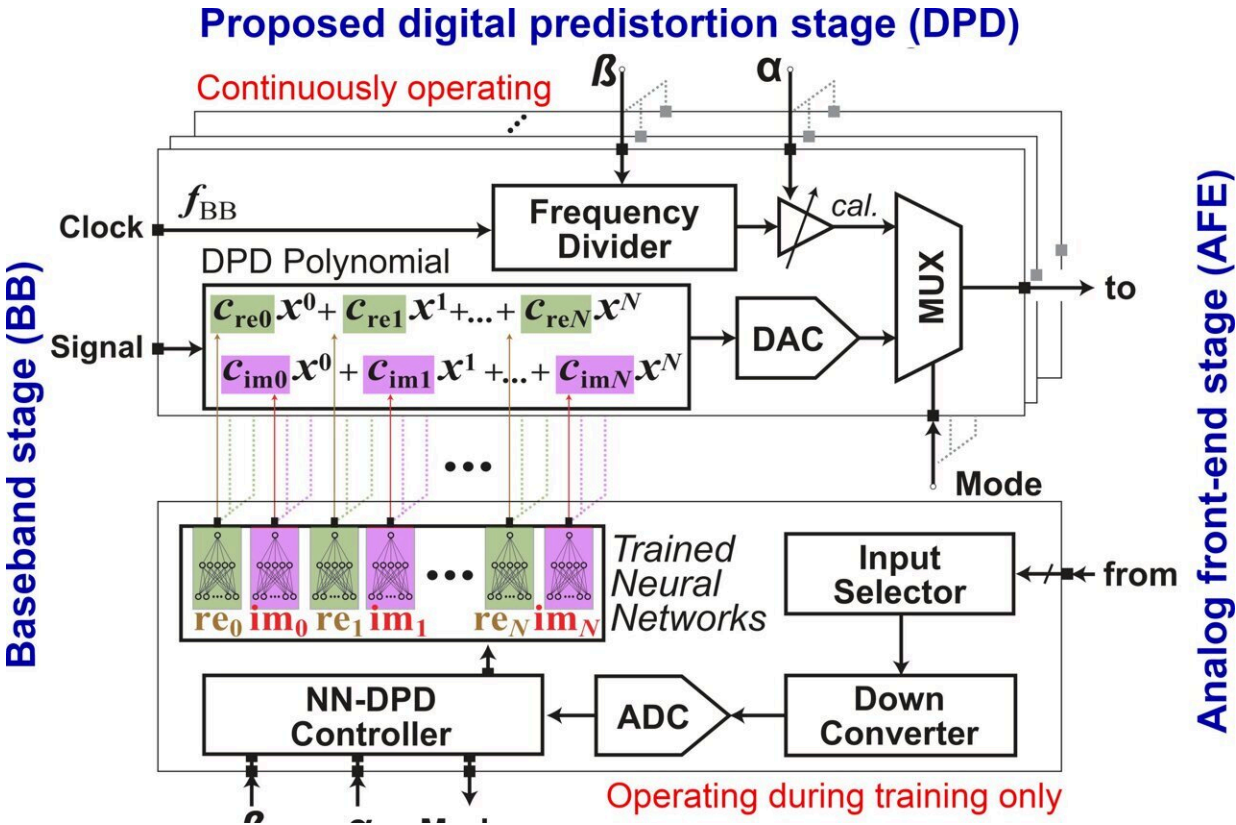# A new approach to using neural networks for low-power digital pre-distortion in mmWave systems

May 10 2024

Overview of the proposed architecture for digital pre-distortion (DPD), connected in between the baseband layer (BB) and the analog front-end (AFE). The top part operates continuously during transmission, whereas the bottom part

is only involved in updating the calibration coefficients. Credit: Ludovico Minati / Tokyo Tech

In a study published in the journal *IEICE Electronics Express*, researchers present a neural network digital pre-distortion (DPD) for mmWave RF-PAs.

In the world around us, a quiet but very important evolution has been taking place in engineering over the last decades. As technology evolves, it becomes increasingly clear that building devices that are physically as close as possible to being perfect is not always the right approach. That's because it often leads to designs that are very expensive, complex to build, and power-hungry.

Engineers, especially electronic engineers, have become skilled in using highly imperfect devices in ways that allow them to behave close enough to the ideal case to be successfully applicable. Historically, a well-known example is that of disk drives, where advances in control systems have made it possible to achieve incredible densities while using electromechanical hardware littered with imperfections, such as nonlinearities and instabilities of various kinds.

