Optimization algorithms in real time analog and digital nonlinear models

IRF department Moscow R&D centre

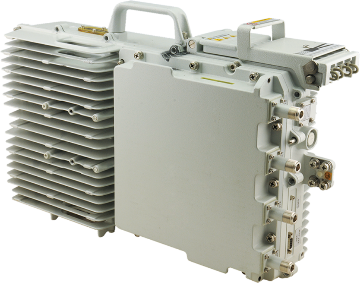
Wireless Research Division
Fixed Network Research Division
WiFi Research Division

Overview

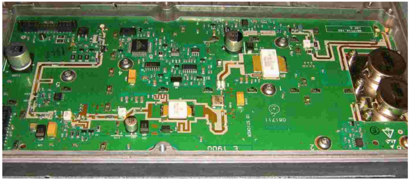
Introduction

Macro Cell Tower

Remote Radio Head (RRH)



Microwave Power Amplifier (PA)



Efficiency Increasing Problem

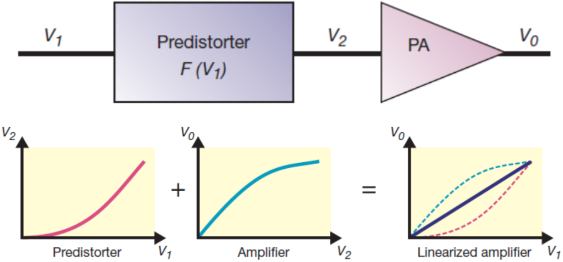
For example, if the efficiency of the amplifier is 10% then 100W power amplifier will consume 1 kW and 900W, we must dissipate in the form of heat.

But, if the efficiency of the amplifier is 50% then 100W power amplifier will consume 200W and we must dissipate in the form of heat only 100W.

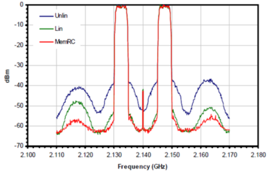
On the other hand, PA efficiency increasing means its nonlinearity increasing. As result we will have out-of-band radiation (intermodulation distortions) and we cannot to use such equipment because we will interfere with neigh-boring channels.

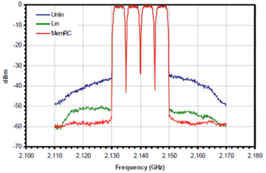
Digital Pre-distortion (DPD)

PA linearization (DPD) principle

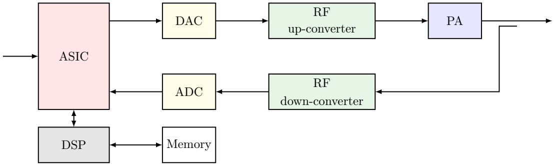


PA output spectrum after DPD





Digital pre-distortion algorithm implementation in hardware (ASIC)



PA Model Representation

PA Model is Multivariable Function :

$$y(k) = F\left[x(k-M), \ldots, x(k), \ldots, x(k+M)\right],$$

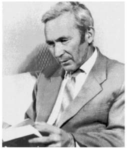
$x(k)$ – input samples (complex), $y(k)$ – output samples (complex), $M = 7 \ldots 15$ – model memory.

Hilbert's Thirteenth Problem

...it is probable that the root of the equation of the seventh degree is a function of its coefficients which does not belong to this class of functions capable of nomographic construction, i. e., that it cannot be constructed by a finite number of insertions of functions of two arguments. In order to prove this, the proof would be necessary that the equation of the seventh degree

$$t^7 + xt^3 + yt^2 + zt + 1 = 0$$

is not solvable with the help of any continuous functions of only two arguments



Kolmogorov-Arnold Representation Theorem

Let $f : \mathbb{I}^n := [0,1]^n \to \mathbb{R}$ be an arbitrary multivariate continuous function. Then it has the representation

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^{n} \psi_{q,p}(x_p) \right)$$

with continuous one-dimensional outer and inner functions Φ_q and $\psi_{q,p}$.

