Machine Learning and Communication

Keywords: error correction, deep neural networks, auto-encoders

1 Context

Since Shannon's enunciation of the channel capacity formula in 1948, which defines the maximum rate at which a communication between two terminals can occur, the main research directions in error correction were about the design of capacity-achieving codes, i.e., codes that yield an arbitrarily small probability of error at large blocklengths [2]. As of today, four families of linear block codes proved to be capacity achieving, namely, Multi-level codes (Ten-Brink Wachsmann 1999), Turbo Codes (Glavieux and Bérou 1990), LDPC codes (Gallager 1960), and very lately, Polar codes (Arikan 2008). Most these codes have satisfactory performances and error scaling behaviour at large blocklengths n. However, for delay considerations in next generation communication system, the packets length is foreseen to be decreased to within very few symbols per frequency resource (in URLLC, 802.11p) in order to host IoT applications. This brings about the question of how to design codes which are close to optimal already at short blocklengths with a remarkably low decoding latency.

When most decoding algorithms in litterature are structured decoders (trellis based, Tanner graph based, ...), an interesting set of recent works (see [3]) showed that resorting to machine-learning based decoding algorithms can outperform structured algorithms in terms of latency and still perform close to optimal for short packets. These algorithms are based on *deep neural networks* (DNN) [1] which, with the advent of powerful computing units (GPUs) and thanks to their deep structure, proved lately to be efficient universal function approximators and thus, have a great appeal in error correction theory. The promising idea behind using neural networks in communication theory is to transform the block-based communication system into an end-to-end optimized autoencoder with a stochastic layer to model the channel statistics.

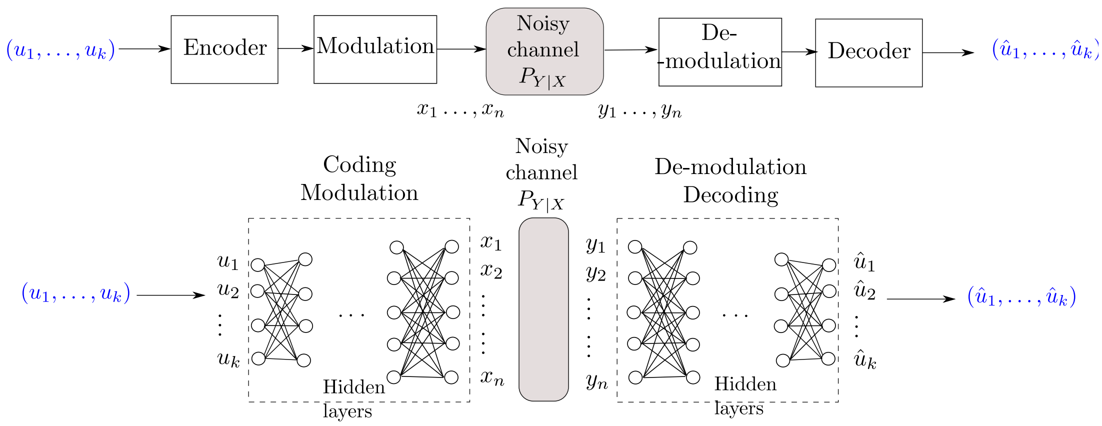

Figure 1: From functional separation to auto-encoders

Such novel machine learning based approaches question the traditional functional separation, however, they are lacking in maturity since the optimization of the different parameters is, to date, mainly heuristic and not formalized. Thus, most recent research results are aiming at formalizing the design of "deep communication systems" as in the present project.

2 Problem statement

The research assignment consists in implementing a linear block code (encoder and decoder) end-to-end resorting to deep neural networks and, more specifically, to auto-endoders.

- First develop a neural network based decoder "deep decoder" (already implemented in literature)
- Develop an autoencoder with end-to-end error correction coding and decoding
- Assess the weight distribution and design parameters of the deep network
- Investigate different channel modesl (AWGN, BEC, BSC)

Existing toolboxes on Matlab (or Python) should allow a fast implementation of the different steps of the research project, however, the optimization of the different parameters will be of crucial importance.

3 Misc

Disciplines: digital communications, machine learning

Tools: Matlab or Python (Theano), and LaTeX

Required curriculum: Supaéro, Master in aerospace engineering

4 Contact

Supervisor: meryem.benammar@isae-supaero.fr

References

[1] David Kriesel. *A Brief Introduction to Neural Networks.* 2007.

[2] Shu Lin and Daniel J. Costello. *Error Control Coding, Second Edition.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2004.

[3] Timothy J O'Shea and Jakob Hoydis. An introduction to machine learning communications systems. *arXiv preprint arXiv:1702.00832*, 2017.