Abstract

We derive $L_\mu^r$-bracketing metric and sup-norm metric entropy rates of various subsets of general function spaces defined over $\mathbb{R}^d$ or, more generally, over Borel sets $\Omega$, exploiting results by Haroske and Triebel (1994, 2004). The function spaces covered are of (fractional) Sobolev, Besov, Hölder, Triebel, and Bessel Potential type. Applications to the theory of empirical processes are discussed. In particular, we show that (weighted) norm balls in the above mentioned spaces are Donsker classes uniformly in various sets $\mathcal{P}$ of probability measures.

MSC 2000 subject classification. Primary 60F17; secondary 46E35

Key words and phrases. Metric entropy with bracketing, uniform metric entropy, Hellinger metric, Sobolev, Besov, Hölder, Triebel, and Bessel Potential spaces, uniform Donsker class, Glivenko-Cantelli class, functions of bounded $p$-variation.

1 Introduction

In the theory of empirical processes (see, e.g., Dudley (1999)) the size of the function class indexing the empirical process plays a central role. The size, more precisely, the degree of compactness of the function class in some relevant topology, is measured by concepts like metric entropy (with or without bracketing). For example, many limit theorems for empirical processes (e.g., Glivenko-Cantelli and Donsker type results) are based on entropy conditions (see, e.g., the monographs by Pollard (1984), van der Vaart and Wellner (1996), and Dudley (1999)). The bracketing metric entropy concept has proved to be particularly useful in this context (see Dudley (1978, 1984), Alexander (1984), Ossiander (1987), Andersen, Giné, Ossiander, and Zinn (1988)). Another prototypical application of bracketing metric entropy is in the study of convergence rates and lower risk bounds of statistical estimators (e.g., Birgé and Massart (1993), van de Geer (1993), Wong and Shen (1995), summarized in van de Geer (2000)).

Classical results on (sup-norm) metric entropy bounds for ‘smooth’ function classes on bounded subsets of Euclidean space can be found in Kolmogorov and Tihomirov (1961) and Birman and Solomyak (1967); for a more recent contribution see also Birgé and Massart (2000). Bracketing metric entropy bounds for such function classes can then immediately be obtained from these sup-norm metric entropy bounds (cf. van der Vaart and Wellner (1996), Dudley (1999), van de Geer (2000)). The only result regarding bracketing metric entropy bounds for classes of ‘smooth’ functions with unbounded support we are aware of is given in van der Vaart (1994), which covers only a quite specific class of functions obtained from suitably pasting together functions that
are of Hölder-type on bounded subsets; see Section 3.3 for further discussion. [Also note that bracketing metric entropy bounds for classes of functions of bounded variation that are defined on the real line are available; see van de Geer (1991) and Remark 2 below.]

In the present paper we attempt a unified approach to bounding the bracketing metric entropy of a variety of function classes such as subsets of fractional Sobolev, Besov, Hölder, Triebel, and Bessel Potential spaces defined on \( \mathbb{R}^d \) or on arbitrary Borel subsets \( \Omega \) thereof. In this course we exploit results from the Fourier-analysis of such spaces (Haroske and Triebel (1994, 2004), Haroske (1995), Edmunds and Triebel (1996)). In Section 2 we obtain sup-norm metric entropy, \( L^r(\mu) \)-bracketing metric entropy, and uniform metric entropy bounds for (weighted) balls in the above mentioned spaces. Here the measure \( \mu \) is not necessarily a finite measure. The entropy rates are always of order \( \varepsilon^{-\delta} \) for some positive \( \delta \), where \( \delta \) depends on the ‘degree of smoothness’ of the function class and is connected to the measure \( \mu \) via a suitable integrability condition. These results also allow for classes of unbounded functions. A bound for the Hellinger bracketing metric entropy for classes of functions with bounded support is also given in Section 2.

Applications of the results in Section 2 to empirical process theory are discussed in Section 3. In particular, we show that (weighted) norm balls in Sobolev, Besov, Hölder, Triebel, and Bessel Potential spaces are Glivenko-Cantelli or Donsker classes uniformly in various sets \( \mathcal{P} \) of probability measures. These results are complemented by a uniform Donsker class result that is based on Dudley (1992) together with a theorem establishing the imbedding of certain Besov spaces into the space of functions of bounded \( p \)-variation, which is proved in the Appendix (Proposition 3). In the Appendix we also collect some technical background needed in the main body of the paper. Statistical applications of the results in the paper will be reported elsewhere.

### 2 Bracketting metric entropy bounds

**Definition 1** For a (non-empty) subset \( \mathcal{F} \) of a normed space \( (X, \| \cdot \|_X) \), let \( N(\varepsilon, \mathcal{F}, \| \cdot \|_X) \) denote the minimal covering number, i.e., the minimal number of closed balls of radius \( \varepsilon, 0 < \varepsilon < \infty \), (w.r.t. \( \| \cdot \|_X \)) needed to cover \( \mathcal{F} \). In accordance, let \( H(\varepsilon, \mathcal{F}, \| \cdot \|_X) = \log N(\varepsilon, \mathcal{F}, \| \cdot \|_X) \) be the metric entropy of the set \( \mathcal{F} \), where \( \log \) denotes the natural logarithm.

For a real-valued function \( h \) on a (non-empty) Borel subset \( \Omega \) of \( \mathbb{R}^d \), \( d \in \mathbb{N} \), and a (non-negative) Borel-measure \( \mu \) on \( \Omega \), we set \( \| h \|_{r, \mu} = \left( \int_\Omega |h|^r \, d\mu \right)^{1/r} \) for \( 1 \leq r \leq \infty \) with the understanding that \( \| h \|_{\infty, \mu} \) is the essential supremum of \( h \) w.r.t. \( \mu \). As usual, we denote by \( L^r(\mu) \) (or sometimes by \( L^r(\Omega, \mu) \)) the set of all real-valued Borel-measurable functions \( h \) on \( \Omega \) satisfying \( \| h \|_{r, \mu} < \infty \); furthermore, \( L^0(\Omega) \) denotes the set of real-valued Borel-measurable functions on \( \Omega \). We always use the term Borel measure to mean a nonnegative Borel measure, not necessarily finite or \( \sigma \)-finite.

**Definition 2** Let \( \Omega \) be a (non-empty) Borel subset of \( \mathbb{R}^d \). Given two Borel-measurable functions \( l, u : \Omega \to \mathbb{R} \), the bracket \([l, u]\) is the set of all functions \( f : \Omega \to \mathbb{R} \) with \( l \leq f \leq u \). Given a Borel-measure \( \mu \) on \( \Omega \) and \( 1 \leq r \leq \infty \), the \( L^r(\mu) \)-size of the bracket \([l, u]\) is defined as \( \| u - l \|_{r, \mu} \). The \( L^r(\mu) \)-bracketing number \( N_r(\varepsilon, \mathcal{F}, \| \cdot \|_{r, \mu}) \) of a (non-empty) set \( \mathcal{F} \subseteq L^0(\Omega) \) is the minimal number of brackets of \( L^r(\mu) \)-size less than or equal to \( \varepsilon, 0 < \varepsilon < \infty \), necessary to cover \( \mathcal{F} \). The logarithm of the bracketing number is called the \( L^r(\mu) \)-bracketing metric entropy \( H_r(\varepsilon, \mathcal{F}, \| \cdot \|_{r, \mu}) \).

