This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Nonlinear equalization of Hammerstein OFDM systems

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1109/TSP.2014.2355773>

PUBLISHER

© IEEE

VERSION

AM (Accepted Manuscript)

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Hong, Xia, Sheng Chen, Yu Gong, and Chris J. Harris. 2019. "Nonlinear Equalization of Hammerstein OFDM Systems". figshare. <https://hdl.handle.net/2134/25653>.

Nonlinear Equalization of Hammerstein OFDM Systems

Xia Hong, *Senior Member, IEEE*, Sheng Chen, *Fellow, IEEE*, Yu Gong, *Member, IEEE*, and Chris J. Harris

Abstract

A practical orthogonal frequency-division multiplexing (OFDM) system can generally be modelled by the Hammerstein system that includes the nonlinear distortion effects of the high power amplifier (HPA) at transmitter. In this contribution, we advocate a novel nonlinear equalisation scheme for OFDM Hammerstein systems. We model the nonlinear HPA, which represents the static nonlinearity of the OFDM Hammerstein channel, by a B-spline neural network, and we develop a highly effective alternating least squares algorithm for estimating the parameters of the OFDM Hammerstein channel, including channel impulse response (CIR) coefficients and the parameters of the B-spline model. Equalisation of the OFDM Hammerstein channel can then be accomplished by the usual one-tap linear equalisation as well as the inversion of the estimated B-spline neural network model. We propose to use an efficient Gauss-Newton algorithm for the latter inversion task. The effectiveness of our nonlinear equalisation scheme for OFDM Hammerstein channels is demonstrated by simulation results.

Index Terms

Orthogonal frequency-division multiplexing, nonlinear high power amplifier, Hammerstein channel, equalisation, B-spline neural networks, De Boor algorithm

I. INTRODUCTION

Orthogonal frequency-division multiplexing (OFDM) [1], [2] has found its way into numerous recent wireless network standards, owing to its virtues of resilience to frequency selective fading

X. Hong is with School of Systems Engineering, University of Reading, Reading RG6 6AY, U.K. (E-mail: x.hong@reading.ac.uk).

S. Chen and C.J. Harris are with Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK (E-mails: sqc@ecs.soton.ac.uk, cjh@ecs.soton.ac.uk). S. Chen is also with Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia.

Y. Gong is with the School of Electronic, Electrical and Systems Engineering, Loughborough University, Loughborough LE11 3TU, UK (E-mail: Y.Gong@lboro.ac.uk).

channels. Both the modulation and demodulation operations of an OFDM system facilitate convenient low-complexity hardware implementations with the aid of the inverse fast Fourier transform (IFFT) and fast Fourier transform (FFT) operations. However, OFDM signals are notoriously known to have high peak to average power ratios, and a transmitted OFDM signal can be seriously distorted by the high power amplifier (HPA) at the transmitter, which exhibits nonlinear saturation characteristics [3]–[7]. Thus, the nonlinearities of the HPA at transmitter will significantly degrade the OFDM system's achievable bit error rate (BER) performance, and it is particularly critical to compensate the nonlinear distortions of the HPA in the design of a OFDM wireless system.

An effective approach to compensate the nonlinear distortions of HPA is to implement a digital predistorter at the transmitter, which is capable of achieving excellent performance, and various predistorter techniques have been developed [8]–[14]. Implementing the predistorter is attractive for the downlink, where the base station (BS) transmitter has the sufficient hardware and software capacities to accommodate the hardware and computational requirements for implementing digital predistorter. In the uplink, however, implementing predistorter at transmitter is difficult, because it is much more challenging for a pocket-size handset to absorb the additional hardware and computational complexity. Alternatively, the nonlinear distortions of the transmitter HPA can be dealt with at the BS receiver, which has sufficient hardware and software resources. With the nonlinear HPA at transmitter, the channel is a complex-valued (CV) nonlinear Hammerstein system and, moreover, the received signal is further impaired by the channel additive white Gaussian noise (AWGN). Therefore, inversion or equalisation of the OFDM Hammerstein channel is not a trivial task.

Against this background, in this paper, we develop a highly effective nonlinear equalisation scheme for OFDM Hammerstein channels based on the B-spline neural network. The reason that we adopt the B-spline neural network is because it has been demonstrated to be very effective in identification and inversion of CV Wiener systems [14], [15]. Specifically, we propose an efficient alternating least squares (ALS) identification algorithm for estimating the channel impulse response (CIR) coefficients together with the parameters of the B-spline neural network that models the HPA static nonlinearity of the OFDM Hammerstein channel. As linear equalization is naturally accomplished in OFDM systems by a simple yet effective one-tap equalisation in frequency domain (FD), nonlinear equalization of the OFDM Hammerstein

channel only additionally involves the inversion of the estimated B-spline neural network, and we propose a computationally efficient Gauss-Newton algorithm to solve this two-dimensional inverse problem. Simulation results are presented to demonstrate the effectiveness of our proposed B-spline neural network based nonlinear equalisation scheme for OFDM Hammerstein channels.

Throughout our discussion, a CV number $x \in \mathbb{C}$ is represented either by the rectangular form $x = x_R + \mathbf{j} \cdot x_I$, where $\mathbf{j} = \sqrt{-1}$, while $x_R = \Re[x]$ and $x_I = \Im[x]$ denote the real and imaginary parts of x , or alternatively by the polar form $x = |x| \cdot e^{\mathbf{j}\angle x}$ with $|x|$ denoting the amplitude of x and $\angle x$ its phase. The vector or matrix transpose and conjugate transpose operators are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively, while $(\cdot)^{-1}$ stands for the inverse operation and the expectation operator is denoted by $\mathbb{E}\{\cdot\}$. Furthermore, $\mathbf{I}$ denotes the identity matrix with an appropriate dimension, and $\text{diag}\{x_0, x_1, \dots, x_{n-1}\}$ is the diagonal matrix with $x_0, x_1, \dots, x_{n-1}$ as its diagonal elements.

