# Ellipsoid Methods for Metric Entropy Computation

Thomas Allard
tallard@mins.ee.ethz.ch

Helmut Bölcskei
hboelcskei@ethz.ch

**Abstract**

We present a new methodology for the characterization of the metric entropy of infinite-dimensional ellipsoids with exponentially decaying semi-axes. This procedure does not rely on the explicit construction of coverings or packings and provides a unified framework for the derivation of the metric entropy of a wide variety of analytic function classes, such as periodic functions analytic on a strip, analytic functions bounded on a disk, and functions of exponential type. In each of these cases, our results improve upon the best known results in the literature.

## 1 Introduction

The concept of metric entropy has traditionally played a significant role in various domains of mathematics such as non-linear approximation theory [1, 2, 3], mathematical physics [4], statistical learning theory [5, 6, 7], and empirical process theory [8, 9]. Recent advances in machine learning theory, more specifically in deep learning, have led to renewed interest in methods for metric entropy computation. Indeed, metric entropy is at the heart of the approximation-theoretic characterization of deep neural networks [7, 10, 11]. 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. It has therefore become common practice to resort to characterizations of the asymptotic behavior as the covering ball radius approaches zero. Even this more modest endeavor has turned out daunting in most cases. A sizeable body of corresponding research exists in the literature [12, 13, 14]. The work of Donoho [15] constitutes a canonical example of the type of asymptotic results sought; specifically, it provides the exact expression for the leading term in the asymptotic expansion of the metric entropy of unit balls in Sobolev spaces.

The methods for characterizing the asymptotic behavior of metric entropy available in the literature are usually highly specific to the function class under consideration. The survey [16, Chapter 7] illustrates this point in the context of complex-analytic functions. An in-depth study of the variety of methods available in the literature leads to the insight that the underlying ellipsoidal structure of the function classes considered can be exploited to formulate a comprehensive methodology and en passant improve many of the best known results. This will, in fact, be the main goal of the present paper. A first step toward such a general method was made in [17, 18] by computing the metric entropy of infinite-dimensional ellipsoids with semi-axes of regularly varying (typically going to zero

The authors gratefully acknowledge support by the Lagrange Mathematics and Computing Research Center, Paris, France.

as the inverse of a polynomial) or slowly varying (typically going to zero as the inverse of a logarithm) lengths. This approach was applied in [17] to recover the above mentioned result by Donoho.

An analogous result for exponentially decaying lengths of the ellipsoid semi-axes does not seem to be available in the literature. We fill this gap and develop a general three-step procedure for the computation of the asymptotic behavior of the metric entropy of a wide variety of function classes. The versatility of our theory is illustrated by applying it to classes of analytic functions on a strip, analytic functions bounded on a disk, and entire functions of exponential type; in each of these cases, we improve upon the best known results in the literature.

These specific improvements along with the general methodology developed in the paper can find practical applications wherever the precise quantitative behavior of metric entropy, as opposed to its asymptotic scaling behavior only, is relevant. Examples include determining the minimum required sampling rate in A/D conversion [19] or the minimum size of deep neural networks needed to learn given function classes [10].

# 2 A Simple Example

The purpose of this section is to present a simple example which elucidates the main ideas the paper is built on. We start with the formal definition of metric entropy.

**Definition 1** (Metric entropy)**.** *Let $(\mathcal{X}, d)$ be a metric space and $\mathcal{K} \subseteq \mathcal{X}$ a compact set. An $\varepsilon$-covering of $\mathcal{K}$ with respect to the metric $d$ is a set $\{x_1, \dots, x_N\} \subseteq \mathcal{X}$ such that for each $x \in \mathcal{K}$, there exists an $i \in \{1, \dots, N\}$ so that $d(x, x_i) \leq \varepsilon$. The $\varepsilon$-covering number $N(\varepsilon; \mathcal{K}, d)$ is the cardinality of a smallest such $\varepsilon$-covering. The metric entropy of $\mathcal{K}$ is*

$$H(\varepsilon; \mathcal{K}, d) \coloneqq \log_2 N(\varepsilon; \mathcal{K}, d).$$

The example we consider is the most basic of examples, namely the computation of the metric entropy of the unit interval $[0, 1]$ equipped with the metric induced by the absolute value $|\cdot|$. Although this example has been studied exhaustively in the literature, the vantage point we take, namely in terms of binary expansions, seems novel. Before developing our idea, we briefly recall the main arguments usually employed to establish the result, following the exposition in [6]. Specifically, one wishes to prove that

$$N(\varepsilon; [0, 1], |\cdot|) = \left\lceil \frac{1}{2\varepsilon} \right\rceil. \tag{1}$$

This is typically done by establishing

$$N(\varepsilon; [0, 1], |\cdot|) \leq \left\lceil \frac{1}{2\varepsilon} \right\rceil \quad \text{and} \quad N(\varepsilon; [0, 1], |\cdot|) \geq \left\lceil \frac{1}{2\varepsilon} \right\rceil, \tag{2}$$

which, when combined, yield the desired result. The upper bound in (2) is obtained by constructing an explicit $\varepsilon$-covering of $[0, 1]$, specifically $\{x_i\}_{i=1}^N$ with $x_i \coloneqq 2\,(i - 1/2)\,\varepsilon$, for $i = 1, \dots, N$, and $N = \lceil 1/(2\varepsilon) \rceil$. It follows by inspection that, for every $x \in [0, 1]$, there exists an $i$ such that $|x - x_i| \leq \varepsilon$. Fig. 1 depicts an example.

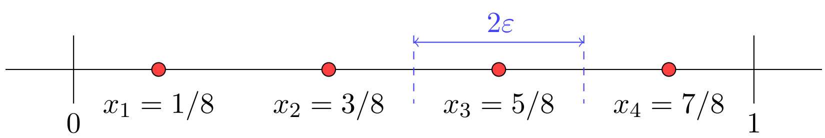

Figure 1: Covering of the unit interval for $\varepsilon = 1/8$ and $N = 4$. The interval $[0, 1]$ is divided into four subintervals of length $2\varepsilon$ centered at $x_1$, $x_2$, $x_3$, and $x_4$, one of which is indicated in blue.

The lower bound in (2) is found through a volume argument as follows. If the interval $[0, 1]$ can be covered with $N(\varepsilon; [0, 1], |\cdot|)$ intervals of size $2\varepsilon$, we must have

$$2\varepsilon N(\varepsilon; [0, 1], |\cdot|) \geq 1.$$

Since $N(\varepsilon; [0, 1], |\cdot|)$ is an integer, this can be refined to

$$N(\varepsilon; [0, 1], |\cdot|) \geq \left\lceil \frac{1}{2\varepsilon} \right\rceil,$$

which is the desired lower bound, thereby finalizing the proof of (1). As a byproduct, it has also been established that the $\varepsilon$-covering $\{x_1, \dots, x_N\}$ constructed above is optimal, in the sense that there is no $\varepsilon$-covering with fewer elements; and we have identified a natural mapping that takes every $x \in [0, 1]$ to its nearest $x_i$.

We now view the derivation of (1) just presented through the prism of binary expansions of real numbers. A binary expansion of $x \in [0, 1]$ is a sequence $\{b_n\}_{n\in\mathbb{N}^*} \in \{0, 1\}^{\mathbb{N}^*}$ of digits, indexed by the positive integers $\mathbb{N}^*$, satisfying

$$x = \sum_{n\in\mathbb{N}^*} b_n 2^{-n}. \tag{3}$$

Note that binary expansions are unique up to the trivial 2-ambiguity that identifies, e.g., $x = 1/2$ with the sequence $b = 10000...$ and the alternative $\tilde{b} = 01111....$ For simplicity of exposition, we assume that $\varepsilon$ is an inverse power of 2, i.e., $\varepsilon = 2^{-k-1}$, for some $k \in \mathbb{N}$, and hence $\log(1/(2\varepsilon)) = k$. For this choice, the $\varepsilon$-covering constructed above divides the interval $[0, 1]$ into the $N = 2^k$ sub-intervals $I_i \coloneqq [(i-1)2^{-k}, i2^{-k}]$, $i = 1, \dots, N$, with the interesting property that, for given $i$, the binary expansions of the points in $I_i$ coincide in their first $k$ digits[1]. Put differently, one can view the construction of the optimal covering above, together with its corresponding natural mapping, as follows: Given a number $x \in [0, 1]$, (i) write its binary expansion, (ii) discard all the digits after the $k$-th one, and (iii) append a 1 to the remaining digits to obtain the covering element $x$ is mapped to. This three-step procedure can be illustrated for $k = 2$ (also see Fig. 2) as follows:

$$x \xrightarrow{(i)} 0.\underbrace{b_1\, b_2}_{\text{first } k \text{ digits}}\ b_3\, b_4\, b_5\, b_6 \dots \xrightarrow{(ii)} 0.\, b_1\, b_2 \xrightarrow{(iii)} x_i \coloneqq_2 0.\, b_1\, b_2\, 1.$$

We can now formally view (3) as an expansion of the objects in the set under consideration into a series that exhibits exponential decay, according to $2^{-n}$ here. The corresponding expansion coefficients, the digits $b_n$, can take on values from the set $\{0, 1\}$. Both the construction of the covering ball centers $x_i$ and the mapping of the elements in $[0, 1]$ to the closest covering ball center are effected by thresholding the series expansion after $k = \log(1/(2\varepsilon))$ terms. Equivalently, this means that the approximation error decays exponentially in the number of terms $k$ retained.

[1] The interval boundary $i2^{-k}$ is identified with the alternative binary expansion $\tilde{b}$.

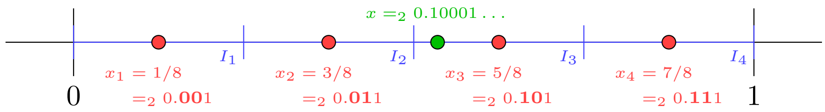


Figure 2: Covering of the unit interval, for $k = 2$.

The present paper builds on a vast generalization of this viewpoint. More precisely, given a class of functions, we follow an analogous three-step procedure by first identifying an expansion of the elements $x$ in the class in terms of an exponentially decaying coefficient sequence $\{x_n\}_{n\in\mathbb{N}^*}$. This exponential decay property is formally expressed through a condition of the form

$$\left(\sum_{n\in\mathbb{N}^*} |x_n/\mu_n|^p\right)^{1/p} \leq 1, \tag{4}$$

where $p \in [1, \infty)$ and $\{\mu_n\}_{n\in\mathbb{N}^*}$ is an exponentially decaying sequence. For $p = \infty$, we use

$$\sup_{n\in\mathbb{N}^*} |x_n/\mu_n| \leq 1. \tag{5}$$

Conditions of the form (4) or (5) define infinite-dimensional ellipsoids with semi-axes $\{\mu_n\}_{n\in\mathbb{N}^*}$. In the binary expansion case considered above, we would have (5) with $x_n = b_n 2^{-n}$, $\mu_n = 2^{-n}$, for $n \in \mathbb{N}^*$, thanks to $b_n \in \{0, 1\}$. In the second step, we reduce the problem to the analysis of finite-dimensional ellipsoids by truncating the expansion. The choice of the number of terms to be retained is informed by the exponential decay of the ellipsoid semi-axes and should depend logarithmically on $1/\varepsilon$; this will, in turn, yield an approximation error on the order of $\varepsilon$. In the third step, the metric entropy of the finite-dimensional ellipsoid obtained through truncation is derived, specifically through volume arguments. This would not be possible in the infinite-dimensional case where volumes are infinite. In summary, we are hence spared from the often tedious task of explicitly constructing coverings and packings as done in traditional approaches (see e.g. [16, Chapter 7] or [2, Chapter 10]). In fact, when the coefficients $x_n$ are real- or complex-valued, as will be the case in general, it is actually unclear how to infer explicit coverings or packings within our approach. In the binary expansion example considered above this was made possible by the finite-alphabet property of the $x_n = b_n 2^{-n}$. In summary, while we have a very general methodology which allows to improve upon best-known metric entropy results, our approach is not constructive in the sense of identifying optimal coverings and packings.

We conclude this section by defining notation and terminology. $\mathbb{N}$ and $\mathbb{N}^*$ stand for the set of natural numbers including and, respectively, excluding zero, $\mathbb{Z}$ is the set of integers. $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and, respectively, complex numbers. We use the generic notation $\mathbb{K}$ to mean that a statement applies both for $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$. It will further be convenient to introduce the notation

$$\sigma_{\mathbb{K}}(d) \coloneqq \begin{cases} d, & \text{if } \mathbb{K} = \mathbb{R}, \\ 2d, & \text{if } \mathbb{K} = \mathbb{C}. \end{cases} \tag{6}$$

For $d \in \mathbb{N}^*$, we denote the geometric mean of the set $\{\mu_1, \ldots, \mu_d\}$ of positive real numbers by

$$\bar{\mu}_d \coloneqq \prod_{n=1}^{d} \mu_n^{1/d}.$$

We refer to the closed ball with center $x$, of radius $r$, and with respect to the metric $\rho$ by $B_\rho(x;r)$. The subscript $\rho$ will be omitted when there is no room for confusion. When the metric $\rho$ is induced by the usual $p$-norm, we simply write $\mathcal{B}_p$ for the unit ball $B_\rho(0\,;1)$. $\log(\cdot)$ stands for the logarithm to base 2 and $\ln(\cdot)$ is the natural logarithm. For $k \in \mathbb{N}^*$, we denote the $k$-fold iterated logarithm by

$$\log^{(k)}(\cdot) := \underbrace{\log \circ \cdots \circ \log}_{k \text{ times}}(\cdot).$$

Finally, when comparing the asymptotic behavior of the functions $f$ and $g$ as $x \to \ell$, with $\ell \in \mathbb{R} \cup \{-\infty, \infty\}$, we use the notation

$$f = o_{x\to\ell}(g) \iff \lim_{x\to\ell} \frac{f(x)}{g(x)} = 0 \quad \text{and} \quad f = \mathcal{O}_{x\to\ell}(g) \iff \lim_{x\to\ell} \left|\frac{f(x)}{g(x)}\right| \leq C,$$

for some constant $C > 0$. We further indicate asymptotic equivalence according to

$$f \sim_{x\to\ell} g \iff \lim_{x\to\ell} \frac{f(x)}{g(x)} = 1.$$

# 3 Metric Entropy of Ellipsoids

As mentioned earlier, we shall characterize the metric entropy of infinite-dimensional ellipsoids by reduction to finite-dimensional ellipsoids while controlling the resulting approximation error. We begin by formally introducing finite-dimensional ellipsoids.

**Definition 2** (Finite-dimensional ellipsoids)**.** *Let $d \in \mathbb{N}^*$ and $p \in [1, \infty]$. To a given set $\{\mu_1, \dots, \mu_d\}$ of strictly positive real numbers, we associate the $p$-ellipsoid norm $\|\cdot\|_{p,\mu}$ on $\mathbb{K}^d$, defined as*

$$\|\cdot\|_{p,\mu} \colon x \in \mathbb{K}^d \mapsto \begin{cases} \left(\sum_{n=1}^{d} |x_n/\mu_n|^p\right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \max_{1\leq n\leq d} |x_n/\mu_n|, & \text{if } p = \infty. \end{cases}$$

*The finite-dimensional $p$-ellipsoid is the unit ball in $\mathbb{K}^d$ with respect to the norm $\|\cdot\|_{p,\mu}$ and is denoted by*

$$\mathcal{E}_p^d\Big(\{\mu_n\}_{n=1}^d\Big) := \Big\{x \in \mathbb{K}^d \mid \|x\|_{p,\mu} \leq 1\Big\}.$$

*The numbers $\{\mu_1, \dots, \mu_d\}$ are referred to as the semi-axes of the ellipsoid $\mathcal{E}_p^d(\{\mu_n\}_{n=1}^d)$. We simply write $\mathcal{E}_p^d$ when the choice of semi-axes is clear from the context. For simplicity of exposition, and without loss of generality, we assume that the semi-axes $\{\mu_1, \dots, \mu_d\}$ are arranged in non-increasing order, i.e.,*

$$\mu_1 \geq \cdots \geq \mu_d > 0.$$

As already mentioned, our approach is based on volume arguments. Specifically, we shall exploit the fact that the total volume occupied by the covering balls of a given set is larger than or equal to the volume of the set. Conversely, when packing balls into a set, the total volume occupied by these balls must be less than or equal to the volume of a slightly augmented version of the set. This intuition is formalized in [6, Lemma 5.7] for $x \in \mathbb{R}^d$ (the case $x \in \mathbb{C}^d$ can be handled similarly by identifying $\mathbb{C}$ with $\mathbb{R}^2$), which we restate here for completeness.

**Lemma 3** (Volume estimates)**.** *Let $d \in \mathbb{N}^*$ and fix $\varepsilon > 0$. Consider the norms $\|\cdot\|$ and $\|\cdot\|'$ on $\mathbb{K}^d$ and let $\mathcal{B}$ and $\mathcal{B}'$ be their respective unit balls. Then, the $\varepsilon$-covering number $N(\varepsilon;\mathcal{B},\|\cdot\|')$ satisfies*

$$\frac{\mathrm{vol}_{\mathbb{K}}(\mathcal{B})}{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}')} \le N\big(\varepsilon;\mathcal{B},\|\cdot\|'\big)\,\varepsilon^{\sigma_{\mathbb{K}}(d)} \le 2^{\sigma_{\mathbb{K}}(d)}\frac{\mathrm{vol}_{\mathbb{K}}\big(\mathcal{B}+\frac{\varepsilon}{2}\mathcal{B}'\big)}{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}')}. \tag{7}$$

Volume ratios akin to those in (7) will appear regularly in our analyses, in particular with $\|\cdot\|$ and $\|\cdot\|'$ given by $p-$ and $q-$norms, for some $p,q \in [1,\infty]$. It will hence turn out convenient to introduce the notation

$$V_{p,q}^{\mathbb{K},d} := \frac{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}_p)}{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}_q)}. \tag{8}$$

Such volume ratios have been studied extensively in the literature, see e.g. [20] and [21, Chapter 3] for representative references in the context of metric entropy.

Now, observe that, by taking logarithms in (7), one can readily deduce the metric entropy scaling behavior of a $d$-dimensional unit ball $\mathcal{B}$ according to

$$H\big(\varepsilon;\mathcal{B},\|\cdot\|'\big) \sim_{\varepsilon\to 0} \sigma_{\mathbb{K}}(d)\log\big(\varepsilon^{-1}\big). \tag{9}$$

Intuitively, this scaling quantifies that, in order to encode an element of $\mathcal{B}$ with precision $\varepsilon$, one needs to quantize each of its components using $\log(\varepsilon^{-1})$ bits.

Recalling that finite-dimensional $p$-ellipsoids are unit balls with respect to $\|\cdot\|_{p,\mu}$-norms, it follows that their metric entropy behavior is, in principle, characterized by (9). However, bringing out the dependency on the semi-axes $\{\mu_1,\dots,\mu_d\}$ and getting good constants in lower and upper bounds on metric entropy, ideally even sharp results, requires significantly more work. The following two theorems, whose proofs have been relegated to Appendix C.1, address these issues and constitute our main results in this context.

