Nonlinear phenomena in beamforming

Aarne Mämmelä, Markku Kiviranta, Henna Paaso

VTT Technical Research Centre of Finland, P.O. Box 1100, FI-90571 Oulu, Finland,
e-mail Aarne.Mammela@vtt.fi, Markku.Kiviranta@vtt.fi, Henna.Paaso@vtt.fi

Abstract

In this paper we present a summary of nonlinear system level problems in beamforming. The paper summarizes first the nonideal phenomena in power amplifiers, mixers, and oscillators. In digital predistortion inversion of the nonlinearity is needed. Not all linear and nonlinear systems are invertible or they may have only a preinverse or postinverse. Commutation is needed in indirect learning architectures for predistortion. The spatial and temporal order of compensation of different distortions is discussed. Some numerical results are also included.

1. Introduction

The analog parts of the system are in general nonideal including power amplifiers, mixers, and oscillators [1-3]. The changes in the nonlinear properties are usually slow and caused by for example temperature changes. In some cases where the transmitted power is changing rapidly due to power control, the temperature may also change quite rapidly. Such cases are out of the scope of our discussion. Our numerical results are for beamforming but similar phenomena must be taken into account in all multi-antenna systems. Our results show that in beamforming each branch should be separately compensated, otherwise the antenna pattern is distorted. Therefore we consider first the compensation of only a single branch. There are many approaches for compensation, but our focus is in digital predistortion. A feedback signal is needed from the output of the nonlinear parts, otherwise they cannot be predistorted.

Power amplifiers are nonlinear components. Nonlinearities cause unwanted effects including spectral spreading, intersymbol interference, and constellation warping. In wideband systems the nonlinear power amplifier may have memory, which is caused by the reactive elements in the amplifier. The memory causes frequency selectivity, which can be seen when considering the amplitude modulation/amplitude modulation (AM/AM) and amplitude modulation/phase modulation (AM/PM) characteristics, which are frequency dependent. Some nonlinear models are separable. For example, in the Wiener model the nonlinear model with memory is divided into two parts, the linear part with memory and the nonlinear part without memory [1]. In the Hammerstein model the order of the parts is reversed. A common method to avoid nonlinear effects is to use constant envelope signals. However, we will show that this is not enough if the nonlinearity has memory. Furthermore, many modern modulation methods are not constant envelope signals, for example quadrature amplitude modulation (QAM) and orthogonal frequency division multiplexing (OFDM).

Mixers have two nonideal properties including self-mixing and in-phase/quadrature-phase (I/Q) imbalance. In self-mixing the two input signals are mixed together due to signal leakage between the input ports and this results in a DC component at the output of the mixer. The I/Q imbalance is caused by the amplitude and phase imbalance in the mixer. The model is linear after the signal is divided into I and Q components. The division into the two components is a nonlinear operation. The oscillators have two nonideal phenomena including frequency offset and phase jitter. The frequency offset is usually slowly changing and therefore it can be quite easily compensated. The phase jitter may have a strong effect in cases where the symbols are rather long, for example in OFDM systems.

2. Signal design for nonlinear systems

One approach to minimize nonlinear effects in power amplifiers is to use large enough backoff or constant envelope signals. Backoff is defined as the ratio of the maximum output power and the average output power and it is usually expressed in decibels. A large backoff improves the linearity of the amplifier as seen by the signal, but on

the other hand, the amplifier is also inefficient. The backoff can be reduced by using constant envelope signals, for example minimum shift keying (MSK). In that case a nonlinear memoryless system changes the amplitude and phase linearly and no predistortion is required. However, if the nonlinearity has memory as for example in the Wiener model, the situation becomes more complicated. Memory implies some kind of filtering and the constant envelope property is destroyed. In the Hammerstein model such a problem does not exist. In general it is not enough that the signal has a constant envelope when the nonlinear system has some memory and some kind of predistortion is needed. Many modern systems are using nonconstant envelope signals. This is another reason why predistortion is used. Therefore for a minimum backoff and a maximum efficiency of the power amplifier the envelope should be as constant as possible. A possibility is to use phase or frequency modulator at the output of an OFDM system [4].