II. OFDM HAMMERSTEIN CHANNEL MODEL

We consider the OFDM system with N subcarriers and employing the M -quadrature amplitude modulation (QAM). The s th FD OFDM symbol vector is expressed as

$$\mathbf{X}[s] = [X_0[s] \ X_1[s] \ \cdots \ X_{N-1}[s]]^T, \quad (1)$$

where $X_n[s]$, $0 \leq n \leq N-1$, denotes the CV data symbol at the n th subcarrier, which takes the values from the M -QAM symbol set

$$\mathbb{X} = \{d(2l - \sqrt{M} - 1) + \mathbf{j} \cdot d(2q - \sqrt{M} - 1), 1 \leq l, q \leq \sqrt{M}\}, \quad (2)$$

where $2d$ is the minimum distance between symbol points. l, q are integers. For notational simplification, we will drop the OFDM symbol index $[s]$ in the sequel. Feeding $\mathbf{X}$ through the N -point IFFT based modulator yields the time-domain (TD) OFDM signal

$$x_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X_n e^{\mathbf{j}2\pi nk/N}, \quad 0 \leq k \leq N-1. \quad (3)$$

Define the FFT matrix $\mathbf{F} \in \mathbb{C}^{N \times N}$ given by

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-\mathbf{j}2\pi/N} & \cdots & e^{-\mathbf{j}2\pi(N-1)/N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-\mathbf{j}2\pi(N-1)/N} & \cdots & e^{-\mathbf{j}2\pi(N-1)(N-1)/N} \end{bmatrix}, \quad (4)$$

which has the orthogonal property of $\mathbf{F}^H \mathbf{F} = \mathbf{F} \mathbf{F}^H = \mathbf{I}$, and let

$$\mathbf{x} = [x_0 \ x_1 \ \cdots \ x_{N-1}]^T. \quad (5)$$

Then, the IFFT based modulation operation (3) can be expressed concisely by

$$\mathbf{x} = \mathbf{F}^H \mathbf{X}. \quad (6)$$

After adding the cyclic prefix (CP) of length N_{CP} to $\mathbf{x}$, the resultant TD OFDM signal is amplified by the HPA and transmitted over the channel whose channel impulse response (CIR) has the length of $L_{\text{CIR}} \leq N_{\text{CP}}$. For notational simplification, we will omit the CP addition operation performed on $\mathbf{x}$. Thus, the actually transmitted TD signal vector $\mathbf{w} = [w_0 \ w_1 \ \cdots \ w_{N-1}]^T$ is defined by

$$w_k = \Psi(x_k), \ 0 \leq k \leq N-1, \quad (7)$$

where $\Psi(\cdot)$ represents the CV static nonlinearity of the transmitter HPA. We consider the solid state power amplifier [6], [7], whose nonlinearity $\Psi(\cdot)$ is constituted by the HPA's amplitude response $A(r)$ and phase response $\Upsilon(r)$ given by

$$A(r) = \frac{g_a r}{\left(1 + \left(\frac{g_a r}{A_{\text{sat}}}\right)^{2\beta_a}\right)^{\frac{1}{2\beta_a}}}, \quad (8)$$

$$\Upsilon(r) = \frac{\alpha_\phi r^{q_1}}{1 + \left(\frac{r}{\beta_\phi}\right)^{q_2}}, \quad (9)$$

where r denotes the amplitude of the input to the HPA, g_a is the small gain signal, β_a is the smoothness factor and A_{sat} is the saturation level, while the parameters of the phase response, α_ϕ , β_ϕ , q_1 and q_2 , are adjusted to match the specific amplifier's characteristics. The NEC GaAs power amplifier used in the standardization [6], [7] has the parameter set

$$g_a = 19, \beta_a = 0.81, A_{\text{sat}} = 1.4; \alpha_\phi = -48000, \beta_\phi = 0.123, q_1 = 3.8, q_2 = 3.7. \quad (10)$$

Hence, given the input $x_k = |x_k| \cdot e^{j\angle x_k}$, the output of the HPA can be expressed as

$$w_k = A(|x_k|) \cdot e^{j(\angle x_k + \Upsilon(|x_k|))}. \quad (11)$$

The operating status of the HPA may be specified by the output back-off (OBO), which is defined as the ratio of the maximum output power P_{max} of the HPA to the average output power P_o of the signal at the HPA output, given by

$$\text{OBO} = 10 \cdot \log_{10} \frac{P_{\text{max}}}{P_o}. \quad (12)$$

The smaller OBO is, the more the HPA is operating into the nonlinear saturation region.

Denote $\mathbf{h} = [h_0 \ h_1 \ \cdots \ h_{L_{\text{CIR}}}]^T$ as the vector of channel impulse response (CIR) coefficients. It is assumed $h_0 = 1$ because if this is not the case, h_0 can always be absorbed into the CV static nonlinearity $\Psi(\bullet)$, and the channel impulse response coefficients are re-scaled as h_i/h_0 for $0 \leq i \leq L_{\text{CIR}}$. At the receiver, after the CP removal, the channel-impaired received signals y_k is given by

$$y_k = \sum_{i=0}^{L_{\text{CIR}}} h_i w_{k-i} + e_k \quad k = 0, \dots, N-1 \quad (13)$$

in which $w_{k-i} = w_{N+k-i}$ for $k < i$, $e_k = e_{R_k} + j \cdot e_{I_k}$ is the TD AWGN with $E\{e_{R_k}^2\} = E\{e_{I_k}^2\} = \sigma^2$. Because $N_{\text{CP}} \geq L_{\text{CIR}}$, the CP removal at the receiver automatically cancels the inter block interference and transfers the linear convolution channel into the circular one. y_k and h_i are respectively passed to the N -point FFT processor to yield the FD received OFDM vector $\mathbf{Y} = [Y_0 \ Y_1 \ \cdots \ Y_{N-1}]^T$ and the frequency domain channel transfer function coefficients (FDCTFCs) H_n , $n = 0, \dots, N-1$, satisfying the relation

$$Y_n = H_n W_n + \Xi_n, \quad 0 \leq n \leq N-1, \quad (14)$$

where $\Xi_{R_n} + j \cdot \Xi_{I_n}$ is the FD n th subcarrier AWGN with $E\{\Xi_{R_n}^2\} = E\{\Xi_{I_n}^2\} = \sigma^2$, and

$$\mathbf{W} = [W_0 \ W_1 \ \cdots \ W_{N-1}]^T = \mathbf{F}\mathbf{w} \quad (15)$$

is the N -point FFT of $\mathbf{w}$. Note that $\mathbf{w}$ is unobservable and therefore $\mathbf{W}$ is unavailable.

If we denote $\mathbf{\Xi} = [\Xi_0 \ \Xi_1 \ \cdots \ \Xi_{N-1}]^T$, the received FD OFDM signal (13) can be expressed concisely as

$$\mathbf{Y} = \text{diag}\{H_0, H_1, \dots, H_{N-1}\} \mathbf{W} + \mathbf{\Xi} = \text{diag}\{H_0, H_1, \dots, H_{N-1}\} \mathbf{F}\mathbf{w} + \mathbf{\Xi}. \quad (16)$$

