Technion - Israel Institute of Technology
Department of Electrical Engineering

Signal and Image Processing Laboratory

Parameter Identification of a Class of Nonlinear Systems for Robust Audio Watermarking Verification

Shay Mizrachi and David Malah

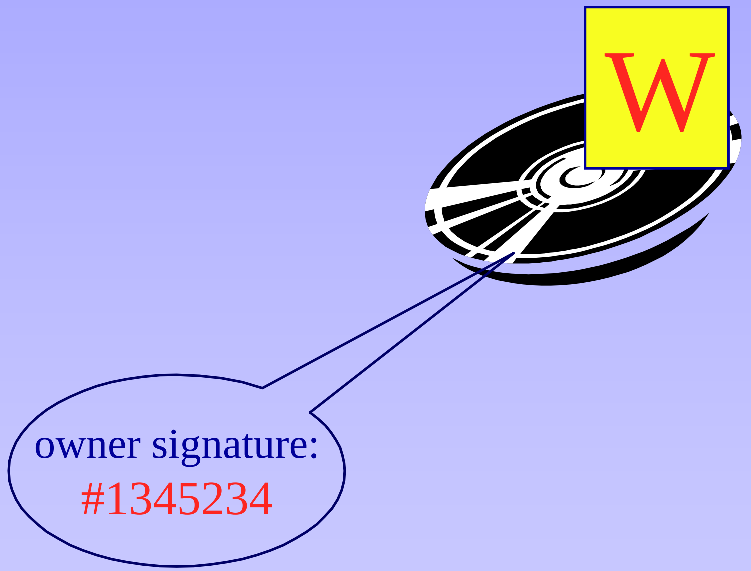
The 22nd convention of IEEE Israel section – 1/12/2002

Outline

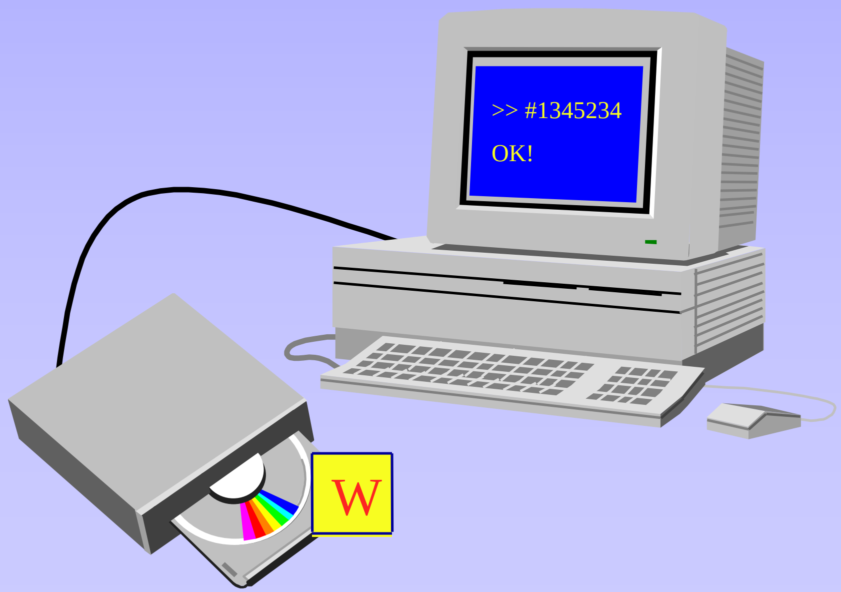
- Digital watermarking
  - Embedding and detection of a digital watermarking.
  - Possible attacks and the modification needed in the detection mechanism.
- Estimation methods for nonlinear system
  - Volterra method
  - piecewise linear approximation
  - Two proposed method
- Results and Conclusions.

