# Ellipsoid Methods for Metric Entropy Rates Computations

**Thomas Allard**
ETH Zürich, Switzerland
tallard@ethz.ch

Historically, metric entropy has played a significant role in various domains such as non-linear approximation [1–3], statistical learning theory [4–6], and empirical processes theory [7, 8]. Recent advances in machine learning theory, and more specifically in deep learning, have led to renewed interest in the concept of metric entropy. Indeed, metric entropy is at the heart of the study of the approximation-theoretic properties and the statistical learning capabilities of deep neural networks [6, 9]. However, computing the precise value of the metric entropy of a given function class turns out to be notoriously difficult in general; exact expressions are available only in very few simple cases. For this reason, it has become common practice to resort to characterizations of the asymptotic behavior of metric entropy only. Even this more modest endeavor has turned out daunting in most cases. Consequently, a sizeable body of research exists in the literature; the survey [10] illustrates the multiplicity of methods employed. Through this survey, one is led to the insight that many of the standard methods implicitly rely on the computation of metric entropy for infinite-dimensional ellipsoids, highlighting the importance of developing a general method for this specific case. A first step toward such a general method was made in [11, 12]. Building on their result, we present a new method for the derivation of the asymptotic behavior of the metric entropy for infinite-dimensional ellipsoids. We further argue that our results provide a unified framework for the derivation of the metric entropy rates of a wide variety of function classes, such as Besov spaces, Modulation spaces, Sobolev spaces, and various classes of analytic functions, thereby retrieving and improving standard results.

[1] Lorentz, G. G. (1966). Metric entropy and approximation.

[2] Lorentz, G. G. (1966). Approximation of Functions, Athena Series. Selected Topics in Mathematics.

[3] Carl, B., & Stephani, I. (1990). Entropy, Compactness and the Approximation of Operators (Cambridge Tracts in Mathematics). Cambridge: Cambridge University Press.

[4] Haussler, D. (1992). Decision theoretic generalizations of the PAC model for neural net and other learning applications. Information and Computation, 100(1), 78-150.

[5] Wainwright, M. J. (2019). High-dimensional statistics: A non-asymptotic viewpoint (Vol. 48). Cambridge university press.

[6] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding Machine Learning: From Theory to Algorithms. Cambridge: Cambridge University Press.

[7] Dudley, R.M. (1984). A course on empirical processes. In: Hennequin, P.L. (eds) École d'Été de Probabilités de Saint-Flour XII - 1982. Lecture Notes in Mathematics, vol 1097. Springer, Berlin, Heidelberg.

[8] Pollard, D. (1990). Empirical processes: theory and applications. NSF-CBMS Regional Conf. Ser. Probab. Statist., 2: 86pp

[9] Elbrächter, D., Perekrestenko, D., Grohs, P., & Bölcskei, H. (2021). Deep neural network approximation theory. IEEE Transactions on Information Theory, 67(5), 2581-2623.

[10] Tikhomirov, V.M. (1993). -Entropy and -Capacity of Sets In Functional Spaces. In: Shiryayev, A.N. (eds) Selected Works of A. N. Kolmogorov. Mathematics and Its Applications, vol 27. Springer, Dordrecht.

[11] Luschgy, H., & Pages, G. (2004). Sharp asymptotics of the Kolmogorov entropy for Gaussian measures. Journal of Functional Analysis, 212(1), 89-120.

[12] Graf, S., & Luschgy, H. (2004). Sharp asymptotics of the metric entropy for ellipsoids. Journal of Complexity, 20(6), 876-882.

*Joint work with Prof. Dr. Helmut Bölcskei (ETH Zürich).*