compatible to the considered nonlinear detection scenario. Recall that the recursive updating of its importance density utilizes the information of *past* particles and observations. The PMF-based scheme, however, fails to construct an importance density by incorporating the historical information.

D. Complexity Analysis and Algorithm Simplification

As the suggested scheme essentially exploits simulation-based techniques to solve signal processing problems, the involved computational complexity is relatively high. For example, during the SIS process of eq. (27), the required number of multiplications may even approach $\mathcal{O}[q_1 L^2 + \kappa_1(n_b) + I\kappa_2(n_b)]$. The first term, $\mathcal{O}(q_1 L^2)$, mainly accounts for the computation of channel statistics. The second term, $\mathcal{O}[\kappa_1(n_b)]$, represents the additional complexity of evaluating the mean and variance of the proposal density, where $\kappa_1(n_b) \propto q_2 n_b + q_3[log(n_b)]^2$ is the complexity of evaluating related mathematical functions and n_b is the number of digital-bits of precision. The third term denotes the complexity of calculating the proposal density, where $\kappa_2(n_b) \propto [log(n_b)]^2$ is mainly the complexity of evaluating Gaussian functions [38]. q_1, q_2 and q_3 are three constants.

V. Experimental Simulations and Performance Evaluations

In this section, we will evaluate the performance of the proposed joint estimation algorithm. In the simulation, the nonlinear PA model regulated by the IEEE 802.11ad TG is used. Without loss of generality, a simplified channel of $L = 3$ is adopted, i.e., the strong LOS component is assumed in mm-Wave communications given the widely adopted beamforming techniques [31], [33]. Correspondingly, the channel mean is configured to $|\bar{\mathbf{h}}| = [1\ 0.1\ 0.001]^T$ and its covariance matrix is $\mathbf{\Sigma} = diag\{\delta^2\ \delta^2\ \delta^2\}$. The size of particles is set to $I = 20$. Each BER curve is obtained from the average on 20 independent realizations of random channel responses and, in each realization, total 20 000 bits are transmitted.



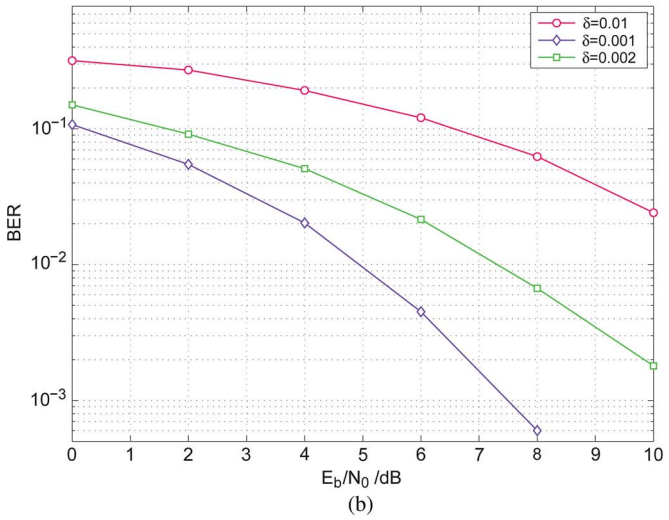
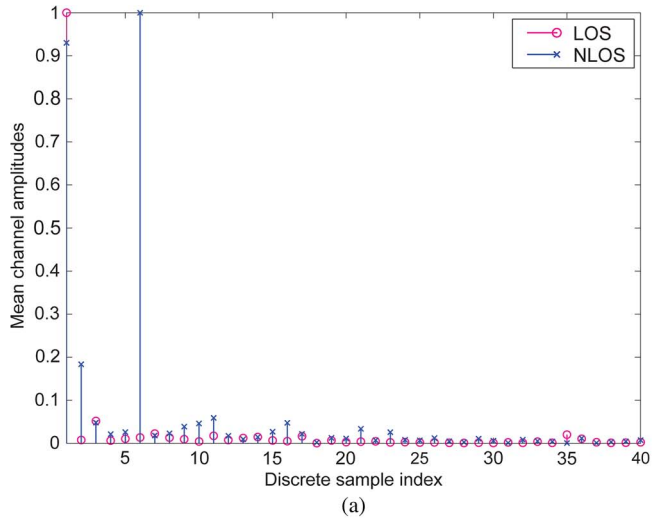


Fig. 4. Detection performance with the different channel covariance δ. (a) QPSK modulation. (b) 16QAM modulation. Note that the OBO value is set to 6 dB.



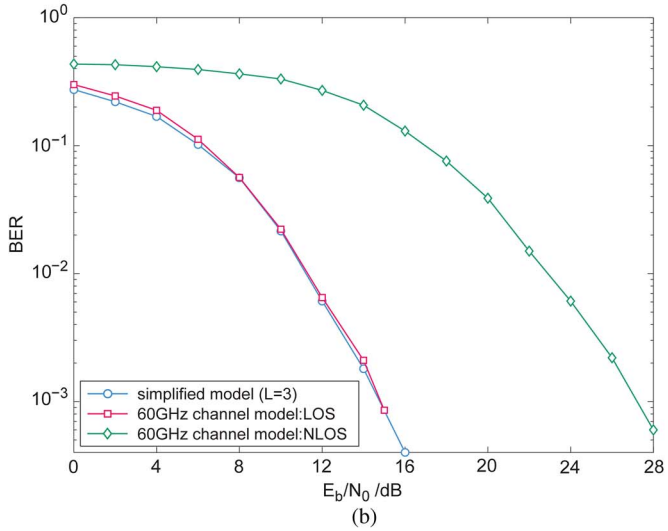


Fig. 5. (a) Simulated response of IEEE 802.15.3c channel model. (b) Detection Performance of 16QAM modulation signals. Notice that the OBO value is set to 6 dB.