PA Model Examples

Volterra Model:

$$y(n) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m_1=0}^{M} \sum_{m_2=m_1}^{M} \cdots \sum_{m_{(p+1)/2}=m_{(p-1)/2}}^{M} \sum_{m_{(p+3)/2}=0}^{M} \cdots$$

$$\cdots \sum_{m_p=m_{p-1}}^{M} h_p(m_1, m_2, \ldots, m_p) \prod_{i=1}^{(p+1)/2} u(n-m_i) \prod_{j=(p+3)/2}^{p} u^*(n-m_j)$$

Memory Polynomial model:

$$y(n) = \sum_{p=1}^{P} \sum_{m=0}^{M-1} a_{pm} x(n-m)|x(n-m)|^{p-1}$$

PA Model Examples

Extended Memory Polynomial model:

$$y(n) = \sum_{p=1}^{P} \sum_{m=0}^{M} a_{pm} u(n-m) |u(n-m)|^{p-1} +$$

$$+ \sum_{p=2}^{P} \sum_{m=0}^{M-1} \sum_{\substack{g=-m \\ g \neq 0}}^{M-1} b_{pmg} u(n-m) |u(n-m-g)|^{p-1}$$

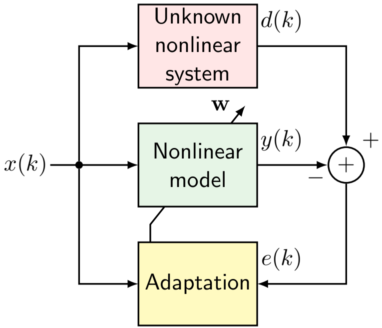
Wiener Model:

$$y(n) = \sum_{p=1}^{P} a_p \left[\sum_{m=0}^{M} h_m u(n-m) \right] \sum_{m=0}^{M} |h_m u(n-m)|^{p-1}$$

Hammerstein Model:

$$y(n) = \sum_{m=0}^{M} h_m \left[\sum_{p=1}^{P} a_p u(n-m) |u(n-m)|^{p-1} \right]$$

Nonlinear system identification problem



$x(k)$ – input signal samples
$y(k)$ – nonlinear model output
$d(k)$ – unknown system output
$e(k) = d(k) - y(k)$ error signal
$\mathbf{w}$ – nonlin. model coeff. vector

Nonlinear model with linear part:

$$y(k) = \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} c_{p,q} \underbrace{x(k - r_p)\Phi_q\big(|x(k - s_p)|\big)}_{u_{q+pQ}(k)}$$

P – number of nonlinear terms;
Q – number of basic functions $\Phi_q(\ )$;
$c_{p,q}$ – model coefficients;
$r_p,\ s_p$ – delays of signal $x(k)$ and magnitude $|x(k)|$.
We can send S samples of the input signal $x(k)$, $k = 0 \ldots S$ through the unknown system and capture a desired signal vector:

$$\mathbf{d} = \big[d(0),\ d(1),\ \ldots\ d(S-1)\big]^T.$$

Nonlinear system identification problem

Vector $\mathbf{u}^T(k)$ of nonlinearities and vector of coefficients $\mathbf{w}$ contains PQ elements:

$$\mathbf{u}^T(k) = \left[u_0(k),\ u_1(k), \ldots, u_{PQ-1}(k)\right],$$

$$\mathbf{w} = \left[c_{0,0},\ c_{0,1}, \ldots, c_{P-1,Q-1}\right]^T,$$

$$w_{q+pQ} = c_{p,q}, \quad p = 0 \ldots P-1, \quad q = 0 \ldots Q-1.$$

Matrix $\mathbf{U}$ of nonlinearities for $k = 0 \ldots S$ samples has S rows and PQ columns:

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}^T(0) \\ \mathbf{u}^T(1) \\ \mathbf{u}^T(2) \\ \vdots \\ \mathbf{u}^T(S-1) \end{bmatrix} = \begin{bmatrix} u_0(0) & u_1(0) & \ldots & u_{PQ-1}(0) \\ u_0(1) & u_1(1) & \ldots & u_{PQ-1}(1) \\ u_0(2) & u_1(2) & \ldots & u_{PQ-1}(2) \\ \vdots & \vdots & \ddots & \vdots \\ u_0(S-1) & u_1(S-1) & \ldots & u_{PQ-1}(S-1) \end{bmatrix}$$

Nonlinear system identification problem

Then output of nonlinear model can be written as

$$\mathbf{y} = \mathbf{U}\mathbf{w} \quad \text{where} \quad \mathbf{y} = \left[y(0),\ y(1), \ldots, y(S-1)\right]^T.$$

Goal of adaptation algorithm is minimization of cost function

$$J(\mathbf{w}) = \|\mathbf{d} - \mathbf{U}\mathbf{w}\|_2^2 = (\mathbf{d} - \mathbf{U}\mathbf{w})^H (\mathbf{d} - \mathbf{U}\mathbf{w}) \to \min$$

Traditional tool to solve problem is least squares (LS) method, based on solving system of linear equations

$$\mathbf{w} = \left(\mathbf{U}^H\mathbf{U}\right)^{-1}\mathbf{U}^H\mathbf{d} = \mathbf{V}^{-1}\mathbf{r},$$

where $\mathbf{V} = \mathbf{U}^H\mathbf{U}$ – correlation matrix of input signal nonlinearities, $\mathbf{r} = \mathbf{U}^H\mathbf{d}$ – cross correlation vector between input nonlinearities and desired signals.

