Development of Nonlinear Lattice-Hammerstein Filters for Gaussian Signals

M. Eghtedari* and M.-H. Kahaei*

Abstract: In this paper, the nonlinear lattice-Hammerstein filter and its properties are derived. It is shown that the error signals are orthogonal to the input signal and also backward errors of different stages are orthogonal to each other. Numerical results confirm all the theoretical properties of the lattice-Hammerstein structure.

Keywords: Hammerstein Series, Lattice Structure, Nonlinear Filter.

1 Introduction

Linear filters have widely been applied to different applications due to simplicity of implementation. From conceptual and implementation points of view, the impulse response of such systems completely characterizes the system behavior. Nevertheless, there are many practical situations where the use of nonlinear structure is inevitable and incorporation of linear structures may lead to inaccurate results. As such, data transmission lines, channel equalizers, echo cancellers, system identifiers, saturated power amplifiers, OFDM systems, and signal detectors may be mentioned [1] to [4]. This accordingly has motivated more research on developing new nonlinear techniques such as neural networks, order-statistics filters, homomorphic filters, and polynomial filters [5].

The latter technique embeds those nonlinear systems whose input-output signals are related through a truncated series expansion. In this regard, the Volterra series expansion has received specific attention in the literature due to its great potential in nonlinear modeling scenarios [2]. However, the major difficulty with these methods is their massive computational burden which may make their implementation impractical. More specifically, the main limitation with the Volterra expression is the large number of coefficients involved in computations. Moreover, it is normally difficult to analyze Volterra filters.

An alternative nonlinear expression is the Hammerstein series with a much simpler structure for nonlinear modeling [1] and [15] to [17]. On the other hand, among linear structures, the lattice filter has thoroughly been studied in the literature. Individual properties of this filter such as modularity, ease of computations, providing a Gram-Schmidt type of orthogonal signals, and etc., has made it an appealing structure in many applications [6] and [9] and [14]. To exploit the lattice properties in nonlinear structures, a lattice-Volterra filter has already been addressed [10] and [11]. This filter is computationally extremely complicated.

The aim of this paper is twofold: first, to present a practical substitution for the lattice-Volterra filter with much less computations; secondly, to develop a new nonlinear structure whose properties are analytically tractable and also benefits of the lattice-type filter properties. In this paper, the Lattice-Hammerstein filter is developed and its properties are theoretically investigated.

The paper is organized as follows. In Section 2, Hammerstein prediction-error filters are derived. The lattice-Hammerstein filter and its properties are presented in Section 3. Simulation results are presented in Section 4 and Section 5 concludes the paper.

2 Hammerstein Prediction-Error Filters

The truncated Hammerstein series expansion is given by [1]

\[ y(n) = \sum_{i=0}^{M-1} h_i(i)x(n-i) + \sum_{i=0}^{M-1} h_2(i)x^2(n-i) + \cdots + \sum_{i=0}^{M-1} h_P(i)x^P(n-i) \]

(1)

where \( x(n) \) and \( y(n) \) respectively show the input and output signals, \( M \) expresses the number of system memories, \( P \) presents the nonlinearity degree, and \( h_i(*) \) denotes polynomial coefficients. Then, the input matrix is given based on the input vector as:

\[ X_M(n) = [x(n), x(n-1), \cdots, x(n-M+1)] \]

(2)
where

\[ x(n-i) = \left[ x(n-i), x^2(n-i), \ldots, x^p(n-i) \right]^T \quad i = 0 \sim M - 1 \quad (3) \]

The error vector of a Hammerstein forward prediction-error filter (see Fig. 1) is defined based on Eq. (3) as

\[ f_m(n) = x(n) - \hat{x}_m(n) = x(n) - \sum_{i=1}^{m} A_{m,i} x(n-i) \quad (4) \]

where \( \hat{x}_m(n) \) shows the forward prediction signal and

\[ A_{m,i}(n) = \begin{bmatrix} a_{m,i}^{11} & a_{m,i}^{12} & \cdots & a_{m,i}^{1P} \\ a_{m,i}^{21} & a_{m,i}^{22} & \cdots & a_{m,i}^{2P} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,i}^{p1} & a_{m,i}^{p2} & \cdots & a_{m,i}^{pp} \end{bmatrix} \quad (5) \]

with \( a_{m,i}^{kl} \ k, l = 1, 2, \ldots, P \) and \( i = 1, 2, \ldots, m \) being the elements of the forward prediction-error coefficient matrix.