A. *Different Multipath Channels*

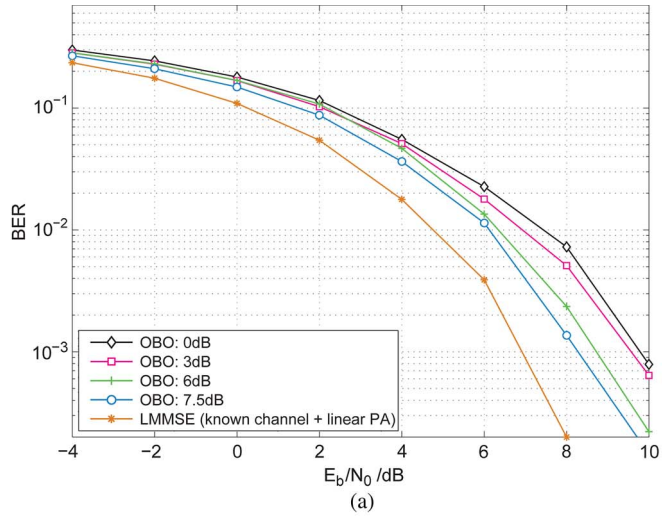
With different channel variance δ^2, the numerically derived BER curves of QPSK signals are illustrated in Fig. 4(a). In this experiment, an OBO value of 6 dB is adopted. As shown by simulation results, the covariance matrix (or the variance δ^2) may have a significant impact on the performance of joint detections. The blind detection performance will decline with the increase of channel uncertainty (i.e., δ^2). When the standard deviation δ is 0.001, the desired E_b/N_0 is about 5 dB when BER drops below 10^{-2}. In comparison, when the value of δ increases to 0.01, the target E_b/N_0 will be remarkably raised to 6.25 dB to acquire the comparable BER.

Compared with QPSK signals, the 16QAM modulation seems to be more sensitive to multipath interference and residual errors. In the experimental results shown by Fig. 4(b), the prescribed OBO value is also 6 dB. It is observed that, when δ increases from 0.002 to 0.01, the archived BER will remarkably increase from 6.0×10^{-4} to 6.2×10^{-2} if E_b/N_0 is configured to 8 dB.

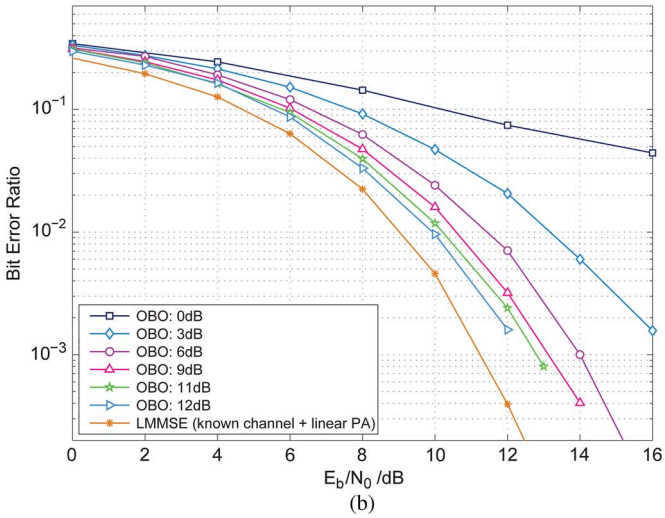
To evaluate the joint detection performance in realistic mm-Wave communications, the standard channel model regulated by the IEEE 802.15.3c task group (TG) is further considered. The CM3 is used for the LOS case, and the mean number of clusters is set to $N = 6$. For the NLOS case where the LOS path is obscured, the CM8 channel is adopted. The mean channel amplitudes generated by LOS and NLOS models have been demonstrated by Fig. 5(a). We may note that the first path of the LOS scenario, which is significantly stronger than other MPCs (with $\Delta K = 24$ dB), will become completely dominant. In this case, the detection performance is comparative to that of the simplified channel (e.g., $N = 1$, $L = 3$), as illustrated by Fig. 5(b). Nevertheless, for realistic NLOS scenarios, the detection performance will be degraded tremendously.

B. *Different Operational Conditions*

The influences on detection performance from nonlinear PA are also investigated. The channel mean is configured by $|\bar{\mathbf{h}}| = [1\ 0.1\ 0.001]^T$ and δ is 0.01. From the BER curve of QPSK signals in Fig. 6(a), it is observed that, even if the operational power is configured to the saturation point (or the OBO is 0 dB), the distorted signals can be still recovered if a sufficiently large E_b/N_0 is provided. Meanwhile, the adoption of OBO may bring some benefits. To be specific, when the OBO is increased



(a)



(b)

Fig. 6. Detection performance with the different OBO values. (a) QPSK signals. (b) 16QAM signals. The channel mean amplitude is $|\bar{\mathbf{h}}| = [1\ 0.1\ 0.001]^T$.

from 3 dB to 6 dB, the desired E_b/N_0 will decrease slightly from 9.5 dB to 8.6 dB when BER drops below 10^{-3}. If the OBO is 7.5 dB, the gap between the proposed scheme and the LMMSE equalization (with known channel response and linear PA) is about 1.7 dB.

The detection performance of 16QAM, with different degree of PA distortions, is illustrated by Fig. 6(b). When the operation voltage is 0.58 V (or the OBO value is 0 dB), the BER will be gradually reduced, indicating the distortion both from nonlinear PA and linear multipath ISI can be eliminated with the increasing of E_b/N_0. As we will see later, the BER floor may not be avoided by traditional schemes, e.g., joint MAP scheme assuming a linear PA. Thus, the nonlinear equalization scheme may be efficiently applied to mm-Wave systems, yet at the cost of the high power consumption and the greatly reduced energy efficiency. In practice, furthermore, the OBO can be combined to enhance the efficiency of the designed algorithm. From Fig. 6(b), the adoption of OBO may noticeably reduce the desired E_b/N_0 and thereby promote the energy efficiency. It is noteworthy that, when the OBO value surpasses 6 dB, the benefit will become limited. As shown from numerical results, the detection performance will be converged finally, which is attributed to residual errors of multipath ISI.