**Theorem 4.** *Let $d \in \mathbb{N}^*$ and $p,q \in [1,\infty]$. Then, there exists a positive real constant $\underline{\kappa}$ independent of $d$ such that, for all $\varepsilon > 0$, the covering number of the ellipsoid $\mathcal{E}_p^d$ in $\|\cdot\|_q$-norm satisfies*

$$N\Big(\varepsilon;\mathcal{E}_p^d,\|\cdot\|_q\Big)^{\frac{1}{\sigma_{\mathbb{K}}(d)}}\,\varepsilon \ge \underline{\kappa}\,\sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\bar{\mu}_d,$$

*where $\bar{\mu}_d$ is the geometric mean of the semi-axes of $\mathcal{E}_p^d$.*

**Theorem 5.** *Let $d \in \mathbb{N}^*$, let $p,q \in [1,\infty]$, and assume that*

$$H\Big(\varepsilon;\mathcal{E}_p^d,\|\cdot\|_q\Big) \ge 2\,\sigma_{\mathbb{K}}(d), \quad \text{for all } \varepsilon \in [0,2\mu_d], \tag{10}$$

*where $\mu_d$ is the smallest semi-axis of $\mathcal{E}_p^d$. Then, there exists a positive real constant $\overline{\kappa}$ independent of $d$ such that, for all $\varepsilon \in [0,2\mu_d]$, the covering number of the ellipsoid $\mathcal{E}_p^d$ in $\|\cdot\|_q$-norm satisfies*

$$N\Big(\varepsilon;\mathcal{E}_p^d,\|\cdot\|_q\Big)^{\frac{1}{\sigma_{\mathbb{K}}(d)}}\,\varepsilon \le \overline{\kappa}\,\sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\bar{\mu}_d,$$

*where $\bar{\mu}_d$ is the geometric mean of the semi-axes of $\mathcal{E}_p^d$.*

Obviously, we must have

$$\overline{\kappa} \geq \underline{\kappa}. \tag{11}$$

Note that the upper bound in Theorem 5 requires the additional assumption $\varepsilon \in [0, 2\mu_d]$. This can be explained as follows. If $\varepsilon$ is large compared to $\mu_d$ (recall that the semi-axes are ordered), the problem of covering the $d$-dimensional ellipsoid can be reduced to that of covering a lower-dimensional version thereof. This insight is illustrated, for $d = 2$, in Figure 3.

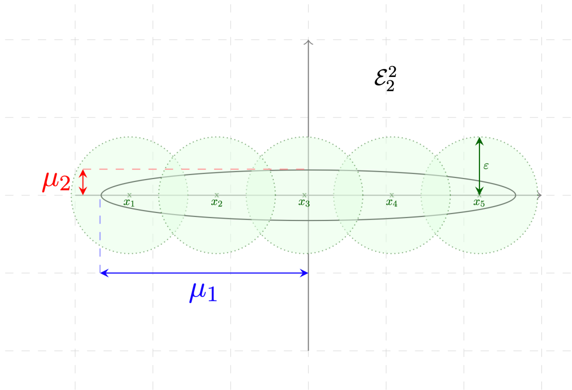


Figure 3: Reduction of a two-dimensional covering problem to one dimension by lining up $\varepsilon$-balls, with $\varepsilon > 2\mu_2$, along the $x$-axis.

Equipped with Theorems 4 and 5, we are ready to study infinite-dimensional ellipsoids.

**Definition 6** (Infinite-dimensional ellipsoids)**.** *Let $p \in [1, \infty]$. To a given bounded sequence $\{\mu_n\}_{n\in\mathbb{N}^*}$ of positive real numbers, we associate the ellipsoid norm $\|\cdot\|_{p,\mu}$ on $\ell^p(\mathbb{N}^*; \mathbb{K})$, defined as*

$$\|\cdot\|_{p,\mu} \colon x \in \ell^p(\mathbb{N}^*; \mathbb{K}) \mapsto \begin{cases} \left(\sum_{n\in\mathbb{N}^*} |x_n/\mu_n|^p\right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \sup_{n\in\mathbb{N}^*} |x_n/\mu_n|, & \text{if } p = \infty. \end{cases}$$

*The infinite-dimensional $p$-ellipsoid is the unit ball in $\ell^p(\mathbb{N}^*; \mathbb{K})$ with respect to the norm $\|\cdot\|_{p,\mu}$ and is denoted by*

$$\mathcal{E}_p(\{\mu_n\}_{n\in\mathbb{N}^*}) \coloneqq \{x \in \ell^p(\mathbb{N}^*; \mathbb{K}) \mid \|x\|_{p,\mu} \leq 1\}.$$

*The elements of $\{\mu_n\}_{n\in\mathbb{N}^*}$ are referred to as the semi-axes of the ellipsoid $\mathcal{E}_p(\{\mu_n\}_{n\in\mathbb{N}^*})$. We simply write $\mathcal{E}_p$ when the choice of the semi-axes is clear from the context. For simplicity of exposition, and without loss of generality, we assume that the semi-axes $\{\mu_n\}_{n\in\mathbb{N}^*}$ are arranged in non-increasing order, i.e.,*

$$\mu_1 \geq \dots \geq \mu_d \geq \dots > 0.$$

Note that $x \in \ell^p(\mathbb{N}^*; \mathbb{K})$ does not necessarily imply $\|x\|_{p,\mu} < \infty$. However, the converse is true, namely every $x$ with finite ellipsoid norm $\|x\|_{p,\mu}$ must be in $\ell^p(\mathbb{N}^*; \mathbb{K})$. Consequently, defining the ellipsoid $\mathcal{E}_p(\{\mu_n\}_{n\in\mathbb{N}^*})$ as a subset of $\ell^p(\mathbb{N}^*; \mathbb{K})$ is without loss of generality.

Restraining the semi-axes to be positive does not come at a loss of generality as the dimensions corresponding to semi-axes that are equal to zero can simply be removed from consideration. In particular, when the semi-axes equal zero beyond a certain index, we are back to the case of finite-dimensional ellipsoids. As announced in the introduction, we restrict ourselves to ellipsoids with semi-axes of exponential decay formalized as follows.

**Definition 7** (Exponentially decaying semi-axes)**.** *We say that a smooth (i.e., infinitely differentiable) function*

$$\psi\colon (0,\infty) \to \mathbb{R}$$

*is a decay rate function with parameter* $t^* \geq 0$ *if*

$$\psi((t^*,\infty)) = (0,\infty) \quad \textit{and} \quad t \mapsto \frac{\psi(t)}{t} \textit{ is non-decreasing on } (t^*,\infty). \tag{12}$$

*Further, the ellipsoid* $\mathcal{E}_p(\{\mu_n\}_{n\in\mathbb{N}^*})$ *has* $\psi$*-exponentially decaying semi-axes if*

$$\mu_n \coloneqq c_0 \exp\{-\ln(2)\,\psi(n)\}, \quad \textit{for all } n \in \mathbb{N}^* \textit{ and some } c_0 > 0. \tag{13}$$

Concrete examples of decay rate functions $\psi$ will be considered in Corollaries 12-14. We note that every decay rate function $\psi$ must be invertible on $(t^*,\infty)$; see Lemma 31 in Appendix B for a formal statement and proof of this fact. The inverse

$$\psi^{(-1)}\colon (0,\infty) \to (t^*,\infty)$$

of $\psi$ plays a central role in our main result Theorem 9. Before stating this result, we need to introduce two auxiliary functions.

**Definition 8** ($\psi$-average and $\psi$-difference functions)**.** *Let* $\psi$ *be a decay rate function. We define the* $\psi$*-average function according to*

$$\delta(d) \coloneqq \frac{1}{d}\sum_{n=1}^{d}(\psi(d) - \psi(n)), \quad \textit{for all } d \geq 1,$$

*and the* $\psi$*-difference function as*

$$\zeta(d) \coloneqq \psi(d) - \psi(d-1), \quad \textit{for all } d \geq 2.$$

The main properties of the $\psi$-average and the $\psi$-difference function are summarized in Appendix B.

The next theorem constitutes the main result of the paper and characterizes the asymptotic behavior of the metric entropy of infinite-dimensional ellipsoids with $\psi$-exponentially decaying semi-axes.

**Theorem 9** (Metric entropy of infinite-dimensional ellipsoids)**.** *Let* $p, q \in [1,\infty]$ *and let* $\psi$ *be a decay rate function. There exists an integer-valued function* $d_\varepsilon$ *which can be written in the form*

$$d_\varepsilon = \psi^{(-1)}\left[\log(\varepsilon^{-1}) + \log(c_0)\right] + O_{\varepsilon\to 0}(1), \tag{14}$$

*with* $c_0$ *as per (13), such that the metric entropy of the infinite-dimensional ellipsoid* $\mathcal{E}_p$ *with* $\psi$*-exponentially decaying semi-axes* $\{\mu_n\}_{n\in\mathbb{N}^*}$ *satisfies*

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) = \sigma_{\mathbb{K}}(d_\varepsilon)\left\{\delta(d_\varepsilon) + \left(\frac{1}{q} - \frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon)) + O_{d_\varepsilon\to\infty}(\zeta(d_\varepsilon))\right\}, \tag{15}$$

*where* $\delta(\cdot)$ *and* $\zeta(\cdot)$ *are the* $\psi$*-average and the* $\psi$*-difference function, respectively.*

We note that the metric entropy of ellipsoids with exponentially decaying semi-axes does not seem to have been characterized before in the literature. The proof of Theorem 9 follows steps two and three of the three-step procedure described in the introduction. Concretely, given an infinite-dimensional ellipsoid $\mathcal{E}_p(\{\mu_n\}_{n\in\mathbb{N}^*})$ to be covered by balls of radius $\varepsilon > 0$, the problem is reduced to covering a finite-dimensional ellipsoid of dimension $d_\varepsilon$. One then applies Theorems 4 and 5 to this $d_\varepsilon$-dimensional ellipsoid to obtain (15). The choice of $d_\varepsilon$ is informed by the idea that semi-axes with indices beyond $d_\varepsilon$ should asymptotically be on the order of $\varepsilon$, i.e.,

$$\mu_{d_\varepsilon} \sim_{\varepsilon\to 0} \varepsilon,$$

or, equivalently using $\mu_{d_\varepsilon} = c_0 \exp\{-\ln(2)\,\psi(d_\varepsilon)\}$, we want

$$d_\varepsilon \sim_{\varepsilon\to 0} \psi^{(-1)}\big[\log\big(\varepsilon^{-1}\big) + \log(c_0)\big],$$

which is (14).

*Proof.* We first introduce, for all $\varepsilon > 0$, the effective dimension of the ellipsoid $\mathcal{E}_p$ with respect to $\|\cdot\|_q$ as

$$d_\varepsilon := \min\Big\{k \in \mathbb{N}^* \mid N(\varepsilon;\mathcal{E}_p,\|\cdot\|_q)^{\frac{1}{\sigma_{\mathbb{K}}(k)}}\,\mu_k \le \overline{\kappa}\,\sigma_{\mathbb{K}}(k)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\bar{\mu}_k\Big\}, \tag{16}$$

where $\overline{\kappa}$ is the constant in the statement of Theorem 5. It is formally established in Lemma 47 (see Appendix C.6.18) that the effective dimension is well-defined for all $\varepsilon > 0$ and approaches infinity as $\varepsilon \to 0$. Taking the logarithm in (16) yields

$$\begin{aligned}
&\sigma_{\mathbb{K}}(d_\varepsilon - 1)\Bigg\{\log(\overline{\kappa}) + \left(\frac{1}{q}-\frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon-1)) \qquad (17)\\
&\qquad\qquad + \frac{1}{d_\varepsilon - 1}\sum_{n=1}^{d_\varepsilon-1}\log\left\{\frac{\mu_n}{\mu_{d_\varepsilon-1}}\right\}\Bigg\}\\
&< H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q)\\
&\le \sigma_{\mathbb{K}}(d_\varepsilon)\Bigg\{\log(\overline{\kappa}) + \left(\frac{1}{q}-\frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon)) + \frac{1}{d_\varepsilon}\sum_{n=1}^{d_\varepsilon}\log\left\{\frac{\mu_n}{\mu_{d_\varepsilon}}\right\}\Bigg\}. \qquad (18)
\end{aligned}$$

Using

$$\delta(d) = \frac{1}{d}\sum_{n=1}^{d}(\psi(d)-\psi(n)) \overset{(13)}{=} \frac{1}{d}\sum_{n=1}^{d}\log\left\{\frac{\mu_n}{\mu_d}\right\},$$

and identifying the term $\log(\overline{\kappa})$ as $O_{d_\varepsilon\to\infty}(1)$, we can rewrite the upper bound in (17)-(18) according to

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) \le \sigma_{\mathbb{K}}(d_\varepsilon)\left\{\delta(d_\varepsilon) + \left(\frac{1}{q}-\frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon)) + O_{d_\varepsilon\to\infty}(1)\right\}. \tag{19}$$

The lower bound in (17)-(18) can equivalently be expressed as

$$\begin{aligned}
&\sigma_{\mathbb{K}}(d_\varepsilon - 1)\left\{\delta(d_\varepsilon - 1) + \left(\frac{1}{q}-\frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon - 1)) + O_{d_\varepsilon\to\infty}(1)\right\}\\
&= \sigma_{\mathbb{K}}(d_\varepsilon)\left\{\delta(d_\varepsilon) + \left(\frac{1}{q}-\frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon)) - \zeta(d_\varepsilon)\,(1 - 1/d_\varepsilon) + O_{d_\varepsilon\to\infty}(1)\right\},
\end{aligned}$$

where we used

$$\sigma_{\mathbb{K}}(d_\varepsilon - 1)\log(\sigma_{\mathbb{K}}(d_\varepsilon - 1)) = \sigma_{\mathbb{K}}(d_\varepsilon)(\log(\sigma_{\mathbb{K}}(d_\varepsilon)) + O_{d_\varepsilon\to\infty}(1))$$

together with the identity

$$\frac{\sigma_{\mathbb{K}}(d)\,\delta(d) - \sigma_{\mathbb{K}}(d-1)\,\delta(d-1)}{\sigma_{\mathbb{K}}(d)} = \zeta(d)\,(1-1/d), \quad \text{for all } d \geq 2,$$

established in Lemma 32 (see Appendix B). With

$$1 = O_{d\to\infty}(\zeta(d)), \quad \text{and therefore} \quad d = O_{d\to\infty}(d\,\zeta(d)), \tag{20}$$

which follows from Lemma 33 (see Appendix B), we can further rewrite the lower bound in (17)-(18) according to

$$\sigma_{\mathbb{K}}(d_\varepsilon)\left\{\delta(d_\varepsilon) + \left(\frac{1}{q} - \frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon)) + O_{d_\varepsilon\to\infty}(\zeta(d_\varepsilon))\right\}. \tag{21}$$

Combining the upper bound (19) with the lower bound (21) and using once more (20), we obtain the desired final result according to

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) = \sigma_{\mathbb{K}}(d_\varepsilon)\left\{\delta(d_\varepsilon) + \left(\frac{1}{q} - \frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d_\varepsilon)) + O_{d_\varepsilon\to\infty}(\zeta(d_\varepsilon))\right\}.$$

It remains to establish that $d_\varepsilon$ can, indeed, be expressed as in (14). This will be accomplished with the help of the following two lemmata.

**Lemma 10.** *There exists a positive real constant $\underline{\mathfrak{c}} \leq 1$ independent of $\varepsilon$ such that*

$$\varepsilon \geq \underline{\mathfrak{c}}\,\mu_{d_\varepsilon}, \quad \textit{for all } \varepsilon > 0,$$

*where $d_\varepsilon$ is as defined in (16).*

*Proof.* See Appendix C.6.1. □

**Lemma 11.** *There exist a positive real number $\varepsilon^*$ and a constant $\overline{\mathfrak{c}} \geq 1$ such that, for all $\varepsilon \in (0,\varepsilon^*)$,*

$$\varepsilon \leq \overline{\mathfrak{c}}\,\mu_{d_\varepsilon - 1},$$

*where $d_\varepsilon$ is as defined in (16).*

*Proof.* See Appendix C.6.2. □

Let $\underline{\mathfrak{c}}$, $\overline{\mathfrak{c}}$, and $\varepsilon^* > 0$ be as in Lemmata 10 and 11. Choose $\varepsilon \in (0,\varepsilon^*)$ small enough for

$$\log(\varepsilon^{-1}) + \log(c_0) + \log(\underline{\mathfrak{c}}) > 0 \quad \text{and} \quad d_\varepsilon - 1 > t^* \tag{22}$$

to hold. It further follows from Lemmata 10 and 11 that

$$\log(\mu_{d_\varepsilon}) + \log(\underline{\mathfrak{c}}) \leq \log(\varepsilon) \leq \log(\mu_{d_\varepsilon - 1}) + \log(\overline{\mathfrak{c}}). \tag{23}$$

On the other hand we have from (13)

$$\log(\mu_d) = -\psi(d) + \log(c_0), \quad \text{for all } d \in \mathbb{N}^*. \tag{24}$$

Combining (23) and (24) yields

$$\psi(d_\varepsilon - 1) - \log(\bar{\mathfrak{c}}) \leq \log\left(\varepsilon^{-1}\right) + \log(c_0) \leq \psi(d_\varepsilon) - \log(\underline{\mathfrak{c}}),$$

or, equivalently,

$$\begin{cases} d_\varepsilon \geq \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0) + \log(\underline{\mathfrak{c}})\right], \text{ and} \\ d_\varepsilon \leq \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0) + \log(\bar{\mathfrak{c}})\right] + 1, \end{cases} \tag{25}$$

where we used that the second part of (22) together with Lemma 31 ensures that $\psi$ can, indeed, be inverted, and the first part of (22) guarantees that both bounds are well defined. Using the subadditivity of $\psi^{(-1)}$ (cf. Lemma 34 in Appendix B) and recalling that $\underline{\mathfrak{c}} \leq 1 \leq \bar{\mathfrak{c}}$, we get from (25)

$$\begin{cases} d_\varepsilon \geq \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0)\right] - \psi^{(-1)}[-\log(\underline{\mathfrak{c}})], \text{ and} \\ d_\varepsilon \leq \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0)\right] + \psi^{(-1)}[\log(\bar{\mathfrak{c}})] + 1, \end{cases}$$

which, in turn, allows us to finish the proof by concluding that

$$d_\varepsilon = \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0)\right] + O_{\varepsilon\to 0}(1).$$

□

For later reference, we next collect consequences of Theorem 9 for specific decay rate functions $\psi$. The first one, Corollary 12, treats the case of linear functions $\psi$; while Corollaries 13 and 14 pertain to $\psi$-functions growing super-linearly.

When $\psi$ is a linear function, the analogy with binary expansions, studied in Section 2, suggests that the effective dimension $d_\varepsilon$ grows logarithmically in $\varepsilon^{-1}$. Combining this observation with (9), we expect the leading term of the metric entropy expression to scale according to $\log^2\left(\varepsilon^{-1}\right)$. The next corollary formalizes this insight.

**Corollary 12.** *Let $p, q \in [1, \infty]$, $c > 0$, and*

$$\psi\colon t \in (0, \infty) \mapsto c\,t \in \mathbb{R}.$$

*The metric entropy of the infinite-dimensional ellipsoid $\mathcal{E}_p$ with $\psi$-exponentially decaying semi-axes satisfies*

$$\begin{aligned} H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) = &\frac{\alpha}{2c} \log^2\left(\varepsilon^{-1}\right) \\ &+ \frac{\alpha}{c}\left(\frac{1}{q} - \frac{1}{p}\right) \log\left(\varepsilon^{-1}\right) \log^{(2)}\left(\varepsilon^{-1}\right) + O_{\varepsilon\to 0}\left(\log\left(\varepsilon^{-1}\right)\right), \end{aligned}$$

*where $\alpha = 1$ if $\mathbb{K} = \mathbb{R}$ and $\alpha = 2$ if $\mathbb{K} = \mathbb{C}$.*

*Proof.* See Appendix C.2. □

We now turn to decay rate functions of super-linear growth, whose corresponding metric entropy asymptotics will be expressed in terms of the Lambert $W$-function, the main properties of which are recalled in Appendix A.