Nonlinear system identification problem

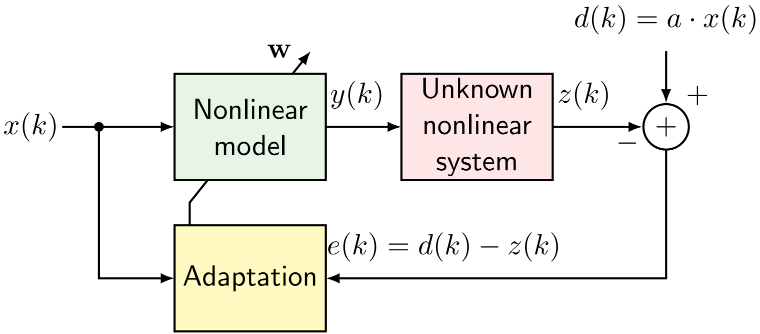
Our current nonlinear models are big and can contains over 10000 coefficients.

Despite this, we cannot achieve the required residual error $e(k)$ level for some cases.

We have a special test that allows us to get the potential accuracy of identification and we know that the residual identification error can be reduced even up to 20 times

Is it possible to find alternative nonlinear model, with better performance and low complexity?

Nonlinear system inversion problem



$x(k)$ – input signal samples
$y(k)$ – nonlinear model output
$z(k)$ – unknown nonlin. system output
$d(k) = a \cdot x(k)$ – desired signal, $a = \text{const}$
$e(k) = d(k) - z(k)$ error signal
$\mathbf{w}$ – nonlin. model coeff. vector

Nonlinear system inversion problem

Nonlinear model has the same structure as previous:

$$y(k) = \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} c_{p,q} \underbrace{x(k-r_p)\Phi_q\big(\,|x(k-s_p)|\,\big)}_{u_{q+pQ}(k)}$$

In that context nonlinear model works as pre-compensator for unknown nonlinear model. Goal of adaptation algorithm is different:

$$J(\mathbf{w}) = \|\mathbf{d}-\mathbf{z}\|_2^2 = (a\mathbf{x}-\mathbf{z})^H\,(a\mathbf{x}-\mathbf{z}) \to \min,$$

where $\mathbf{z}$ – vector with output samples of unknown system, $\mathbf{x}$ – vector of original signal samples, a – some constant.

$$\mathbf{z} = \big[z(0),\ z(1),\ \ldots\ z(S-1)\big]^T$$
$$\mathbf{x} = \big[x(0),\ x(1),\ \ldots\ x(S-1)\big]^T$$

Nonlinear system inversion problem

This problem can be solved by iterative LS method. To achieve acceptable performance, adaptation algorithm requires several (usually 10..15) iterations to full convergence. Function for minimization on t-th iteration can be written as

$$J\left(\mathbf{w}^{(t)}\right) = \left\|\mathbf{U}^{(t-1)}\mathbf{w}^{(t-1)} + \frac{1}{a}\mathbf{e}^{(t)} - \mathbf{U}^{(t)}\mathbf{w}^{(t)}\right\|_2^2,$$

where $\mathbf{e}^{(t)}$ – vector with samples of residual error

$$\mathbf{e}^{(t)} = \left[e(0),\ e(1),\ \ldots\ e(S-1)\right]^T$$

Nonlinear system inversion problem

Hence LS equations can be transformed to

$$\mathbf{w}^{(t)} = \left(\mathbf{U}^{(t)^H}\mathbf{U}^{(t)}\right)^{-1}\mathbf{U}^{(t)^H}\left(\frac{1}{a}\mathbf{e}^{(t)} + \mathbf{U}^{(t-1)}\mathbf{w}^{(t-1)}\right) \approx$$

$$\approx \left(\mathbf{U}^{(t)^H}\mathbf{U}^{(t)}\right)^{-1}\left(\frac{1}{a}\mathbf{U}^{(t)^H}\mathbf{e}^{(t)} + \mathbf{U}^{(t-1)^H}\mathbf{U}^{(t-1)}\mathbf{w}^{(t-1)}\right) =$$

$$= \mathbf{V}^{(t)^{-1}}\left(\frac{1}{a}\mathbf{U}^{(t)^H}\mathbf{e}^{(t)} + \mathbf{V}^{(t-1)}\mathbf{w}^{(t-1)}\right),$$

where $\mathbf{V}^{(t)}$ and $\mathbf{V}^{(t-1)}$ correlation matrices of input signal nonlinearities for current and previous iterations; $\mathbf{w}^{(t)}$ and $\mathbf{w}^{(t-1)}$ – coefficients for current and previous iterations.