Given the FDCTFCs H_n for $0 \leq n \leq N-1$, the FD one-tap equalisation can be carried out. The zero-forcing equalisation, for example, is given by

$$\widetilde{W}_n = \frac{Y_n}{H_n}, \quad 0 \leq n \leq N-1. \quad (17)$$

If the HPA $\Psi(\cdot)$ at the transmitter were linear, $\widetilde{Y}_n$ would be an estimate of the transmitted data symbol X_n . But $\Psi(\cdot)$ is nonlinear, and the linear equalisation (16) alone is no longer sufficient for estimating $\mathbf{X}$. If the nonlinearity $\Psi(\cdot)$ is known and it is invertible, then the effects of $\Psi(\cdot)$ can be compensated by inverting it. Specifically, let us define $\widetilde{\mathbf{W}} = [\widetilde{W}_0 \ \widetilde{W}_1 \ \cdots \ \widetilde{W}_{N-1}]^T$ and

$$\mathbf{\Psi}(\mathbf{x}) = [\Psi(x_0) \ \Psi(x_1) \ \cdots \ \Psi(x_{N-1})]^T. \quad (18)$$

Performing the IFFT on $\tilde{\mathbf{Y}}$ yields

$$\tilde{\mathbf{w}} = [\tilde{w}_0 \ \tilde{w}_1 \cdots \tilde{w}_{N-1}]^T = \mathbf{F}^H \tilde{\mathbf{W}} = \boldsymbol{\Psi}(\mathbf{x}) + \mathbf{F}^H \boldsymbol{\Xi}. \quad (19)$$

Thus, an estimate of the TD OFDM signal $\mathbf{x}$ is given by

$$\hat{\mathbf{x}} = \boldsymbol{\Psi}^{-1}(\tilde{\mathbf{w}}) = [\Psi^{-1}(\tilde{w}_0) \ \Psi^{-1}(\tilde{w}_1) \cdots \Psi^{-1}(\tilde{w}_{N-1})]^T, \quad (20)$$

which further yields the estimate of the FD OFDM symbol vector $\mathbf{X}$ by

$$\hat{\mathbf{X}} = \mathbf{F} \hat{\mathbf{x}}. \quad (21)$$

III. NONLINEAR EQUALISATION OF OFDM HAMMERSTEIN SYSTEM

As discussed in the previous section, the reliable detection of the transmitted OFDM data symbols depends on the ability of estimating the FDCTFCs H_n and the CV static nonlinearity $\Psi(\cdot)$ of the transmitter HPA as well as the ability of inverting $\Psi(\cdot)$. Note that the CV HPA's nonlinearity, (8) and (9), is completely unknown to the receiver and $\mathbf{w}$ is unobserved. We adopt the CV B-spline neural network [15], [16] to represent the mapping $\hat{w} = \hat{\Psi}(x_R + j \cdot x_I) : \mathbb{C} \rightarrow \mathbb{C}$ that is the estimate of the underlying CV nonlinear function $\Psi(\cdot)$. We then propose an efficient algorithm for jointly estimating H_n and $\Psi(\cdot)$ based on the CV B-spline modelling of $\Psi(\cdot)$. A significant advantage of our CV B-spline neural network is that its inversion can be effectively calculated [15]. Before introducing the B-spline modelling of $\Psi(\cdot)$, we point out that the HPA $\Psi(\cdot)$ of (8) and (9) satisfies the following conditions.

- 1) $\Psi(\cdot)$ is a one to one mapping, i.e. it is an invertible and continuous function.
- 2) x_R and x_I are upper and lower bounded by some finite and known real values, where

$$x = x_R + j \cdot x_I \text{ denotes the input to the HPA } \Psi(\cdot).$$

According to the property 2), we assume that $U_{\min} < x_R < U_{\max}$ and $V_{\min} < x_I < V_{\max}$, where $U_{\min}$, $U_{\max}$, $V_{\min}$ and $V_{\max}$ are known finite real values.

A. Complex-valued B-spline neural network

A set of univariate B-spline basis functions based on x_R is parametrised by the order $(P_o - 1)$ of a piecewise polynomial and a knot sequence which is a set of values defined on the real line that break it up into a number of intervals. To have N_R basis functions, the knot sequence is

specified by $(N_R + P_o + 1)$ knot values, $\{U_0, U_1, \dots, U_{N_R+P_o}\}$, with

$$\begin{aligned} U_0 < U_1 < \dots < U_{P_o-2} < U_{P_o-1} = U_{\min} < U_{P_o} < \dots < \\ U_{N_R} < U_{N_R+1} = U_{\max} < U_{N_R+2} < \dots < U_{N_R+P_o}. \end{aligned} \quad (22)$$

At each end, there are $P_o - 1$ “external” knots that are outside the input region and one boundary knot. As a result, the number of “internal” knots is $N_R + 1 - P_o$. Given the set of predetermined knots (22), the set of N_R B-spline basis functions can be formed by using the De Boor recursion [17], yielding for $1 \leq l \leq N_R + P_o$,

$$B_l^{(\mathbb{R},0)}(x_R) = \begin{cases} 1, & \text{if } U_{l-1} \leq x_R < U_l, \\ 0, & \text{otherwise,} \end{cases} \quad (23)$$

as well as for $l = 1, \dots, N_R + P_o - p$ and $p = 1, \dots, P_o$,

$$B_l^{(\mathbb{R},p)}(x_R) = \frac{x_R - U_{l-1}}{U_{p+l-1} - U_{l-1}} B_l^{(\mathbb{R},p-1)}(x_R) + \frac{U_{p+l} - x_R}{U_{p+l} - U_l} B_{l+1}^{(\mathbb{R},p-1)}(x_R). \quad (24)$$

The derivatives of the basis functions $B_l^{(\mathbb{R},P_o)}(x_R)$ for $1 \leq l \leq N_R$ can also be computed recursively according to

$$\frac{dB_l^{(\mathbb{R},P_o)}(x_R)}{dx_R} = \frac{P_o}{U_{P_o+l-1} - U_{l-1}} B_l^{(\mathbb{R},P_o-1)}(x_R) - \frac{P_o}{U_{P_o+l} - U_l} B_{l+1}^{(\mathbb{R},P_o-1)}(x_R). \quad (25)$$

The De Boor recursion is illustrated in Fig. ???. $P_o = 3$ to 5 is sufficient for most practical applications. The number of B-spline basis functions should be chosen to be sufficiently large to provide accurate approximation capability but not too large as to cause overfitting and to impose unnecessary computational complexity. The internal knots may be uniformly spaced in the interval $[U_{\min}, U_{\max}]$. The extrapolation capability of the B-spline model is influenced by the choice of the external knots. Note that there exist no data for $x_R < U_{\min}$ and $x_R > U_{\max}$ in identification but it is desired that the B-spline model has certain extrapolating capability outside the interval $[U_{\min}, U_{\max}]$. The external knots can be set empirically to meet the required extrapolation capability.