**Corollary 13.** *Let* $p, q \in [1, \infty]$, $c, c' > 0$, *and*

$$\psi \colon t \in [1, \infty) \mapsto c\, t\, (\log(t) - c') \in \mathbb{R}.$$

*The metric entropy of the infinite-dimensional ellipsoid* $\mathcal{E}_p$ *with* $\psi$*-exponentially decaying semi-axes satisfies*

$$H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) = \frac{\alpha\, c\, 2^{2c'-1}}{\ln(2)} \exp\{2\beta(\varepsilon)\} \left(\beta(\varepsilon) + \frac{1}{2}\right)\left(1 + O_{\beta(\varepsilon)\to\infty}(\exp\{-\beta(\varepsilon)\})\right),$$

*where* $\alpha = 1$ *if* $\mathbb{K} = \mathbb{R}$, $\alpha = 2$ *if* $\mathbb{K} = \mathbb{C}$, *and*

$$\beta(\varepsilon) := W\left(\frac{\ln(\varepsilon^{-1})}{2^{c'} c}\right),$$

*with* $W(\cdot)$ *denoting the Lambert* $W$*-function.*

*Proof.* See Appendix C.3. □

The next result exploits properties of the Lambert $W$-function to find the explicit asymptotic behavior of the metric entropy in Corollary 13 up to second order. We remark that inspection of the proof reveals that it is actually possible to obtain the asymptotic behavior up to arbitrary order. Concretely, this can be done by continuing the asymptotic expansion of $\psi^{(-1)}$ in Lemma 45 using more terms from Lemma 29. For clarity of exposition, however, we decided to limit the formal statement of the result to second order.

**Corollary 14.** *Let* $p, q \in [1, \infty]$, $c, c' > 0$, *and*

$$\psi \colon t \in [1, \infty) \mapsto c\, t\, (\log(t) - c') \in \mathbb{R}.$$

*The metric entropy of the infinite-dimensional ellipsoid* $\mathcal{E}_p$ *with* $\psi$*-exponentially decaying semi-axes satisfies*

$$H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) = \frac{\alpha \log^2(\varepsilon^{-1})}{2\, c \log^{(2)}(\varepsilon^{-1})}\left(1 + \frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})} + o_{\varepsilon\to 0}\left(\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\right)\right).$$

*Proof.* See Appendix C.3. □

We conclude by pointing out that, both in Corollary 13 and Corollary 14, the asymptotic behavior of metric entropy does not depend on the parameters $p$ and $q$.

# 4 Applications to Complex Analytic Functions

We now turn to the application of our general results to classes of analytic functions whose asymptotic metric entropy behavior has been characterized before in the literature. In each of these cases, we improve upon the best known results, in some cases significantly so. All of these best known results are based on evaluating the cardinalities of explicitly constructed coverings and packings. The corresponding proofs are hence often tedious and highly specific to the function class under consideration.

## 4.1 Periodic Functions Analytic on a Strip

We first consider a class of functions that are periodic on the real line and can be analytically continued to a strip in the complex plane. The formal definition is as follows.

**Definition 15** (Periodic functions analytic on a strip)**.** *Let $M$ and $s$ be positive real numbers. We denote by $\mathcal{A}_s(M)$ the class of functions $f\colon \mathbb{R} \to \mathbb{C}$ which are $2\pi$-periodic and can be analytically continued to the domain*

$$\mathcal{S} := \{z = x + iy \in \mathbb{C} \mid x \in \mathbb{R} \text{ and } |y| < s\},$$

*such that*

$$\sup_{z \in \mathcal{S}} |f(z)| \leq M.$$

Note that, by the identity theorem (see e.g. [22, Corollary 10.18]), $f(z)$ in Definition 15 is $2\pi$-periodic on the entire strip $\mathcal{S}$, i.e., $f(z + 2\pi) = f(z), z \in \mathcal{S}$.

The metric entropy of the class $\mathcal{A}_s(M)$ endowed with the metric

$$d_{2\pi}\colon (f_1, f_2) \longmapsto \sup_{x \in [0, 2\pi]} |f_1(x) - f_2(x)|$$

was characterized in [16, Chapter 7, Section 2.4] according to

$$H(\varepsilon; \mathcal{A}_s(M), d_{2\pi}) = \frac{2\ln(2)}{s} \log^2(\varepsilon^{-1}) + O_{\varepsilon \to 0}\Big(\log(\varepsilon^{-1}) \log^{(2)}(\varepsilon^{-1})\Big). \tag{26}$$

The standard technique for deriving this result is to construct covering and packing elements via an adequate choice of Fourier series coefficients. We next demonstrate that our general approach improves upon (26) by providing a more precise characterization of the $O_{\varepsilon \to 0}(\cdot)$ term.

**Theorem 16.** *Let $M$ and $s$ be positive real numbers. The metric entropy of the class $\mathcal{A}_s(M)$ equipped with the metric $d_{2\pi}$ can be expressed as*

$$H(\varepsilon; \mathcal{A}_s(M), d_{2\pi}) = \frac{2\ln(2)}{s} \Big[\log^2(\varepsilon^{-1}) + \log(\varepsilon^{-1}) \log^{(2)}(\varepsilon^{-1}) (\gamma(\varepsilon) + o_{\varepsilon \to 0}(1))\Big],$$

*with $\gamma(\cdot)$ a function satisfying $|\gamma(\varepsilon)| \leq 1$, for all $\varepsilon > 0$.*

Theorem 16 improves upon (26) as follows. While (26) states that there exists $K > 0$, possibly depending on $s$ and $M$, such that

$$\lim_{\varepsilon \to 0} \left| \frac{\frac{s}{2\ln(2)} H(\varepsilon; \mathcal{A}_s(M), d_{2\pi}) - \log^2(\varepsilon^{-1})}{\log(\varepsilon^{-1}) \log^{(2)}(\varepsilon^{-1})} \right| \leq K,$$

Theorem 16 establishes that the constant $K$ can be taken to be equal to 1, independently of $s$ and $M$.

*Proof.* The elements of $\mathcal{A}_s(M)$ are $2\pi$-periodic functions and can hence be represented in terms of Fourier series according to

$$f(x) = \sum_{k=-\infty}^{\infty} a_k e^{ikx}, \quad \text{with } a_k := \frac{1}{2\pi} \int_0^{2\pi} f(x)\, e^{-ikx}\, dx, \quad \text{for all } k \in \mathbb{Z}.$$

We shall use the shorthand $a := \{a_k\}_{k\in\mathbb{Z}}$. Following our three-step procedure, we first establish that the body formed by the Fourier series coefficients is of ellipsoidal structure, more specifically it can be inscribed and circumscribed with ellipsoids. There is a minor technical adjustment we need to make before we can proceed. Specifically, so far we only considered one-sided ellipsoids, that is, ellipsoids indexed by $\mathbb{N}^*$ rather than $\mathbb{Z}$. The Fourier series coefficients $\{a_k\}_{k\in\mathbb{Z}}$ are, however, two-sided sequences. To consolidate this matter, we define the coefficients

$$\tilde{a}_1 := a_0, \quad \tilde{a}_{2n} := a_n, \quad \text{and} \quad \tilde{a}_{2n+1} := a_{-n}, \quad \text{for all } n \in \mathbb{N}^*, \tag{27}$$

together with the map

$$\iota\colon f \in \mathcal{A}_s(M) \mapsto \{\tilde{a}_n\}_{n\in\mathbb{N}^*}.$$

The ellipsoidal structure of the body formed by the Fourier series coefficients is formalized in the following lemma.

**Lemma 17** (Ellipsoidal structure of $\iota(\mathcal{A}_s(M))$)**.** *Let $M$ and $s$ be positive real numbers and set*

$$\mu_n^{(1)} := Me^{-\frac{sn}{2}} \quad \textit{and} \quad \mu_n^{(2)} := \sqrt{2}Me^{-\frac{s(n-1)}{2}}, \quad \textit{for all } n \in \mathbb{N}^*. \tag{28}$$

*Then, we have*

$$\mathcal{E}_1\left(\left\{\mu_n^{(1)}\right\}_{n\in\mathbb{N}^*}\right) \subseteq \iota(\mathcal{A}_s(M)) \subseteq \mathcal{E}_2\left(\left\{\mu_n^{(2)}\right\}_{n\in\mathbb{N}^*}\right).$$

*Proof.* See Appendix C.6.3. □

We next relate the metric $d_{2\pi}$ to $\ell^q$-metrics in sequence spaces.

**Lemma 18.** *Let $f_1$ and $f_2$ be $2\pi$-periodic functions with Fourier series coefficients $a^{(1)}$ and $a^{(2)}$, respectively, such that*

$$f_1(x) = \sum_{k=-\infty}^{\infty} a_k^{(1)} e^{ikx} \quad \textit{and} \quad f_2(x) = \sum_{k=-\infty}^{\infty} a_k^{(2)} e^{ikx}.$$

*Then, we have*

$$\left\|a^{(1)} - a^{(2)}\right\|_{\ell^2(\mathbb{Z})} \leq d_{2\pi}(f_1, f_2) \leq \left\|a^{(1)} - a^{(2)}\right\|_{\ell^1(\mathbb{Z})}.$$

*Proof.* See Appendix C.4.1. □

Combining Lemma 17 with Lemma 18, we obtain

$$H\left(\varepsilon; \mathcal{E}_1\left(\left\{\mu_n^{(1)}\right\}_{n\in\mathbb{N}^*}\right), \|\cdot\|_2\right) \leq H(\varepsilon; \mathcal{A}_s(M), d_{2\pi}) \tag{29}$$

$$\leq H\left(\varepsilon; \mathcal{E}_2\left(\left\{\mu_n^{(2)}\right\}_{n\in\mathbb{N}^*}\right), \|\cdot\|_1\right), \tag{30}$$

where $\{\mu_n^{(1)}\}_{n\in\mathbb{N}^*}$ and $\{\mu_n^{(2)}\}_{n\in\mathbb{N}^*}$ are according to (28). We now wish to apply Corollary 12 as a proxy for the second and third steps of our procedure. To this end, first observe that, with $\psi(t) = \frac{s}{2\ln(2)}t$, we have

$$\mu_n^{(1)} = M\exp\{-\ln(2)\,\psi(n)\} \quad \text{and} \quad \mu_n^{(2)} = \sqrt{2}Me^{s/2}\exp\{-\ln(2)\,\psi(n)\},$$

for all $n \in \mathbb{N}^*$. This shows that the semi-axes $\{\mu_n^{(1)}\}_{n\in\mathbb{N}^*}$ and $\{\mu_n^{(2)}\}_{n\in\mathbb{N}^*}$ are $\psi$-exponentially decaying. We are now in a position to apply Corollary 12 with $c = \frac{s}{2\ln(2)}$ and $\alpha = 2$ to get

$$H\Big(\varepsilon; \mathcal{E}_2\Big(\Big\{\mu_n^{(2)}\Big\}_{n\in\mathbb{N}^*}\Big), \|\cdot\|_1\Big) \tag{31}$$

$$= \frac{2\ln(2)}{s}\log^2\big(\varepsilon^{-1}\big) + \frac{2\ln(2)}{s}\log\big(\varepsilon^{-1}\big)\log^{(2)}\big(\varepsilon^{-1}\big)(1 + o_{\varepsilon\to 0}(1)) \tag{32}$$

and

$$H\Big(\varepsilon; \mathcal{E}_1\Big(\Big\{\mu_n^{(1)}\Big\}_{n\in\mathbb{N}^*}\Big), \|\cdot\|_2\Big) \tag{33}$$

$$= \frac{2\ln(2)}{s}\log^2\big(\varepsilon^{-1}\big) + \frac{2\ln(2)}{s}\log\big(\varepsilon^{-1}\big)\log^{(2)}\big(\varepsilon^{-1}\big)(-1 + o_{\varepsilon\to 0}(1)). \tag{34}$$

Inserting (31)-(32) and (33)-(34) in (29)-(30) yields the desired result

$$H(\varepsilon; \mathcal{A}_s(M), d_{2\pi}) = \frac{2\ln(2)}{s}\Big[\log^2\big(\varepsilon^{-1}\big) + \log\big(\varepsilon^{-1}\big)\log^{(2)}\big(\varepsilon^{-1}\big)(\gamma(\varepsilon) + o_{\varepsilon\to 0}(1))\Big],$$

with $|\gamma(\varepsilon)| \leq 1$, for all $\varepsilon > 0$. □

We conclude by noting that thresholding infinite-dimensional ellipsoids in the Fourier series sequence space and covering through the resulting finite-dimensional ellipsoids implicitly amounts to covering the class $\mathcal{A}_s(M)$ by finite Fourier series expansions, i.e., by trigonometric polynomials.

## 4.2 Functions Bounded on a Disk

We next consider functions that are analytic and bounded on a disk, formally defined as follows.

**Definition 19** (Analytic functions bounded on a disk)**.** *Let $M$ and $r'$ be positive real numbers. We denote by $\mathcal{A}(r'; M)$ the class of functions $f$ that are analytic on the open disk $D(0; r')$ in the complex plane centered at $0$ and of radius $r'$, and satisfy*

$$\sup_{z\in D(0;r')} |f(z)| \leq M.$$

Based on our three-step procedure, we now characterize the asymptotic behavior of the metric entropy of the class $\mathcal{A}(r'; M)$ endowed with the metric

$$d_r \colon (f_1, f_2) \longmapsto \sup_{z\in D(0;r)} |f_1(z) - f_2(z)|, \tag{35}$$

where we will assume throughout that $0 < r < r'$.

The metric entropy of this function class is of relevance, inter alia, in control theory and signal processing [23, 24] as well as in neural network theory [25], with the best-known result due to Vituškin [2, Chapter 10, Theorem 4] given by

$$H\big(\varepsilon; \mathcal{A}(r'; M), d_r\big) = \frac{\log^2\big(\varepsilon^{-1}\big)}{\log(r'/r)} + \mathcal{O}_{\varepsilon\to 0}\Big(\log\big(\varepsilon^{-1}\big)\log^{(2)}\big(\varepsilon^{-1}\big)\Big). \tag{36}$$

The standard technique for deriving this result is based on explicit constructions of coverings and packings via adequate choices of Taylor series coefficients of functions in $\mathcal{A}(r'; M)$. Again, we shall show that our general approach improves upon (36) by providing a more precise characterization of the second-order term, and does so without resorting to the explicit construction of coverings and packings.

**Theorem 20.** *Let $r$, $r'$, and $M$ be positive real numbers such that $r' > r$. The metric entropy of the class $\mathcal{A}(r';M)$ equipped with the metric $d_r$ can be expressed as*

$$H\big(\varepsilon;\mathcal{A}(r';M),d_r\big) = \frac{\log^2\big(\varepsilon^{-1}\big)}{\log(r'/r)} + \frac{\log\big(\varepsilon^{-1}\big)}{\log(r'/r)}\log^{(2)}\big(\varepsilon^{-1}\big)\big(\gamma(\varepsilon) + o_{\varepsilon\to 0}(1)\big),$$

*with $\gamma(\cdot)$ a function satisfying $|\gamma(\varepsilon)| \le 1$, for all $\varepsilon > 0$.*

Theorem 20 improves upon (36) as follows. While (36) states that there exists $K > 0$, possibly depending on $r, r'$, and $M$, such that

$$\lim_{\varepsilon\to 0}\left|\frac{\log(r'/r)\,H(\varepsilon;\mathcal{A}(r';M),d_r) - \log^2\big(\varepsilon^{-1}\big)}{\log(\varepsilon^{-1})\log^{(2)}(\varepsilon^{-1})}\right| \le K,$$

Theorem 20 establishes that the constant $K$ can be taken to be equal to 1, independently of $r, r'$, and $M$.

*Proof.* Let $\{a_k\}_{k\in\mathbb{N}}$ be the sequence of Taylor series coefficients of $f$, which exists by analyticity, i.e.,

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \quad \text{for all } z \in D(0;r').$$

We start by defining the embedding

$$\iota\colon f \in \mathcal{A}(r';M) \mapsto \{\tilde{a}_k\}_{k\in\mathbb{N}} \in \ell^\infty(\mathbb{N}), \quad \text{with } \tilde{a}_k := a_k r^k, \tag{37}$$

where $\{\tilde{a}_k\}_{k\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$ is a direct consequence of Cauchy's estimate (47). As in the proof of Theorem 16, we next inscribe and circumscribe $\iota(\mathcal{A}(r';M))$ with ellipsoids.

**Lemma 21** (Ellipsoidal structure of $\mathcal{A}(r';M)$)**.** *Let $r$, $r'$, and $M$ be positive real numbers such that $r' > r$, and let*

$$\mu_k := M\exp\big\{-k\ln(r'/r)\big\}, \quad \textit{for all } k \in \mathbb{N}. \tag{38}$$

*Then, we have*

$$\mathcal{E}_1(\{\mu_k\}_{k\in\mathbb{N}}) \subseteq \iota\big(\mathcal{A}(r';M)\big) \subseteq \mathcal{E}_2(\{\mu_k\}_{k\in\mathbb{N}}).$$

*Proof.* See Appendix C.6.4. □

We next relate the metric $d_r$ to $\ell^q$-metrics in sequence spaces.

**Lemma 22.** *Let $r$ be a positive real number and let $f_1, f_2$ be functions that are analytic on $D(0;r)$, with respective Taylor series expansions*

$$f_1(z) = \sum_{k=0}^{\infty} a_k^{(1)} z^k \quad \textit{and} \quad f_2(z) = \sum_{k=0}^{\infty} a_k^{(2)} z^k.$$

*Then, we have*

$$\left\|\tilde{a}^{(1)} - \tilde{a}^{(2)}\right\|_{\ell^2} \le d_r(f_1,f_2) \le \left\|\tilde{a}^{(1)} - \tilde{a}^{(2)}\right\|_{\ell^1}, \tag{39}$$

*where*

$$\tilde{a}^{(1)} := \left\{a_k^{(1)} r^k\right\}_{k\in\mathbb{N}} \quad \textit{and} \quad \tilde{a}^{(2)} := \left\{a_k^{(2)} r^k\right\}_{k\in\mathbb{N}}.$$

*Proof.* See Appendix C.4.2. □

Combining Lemma 21 with Lemma 22, we obtain

$$H(\varepsilon;\mathcal{E}_1(\{\mu_k\}_{k\in\mathbb{N}}),\|\cdot\|_2)\leq H\big(\varepsilon;\mathcal{A}(r';M),d_r\big)\leq H(\varepsilon;\mathcal{E}_2(\{\mu_k\}_{k\in\mathbb{N}}),\|\cdot\|_1), \tag{40}$$

where $\{\mu_k\}_{k\in\mathbb{N}}$ is as per (38). We can now apply Corollary 12 with $c=\log(r'/r)$ to obtain

$$H(\varepsilon;\mathcal{E}_2(\{\mu_k\}_{k\in\mathbb{N}}),\|\cdot\|_1)=\frac{\log^2\big(\varepsilon^{-1}\big)}{\log(r'/r)} \tag{41}$$

$$+\frac{2}{\log(r'/r)}\log\big(\varepsilon^{-1}\big)\log^{(2)}\big(\varepsilon^{-1}\big)\left(\frac{1}{2}+o_{\varepsilon\to 0}(1)\right) \tag{42}$$

and

$$H(\varepsilon;\mathcal{E}_1(\{\mu_k\}_{k\in\mathbb{N}}),\|\cdot\|_2)=\frac{\log^2\big(\varepsilon^{-1}\big)}{\log(r'/r)} \tag{43}$$

$$+\frac{2}{\log(r'/r)}\log\big(\varepsilon^{-1}\big)\log^{(2)}\big(\varepsilon^{-1}\big)\left(-\frac{1}{2}+o_{\varepsilon\to 0}(1)\right). \tag{44}$$

Inserting (41)-(42) and (43)-(44) in (40) yields the desired result

$$H\big(\varepsilon;\mathcal{A}(r';M),d_r\big)=\frac{\log^2\big(\varepsilon^{-1}\big)}{\log(r'/r)}+\frac{\log\big(\varepsilon^{-1}\big)}{\log(r'/r)}\log^{(2)}\big(\varepsilon^{-1}\big)(\gamma(\varepsilon)+o_{\varepsilon\to 0}(1)),$$

with $|\gamma(\varepsilon)|\leq 1$, for all $\varepsilon>0$. □

We conclude by noting that thresholding infinite-dimensional ellipsoids in the Taylor series sequence space and covering through the resulting finite-dimensional ellipsoids implicitly amounts to covering the class $\mathcal{A}(r';M)$ by finite Taylor series expansions, i.e., by polynomials, in a complex variable.