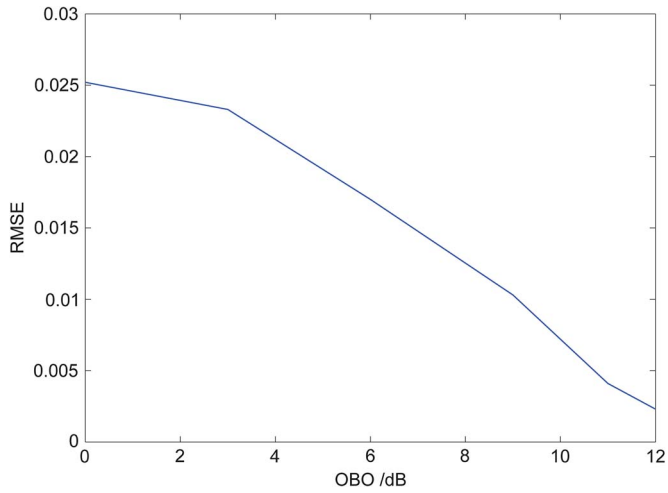


Fig. 7. RMSE performance of 16QAM signals under different OBO values. The E_b/N_0 is set to 10 dB.

The estimation performance of multipath channels with different OBOs is also investigated. The root-mean-square error (RMSE) of channel estimations is numerically derived from experimental simulations. During the experiment, the mean channel amplitude is configured to $[1\ 0.1\ 0.001]^T$ and the deviation is $\delta = 0.01$. It is seen from Fig. 7 that the RMSE of channel estimations will be decreased with the increase of OBO. In practice, with the larger OBO, signal constellations would experience less nonlinear distortions and, therefore, the detection performance of unknown symbols, accompanying the estimation accuracy of multipath channels, may be promoted.

C. Performance of Inaccurate PA Model

It is noted that the designed joint estimation algorithm will depend on the parametric PA model, i.e., the AM-AM and AM-PM model, as the proposal density may involve the first-order derivative of nonlinear observation functions. In practice, the PF-based method may fail to derive joint estimations of parametric PA models and multipath responses, since this process involves intractable high-dimensional marginalization. Thus, the designed algorithm may be not adaptable to realistic PA variations.

To make the designed scheme more applicable to realistic scenarios, the proposed algorithm in the presence of inaccurate PA models is evaluated. In numerical experiments, the parametric PA model will vary randomly from different operating environments (e.g., chip heating). However, the mean of PA parameters, i.e., $\boldsymbol{\nu} \triangleq (\bar{\boldsymbol{G}}, \bar{\boldsymbol{\Psi}})$, is known as *a priori* by receivers, which may be justified by the fact a PA model is formulated practically based on the common (or average) chip heating. For simplicity, PA's parameters are assumed to follow the i.i.d. Gaussian distribution, i.e., $(\boldsymbol{G}, \boldsymbol{\Psi}) \sim \mathcal{N}\{\boldsymbol{\nu}, \boldsymbol{\Sigma}_p\}$. The nth element of the covariance matrix, i.e., $\sigma_p(n) = \boldsymbol{\Sigma}_p(n,n)$, accounts for the random variance of the nth PA parameter, which is assumed to remain proportional to $\boldsymbol{\nu}(n) \times \varrho$, where ϱ is referred to as the *relative mismatch ratio*. Based on experimental simulations, the detection performance of different relative mismatch ratios have been shown in Fig. 8. In the analysis, the 16QAM signal is considered. The mean channel

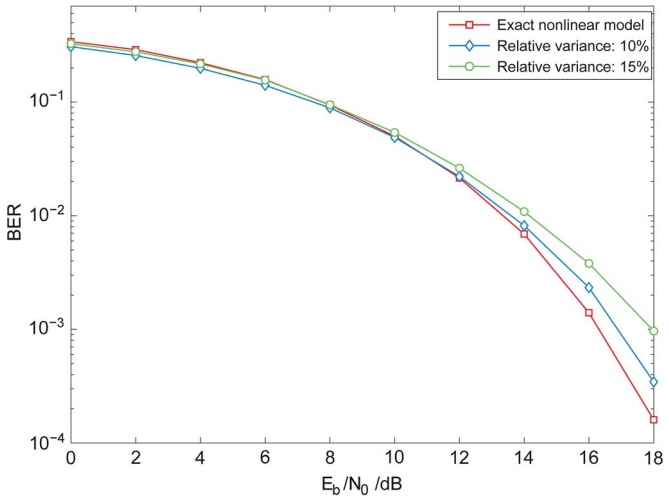


Fig. 8. Detection performance of realistic PA of inaccurate parametric model.

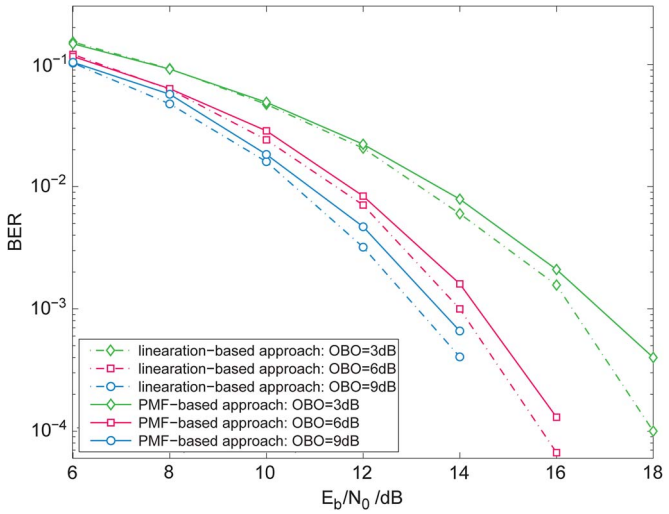


Fig. 9. Detection performance of the local-linearization based scheme and the PMF-based scheme.

amplitude is configured to $[1\ 0.1\ 0.001]^T$. It is seen that the mismatch between the assumed parametric PA model and a realistic PA model will degrade the detection performance. The greater the relative mismatch ratio, the worse the BER performance. For example, if a relative mismatch ratio of 15% is considered, the BER performance may be deteriorated even by an order of magnitude. Although we may draw a conclusion that, for a relatively small mismatch (e.g., the relative mismatch ratio less than 10%), the designed scheme could be applied to realistic scenarios, it has to be emphasized that the stability of parametric PA models is of importance to the proposed joint estimation algorithm.