A similar problem has been emerging for radio communication systems. As the carrier frequencies keep increasing and channel packing becomes more and more dense, the requirements in terms of linearity for the radio-frequency power amplifiers (RF-PAs) used in telecommunication systems have become more stringent. Traditionally, the best linearity is provided by designs known as "Class A," which sacrifice great amounts of power to maintain operation in a region where transistors respond in

the most linear possible way.

On the other hand, highly energy-efficient designs are affected by nonlinearities that render them unstable without suitable correction. The situation has been getting worse because the modulation systems used by the latest cellular systems have a very high power ratio between the lowest- and highest-intensity symbols. Specific RF-PA types such as Doherty amplifiers are highly suitable and power-efficient, but their native non-linearity is not acceptable.

Over the last two decades, high-speed digital signal processing has become widely available, economical, and power-efficient, leading to the emergence of algorithms allowing the real-time correction of amplifier non-linearities through intentionally "distorting" the signal in a way that compensates the amplifier's physical response.

These algorithms have become collectively known as digital pre-distortion (DPD), and represent an evolution of earlier implementations of the same approach in the analog domain. Throughout the years, many types of DPD algorithms have been proposed, typically involving real-time feedback from the amplifier through a so-called "observation signal," and fairly intense calculations.

While this approach has been instrumental to the development of third- and fourth-generation cellular networks (3G, 4G), it falls short of the emerging requirements for fifth-generation (5G) networks, due to two reasons. First, dense antenna arrays are subject to significant disturbances between adjacent elements, known as cross-talking, making it difficult to obtain clean observation signals and causing instability.

The situation is made considerably worse by the use of ever-increasing frequencies. Second, dense arrays of antennas require very low-power solutions, and this is not compatible with the idea of complex processing

taking place for each individual element.

"We came up with a solution to this problem starting from two well-established mathematical facts. First, when a non-linearity is applied to a sinusoidal signal, it distorts it, leading to the appearance of new frequencies. Their intensity provides a sort of signature, that, if the non-linearity is a polynomial, is almost univocally associated with a set of coefficients. Second, multi-layer neural networks, of the early kinds, introduced decades ago, are universal function approximations, therefore, are capable of learning such an association, and inverting it," explains Prof. Ludovico Minati, leading inventor of the patent on which the study is based and formerly a specially-appointed associate professor at Tokyo Tech.

The most recent types of RF-PAs based on CMOS technology, even when they are heavily nonlinear, tend to have a relatively simple response, free from memory effects.

"This implies that the DPD problem can be reduced to finding the coefficients of a suitable polynomial, in a way that is quick and stable enough for real-world operation," explains Dr. Aravind Tharayil Narayanan, lead author of the study. Through a dedicated hardware architecture, the engineers at the Nano Sensing Unit of Tokyo Tech were able to implement a system that automatically determines the polynomial coefficients for DPD, based on a limited amount of data that could be acquired within the course of a few milliseconds.

Performing calibration in the "foreground," that is, one path at a time, reduces issues related to cross-talk and greatly simplifies the design. While there is no observation signal needed, the calibration can adjust itself to varying conditions through the inputs of additional signals, such as die temperature, power supply voltage, and settings of the phase shifters and couplers connecting the antenna. While standards

compliance may pose some limitations, the approach is in principle widely applicable.

"Because there is very limited processing happening in real-time, the hardware complexity is truly reduced to a minimum, and the power efficiency is maximized. Our results prove that this approach could in principle be sufficiently effective to support the most recent emerging standards. Another very convenient feature is that a considerable amount of hardware can be shared between elements, which is particularly convenient in dense array designs," says Prof. Hiroyuki Ito, head of the Nano Sensing Unit of TokyoTech where the technology was developed.

As a part of an industry-academia collaboration effort, the authors were able to test the concept on realistic, leading-edge hardware operating at 28 GHz provided by Fujitsu Limited, working in close collaboration with a team of engineers in the Product Planning Division of the Mobile System Business Unit. Future work will include large-scale implementation using dedicated ASIC designs, detailed standards compliance analysis and realistic benchmarking on the field under a variety of settings.

An international PCT application for the methodology and design has been filed.

Provided by Tokyo Institute of Technology