## 4.3 Functions of Exponential Type

We finally consider functions of exponential type [22, Chapter 19] defined as follows.

**Definition 23** (Functions of exponential type)**.** *An entire function $f$ is said to be of exponential type if there exist positive real constants $A$ and $C$ such that, for all $z\in\mathbb{C}$,*

$$|f(z)|\leq Ce^{A|z|}. \tag{45}$$

*We write $\mathcal{F}_{exp}^{A,C}$ for the class of entire functions of exponential type with constants $A$ and $C$.*

Functions of exponential type appear naturally in many practical applications, mainly as they can be identified, through the Paley-Wiener theorem [22, Theorem 19.3], with band-limited functions. For instance, in [19] the metric entropy (rate) of band-limited signals plays a fundamental role in assessing the ultimate performance limits of analog-to-digital converters.

The best known result on the metric entropy of $\mathcal{F}_{\text{exp}}^{A,C}$ is [16, Chapter 7, Theorem XX]

$$H\big(\varepsilon;\mathcal{F}_{\text{exp}}^{A,C},d_1\big)\sim_{\varepsilon\to 0}\frac{\log^2\big(\varepsilon^{-1}\big)}{\log^{(2)}\big(\varepsilon^{-1}\big)}, \tag{46}$$

where $d_1$ is the metric defined in (35) under the choice $r=1$. We improve upon (46) as follows.

**Theorem 24.** *Let $A$ and $C$ be positive real constants. The metric entropy of the class $\mathcal{F}_{exp}^{A,C}$ (as per Definition 23) equipped with the metric $d_1$ satisfies*

$$\begin{aligned}
&\left(\frac{eA}{2}\right)^2 \frac{\exp\{2\beta_1(\varepsilon)\}}{\ln(2)}\left[\beta_1(\varepsilon)+\frac{1}{2}\right]\left[1+O_{\beta_1(\varepsilon)\to\infty}(\exp\{-\beta_1(\varepsilon)\})\right] \\
&\leq H\left(\varepsilon;\mathcal{F}_{exp}^{A,C},d_1\right) \\
&\leq (eA)^2 \frac{\exp\{2\beta_2(\varepsilon)\}}{\ln(2)}\left[\beta_2(\varepsilon)+\frac{1}{2}\right]\left[1+O_{\beta_2(\varepsilon)\to\infty}(\exp\{-\beta_2(\varepsilon)\})\right],
\end{aligned}$$

*where*

$$\beta_1(\varepsilon) := W\left(\frac{2\ln(\varepsilon^{-1})}{e\,A}\right) \quad \text{and} \quad \beta_2(\varepsilon) := W\left(\frac{\ln(\varepsilon^{-1})}{e\,A}\right),$$

*with $W$ denoting the Lambert $W$-function (cf. Appendix A).*

*Proof.* See Appendix C.5. □

It is not immediate that the characterization provided in Theorem 24 indeed constitutes an improvement over (46). We therefore employ Theorem 24 to make the second order term in the asymptotic expansion of the metric entropy of $\mathcal{F}_{\exp}^{A,C}$ explicit.

**Corollary 25.** *Let $A$ and $C$ be positive real constants. The metric entropy of the class $\mathcal{F}_{exp}^{A,C}$ (as per Definition 23) equipped with the metric $d_1$ satisfies*

$$H\left(\varepsilon;\mathcal{F}_{exp}^{A,C},d_1\right) = \frac{\log^2(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\left(1+\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}+o_{\varepsilon\to 0}\left(\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\right)\right).$$

*Proof.* See Appendix C.5. □

**Acknowledgments.** The authors would like to thank V. Abadie for drawing their attention to the Lambert $W$-function.

# References

[1] G. G. Lorentz, "Metric entropy and approximation," *Bulletin of the American Mathematical Society*, vol. 72, no. 6, pp. 903–937, 1966.

[2] ——, *Approximation of functions*, ser. Athena Series. American Mathematical Society, 1966, vol. 48.

[3] B. Carl and I. Stephani, *Entropy, Compactness and the Approximation of Operators*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, 1990.

[4] M. Franceschetti, *Wave Theory of Information*. Cambridge University Press, 2017.

[5] D. Haussler, "Decision theoretic generalizations of the pac model for neural net and other learning applications," *Information and Computation*, vol. 100, no. 1, pp. 78–150, 1992.

[6] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[7] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms.* Cambridge University Press, 2014.

[8] R. M. Dudley, H. Kunita, and F. Ledrappier, *École d'été de probabilités de Saint-Flour XII - 1982*, ser. Lecture Notes in Mathematics, P. L. Hennequin, Ed. Springer Berlin Heidelberg, 1984, vol. 1097.

[9] D. Pollard, "Empirical processes: Theory and applications," *NSF-CBMS Regional Conference Series in Probability and Statistics*, vol. 2, pp. i–86, 1990.

[10] D. Elbrächter, D. Perekrestenko, P. Grohs, and H. Bölcskei, "Deep neural network approximation theory," *IEEE Trans. Inform. Theory*, vol. 67, no. 5, pp. 2581–2623, 2021.

[11] R. DeVore, B. Hanin, and G. Petrova, "Neural network approximation," *Acta Numerica*, vol. 30, pp. 327–444, 2021.

[12] A. N. Kolmogorov, "On certain asymptotic characteristics of completely bounded metric spaces," *Dokl. Akad. Nauk*, vol. 108, pp. 385–388, 1956.

[13] B. S. Mityagin, "Approximate dimension and bases in nuclear spaces," *Russian Mathematical Surveys*, vol. 16, no. 4, pp. 59–127, 1961.

[14] A. Vitushkin, *Theory of the Transmission and Processing of Information.* Pergamon Press, 1961.

[15] D. L. Donoho, *Counting bits with Kolmogorov and Shannon.* Department of Statistics, Stanford University, 2000.

[16] Shiryayev, *Selected Works of A. N. Kolmogorov*, ser. Mathematics and Its Applications. Springer Netherlands, 1993, vol. 27.

[17] H. Luschgy and G. Pagès, "Sharp asymptotics of the kolmogorov entropy for gaussian measures," *Journal of Functional Analysis*, vol. 212, no. 1, pp. 89–120, 2004.

[18] S. Graf and H. Luschgy, "Sharp asymptotics of the metric entropy for ellipsoids," *Journal of Complexity*, vol. 20, no. 6, pp. 876–882, 2004.

[19] I. Daubechies, R. DeVore, C. Güntürk, and V. Vaishampayan, "A/d conversion with imperfect quantizers," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 874–885, 2006.

[20] H. Kempka and J. Vybíral, "Volumes of unit balls of mixed sequence spaces," *Mathematische Nachrichten*, vol. 290, no. 8-9, pp. 1317–1327, 2017.

[21] D. E. Edmunds and H. Triebel, *Function Spaces, Entropy Numbers, Differential Operators*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, 1996.

[22] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill, 1987.

[23] G. Zames, "On the metric complexity of causal linear systems: Estimates of $\varepsilon$-entropy and $\varepsilon$-dimension," *IEEE Trans. Automat. Contr.*, vol. 24, no. 2, pp. 222–230, 1979.

[24] G. Zames and J. Owen, "A note on metric dimension and feedback in discrete time," *IEEE Trans. Automat. Contr.*, vol. 38, no. 4, pp. 664–667, 1993.

[25] C. Hutter, R. Gül, and H. Bölcskei, "Metric entropy limits on recurrent neural network learning of linear dynamical systems," *Applied and Computational Harmonic Analysis*, vol. 59, pp. 198–223, 2022.

[26] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, "On the lambert w function," *Advances in Computational mathematics*, vol. 5, pp. 329–359, 1996.

[27] N. G. De Bruijn, *Asymptotic methods in analysis.* Courier Corporation, 1981, vol. 4.

[28] T. Allard and H. Bölcskei, "Entropy of compact operators for sampling theorems and sobolev spaces," -, 2024.

[29] G. G. Lorentz, M. von Golitschek, and Y. Makovoz, *Constructive Approximation: Advanced Problems*, ser. Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg, 1996.

[30] C. Schütt, "Entropy numbers of diagonal operators between symmetric Banach spaces," *Journal of Approximation Theory*, vol. 40, no. 2, pp. 121–128, 1984.

[31] W. Rudin *et al.*, *Principles of mathematical analysis.* McGraw-hill New York, 1964, vol. 3.

# A Complements

This appendix collects known results that will be used frequently in the paper along with auxiliary results.

**Lemma 26** (Cauchy's estimate)**.** *[22, Theorem 10.26] Let $r$ and $M$ be positive real numbers and let $f$ be analytic on $D(0;r)$ with Taylor series coefficients $\{c_k\}_{k\in\mathbb{N}}$. If $|f(z)| \leq M$, for all $z \in D(0;r)$, then*

$$|c_k| \leq \frac{M}{r^k}, \quad \text{for all } k \in \mathbb{N}. \tag{47}$$

We next recall results on the Lambert $W$-function.

**Definition 27** (Lambert $W$-function)**.** *The Lambert $W$-function is defined as the unique function satisfying*

$$W(x)\, e^{W(x)} = x, \quad \text{for all } x \geq 0.$$

**Lemma 28.** *Let $a$ and $b$ be positive real numbers. Then, the equation*

$$x = ae^{-x} + b$$

*has a unique solution, which is given by*

$$x = b + W\Big(ae^{-b}\Big).$$

**Lemma 29.** *The Lambert $W$-function satisfies the relation*

$$W(x) = \ln(x) - \ln^{(2)}(x) + \sum_{n=0}^{\infty}\sum_{m=1}^{\infty} \frac{(-1)^n s(n+m, n+1)}{m!} \frac{\left(\ln^{(2)}(x)\right)^m}{\ln^{n+m}(x)},$$

*for all $x > 0$, where $s(n+m, n+1)$ denotes the Stirling numbers of the first kind. In particular, one has the asymptotic behavior*

$$W(x) = \ln(x) - \ln^{(2)}(x) + \frac{\ln^{(2)}(x)}{\ln(x)}(1 + o_{x\to\infty}(1)). \tag{48}$$

We refer to [26] for further material on the Lambert $W$-function and to [27, Chapter 2.4] for a proof of Lemma 29. Lemma 28 can be verified by direct substitution.

Finally, we shall need the following auxiliary result.

**Lemma 30.** *It holds that*

$$1 \le \sup_{k\in\mathbb{N}^*} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k} < \infty.$$

*Proof.* See Appendix C.6.5. ☐

# B Properties of Decay Rate Functions

**Lemma 31.** *Let $\psi$ be a decay rate function with parameter $t^*$. Then, $\psi$ is strictly increasing and invertible on $(t^*, \infty)$.*

*Proof.* Fix $t_1, t_2 \in (t^*, \infty)$ such that $t_1 < t_2$. We then have

$$\psi(t_2) - \psi(t_1) = t_2\left[\frac{\psi(t_2)}{t_2} - \frac{\psi(t_1)}{t_2}\right] > t_2\left[\frac{\psi(t_2)}{t_2} - \frac{\psi(t_1)}{t_1}\right] \ge 0. \tag{49}$$

The inequality (49) directly implies that $\psi$ is strictly increasing on $(t^*, \infty)$. Combined with the continuity of $\psi$ on $(t^*, \infty)$, which follows by definition, this allows us to conclude that $\psi$ is invertible on $(t^*, \infty)$. ☐

**Lemma 32.** *Let $\psi$ be a decay rate function. The $\psi$-average function $\delta$ and the $\psi$-difference function $\zeta$ are related according to*

$$\frac{\sigma_{\mathbb{K}}(d)\,\delta(d) - \sigma_{\mathbb{K}}(d-1)\,\delta(d-1)}{\sigma_{\mathbb{K}}(d)} = \zeta(d)\,(1 - 1/d), \quad \textit{for all } d \ge 2.$$

*Proof.* We fix $d \ge 2$ and note the following chain of identities

$$\begin{aligned} d\,\delta(d) - (d-1)\,\delta(d-1) &= \sum_{n=1}^{d}(\psi(d) - \psi(n)) - \sum_{n=1}^{d-1}(\psi(d-1) - \psi(n)) \\ &= d\,\psi(d) - (d-1)\,\psi(d-1) + \sum_{n=1}^{d-1}\psi(n) - \sum_{n=1}^{d}\psi(n) \\ &= d\,\psi(d) - (d-1)\,\psi(d-1) - \psi(d) \\ &= (d-1)(\psi(d) - \psi(d-1)) \\ &= (d-1)\,\zeta(d) = d\,\zeta(d)\,(1 - 1/d). \end{aligned}$$

Rewriting this relation according to

$$\frac{\sigma_{\mathbb{K}}(d)\,\delta(d) - \sigma_{\mathbb{K}}(d-1)\,\delta(d-1)}{\sigma_{\mathbb{K}}(d)} = \zeta(d)\,(1-1/d)$$

finishes the proof. □

**Lemma 33.** *Let $\psi$ be a decay rate function with parameter $t^*$. There exists a positive real number $\kappa$ such that the $\psi$-difference function $\zeta$ satisfies*

$$\zeta(d) \geq \kappa, \quad \textit{for all } d > t^* + 2.$$

*Proof.* We fix $d > t^* + 2$ and start by rewriting $\zeta$ according to

$$\begin{aligned} \zeta(d) &= \psi(d) - \psi(d-1) && (50)\\ &= d\left(\frac{\psi(d)}{d} - \frac{\psi(d-1)}{d-1}\left(1 - \frac{1}{d}\right)\right)\\ &= d\left(\frac{\psi(d)}{d} - \frac{\psi(d-1)}{d-1}\right) + \frac{\psi(d-1)}{d-1}. && (51)\end{aligned}$$

As the function

$$t \in (t^*, \infty) \mapsto \frac{\psi(t)}{t} \in (0, \infty)$$

is non-decreasing by definition, we have

$$\frac{\psi(d)}{d} - \frac{\psi(d-1)}{d-1} \geq 0 \quad \text{and} \quad \frac{\psi(d-1)}{d-1} \geq \frac{\psi(\lfloor t^*+1 \rfloor)}{\lfloor t^*+1 \rfloor}.$$

Putting everything together, we obtain

$$\zeta(d) = d\left(\frac{\psi(d)}{d} - \frac{\psi(d-1)}{d-1}\right) + \frac{\psi(d-1)}{d-1} \geq \frac{\psi(\lfloor t^*+1 \rfloor)}{\lfloor t^*+1 \rfloor},$$

which establishes the desired result upon setting

$$\kappa := \frac{\psi(\lfloor t^*+1 \rfloor)}{\lfloor t^*+1 \rfloor}.$$

□

**Lemma 34** (Subadditivity of $\psi^{(-1)}$)**.** *The inverse $\psi^{(-1)}$ of the decay rate function $\psi$ is subadditive, i.e., for all $a, b > 0$, we have*

$$\psi^{(-1)}(a+b) \leq \psi^{(-1)}(a) + \psi^{(-1)}(b).$$

*Proof.* We proceed in two steps, the first of which analyzes the variations of the function

$$u \mapsto \frac{\psi^{(-1)}(u)}{u}.$$

To this end, we compute

$$\begin{aligned} \left[\frac{\psi^{(-1)}(u)}{u}\right]' &= \left[\frac{\left[\psi'\left(\psi^{(-1)}(u)\right)\right]^{-1}}{u} - \frac{\psi^{(-1)}(u)}{u^2}\right] && (52)\\ &= \frac{\left[\psi^{(-1)}(u)\right]^2}{u^2\,\psi'\left(\psi^{(-1)}(u)\right)}\left[\frac{u}{\left[\psi^{(-1)}(u)\right]^2} - \frac{\psi'\left(\psi^{(-1)}(u)\right)}{\psi^{(-1)}(u)}\right]. && (53)\end{aligned}$$

Changing variables according to

$$t = \psi^{(-1)}(u),$$

(52)-(53) can be rewritten as

$$\left[\frac{\psi^{(-1)}(u)}{u}\right]' = \frac{t^2}{[\psi(t)]^2\,\psi'(t)}\left[\frac{\psi(t)}{t^2} - \frac{\psi'(t)}{t}\right] \tag{54}$$

$$= -\frac{t^2}{[\psi(t)]^2\,\psi'(t)}\left[\frac{\psi(t)}{t}\right]' \leq 0, \tag{55}$$

where the inequality holds for all $t \in (t^*, \infty)$ as a direct consequence of $\psi$ being strictly increasing (cf. Lemma 31) and satisfying the properties (12) in Definition 7. It hence follows from (54)-(55) that

$$u \mapsto \frac{\psi^{(-1)}(u)}{u} \tag{56}$$

is non-increasing on $(0, \infty)$.

For the second step, fix $a, b > 0$ and observe that

$$a\,\frac{\psi^{(-1)}(a+b)}{a+b} \leq \psi^{(-1)}(a) \quad \text{and} \quad b\,\frac{\psi^{(-1)}(a+b)}{a+b} \leq \psi^{(-1)}(b). \tag{57}$$

The desired result now follows by application of these two inequalities according to

$$\psi^{(-1)}(a+b) = a\,\frac{\psi^{(-1)}(a+b)}{a+b} + b\,\frac{\psi^{(-1)}(a+b)}{a+b} \overset{(57)}{\leq} \psi^{(-1)}(a) + \psi^{(-1)}(b).$$

□

# C Proofs

## C.1 Proofs of Theorems 4 and 5

We start with preparatory material needed in both proofs. First, observe that $\mathcal{E}_p^d$ is the image of $\mathcal{B}_p$ under the diagonal matrix

$$A_\mu = \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_d \end{pmatrix}. \tag{58}$$

Indeed, we have

$$\begin{aligned} \mathcal{E}_p^d &= \left\{x \in \mathbb{K}^d \mid \|x\|_{p,\mu} \leq 1\right\} \\ &= \left\{x \in \mathbb{K}^d \mid \|z\|_p \leq 1, \text{ such that } x_n = \mu_n z_n, \text{ for all } n \in \{1, \dots, d\}\right\} \\ &= \left\{A_\mu z \mid z \in \mathbb{K}^d \text{ and } \|z\|_p \leq 1\right\} = A_\mu \mathcal{B}_p. \end{aligned}$$

Next, note that

$$\det(A_\mu) = \bar{\mu}_d^{\sigma_{\mathbb{K}}(d)}, \tag{59}$$

where we identified $\mathbb{C}$ with $\mathbb{R}^2$. It now follows as an immediate consequence of $\mathcal{E}_p^d = A_\mu \mathcal{B}_p$ that the volumes of the ellipsoid $\mathcal{E}_p^d$ and the ball $\mathcal{B}_p$ are related according to

$$\mathrm{vol}_{\mathbb{K}}(\mathcal{E}_p^d) = \mathrm{vol}_{\mathbb{K}}(\mathcal{B}_p)\det(A_\mu) \overset{(59)}{=} \mathrm{vol}_{\mathbb{K}}(\mathcal{B}_p)\,\bar{\mu}_d^{\sigma_{\mathbb{K}}(d)}. \tag{60}$$

We conclude with a lemma on volume ratios.