D. Linearization Scheme vs PMF Method

In this experiment, both the linearization-based scheme and the PMF-based method are studied in the context of mm-Wave nonlinear equalization and signal detection. It is seen from numerical results in Fig. 9 that, with small OBOs and high SNRs, the linearization-based scheme may surpass the PMF-based method. For example, a detection gain of 1 dB can be achieved by the linearization method when the OBO is 3 dB and

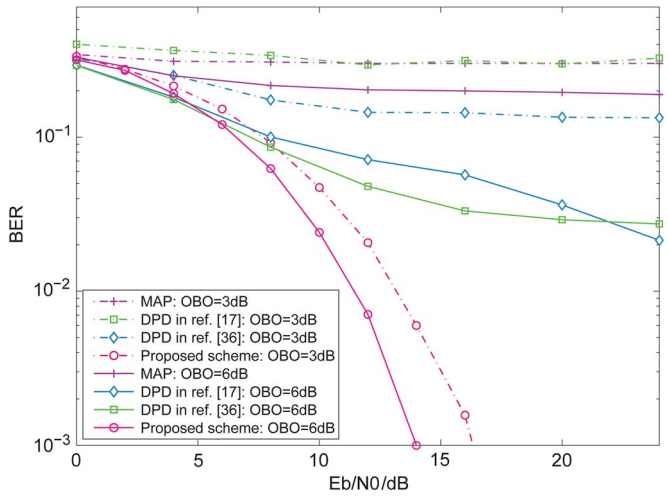


Fig. 10. Performance comparisons with the joint MAP estimation method (with LMMSE blind equalization) and transmitter-end DPD techniques.

the BER drops to 4×10^{-4}. To fully promote the detection performance, it seems that the linearization approach is relatively more effective in dealing with received symbols, which exhibit some level memory introduced by the PA nonlinearity. An important characteristic of the linearization method is that, with the full exploration of historical information, the importance density will be evolved sequentially, which may exploit the underlying memory more thoroughly and thereby may be more attractive to mm-Wave joint estimation (e.g., with dependent symbols/observations). For the PMF-based method, by passing independent symbols through a nonlinear PA, its importance density focuses primarily on the current information and, unfortunately, may underuse the involved memory effect.

E. Different Methods

For comparative analysis, some other existing schemes, i.e., the linear equalization algorithm and the DPD technique, are further investigated.

1) Blind Linear Equalization: To blindly equalize the multipath ISI channel, a joint MAP estimation scheme may be suggested. As in [21], the multipath response is sequentially estimated with the LMMSE criterion. Note that, like most blind equalization schemes, the linear PA has to be assumed in realizations.

2) Transmitting Pre-Distortion: First, the PA identification process of the DPD is carried out, in which the polynomial memory model is adopted. To be specific, the 7th odd-order nonlinearity is considered and the number of delay-taps are 3 [17]. This process is implemented with the assistance of a sequence of *a priori* training sequence. The length of training sequence is set to 512. After this supervised identification process, an inverse system of the identified PA model can be extracted which will be used to perform pre-distortion [17], [39]. It should be noteworthy that, in mm-Wave receivers, the LMMSE equalization and MAP estimation are also used other than the transmitting pre-distortion [21].

With regards to the joint MAP method [21], it seems from Fig. 10 that the estimation performance of 16QAM signals will be degraded dramatically. Since such traditional methods

may fail to deal with nonlinear distortions, the BER floor is relatively obvious in the presence of realistic PA nonlinearity (even if the OBO is set to 3 dB). For DPD techniques [17], [39], notice that the nonlinearity calibration will be conducted in transmitter-ends, which may also lead to complex computations and implementations in transmitters, especially for low-power and low-cost mm-Wave applications. Even so, its performance seems still to be less competitive when OBO $= 6$ dB. In comparison, by mitigating the BER floor and improving detection performance remarkably, the proposed nonlinear equalization and signal detection algorithm may effectively address realistic PA distortions.

F. Other Considerations

Except for the advantage of detection performance, the suggested scheme may be implemented completely in the baseband of receiver-ends. Thus, we may further analyze its potential advantages in some technical terms, e.g., energy consumption and chip designing. It has to be noted that, in practice, such technical quantities may vary with specific designing philosophies, e.g., the built-in or external linearization [40], the digital (baseband) or RF pre-distortion (RPD) [41], the adaptive feedback or LUT methods [42], [43], etc. 1) With regards to the power consumption, taking the common DPDs for example, a pre-distortion modular may conservatively account for 30% of total power consumption of a broadband transmitter [44]. When the most complicated pre-distortion structure is considered [41], i.e., involving the down-converter, ADC, digital signal processing, digital to analog converter (DAC) and DPD modular, the power consumption of pre-distortion may become more considerable, typically >30% of total power consumption. Note that, for some other simpler methods (e.g., LUT), the power consumption may be not so high. 2) With regards to the chip area, it seems also hard to give a unified evaluation. Taking the RPD for example, it has been shown that the pre-distortion modular may usually cover $5 \sim 15\%$ of the chip area. For the DPD technique, a pre-distorter may even account for >30% of the chip area usage, e.g., the LUT technique with multiple on-chip static random-access memories (SRAMs) [45].

On the other hand, the power consumption of the receiving procedure will be increased due to additional nonlinear estimation, since the baseband signal processing will be realized in receiver-ends. Although it may be hard to figure out additional consumption accurately, it seems that the introduction of nonlinear estimation may not increase the total power of receivers significantly. As in most joint estimation methods [21], [24], in fact, the estimation of multipath channels and unknown symbols has already consumed much of energy.