In the above definitions it is implicitly understood that \( N(\varepsilon, \mathcal{F}, \| \cdot \|_X) \) and \( N_r(\varepsilon, \mathcal{F}, \| \cdot \|_{r, \mu}) \) are finite. For two real-valued functions \( a(\cdot) \) and \( b(\cdot) \), we write \( a(\varepsilon) \lesssim b(\varepsilon) \) if there exists a positive (finite) constant \( c \) not depending on \( \varepsilon \) such that \( a(\varepsilon) \leq cb(\varepsilon) \) holds for all \( \varepsilon > 0 \). If
$a(\varepsilon) \lesssim b(\varepsilon)$ and $b(\varepsilon) \lesssim a(\varepsilon)$ both hold we write $a(\varepsilon) \sim b(\varepsilon)$. (In abuse of notation, we shall also use this notation for sequences $a_k$ and $b_k$, $k \in \mathbb{N}$.) Furthermore, for two norms $\| \cdot \|_{X,1}$ and $\| \cdot \|_{X,2}$ on a vector space $X$, we write $\| \cdot \|_{X,1} \lesssim \| \cdot \|_{X,2}$ if $\| \cdot \|_{X,1} \leq c \| \cdot \|_{X,2}$ for a (finite) positive constant $c$, and we write $\| \cdot \|_{X,1} \sim \| \cdot \|_{X,2}$ if the norms are equivalent.

In what follows, let $C(\mathbb{R}^d)$ be the vector space of bounded continuous functions on $\mathbb{R}^d$ normed by the sup-norm $\| \cdot \|_{\infty}$. Furthermore, let $UC(\mathbb{R}^d)$ be the subspace of $C(\mathbb{R}^d)$ that consists of all uniformly continuous functions $f$ again equipped with the sup-norm. Attaching a subscript 0 to any of these two spaces denotes the respective subspaces of functions satisfying $\lim_{|x| \to \infty} f(x) = 0$ where $\| \cdot \|$ always denotes the Euclidean norm. To control tail behavior, we introduce the polynomial weighting function $(x)^\beta = (1 + \|x\|^2)^{\beta/2}$ parameterized by $\beta \in \mathbb{R}$, where $x$ is an element of $\mathbb{R}^d$.

### 2.1 Fractional Sobolev Spaces

For $1 \leq p < \infty$ and $s$ a nonnegative integer, let

$$W^{s,p}(\mathbb{R}^d) = \left\{ f \in L^0(\mathbb{R}^d) : \sum_{0 \leq |\alpha| \leq s} \| D^\alpha f \|_{p,\lambda} < \infty \right\}$$

be the classical Sobolev space of functions which have partial derivatives (in the generalized sense of distributions) up to and including order $s$ that belong to $L^p(\mathbb{R}^d, \lambda)$; see, e.g., Adams (1975, 3.1). Here $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index of nonnegative integers $\alpha_i$, $|\alpha| = \sum_{i=1}^d \alpha_i$, and $D^\alpha = \frac{\partial^{\alpha_i}}{(\partial x_1)^{\alpha_1} \ldots (\partial x_d)^{\alpha_d}}$ denotes the partial differential operator of order $|\alpha|$. Furthermore,

$$\| f \|_{s,p,\lambda} = \sum_{0 \leq |\alpha| \leq s} \| D^\alpha f \|_{p,\lambda}$$

defines the Sobolev (semi)norm on $W^{s,p}(\mathbb{R}^d)$.

The concept of classical Sobolev spaces with $s$ a nonnegative integer can be extended to so-called fractional Sobolev spaces where $s$ is real and nonnegative. These spaces will again be denoted by $W^{s,p}(\mathbb{R}^d)$, where $\| \cdot \|_{s,p,\lambda}$ again denotes the respective (semi)norm. For a detailed definition of fractional Sobolev spaces we refer the reader to Adams (1975, 7.35, 7.48). Since these definitions are somewhat involved we do not reproduce them here. We only mention that

$$W^{s,p}(\mathbb{R}^d) \subseteq W^{t,p}(\mathbb{R}^d) \subseteq W^{0,p}(\mathbb{R}^d) = L^p(\mathbb{R}^d, \lambda) \text{ for } s \geq t \geq 0.$$ 

With each element $f$ of $W^{s,p}(\mathbb{R}^d)$ any element of the entire equivalence class of functions in $L^0(\mathbb{R}^d)$ that are a.e. equal to $f$ also belongs to $W^{s,p}(\mathbb{R}^d)$. However, in case $s > d/p$, each such equivalence class contains exactly one continuous function $f$, and we shall then always view $W^{s,p}(\mathbb{R}^d)$ as only consisting of these continuous representatives; furthermore, any such $f$ is bounded and satisfies $\lim_{|x| \to \infty} f(x) = 0$. As a consequence, the Sobolev (semi)norm then actually is a norm. While derivatives of functions belonging to a Sobolev space are understood in the generalized sense, in case $s > d/p$ each continuous representative $f$ possesses in fact classical partial derivatives of order (at least) $s - d/p$. Therefore, the quantity $s - d/p$ (rather than $s$) is called the differential dimension of the Sobolev space. For convenience of the reader these well-known properties are proved in the Appendix, see Proposition 3, where we also show that any function in $W^{s,p}(\mathbb{R})$ is of bounded p-variation in case $s > 1/p$.

For $s - d/p > 0$ we shall consider the function class

3
\[ \mathcal{D}^{\beta}_{s,p} = \left\{ f \in L^0(\mathbb{R}^d) : \left\| f \cdot \langle x \rangle^{\beta} \right\|_{s,p,\lambda} \leq b \right\} \quad (0 \leq b < \infty) \]

with \(1 \leq p < \infty\) and \(\beta \in \mathbb{R}\). We shall suppress the dependence of \(\mathcal{D}^{\beta}_{s,p}\) on the radius \(b\) in the notation. We stress that we use the convention discussed in the preceding paragraph, i.e., that each \(f \cdot \langle x \rangle^{\beta} \in W^{s,p}(\mathbb{R}^d)\) is continuous, and hence so is any \(f \in \mathcal{D}^{\beta}_{s,p}\). Note that \(\mathcal{D}^{\beta}_{s,p} \subseteq W^{s,p}(\mathbb{R}^d)\) for \(\beta \geq 0\), but not for \(\beta < 0\). [As a point of interest we further note that, regardless of the value of \(\beta\), the set \(\mathcal{D}^{\beta}_{s,p}\) is contained in a norm ball in the weighted Sobolev space with weighting \(d\nu = \langle x \rangle^{\beta} \, dx\); conversely, any norm ball in this weighted Sobolev space is contained in a set of the form \(\mathcal{D}^{\beta}_{s,p}\). This is so since the weighted Sobolev norm \(\|\cdot\|_{s,p,\nu}\) is equivalent to the norm \(\left\| \langle \cdot \rangle^\beta \right\|_{s,p,\lambda}\), see Schmeisser and Triebel (1987), 5.1.4/6 and 5.1.4/7.]