**Lemma 35.** *Let $d \in \mathbb{N}^*$ and $p, q \in [1, \infty]$. The volume ratio $V_{p,q}^{\mathbb{K},d}$ satisfies*

$$\log\left\{\left(V_{p,q}^{\mathbb{K},d}\right)^{\frac{1}{\sigma_{\mathbb{K}}(d)}}\right\} = \left(\frac{1}{q} - \frac{1}{p}\right)\log(\sigma_{\mathbb{K}}(d)) + O_{d\to\infty}(1).$$

*Proof.* See Appendix C.6.6. □

## Proof of Theorem 4

The proof of Theorem 4 is effected by applying Lemma 3 with $\mathcal{B} = \mathcal{E}_p^d$, $\mathcal{B}' = \mathcal{B}_q$, $\|\cdot\| = \|\cdot\|_{p,\mu}$, and $\|\cdot\|' = \|\cdot\|_q$, according to

$$N\left(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\right)\varepsilon^{\sigma_{\mathbb{K}}(d)} \geq \frac{\mathrm{vol}_{\mathbb{K}}(\mathcal{E}_p^d)}{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}_q)} \overset{(60)}{=} V_{p,q}^{\mathbb{K},d}\,\bar{\mu}_d^{\sigma_{\mathbb{K}}(d)}, \tag{61}$$

and observing that Lemma 35 then implies the existence of a constant $\underline{\kappa} > 0$ independent of $d$ such that

$$\left(V_{p,q}^{\mathbb{K},d}\right)^{\frac{1}{\sigma_{\mathbb{K}}(d)}} \geq \underline{\kappa}\,\sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)}. \tag{62}$$

Using (62) in (61) yields the desired result. We additionally note that, in the case $p = q$, one can take $\underline{\kappa} = 1$. This observation is exploited in [28, Theorem 5].

## Proof of Theorem 5

We split the proof of Theorem 5 into the cases $p \geq q$ and $p < q$.

*Case $p \geq q$.* With a view towards application of Lemma 3 with $\mathcal{B} = \mathcal{E}_p^d$ and $\mathcal{B}' = \mathcal{B}_q$, we first consider the set

$$\mathcal{E}_p^d + \frac{\varepsilon}{2}\mathcal{B}_q$$

and note that owing to $\mathcal{B}_q \subseteq \mathcal{B}_p$ and $\varepsilon \leq 2\mu_d$, it holds that

$$\mathcal{E}_p^d + \frac{\varepsilon}{2}\mathcal{B}_q \subseteq \mathcal{E}_p^d + \frac{\varepsilon}{2}\mathcal{B}_p \subseteq \mathcal{E}_p^d + \mu_d\mathcal{B}_p. \tag{63}$$

Next, as the semi-axes $\mu_1, \ldots, \mu_d$ are non-increasing, we obtain

$$\mathcal{E}_p^d + \mu_d\mathcal{B}_p \subseteq \mathcal{E}_p^d + \mathcal{E}_p^d = 2\mathcal{E}_p^d, \tag{64}$$

where the equality is thanks to the convexity of the ellipsoid $\mathcal{E}_p^d$ (recall that $p \geq 1$ by assumption). Combining (63) and (64) now yields

$$\mathrm{vol}_{\mathbb{K}}\left(\mathcal{E}_p^d + \frac{\varepsilon}{2}\mathcal{B}_q\right) \leq 2^{\sigma_{\mathbb{K}}(d)}\,\mathrm{vol}_{\mathbb{K}}\left(\mathcal{E}_p^d\right). \tag{65}$$

Applying the upper bound from Lemma 3, with $\mathcal{B} = \mathcal{E}_p^d$, $\mathcal{B}' = \mathcal{B}_q$, $\|\cdot\| = \|\cdot\|_{p,\mu}$, and $\|\cdot\|' = \|\cdot\|_q$, we get

$$N\Big(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\Big)\, \varepsilon^{\sigma_{\mathbb{K}}(d)} \leq 2^{\sigma_{\mathbb{K}}(d)} \frac{\mathrm{vol}_{\mathbb{K}}\big(\mathcal{E}_p^d + \frac{\varepsilon}{2}\mathcal{B}_q\big)}{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}_q)}$$
$$\overset{(65)}{\leq} 4^{\sigma_{\mathbb{K}}(d)} \frac{\mathrm{vol}_{\mathbb{K}}(\mathcal{E}_p^d)}{\mathrm{vol}_{\mathbb{K}}(\mathcal{B}_q)} \overset{(60)}{=} 4^{\sigma_{\mathbb{K}}(d)}\, V_{p,q}^{\mathbb{K},d}\, \bar{\mu}_d^{\sigma_{\mathbb{K}}(d)}. \tag{66}$$

From Lemma 35 we know that there exists a constant $C > 0$ that is independent of $d$ and such that

$$\Big(V_{p,q}^{\mathbb{K},d}\Big)^{\frac{1}{\sigma_{\mathbb{K}}(d)}} \leq C\, \sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)},$$

which, when used in (66), leads to

$$N\Big(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\Big)^{\frac{1}{\sigma_{\mathbb{K}}(d)}} \varepsilon \leq 4\, C\, \sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)} \bar{\mu}_d. \tag{67}$$

Setting $\overline{\kappa} := 4\,C$ yields the desired result.

*Case* $p < q$. The proof proceeds through arguments based on entropy numbers. We first recall that, for $m \in \mathbb{N}^*$, the $m$-th entropy number $e_m(T)$ of a linear operator $T\colon (\mathbb{K}^d, \|\cdot\|) \to (\mathbb{K}^d, \|\cdot\|')$ is defined according to

$$e_m(T) := \inf\big\{\varepsilon > 0 \mid H\big(\varepsilon; T(\mathcal{B}), \|\cdot\|'\big) \leq m\big\},$$

where $\mathcal{B}$ denotes the unit ball in $\mathbb{K}^d$ w.r.t. the norm $\|\cdot\|$ (see [21, Definition 1.3.1]). We are particularly interested in the mapping

$$T_{p,q,\mu} := \mathrm{id}\colon (\mathbb{K}^d, \|\cdot\|_{p,\mu}) \to (\mathbb{K}^d, \|\cdot\|_q), \tag{68}$$

that is, the identity operator between $\mathbb{K}^d$ equipped with the $p$-ellipsoid norm $\|\cdot\|_{p,\mu}$ and $\mathbb{K}^d$ equipped with the $q$-norm $\|\cdot\|_q$. With this choice, we have

$$e_m(T_{p,q,\mu}) := \inf\Big\{\varepsilon > 0 \mid H\Big(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\Big) \leq m\Big\}.$$

Additionally, we fix $m$ according to

$$m := \frac{1}{2}\lfloor H\Big(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\Big)\rfloor,$$

so that trivially

$$\varepsilon \leq e_{2m}(T_{p,q,\mu}) \quad \text{and} \quad N\Big(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\Big) \leq 2^{2m+1}. \tag{69}$$

Next, note that the operator $T_{p,q,\mu}$ can be factorized according to

$$T_{p,q,\mu} = \mathrm{id}_{p,q} \circ A_\mu,$$

where $\mathrm{id}_{p,q}$ refers to the identity operator between $\mathbb{K}^d$ equipped with the $p$-norm $\|\cdot\|_p$ and $\mathbb{K}^d$ equipped with the $q$-norm $\|\cdot\|_q$, and $A_\mu$ was defined in (58). Using the multiplicativity property of entropy numbers (see e.g. [29, Equation 15.7.2]), we get

$$e_{2m}(\mathrm{id}_{p,q} \circ A_\mu) \leq e_m(\mathrm{id}_{p,q})\, e_m(A_\mu), \tag{70}$$

which allows us to reduce the problem of upper-bounding the entropy number $e_{2m}(T_{p,q,\mu})$ to that of upper-bounding the entropy numbers $e_m(\mathrm{id}_{p,q})$ and $e_m(A_\mu)$. Under the assumption (10) and using the fact that $\sigma_{\mathbb{K}}(d)$ is integer, one has $m \geq \sigma_{\mathbb{K}}(d)$ and hence the classical result [30, Theorem 1] on the entropy number of diagonal operators applies, which can be reformulated in our setting according to

$$e_m(\mathrm{id}_{p,q}) \leq C_1 2^{-\frac{m}{\sigma_{\mathbb{K}}(d)}} \sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)}, \tag{71}$$

where $C_1$ is a positive numerical constant. Moreover, by setting $p = q$ in (67), we get

$$e_m(A_\mu) \leq C_2 2^{-\frac{m}{\sigma_{\mathbb{K}}(d)}} \bar{\mu}_d, \tag{72}$$

where $C_2$ is a positive numerical constant. Now combining (71) and (72) in (70) yields

$$e_{2m}(T_{p,q,\mu}) \leq C_0 2^{-\frac{2m}{\sigma_{\mathbb{K}}(d)}} \sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)} \bar{\mu}_d, \tag{73}$$

with $C_0$ a positive numerical constant. Finally, using (69) in combination with (73) results in the desired bound according to

$$\begin{aligned} N\Big(\varepsilon; \mathcal{E}_p^d, \|\cdot\|_q\Big)^{\frac{1}{\sigma_{\mathbb{K}}(d)}} \varepsilon &\overset{(69)}{\leq} 2^{\frac{2m+1}{\sigma_{\mathbb{K}}(d)}} e_{2m}(T_{p,q,\mu}) \\ &\overset{(73)}{\leq} 2^{\frac{1}{\sigma_{\mathbb{K}}(d)}} C_0 \, \sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)} \bar{\mu}_d \leq \overline{\kappa} \, \sigma_{\mathbb{K}}(d)^{\left(\frac{1}{q}-\frac{1}{p}\right)} \bar{\mu}_d, \end{aligned}$$

where we set $\overline{\kappa} := 2\, C_0$ and used $2^{\frac{1}{\sigma_{\mathbb{K}}(d)}} \leq 2$ in the last inequality. This concludes the proof.

## C.2 Proof of Corollary 12

First, note that $\psi(t) = ct$ trivially satisfies the defining properties of decay rate functions with the choice $t^* = 0$ (cf. Definition 7). We can therefore apply Theorem 9 to get

$$H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) = \sigma_{\mathbb{C}}(d_\varepsilon)\left\{\delta(d_\varepsilon) + \left(\frac{1}{q}-\frac{1}{p}\right)\log(\sigma_{\mathbb{C}}(d_\varepsilon)) + O_{d_\varepsilon\to\infty}(\zeta(d_\varepsilon))\right\}, \tag{74}$$

with $d_\varepsilon$ satisfying

$$d_\varepsilon = \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0)\right] + O_{\varepsilon\to 0}(1) = \frac{\log\left(\varepsilon^{-1}\right)}{c} + O_{\varepsilon\to 0}(1). \tag{75}$$

With Definition 8, we obtain the $\psi$-average function according to

$$\delta(d) = \frac{1}{d}\sum_{n=1}^{d}(cd - cn) = c\,\frac{d-1}{2}, \qquad \text{for all } d \in \mathbb{N}^*, \tag{76}$$

together with the $\psi$-difference function

$$\zeta(d) = \psi(d) - \psi(d-1) = c, \qquad \text{for all } d \geq 2. \tag{77}$$

Finally, observe that $\sigma_{\mathbb{C}}(d) = \alpha\, d$, for all $d \in \mathbb{N}^*$. Using (75), (76), and (77) in (74), the desired result follows according to

$$\begin{aligned} H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) &= \alpha\, d_\varepsilon \left\{ \frac{c}{2} d_\varepsilon + \left(\frac{1}{q} - \frac{1}{p}\right) \log(d_\varepsilon) + O_{d_\varepsilon \to \infty}(1) \right\} \\ &= \frac{\alpha\, c}{2}\, d_\varepsilon^2 + \alpha \left(\frac{1}{q} - \frac{1}{p}\right) d_\varepsilon \log(d_\varepsilon) + O_{d_\varepsilon \to \infty}(d_\varepsilon) \\ &= \frac{\alpha\, \log^2\left(\varepsilon^{-1}\right)}{2\, c} \\ &\quad + \frac{\alpha}{c} \left(\frac{1}{q} - \frac{1}{p}\right) \log\left(\varepsilon^{-1}\right) \log^{(2)}\left(\varepsilon^{-1}\right) + O_{\varepsilon \to 0}\left(\log\left(\varepsilon^{-1}\right)\right). \end{aligned}$$

## C.3 Proofs of Corollaries 13 and 14

First, note that $\psi(t) = ct(\log(t) - c')$ verifies the properties of decay rate functions with the choice $t^* := 2^{c'}$. We can therefore apply Theorem 9 with $\mathbb{K} = \mathbb{C}$ to get

$$H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) = \alpha\, d_\varepsilon \{\delta(d_\varepsilon) + O_{\varepsilon \to 0}(\log(d_\varepsilon) + \zeta(d_\varepsilon))\}, \tag{78}$$

where

$$d_\varepsilon = \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right) + \log(c_0)\right] + O_{\varepsilon \to 0}(1) = \psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right)\right] + O_{\varepsilon \to 0}(1). \tag{79}$$

Here, we made use of the sublinearity of $\psi^{(-1)}$ to rid ourselves of the $\log(c_0)$ term.

The following two lemmata are needed in the study of the asymptotic behavior of the right-hand-side in (78).

**Lemma 36** (Asymptotic behavior of the $\psi$-difference function)**.** *Let $c, c' > 0$ and consider the function $\psi\colon t \mapsto ct(\log(t) - c')$. The $\psi$-difference function $\zeta$ scales according to*

$$\zeta(d) = O_{d \to \infty}(\log(d)).$$

*Proof.* See Appendix C.6.7. □

**Lemma 37** (Asymptotic behavior of the $\psi$-average function)**.** *Let $c, c' > 0$ and consider the function $\psi\colon t \mapsto ct(\log(t) - c')$. The $\psi$-average function $\delta$ scales according to*

$$\delta(d) = \frac{cd \log(d)}{2} + \frac{c\, d\, (\ln^{-1}(2) - 2\, c')}{4} + O_{d \to \infty}(\log(d)).$$

*Proof.* See Appendix C.6.8. □

Using Lemma 36, we can now express (78) according to

$$H(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) = \alpha\, d_\varepsilon \{\delta(d_\varepsilon) + O_{\varepsilon \to 0}(\log(d_\varepsilon))\}. \tag{80}$$

Furthermore, direct application of Lemma 37 yields

$$\delta(d_\varepsilon) = \frac{c\, d_\varepsilon \log(d_\varepsilon)}{2} + \frac{c\, d_\varepsilon\, (\ln^{-1}(2) - 2\, c')}{4} + O_{\varepsilon \to 0}(\log(d_\varepsilon)),$$

so that (80) becomes

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) = \frac{\alpha c d_\varepsilon^2 \log(d_\varepsilon)}{2} + \frac{\alpha c d_\varepsilon^2(\ln^{-1}(2)-2c')}{4} + O_{\varepsilon\to 0}(d_\varepsilon \log(d_\varepsilon)) \tag{81}$$

$$= \frac{\alpha\, c}{2}\, d_\varepsilon^2 \left[\log(d_\varepsilon) + \frac{\ln^{-1}(2) - 2\,c'}{2}\right]\left[1 + O_{\varepsilon\to 0}\left(\frac{1}{d_\varepsilon}\right)\right]. \tag{82}$$

Inserting (79) in (81)-(82), we get

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) = \frac{\alpha\, c}{2}\left(\psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right)\right] + O_{\varepsilon\to 0}(1)\right)^2 \tag{83}$$

$$\times \left[\log\left(\psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right)\right] + O_{\varepsilon\to 0}(1)\right) + \frac{\ln^{-1}(2) - 2\,c'}{2}\right]$$

$$\times \left[1 + O_{\varepsilon\to 0}\left(\frac{1}{\psi^{(-1)}[\log(\varepsilon^{-1})]}\right)\right]$$

$$= \frac{\alpha\, c}{2}\left(\psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right)\right]\right)^2$$

$$\times \left[\log\left(\psi^{(-1)}\left[\log\left(\varepsilon^{-1}\right)\right]\right) + \frac{\ln^{-1}(2) - 2\,c'}{2}\right]$$

$$\times \left[1 + O_{\varepsilon\to 0}\left(\frac{1}{\log\left(\psi^{(-1)}[\log(\varepsilon^{-1})]\right)}\right)\right]. \tag{84}$$

Next, we characterize the inverse of the decay rate function $\psi(t) = ct(\log(t) - c')$.

**Lemma 38.** *Let $c, c' > 0$. The inverse of the function $\psi\colon t \mapsto ct(\log(t) - c')$ is given by*

$$\psi^{(-1)}\colon u \longmapsto \exp\left\{c'\ln(2) + W\left(\frac{u\ln(2)}{c}e^{-c'\ln(2)}\right)\right\},$$

*where $W$ denotes the Lambert $W$-function.*

*Proof.* See Appendix C.6.9. □

Using Lemma 38 in (83)-(84), we obtain

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) = \frac{\alpha\, c\, 2^{2c'-1}}{\ln(2)}\ \exp\left\{2W\left(\frac{\ln\left(\varepsilon^{-1}\right)}{2^{c'}c}\right)\right\}\left[W\left(\frac{\ln\left(\varepsilon^{-1}\right)}{2^{c'}c}\right) + \frac{1}{2}\right]$$

$$\times \left[1 + O_{\varepsilon\to 0}\left(\exp\left\{-W\left(\frac{\ln\left(\varepsilon^{-1}\right)}{2^{c'}c}\right)\right\}\right)\right],$$

which concludes the proof of Corollary 13.

For the proof of Corollary 14, we first need a characterization of the asymptotic behavior of $\psi^{(-1)}$.