Similarly, a set of univariate B-spline basis functions based on x_I can be established. Suppose that the order of the piecewise polynomial is again $(P_o - 1)$ and there are N_I basis functions. Then the knot vector is defined on the imaginary line in a similar manner, which is specified by the $(N_I + P_o + 1)$ knot values, $\{V_0, V_1, \dots, V_{N_I+P_o}\}$. Specifically,

$$\begin{aligned} V_0 < V_1 < \dots < V_{P_o-2} < V_{P_o-1} = V_{\min} < V_{P_o} < \dots \\ < V_{N_I} < V_{N_I+1} = V_{\max} < V_{N_I+2} < \dots < V_{N_I+P_o}. \end{aligned} \quad (26)$$

Again, at each end, there are $P_o - 1$ external knots that are outside the input region and one boundary knot. Thus, the number of internal knots is $N_I + 1 - P_o$. Similarly, the set of N_I B-spline basis functions are constructed by the De Boor recursion [17] as, for $1 \leq m \leq N_I + P_o$,

$$B_m^{(\Im,0)}(x_I) = \begin{cases} 1, & \text{if } V_{m-1} \leq x_I < V_m, \\ 0, & \text{otherwise,} \end{cases} \quad (27)$$

as well as for $m = 1, \dots, N_I + P_o - p$ and $p = 1, \dots, P_o$,

$$B_m^{(\Im,p)}(x_I) = \frac{x_I - V_{m-1}}{V_{p+m-1} - V_{m-1}} B_m^{(\Im,p-1)}(x_I) + \frac{V_{p+m} - x_I}{V_{p+m} - V_m} B_{m+1}^{(\Im,p-1)}(x_I), \quad (28)$$

while the derivatives of the B-spline basis functions $B_m^{(\Im,P_o)}(x_I)$ for $1 \leq m \leq N_I$ are given recursively by

$$\frac{dB_m^{(\Im,P_o)}(x_I)}{dx_I} = \frac{P_o}{V_{P_o+m-1} - V_{m-1}} B_m^{(\Im,P_o-1)}(x_I) - \frac{P_o}{V_{P_o+m} - V_m} B_{m+1}^{(\Im,P_o-1)}(x_I). \quad (29)$$

Using the tensor product between the two sets of univariate B-spline basis functions [18], $B_l^{(\Re,P_o)}(x_R)$ for $1 \leq l \leq N_R$ and $B_m^{(\Im,P_o)}(x_I)$ for $1 \leq m \leq N_I$, a set of new B-spline basis functions $B_{l,m}^{(P_o)}(x)$ can be formed and used in the CV B-spline neural network, giving rise to

$$\hat{w} = \hat{\Psi}(x) = \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_{l,m}^{(P_o)}(x) \theta_{l,m} = \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_l^{(\Re,P_o)}(x_R) B_m^{(\Im,P_o)}(x_I) \theta_{l,m}, \quad (30)$$

where $\theta_{l,m} = \theta_{R_{l,m}} + j \cdot \theta_{I_{l,m}} \in \mathbb{C}$, $1 \leq l \leq N_R$ and $1 \leq m \leq N_I$, are the CV weights.

Consider now using the CV B-spline neural network (30) to approximate the HPA nonlinearity $\Psi(\cdot)$ over one OFDM symbol $\mathbf{x}$. Firstly, define the overall parameter vector $\boldsymbol{\theta} \in \mathbb{C}^{N_B}$, where $N_B = N_R \cdot N_I$, of the B-spline model (30) as

$$\boldsymbol{\theta} = [\theta_{1,1} \ \theta_{1,2} \ \dots \ \theta_{l,m} \ \dots \ \theta_{N_R,N_I}]^T, \quad (31)$$

and the B-spline basis function matrix $\mathbf{B} \in \mathbb{R}^{N \times N_B}$ as

$$\mathbf{B} = \begin{bmatrix} B_{1,1}^{(P_o)}(x_0) & B_{1,2}^{(P_o)}(x_0) & \dots & B_{N_R,N_I}^{(P_o)}(x_0) \\ B_{1,1}^{(P_o)}(x_1) & B_{1,2}^{(P_o)}(x_1) & \dots & B_{N_R,N_I}^{(P_o)}(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ B_{1,1}^{(P_o)}(x_{N-1}) & B_{1,2}^{(P_o)}(x_{N-1}) & \dots & B_{N_R,N_I}^{(P_o)}(x_{N-1}) \end{bmatrix}. \quad (32)$$

Then the B-spline model (30) over $\mathbf{x}$ can be represented concisely by

$$\hat{\mathbf{w}} = \mathbf{B}\boldsymbol{\theta} \quad (33)$$

where $\hat{\mathbf{w}} = [\hat{w}_0 \ \hat{w}_1 \ \dots \ \hat{w}_{N-1}]^T$ with $\hat{w}_n = \hat{\Psi}(x_n)$.

B. Identification of OFDM Hammerstein channel

The identification of OFDM Hammerstein channel involves estimating the nonlinearity $\Psi(\bullet)$, and channel impulse response coefficients h_i . Note that by applying N -point FFT to h_i we obtain H_n . Consider the estimation of θ and $\mathbf{h}$ based on a block of training sequence, e.g. the first K samples from part ($K < N$) or whole OFDM symbol period ($K = N$). The identification task can be formulated as the one that minimises the cost function

$$J(\mathbf{h}, \theta) = \frac{1}{K} \sum_{k=0}^K |e_k|^2 = \frac{1}{K} \sum_{k=0}^{K-1} |y_k - \hat{y}_k|^2, \text{ s.t. } h_0 = 1 \quad (34)$$

in which the model prediction $\hat{y}_k$ is given by

$$\hat{y}_k = \sum_{i=0}^{L_{\text{CIR}}} h_i w_{k-i} = \sum_{i=0}^{L_{\text{CIR}}} h_i \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_{l,m}^{(P_o)}(x_{k-i}) \theta_{l,m}, \quad (35)$$

where $x_{k-i} = x_{k-i+N}$, if $k < i$. Note that (36) can be viewed two different linear regression models, i.e. one is with respect to $\mathbf{h}$ when fixing θ and the other is with respect to θ given a fixed $\mathbf{h}$, each problem having a close form solution.

Specifically letting $\mathbf{e} = [e_0, \dots, e_{K-1}]^T$, and over the training data set, the system can be represented as

$$\mathbf{y} = \mathbf{P}\mathbf{h} + \mathbf{e} = \mathbf{Q}\theta + \mathbf{e} \quad (36)$$

where $\mathbf{y} = [y_0 \ y_1 \ \dots \ y_{K-1}]^T$ is the received training sample vector, and the regression matrices are

$$\mathbf{P} = \begin{bmatrix} w_0 & w_{N-1} & \cdots & w_{N-L_{\text{CIR}}} \\ \vdots & \vdots & \vdots & \vdots \\ w_k & w_{k-1} & \cdots & w_{k-L_{\text{CIR}}-1} \\ \vdots & \vdots & \vdots & \vdots \\ w_{K-1} & w_{K-2} & \cdots & w_{K-L_{\text{CIR}}-1} \end{bmatrix} \in \mathbb{C}^{K \times (L_{\text{CIR}}+1)}. \quad (37)$$

and

$$\mathbf{Q} = \begin{bmatrix} \varphi_{1,1}(0) & \cdots & \varphi_{l,m}(0) & \cdots & \varphi_{N_R, N_I}(0) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \varphi_{1,1}(k) & \cdots & \varphi_{l,m}(k) & \cdots & \varphi_{N_R, N_I}(k) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \varphi_{1,1}(K) & \cdots & \varphi_{l,m}(K) & \cdots & \varphi_{N_R, N_I}(K) \end{bmatrix} \in \mathbb{C}^{K \times N_B}. \quad (38)$$

respectively, in which $\varphi_{l,m}(k) = 1 + \sum_{i=1}^{L_{\text{CIR}}} h_i B_{l,m}^{(P_o)}(x_{k-i})$, with $x_{k-i} = x_{k-i+N}$, if $k < i$.