Obviously, \(\mathcal{D}^{\beta}_{s,p}\) is a convex set. Finally, there exists a finite positive constant \(K\) such that \(K \langle x \rangle^{-\beta}\) is an envelope for \(\mathcal{D}^{\beta}_{s,p}\), i.e., \(|f(x)| \leq K \langle x \rangle^{-\beta}\) for all \(x \in \mathbb{R}^d\) and all \(f \in \mathcal{D}^{\beta}_{s,p}\); see Proposition 3 in the Appendix.

The following results will be given for fractional spaces with \(s > 0\). A reader only familiar with classical Sobolev spaces can ignore this extra generality and always think of \(s\) as an integer. We start with the following proposition which is a basic result that bounds the sup-norm metric entropy.

**Proposition 1** Let \(1 \leq p < \infty\), \(\beta > 0\), and \(s - d/p > 0\). Then \(\mathcal{D}^{\beta}_{s,p}\) is a relatively compact subset of the normed space \((UC_0(\mathbb{R}^d), \| \cdot \|_{\infty})\). For \(\beta > s - d/p\) we have

\[ H \left( \varepsilon, \mathcal{D}^{\beta}_{s,p}, \| \cdot \|_{\infty} \right) \leq \varepsilon^{-d/s}, \]

and for \(\beta < s - d/p\) we have

\[ H \left( \varepsilon, \mathcal{D}^{\beta}_{s,p}, \| \cdot \|_{\infty} \right) \leq \varepsilon^{-(\beta/(d+1/p))^{-1}}. \]

**Proof.** Observe that \(W^{s,p}(\mathbb{R}^d)\) coincides with the Besov space \(B^s_{pq}(\mathbb{R}^d)\) for non-integer \(s > 0\) by Triebel (1983, 2.5.7/9). For integer \(s > 0\), the space \(W^{s,p}(\mathbb{R}^d)\) is imbedded into the Besov space \(B^s_{pq}(\mathbb{R}^d)\) by Triebel (1983, 2.5.7/10). Hence, in both cases, the set \(\mathcal{D}^{\beta}_{s,p}\) is contained in a \(\left\| \langle \cdot \rangle \right\|_{B^s_{pq},\lambda}\)-bounded subset of a Besov space \(B^s_{pq}(\mathbb{R}^d)\) for suitable \(1 \leq q \leq \infty\). The result now follows from the first part of Theorem 2 in Section 2.2. ■

A direct metric entropy bound for the case \(\beta = s - d/p\) is involved. However, one can obtain a simple metric entropy bound for \(\mathcal{D}^{\beta}_{s,p}\) in this case by viewing it as a subset of \(\mathcal{D}^{\beta}_{t,p}\) with \(t < s\), and by applying the above proposition to \(\mathcal{D}^{\beta}_{t,p}\). (A similar remark applies to all the results below and will not be repeated.)

In the theory of empirical processes often bounds on the \(L^r(\mu)\)-bracketing metric entropy \((1 \leq r \leq \infty)\) are of interest. The following theorem builds on Proposition 1, but in contrast to this result, it also covers function classes containing unbounded functions, i.e., \(\mathcal{D}^{\beta}_{s,p}\) for \(\beta < 0\). Furthermore, in case \(\beta = 0\), it delivers bounds for the \(L^r(\mu)\)-bracketing metric entropy of an (unweighted) Sobolev ball in \(W^{s,p}(\mathbb{R}^d)\).

**Theorem 1** Let \(1 \leq p < \infty\), \(\beta \in \mathbb{R}\), and \(s - d/p > 0\). Suppose that for some Borel measure \(\mu\) on \(\mathbb{R}^d\) and some \(1 \leq r \leq \infty\) the moment condition \(\left\| \langle x \rangle^{\gamma - \beta} \right\|_{r,\mu} < \infty\) holds for some \(\gamma > 0\). If \(\gamma > s - d/p\), we have

\[ H_{\| \cdot \|} \left( \varepsilon, \mathcal{D}^{\beta}_{s,p}, \| \cdot \|_{r,\mu} \right) \lesssim \varepsilon^{-d/s}; \]

4
if $\gamma < s - d/p$, we have

$$H_{\varepsilon} \left( \varepsilon, \mathcal{D}^m_{s,p} \| \mathcal{L}_{r,\mu} \right) \lesssim \varepsilon^{-(\gamma/d+1/p)-1}.$$ 

**Proof.** By the same reasoning as in the proof of Proposition 1, the set $\mathcal{D}^m_{s,p}$ is contained in a $\left\| (\cdot) \langle x \rangle^2 \|_{B_{s,p,q,\Lambda}} \right.$-bounded subset of a Besov space $B_{pq}^s(\mathbb{R}^d)$ for suitable $1 \leq q \leq \infty$. The result now follows from the second part of Theorem 2 in Section 2.2. 

The moment condition constitutes a trade-off between the behavior of the envelope $K \langle x \rangle^{-\beta}$ for large $\| x \|$ and the tail-behavior of the measure $\mu$. While the moment condition becomes more lenient the smaller $\beta$ is, the rate bound deteriorates as $\beta$ becomes smaller. If $\beta > 0$ and $\mu$ is a finite measure (e.g., probability measure) or $r = \infty$, the moment condition is always satisfied at least for $\gamma = \beta$. If $\mu$ has an exponential moment, the condition $\| \langle x \rangle^{\gamma-\beta} \|_{r,\mu} < \infty$ is of course satisfied for any $\gamma$ and $\beta$ (and $r < \infty$). If $\mu$ is Lebesgue-measure, the moment condition holds for any $\gamma$ satisfying $0 < \gamma < \beta - d/r$.

While Theorem 1 above provides bracketing metric entropy bounds for given measure $\mu$, uniform versions of this theorem will be discussed in Section 2.3. Furthermore, sometimes bracketing metric entropy bounds for sets that are obtained from $\mathcal{D}^m_{s,p}$ by multiplication with a function $h$ are of interest. Such bounds can be obtained from a combination of Theorem 1 and Lemma 1 in Section 2.4 below.

**Remark 1** Unweighted Sobolev balls

In case $\beta = 0$, Theorem 1 imposes the moment condition $\| \langle x \rangle^\gamma \|_{r,\mu} < \infty$ for some $\gamma > 0$, which is not satisfied for every probability measure. If the entropy calculations and the imbedding theory of Haroske and Triebel (1994) used in the present paper could be generalized to weighting functions more general than polynomial ones, it might be possible to avoid any moment conditions and still obtain a polynomial bound for the bracketing metric entropy of $\mathcal{D}^d_{s,p}$ for each given probability measure $\mu$.□

**Remark 2** In the particular case $s \geq 1$, $p = 1$, $d = 1$, the Sobolev ball $\mathcal{D}^d_{s,1}$ is a bounded subset of the space $V_1(\mathbb{R})$ of functions of bounded variation (see the argument following (10) in the Appendix). An $L^2(\mu)$-bracketing metric entropy bound of order $\varepsilon^{-1}$ for bounded subsets of the space $V_1(\mathbb{R})$ is given in van de Geer (1991) for $\mu$ an arbitrary probability measure. In the case $s > 1$, $p = 1$, $d = 1$ Theorem 1 above provides a better rate for the $L^2(\mu)$-bracketing metric entropy bound at the expense of introducing a moment condition of order $2\gamma$, $\gamma > s - 1 > 0$. Furthermore, van de Geer’s approach does not lend itself to generalization beyond $d = 1$. More importantly, note that for general $s$, $p$ (still keeping $d = 1$) the Sobolev ball $\mathcal{D}^d_{s,p}$ is not necessarily contained in the space $V_1(\mathbb{R})$ of functions of bounded variation.□

### 2.1.1 Fractional Sobolev Spaces on Subsets of $\mathbb{R}^d$

Often classes of ‘smooth’ functions defined on Borel subsets $\Omega$ of $\mathbb{R}^d$, e.g., on cubes or other convex subsets of $\mathbb{R}^d$, are of interest. Lemma 1 in Section 2.4 allows one immediately to obtain results for such function classes as long as they are defined through restricting the elements of a set of the form $\mathcal{D}^d_{s,p}$ to the set $\Omega$.