Digital Watermarking
For Copyright Protection

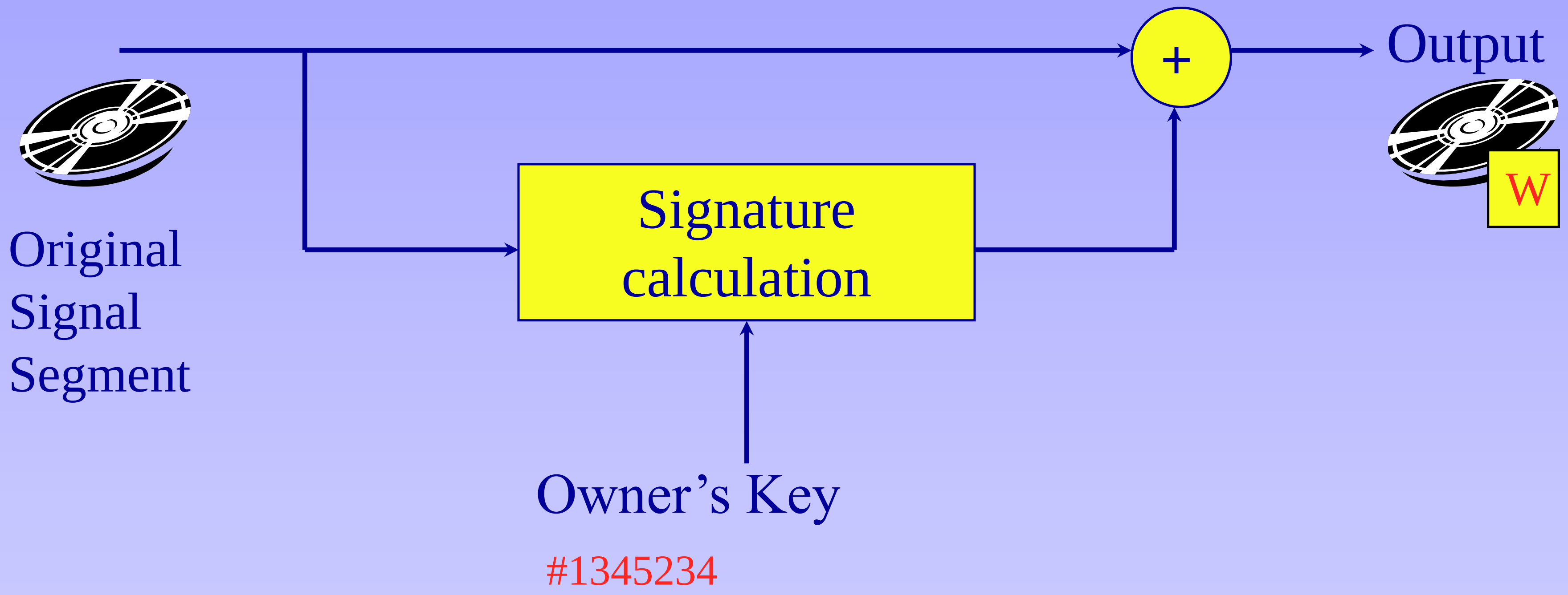
Signature embedding



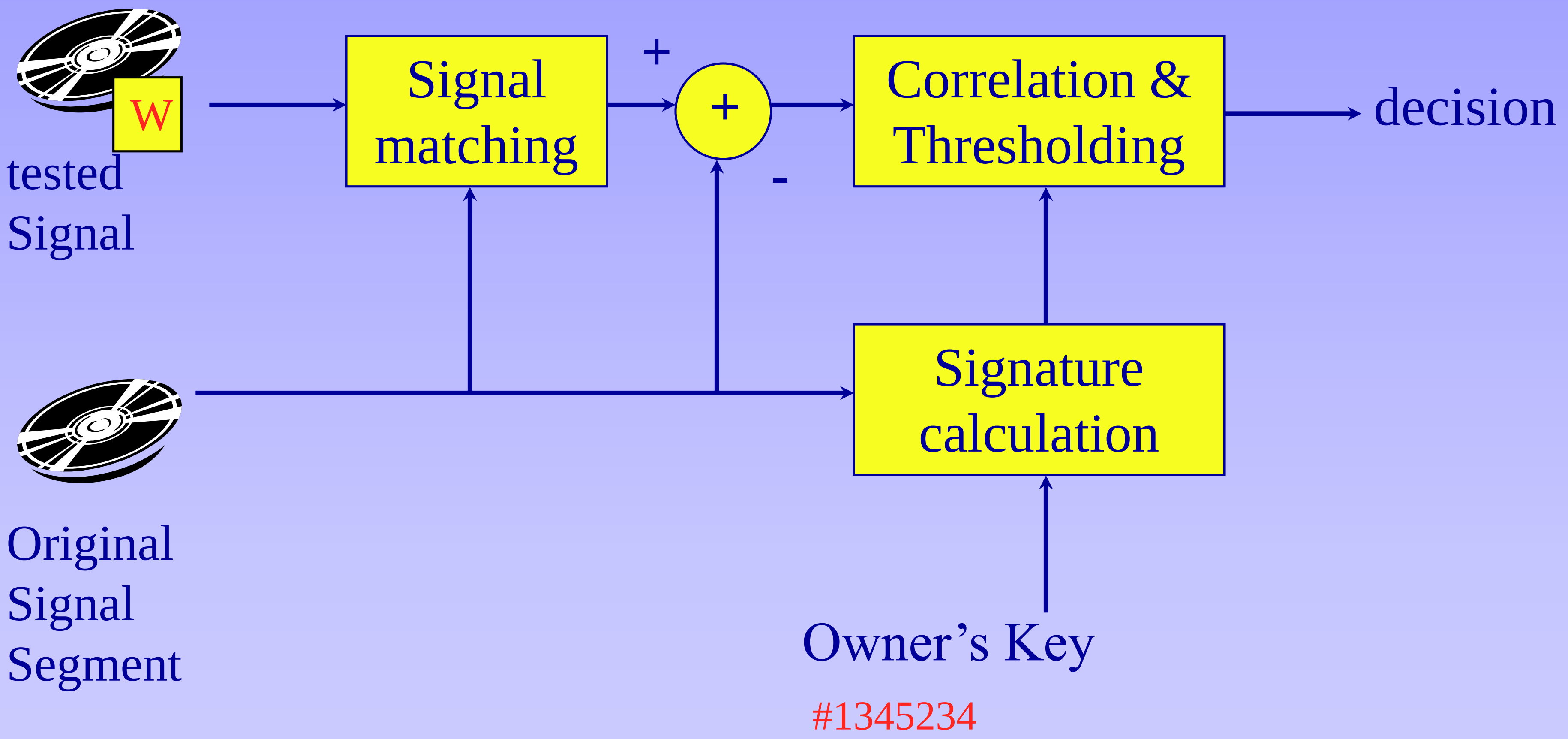
Signature detection



Signature Embedding Mechanism



Detection Mechanism



Work Goals:

Dealing with ownership claims of attackers.

To do so:

- Find the characteristics of the attacker's system.
- Modify the detection system in order to increase the signature detection probability for these attacks.

Fundamental assumption:

The attack is limited in the sense of preservation audio quality.

Attacks

<u>Naive attacks:</u>

- Gain
- Coping part of a file
- Equalization
- Compression
- …