Nonlinear system inversion questions

1) How to reduce number of iterations, required for convergence to steady state?

2) How to guarantee smooth and robust convergence in the presence of significant delay between estimation of $\mathbf{U}^{(t)^H}\mathbf{e}^{(t)}$ and update of $\mathbf{w}^{(t)}$?

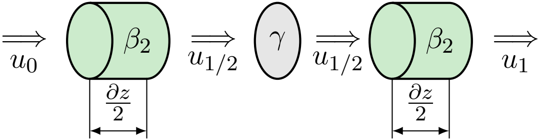
3) How to compress vector $\mathbf{U}^{(t)^H}\mathbf{e}^{(t)}$ in order to simplify its transmission?

Fiber Optic math problems introduction

Mathematical model of optic fiber is described by nonlinear Schrodinger equation (NLSE) (1) for modelling of light beam propagation in dispersion nonlinear environment along optical fiber.

$$\frac{\partial A}{\partial z} + \beta_1 \frac{\partial A}{\partial t} + \frac{j}{2}\beta_2 \frac{\partial^2 A}{\partial t^2} - \frac{1}{6}\beta_3 \frac{\partial^3 A}{\partial t^3} = j\gamma |A|^2 A - \frac{\alpha}{2} A. \qquad (1)$$

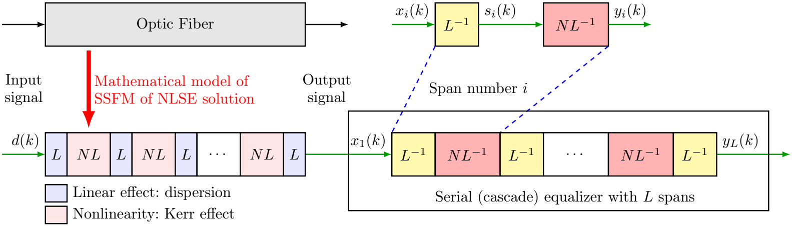
Usually, equation (1) is solved by Split–Step Fourier Method.



The main idea is in iterative (cascade or serial) splitting common NLSE into two parts. First part corresponds to dispersion and attenuation loss, second part corresponds to nonlinear effects.

Serial Model Parameters Estimation

Model of equalizer is a reversed model of split–step Fourier decomposition. It consists of serial similar blocks of linear and nonlinear transformations. So, schematic structure of serial equalizer for optic fiber link is



Serial Model Parameters Estimation

Mathematical model of optic fiber is below

$$\mathbf{output} = L\Bigg(f\bigg(L\Big(\ldots f\big(L(\mathbf{input})\big)\ldots\Big)\bigg)\Bigg)$$

where $\mathbf{output}$ and $\mathbf{input}$ – vectors of respective signals; $L(\ )$ and $f(\ )$ - unitary operators: linear and nonlinear effects of fiber. Mathematical model of equalizer is a serial model with recurrent equations:

$$\begin{cases} x_1(k) - \text{known for } k = 1 \ldots K; \\ s_i(k) = \sum_{m=-M}^{M} h_{i,m} x_i(k-m); \\ y_i(k) = s_i(k) \exp\left(-j\gamma_i |s_i(k)|^2\right); \\ x_{i+1}(k) = y_i(k); \quad i = 1 \ldots L. \end{cases}$$

where $K, L, M \in \mathbb{N}$, $K >> L, K >> M$; $h_{i,m} \in \mathbb{C}$ and $\gamma_i \in \mathbb{R}$ – unknown parameters of model; $x_i(k), s_i(k), y_i(k) \in \mathbb{C}$.