To compute the optimal coefficients in \textit{mean-square error} sense, the cost function is defined for forward errors as

\[ P_m^f = E \left[ f_m^T(n)f_m(n) \right] \quad (6) \]

Setting the partial derivatives of \( P_m^f \) with respect to \( a_{m,i}^{kl} \) to zero, the normal equation is obtained as

\[ A_m^T = R^{-1}P^f \quad (7) \]

with

\[ A_m = \begin{bmatrix} a_{m,1}^1 & a_{m,2}^1 & \cdots & a_{m,n}^1 \\ a_{m,1}^2 & a_{m,2}^2 & \cdots & a_{m,n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}^P & a_{m,2}^P & \cdots & a_{m,n}^P \end{bmatrix}_{mp \times mp} \]

\[ R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1P} \\ R_{21} & R_{22} & \cdots & R_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ R_{p1} & R_{p2} & \cdots & R_{pp} \end{bmatrix}_{mp \times mp} \quad (8) \]

\[ P = \begin{bmatrix} P_{11}^f & P_{12}^f & \cdots & P_{1P}^f \\ P_{21}^f & P_{22}^f & \cdots & P_{2P}^f \\ \vdots & \vdots & \ddots & \vdots \\ P_{p1}^f & P_{p2}^f & \cdots & P_{pp}^f \end{bmatrix}_{mp \times mp} \]

where we call \( R \) and \( P^f \) as the \textit{combined} correlation and cross-correlation matrices whose elements are respectively given by

\[ p_{i,j}^f = [r_{i1}(0), r_{i1}(1), \ldots, r_{i1}(m)]^T, \]

\[ R_{il} = \begin{bmatrix} r_{i1}(0) & \cdots & r_{i1}(m-1) \\ r_{i1}(-1) & \cdots & r_{i1}(m-2) \\ \vdots & \ddots & \vdots \\ r_{i1}(-m+1) & \cdots & r_{i1}(0) \end{bmatrix}_{m \times m} \quad (9) \]

\[ a_{il}^m = \begin{bmatrix} a_{i1}^{11} & a_{i1}^{12} & \cdots & a_{i1}^{1P} \\ a_{i1}^{21} & a_{i1}^{22} & \cdots & a_{i1}^{2P} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1}^{p1} & a_{i1}^{p2} & \cdots & a_{i1}^{pp} \end{bmatrix}_{mp} \]

\[ x(n) \]

\[ x(n-1) \]

\[ x(n-2) \]

\[ x(n-m) \]

\[ \hat{x}_m(n) \]

\[ f_m(n) \]

Fig. 1 A Hammerstein forward prediction-error filter with the nonlinear degree of \( P \) and \( m \) memories.
with \( r_k(i) = \mathbb{E}[x(n)x(n-i)] \) denoting the correlation function and assuming that \( x(n) \) is stationary. Applying the same approach, the normal equation for backward-error coefficients is obtained as

\[
G_m^T = R^{-1}P^b
\]  
(10)

where

\[
p^b = \begin{bmatrix} p_{b1}^b & p_{b2}^b & \cdots & p_{bP}^b \\
p_{b12}^b & p_{b22}^b & \cdots & p_{bP2}^b \\
\vdots & \vdots & \ddots & \vdots \\
p_{b1P}^b & p_{b2P}^b & \cdots & p_{bPP}^b 
\end{bmatrix},
\]

and

\[
g^m_i = \begin{bmatrix} g_{m1}^i & g_{m2}^i & \cdots & g_{mP}^i \\
g_{m12}^i & g_{m22}^i & \cdots & g_{mP2}^i \\
\vdots & \vdots & \ddots & \vdots \\
g_{m1P}^i & g_{m2P}^i & \cdots & g_{mPP}^i 
\end{bmatrix}
\]  
(11)

while the error vector of a Hammerstein backward prediction-error filter (see Fig. 2), \( b_m(n) = [b_{m,1}(n), b_{m,2}(n), \ldots, b_{m,P}(n)]^T \) with \( m \) memories is obtained by

\[
b_m(n) = x(n-m) - \sum_{i=1}^{m} G_{m,i} x(n-i+1) = x(n-m) - \hat{x}_m^b(n)
\]  
(12)

where the backward coefficients matrix is defined as

\[
G_{m,i} = \begin{bmatrix} g_{m1}^{i1} & g_{m1}^{i2} & \cdots & g_{m1}^{iP} \\
g_{m2}^{i1} & g_{m2}^{i2} & \cdots & g_{m2}^{iP} \\
\vdots & \vdots & \ddots & \vdots \\
g_{mP}^{i1} & g_{mP}^{i2} & \cdots & g_{mP}^{iP} 
\end{bmatrix}
\]  
(13)