We adopt the following iterative procedure of alternating the LS estimation of $\mathbf{h}$ and the LS estimation of $\boldsymbol{\theta}$, which is a coordinate gradient descent algorithm [57], [58] that guarantees to converge fast to an unbiased estimate of $\mathbf{h}$ and $\boldsymbol{\theta}$ jointly in our case. Note that unlike a generic coordinate gradient descent algorithm, in our case we have the close-form solutions for both $\mathbf{h}$ and $\boldsymbol{\theta}$.

Initialisation.

Initialize $w_k = x_k$ in $\mathbf{P}$ of (38). Calculate $\mathbf{h}$ as the LS estimate given by

$$\hat{\mathbf{h}}^{(0)} = (\mathbf{P}^H \mathbf{P})^{-1} \mathbf{P}^H \mathbf{y}. \quad (39)$$

Obtain $\hat{\mathbf{h}}^{(0)}$ by normalizing $h_i \leftarrow h_i/h_0$ for $0 \leq i \leq L_{\text{CIR}} - 1$.

Alternating LS estimation. For $1 \leq \tau \leq \tau_{\text{max}}$, where τ_{max} is the maximum number of iterations, perform:

a) Fix $\mathbf{h}$ as $\hat{\mathbf{h}}^{(\tau-1)}$ in $\mathbf{Q}$ of (38). The LS estimate of $\hat{\boldsymbol{\theta}}^{(\tau)}$ is readily given by

$$\hat{\boldsymbol{\theta}}^{(\tau)} = (\mathbf{Q}^H \mathbf{Q})^{-1} \mathbf{Q}^H \mathbf{y}. \quad (40)$$

b) For $\mathbf{P}$ of (37), fix w_i as its resultant estimate from (30) using $\hat{\boldsymbol{\theta}}^{(\tau)}$. Calculate

$$\hat{\mathbf{h}}^{(\tau)} = (\mathbf{P}^H \mathbf{P})^{-1} \mathbf{P}^H \mathbf{y}. \quad (41)$$

Normalize $h_i \leftarrow h_i/h_0$ for $0 \leq i \leq L_{\text{CIR}}$.

A few iterations, i.e. a very small τ_{max} , are sufficient for the above alternating LS estimation procedure to converge to a joint unbiased estimate of $\mathbf{h}$ and $\boldsymbol{\theta}$ that is at least a local minimum solution for minimising the cost function (34).

C. Inversion of OFDM Hammerstein channel's static nonlinear function

The inverse of Hammerstein system's static nonlinear function based on B-spline neural network was introduced in [56], and this is described in below for completeness for solving (20).

The CV B-spline neural network (30) can obviously be decomposed as the following two RV B-spline neural networks

$$\hat{w}_R = \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_l^{(\Re, P_o)}(x_R) B_m^{(\Im, P_o)}(x_I) \theta_{R_{l,m}}, \quad (42)$$

$$\hat{w}_I = \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_l^{(\Re, P_o)}(x_R) B_m^{(\Im, P_o)}(x_I) \theta_{I_{l,m}}. \quad (43)$$

Given the CV Hammerstein system's static nonlinearity $\Psi(\bullet)$, we wish to compute its inverse defined by $x(k) = \Psi^{-1}(w(k))$. This task is identical to find the CV root of $w(k) = \Psi(x(k))$, given $w(k)$. In Subsection III-A, the estimate $\widehat{\Psi}(\bullet)$ for $\Psi(\bullet)$ has been obtained based on the CV B-spline neural network. We now show that $\widehat{\Psi}^{-1}(\bullet)$ can be effectively obtained. Given $\widehat{\Psi}(\bullet)$, we have

$$\widehat{w}_R(k) = \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_l^{(\Re, P_o)}(x_R(k)) B_m^{(\Im, P_o)}(x_I(k)) \widehat{\theta}_{Rl,m}, \quad (44)$$

$$\widehat{w}_I(t) = \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_l^{(\Re, P_o)}(x_R(k)) B_m^{(\Im, P_o)}(x_I(k)) \widehat{\theta}_{Il,m}, \quad (45)$$

where $\widehat{\theta}_{l,m} = \widehat{\theta}_{Rl,m} + j \cdot \widehat{\theta}_{Il,m}$. Define $\zeta(k) = w(k) - \widehat{w}(k)$ and the squared error (SE)

$$SE(k) = \zeta_R^2(k) + \zeta_I^2(k). \quad (46)$$

If $S(k) = 0$, then $x(k)$ is the CV root of $w(k) = \widehat{\Psi}(x(k))$. Thus, the task is equivalent to the one that minimises the SE (46). We propose to use the following Gauss-Newton algorithm to solve this optimisation problem with the aid of the efficient De Boor algorithm.

By denoting the iteration step with the superscript (t) and giving a random initialisation of $x^{(0)}(k)$ that satisfies $U_{\min} < x_R^{(0)}(k) < U_{\max}$ and $V_{\min} < x_I^{(0)}(k) < V_{\max}$, the iterative procedure is given by

$$\begin{aligned} \begin{bmatrix} x_R^{(t)}(k) \\ x_I^{(t)}(k) \end{bmatrix} &= \begin{bmatrix} x_R^{(t-1)}(k) \\ x_I^{(t-1)}(k) \end{bmatrix} - \eta \left((\mathbf{J}_x^{(t)})^T \mathbf{J}_x^{(t)} \right)^{-1} (\mathbf{J}_x^{(t)})^T \\ &\quad \times \begin{bmatrix} \zeta_R^{(t-1)}(k) \\ \zeta_I^{(t-1)}(k) \end{bmatrix}, \end{aligned} \quad (47)$$

where $\eta > 0$ is the step size, $\zeta^{(t)}(k) = w(k) - \widehat{w}^{(t)}(k)$ with $\widehat{w}^{(t)}(k) = \widehat{\Psi}(x^{(t)}(k))$. During the iterations, $x_R^{(t)}(k)$ is reset as $U_{\min}$ or $U_{\max}$ if it goes out the bound. Similarly $x_I^{(t)}(k)$ is reset as $V_{\min}$ or $V_{\max}$. $\mathbf{J}_x^{(t)}$ is the 2×2 Jacobian matrix given by

$$\mathbf{J}_x^{(t)} = \begin{bmatrix} \frac{\partial \zeta_R(k)}{\partial x_R(k)} & \frac{\partial \zeta_R(k)}{\partial x_I(k)} \\ \frac{\partial \zeta_I(k)}{\partial x_R(k)} & \frac{\partial \zeta_I(k)}{\partial x_I(k)} \end{bmatrix}_{|x(k)=x^{(t)}(k)}. \quad (48)$$