3. Invertibility and commutability

Digital predistortion is done by inverting the linear and nonlinear distortions. Predistortion is used in the transmitter because we are interested in digital implementations and it is best to compensate the nonlinear properties near the place where they are produced. If postdistortion or equalization would be done in the receiver, we would have a much more complicated problem since the channel is usually time-variant. Our system model would then be nonlinear and time-variant. The invertibility of the systems is also an important question. In some indirect learning architectures the postinverse is first found and after that the order of the amplifier and the postinverse is changed so that the postinverse is used as a preinverse. Thus the commutability is an important problem [5].

A system is linear if the superposition theorem is valid. Linear systems can be presented with an equation $\mathbf{s} = \mathbf{H}\boldsymbol{\theta}$ where $\boldsymbol{\theta}$ is the input signal vector, $\mathbf{s}$ is the output signal vector, and $\mathbf{H}$ is the known $K \times M$ transformation matrix representing the system and does not depend on $\boldsymbol{\theta}$ [6]. An example of a linear system is a multi-antenna system without any nonlinear elements. If $\mathbf{H}$ is a nonsingular square matrix, the transformation is bijective, and an inverse $\mathbf{H}^{-1}$ exists. In that case the original matrix and its inverse are commutable, i.e., $\mathbf{H}\mathbf{H}^{-1} = \mathbf{H}^{-1}\mathbf{H} = \mathbf{I}$ where $\mathbf{I}$ is the unity matrix.

In most practical cases $\mathbf{H}$ is not a square matrix and no inverse exists. In that case we can still use a pseudoinverse $\mathbf{H}^{+}$ [7]. It is not a true inverse since it is not commutable with the original matrix. If $K > M$, the transformation $\mathbf{H}$ is injective and we can find a pseudoinverse $\mathbf{H}^{+}$. We call it a left inverse since $\mathbf{H}^{+}\mathbf{H} = \mathbf{I}$. It is also called the postinverse since the pseudoinverse is used after the original transformation. If $K < M$, the transformation is surjective and the pseudoinverse is called a preinverse or right inverse since $\mathbf{H}\mathbf{H}^{+} = \mathbf{I}$.

Just as in the linear case a nonlinear system $\mathbf{s} = \mathbf{H}(\boldsymbol{\theta})$ is invertible if it is a bijection. The inverse $\mathbf{H}^{-1}(\boldsymbol{\theta})$ has the property $\mathbf{H}[\mathbf{H}^{-1}(\boldsymbol{\theta})] = \mathbf{H}^{-1}[\mathbf{H}(\boldsymbol{\theta})] = \boldsymbol{\theta}$, i.e., commutation is valid for the inverse. In general a nonlinear system can be inverted only iteratively [6]. It is well known that the iterative algorithms have convergence problems due to local optima, and no general solution to those problems is known. If $\mathbf{H}(\boldsymbol{\theta})$ is an injective, it has a postinverse, and if it is surjective, it has a preinverse. The preinverse and postinverse are not in general identical unless the transformation is a bijection. Usually it is quite difficult to study the invertibility of nonlinear systems and to find the inverse. Because of saturation effects nonlinear systems are usually invertible only for a finite input amplitude range where the transformation corresponds to a bijection [2]. Thus the best we can do is to approximate a soft limiter because saturation effects cannot be ideally compensated.