**Corollary 1** Let $\Omega$ be a (non-empty) Borel subset of $\mathbb{R}^d$, let $1 \leq p < \infty$, $\beta \in \mathbb{R}$, and $s - d/p > 0$. Let $\mathcal{D}^d_{s,p}[\Omega]$ be the set of restrictions $f|\Omega$ of elements $f \in \mathcal{D}^d_{s,p}$ to the set $\Omega$. Furthermore, let $\mu$ be some Borel measure on $\Omega$. 

5
1. Suppose the moment condition $\| (x)^{\gamma - \beta} \|_{r,\mu} < \infty$ holds for some $\gamma > 0$ and for some $1 \leq r \leq \infty$. If $\gamma > s - d/p$, we have

$$H_{\| \epsilon \|_{\gamma, \Omega}}, \mathcal{D}_{s,p}^\beta|, \| \|_{r,\mu} \|^\gamma_{r,\mu} \lesssim \epsilon^{-d/s}.$$ 

if $\gamma < s - d/p$, we have

$$H_{\| \epsilon \|_{\gamma, \Omega}}, \mathcal{D}_{s,p}^\beta|, \| \|_{r,\mu} \|^\gamma_{r,\mu} \lesssim \epsilon^{-(\gamma/d + 1/p)^{-1}}.$$ 

2. Suppose $\Omega$ is a bounded set. If $r < \infty$ and $\mu(\Omega) < \infty$, or if $r = \infty$ then

$$H_{\| \epsilon \|_{\gamma, \Omega}}, \mathcal{D}_{s,p}^\beta|, \| \|_{r,\mu} \|^\gamma_{r,\mu} \lesssim \epsilon^{-d/s}.$$ 

**Proof.** To prove the first part, identify $\mathcal{D}_{s,p}^\beta|_s \Omega$ with $\mathcal{D}_{s,p}^\beta|_s \Omega$ and view the measure $\mu$ as a Borel measure on $\mathbb{R}^d$ with $\mu(\mathbb{R}^d \setminus \Omega) = 0$. Now apply the second part of Lemma 1 in Section 2.4 below with $\Omega' = \mathbb{R}^d$ and $h = 1_{\Omega'}$. For the proof of the second part of the corollary choose some $\beta' > s - d/p$ and observe that $\mathcal{D}_{s,p}^\beta|_s \Omega \subseteq \mathcal{D}_{s,p}^{\beta'}|_s \Omega$, possibly with different radii $b$ and $b'$, since $\Omega$ is a bounded set. Set $\gamma = \beta'$ and apply the first part of the corollary to the set $\mathcal{D}_{s,p}^\beta|_s \Omega$. Note that now the moment condition

$$\| (x)^{\gamma - \beta} \|_{r,\mu} = \| 1 \|_{r,\mu} < \infty$$

is always satisfied. 

However, for $\Omega$ a domain, i.e., an open subset of $\mathbb{R}^d$, the classical definition of a Sobolev space for $s$ a nonnegative integer is not given via restricting the elements of $W^{s,p}(\mathbb{R}^d)$ to $\Omega$, but is given “intrinsically” by

$$W^{s,p}(\Omega) = \left\{ f \in L^0(\Omega) : \sum_{0 \leq |\alpha| \leq s} \| D^\alpha f \|_{p,\lambda|\Omega} < \infty \right\},$$

where $\lambda|\Omega$ denotes Lebesgue measure on $\Omega$ and where again $D^\alpha$ denotes the partial derivative in the generalized sense of distributions. For positive non-integer $s$ an extension of this definition to fractional Sobolev spaces on $\Omega$ is again available (Adams (1975, 7.35, 7.48)), the norm being denoted by $\| \cdot \|_{s,p,\lambda|\Omega}$. As in the case $\Omega = \mathbb{R}^d$, if $\Omega$ is sufficiently well-behaved in the sense of possessing the cone-property (Adams (1975, 4.3)), it is again true for $s - d/p > 0$ that the Sobolev space $W^{s,p}(\Omega)$ can be viewed as a subset of the set of bounded continuous functions on $\Omega$ in the sense as discussed in Section 2.1 (Adams (1975), Theorem 7.57).

For $s - d/p > 0$ consider the function class

$$\mathcal{D}_{s,p}^\beta(\Omega) = \left\{ f \in L^0(\Omega) : \| f \cdot (x)^{\beta} \|_{s,p,\lambda|\Omega} < b \right\}$$

with $1 \leq p < \infty$ and $\beta \in \mathbb{R}$. In light of the preceding discussion, any $f \in \mathcal{D}_{s,p}^\beta(\Omega)$ is continuous if $\Omega$ possesses the cone-property.

Suppose now $\Omega$ possesses the cone-property and that $s - d/p > 0$. Whereas the “extrinsically” defined space $W^{s,p}(\mathbb{R}^d)/\Omega$ obtained by restricting the elements of $W^{s,p}(\mathbb{R}^d)$ to $\Omega$ is obviously contained in the “intrinsically” defined space $W^{s,p}(\Omega)$, we are not aware of a result that would
establish the reverse inclusion for Ω possessing the cone-property only. Hence, we are not guaran-
teed that the set $D_{s,p}^\beta(\Omega)$ is contained in any set of the form $D_{s,p}^\beta |\Omega$, and thus we cannot take recourse to Corollary 1. However, if Ω happens to be such that the spaces $(W^{s,p}(\Omega), \|\cdot\|_{s,p,\lambda})$ and $(W^{s,p}(\mathbb{R}^d)|\Omega, \|\cdot\|_{s,p,\lambda})$ coincide (e.g., Ω has appropriate further geometrical properties), where the norm on the latter space is given by

$$\|f\|_{s,p,\lambda}|\Omega| = \inf \left\{ \|g\|_{s,p,\lambda} : g \in W^{s,p}(\mathbb{R}^d) : g|\Omega = f \right\},$$

then we obtain the result

$$H_{[\epsilon]} \left( \varepsilon, D_{s,p}^\beta(\Omega), \|\cdot\|_{r,\mu} \right) \leq H_{[\epsilon]} \left( \varepsilon, D_{s,p}^\beta |\Omega|, \|\cdot\|_{r,\mu} \right),$$

where the radii in the definition of the sets $D_{s,p}^\beta(\Omega)$ and $D_{s,p}^\beta |\Omega|$ may be different. [To see this note that both spaces are Banach-spaces and are imbedded in $L^\infty(\lambda |\Omega|)$. Hence, it follows from the closed graph theorem that the set-theoretic equality already implies that the respective norms are equivalent; cf. the first paragraph of the Appendix. Consequently, $D_{s,p}^\beta(\Omega)$ is contained in $D_{s,p}^\beta |\Omega|$ for $\beta = 0$, which immediately gives the result for general $\beta$.] The above inequality together with Corollary 1 can then immediately be used to deliver rates for the bracketing metric entropy of $D_{s,p}^\beta(\Omega)$.