The entries in (48) are given by

$$\left\{ \begin{aligned} \frac{\partial \zeta_R(k)}{\partial x_R(k)} &= - \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} \frac{dB_l^{(\Re, P_o)}(x_R(k))}{dx_R(k)} B_m^{(\Im, P_o)}(x_I(k)) \hat{\theta}_{Rl,m}, \\ \frac{\partial \zeta_R(k)}{\partial x_I(k)} &= - \sum_{l=1}^{M_R} \sum_{m=1}^{N_I} B_l^{(\Re, P_o)}(x_R(k)) \frac{dB_m^{(\Im, P_o)}(x_I(k))}{dx_I(k)} \hat{\theta}_{Rl,m}, \\ \frac{\partial \zeta_I(k)}{\partial x_R(k)} &= - \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} \frac{dB_l^{(\Re, P_o)}(x_R(k))}{dx_R(k)} B_m^{(\Im, P_o)}(x_I(k)) \hat{\theta}_{Il,m}, \\ \frac{\partial \zeta_I(k)}{\partial x_I(k)} &= - \sum_{l=1}^{N_R} \sum_{m=1}^{N_I} B_l^{(\Re, P_o)}(x_R(k)) \frac{dB_m^{(\Im, P_o)}(x_I(k))}{dx_I(k)} \hat{\theta}_{Il,m}, \end{aligned} \right. \quad (49)$$

for which the De Boor algorithm, (23)–(25) and (27)–(29), can be used for their calculation efficiently. The algorithm is terminated when $SE(k) < \rho$, where ρ is a preset required precision, e.g. $\rho = 10^{-8}$, or when t reaches a predetermined maximum value. The step size η of the Gauss-Newton algorithm can be set to a relatively large value. As this is a two-dimensional problem, ten iterations are often sufficient for the Gauss-Newton algorithm to converge.

IV. SIMULATION STUDY

We consider Hammerstein OFDM Systems in which the static nonlinearity described by the HPA (8)-10) is employed using 64-QAM and $N = 2048$ subcarriers. *{ Alex, Check A Rayleigh mutlipath fading channel with an exponentially decreasing power delay profile is used. The channel spread is 10 and channel regression factor is set as 3 }??}* We used a full OFDM symbol period $K = N = 2048$ training samples in the joint estimation of the channel impulse coefficient vector $\mathbf{h}$ and the B-spline network parameter vector $\boldsymbol{\theta}$. The piecewise cubic polynomial ($P_o = 4$) was chosen as the B-spline basis function, and the number of B-spline basis functions was set to eight. The input signals were generated based on the settings of $OBO = 5$ DB and $OBO = 3$ DB respectively. We define E_b as the average power of the input signal x_n to the HPA, and $N_o = 2\sigma^2$. The empirically determined knot sequences for different HPA operating conditions are listed in Table. I. Observe from Table II that the identification of the linear subsystem in the Hammerstein channel was achieved with high accuracy for four conditions. The modeling results of B-spline CV neural network for HPA static nonlinearity are shown in 1, demonstrating that identification algorithm are successful. The BER performance of the proposed nonlinear equalizer of Hammerstein OFDM systems are plotted in Fig. 2, and the proposed nonlinear equalizer is further illustrated using Fig. 3.

Fig. 1. Comparison of the HPA's static nonlinearity $\Psi(\bullet)$ and the estimated static nonlinearity $\hat{\Psi}(\bullet)$ under (a) OBO= 5 dB, $E_b/N_o = 0$ dB; (b) OBO= 5 dB, $E_b/N_o = 10$ dB; (c) OBO= 3 dB, $E_b/N_o = 0$ dB; (d) OBO= 3 dB, $E_b/N_o = 10$ dB;

V. CONCLUSIONS

.....

REFERENCES

- [1] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, Vol. 28, no. 5, pp. 5–14, May 1990.
- [2] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs, and Broadcasting*. Chichester, UK: Wiley, 2003.
- [3] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Communications*, vol. COM-29, no. 11, pp.1715–1720, Nov. 1981.

TABLE I
EMPIRICALLY DETERMINED KNOT SEQUENCE.

OBO	$U_{\min}, V_{\min}$	$U_{\max}, V_{\max}$	Knot sequence
5 dB	-0.15	0.15	-10, -9, -0.3, -0.1, -0.05, -0.02, 0, 0.02, 0.05, 0.1, 0.3, 9, 10
3 dB	-0.3	0.3	-10, -9, -0.3, -0.15, -0.05, -0.02, 0, 0.02, 0.05, 0.15, 0.3, 9, 10

Fig. 2. The bit error rate versus channel SNR performance.

TABLE II
IDENTIFICATION RESULTS FOR $\mathbf{h}$ OF THE HAMMERSTEIN CHANNEL.

	True Parameters	Parameter estimate under			
		$E_b/N_o = 0$ dB $OBO = 5$ dB	$E_b/N_o = 10$ dB $OBO = 5$ dB	$E_b/N_o = 0$ dB $OBO = 3$ dB	$E_b/N_o = 10$ dB $OBO = 3$ dB
h_0	1	1	1	1	1
h_1	$-0.2145 - j0.1867$	$-0.2143 - j0.1872$	$-0.2144 - j0.1869$	$-0.2141 - j0.1874$	$-0.2143 - j0.1869$
h_2	$0.0399 + j0.3675$	$0.0396 + j0.3684$	$0.0399 + j0.3678$	$0.0396 + j0.3686$	$0.0399 + j0.3678$
h_3	$-0.0900 + j0.4053$	$-0.0903 + j0.4051$	$-0.0901 + j0.4052$	$-0.0902 + j0.4051$	$-0.0895 + j0.4052$
h_4	$-0.0893 + j0.1287$	$-0.0895 + j0.1265$	$-0.0894 + j0.1280$	$-0.0896 + j0.1261$	$-0.0895 + j0.1279$
h_5	$-0.1117 + j0.3035$	$-0.1131 + j0.3035$	$-0.1122 + j0.3035$	$-0.1134 + j0.3035$	$-0.1122 + j0.3035$
h_6	$-0.0766 - j0.0264$	$-0.0765 - j0.0269$	$-0.0766 - j0.0265$	$-0.0765 - j0.0269$	$-0.0766 - j0.0265$
h_7	$0.0623 - j0.0668$	$0.0621 - j0.0668$	$0.0623 - j0.0668$	$0.0621 - j0.0670$	$0.0623 - j0.0669$
h_8	$0.0282 + j0.0324$	$0.0283 + j0.0325$	$0.0282 + j0.0325$	$0.0284 + j0.0326$	$0.0282 + j0.0325$
h_9	$-0.0395 - j0.0291$	$-0.0393 - j0.0304$	$-0.0394 - j0.0295$	$-0.0391 - j0.0308$	$-0.0393 - j0.0297$