<u>Sophisticated attacks:</u>

- Nonlinear transforms
- All-pass filters
  - Fixed and time varying
- Equalization
- Noise
  (White or Colored – perceptually based)
- Time scale modification
- Echo
- …

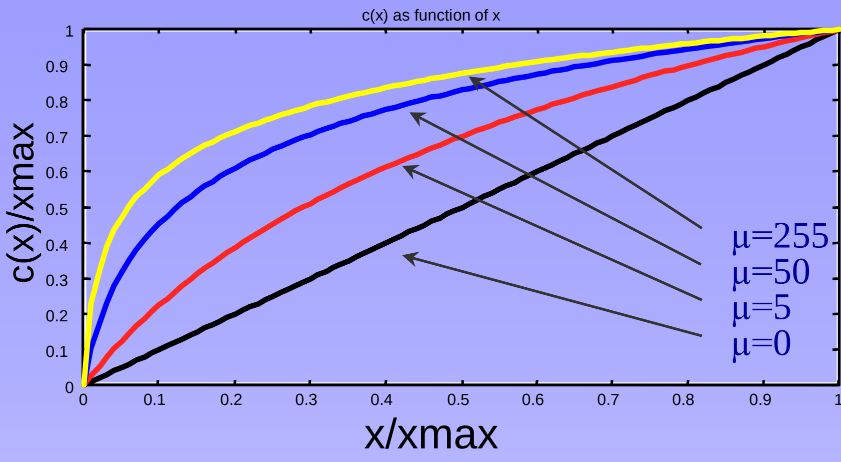
These attacks should not decrease the audio quality!

All-pass filter effect

- The effect is not heard even for big number of filter taps (hundreds)
- This distortion reduces the correlation measure to ~0.4
- May be corrected by using LMS to estimate the filter – correlation measure is increased to >0.9

μ-Law Nonlinear Distortion

We used μLaw distortion with μ=5.
(the hearing threshold is μ=~3.)