Serial Model Parameters Estimation Questions

1) How to define the parameters $h_{i,m}$ and γ_i to satisfy:

$$\sum_{k=1}^{K} |y_L(k) - d(k)|^2 \to \min, \tag{2}$$

$L, M \in \mathbb{N},\ x_1(k)$ and $d(k) \in \mathbb{C}$ are known for $k = 1 \dots K$.
2) How to solve (2) with constrains: $h_{i_1,m} = h_{i_2,m}$, $\forall i_1, i_2$ and $\gamma_{i_1} = \gamma_{i_2}\ \forall i_1, i_2$.
3) How to define the optimal number of span L_{opt}:

$$\sum_{k=1}^{K} |y_{L_{opt}}(k) - d(k)|^2 \to \min,$$

4) How to minimize number of span $L < L_{opt}$ with fixed performance degradation δ:

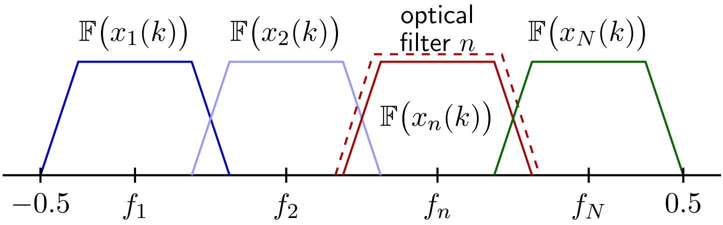
$$\sum_{k=1}^{K} |y_L(k) - d(k)|^2 - \sum_{k=1}^{K} |y_{L_{opt}}(k) - d(k)|^2 \leq \delta$$

Multidimensional processing of optical signals

Optical signal with complex envelop $x(k) \in \mathbb{C}$, $k = 1, 2, \dots K$ time domain index, propagates through fiber channel with model (1).
Signal $x(k)$ consists of narrowband signals $x_n^{bb}(k) \in \mathbb{C}$, $n = 1, 2, \dots N$ as describes in equation (3), where f_n – carrier frequency of channel n:

$$x(k) = \sum_{n=1}^{N} x_n(k), \quad x_n(k) = x_n^{bb} \cdot \exp\left(j 2\pi f_n k\right) \tag{3}$$

Discrete Fourier transform $\mathbb{F}(\;)$ of signals $x(k)$ and $x_n(k)$



Multidimensional processing of optical signals

Optical filter n is applied at receiver before ADC. It separates respectively signal $\tilde{x}_n^{bb}(k) \in \mathbb{C}$, $n = 1, 2, \ldots N$ with distortions. One of the way, how to compensate distortion and to recover signals $x_n^{bb}(k) \in \mathbb{C}$, $n = 1, 2, \ldots N$ is high order multi-dimension model, for example:

$$y_{n_0}(k) = \sum_{n=1}^{N} \sum_{m=0}^{M-1} \alpha_{n_0,m,n} \tilde{x}_n^{bb}(k - m + D) +$$

$$+ \sum_{n=1}^{N} \sum_{n_1=1}^{N} \cdots \sum_{n_N=n|N-1|}^{N} \sum_{r_{n_0}=\min}^{\max} x_{n_0}(k - r_{n_0}) \times$$

$$\times \sum_{s_{n_1}} \sum_{s_{n_2}} \cdots \sum_{s_{n_N}} \beta_{\substack{n_0,n_1,\ldots,n_N \\ s_{n_1},s_{n_2},\ldots,s_{n_N}}} \times$$

$$\times G_{\substack{n_0,n_1,\ldots,n_N \\ s_{n_1},s_{n_2},\ldots,s_{n_N}}} \left(x_1(k - s_{n_1}), x_2(k - s_{n_2}), \ldots, x_N(k - s_{n_N}) \right)$$

Multidimensional processing of optical signals

$y_{n_0}(k) \in \mathbb{C}$, $n = 1, 2, \ldots N$ – recovered signal of $x_n^{bb}(k)$ respectively;

$G_{\substack{n_0, n_1, \ldots, n_N \\ s_{n_1}, s_{n_2}, \ldots, s_{n_N}}}\big(x_1, x_2, \ldots, x_N\big)$ – multi-dimension nonlinear functions, which can be represented as polynomial, spline or other simplification.

If we will consider $N = 4$, each memory 11, that number of functions $G_{\substack{n_0, n_1, \ldots, n_N \\ s_{n_1}, s_{n_2}, \ldots, s_{n_N}}}\big(x_1, x_2, \ldots, x_N\big)$ will be $4 \cdot 4^4 \cdot 11^4 \approx 15 \cdot 10^6$. If each function will be unpacked in series, that number of coefficients will be huge.

How to reduce order and dimensions of model?

Thank You

www.huawei.com