It is beneficial if the system would be separable and the system $\mathbf{H}(\boldsymbol{\theta})$ could be separated into two parts $\mathbf{H}_1(\boldsymbol{\theta})$ and $\mathbf{H}_2(\boldsymbol{\theta})$. We use the theory of composite functions, which are defined as $\mathbf{s} = \mathbf{H}(\boldsymbol{\theta}) = \mathbf{H}_2[\mathbf{H}_1(\boldsymbol{\theta})]$, where $\mathbf{H}_1(\boldsymbol{\theta})$ is the inner nonlinear function and $\mathbf{H}_2(\boldsymbol{\theta})$ is the outer nonlinear function, which are not in general commutable. If all the inverses exist, the inverse is $\mathbf{H}^{-1}(\boldsymbol{\theta}) = \mathbf{H}_1^{-1}[\mathbf{H}_2^{-1}(\boldsymbol{\theta})]$. For example in a postinverse, the function $\mathbf{H}_2(\boldsymbol{\theta})$ must be inverted first before the function $\mathbf{H}_1(\boldsymbol{\theta})$ can be inverted. In some cases the system is nonlinear in some parameters and linear in some other parameters [6]. This idea has been applied in [8] by using a memory polynomial model for the power amplifier. The memory polynomials are based on Volterra series and therefore they have both similar problems [2]. For example, the transformation must be smooth enough so that the memory polynomials converge.

An interesting separable nonlinear system is the transformation of the form $\mathbf{s} = \mathbf{H\theta} + \mathbf{w}$ where $\mathbf{w}$ is a known vector. Although such a transformation is often called linear, strictly speaking it is not linear since in general it does not follow the superposition principle. The transformation is in fact affine. It can be separated into a linear transformation $\mathbf{H}_1(\boldsymbol{\theta}) = \mathbf{H\theta}$ and a nonlinear transformation $\mathbf{H}_2(\boldsymbol{\theta}) = \boldsymbol{\theta} + \mathbf{w}$. For example, $\mathbf{H}_2^{-1}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \mathbf{w}$.

The spatial and temporal orders of predistortion of nonideal blocks are presented in [9]. In the predistorter the corresponding blocks have a reverse order compared to the nonlinear model. In this way each nonideality can be compensated separately. In addition, the linear blocks are usually predistorted first with a weak training signal, which does not "see" the nonlinear effects, and the nonlinear blocks are predistorted with a strong training signal.

4. Nonlinear phenomena in beamforming

We consider the effects of nonlinear power amplifiers and nonideal mixers and oscillators in beamforming. The original results are presented in [10]. By creating beams and nulls in the antenna array system, we can increase the gain in the direction of wanted signals and decrease the gain in the direction of interference and noise. Nonlinear distortions cause, however, impairments in beamwidth, directivity, sidelobe level, null depth, and null direction. A simple way to make beamforming is to use RF phase shifters at each antenna element. The linear array consists of N identical radiators which are equally spaced by a distance d [11]. When λ is wavelength, the beam can be steered to the wanted direction θ_o by changing progressive phase shift $\alpha_z = (2\pi/\lambda)d\sin\theta_o$.

The array factor represents the far-field radiation pattern of an array of isotropically radiating elements. The ideal two dimensional array factor of a phased array is given by $F_a(\theta) = \sum_{n=1}^{N} e^{j(n-1)(kd\sin\theta - \alpha_z)}$, where $k = 2\pi/\lambda$ and θ is the elevation angle. For a linear array $F_a(\theta)$ is rotationally symmetric with respect to its axis. If the antenna element spacing d is too large compared to the wavelength λ, a second main lobe can appear in the antenna array factor. The extra main lobes are referred to as grating lobes. To prevent grating lobes we must select $d < \lambda/2$. Thus the element spacing has to be precise.

Amplitude distortions in the input of an antenna array or in the individual antenna element generate errors in the level of the main lobe. There is distortion in the direction of the main lobe if there is an error in the progressive phase shift α_z.The statements are demonstrated in Fig. 1. The simulated array factor in the middle in Fig. 1 verifies that the direction of the main lobe is changed if the progressive phase shift is not the same between all the antenna elements. The right-hand simulated array factor in Fig. 1 proves that the level of the main lobe is decreased, but the direction is not changed if the antenna elements have some amplitude distortion. In practice, there are always some differences between antenna elements even if they are designed to be similar. This may mean nonideal RF phase shifters and nonlinear power amplifiers. When the oscillator exhibits phase noise in the RF mixer at the input of the antenna array, the progressive phase shift remains the same between all the antenna elements, and thus there is no error in the array factor. Finally, even though the array factors seem to be only slightly distorted, the signal constellation can be severely distorted. As a consequence the bit error rate may be significantly degraded. Since the signal is distorted, this may have indirect effects also on synchronization and estimation in the receiver.