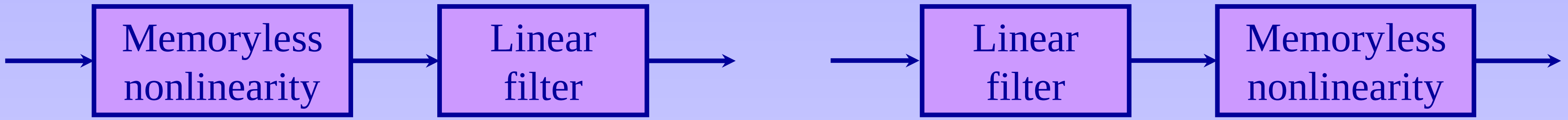
- This distortion reduces the correlation measure to ~0.95
- Adding this nonlinear distortion before/after all-pass filtering reduces the correlation to 0.36/0.38 (μ=5)
- while the allpass filter effect could be compensated by the regular LMS, the combined linear filter and nonlinearity requires a different treatment

μ-Law Nonlinear Distortion – cont'

- The solution:

  insert a nonlinearity into the distortion model.

Considered distortion models:



Possible Model Estimation

- Volterra series method:

provides a method for describing the input-output relationship for a nonlinear device with memory:

$$DC + \sum_i h_i \cdot x_i + \sum_i \sum_j h_{ij} \cdot x_i \cdot x_j + \sum_i \sum_j \sum_k h_{ijk} \cdot x_i \cdot x_j \cdot x_k + \ldots$$

-Volterra series of N taps (filter memory) and order 2 requires $N+N^2$ Coefficients.

- LMS based scheme for parameter estimation (the h's)

- Convergence rate is much slower for the Volterra method compared to LMS (Linear filter) !

- Special problem when dealing with a time varying all-pass filter.

Proposed solution

Assumptions:

- The nonlinearity is not high
- Antisymmetric

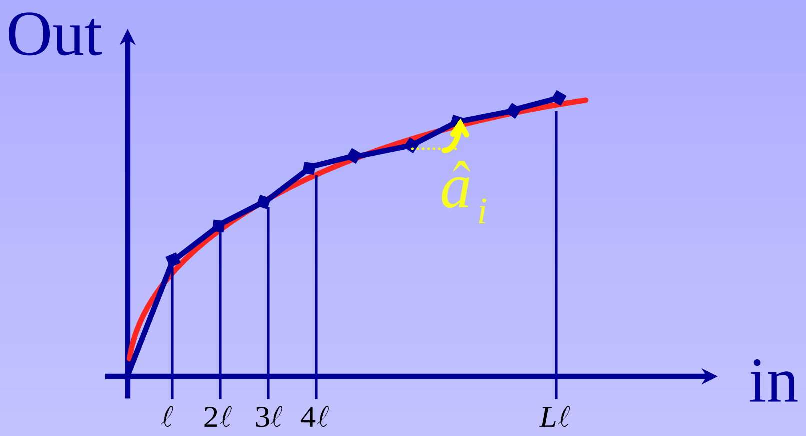
Model: Piecewise linear approximation.

Piecewise Approximation of Nonlinearity

$$out = \hat{f}\left(in\right)$$

Where:

$$\hat{f}\left(x\right) = \begin{cases} \hat{f}_1\left(x\right) & 0 \le x \le \ell \\ \hat{f}_2\left(x\right) & \ell \le x \le 2\ell \\ \vdots & \\ \hat{f}_L\left(x\right) & \left(L-1\right)\cdot\ell \le x \le L\cdot\ell \end{cases}$$



$$\hat{f}_i\left(x\right) = \hat{a}_i \cdot \left(x - \left(i-1\right)\cdot\ell\right) + \ell \cdot \sum_{k=0}^{i-1} \hat{a}_k \qquad \left(\hat{a}_0 = 0\right)$$

Nonlinearity Estimation

Two methods:

- Using LS criterion and (x,y) pairs to estimate the slope coefficients.
- Adaptive system for estimating the slope coefficients using sample by sample adaptation.

Nonlinearity Estimation – LS Criterion

For each segment calculate $\hat{a}_i$ using the *i-th* segment samples: $(x,y)\in\{(x,y):\ (i-1)\cdot\ell\le x\le i\cdot\ell\}$

Begin with $\hat{a}_1$, continue with $\hat{a}_2,\hat{a}_3,\ldots,\hat{a}_L$.