Sufficient geometrical conditions on Ω (implying the cone-property) such that $W^{s,p}(\Omega)$ and $W^{s,p}(\mathbb{R}^d)|\Omega$ coincide are available in the literature. First, if Ω is an open half-space, this coincidence holds for any real $s > 0, 1 < p < \infty$ (Adams, 1975, 4.26, 7.41). Second, if Ω is a $C^\infty$-domain in the sense of Triebel (1983, 3.2.1), this coincidence holds for any real $s > 0, 1 < p < \infty$; if $s$ is non-integer then it also holds for $p = 1$ (Triebel, 1983, 3.4.2)). Similar results for domains Ω satisfying weaker geometrical properties (e.g., various stronger versions of the cone-property) can be deduced from Calderón-type extension theorems; e.g., Stein (1970, Chapter 6), Adams (1975, 4.32), Triebel (1995, 4.2.3).

**Remark 3** 'Sobolev' classes on not necessarily open subsets of $\mathbb{R}^d$

Sobolev spaces are canonically defined on open sets. In applications in the probability and statistics literature often ‘smooth’ function classes defined on a non-open Borel set $\bar{\Omega}$ (e.g., on a closed cube $\times_{i=1}^d [a_i, b_i]$, or on a ‘semi-closed’ cube $\times_{i=1}^d (a_i, b_i]$) are of interest. If such function classes are defined “extrinsically”, i.e., via restriction from $W^{s,p}(\mathbb{R}^d)$, Corollary 1 can of course be used directly. Alternatively, if one wishes to define such function classes in a more “intrinsic” fashion, one can proceed as follows (for $s > d/p$): Suppose one can find a domain Ω with the cone-property such that $\Omega \subseteq \bar{\Omega} \subseteq cl(\Omega)$, where $cl(\cdot)$ denotes the closure, and such that $W^{s,p}(\Omega)$ and $W^{s,p}(\mathbb{R}^d)|\Omega$ coincide. Then one can define $D_{s,p}^\beta(\bar{\Omega})$ simply as the set of all continuous functions $f : \bar{\Omega} \to \mathbb{R}$ the restrictions $f|\Omega$ of which belong to $D_{s,p}^\beta(\Omega)$. We then immediately have for $1 \leq r \leq \infty$ and any Borel measure $\mu$ on $\bar{\Omega}$ that

$$H_{[\epsilon]} \left( \varepsilon, D_{s,p}^\beta(\bar{\Omega}), \|\cdot\|_{r,\mu} \right) \leq H_{[\epsilon]} \left( \varepsilon, D_{s,p}^\beta(\Omega), \|\cdot\|_{r,\mu} \right).$$

The r.h.s. of the above inequality can then directly be bounded by Corollary 1. □

### 2.1.2 Hellinger-Bracketing Metric Entropy

For obtaining convergence rates and minimax risk bounds of estimators, bounds on the *Hellinger-bracketing metric entropy* of the underlying function classes are of importance. For the case of a bounded (non-empty) Borel set $\Omega \subseteq \mathbb{R}^d$ define the set $F_{s,p,\zeta} = \{ f \in D_{s,p}^\beta |\Omega| : \inf_{x \in \Omega} f(x) \geq \zeta \}$, which is non-empty if $\zeta > 0$ is sufficiently small. We obtain a bound for the Hellinger-bracketing metric entropy, i.e., for $H_{[\epsilon]}(\varepsilon, F_{s,p,\zeta}, \|\cdot\|_{2,\lambda}|\Omega)$, as follows.
Corollary 2. Let \( \tilde{F}_{s,p,\zeta} = \{ \tilde{f} \in \mathcal{F} : \tilde{f} = f^{1/2}, f \in F_{s,p,\zeta} \} \) with \( \zeta > 0, 1 \leq p < \infty \), and \( s - d/p > 0 \).
If \( \tilde{F}_{s,p,\zeta} \) is non-empty, the Hellinger-bracketing metric entropy of \( F_{s,p,\zeta} \) satisfies
\[
H_1(\varepsilon, \tilde{F}_{s,p,\zeta}, \| \cdot \|_{2,\lambda}) \lesssim \varepsilon^{-d/s}.
\]

Proof. Assume that \( \tilde{F}_{s,p,\zeta} \) is non-empty. For \( \varepsilon > 0 \) set \( \varepsilon' = \varepsilon \sqrt{\zeta/2} \). Let \( [a_i, b_i], i = 1, \ldots, N \) be brackets covering \( D_{s,p}^0(\Omega) \) with \( L^2(\lambda) \)-size less than or equal to \( \varepsilon' \). If a bracket \( [a_i, b_i] \) does not meet \( \tilde{F}_{s,p,\zeta} \), drop it from the list. Otherwise, the bracket \( [a_i, b_i] \) can be modified into a bracket \( [a'_i, b'_i] \) with \( b'_i = \max(a_i, \zeta/2) \), which also has size less than or equal to \( \varepsilon' \). The remaining modified brackets obviously cover \( \tilde{F}_{s,p,\zeta} \). Define the brackets \( [\sqrt{b_i'}, \sqrt{a_i}] \) which cover \( \tilde{F}_{s,p,\zeta} \). Furthermore, the elementary inequality \( \sqrt{b_i} - \sqrt{a_i} \leq \sqrt{2/\zeta} |a_i - b_i| \) shows that the \( L^2(\lambda) \)-size of these brackets is less than or equal to \( \varepsilon \). Thus
\[
H_1(\varepsilon, \tilde{F}_{s,p,\zeta}, \| \cdot \|_{2,\lambda}) \leq H_1(\varepsilon', D_{s,p}^0(\Omega), \| \cdot \|_{2,\lambda}) \lesssim \varepsilon^{-d/s}
\]
where the last bound follows from the second part of Corollary 1. \( \blacksquare \)

Clearly, a result like Corollary 2 for \( \zeta = 0 \) cannot be obtained by the same method of proof for obvious reasons. For \( \Omega \) with infinite Lebesgue measure the set \( F_{s,p,\zeta} \) is empty for any \( \zeta > 0 \), and the case \( \zeta = 0 \) can again not be treated for the same obvious reasons.

2.2 Besov Spaces

In recent years, classes of functions that belong to a Besov space have attracted interest, for example, in the wavelet literature (e.g., Donoho and Johnstone (1998), Birgé and Massart (2000)) the interest resulting from the usefulness (and in fact origin) of Besov spaces in approximation theory. In this section we provide bounds for the bracketing metric entropy of several subsets of Besov spaces. Since the results are parallel to the results in Section 2.1, we summarize them in one theorem and shall omit discussion that is similar.

