Approximation speed of quantized *vs.* unquantized ReLU neural networks and beyond.

Antoine Gonon
Univ Lyon, EnsL, UCBL, CNRS, Inria, LIP, F-69342, LYON Cedex 07, France.
antoine.gonon@ens-lyon.fr

Neural networks are used to approximate functions with success in many applications. In line with the works [1, 2, 3], we are interested in understanding their approximation power in practice and in theory. Regarding practical applications, a key question is to be able to compare approximation properties of quantized versus unquantized neural networks. Another important question is to better understand non-trivial situations where neural networks can be expected (or not) to have better approximation properties than the best known approximation methods, quantized or not.

We are concerned with quantitatively characterizing the optimal polynomial speed $\gamma^{*\text{approx}}(\mathcal{C}|\Sigma)$ at which all functions of a set $\mathcal{C}$, subset of a pseudo-metric space $\mathcal{F}$, can be approximated by a sequence $\Sigma = (\Sigma_M)_{M\in\mathbb{N}}$ of sets Σ_M of "simpler" functions, such as ones that can be represented by polynomials of degree M, or ReLU neural networks with M non-zero parameters. We introduce a new property of the sequence Σ, called ∞-encodability, which forbids degenerate cases, where for example $\Sigma_1 = \mathcal{C}$ is already so rich that it yields unreasonable approximation rates. We show that:

(i) if Σ is ∞-encodable, then the Kolmogorov-Donoho complexity $\gamma^{*\text{encod}}(\mathcal{C})$, which measures the best polynomial asymptotic speed at which $\mathcal{C}$ can be *encoded* as binary sequences, and which is known for many classical functions sets such as balls of Sobolev spaces, bounds from above $\gamma^{*\text{approx}}(\mathcal{C}|\Sigma)$;

(ii) many sequences $\Sigma = (\Sigma_M)_{M\in\mathbb{N}}$ are ∞-encodable: when Σ_M contains M-terms linear combination of a dictionary, with boundedness conditions on the coefficients, or when Σ_M is Lipschitz-parameterized by some bounded set in finite dimension, the latter includes the case of ReLU neural networks for which we identify "simple" sufficient conditions on the considered architectures for this to hold;

(iii) when ∞-encodability is inherited from Lipschitz-parameterization, a simple quantization scheme turns Σ into a quantized sequence whose elements can be represented in a computer, attaining the same polynomial approximation speed as Σ on *every set* $\mathcal{C}$.

In light of point (ii), point (i) unifies and generalizes [2, Theorem VI.4][3, Theorem 5.24][4, Proposition 11].

Our framework applied to ReLU neural networks guarantees that uniformly quantized sparse ReLU networks with standard growth assumption on sparsity, depth and weight magnitudes, approximate every set of functions $\mathcal{C}$ with the same polynomial rate as their unquantized version. It also shows that approximation methods based on an ∞-encodable sequence defined with ReLU neural networks share a common upper-bound on approximation rates with other classical approximation methods also based on ∞-encodable sequences. As a consequence, given $\mathcal{C}$, if an ∞-encodable sequence is known such that $\gamma^{*\text{approx}}(\mathcal{C}|\Sigma) = \gamma^{*\text{encod}}(\mathcal{C})$, then no improved approximation rate can be hoped for using ReLU networks.

Joint work with: Nicolas Brisebarre, Rémi Gribonval, Elisa Riccietti.

References

[1] Ronald DeVore, Boris Hanin, and Guergana Petrova. Neural network approximation. *Acta Numer.*, 30:327–444, 2021.

[2] Dennis Elbrächter, Dmytro Perekrestenko, Philipp Grohs, and Helmut Bölcskei. Deep neural network approximation theory. *IEEE Trans. Inf. Theory*, 67(5):2581–2623, 2021.

[3] Philipp Grohs. Optimally sparse data representations. In *Harmonic and applied analysis*, Appl. Numer. Harmon. Anal., pages 199–248. Birkhäuser/Springer, Cham, 2015.

[4] Gérard Kerkyacharian and Dominique Picard. Entropy, universal coding, approximation, and bases properties. *Constr. Approx.*, 20(1):1–37, 2004.