Note: D_i depends on $\hat{a}_1,\ \hat{a}_2,\ \ldots\ ,\ \hat{a}_i$

Distortion for the i-th segment:

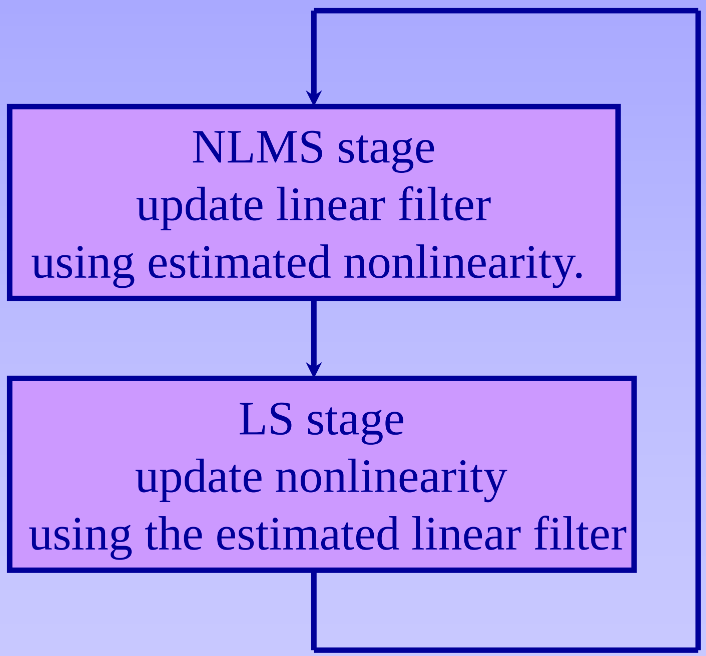
$$D_i=\sum_{(x,y)\in\{(x,y)\}_i}\left(\hat{a}_i\cdot\left(x-(i-1)\cdot\ell\right)+\ell\cdot\sum_{k=1}^{i-1}\hat{a}_k-y\right)^2$$

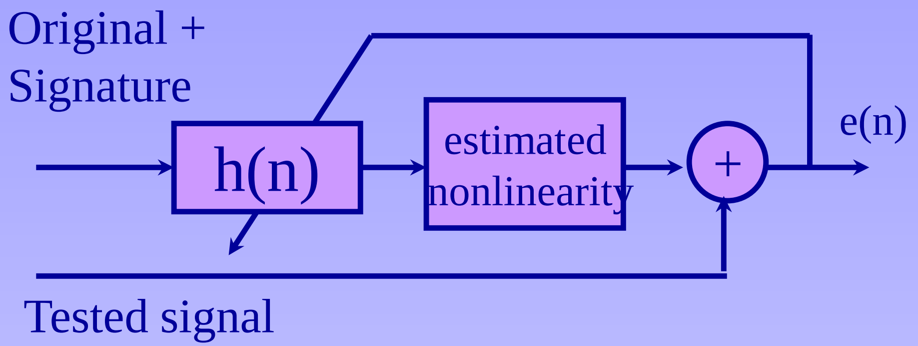
gives:

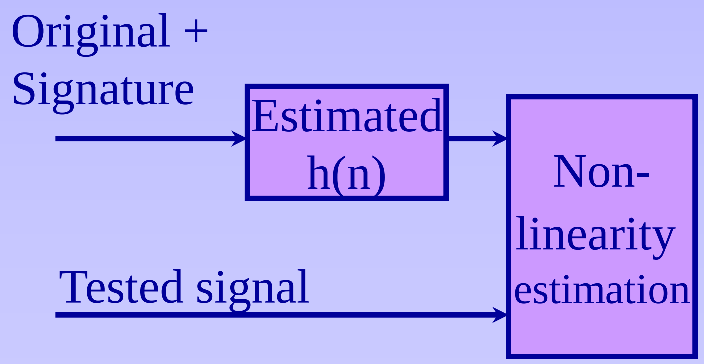
$$\hat{a}_i=\frac{\displaystyle\sum_{(x,y)\in\{(x,y)\}_i}\left(y-\ell\cdot\sum_{k=1}^{i-1}\hat{a}_k\right)\cdot\left(x-(i-1)\cdot\ell\right)}{\displaystyle\sum_{(x,y)\in\{(x,y)\}_i}\left(x-(i-1)\cdot\ell\right)^2}\qquad i=1,\ldots,L$$