For \( 0 < s < \infty, 1 \leq p \leq \infty \), and \( 1 \leq q \leq \infty \) let \( B_{s,p}^q(\mathbb{R}^d) \) denote the Besov spaces over \( \mathbb{R}^d \) (semi)norned by \( \| \cdot \|_{B_{s,p,q,\lambda}} \); for the definition see, e.g., Triebel (1983, 2.3.1/5); for a more commonly used equivalent definition see also Triebel (1983, 2.2.2/9). Note that \( B_{s,\infty}^\infty \) coincides with the Hölder-Zygmund space (see Triebel (1983), Theorem 2.5.7). Besov spaces have many characteristic properties in common with Sobolev spaces. Most importantly, if \( s - d/p > 0 \), every \( f \in B_{s,p}^q(\mathbb{R}^d) \) can be viewed as a bounded continuous function that has classical partial derivatives of order at least \( s - d/p \), a convention we shall always adopt (and renders the seminorm a norm); furthermore, every such \( f \) vanishes at infinity (unless \( p = \infty \)). See Proposition 3 in the Appendix for more detail, where we also show that any function in \( B_{s,p}^p(\mathbb{R}) \) is of bounded \( p \)-variation in case \( s > 1/p \) (and \( p < \infty \)). Also note that \( B_{s,p}^p(\mathbb{R}^d) = W^{s,p}(\mathbb{R}^d) \) for \( s \) non-integer (and \( p < \infty \)).

For \( s - d/p > 0 \) consider the function class
\[
B_{s,p,q}^\beta = \left\{ f \in L^0(\mathbb{R}^d) : \left\| f \cdot (x)^\beta \right\|_{B_{s,p,q,\lambda}} \leq b \right\} \quad (0 \leq b < \infty)
\]
with \( 1 \leq p \leq \infty, 1 \leq q \leq \infty \), and \( \beta \in \mathbb{R} \). In view of the convention discussed in the preceding paragraph each \( f \cdot (x)^\beta \in B_{s,p}^p(\mathbb{R}^d) \) is continuous, and hence so is any \( f \in B_{s,p,q}^\beta \). Note that \( B_{s,p,q}^\beta \subseteq B_{s,p}^p(\mathbb{R}^d) \) for \( \beta \geq 0 \), but not for \( \beta < 0 \). Again, there exists a finite positive constant \( K \) such that \( K \langle x \rangle^{-\beta} \) is an envelope for \( B_{s,p,q}^\beta \); see Proposition 3 in the Appendix.
Theorem 2  Let $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $\beta \in \mathbb{R}$, and $s - d/p > 0$. Then the following statements hold:

1. Let $\beta > 0$. Then $\mathcal{B}_{s,p,q}^\beta$ is a relatively compact subset of the normed space $(UC_0(\mathbb{R}^d), \|\cdot\|_\infty)$. Furthermore, for $\beta > s - d/p$ we have
   \[ H(\varepsilon, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_\infty) \lesssim \varepsilon^{-d/s}, \]
   and for $\beta < s - d/p$ we have
   \[ H(\varepsilon, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_\infty) \lesssim \varepsilon^{-(\beta/d + 1/p)^{-1}}. \]

2. Suppose that for some Borel measure $\mu$ on $\mathbb{R}^d$ and some $1 \leq r \leq \infty$ the moment condition $\| (x)^{\gamma - \beta} \|_{r,\mu} < \infty$ holds for some $\gamma > 0$. If $\gamma > s - d/p$, we have
   \[ H[1](\varepsilon, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{r,\mu}) \lesssim \varepsilon^{-d/s}, \]
   if $\gamma < s - d/p$, we have
   \[ H[1](\varepsilon, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{r,\mu}) \lesssim \varepsilon^{-(\gamma/d + 1/p)^{-1}}. \]

Proof. By (7)-(8) in the Appendix, the set $\mathcal{B}_{s,p,q}^\beta$ is contained in a bounded subset of the weighted Besov space $B_{qs}^\beta(\mathbb{R}^d, (x)^{\beta})$. To prove the first claim in Part 1 of the theorem, apply Proposition 2 in the Appendix with $\gamma = \beta > 0$. This together with Triebel (1983, 2.5.7/2) gives the following chain of imbeddings
\[ B_{qs}^\beta(\mathbb{R}^d, (x)^{\beta}) \hookrightarrow B_{2\infty}^0(\mathbb{R}^d) \hookrightarrow (UC(\mathbb{R}^d), \|\cdot\|_\infty). \]
Since the first imbedding is compact, the composite imbedding is also compact. Hence, $\mathcal{B}_{s,p,q}^\beta$ is a relatively compact subset of $UC(\mathbb{R}^d)$; by the second part of Proposition 3 it even belongs to the closed subspace $UC_0(\mathbb{R}^d)$. This proves this claim.

Proposition 2 in the Appendix gives the entropy number $e(k, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{B(\beta - \gamma), 0, \infty, 1, \lambda})$ of $\mathcal{B}_{s,p,q}^\beta$ in the space $B_{2\infty}^\beta(\mathbb{R}^d, (x)^{\beta - \gamma})$ which we abbreviate here by $e_k$. Without loss of generality we have here set the radius $b$ of $\mathcal{B}_{s,p,q}^\beta$ equal to one. Observe first that the sequence $e_k$ is nonincreasing, converges to zero, and satisfies $0 < e_k < \infty$. Therefore, for every $k \in \mathbb{N}$ there exists a unique non-negative integer $l(k)$ such that $e_k = e_{k+l(k)}$ and $e_k > e_{k+1(l(k)+1)}$. Hold from the definition of the covering numbers it follows for every $\eta > e_{k+l(k)+1}$ that
\[ \log_2 N(\eta, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{B(\beta - \gamma), 0, \infty, 1, \lambda}) \leq k + l(k) \]

and thus
\[ \log_2 N(e_k, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{B(\beta - \gamma), 0, \infty, 1, \lambda}) \leq k + l(k). \]

Fix $0 < \varepsilon \leq e_1$. Then there exists a unique index $k = k(\varepsilon) \geq 2$ such that $e_k < \varepsilon \leq e_{k-1}$. Consequently,
\[ \log_2 N(\varepsilon, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{B(\beta - \gamma), 0, \infty, 1, \lambda}) \leq \log_2 N(e_k, \mathcal{B}_{s,p,q}^\beta, \|\cdot\|_{B(\beta - \gamma), 0, \infty, 1, \lambda}) \leq k + l(k). \]
By Proposition 2 in the Appendix, there exists a positive and finite constant \( c \) such that 
\( e_k = e_{k+l(k)} \leq c(k+l(k))^{-\alpha} \) holds, where either \( \alpha = s/d \) or \( \alpha = \gamma/d + 1/p \). Consequently,

\[
\log_2 N(\varepsilon, B^3_{s,p,q}, \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \leq k + l(k) \leq c^{1/\alpha} e_k^{-1/\alpha}.
\]

Similarly, there exists a positive finite constant \( c' \) such that \( c'k^{-\alpha} \leq e_k \). Hence,

\[
\log_2 N(\varepsilon, B^3_{s,p,q}, \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \leq \frac{(c/c')^{1/\alpha} k}{(c/c')^{1/\alpha}[k - 1 + 1]} \leq \frac{(c/c')^{1/\alpha}[e^{1/\alpha} e_k^{-1/\alpha} + 1]}{(c/c')^{1/\alpha}[e^{1/\alpha} e_k^{-1/\alpha} + 1]}
\]

where in the final two steps we have used the relations \( e_{k-1} \leq c(k-1)^{-\alpha} \) and \( e_{k-1} \geq \varepsilon \). This establishes

\[
H(\varepsilon, B^3_{s,p,q}, \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \leq C_1 \varepsilon^{-1/\alpha}
\]

for \( 0 < \varepsilon \leq e_1 \) and a suitable real number \( C_1 \). In fact (1) holds for every \( \varepsilon > 0 \), since the metric entropy is zero for \( \varepsilon > e_1 \) as \( e_1 \) is the operator norm of the imbedding.