VI. Conclusion

A nonlinear equalization and signal detection scheme is proposed to combat the realistic PA nonlinearity of 5G mm-Wave communications, which is designed particularly for the baseband processing in receiver-ends. With such a new paradigm, which remains in contrast to the widely used DPDs in transmitter-ends, the received symbols contaminated by both nonlinear effects and multipath fading are blindly calibrated. By resorting to the Monte-Carlo SIS based numerical approximation approach, even with realistic nonlinear distortions, the unknown symbols and the multipath channel can be estimated jointly. A generalized PF scheme, embedded with a local linearization model, is further suggested to effectively cope with encountered nonlinear observations. Thus, the involved posterior density is derived numerically. Simulation results validate the proposed algorithm. It is shown that the new scheme can enhance the detection performance of seriously distorted signals and may also simplify the implementation complexity of transmitters by excluding sophisticated DPDs, which may provide a promising nonlinear signal detection framework for the emerging 5G mm-Wave communications.

Appendix A
Proof of Proposition 1

To derive the analytic importance distribution, we firstly construct a new variable

$$\tilde{\mathbf{x}}_k \triangleq \left.\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \times \mathbf{x}_k \tag{39}$$

$$= y_k - g\left(f(\mathbf{x}_{k-1})\right) + \left.\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} f(\mathbf{x}_{k-1}) - n_k. \tag{40}$$

Given the *a priori* Gaussian distribution of n_k, then we may have eqs. (41) and (42), shown at the bottom of the page. Notice from eq. (42), such an instrumental variable $\tilde{\mathbf{x}}_k$ follows a Gaussian distribution, i.e., $\tilde{\mathbf{x}}_k \sim \mathcal{N}(\tilde{m}, \tilde{\sigma})$, with the mean $\tilde{m} = g(f(\mathbf{x}_{k-1})) - \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \times f(\mathbf{x}_{k-1})$ and the variance $\tilde{\sigma} = \sigma$.

As the linearization is performed at the point $\mathbf{x}_k = f(\mathbf{x}_{k-1})$, then following the state transition in (6), at time index k the term $f(\mathbf{x}_{k-1}) \triangleq \mathbf{T}\mathbf{x}_{k-1}$ is only related with $\mathbf{x}_{0:k-1}$ which may be further replaced properly by using the derived particles $\mathbf{x}_{0:k-1}^{(i)}$. As a consequence, the term $\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}|_{\mathbf{x}_k=f(\mathbf{x}_{k-1}^{(i)})}$ can be regarded as a constant. Taking the fact $\mathbf{x}_k = \tilde{\mathbf{x}}_k \times \left(\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}|_{\mathbf{x}_k=f(\mathbf{x}_{k-1}^{(i)})}\right)^{-1}$

$$p\left(\tilde{\mathbf{x}}_k | y_k, \mathbf{x}_{0:k-1}, \mathbf{h}, (\mathbf{G}, \mathbf{\Psi})\right) \propto exp\left(-\frac{1}{\sigma^2}\left| y_k - g\left(f(\mathbf{x}_{k-1})\right) + \left.\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \times f(\mathbf{x}_{k-1})\right|^2\right) \tag{41}$$

$$\sim \mathcal{N}\left(g\left(f(\mathbf{x}_{k-1})\right) - \left.\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k)}\right|_{\mathbf{x}_k=f(\mathbf{x}_{k-1})} \times f(\mathbf{x}_{k-1}), \sigma^2\right) \tag{42}$$

into account, we may derive the analytic form of importance function as following:

$$p\left(\mathbf{x}_k|\mathbf{x}_{0:k-1}^{(i)}, y_{0:k}, \mathbf{h}, (\mathbf{G}, \mathbf{\Psi})\right) \sim \mathcal{N}(m_*, \sigma_*).$$

That is, the optimal importance function also follows the Gaussian distribution, with its mean and variance defined by:

$$m_* = \tilde{m} \times \left(\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k}\bigg|_{\mathbf{x}_k = f\left(\mathbf{x}_{k-1}^{(i)}\right)} \right)^{-1}, \tag{43}$$

and

$$\sigma_*^{(-1)} = \tilde{\sigma}^{-1} \times \left(\frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k}\bigg|_{\mathbf{x}_k = f\left(\mathbf{x}_{k-1}^{(i)}\right)} \right)^H \frac{\partial g(\mathbf{x}_k)}{\partial \mathbf{x}_k}\bigg|_{\mathbf{x}_k = f\left(\mathbf{x}_{k-1}^{(i)}\right)}, \tag{44}$$

respectively. Then, further manipulations on eqs. (43) and (44) will lead to eqs. (28) and (29).

Appendix B
Derivation of Eqs. (34), (35)

Following eq. (21), we may easily have:

$$\begin{aligned} p(\mathbf{h}|\mathbf{X}_k, y_{0:k}) &\propto p\left(y_{0:k}|\mathbf{h}, \mathbf{X}_k, (\mathbf{G}, \mathbf{\Psi})\right) p(\mathbf{h}) \\ &\propto exp\left(-\frac{1}{\sigma^2} \sum_{j=0}^{k} |y_j - g(\mathbf{x}_j)|^2 \right) \\ &\quad \times exp\left[-(\mathbf{h} - \mathbf{h}_0)^H \mathbf{\Sigma}_0^{-1} (\mathbf{h} - \mathbf{h}_0) \right]. \end{aligned}$$

By utilizing the detachable property between the linear multipath propagations and nonlinear effect, and simultaneously by fully exploiting the derived new estimation $\hat{\mathbf{x}}_k$, the first term in above equation can be further approximated by:

$$\begin{aligned} &p\left(y_{0:k}|\mathbf{h}, \mathbf{X}_k, (\mathbf{G}, \mathbf{\Psi})\right) \\ &\quad \simeq \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{1}{\sigma^2} \sum_{j=0}^{k} \left| y_j - \mathbf{h}^H \hat{\mathbf{x}}_j^\dagger \right|^2 \right). \end{aligned} \tag{45}$$