Estimating The Complete Distortion Model (LS Criterion)

- Two stage process:

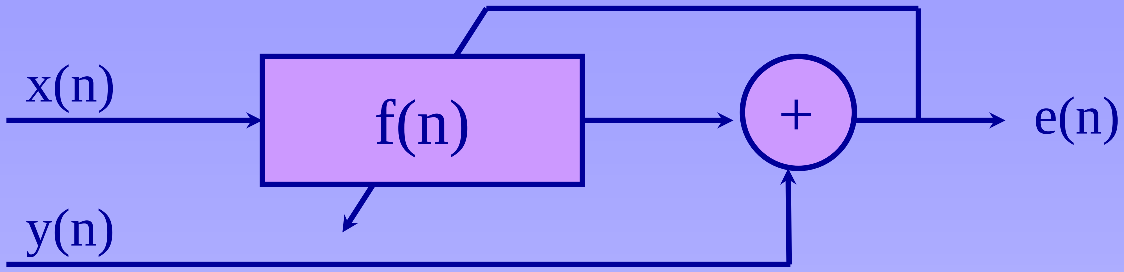






Typically, convergence is achieved in 2-3 iterations

Nonlinearity Estimation – Adaptive Method



For each sample pair *x(n)*,*y(n)* update $\left\{\hat{a}_i\right\}_{i=1}^{L}$ as follows:

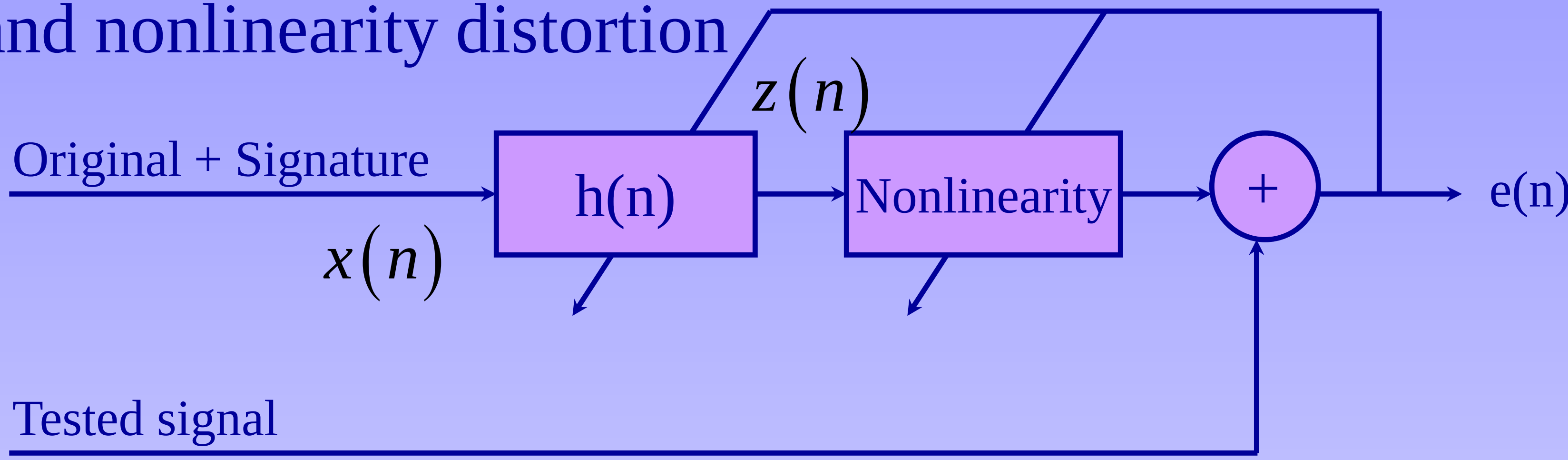
$$\hat{a}_j^{m+1} = \hat{a}_j^{m} + \begin{cases} \mu \cdot e(n) \cdot \left(x(n) - \left(i\left(x(n)\right) - 1\right) \cdot \ell\right) & \text{if } j = i\left(x(n)\right) \\ \mu \cdot e(n) \cdot \ell & \text{if } j < i\left(x(n)\right) \\ 0 & \text{elsewhere} \end{cases}$$

$$j = 1, \ldots, L$$

When *i(x)* denotes the segment that includes *x*.