We derive the lattice-Hammerstein structure using the above forward and backward prediction-error filters for Gaussian inputs for which the higher-order statistics are expressed in terms of second-order statistics as [8]

\[
E\{x_i x_{i+1} \ldots x_N\} = \begin{cases} 0; & \text{N odd} \\ \sum_{i=1}^{N} E\{x_i x_j\} ; & \text{N even} \end{cases}
\]  
(14)

As a result, we can write for the autocorrelation function: \( r_k(i) = r_k(-i) = r_k(i) \) for \( k, l = 1 \rightarrow P \) and correspondingly for (7) and (10), \( R = R^T, A_m = A_m^T, G_m = G_m^T \).

By expansion of Eq. (7) and Eq. (10), one can easily obtain the relationship between forward and backward predictor-error coefficients as

\[
\begin{cases}
a_{m,i}^{n} = g_{m,i}^{n} + g_{m,i}^{n-1} i = 1 \sim m \\
g_{m,i}^{n} = a_{m,i}^{n} \quad \quad \quad \quad i = 1 \sim m
\end{cases}
\]  
(15)

In the sequel, we investigate some properties of forward and backward-error signals. To elaborate, by defining the backward error power as \( P_b^m = E\{b_m(n)^T\} \) and differentiating it with respect to the elements of Eq. (13) and doing necessary manipulations, we get

\[
E\{x^b(n-i+1) x^b(n-i+1), \ldots, x^b(n-i+1) b_m(n)\} = 0
\]  
(16)

for \( k_1, k_2, \ldots, k_p = 1 \sim P, i = 1 \sim m \)

which for \( k_1 = 1 \), \( k_2 = 2 \), ..., \( k_p = P \), reduces to Property 1 of the lattice-Hammerstein filter as

\[
E\{x(n-i+1) x(n-i+1), \ldots, x(n-i+1) b_m(n)\} = 0
\]  
(17)

This means that backward errors are orthogonal to the input signal. Applying the same procedure to the forward prediction-errors, one can similarly show Property 2 as

\[
E\{x^f(n-i) f_m(n)\} = 0 \quad i = 1 \sim m
\]  
(18)

Moreover, using Eq. (12) and incorporating Eq. (17) results in the important orthogonal property of backward errors (Property 3) as

\[
E\{b_{m_1}^T(n) b_{m_2}(n)\} = E\{x(n-m_1) - \sum_{i=1}^{m_1} G_{m_1,i} x(n-i+1)\}^T b_{m_2}(n) = 0
\]  
(19)

for \( m_1 \neq m_2 \) which expresses that backward errors of different stages \( (m_1 \neq m_2) \) are orthogonal to each other. These results are similar to those reported for linear lattice filters.
3 Derivation of Lattice-Hammerstein Filters

Analogous to linear lattice filters, we combine the forward and backward prediction errors to develop the lattice-Hammerstein structure shown in Fig. 3 [13]. Each module of this filter presented in Fig. 3(b) contains two input vectors from the previous module, two output vectors, and two lattice-Hammerstein coefficient matrices. The modularity of the resulting filter is in general similar to that of the linear lattice filter except that here the coefficient matrices are composed of different components based on Hammerstein series coefficients and also the inputs of the first module is a combination of the input signal \( x(n) \).

To derive these coefficients, we start by writing Eq. (4) for the \((m+1)\)-th stage as

\[
f_{m+1}(n) = \sum_{i=1}^{m} A_{m,i} x(n-i) + K_{m+1} b_m(n-1)
\]  

(20)

where \( A_{m,i} x(n-i) = A_{m+1,i} + A_{m+1,m} G_{m,i} \) and the \((m+1)\)-th forward prediction-error coefficient \( A_{m+1,m+1} \) has been taken out. In vector forms, (20) can be rewritten as

\[
f_{m+1}(n) = x(n) - W^T Z(n)
\]  

(21)

where

\[
Z(n) = [x(n-1), x(n-2), ..., x(n-m), b_m(n-1)]
\]  

(22)

\[
W = [A'_{m,1}, A'_{m,2}, ..., A'_{m,m}, K'_{m+1}]
\]  

Fig. 2 A Hammerstein backward prediction-error filter with the nonlinear degree of P and m memories.