**Lemma 39.** *Let $c, c' > 0$. The inverse function of $\psi\colon t \mapsto ct(\log(t) - c')$ satisfies the relation*

$$\psi^{(-1)}(u)^2 \log\left(\psi^{(-1)}(u)\right) = \frac{u^2}{c^2\log(u)}\left[1 + \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right].$$

*Proof.* See Appendix C.6.10. □

Now, starting from (83)-(84) and using Lemma 39 with $u = \log\left(\varepsilon^{-1}\right)$ together with the observation

$$O_{\varepsilon\to 0}\left(\frac{1}{\log\left(\psi^{(-1)}[\log(\varepsilon^{-1})]\right)}\right) = O_{\varepsilon\to 0}\left(\frac{1}{\log^{(2)}(\varepsilon^{-1})}\right) = o_{\varepsilon\to 0}\left(\frac{\log^{(3)}\left(\varepsilon^{-1}\right)}{\log^{(2)}\left(\varepsilon^{-1}\right)}\right),$$

which employs Lemma 45, we obtain

$$H(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) = \frac{\alpha\,\log^2\left(\varepsilon^{-1}\right)}{2\,c\log^{(2)}(\varepsilon^{-1})}\left[1+\frac{\log^{(3)}\left(\varepsilon^{-1}\right)}{\log^{(2)}\left(\varepsilon^{-1}\right)}+o_{\varepsilon\to 0}\left(\frac{\log^{(3)}\left(\varepsilon^{-1}\right)}{\log^{(2)}\left(\varepsilon^{-1}\right)}\right)\right].$$

This concludes the proof of Corollary 14.

## C.4 Proofs of Lemmata 18 and 22

### C.4.1 Proof of Lemma 18

We first observe that, for all $x \in \mathbb{R}$,

$$|f_1(x) - f_2(x)| = \left|\sum_{k=-\infty}^{\infty} a_k^{(1)} e^{ikx} - \sum_{k=-\infty}^{\infty} a_k^{(2)} e^{ikx}\right| \leq \sum_{k=-\infty}^{\infty}\left|a_k^{(1)} - a_k^{(2)}\right| = \left\|a^{(1)} - a^{(2)}\right\|_{\ell_1(\mathbb{Z})}.$$

Upon taking the supremum over all $x \in \mathbb{R}$, we hence get the desired upper bound

$$d_{2\pi}(f_1, f_2) \leq \left\|a^{(1)} - a^{(2)}\right\|_{\ell_1(\mathbb{Z})}.$$

The sought lower bound is obtained through Parseval's identity according to

$$\left\|a^{(1)} - a^{(2)}\right\|_{\ell_2(\mathbb{Z})}^2 = \sum_{k=-\infty}^{\infty}\left|a_k^{(1)} - a_k^{(2)}\right|^2 = \frac{1}{2\pi}\int_0^{2\pi}|f_1(x) - f_2(x)|^2\,dx \leq \sup_{x\in\mathbb{R}}|f_1(x) - f_2(x)|^2 = d_{2\pi}(f_1, f_2)^2.$$

This concludes the proof.

### C.4.2 Proof of Lemma 22

We start by rewriting the metric $d_r$ according to

$$d_r(f_1, f_2) = \sup_{z\in D(0;r)}|f_1(z) - f_2(z)| = \sup_{z\in D(0;r)}\left|\tilde{f}_1\left(\frac{z}{r}\right) - \tilde{f}_2\left(\frac{z}{r}\right)\right| = \sup_{z\in D(0;1)}\left|\tilde{f}_1(z) - \tilde{f}_2(z)\right|,$$

where $\tilde{f}_1$ and $\tilde{f}_2$ are analytic functions defined on the unit disk $D(0\,;1)$ via their power series expansions

$$\tilde{f}_1(z) = \sum_{k=0}^{\infty} \tilde{a}_k^{(1)} z^k \quad \text{and} \quad \tilde{f}_2(z) = \sum_{k=0}^{\infty} \tilde{a}_k^{(2)} z^k.$$

Applying [25, Theorem 3.1], we obtain

$$d_r(f_1, f_2) = \sup_{\|x\|_{\ell^2}=1} \left\| \left( \tilde{a}^{(1)} - \tilde{a}^{(2)} \right) \star x \right\|_{\ell^2}, \tag{85}$$

where $\star$ denotes the convolution operator (for sequences). The desired upper bound now follows directly from (85) by Young's convolution inequality according to

$$\begin{aligned} d_r(f_1, f_2) &= \sup_{\|x\|_{\ell^2}=1} \left\| \left( \tilde{a}^{(1)} - \tilde{a}^{(2)} \right) \star x \right\|_{\ell^2} \\ &\leq \sup_{\|x\|_{\ell^2}=1} \|x\|_{\ell^2} \left\| \tilde{a}^{(1)} - \tilde{a}^{(2)} \right\|_{\ell^1} = \left\| \tilde{a}^{(1)} - \tilde{a}^{(2)} \right\|_{\ell^1}. \end{aligned}$$

The sought lower bound follows from (85) by particularizing the choice of $x$ to

$$x_0 \coloneqq (1, 0, 0, \dots) \in \ell^2,$$

so that

$$\begin{aligned} d_r(f_1, f_2) &= \sup_{\|x\|_{\ell^2}=1} \left\| \left( \tilde{a}^{(1)} - \tilde{a}^{(2)} \right) \star x \right\|_{\ell^2} \\ &\geq \left\| \left( \tilde{a}^{(1)} - \tilde{a}^{(2)} \right) \star x_0 \right\|_{\ell^2} = \left\| \tilde{a}^{(1)} - \tilde{a}^{(2)} \right\|_{\ell^2}. \end{aligned}$$

This concludes the proof.

## C.5 Proofs of Theorem 24 and Corollary 25

We start by developing material that is pertinent to both proofs. Let $\{a_k\}_{k\in\mathbb{N}}$ be the sequence of Taylor series coefficients of

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \quad \text{for all } z \in \mathbb{C}.$$

We start by defining the embedding

$$\iota\colon f \in \mathcal{F}_{\exp}^{A,C} \mapsto \{a_k\}_{k\in\mathbb{N}} \in \ell^\infty(\mathbb{N}),$$

where $\{a_k\}_{k\in\mathbb{N}} \in \ell^\infty(\mathbb{N})$ follows from arguments similar to those employed in the proof of Lemma 40. We wish to circumscribe and inscribe $\iota(\mathcal{F}_{\exp}^{A,C})$ with ellipsoids.

**Lemma 40.** *Let $A$ and $C$ be positive real constants. It holds that*

$$\iota\big(\mathcal{F}_{exp}^{A,C}\big) \subseteq \mathcal{E}_\infty(\{\mu_k\}_{k\in\mathbb{N}}),$$

*where*

$$\begin{cases} \mu_0 = C, \\ \mu_k = C \exp\{-k(\ln(k) - 1 - \ln(A))\}, & \text{for } k \geq 1. \end{cases} \tag{86}$$

*Proof.* See Appendix C.6.11. □

**Lemma 41.** *Let $A$ and $C$ be positive real constants and set*

$$\tilde{A} := \frac{A}{2} \quad \text{and} \quad \tilde{C} := C\left[\sup_{k\in\mathbb{N}^*} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k}\right]^{-1}. \tag{87}$$

*Then,*

$$\mathcal{E}_\infty(\{\tilde{\mu}_k\}_{k\in\mathbb{N}}) \subseteq \iota\big(\mathcal{F}_{exp}^{A,C}\big),$$

*where*

$$\begin{cases} \tilde{\mu}_0 = \tilde{C}, \\ \tilde{\mu}_k = \tilde{C}\exp\Big\{-k\Big(\ln(k) - 1 - \ln(\tilde{A})\Big)\Big\}, \quad \text{for } k \geq 1. \end{cases} \tag{88}$$

*Proof.* See Appendix C.6.12. □

Combining Lemmata 40 and 41 with Lemma 22, yields

$$H(\varepsilon; \mathcal{E}_\infty(\{\tilde{\mu}_k\}_{k\in\mathbb{N}}), \|\cdot\|_2) \leq H\big(\varepsilon; \mathcal{F}_{\text{exp}}^{A,C}, d_1\big) \tag{89}$$

$$\leq H(\varepsilon; \mathcal{E}_\infty(\{\mu_k\}_{k\in\mathbb{N}}), \|\cdot\|_1). \tag{90}$$

We are now ready to proceed to the proof of Theorem 24.

### C.5.1 Proof of Theorem 24

Observe that, from (88) and under the convention $0\log(0) = 0$, we have

$$\tilde{\mu}_k = \tilde{C}\exp\Big\{-k\Big(\ln(k) - 1 - \ln(\tilde{A})\Big)\Big\} = \tilde{C}\exp\{-\ln(2)\,\psi(k)\},$$

for all $k \in \mathbb{N}$, with

$$\psi(k) = c\,k\,(\log(k) - c'), \quad \text{where } c = 1 \text{ and } c' = \log\Big(e\tilde{A}\Big).$$

Application of Corollary 13 with $c = 1$ and $c' = \log\Big(e\tilde{A}\Big)$ now yields

$$H(\varepsilon; \mathcal{E}_\infty(\{\tilde{\mu}_k\}_{k\in\mathbb{N}}), \|\cdot\|_2) \tag{91}$$

$$= \Big(e\tilde{A}\Big)^2 \frac{\exp\{2\beta_1(\varepsilon)\}}{\ln(2)}\left[\beta_1(\varepsilon) + \frac{1}{2}\right]\big[1 + O_{\beta_1(\varepsilon)\to\infty}(\exp\{-\beta_1(\varepsilon)\})\big], \tag{92}$$

where

$$\beta_1(\varepsilon) := W\left(\frac{\ln(\varepsilon^{-1})}{e\,\tilde{A}}\right).$$

Likewise, upon application of Corollary 13 with $c = 1$ and $c' = \log(eA)$, we obtain

$$H(\varepsilon; \mathcal{E}_\infty(\{\mu_k\}_{k\in\mathbb{N}}), \|\cdot\|_1) \tag{93}$$

$$= (eA)^2 \frac{\exp\{2\beta_2(\varepsilon)\}}{\ln(2)}\left[\beta_2(\varepsilon) + \frac{1}{2}\right]\big[1 + O_{\beta_2(\varepsilon)\to\infty}(\exp\{-\beta_2(\varepsilon)\})\big], \tag{94}$$

where

$$\beta_2(\varepsilon) := W\left(\frac{\ln(\varepsilon^{-1})}{e\,A}\right).$$

Using (91)-(92) and (93)-(94) in (89)-(90), yields

$$\begin{aligned}
&\left(e\tilde{A}\right)^2 \frac{\exp\{2\beta_1(\varepsilon)\}}{\ln(2)}\left[\beta_1(\varepsilon)+\frac{1}{2}\right]\left[1+O_{\beta_1(\varepsilon)\to\infty}(\exp\{-\beta_1(\varepsilon)\})\right]\\
&\le H\left(\varepsilon;\mathcal{F}_{\exp}^{A,C},d_1\right)\\
&\le (eA)^2\frac{\exp\{2\beta_2(\varepsilon)\}}{\ln(2)}\left[\beta_2(\varepsilon)+\frac{1}{2}\right]\left[1+O_{\beta_2(\varepsilon)\to\infty}(\exp\{-\beta_2(\varepsilon)\})\right],
\end{aligned}$$

thereby concluding the proof.

### C.5.2 Proof of Corollary 25

Upon application of Corollary 14 with $c = 1$ and $c' = \log(e\tilde{A})$, we obtain

$$\begin{aligned}
&H(\varepsilon;\mathcal{E}_\infty(\{\tilde{\mu}_k\}_{k\in\mathbb{N}}),\|\cdot\|_2) &(95)\\
&= \frac{\log^2(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\left(1+\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}+o_{\varepsilon\to 0}\left(\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\right)\right). &(96)
\end{aligned}$$

Likewise, Corollary 14 with $c = 1$ and $c' = \log(eA)$, yields

$$\begin{aligned}
&H(\varepsilon;\mathcal{E}_\infty(\{\mu_k\}_{k\in\mathbb{N}}),\|\cdot\|_1) &(97)\\
&= \frac{\log^2(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\left(1+\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}+o_{\varepsilon\to 0}\left(\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\right)\right). &(98)
\end{aligned}$$

Using (95)-(96) and (97)-(98) in (89)-(90), results in

$$H\left(\varepsilon;\mathcal{F}_{\exp}^{A,C},d_1\right) = \frac{\log^2(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\left(1+\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}+o_{\varepsilon\to 0}\left(\frac{\log^{(3)}(\varepsilon^{-1})}{\log^{(2)}(\varepsilon^{-1})}\right)\right),$$

thereby concluding the proof.

## C.6 Proofs of Auxiliary Results

### C.6.1 Proof of Lemma 10

We argue that covering the ellipsoid $\mathcal{E}_p$ by balls of radius $\varepsilon$ essentially reduces to covering the corresponding finite-dimensional ellipsoid $\mathcal{E}_p^{d_\varepsilon}$ obtained by retaining the first $d_\varepsilon$ semi-axes of $\mathcal{E}_p$. More precisely, we combine the inequality

$$N(\varepsilon;\mathcal{E}_p,\|\cdot\|_q) \ge N\left(\varepsilon;\mathcal{E}_p^{d_\varepsilon},\|\cdot\|_q\right)$$

with Theorem 4 to obtain

$$N(\varepsilon;\mathcal{E}_p,\|\cdot\|_q)^{\frac{1}{\sigma_{\mathbb{K}}(d_\varepsilon)}}\,\varepsilon \ge \underline{\kappa}\,\sigma_{\mathbb{K}}(d_\varepsilon)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\bar{\mu}_{d_\varepsilon}.$$

Now, using the definition of $d_\varepsilon$ in (16), this yields

$$\varepsilon \ge \underline{\kappa}\,\overline{\kappa}^{-1}\,\mu_{d_\varepsilon},$$

which, upon setting $\underline{\mathfrak{c}} := \underline{\kappa}\,\overline{\kappa}^{-1}$ and noting that (11) then implies $\underline{\mathfrak{c}} \le 1$, concludes the proof.

### C.6.2 Proof of Lemma 11

The proof is effected by reducing the infinite-dimensional covering problem to a finite-dimensional one followed by application of Theorem 5.

**Lemma 42.** *Let $p, q \in [1, \infty]$ and $\rho > 0$ be real numbers, and let $d \geq 1$ be an integer. Let $\mathcal{E}_p$ be an infinite-dimensional ellipsoid with exponentially decaying semi-axes $\{\mu_n\}_{n\in\mathbb{N}^*}$, and let $\mathcal{E}_p^d$ be the d-dimensional ellipsoid obtained by retaining the first d semi-axes of $\mathcal{E}_p$. Then, there exist an integer $d^*$ and a constant $K \geq 1$ not depending on d, such that, for all $d \geq d^*$, it holds that*

$$N\Big(\rho\,;\mathcal{E}_p^d, \|\cdot\|_q\Big) \geq N(\bar{\rho}\,;\mathcal{E}_p, \|\cdot\|_q), \quad \textit{with } \bar{\rho} := \begin{cases} \left(\rho^q + K\mu_{d+1}^q\right)^{1/q}, & \textit{if } 1 \leq q < \infty, \\ \max\{\rho, \mu_{d+1}\}, & \textit{if } q = \infty. \end{cases}$$

*Proof.* See Appendix C.6.13. □

Now, let $d^*$ and $K \geq 1$ be as in Lemma 42 and let $\varepsilon_1^* > 0$ be such that $d_{\varepsilon_1^*} - 1 \geq d^*$. We can further assume, without loss of generality, that

$$\varepsilon > K^{1/q}\,\mu_{d_\varepsilon}, \tag{99}$$

as, otherwise, the statement of Lemma 11 follows trivially by setting $\bar{\mathfrak{c}} := K^{1/q} \geq 1$. Next, we define the radius

$$\rho := \begin{cases} \left(\varepsilon^q - K\mu_{d_\varepsilon}^q\right)^{1/q}, & \text{if } 1 \leq q < \infty, \\ \varepsilon, & \text{if } q = \infty, \end{cases} \tag{100}$$

which is positive by assumption (99). As a direct consequence of (100), we obtain

$$\varepsilon = \bar{\rho} := \begin{cases} \left(\rho^q + K\mu_{d_\varepsilon}^q\right)^{1/q}, & \text{if } 1 \leq q < \infty, \\ \varepsilon, & \text{if } q = \infty, \end{cases}$$

so that $N(\varepsilon;\mathcal{E}_p, \|\cdot\|_q) = N(\bar{\rho}\,;\mathcal{E}_p, \|\cdot\|_q)$. Applying Lemma 42 with $d = d_\varepsilon - 1$, then yields

$$N(\varepsilon;\mathcal{E}_p, \|\cdot\|_q) \leq N\Big(\rho;\mathcal{E}_p^{d_\varepsilon-1}, \|\cdot\|_q\Big). \tag{101}$$

Our strategy now consists of first upper-bounding $\rho$ and then using the definition (100) to obtain a corresponding bound on $\varepsilon$. Specifically, we employ Theorem 5 to obtain

$$\rho \leq \frac{\overline{\kappa}\,\sigma_{\mathbb{K}}(d_\varepsilon - 1)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\bar{\mu}_{d_\varepsilon-1}}{N\Big(\rho;\mathcal{E}_p^{d_\varepsilon-1}, \|\cdot\|_q\Big)^{\frac{1}{\sigma_{\mathbb{K}}(d_\varepsilon-1)}}}. \tag{102}$$

The assumption $\rho \leq 2\mu_{d_\varepsilon-1}$ required by Theorem 5 will be verified below, in (105). Further, by (101) and the fact that $H\big(\varepsilon;\mathcal{E}_p^{d_\varepsilon}, \|\cdot\|_q\big)$ grows superlinearly in $d_\varepsilon$, there exists $\varepsilon_2^* > 0$ such that the hypothesis (10) needed in Theorem 5, that is,

$$H\Big(\rho;\mathcal{E}_p^{d_\varepsilon}, \|\cdot\|_q\Big) \geq 2\,\sigma_{\mathbb{K}}(d_\varepsilon), \quad \text{for all } \rho \in [0, 2\mu_{d_\varepsilon}],$$

is satisfied for all $\varepsilon < \varepsilon_2^*$. Upon application of (101), the bound (102) becomes

$$\rho \leq \frac{\overline{\kappa}\,\sigma_{\mathbb{K}}(d_\varepsilon - 1)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\bar{\mu}_{d_\varepsilon-1}}{N(\varepsilon;\mathcal{E}_p, \|\cdot\|_q)^{\frac{1}{\sigma_{\mathbb{K}}(d_\varepsilon-1)}}}. \tag{103}$$

Moreover, we have

$$\frac{\overline{\kappa}\,\sigma_{\mathbb{K}}(d_\varepsilon-1)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\overline{\mu}_{d_\varepsilon-1}}{N(\varepsilon;\mathcal{E}_p,\|\cdot\|_q)^{\frac{1}{\sigma_{\mathbb{K}}(d_\varepsilon-1)}}} \leq \mu_{d_\varepsilon-1} \tag{104}$$

as a direct consequence of the definition (16) of $d_\varepsilon$. Combining (103) and (104), we now get an upper bound on the radius $\rho$ according to

$$\rho \overset{(103)}{\leq} \frac{\overline{\kappa}\,\sigma_{\mathbb{K}}(d_\varepsilon-1)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\,\overline{\mu}_{d_\varepsilon-1}}{N(\varepsilon;\mathcal{E}_p,\|\cdot\|_q)^{\frac{1}{\sigma_{\mathbb{K}}(d_\varepsilon-1)}}} \overset{(104)}{\leq} \mu_{d_\varepsilon-1}. \tag{105}$$

For all $\varepsilon < \varepsilon^* := \min\{\varepsilon_1^*, \varepsilon_2^*\}$, combining (105) with (100) now yields

$$\varepsilon = \begin{cases} \left(\rho^q + K\mu_{d_\varepsilon}^q\right)^{1/q} \leq \left(\mu_{d_\varepsilon-1}^q + K\mu_{d_\varepsilon}^q\right)^{1/q} \leq \overline{\mathfrak{c}}\mu_{d_\varepsilon-1}, & \text{if } 1 \leq q < \infty, \\ \rho \leq \mu_{d_\varepsilon-1}, & \text{if } q = \infty, \end{cases}$$

where we have used that the semi-axes are non-increasing and we set $\overline{\mathfrak{c}} := (1+K)^{1/q} \geq 1$. This concludes the proof.