- [4] M. Honkanen and S.-G. Häggman, "New aspects on nonlinear power amplifier modeling in radio communication system simulations," in *Proc. PIMRC'97* (Helsinki, Finland), Sept. 1-4, 1997, pp. 844–848.
- [5] C. J. Clark, G. Chrisikos, M. S. Muha, A. A. Moulthrop, and C. P. Silva, "Time-domain envelope measurement technique with application to wideband power amplifier modeling," *IEEE Trans. Microwave Theory and Techniques*, vol. 46, no. 12, pp. 2531–2540, Dec. 1998.
- [6] C.-S. Choi, *et al.*, "RF impairment models 60 GHz band SYS/PHY simulation," Document IEEE 802.15-06-0477-01-003c, Nov. 2006. <https://mentor.ieee.org/802.15/dcn/06/15-06-0477-01-003c-rf-impairment-models-60ghz-band-sysphy-simulation.pdf>
- [7] V. Erceg, *et al.*, "60 GHz impairments modeling," Document IEEE 802.11-09/1213r1, Nov. 2009.
- [8] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Communications*, vol. 52, no. 1, pp. 159–165, Jan. 2004.
- [9] D. Zhou and V. E. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm," *IEEE Trans. Signal Processing*, vol. 55, no. 1, pp. 120–133, Jan. 2007.
- [10] M.-C. Chiu, C.-H. Zeng, and M.-C. Liu, "Predistorter based on frequency domain estimation for compensation of nonlinear distortion in OFDM systems," *IEEE Trans. Vehicular Technology*, vol. 57, no. 2, pp. 882–892, March 2008.
- [11] S. Choi, E.-R. Jeong, and Y. H. Lee, "Adaptive predistortion with direct learning based on piecewise linear approximation of amplifier nonlinearity," *IEEE J. Selected Topics in Signal Processing*, vol. 3, no. 3, pp. 397–404, June 2009.
- [12] V. P. G. Jiménez, Y. Jabrane, A. G. Armada, and B. Ait Es Said, "High power amplifier pre-distorter based on neural-fuzzy systems for OFDM signals," *IEEE Trans. Broadcasting*, vol. 57, no. 1, pp. 149–158, March 2011.
- [13] S. Chen, "An efficient predistorter design for compensating nonlinear memory high power amplifier," *IEEE Trans. Broadcasting*, vol. 57, no. 4, pp. 856–865, Dec. 2011.
- [14] S. Chen, X. Hong, Y. Gong, and C. J. Harris, "Digital predistorter design using B-spline neural network and inverse of De Boor algorithm," *IEEE Trans. Circuits and Systems I*, vol. 60, no. 6, pp. 1584–1594, June 2013.

- [15] X. Hong, S. Chen, and C. J. Harris, "Complex-valued B-spline neural networks for modeling and inverse of Wiener systems," Chapter 9 in: A. Hirose, ed. *Complex-Valued Neural Networks: Advances and Applications*. Hoboken, NJ: IEEE and Wiley, 2012, pp. 209–233.
- [16] X. Hong and S. Chen, "Modeling of complex-valued Wiener systems using B-spline neural network," *IEEE Trans. Neural Networks*, vol 22, no. 5, pp. 818–825, May 2011.
- [17] C. De Boor, *A Practical Guide to Splines*. New York: Springer Verlag, 1978.
- [18] C. J. Harris, X. Hong, and Q. Gan, *Adaptive Modelling, Estimation and Fusion from Data: A Neurofuzzy Approach*. Berlin: Springer-Verlag, 2002.
- [19] E. W. Bai, "An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol. 34, no. 3, pp. 333–338, March 1998.

Fig. 3. The illustration of nonlinear equalizer of Hammerstein OFDM systems for one sample OFDM symbol based on $OBO = 5\text{dB}$, $E_b/N_o = 10\text{ dB}$.

- [20] S. Chen, S. McLaughlin, and B. Mulgrew, "Complex-valued radial basis function network, Part I: Network architecture and learning algorithms," *Signal Processing*, vol. 35, no. 1, pp. 19–31, Jan. 1994.
- [21] S. Chen, S. McLaughlin, and B. Mulgrew, "Complex-valued radial basis function network, Part II: Application to digital communications channel equalisation," *Signal Processing*, vol. 36, no. 2, pp. 175–188, March 1994.
- [22] A. Uncini, L. Vecchi, P. Campolucci, and F. Piazza, "Complex valued neural networks with adaptive spline activation function for digital radio links nonlinear equalization," *IEEE Trans. Signal Processing*, vol. 47, no. 2, pp. 505–514, Feb. 1999.
- [23] T. Kim and T. Adali, "Approximation by fully complex multilayer perceptrons," *Neural Computation*, vol. 15, no. 7, pp. 1641–1666, July 2003.
- [24] C.-C. Yang and N. K. Bose, "Landmine detection and classification with complex-valued hybrid neural network using scattering parameters dataset," *IEEE Trans. Neural Networks*, vol. 16, no. 3, pp. 743–753, May 2005.
- [25] M. B. Li, G. B. Guang, P. Saratchandran, and N. Sundararajan, "Fully complex extreme learning machine," *Neurocomputing*, vol. 68, pp. 306–314, Oct. 2005.
- [26] A. Hirose, *Complex Valued Neural Networks*. Berlin: Springer-Verlag, 2006.
- [27] S. Chen, X. Hong, C. J. Harris, and L. Hanzo, "Fully complex-valued radial basis function networks: Orthogonal least squares regression and classification," *Neurocomputing*, vol. 71, no. 16–18, pp. 3421–3433, Oct. 2008.
- [28] T. Nitta, Ed., *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*. New York: Information Science Reference, 2009.
- [29] A. S. Gangal, P. K. Kalra, and D. S. Chauhan, "Inversion of complex valued neural networks using complex back-propagation algorithm," *Int. J. Mathematics and Computers in Simulation*, vol. 3, no. 1, pp. 1–8, 2009.
- [30] M. Kobayashi, "Exceptional reducibility of complex-valued neural networks," *IEEE Trans. Neural Networks*, vol. 21, no. 7, pp. 1060–1072, July 2010.
- [31] A. Hirose, ed., *Complex-Valued Neural Networks: Advances and Applications*. Hoboken, NJ: John Wiley & Sons, 2012 (in press).
- [32] S. A. Billings, "Identification of nonlinear systems – a survey," *IEE Proc. D*, vol. 127, no. 6, pp. 272–285, Nov. 1980.
- [33] I. W. Hunter and M. J. Korenberg, "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models," *Biological Cybernetics*, vol. 55, nos. 2–3, pp. 135–144, 1986.
- [34] Y. Zhu, "Estimation of an N-L-N Hammerstein-Wiener model," *Automatica*, vol. 38, no. 9, pp. 1607–1614, Sept. 2002.
- [35] J. Schoukens, J. G. Nemeth, P. Crama, Y. Rolain, and R. Pintelon, "Fast approximate identification of nonlinear systems," *Automatica*, vol. 39, no. 7, pp. 1267–1274, July 2003.
- [36] K. Hsu, T. Vincent, and K. Poolla, "A kernel based approach to structured nonlinear system identification part I: algorithms," in *Proc. 14th IFAC Symp. System Identification* (Newcastle, Australia), March 29–31, 2006, 6 pages.
- [37] K. Hsu, T. Vincent, and K. Poolla, "A kernel based approach to structured nonlinear system identification part II: convergence and consistency," in *Proc. 14th IFAC Symp. System Identification* (Newcastle, Australia), March 29–31, 2006, 6 pages.
- [38] W. Greblicki, "Nonparametric identification of Wiener systems," *IEEE Trans. Information Theory*, vol. 38, no. 5, pp. 1487–1493, Sept. 1992.
- [39] A. Kalafatis, N. Arifin, L. Wang, and W. R. Cluett, "A new approach to the identification of pH processes based on the Wiener model," *Chemical Engineering Science*, vol. 50, no. 23, pp. 3693–3701, Dec. 1995.
- [40] A. D. Kalafatis, L. Wang, and W. R. Cluett, "Identification of Wiener-type nonlinear systems in a noisy environment," *Int. J. Control*, vol. 66, no. 7, pp. 923–941, 1997.