With the help of the auxiliary particles $\hat{\mathbf{x}}_k^\dagger$, further treatment on the above relationship leads to:

$$\begin{aligned} p\left(\mathbf{h}|\mathbf{X}_k, y_{0:k}, (\mathbf{G}, \mathbf{\Psi})\right) &\propto exp\left\{ -\mathbf{h}^H \underbrace{\left(\frac{1}{\sigma^2} \sum_{j=0}^{k} \hat{\mathbf{x}}_j^\dagger \hat{\mathbf{x}}_j^{\dagger H} + \mathbf{\Sigma}_0^{-1} \right)}_{\hat{\mathbf{\Sigma}}_k^{-1}} \mathbf{h} \right. \\ &\quad \left. + 2\mathfrak{Re}\left[\mathbf{h}^H \underbrace{\left(\frac{1}{\sigma^2} \sum_{j=0}^{k} \hat{\mathbf{x}}_j^\dagger y_j^* + \mathbf{\Sigma}_0^{-1} \mathbf{h}_0 \right)}_{\hat{\mathbf{\Sigma}}_k^{-1} \hat{\mathbf{h}}_k} \right] \right\} \\ &= exp\left[-(\mathbf{h} - \hat{\mathbf{h}}_k)^H \mathbf{\Sigma}_k^{-1} (\mathbf{h} - \hat{\mathbf{h}}_k) \right]. \end{aligned} \tag{46}$$

From eq. (46), $p(\mathbf{h}|\mathbf{x}_{0:k}, y_{0:k})$ follows also a Gaussian distribution $\mathcal{N}\left(\hat{\mathbf{h}}_k, \hat{\mathbf{\Sigma}}_k\right)$, with the covariance matrix and mean vector determined by $\hat{\mathbf{\Sigma}}_k^{-1} = \sum_{j=0}^{k} \hat{\mathbf{x}}_j^\dagger \hat{\mathbf{x}}_j^{\dagger H} / \sigma^2 + \hat{\mathbf{\Sigma}}_0^{-1}$ and $\hat{\mathbf{h}}_k = \hat{\mathbf{\Sigma}}_0 \left(\sum_{j=0}^{k} \hat{\mathbf{x}}_j^\dagger y_j^* \sigma^2 + \hat{\mathbf{\Sigma}}_0^{-1} \hat{\mathbf{h}}_0 \right)$ respectively. On this basis, we may easily have the recursion of channel covariance matrix by using $\hat{\mathbf{\Sigma}}_k^{-1} = \hat{\mathbf{x}}_k^\dagger \hat{\mathbf{x}}_k^{\dagger H} / \sigma^2 + \hat{\mathbf{\Sigma}}_{k-1}^{-1}$. Then, the Sherman-Marrison matrix inversion lemma can be applied and further manipulation may finally result in eqs. (34) and (35).

Acknowledgment

The authors would like to thank the anonymous reviewers for their thoughtful and constructive remarks that are helpful to improve the quality of this paper.

References

[1] S. K. Yong, P. Xia, and A. Valdes-Garcia, *60 GHz Technology for Gbps WLAN and WPAN*. Chichester, U.K.: Wiley, 2011.

[2] C. H. Park and T. S. Rappaport, "Short-range wireless communications for next-generation networks: UWB, 60 GHz millimeter-wave PAN, Zigbee," *IEEE Wireless Commun. Mag.*, vol. 14, no. 4, pp. 70–78, Aug. 2007.

[3] *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs): Amendment 2: Millimeter-Wave Based Alternative Physical Layer Extension*, IEEE Std. 802.15.3c, 2009.

[4] *IEEE 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 3: Enhancements for Very High Throughput in the 60 GHz Band*, IEEE Std 802.11ad-2012, 2012.

[5] H. Singh *et al.*, "A 60 GHz wireless network for enabling uncompressed video communication," *IEEE Commun. Mag.*, vol. 46, no. 12, pp. 71–78, Dec. 2008.

[6] K. Ohata *et al.*, "1.25 Gbps wireless gigabit ethernet link at 60 GHz-band," in *Proc IEEE MTT-S Int. Microw. Symp. Dig.*, Philadelphia, PA, USA, 2003, vol. 1, pp. 373–376.

[7] S. Kato *et al.*, "Single carrier transmission for multi-gigabit 60-GHz WPAN system," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1466–1478, Oct. 2009.

[8] X. Zhang, L. R. Lu, R. Funada, C. S. Sum, and H. Harada, "Physical layer design and performance analysis on multi-Gbps millimeter-wave WLAN system," in *Proc IEEE ICCS*, Singapore, Nov. 2010, pp. 92–96.

[9] C. R. Anderson and T. S. Rappaport, "In-building wideband partition loss measurements at 2.5 and 60 GHz," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 922–928, May 2004.

[10] M. Lei, C. S. Choi, R. Funada, H. Harada, and S. Kato, "Throughput comparison of multi-Gbps WPAN (IEEE 802.15.3c) PHY layer designs under non-linear 60-GHz power amplifier," in *Proc IEEE Int. Symp. PIMRC*, Athens, Greece, Sep. 2007, pp. 1–5.

[11] E. Perahia *et al.*, IEEE P802.11 Wireless LANs TGad Evaluation Methodology, 2010.

[12] W. Gerhard and R. Knoechel, "Improvement of power amplifier efficiency by reactive Chireix combining, power back-off and differential phase adjustment," in *Proc IEEE MTT-IMS*, San Francisco, CA, USA, Jun. 2006, pp. 1887–1890.

[13] N. Safari, J. P. Tanem, and T. Roste, "A block-based predistortion for high power-amplifier linearization," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 6, pp. 2813–2820, Jun. 2006.

[14] A. Zhu *et al.*, "Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 7, pp. 1524–1534, Jul. 2008.

[15] M. Rawat, K. Rawat, and F. M. Ghannouchi, "Adaptive digital predistortion of wireless power amplifiers/transmitters using dynamic real-valued focused time-delay line neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 95–104, Jan. 2010.