### C.6.3 Proof of Lemma 17

We first define the sequences

$$\mu_n := Me^{-sn} \quad \text{and} \quad \tilde{\mu}_n := Me^{-s\lfloor n/2 \rfloor}, \quad \text{for all } n \in \mathbb{N}, \tag{106}$$

and fix $f \in \mathcal{A}_s(M)$ with Fourier series coefficients $a_k$. Then, we note that

$$\begin{aligned}\sum_{n=1}^{\infty}\left|\frac{\tilde{a}_n}{\tilde{\mu}_n}\right| &= \left|\frac{\tilde{a}_1}{\tilde{\mu}_1}\right| + \sum_{n=1}^{\infty}\left[\left|\frac{\tilde{a}_{2n}}{\tilde{\mu}_{2n}}\right| + \left|\frac{\tilde{a}_{2n+1}}{\tilde{\mu}_{2n+1}}\right|\right] \\ &= \left|\frac{a_0}{\mu_0}\right| + \sum_{n=1}^{\infty}\left[\left|\frac{a_n}{\mu_n}\right| + \left|\frac{a_{-n}}{\mu_n}\right|\right] = \sum_{k\in\mathbb{Z}}\left|\frac{a_k}{\mu_{|k|}}\right| = \left\|\left\{\frac{a_k}{\mu_{|k|}}\right\}_{k\in\mathbb{Z}}\right\|_{\ell_1(\mathbb{Z})},\end{aligned}$$

where $\{\tilde{a}_n\}_{n\in\mathbb{N}^*}$ is as defined in (27). Moreover, $\tilde{\mu}_n \geq \mu_n^{(1)}$ readily implies

$$\left\|\{\tilde{a}_n\}_{n\in\mathbb{N}^*}\right\|_{1,\mu^{(1)}} \geq \sum_{n=1}^{\infty}\left|\frac{\tilde{a}_n}{\tilde{\mu}_n}\right| = \left\|\left\{\frac{a_k}{\mu_{|k|}}\right\}_{k\in\mathbb{Z}}\right\|_{\ell_1(\mathbb{Z})}. \tag{107}$$

The right-hand-side in (107) can be dealt with using the following lemma.

**Lemma 43.** *Let $M$ and $s$ be positive real numbers, and let $f$ be a $2\pi$-periodic function with Fourier series coefficients $\{a_k\}_{k\in\mathbb{Z}}$ satisfying*

$$\left\|\left\{\frac{a_k}{\mu_{|k|}}\right\}_{k\in\mathbb{Z}}\right\|_{\ell_1(\mathbb{Z})} \leq 1,$$

*with the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ as defined in (106). Then, $f \in \mathcal{A}_s(M)$.*

*Proof.* See Appendix C.6.14. □

Combining Lemma 43 with (107), it follows that

$$\mathcal{E}_1\Big(\Big\{\mu_n^{(1)}\Big\}_{n\in\mathbb{N}^*}\Big) \subseteq \iota(\mathcal{A}_s(M)). \tag{108}$$

Next, upon observing that $\sqrt{2}\,\tilde{\mu}_n \leq \mu_n^{(2)}$, for all $n \in \mathbb{N}^*$, we readily obtain

$$\big\|\{\tilde{a}_n\}_{n\in\mathbb{N}^*}\big\|_{2,\mu^{(2)}} \leq \left(\sum_{n=1}^{\infty}\left|\frac{\tilde{a}_n}{\sqrt{2}\,\tilde{\mu}_n}\right|^2\right)^{1/2} = \left\|\left\{\frac{a_k}{\sqrt{2}\,\mu_{|k|}}\right\}_{k\in\mathbb{Z}}\right\|_{\ell_2(\mathbb{Z})}. \tag{109}$$

Next, we show that the right-most expression in (109) can be upper-bounded by 1.

**Lemma 44.** *Let $M$ and $s$ be positive real numbers, and consider $f \in \mathcal{A}_s(M)$ with Fourier series coefficients $\{a_k\}_{k\in\mathbb{Z}}$. Then, we have*

$$\left\|\left\{\frac{a_k}{\sqrt{2}\,\mu_{|k|}}\right\}_{k\in\mathbb{Z}}\right\|_{\ell_2(\mathbb{Z})} \leq 1,$$

*where the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ has been defined in (106).*

*Proof.* See Appendix C.6.15. ☐

Using Lemma 44 together with (109), we can conclude that

$$\iota(\mathcal{A}_s(M)) \subseteq \mathcal{E}_2\Big(\Big\{\mu_n^{(2)}\Big\}_{n\in\mathbb{N}^*}\Big). \tag{110}$$

The proof is finalized by combining (108) and (110).

### C.6.4 Proof of Lemma 21

Let $f$ be analytic on $D(0\,;r')$ with Taylor series coefficients $\{a_k\}_{k\in\mathbb{N}}$, and set $\tilde{a}_k := a_k r^k$, for all $k \in \mathbb{N}$. First, note that

$$d_{r'}(f,0) = \sup_{z\in D(0;r')} |f(z)|. \tag{111}$$

Application of Lemma 22 yields

$$M\left\|\left\{\frac{\tilde{a}_k}{\mu_k}\right\}_{k\in\mathbb{N}}\right\|_{\ell^2} = \left\|\left\{a_k r'^k\right\}_{k\in\mathbb{N}}\right\|_{\ell^2} \leq d_{r'}(f,0)$$
$$\leq \left\|\left\{a_k r'^k\right\}_{k\in\mathbb{N}}\right\|_{\ell^1} = M\left\|\left\{\frac{\tilde{a}_k}{\mu_k}\right\}_{k\in\mathbb{N}}\right\|_{\ell^1}.$$

For $f \in \mathcal{A}(r';M)$, we hence get

$$\left\|\left\{\frac{\tilde{a}_k}{\mu_k}\right\}_{k\in\mathbb{N}}\right\|_{\ell^2} \leq \frac{d_{r'}(f,0)}{M} \leq 1,$$

which, in turn, implies the inclusion relation

$$\iota\big(\mathcal{A}(r';M)\big) \subseteq \mathcal{E}_2(\{\mu_k\}_{k\in\mathbb{N}}). \tag{112}$$

Conversely, assuming that $\{\tilde{a}_k\}_{k\in\mathbb{N}} \in \mathcal{E}_1(\{\mu_k\}_{k\in\mathbb{N}})$, we have

$$d_{r'}(f,0) \leq M\left\|\left\{\frac{\tilde{a}_k}{\mu_k}\right\}_{k\in\mathbb{N}}\right\|_{\ell^1} \leq M,$$

which implies $f \in \mathcal{A}(r';M)$ and hence yields

$$\mathcal{E}_1(\{\mu_k\}_{k\in\mathbb{N}}) \subseteq \iota\big(\mathcal{A}(r';M)\big). \tag{113}$$

Combining (112) and (113) establishes the desired result.

### C.6.5 Proof of Lemma 30

The function

$$g\colon t \mapsto \frac{\sqrt{2\pi t}\, e^{\frac{1}{12t}}}{2^t}$$

is continuous on $[1, \infty)$ and satisfies

$$\lim_{t\to\infty} g(t) = 0.$$

Therefore, $g$ is bounded on $[1, \infty)$, i.e.,

$$\sup_{t\in[1,\infty)} g(t) < \infty.$$

This obviously implies

$$\sup_{k\in\mathbb{N}^*} g(k) < \infty.$$

We finally observe that

$$\sup_{k\in\mathbb{N}^*} g(k) \geq g(1) = \sqrt{\frac{\pi}{2}}\, e^{\frac{1}{12}} \geq 1,$$

which concludes the proof.

### C.6.6 Proof of Lemma 35

We start by noting that for $\mathbb{K} = \mathbb{R}$, it follows from the usual volume formula for unit balls in Euclidean spaces (see e.g. [20, Eq. (1)]) that

$$\left(V_{p,q}^{\mathbb{R},d}\right)^{\frac{1}{\sigma_{\mathbb{R}}(d)}} = \frac{\Gamma(1/p+1)}{\Gamma(1/q+1)} \left(\frac{\Gamma(d/q+1)}{\Gamma(d/p+1)}\right)^{\frac{1}{d}}, \tag{114}$$

where $\Gamma$ denotes the Euler gamma function. By the Stirling formula, we have

$$\Gamma(d/q+1) = \sqrt{\frac{2\pi d}{q}} \left(\frac{d}{qe}\right)^{d/q} (1 + O_{d\to\infty}(1/d)),$$

and

$$\Gamma(d/p+1) = \sqrt{\frac{2\pi d}{p}} \left(\frac{d}{pe}\right)^{d/p} (1 + O_{d\to\infty}(1/d)).$$

Taking ratios yields

$$\frac{\Gamma(d/q+1)}{\Gamma(d/p+1)} = \frac{p^{d/p+1/2}}{q^{d/q+1/2}} \left(\frac{d}{e}\right)^{d(1/q-1/p)} (1 + O_{d\to\infty}(1/d)). \tag{115}$$

Inserting (115) into (114) gives

$$\left(V_{p,q}^{\mathbb{R},d}\right)^{\frac{1}{\sigma_{\mathbb{R}}(d)}} = \frac{\Gamma(1/p+1)\, p^{1/p}}{\Gamma(1/q+1)\, q^{1/q}\, e^{(1/q-1/p)}} d^{(1/q-1/p)} (1 + O_{d\to\infty}(1/d)). \tag{116}$$

Taking the logarithm on both sides of (116) yields the desired result in the case $\mathbb{K} = \mathbb{R}$. Likewise, for $\mathbb{K} = \mathbb{C}$, we use the corresponding volume formula (see e.g. [21, Proposition 3.2.1]) to get

$$\left(V_{p,q}^{\mathbb{C},d}\right)^{\frac{1}{\sigma_{\mathbb{C}}(d)}} = \left(\frac{\Gamma(2/p+1)}{\Gamma(2/q+1)}\right)^{\frac{1}{2}} \left(\frac{\Gamma(2d/q+1)}{\Gamma(2d/p+1)}\right)^{\frac{1}{2d}}. \tag{117}$$

Following the same steps as in the case $\mathbb{K} = \mathbb{R}$ then yields the desired result.

### C.6.7 Proof of Lemma 36

The result follows directly from

$$\begin{aligned}\zeta(d) = \psi(d) - \psi(d-1) &= c\, d \log(d) - c\,(d-1)\log(d-1) - c\, c' \\ &= c\, d \log(d) - c\, d \log(d-1) + c\, \log(d-1) - c\, c' \\ &= -c\, d \log(1 - 1/d) + c\, \log(d-1) - c\, c' \\ &= c\, \log(d) + O_{d\to\infty}(1).\end{aligned}$$

### C.6.8 Proof of Lemma 37

We first note that

$$\begin{aligned}\delta(d) &= \frac{1}{d}\sum_{n=1}^{d}(\psi(d) - \psi(n)) && (118)\\ &= \frac{c}{d}\sum_{n=1}^{d}\big[d(\log(d) - c') - n(\log(n) - c')\big] \\ &= c\, d \log(d) - \frac{c}{d}\sum_{n=1}^{d} n \log(n) - \frac{c\, c'\, d}{2} + O_{d\to\infty}(1). && (119)\end{aligned}$$

Sum-integral comparisons now yield the lower bound

$$\begin{aligned}\sum_{n=1}^{d} n \log(n) \geq \int_{1}^{d} t \log(t)\, dt &= \frac{1}{2}\left[d^2 \log(d) - \frac{d^2}{2\ln(2)} + \frac{1}{2\ln(2)}\right] \\ &= \frac{d^2 \log(d)}{2} - \frac{d^2}{4\ln(2)} + O_{d\to\infty}(1)\end{aligned}$$

and the upper bound

$$\begin{aligned}\sum_{n=1}^{d} n \log(n) \leq \int_{2}^{d+1} t \log(t)\, dt &= \frac{1}{2}\left[(d+1)^2 \log(d+1) - \frac{(d+1)^2}{2\ln(2)} - \frac{2}{\ln(2)}\right] \\ &= \frac{d^2 \log(d)}{2} - \frac{d^2}{4\ln(2)} + O_{d\to\infty}(d \log(d)).\end{aligned}$$

Combining these bounds, we obtain

$$\sum_{n=1}^{d} n \log(n) = \frac{d^2 \log(d)}{2} - \frac{d^2}{4\ln(2)} + O_{d\to\infty}(d \log(d)),$$

which, when inserted into (118)-(119) yields the desired result

$$\delta(d) = \frac{cd \log(d)}{2} + \frac{c\, d\,(\ln^{-1}(2) - 2\, c')}{4} + O_{d\to\infty}(\log(d)).$$

### C.6.9 Proof of Lemma 38

Starting from the expression

$$\left(\psi \circ \psi^{-1}\right)(u) = u, \quad \text{for all } u \in (0,\infty) \text{ and with } \psi(t) = ct(\log(t) - c'),$$

it follows that

$$u = c\,\psi^{(-1)}(u)\Big(\log\Big(\psi^{(-1)}(u)\Big) - c'\Big). \tag{120}$$

Setting

$$x := \ln\Big(\psi^{(-1)}(u)\Big), \tag{121}$$

we can rewrite (120) according to

$$u = c\,e^{x}\left(\frac{x}{\ln(2)} - c'\right), \quad \text{or equivalently,} \quad x = \frac{u\ln(2)}{c}e^{-x} + c'\ln(2). \tag{122}$$

Next, applying Lemma 28 with

$$a = \frac{u\ln(2)}{c} \quad \text{and} \quad b = c'\ln(2),$$

the solution of (122) can be expressed as

$$x = c'\ln(2) + W\left(\frac{u\ln(2)}{c}e^{-c'\ln(2)}\right),$$

which, when combined with the definition (121), yields the desired result

$$\psi^{(-1)}(u) = \exp\left\{c'\ln(2) + W\left(\frac{u\ln(2)}{c}e^{-c'\ln(2)}\right)\right\}.$$

### C.6.10 Proof of Lemma 39

We start with a result on the asymptotic behavior of $\psi^{(-1)}$.

**Lemma 45.** *Let $c, c' > 0$. The inverse of $\psi\colon t \mapsto ct(\log(t) - c')$ satisfies*

$$\psi^{(-1)}(u) = \frac{u}{c\log(u)} + \frac{u\log^{(2)}(u)}{c\log^2(u)} + o_{u\to\infty}\left(\frac{u\log^{(2)}(u)}{\log^2(u)}\right).$$

*Proof.* See Appendix C.6.16. □

We note that by Lemma 45

$$\psi^{(-1)}(u)^2 = \left[\frac{u}{c\log(u)}\right]^2\left[1 + \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right]^2 \tag{123}$$

$$= \left[\frac{u}{c\log(u)}\right]^2\left[1 + 2\,\frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right], \tag{124}$$

as well as

$$\log\Big(\psi^{(-1)}(u)\Big) = \log\left[\frac{u}{c\log(u)}\right] + \log\left[1 + \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right] \tag{125}$$

$$= \log(u) - \log^{(2)}(u) - \log(c) + \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)$$

$$= \log(u)\left[1 - \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right]. \tag{126}$$

We further observe that

$$\left[1+2\,\frac{\log^{(2)}(u)}{\log(u)}+o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right]\left[1-\frac{\log^{(2)}(u)}{\log(u)}+o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right] \tag{127}$$

$$=1+\frac{\log^{(2)}(u)}{\log(u)}+o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right). \tag{128}$$

Combining (123)-(124), (125)-(126), and (127)-(128), we obtain

$$\begin{aligned}\psi^{(-1)}(u)^2\log\left(\psi^{(-1)}(u)\right) &=\left[\frac{u}{c\log(u)}\right]^2\log(u)\left[1+\frac{\log^{(2)}(u)}{\log(u)}+o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right]\\ &=\frac{u^2}{c^2\log(u)}\left[1+\frac{\log^{(2)}(u)}{\log(u)}+o_{u\to\infty}\left(\frac{\log^{(2)}(u)}{\log(u)}\right)\right],\end{aligned}$$

which finalizes the proof.

### C.6.11 Proof of Lemma 40

Given $f\in\mathcal{F}_{\exp}^{A,C}$ with Taylor series coefficients $\{a_k\}_{k\in\mathbb{N}}$, we need to show that

$$\begin{cases}|a_0|\le C,\\ |a_k|\le C\exp\{-k(\ln(k)-1-\ln(A))\}, & \text{for } k\ge 1.\end{cases}$$

From (45) it follows that

$$\sup_{z\in D(0;R)}|f(z)|\le Ce^{AR},\quad \forall R>0. \tag{129}$$

Applying Cauchy's estimate (47) together with (129), we get

$$|a_k|\le\frac{Ce^{AR}}{R^k},\quad \forall R>0.$$

As this bound holds for all values of $R>0$, it holds in particular for the value minimizing the upper bound, or, more specifically, we have

$$|a_k|\le\inf_{R>0}C\exp\{AR-k\ln(R)\}. \tag{130}$$

For $k=0$, we readily get

$$|a_0|\le C,$$

as desired. It remains to consider the case $k\ge 1$. A straightforward calculation reveals that the infimum in (130) is attained at

$$R_k:=\frac{k}{A}>0,$$

which, upon insertion into (130), yields

$$\begin{aligned}|a_k|&\le C\exp\{AR_k-k\ln(R_k)\} &(131)\\ &=C\exp\{-k(\ln(k)-1-\ln(A))\}. &(132)\end{aligned}$$

This concludes the proof.

### C.6.12 Proof of Lemma 41

We first note that $\tilde{C}$ is well-defined owing to Lemma 30. Next, consider a sequence $\{a_k\}_{k\in\mathbb{N}}$ satisfying

$$\begin{cases} |a_0| \le \tilde{C}, \\ |a_k| \le \tilde{C} \exp\Big\{-k\Big(\ln(k) - 1 - \ln\Big(\tilde{A}\Big)\Big)\Big\}, \quad \text{for } k \ge 1. \end{cases} \tag{133}$$

Under the definition

$$f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

we need to show that

$$|f(z)| \le C e^{A|z|}, \quad \text{for all } z \in \mathbb{C}.$$

To this end, fix $z \in \mathbb{C}$ and start with the following chain of inequalities

$$|f(z)| = \left|\sum_{k=0}^{\infty} a_k z^k\right| \le \sum_{k=0}^{\infty} |a_k||z|^k \tag{134}$$

$$\overset{(133)}{\le} \tilde{C} + \sum_{k=1}^{\infty} \tilde{C} \exp\Big\{-k\Big(\ln(k) - 1 - \ln\Big(\tilde{A}\Big)\Big)\Big\}|z|^k$$

$$= \tilde{C}\left(1 + \sum_{k=1}^{\infty} \frac{e^k \tilde{A}^k |z|^k}{k^k}\right). \tag{135}$$

Applying Stirling's inequality, we obtain

$$\frac{e^k \tilde{A}^k |z|^k}{k^k} \le \frac{A^k |z|^k}{k!} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k}, \quad \text{for all } k \ge 1. \tag{136}$$

Using (136) in (134)-(135), we get

$$\begin{aligned} |f(z)| &\le \tilde{C}\left(1 + \sum_{k=1}^{\infty} \frac{A^k |z|^k}{k!} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k}\right) \\ &\le \tilde{C}\left(1 + \sup_{k\in\mathbb{N}^*} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k} \sum_{k=1}^{\infty} \frac{A^k|z|^k}{k!}\right) \\ &\le \tilde{C}\left[\sup_{k\in\mathbb{N}^*} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k}\right] \sum_{k=0}^{\infty} \frac{A^k |z|^k}{k!} = C e^{A|z|}, \end{aligned}$$

where the last inequality relies on the property

$$\sup_{k\in\mathbb{N}^*} \frac{\sqrt{2\pi k}\, e^{\frac{1}{12k}}}{2^k} \ge 1$$

established in Lemma 30. This concludes the proof.