In view of (8) in the Appendix and the first part of Lemma 1 the bound (1) leads to

\[
H(\varepsilon, B^3_{s,p,q}, \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \leq C_2 \varepsilon^{-1/\alpha}
\]

for every \( \varepsilon > 0 \) and a suitable real number \( C_2 \). The imbedding

\[
(B^0_{\infty1}(\mathbb{R}^d), \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \hookrightarrow (UC(\mathbb{R}^d), \| \cdot \|_\infty)
\]

(Triebel (1983, 2.5.7/2)) implies that the identity operator is bounded. It follows that

\[
\| \langle \cdot \rangle \langle x \rangle^{\beta - \gamma} \|_\infty \leq \| \langle \cdot \rangle \langle x \rangle^{\beta - \gamma} \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}
\]

holds (on the space \( B^0_{\infty1}(\mathbb{R}^d, \langle x \rangle^{\beta - \gamma}) \)). Hence, the first part of Lemma 1 gives

\[
H(\varepsilon, B^3_{s,p,q}, \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \leq C_3 \varepsilon^{-1/\alpha}
\]

for every \( \varepsilon > 0 \) and a suitable real number \( C_3 \). This immediately completes the proof of the first part of the theorem upon setting \( \gamma = \beta \) and using the definition of \( \alpha \).

To the prove the second part, let \( B_i, i = 1, ..., N(\varepsilon, B^3_{s,p,q}, \| \cdot \|_{B(\varepsilon^{-1}, 0, \infty, 1, \lambda)}) \), denote closed \( \| \langle \cdot \rangle \langle x \rangle^{\beta - \gamma} \|_\infty \)-balls of radius \( \varepsilon \) covering \( B^3_{s,p,q} \). Observe that each such ball \( B_i \) (with center \( f_i \)) contains all (continuous) functions \( f \) for which

\[
\sup_{x \in \mathbb{R}^d} |f(x) - f_i(x)| \langle x \rangle^{\beta - \gamma} \leq \varepsilon.
\]

These balls define brackets

\[
[f_{i} - \varepsilon \langle x \rangle^{\gamma - \beta}, f_{i} + \varepsilon \langle x \rangle^{\gamma - \beta}]
\]

which obviously cover \( B^3_{s,p,q} \). The \( L^r(\mu) \)-size of such a bracket is given by

\[
\| 2\varepsilon \langle x \rangle^{\gamma - \beta} \|_{r,\mu}
\]
which is finite by assumption. Thus, using (3),
\[ H_1 \left( \varepsilon \|2 (x)^{1 - \beta} \|_{r, \mu}, B_{s, p, q}^\beta \| \|_{r, \mu} \right) \leq H \left( \varepsilon, B_{s, p, q}^\beta, \| \| \right) \leq \varepsilon^{-1/\alpha} \]  
for every \( \varepsilon > 0 \) (provided \( \mu(\mathbb{R}^d) > 0 \)). Inserting the definition of \( \alpha \) and rescaling by \( \|2 (x)^{1 - \beta} \|_{r, \mu} \) delivers the desired result. (If \( \mu(\mathbb{R}^d) = 0 \), the result is trivial since by Proposition 3 in the Appendix the single bracket \( [-K (x)^{1 - \beta}, K (x)^{1 - \beta}] \) covers \( B_{s, p, q}^\beta \) and has size zero.)

The proof in fact establishes a more general bound for the weighted sup-norm metric entropy, cf. (3), for arbitrary \( \beta \in \mathbb{R} \), which delivers the first part of the theorem as a special case upon setting \( \gamma = \beta > 0 \), and provides \( L^r(\mu) \)-bracketing metric entropy bounds for general \( \beta \).

Similarly as discussed at the beginning of Section 2.1.1, the second part of Lemma 1 in Section 2.4 can be used to obtain bracketing metric entropy bounds for Besov-type classes of functions defined on Borel subsets \( \Omega \) of \( \mathbb{R}^d \) as long as they are obtained by restricting the elements of \( B_{s, p, q}^\beta \) to \( \Omega \). In analogy to the results in Sections 2.1.1 and 2.1.2, we give the following proposition:

**Corollary 3** Let \( \Omega \) be a (non-empty) Borel subset of \( \mathbb{R}^d \), let \( 1 \leq p \leq \infty, \ 1 \leq q \leq \infty, \beta \in \mathbb{R} \), and \( s - d/p > 0 \). Let \( B_{s, p, q}^\beta \Omega \) be the set of restrictions \( f | \Omega \) of elements \( f \in B_{s, p, q}^\beta \) to the set \( \Omega \). Furthermore, let \( \mu \) be some Borel measure on \( \Omega \).

1. Suppose the moment condition \( \| (x)^{1 - \beta} \|_{r, \mu} < \infty \) holds for some \( \gamma > 0 \) and for some \( 1 \leq r \leq \infty \). If \( \gamma > s - d/p \), we have
\[ H_1 \left( \varepsilon, B_{s, p, q}^\beta | \Omega, \| \|_{r, \mu} \right) \leq \varepsilon^{-s/d}; \]
if \( \gamma < s - d/p \), we have
\[ H_1 \left( \varepsilon, B_{s, p, q}^\beta | \Omega, \| \|_{r, \mu} \right) \leq \varepsilon^{-\gamma/d + 1/p}. \]

2. Suppose \( \Omega \) is a bounded set. If \( r < \infty \) and \( \mu(\Omega) < \infty \), or if \( r = \infty \) then
\[ H_1 \left( \varepsilon, B_{s, p, q}^\beta | \Omega, \| \|_{r, \mu} \right) \leq \varepsilon^{-d/s}. \]

3. Suppose \( \Omega \) is a bounded set. Let \( F_{s, p, q, \zeta} = \{ f \in B_{s, p, q}^0 | \Omega \ : \ \inf_{x \in \Omega} f(x) \geq \zeta \} \) with \( \zeta > 0 \). Define
\[ \hat{F}_{s, p, q, \zeta} = \{ \hat{f} \ : \ \hat{f} = f^{1/2}, \ f \in F_{s, p, q, \zeta} \}. \]
If \( \hat{F}_{s, p, q, \zeta} \) is non-empty, the Hellinger-bracketing metric entropy of \( F_{s, p, q, \zeta} \) satisfies
\[ H_1 \left( \varepsilon, \hat{F}_{s, p, q, \zeta}, \| \|_{2, \lambda | \Omega} \right) \leq \varepsilon^{-d/s}. \]