5. Conclusion

We summarized some system design solutions that can be used when we need to predistort linear and nonlinear phenomena in a transmitter, including power amplifiers, mixers, and oscillators. We showed that in a wideband transmitter the memory effects must be taken into account. Constant envelope signals do not solve the problems because the envelope may change due to the memory effects. Nonlinear blocks are not in general commutable. The spatial and temporal order of predistortion is crucial. The spatial order should be the reverse to that in the nonlinear model. The linear phenomena should usually be predistorted first and then the nonlinear phenomena. In beamforming the distortions were shown to distort the antenna pattern by changing the direction of the main lobe and by changing the gain of the antenna. This implies that ideally each branch of the multiantenna system must be predistorted separately for accurate results. The remaining challenges include the downconversion for each antenna element for the feedback signal from the amplifier output. Adaptivity can be done off-line if the changes are slow.

There is clearly a tradeoff between complexity and performance when there is a common predistorter or separate ones for all antenna elements.

6. Acknowledgments

This work was done in the TAFE project, partially funded by Tekes, and in the WIPLA project, funded by VTT Technical Research Centre of Finland.

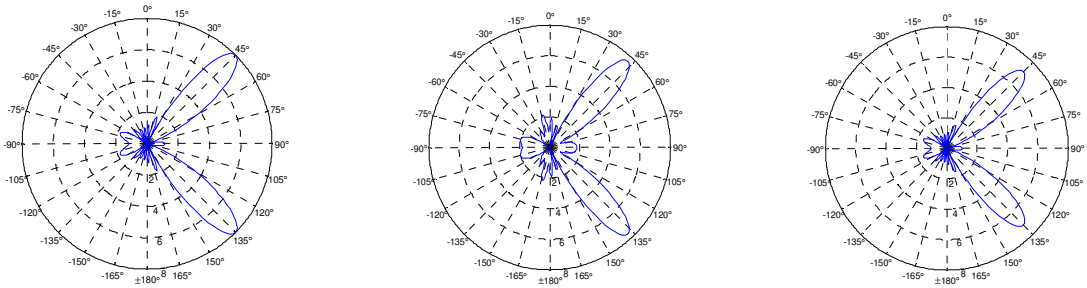

Figure 1. Ideal array factor (left) and distorted array factors when there is error in progressive phase shift (middle) and amplitude distortion in antenna elements (right).

7. References

1. M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology, and Techniques, Second Edition.* New York, Kluwer, 2000.

2. M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems, Second Edition*, Wiley, 2006.

3. M. Kiviranta, A. Mämmelä, S. Boumard, I. Moilanen, T. Sarkkinen, and T. Jämsä, "Real-time simulation of impairments in the analog parts of the transmitter-receiver," *VTC'08 Spring*, Stockholm, Sweden, 2005.

4. M. Kiviranta, A. Mämmelä, D. Cabric, D. Sobel, and R. W. Brodersen, "Constant envelope multicarrier modulation: Performance evaluation in AWGN and fading channel," *MILCOM'05*, Atlantic City, NJ, 2005.

5. A. Mämmelä, "Commutation in linear and nonlinear systems," *Frequenz*, **60**, May/June 2006, pp. 92-94.

6. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993.

7. S. Haykin, *Adaptive Filter Theory, Fourth Edition*, Prentice-Hall, 2001.

8. L. Ding et al. "A robust baseband predistorter contructed using memory polynomials," *IEEE Transactions on Communications*, **52**, January 2004, pp. 159-165.

9. M. Lasanen, A. Kotelba, A. Hekkala, P. Järvensivu, and A. Mämmelä, "Adaptive predistortion architecture for nonideal radio transmitter," *VTC'08 Spring*, Singapore, 2008.

10. M. Kiviranta, A. Mämmelä, H. Paaso, and M. Höyhtyä, "Effects of nonlinear analog parts in beamforming," unpublished.

11. J. Litva and T. K.-Y. Lo, *Digital Beamforming in Wireless Communications*, Artech House, 1996.