[16] G. Montoro, P. L. Gilabert, E. Bertran, A. Cesari, and J. A. Garcła, "An LMS-based adaptive predistorter for cancelling nonlinear memory effects in RF power amplifiers," in *Proc. APMC*, Bangkok, Thailand, Dec. 2007, pp. 1–4.

[17] L. Ding *et al.*, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.
[18] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer-Verlag, 2001.
[19] A. Doucet, "On sequential simulation-based methods for bayesian filtering," *Stat. Comput.*, vol. 10, no. 3, pp. 197–208, Jul. 2000.
[20] P. M. Djuric *et al.*, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, Sep. 2003.
[21] J. Miguez and P. M. Djuric, "Blind equalization of frequency-selective channels by sequential importance sampling," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2738–2748, Oct. 2004.
[22] D. Crisan, "Particle filters: A theoretical perspective," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G. de Freitas, and N. J. Gordon, Eds. New York, NY, USA: Springer-Verlag, 2001, pp. 17–38.
[23] J. Liu and R. Chen, "Blind deconvolution via sequential imputations," *J. Amer. Stat. Assoc.*, vol. 90, no. 430, pp. 567–576, Jun. 1995.
[24] Z. G. Yang and X. D. Wang, "A sequential Monte Carlo blind receiver for OFDM systems in frequency-selective fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 271–280, Feb. 2002.
[25] T. Bertozzi, D. Le Ruyet, G. Rigal, and H. Vu-Thien, "Joint data-channel estimation using the particle filtering on multipath fading," in *Proc. ICT*, Papeete, Tahiti, 2003, pp. 1284–1289.
[26] Y. Huang and P. M. Djuric, "A new importance function for particle filtering and its application to blind detection in flat fading channels," in *Proc. IEEE ICASSP*, 2002, pp. 1617–1620.
[27] J. H. Kotecha and P. M. Djuric, "Sequential Monte Carlo sampling detector for Rayleigh fast-fading channels," in *Proc. IEEE ICASSP*, 2000, pp. 61–64.
[28] J. C. Pedro, N. B. Carvalho, and P. M. Lavrador, "Modeling nonlinear behavior of band-pass memoryless and dynamic systems," in *Proc. IEEE MTTS Int. Microw. Symp. Dig. MTT-IMS*, Philadelphia, PA, USA, Jun. 2003, pp. 2133–2136.
[29] S. K. Yong, "TG3c channel modeling sub-committee final report," IEEE 802.15 TG3c Working Group for Wireless Personal Area Network (WPAN), IEEE 802.15-07-0584-01-003c, May 2007.
[30] K. Sato, H. Sawada, Y. Shoji, and S. Kato, "Channel model for millimeter-wave WPAN," in *Proc. IEEE 18th Int. Symp. PIMRC*, Athens, Greece, Sep. 2007, pp. 1–5.
[31] B. Li, Z. Zhou, W. X. Zou, X. B. Sun, and G. L. Du, "On the efficient beam pattern training for 60 GHz wireless personal area networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 504–515, Feb. 2013.
[32] Z. Y. Xiao, X. G. Xia, D. P. Jin, and N. Ge, "Iterative eigenvalue decomposition and multipath-grouping Tx/Rx joint beamformings for millimeter-wave communications," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1595–1607, Mar. 2015.
[33] B. Li, Z. Zhou, H. J. Zhang, and A. Nallanathan, "Efficient beam-forming training for 60 GHz millimeter-wave communications: A novel numerical optimization framework," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 703–717, Feb. 2014.
[34] J. H. K. Vuolevi, T. Rahkonen, and J. P. A. Manninen, "Measurement technique for characterizing memory effects in RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 8, pp. 1383–1389, Aug. 2001.
[35] H. Ku, M. D. McKinley, and J. S. Kenney, "Extraction of accurate behavior models for power amplifiers with memory effects using two-tone measurements," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Seattle, WA, USA, Jun. 2002, pp. 139–142.
[36] A. D. Zhu and T. J. Brazil, "Behavioral modeling of RF power amplifiers based on pruned volterra series," *IEEE Microw. Wireless Compon. Lett.*, vol. 14, no. 12, pp. 563–565, Dec. 2004.
[37] S. J. Julier and K. U. Jeffery, "A new extension of the Kalman filter to nonlinear systems," in *Proc. SPIE 11th Int. Symp. Aerosp. Defense Sens., Simul. Controls, Multi Sensor Fusion, Tracking Resource Manage. II*, 1997, p. 182.
[38] J. Borwein and P. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*. Hoboken, NJ, USA: Wiley, 1987.
[39] N. Safari, T. Roste, P. Fedorenko, and J. S. Kenney, "An approximation of volterra series using delay envelopes, applied to digital predistortion of RF power amplifiers with memory effects," *IEEE Microw. Wireless Compon. Lett.*, vol. 18, no. 2, pp. 115–117, Feb. 2008.
[40] H. Y. Yang and T. W. Huang, "60 GHz CMOS power amplifier with built-in pre-distortion linearizer," *IEEE Trans. Microw. Theory Techn.*, vol. 21, no. 12, pp. 676–678, Dec. 2011.
[41] *RF Predistortion (RFPD) vs. Digital Predistortion (DPD)*, Tech. Rep. [Online]. Available: http://www.scintera.com/technology/rf-predistortion-rfpd-vs-digital-predistortion-dpd/
[42] A. D. Zhu *et al.*, "Digital predistortion for envelope-tracking power amplifiers using decomposed piecewise volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 10, pp. 2237–2247, Oct. 2008.
[43] K. J. Muhonen, M. Kavehrad, and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: A development and comparison," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1995–1200, Sep. 2000.
[44] *Nujira Sets World Record for ET PA Linearity Without Use of DPD, Technical News*. [Online]. Available: http://www.nujira.com/nujira-sets-world-record-for-et-linearity-without-use-of-dpd-i-364.php
[45] M. S. Alavi, R. B. Staszewski, L. C. N. de Vreede, and J. R. Long, "A wideband 2 13-bit all-digital I/Q RF-DAC," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 4, pp. 732–752, Apr. 2014.