Estimating The Complete Distortion Model (Adaptive Method)

- Use NLMS method to estimate both linear filter and nonlinearity distortion

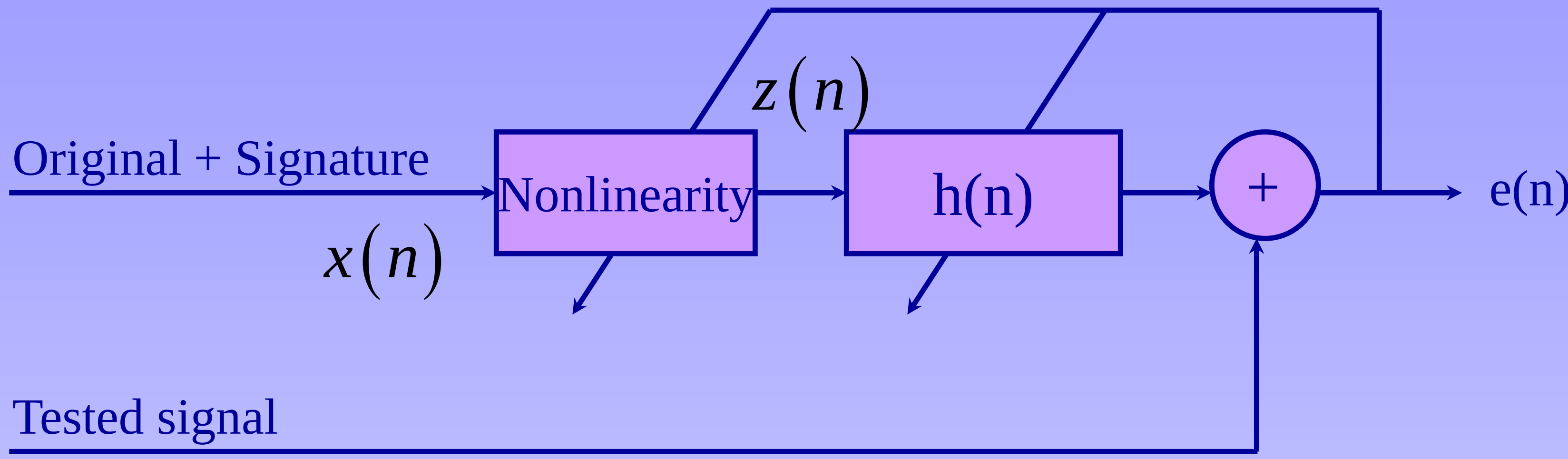


$$\hat{h}_j^{m+1} = \hat{h}_j^m + \mu \cdot e(n) \cdot x(n-j) \cdot \hat{a}_{i(z(n))}^m \qquad j = 1, \ldots, N$$

$$\hat{a}_j^{m+1} = \hat{a}_j^m + \mu \cdot e(n) \cdot \begin{cases} \left(z(n) - \left(i(z(n)) - 1\right) \cdot \ell\right) & \text{if } j = i(z(n)) \\ \ell & \text{if } j < i(z(n)) \\ 0 & \text{elsewhere} \end{cases}$$

$$j = 1, \ldots, L$$

Estimating The Complete Distortion Model (Adaptive Method)



$$\hat{h}_j^{m+1} = \hat{h}_j^m + \mu \cdot e(n) \cdot z(n-j) \qquad j = 1,\ldots,N$$

$$\hat{a}_j^{m+1} = \hat{a}_j^m + \mu \cdot e(n) \cdot \left[\sum_{k=0}^{N-1} \hat{h}_k^m \cdot \begin{cases} \left(x(n-k) - \left(i\left(x(n-k)\right)-1\right)\cdot \ell\right) & \text{if } j = i\left(x(n-k)\right) \\ \ell & \text{if } j < i\left(x(n-k)\right) \\ 0 & \text{elsewhere} \end{cases} \right]$$

$$j = 1,\ldots,L$$

Comparison between LS method and Adaptive Method

- Both methods give about the same results for linear filtering followed by nonlinearity.
- The adaptive method can handle nonlinearity followed by linear filtering is the same manner while LS involves inversing of a linear filter in this case.

Results and Conclusions

- Piecewise linear approximation requires fewer parameters in comparison to the Volterra method.
- We developed update equations that enable estimating the parameters of specific types of nonlinear filters.
- The enhanced watermark detection scheme provides a significant performance improvement (0.8-0.95 vs. 0.36-0.38).