Fig. 3 Block diagram of (a) a lattice-Hammerstein filter with the nonlinear degree of P and M memories, (b) the i-th lattice-Hammerstein module.
The weight vector $\mathbf{W}$ which minimizes $\mathbf{f}_m(n)$ in the mean-square error sense is then obtained as

$$\mathbf{R}_w\mathbf{W} = \mathbf{P}_w,$$  \hspace{1cm} (23)

where $\mathbf{R}_w = \mathbb{E}[\mathbf{Z}(n)\mathbf{Z}^T(n)]$ and $\mathbf{P}_w = \mathbb{E}[\mathbf{x}^T(n)\mathbf{Z}(n)]$. Substituting Eq. (22) in Eq. (23), we obtain the error vector of a Hammerstein forward prediction-error filter with $m+1$ memories as a function of the same filter with $m$ memories as

$$\mathbf{f}_{m+1}(n) = \mathbf{f}_m(n) - \mathbf{K}^f_{m+1}\mathbf{b}_m(n-1)$$  \hspace{1cm} (24)

and also,

$$\mathbf{R}^b_{m}(n-1)\mathbf{K}^f_{m+1} = \mathbb{E}[\mathbf{b}_m(n-1)x^T(n)]$$

$$= \mathbb{E}[\mathbf{b}_m(n-1)\mathbf{f}^T_m(n)]$$  \hspace{1cm} (25)

where $\mathbf{R}^b_{m}(n-1) = \mathbb{E}[\mathbf{b}_m(n-1)\mathbf{b}_m^T(n-1)]$ is the backward error matrix. Note that Eq. (25) is obtained using Eq. (4) and incorporating the orthogonal property of backward prediction errors to the input signal as proved in Eq. (17).

In the same manner, we can express the error vector of a Hammerstein backward prediction-error filter with $m+1$ memories as

$$\mathbf{b}_{m+1}(n) = \mathbf{b}_m(n-1) - \mathbf{K}^b_{m+1}\mathbf{f}_m(n)$$  \hspace{1cm} (26)

and

$$\mathbf{R}^f_{m}(n)\mathbf{K}^b_{m+1} = \mathbb{E}[\mathbf{f}_m(n)x^T(n-m-1)]$$

$$= \mathbb{E}[\mathbf{f}_m(n)\mathbf{b}^T_m(n-1)]$$  \hspace{1cm} (27)

where $\mathbf{R}^f_{m}(n) = \mathbb{E}[\mathbf{f}_m(n)f^T_m(n)]$. Equations (24) and (26) present the lattice-Hammerstein relationships between, forward errors, backward errors, and the corresponding lattice-Hammerstein coefficients in accordance with Fig. 3, while $\mathbf{f}_0(n) = \mathbf{b}_0(n) = \mathbf{x}(n)$. Moreover, using Eq. (25) and Eq. (27), the filter coefficients are computed based on the error signal matrices. Note that $\mathbf{K}^f_{m+1}$ and $\mathbf{K}^b_{m+1}$ are not in general the same.

The relationship between forward and backward error powers, i.e., $\mathbf{R}^f_{m}(n)$ and $\mathbf{R}^b_{m}(n-1)$ is first extended from Eq. (12) and Eq. (4) as assuming that the autocorrelation matrix $\mathbf{R}(i)$ of the input signal given by

$$\mathbf{R}(i) = \mathbb{E}[\mathbf{x}(n)x^T(n-i)]$$

$$= \begin{bmatrix} r_1(i) & r_2(i) & \cdots & r_n(i) \\ r_2(i) & r_2(i) & \cdots & r_n(i) \\ \vdots & \vdots & \ddots & \vdots \\ r_n(i) & r_n(i) & \cdots & r_n(i) \end{bmatrix}$$  \hspace{1cm} (29)

is symmetric for stationary and Gaussian input signals. Changing $i$ to $m+1-i$ in Eq. (28) and considering Eq. (15), the important relation between the power of forward and backward matrices are shown to be equal as

$$\mathbf{R}^f_{m}(n-1) = \mathbf{R}^b_{m}(n)$$  \hspace{1cm} (30)

Noting that the trace of Eq. (30) is equivalent to the power of backward and forward errors, i.e., $P^b_m = \mathbb{E}[\mathbf{b}_m(n)^2]$ and $P^f_m = \mathbb{E}[\mathbf{f}_m(n)^2]$, we can easily see that these powers are equal (Property 4):

$$P^b_m = \text{tr}(\mathbf{R}^b_{m}(n-1)) = \text{tr}(\mathbf{R}^f_{m}(n)) = P^f_m$$ \hspace{1cm} (31)