Bin Li received the bachelor's degree in electrical information engineering from Beijing University of Chemical Technology (BUCT), in 2007 and the Ph.D. degree in communication and information engineering from Beijing University of Posts and Telecommunications (BUPT), in 2013. He joined BUPT in 2013 as a Lecturer at the School of Information and Communication Engineering (SCIE). His current research interests are focused on statistical signal processing algorithms for wireless communications, e.g., ultra-wideband (UWB), wireless sensor networks, millimeter-wave (mm-Wave) communications, and cognitive radios (CRs). He has published more than 60 journal and conference papers. He received the 2011 ChinaCom Best Paper Award and the 2010 and 2011 BUPT Excellent Ph.D Student Foundation Award. He serves as the Co-Chair of the Technical program Committee of the Signal Processing for Communications Symposium of the 2016 IEEE International Conference on Computing, Networking and Communications (IEEE-ICNC'16).

Chenglin Zhao received the bachelor's degree in radio-technology from Tianjin University, in 1986, the master's degree in circuits and systems from Beijing University of Posts and Telecommunications (BUPT), in 1993, and the Ph.D. degree in communication and information system from Beijing University of Posts and Telecommunications, in 1997. At present, he serves as a Professor at Beijing University of Posts and Telecommunications, Beijing, China. His research is focused on emerging technologies of short-range wireless communication, cognitive radios, 60 GHz millimeter-wave communications.

Mengwei Sun received the bachelor's degree in communication engineering from Civil Aviation University of China, Tianjin, China, in 2011. She is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. Her current research interests focus on signal processing for cognitive radios and 60-GHz millimeter-wave communications.

Haijun Zhang received the Ph.D. degree from the School of Information and Communication Engineering (SICE), Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts Telecommunications (BUPT). Currently, he is a Postdoctoral Research Fellow in the Department of Electrical and Computer Engineering, the University of British Columbia (UBC), and also an Assistant Professor in the College of Information Science and Technology, Beijing University of Chemical Technology. From September 2011 to September 2012, he visited the Centre for Telecommunications Research, King's College London, London, U.K., as a joint Ph.D. student and Visiting Research Associate. He has published more than 40 papers and authored two books. He served as Chair of the Technical Program Committee of the Game Theory in Wireless Communications Symposium of the 2014 International Conference on Game Theory for Networks (GAMENETS'14). His current research interests include wireless resource allocation, 5G, heterogeneous small cell networks and ultra-dense networks.

Zheng Zhou (M'05) received the bachelor's degree in electrical engineering from the Harbin Institute of Military Engineering, Harbin, China, in 1967, and the master's and Ph.D. degrees in electrical engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1982 and 1988, respectively.

Supported by the Hong Kong Telecom International Postdoctoral Fellowship, he was a Visiting Research Fellow in the Information Engineering Department, Chinese University of Hong Kong, Hong Kong, from 1993 to 1995. He was also the Vice-Dean in the School of Telecommunication Engineering, BUPT, from 1998 to 2003, and was the Invited Overseas Researcher at the Japan Kyocera DDI Future Communication Research Institute (supported by the Japan Key Technology Center) in 2000. At present, he is a Professor at Beijing University of Posts and Telecommunications, Beijing, China.

Dr. Zhou served as a member of the Technical Subcommittee on Cognitive Networks (TCCN), IEEE Communications Society, the International Steering Committee Member of IEEE ISCIT during 2003–2010 (International Symposium on Communications and Information Technologies), and the TPC co-chair of the IEEE ISCIT 2005. He was also the General Vice Chair of the IEEE ChinaCom 2006 (the first international conference on communications and networking in China), and a steering committee member of IEEE ChinaCom 2007. He is a voting member and contributor of the IEEE 802.15 Task Group (TG3a and TG4a), senior member of China Institution of Communications (CIC), radio application and management technical committee member of CIC, senior member of China Computer Federation (CCF), sensor network technical committee member of CCF, *H*-subcommittee member of China Radio Interference Standard Technology Committee, the General Secretary of China UWB Forum, and the General Secretary of China Bluetooth Forum.

Arumugam Nallanathan (S'97–M'00–SM'05) is a Professor of Wireless Communications in the Department of Informatics, King's College London (University of London). He served as the Head of Graduate Studies in the School of Natural and Mathematical Sciences, King's College London, 2011/12. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include 5G Technologies, Millimeter wave communications, Cognitive Radio and Relay Networks. In these areas, he co-authored nearly 250 papers. He is a co-recipient of the Best Paper Award presented at the 2007 IEEE International Conference on Ultra-Wideband (ICUWB2007). He is a Distinguished Lecturer of IEEE Vehicular Technology Society.

He is an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and a Guest Editor for IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING: SPECIAL ISSUE ON ADVANCES IN MOBILE AND CLOUD COMPUTING. He was an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2006–2011), IEEE WIRELESS COMMUNICATIONS LETTERS and IEEE SIGNAL PROCESSING LETTERS. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society, Technical Program Co-Chair (MAC track) for IEEE WCNC 2014, Co-Chair for the IEEE GLOBECOM 2013 (Communications Theory Symposium), Co-Chair for the IEEE ICC 2012 (Signal Processing for Communications Symposium), Co-Chair for the IEEE GLOBECOM 2011 (Signal Processing for Communications Symposium), Technical Program Co-Chair for the IEEE International Conference on UWB 2011 (IEEE ICUWB 2011), Co-Chair for the IEEE ICC 2009 (Wireless Communications Symposium), Co-chair for the IEEE GLOBECOM 2008 (Signal Processing for Communications Symposium) and General Track Chair for IEEE VTC 2008. He received the IEEE Communications Society SPCE Outstanding Service Award 2012 and IEEE Communications Society RCC Outstanding Service Award 2014.