- [41] Y. Zhu, "Distillation column identification for control using Wiener model," in *Proc. 1999 American Control Conference* (San Diego, USA), June 2-4, 1999, pp. 3462–3466.
- [42] J. C. Gomez, A. Jutan, and E. Baeyens, "Wiener model identification and predictive control of a pH neutralisation process," *IEE Proc. Control Theory and Applications*, vol. 151, no. 3, pp. 329–338, May 2004.
- [43] I. Skrjanc, S. Blazic, and O. E. Agamennoni, "Interval fuzzy modeling applied to Wiener models with uncertainties," *IEEE Trans. Systems, Man and Cybernetics, Part B*, vol. 35, no. 5, pp. 1092–1095, Oct. 2005.
- [44] A. Hagenblad, L. Ljung, and A. Wills, "Maximum likelihood identification of Wiener models," *Automatica*, vol. 44, no. 11, pp. 2697–2705, Nov. 2008.
- [45] S. A. Billings and S. Y. Fakhouri, "Non-linear system identification using the Hammerstein model," *Int. J. Systems Science*, vol. 10, no. 5, pp. 567–578, May 1979.
- [46] P. Stoica and T. Söderström, "Instrumental-variable methods for identification of Hammerstein systems," *Int. J. Control*, vol. 35, no. 3, pp. 459–476, 1982.
- [47] A. Balestrino, A. Landi, M. Ould-Zmirli, and L. Sani, "Automatic nonlinear auto-tuning method for Hammerstein modelling of electrical drives," *IEEE Trans. Industrial Electronics*, vol. 48, no. 3, pp. 645–655, June 2001.
- [48] J. Turunen, J. T. Tanttu, and P. Loula, "Hammerstein model for speech coding," *EURASIP J. Applied Signal Processing*, vol. 2003, pp. 1238–1249, Jan. 2003.
- [49] S. W. Su, L. Wang, B. G. Celler, A. V. Savkin, and Y. Guo, "Identification and control for heart rate regulation during treadmill exercise," *IEEE Trans. Biomedical Engineering*, vol. 54, no. 7, pp. 1238–1246, July 2007.
- [50] X. Hong and R. J. Mitchell, "Hammerstein model identification algorithm using Bezier-Bernstein approximation," *IET Control Theory and Applications*, vol. 1, no. 1, pp. 1149–1159, April 2007.
- [51] X. Hong, R. J. Mitchell, and S. Chen, "Modeling and control of Hammerstein system using B-spline approximation and the inverse of De Boor algorithm," *Int. J. Systems Science*, vol. 43, no. 10, pp. 1976–1984, Oct. 2012.
- [52] G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide*. Fourth Edition. Boston: Academic Press, 1996.
- [53] T. Kavli, "ASMOD – an algorithm for adaptive spline modelling of observation data," *Int. J. Control*, vol. 58, no. 4, pp. 947–967, 1993.
- [54] M. Brown and C. J. Harris, *Neurofuzzy Adaptive Modelling and Control*. Hemel Hempstead: Prentice Hall, 1994.
- [55] Y. Yang, L. Guo, and H. Wang, "Adaptive statistic tracking control based on two-step neural networks with time delays," *IEEE Trans. Neural Networks*, vol. 20, no. 3, pp. 420–429, March 2009.
- [56] X. Hong, S. Chen, and C. J. Harris, "Complex-valued B-spline neural networks for modelling and inverse of Wiener systems," chapter 9 in: A. Hirose, ed., *Complex-Valued Neural Networks: Advances and Applications*. Hoboken, NJ: John Wiley & Sons, 2012 (in press).
- [57] R. J. Hathaway and J. C. Bezdek, "Grouped coordinate minimization using Newton's method for inexact minimization in one vector coordinate," *J. Optimization Theory and Applications*, vol. 71, no. 3, pp. 503–516, Dec. 1991.
- [58] Z. Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optimization Theory and Applications*, vol. 72, no. 1, pp. 7–35, Jan. 1991.
- [59] B. Igelnik, "Kolmogorov's spline complex network and adaptive dynamic modeling of data," in: T. Nitta, Ed., *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*. New York: Information Science Reference: 2009, pp. 56–78.
- [60] M. Scarpiniti, D. Vigliano, R. Parisi, and A. Uncinis, "Flexible blind signal separation in the complex domain," in: T. Nitta, Ed., *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters*. New York: Information Science Reference, 2009, pp. 284–323.

- [61] S. Haykin, *Adaptive Filter Theory* (2nd Edition). Englewood, NJ: Prentice Hall, 1991.
- [62] L. Hanzo, S. X. Ng, T. Keller, and W. Webb, *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded, Turbo-Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*. Chichester, UK: John Wiley, 2004.
- [63] J. G. Proakis, *Digital communications* (4th edition). McGraw-Hill, 2000.