4 Simulation Results

Using computer simulations, the properties of the lattice-Hammerstein filter derived in previous parts are inspected. The input is a colored Gaussian signal generated by an FIR filter defined by $h=[0.9045 \ 0.7 \ 0.9045]$ whose input is zero-mean white Gaussian noise. The results are averaged over 100 independent trials. The number of input samples is 1000. In the first experiment, the noise variance is 0.0248, the degree of nonlinearity of Hammerstein is $P=2$, and the number of stages (memories) is $M=10$. To verify equations Eq. (31), Eq. (17), Eq. (18), and Eq. (19), power of forward and backward errors are depicted in Figs 4 (a), (b) and (c). In agreement with Eq. (19), Fig. 4 (a) shows that backward errors of different stages are orthogonal, and thus, nonzero at the same stages and zero elsewhere. Also, from Fig. 5 (b), it is seen that in accordance to Eq. (31), the power of forward and backward errors of similar stages are identical. Moreover, Fig. 5 (c) shows that forward/backward errors are orthogonal to the input signal as proved in Eq. (17) and Eq. (18).
In the second experiment, the effect of nonlinearity on the derived results is investigated. This is carried out by changing \( P \) from 2 to 6. As shown in Fig. 5 (a), forward and backward errors powers are again identical. Moreover, if we compare Fig. 4 (b) with Fig. 5 (b), it is observed that the latter powers are the same for different degrees of nonlinearities, \( i.e., P=2 \ & 6 \). This may be justified by noting the fact that in both experiments the same input signal is applied, and therefore, higher degrees of nonlinearity are not effectively involved in the signal modeling process. This has been led to very close error powers.

**Fig. 4** Properties of lattice-Hammerstein coefficients for \( P=2, M=10 \), and variance of 0.0248, (a) orthogonal property of backward errors, (b) power of forward & backward errors, and (c) orthogonal property of forward/backward errors to the input signal.
In the third experiment, the role of the number of memories, \( M \) (stages) and the input energy are studied. In this case, we have \( P=2 \), and \( M \) changes from 10 to 7. Figs 6 (a), (b) and (c) shows the results for the input variance of 0.0248. Fig. 6 (a) once more confirms Eq. (31). Also, due to the orthogonal property of backward errors shown in Fig. 6 (c), as proved in Eq. (19), each lattice module acts separately [17] in the sense that the outputs of each stage are only determined by the coefficients and inputs of the same stage.

![Graph](attachment://graph1.jpg)

**Fig. 5** Properties of lattice-Hammerstein coefficients for \( P=6, M=10 \), and variance of 0.0248, (a) power of forward & backward errors and (b) orthogonal property of backward errors.

To see the effect of the input energy on the filter performance, the noise variance is increased to 2.48. It is seen that the previous properties observed in Figs 6 (b) and (d) still remain. From Figs. 6 (b) and (d) compared to Figs 6 (a) and (c), we can see that, as expected, the errors energies in accordance to inputs energy have been increased.

![Graph](attachment://graph2.jpg)

To more investigate the correlation effect on the coefficients, in the forth experiment, we have generated a less correlated signal with a variance of 2.48, while the other conditions are similar to those of the previous experiment. Comparing Figs 7 (a) and (b) with Figs 6 (b) and (d) previous results are confirmed.
4 Conclusions
The nonlinear lattice-Hammerstein filter was developed. In this regard, the forward and backward Hammerstein prediction-error filters and the corresponding relationships were analytically derived. Various properties of the lattice-Hammerstein filter were proved. It was shown that the powers of forward and backward error signals are identical and orthogonal to the input signal and backward errors of different stages are orthogonal to each other. All the theoretical results were verified using extensive computer simulations. The computational burden of the proposed filter is much less than that of the lattice-Volterra filter. This makes the lattice-Hammerstein filter a reasonable candidate for nonlinear modeling problems.
References


Mostafa Eghtedari received the B.Sc. degree from Isfahan University of Technology, Isfahan, Iran, in 2002, in Electrical Engineering and the M.Sc. degree from Iran University of Science and Technology, Tehran, Iran, in 2005, in telecommunication engineering. His research interests are in the areas of digital signal processing and adaptive filtering.

Mohammad-Hossein Kahaei received the B.Sc. degree from Isfahan University of Technology, Isfahan, Iran, in 1986, the M.Sc. degree from the University of the Ryukyus, Okinawa, Japan, in 1994, and the Ph.D. degree in signal processing from the School of Electrical and Electronic Systems Engineering, Queensland University of Technology, Brisbane, Australia, in 1998. He joined the Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran, in 1999. His research interests are signal processing with primary emphasis on adaptive filters theory, estimation, tracking, and interference cancellation.