### C.6.13 Proof of Lemma 42

Let $\{x^1, \dots, x^N\}$ be a $\rho$-covering of $\mathcal{E}_p^d$ with respect to the $\|\cdot\|_q$-norm. We introduce the vector obtained by completing the coordinates of $x^i$ with infinitely many zeros as

$$\bar{x}^i := (x_1^i, \dots, x_d^i, 0, \dots), \quad \text{for all } i \in \{1, \dots, N\}.$$

The proof is then effected by showing that $\{\bar{x}^1, \dots, \bar{x}^N\}$ is a $\bar{\rho}$-covering of $\mathcal{E}_p$. To this end, take $\bar{x} \in \mathcal{E}_p$ and define the vector

$$x := (\bar{x}_1, \dots, \bar{x}_d)$$

obtained by retaining the first $d$ components of $\bar{x}$. By definition, $x \in \mathcal{E}_p^d$. As $\{x^1, \dots, x^N\}$ is a $\rho$-covering of $\mathcal{E}_p^d$ in the $\|\cdot\|_q$-norm, there exists an index $j \in \{1, \dots, N\}$ such that

$$\left\|x - x^j\right\|_q \leq \rho. \tag{137}$$

Now, consider the case $q \in [1, \infty)$, and observe that

$$\left\|\bar{x} - \bar{x}^j\right\|_q^q = \sum_{n=1}^{d} \left|\bar{x}_n - \bar{x}_n^j\right|^q + \sum_{n=d+1}^{\infty} |\bar{x}_n|^q. \tag{138}$$

The first term on the right-hand-side of (138) equals $\|x - x^j\|_q^q$ and can hence be upper-bounded by $\rho^q$ owing to (137). We next observe that, thanks to $\bar{x} \in \mathcal{E}_p$, $|\bar{x}_n| \leq \mu_n$, for $n \in \mathbb{N}^*$. This immediately implies

$$\sum_{n=d+1}^{\infty} |\bar{x}_n|^q \leq \sum_{n=d+1}^{\infty} \mu_n^q = \mu_{d+1}^q \sum_{n=d+1}^{\infty} \left(\frac{\mu_n}{\mu_{d+1}}\right)^q. \tag{139}$$

We will need to further upper-bound the right-hand-side in (139), which will be effected through the following result.

**Lemma 46.** *Let $\psi$ be a decay rate function. For every sequence $\{\mu_n\}_{n\in\mathbb{N}^*}$ of $\psi$-exponentially decaying semi-axes, there exist a positive real number $c$ and an $n^* \in \mathbb{N}^*$ such that, for all pairs of integers $n$ and $m$ satisfying $n \geq m \geq n^*$, it holds that*

$$\frac{\mu_n}{\mu_m} \leq 2^{c(m-n)}.$$

*Proof.* See Appendix C.6.17. □

With $n^*$ according to Lemma 46, setting $d^* := n^* - 1$, it follows from Lemma 46 that

$$\sum_{n=d+1}^{\infty} \left(\frac{\mu_n}{\mu_{d+1}}\right)^q \leq \sum_{n=d+1}^{\infty} 2^{cq(d+1-n)} = \sum_{n=0}^{\infty} 2^{-cqn} = \frac{2^{cq}}{2^{cq} - 1}, \tag{140}$$

for all $d \geq d^*$. Upon defining the constant

$$K := \frac{2^{cq}}{2^{cq} - 1} \geq 1,$$

we get, from (139) and (140), the bound

$$\sum_{n=d+1}^{\infty} |\bar{x}_n|^q \leq K \mu_{d+1}^q.$$

Consequently, (138) becomes

$$\left\|\bar{x} - \bar{x}^j\right\|_q^q \leq \rho^q + K \mu_{d+1}^q = \bar{\rho}^q,$$

which shows that $\{\bar{x}^1, \dots, \bar{x}^N\}$ is a $\bar{\rho}$-covering of $\mathcal{E}_p$. Finally noting that $K$ does not depend on $d$ as required, the result is established for $q \in [1, \infty)$. For $q = \infty$, we have

$$\left\|\bar{x} - \bar{x}^j\right\|_q = \sup_{n\in\mathbb{N}^*} \left|\bar{x}_n - \bar{x}_n^i\right| = \max\left\{\max_{n=1,\dots,d} \left|\bar{x}_n - \bar{x}_n^i\right|, \sup_{n\geq d+1} |\bar{x}_n|\right\}. \tag{141}$$

As $\max_{n=1,\dots,d} \left|\bar{x}_n - \bar{x}_n^i\right| \leq \rho$ thanks to (137) and $\sup_{n\geq d+1} |\bar{x}_n| \leq \mu_{d+1}$, it follows that $\{\bar{x}^1, \dots, \bar{x}^N\}$ is a $\max\{\rho, \mu_{d+1}\}$-covering of $\mathcal{E}_p$.

### C.6.14 Proof of Lemma 43

Consider the analytic extension of $f$ to $\mathcal{S}$ (recall Definition 15), i.e.,

$$g\colon z \in \mathcal{S} \mapsto \sum_{k=-\infty}^{\infty} a_k e^{ikz},$$

and write $z = x + iy$, with $|y| < s$. Then, it follows from

$$\sum_{k=-\infty}^{\infty} \left| a_k e^{ik(x+iy)} \right| \leq \sum_{k=-\infty}^{\infty} |a_k| e^{s|k|} = M \left\| \left\{ \frac{a_k}{\mu_{|k|}} \right\}_{k\in\mathbb{Z}} \right\|_{\ell_1(\mathbb{Z})} \leq M,$$

that the infinite series defining $g$ is absolutely and uniformly convergent on $\mathcal{S}$, which implies analyticity of $g$ on $\mathcal{S}$. In particular, as $g = f$ on $\mathbb{R} \subseteq \mathcal{S}$, we have that $g$ is the analytic continuation of $f$ to $\mathcal{S}$. It remains to prove that $\sup_{z\in\mathcal{S}} |g(z)| \leq M$. To this end, fix $s' \in \mathbb{R}$ such that $|s'| < s$ and apply Lemma 48 which yields

$$|g(x + is')| = \left| \sum_{k=-\infty}^{\infty} a_k e^{ikx - ks'} \right| \leq \sum_{k=-\infty}^{\infty} |a_k| e^{s|k|} \tag{142}$$

$$= M \left\| \left\{ \frac{a_k}{\mu_{|k|}} \right\}_{k\in\mathbb{Z}} \right\|_{\ell_1(\mathbb{Z})} \leq M. \tag{143}$$

Taking the supremum in (142)-(143), according to

$$\sup_{z\in\mathcal{S}} |g(z)| = \sup_{x\in\mathbb{R}} \sup_{|s'|<s} |f(x + is')| \leq M,$$

concludes the proof.

### C.6.15 Proof of Lemma 44

Fix $s' \in (0; s)$ and define

$$\mu'_n := M e^{-s'n}, \quad \text{for all } n \in \mathbb{N}.$$

We start with the observation that

$$M^2 \left\| \left\{ \frac{a_k}{\mu'_{|k|}} \right\}_{k\in\mathbb{Z}} \right\|_{\ell_2(\mathbb{Z})}^2 = \sum_{k=-\infty}^{\infty} \left| a_k\, e^{|k|s'} \right|^2 \tag{144}$$

$$\leq \sum_{k=-\infty}^{\infty} \left| a_k\, e^{-ks'} \right|^2 + \sum_{k=-\infty}^{\infty} \left| a_k\, e^{ks'} \right|^2. \tag{145}$$

Next, using Parseval's identity combined with Lemma 48, we get

$$\sum_{k=-\infty}^{\infty} \left| a_k\, e^{-ks'} \right|^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| f(x + is') \right|^2 dx \leq \sup_{z\in\mathcal{S}} |f(z)|^2 \leq M^2. \tag{146}$$

Therefore, upon taking the limit $s' \to s$ in (146), the sequence of partial sums $\{\sum_{k=-N}^{N} \left| a_k\, e^{-ks} \right|^2\}_{N\in\mathbb{N}^*}$ is non-decreasing and bounded by $M^2$. By [31, Theorem 3.14], it hence converges to a limit which is also bounded by $M^2$, i.e., we have

$$\sum_{k=-\infty}^{\infty} \left| a_k\, e^{-ks} \right|^2 \leq M^2. \tag{147}$$

Likewise, one can show that

$$\sum_{k=-\infty}^{\infty} \left| a_k \, e^{ks} \right|^2 \leq M^2. \tag{148}$$

Using (147) and (148) in (144)-(145) hence yields

$$\left\| \left\{ \frac{a_k}{\sqrt{2} \mu_{|k|}} \right\}_{k \in \mathbb{Z}} \right\|_{\ell_2(\mathbb{Z})} \leq 1,$$

which concludes the proof.

### C.6.16 Proof of Lemma 45

We start by observing that

$$\begin{aligned}
&\ln\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c \ln\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c} \right)} \right) \\
&= -c' \ln(2) + \ln\left( \frac{u}{c[\log(u) + O_{u\to\infty}(1)]} \right) && (149) \\
&= -c' \ln(2) + \ln\left( \frac{u}{c \log(u)} \left[ 1 + O_{u\to\infty}\left( \frac{1}{\log(u)} \right) \right] \right) \\
&= -c' \ln(2) + \ln\left( \frac{u}{c \log(u)} \right) + O_{u\to\infty}\left( \frac{1}{\log(u)} \right). && (150)
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\frac{\ln^{(2)}\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c} \right)}{\ln\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c} \right)} &= \frac{\ln^{(2)}(u) + O_{u\to\infty}\left( \frac{1}{\ln(u)} \right)}{\ln(u) + O_{u\to\infty}(1)} && (151) \\
&= \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left( \frac{\log^{(2)}(u)}{\log(u)} \right). && (152)
\end{aligned}$$

Combining (149)-(150) and (151)-(152) with (48) in Lemma 29, we obtain

$$\begin{aligned}
&W\left( \frac{u \ln(2)}{c} e^{-c' \ln(2)} \right) && (153) \\
&= \ln\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c \ln\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c} \right)} \right) + \frac{\ln^{(2)}\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c} \right)}{\ln\left( \frac{u \ln(2) \, e^{-c' \ln(2)}}{c} \right)} (1 + o_{u\to\infty}(1)) \\
&= -c' \ln(2) + \ln\left( \frac{u}{c \log(u)} \right) + \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left( \frac{\log^{(2)}(u)}{\log(u)} \right) \\
&= -c' \ln(2) + \ln\left( \frac{u}{c \log(u)} \right) + \ln\left\{ 1 + \frac{\log^{(2)}(u)}{\log(u)} + o_{u\to\infty}\left( \frac{\log^{(2)}(u)}{\log(u)} \right) \right\} \\
&= -c' \ln(2) + \ln\left\{ \frac{u}{c \log(u)} + \frac{u \log^{(2)}(u)}{c \log^2(u)} + o_{u\to\infty}\left( \frac{u \log^{(2)}(u)}{\log^2(u)} \right) \right\}. && (154)
\end{aligned}$$

Using (153)-(154) in the expression for $\psi^{(-1)}$ according to Lemma 38, yields

$$\psi^{(-1)}(u) = \exp\left\{c' \ln(2) + W\left(\frac{u \ln(2)}{c} e^{-c' \ln(2)}\right)\right\}$$
$$= \frac{u}{c \log(u)} + \frac{u \log^{(2)}(u)}{c \log^2(u)} + o_{u\to\infty}\left(\frac{u \log^{(2)}(u)}{\log^2(u)}\right),$$

thereby concluding the proof.

### C.6.17 Proof of Lemma 46

Let $t^*$ be the parameter of $\psi$ and set

$$n^* := \lceil t^* \rceil + 1.$$

Then, from (12), we get

$$\frac{\psi(m)}{m} \leq \frac{\psi(n)}{n},$$

which directly implies

$$\frac{\psi(m) - \psi(n)}{m} \leq \psi(n)\left(\frac{1}{n} - \frac{1}{m}\right),$$

or, equivalently,

$$\psi(n) - \psi(m) \geq m\, \psi(n)\left(\frac{1}{m} - \frac{1}{n}\right). \tag{155}$$

Furthermore, we have

$$\frac{\psi(n)}{n} \geq \frac{\psi(n^*)}{n^*} =: c > 0,$$

which, when used in (155), yields

$$\psi(n) - \psi(m) \geq c\, m\, n\left(\frac{1}{m} - \frac{1}{n}\right) = c\,(n - m).$$

Finally, invoking (13) results in

$$\frac{\mu_n}{\mu_m} = \exp\{-\ln(2)(\psi(n) - \psi(m))\} \leq 2^{c(m-n)},$$

which is the desired result.

### C.6.18 Statement and Proof of Lemma 47

**Lemma 47.** *Let $p, q \in [1, \infty]$ and let $\mathcal{E}_p$ be the infinite-dimensional ellipsoid with semi-axes $\{\mu_n\}_{n\in\mathbb{N}^*}$. Then, the effective dimension*

$$d_\varepsilon = \min\left\{k \in \mathbb{N}^* \mid N(\varepsilon; \mathcal{E}_p, \|\cdot\|_q)^{\frac{1}{\sigma_{\mathbb{K}}(k)}}\, \mu_k \leq \overline{\kappa}\, \sigma_{\mathbb{K}}(k)^{\left(\frac{1}{q}-\frac{1}{p}\right)}\, \bar{\mu}_k\right\},$$

*with $\bar{\mu}_k$ denoting the geometric mean of the first $k$ semi-axes, for all $k \in \mathbb{N}^*$, is well-defined for all $\varepsilon > 0$ and $\lim_{\varepsilon\to 0} d_\varepsilon = \infty$. Here, $\overline{\kappa}$ is the constant defined in Theorem 5.*

*Proof.* To prove that the effective dimension is well-defined, it suffices to show that, for all $\varepsilon > 0$, there exists a $k \in \mathbb{N}^*$ such that

$$N(\varepsilon; \mathcal{E}_p, \|\cdot\|_q) \leq \left[ \overline{\kappa}\, \sigma_{\mathbb{K}}(k)^{\left(\frac{1}{q}-\frac{1}{p}\right)} \frac{\bar{\mu}_k}{\mu_k} \right]^{\sigma_{\mathbb{K}}(k)}, \tag{156}$$

which, in turn, is guaranteed by

$$\left[ \overline{\kappa}\, \sigma_{\mathbb{K}}(k)^{\left(\frac{1}{q}-\frac{1}{p}\right)} \frac{\bar{\mu}_k}{\mu_k} \right]^{\sigma_{\mathbb{K}}(k)} \xrightarrow{k\to\infty} \infty.$$

It is hence sufficient to establish that

$$\log\left(\frac{\bar{\mu}_k}{\mu_k}\right) + \left(\frac{1}{q} - \frac{1}{p}\right) \log(\sigma_{\mathbb{K}}(k)) \xrightarrow{k\to\infty} \infty,$$

which holds as a direct consequence of the identity

$$\log\left(\frac{\bar{\mu}_k}{\mu_k}\right) \overset{(13)}{=} \frac{1}{k} \sum_{n=1}^{k} [\psi(k) - \psi(n)] \geq \kappa(k-1),$$

for some $\kappa > 0$ independent of $k$, with the inequality relying on the assumption that $\psi$ is a decay rate function and therefore grows at least linearly.

Finally, $\lim_{\varepsilon\to 0} d_\varepsilon = \infty$ is an immediate consequence of the right hand side in (156) being finite for all $k \in \mathbb{N}^*$, while the left hand side goes to infinity as $\varepsilon$ approaches zero. ☐

### C.6.19 Statement and Proof of Lemma 48

**Lemma 48.** *Let $s$ be a positive real number and let $s' \in \mathbb{R}$ be such that $|s'| < s$. Consider the $2\pi$-periodic function $f$ analytic on the strip $\mathcal{S}$ of width $s$ from Definition 15. Then, the Fourier series coefficients of the function*

$$x \mapsto f(x + is') \quad \textit{are given by} \quad \left\{ a_k e^{-ks'} \right\}_{k\in\mathbb{Z}}.$$

*Proof.* We first define

$$\gamma \colon t \in [0,1] \mapsto \begin{cases} 8\pi t, & t \in [0, 1/4], \\ 2\pi + is'(4t-1), & t \in [1/4, 1/2], \\ 8\pi(3/4 - t) + is', & t \in [1/2, 3/4], \\ 4is'(1-t), & t \in [3/4, 1], \end{cases}$$

which is a closed contour in $\mathcal{S}$ (see Figure 4).

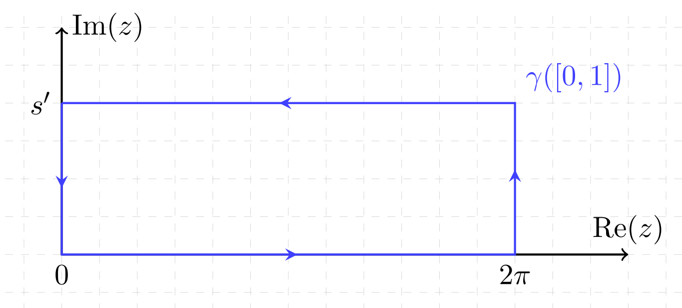


Figure 4: Plot of the contour $\gamma$.

As the function

$$z \mapsto f(z)\, e^{-ikz}$$

is analytic in $\mathcal{S}$ by assumption, we can apply Cauchy's integral theorem on the contour $\gamma$ to get

$$\int_{\gamma([0,1])} f(z)\, e^{-ikz}\, dz = 0, \quad \text{for all } k \in \mathbb{Z}. \tag{157}$$

Using the $2\pi$-periodicity

$$f(z)\, e^{-ikz} = f(z+2\pi)\, e^{-ikz} = f(z+2\pi)\, e^{-ik(z+2\pi)},$$

it follows from (157) that

$$\int_{\gamma([1/4,1/2])} f(z)\, e^{-ikz}\, dz + \int_{\gamma([3/4,1])} f(z)\, e^{-ikz}\, dz = 0, \quad \text{for all } k \in \mathbb{Z}. \tag{158}$$

Combining (157) and (158) yields, for all $k \in \mathbb{Z}$,

$$a_k = \frac{1}{2\pi} \int_0^{2\pi} f(x)\, e^{-ikx}\, dx = 4 \int_{\gamma([0,1/4])} f(z)\, e^{-ikz}\, dz \tag{159}$$

$$= -4 \int_{\gamma([1/2,3/4])} f(z)\, e^{-ikz}\, dz$$

$$= \frac{1}{2\pi} \int_0^{2\pi} f(x+is')\, e^{-ikx+ks'}\, dx. \tag{160}$$

Rearranging terms, we obtain

$$a_k e^{-ks'} = \frac{1}{2\pi} \int_0^{2\pi} f(x+is')\, e^{-ikx}\, dx, \quad \text{for all } k \in \mathbb{Z},$$

which concludes the proof. □