**Proof.** The proof is completely analogous to the proofs of Corollaries 1 and 2. 

The above results are given for “extrinsically” defined Besov spaces. If \( \Omega \) is an open set, again an “intrinsic” definition of a Besov space on \( \Omega \) is possible (Triebel (1992), 5.2). Again, under appropriate geometrical conditions, both definitions coincide (more precisely, the norms are equivalent); see, e.g., Triebel (1983, 3.4.2) and Triebel (1995, 4.2.3). Hence the above results are applicable to subsets of “intrinsically” defined Besov spaces on domains \( \Omega \). We do not repeat the details as they are completely analogous to Section 2.1.1. Also, a remark similar to Remark 3 applies to Besov spaces.
2.2.1 Hölder, Triebel, and Bessel Potential Spaces

We define for integer \( s > 0 \) the space \( C^s(\mathbb{R}^d) \) as the set of all real-valued functions \( f \) on \( \mathbb{R}^d \) for which

\[
\|f\|_{s,\infty} = \sum_{0 \leq |\alpha| \leq s} \|D^\alpha f\|_{\infty}
\]

is finite and the (classical) partial derivatives \( D^\alpha f \) are uniformly continuous for \( |\alpha| = s \). For real, non-integer \( s > 0 \) we define \( C^s(\mathbb{R}^d) \) as the set of all real-valued functions \( f \) on \( \mathbb{R}^d \) for which

\[
\|f\|_{s,\infty} = \sum_{0 \leq |\alpha| \leq [s]} \|D^\alpha f\|_{\infty} + \sum_{|\alpha| = [s]+1} \sup_{x \neq y} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x - y|^{s-[s]}}
\]

is finite. Here \([s]\) denotes the integer part of \( s \) and the derivatives are again to be understood in the classical sense. (For non-integer \( s \) we refer to \( C^s(\mathbb{R}^d) \) as a Hölder space, but we prefer to avoid this terminology in case \( s \) is integer, because there seems to be no universally accepted definition in this case.) For \( \beta \in \mathbb{R} \) define the function class

\[
C^\beta_s = \left\{ f : \|f \cdot \langle x \rangle^\beta\|_{s,\infty} \leq 1 \right\} \quad (0 \leq b < \infty).
\]

Observe that \( C^\beta_s \) is contained in a set of the form \( B^\beta_{s,\infty,\infty} \), see Lemma 3 in the Appendix. For \( \beta > 0 \) it then follows from the first part of Theorem 2 that \( C^\beta_s \) is a relatively compact subset of the normed space \( (UC_0(\mathbb{R}^d), \|\cdot\|_{s,\infty}) \); furthermore, for \( \beta \neq s \), Theorem 2 gives the sup-norm metric entropy bound

\[
H(\varepsilon, C^\beta_s, \|\cdot\|_{s,\infty}) \lesssim \varepsilon^{-d/\min(s,\beta)}.
\]

We note that \( b \langle x \rangle^{-\beta} \) is an envelope of \( C^\beta_s \) which is sharp in the sense that \( b \langle x \rangle^{-\beta} \in C^\beta_s \) holds. If \( \|\langle x \rangle^{-\beta}\|_{r,\mu} < \infty \) holds for some \( \gamma > 0 \), some \( 1 \leq r \leq \infty \), and some Borel measure \( \mu \) on \( \mathbb{R}^d \), we have from the second part of Theorem 2 that for \( \gamma \neq s \)

\[
H(\varepsilon, C^\beta_s, \|\cdot\|_{r,\mu}) \lesssim \varepsilon^{-d/\min(s,\gamma)}.
\]

Of course, also Proposition 2 holds with \( B^\beta_{s,\infty,\infty} \mid \Omega \) replaced by the set \( C^\beta_s \mid \Omega \) consisting of the restrictions of the elements of \( C^\beta_s \) to \( \Omega \). A discussion similar to the one following Proposition 2 also applies here.

All results for Besov spaces carry over to (weighted) Triebel spaces \( F^s_{pq} \) (for a definition see Triebel (1983, 2.3.1/6)). This is so since any Triebel space \( F^s_{pq} \) can be imbedded into a Besov space \( B^s_{pq} \) for \( v \geq p \) (see Edmunds and Triebel (1996, 2.3.3/5)) which implies that any bounded set in \( F^s_{pq} \) is also a bounded set in \( B^s_{pq} \). Since the index \( v \) does not enter the above results for Besov spaces, the constraint \( v \geq p \) has no consequences for the entropy results. We also note that Triebel spaces contain the Bessel Potential spaces \( H^{s,p} \) as special cases, i.e., \( F^s_{p2} = H^{s,p} \) for \( 1 < p < \infty \); see Triebel (1983, 2.5.6/2).

2.3 Uniform (Bracketing) Metric Entropy Bounds

The results in Section 2 imply the following uniform (bracketing) metric entropy bounds for (subsets of) \( B^s_{p,q} \). Since \( D^s_{p,q} \) and \( C^s \) are contained in respective sets of the form \( B^s_{p,q} \) (for suitable radii) the following corollary also applies to these sets. Analogous results hold for balls in Besov spaces defined over Borel sets \( \Omega \) other than \( \mathbb{R}^d \).
Corollary 4 Let $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $\beta \in \mathbb{R}$, and $s - d/p > 0$. Let $\mathcal{M}$ be a (non-empty) family of Borel measures $\mu$ on $\mathbb{R}^d$ such that the condition
\[ \sup_{\mu \in \mathcal{M}} \left\| \langle x \rangle^{\gamma - \beta} \right\|_{r, \mu} < \infty \]
holds for some $\gamma > 0$ and for some $1 \leq r \leq \infty$. Then
\[ \sup_{\mu \in \mathcal{M}} H(\varepsilon, B^d_{s,p,q}, \|\cdot\|_{r, \mu}) \leq \sup_{\mu \in \mathcal{M}} H_{\| \cdot \|_{\infty, \mu}}(2\varepsilon, B^d_{s,p,q}, \|\cdot\|_{r, \mu}) \]
for $\gamma > s - d/p$ and
\[ \varepsilon^{-d/s} \quad \text{for } \gamma < s - d/p \]
holds.

Proof. The first inequality is obvious. Inspection of (4) in the proof of the second part of Theorem 2 immediately gives the second inequality. \( \blacksquare \)

In particular, if $\beta > 0$, we may set $\gamma = \beta$ in which case the uniform moment condition is satisfied, e.g., for $\mathcal{M}$ the set of all probability measures. This is useful as limit theorems for empirical processes are often based on $L^r(\mathcal{P})$-metric entropy bounds uniformly in all probability measures $\mathcal{P}$ (or in all probability measures supported in finitely many points).

### 2.4 Preservation of (Bracketing) Metric Entropy Rates Under Transformations

Obviously, the metric entropy of a set is invariant under homeomorphisms. The following elementary results are of a similar nature and can be useful for deriving bracketing metric or metric entropy bounds for a variety of function classes that can be obtained by certain transformations of function classes with known entropy bounds.

Lemma 1. Let $(X, \| \cdot \|_X)$ and $(Y, \| \cdot \|_Y)$ be normed spaces and let $\mathcal{F}$ be a subset of $X$ with finite metric entropy $H(\varepsilon, \mathcal{F}, \| \cdot \|_X)$ for all $\varepsilon$, $0 < \varepsilon < \infty$. Let the map $A : \mathcal{F} \to (Y, \| \cdot \|_Y)$ satisfy $\|A(x_1) - A(x_2)\|_Y \leq C \|x_1 - x_2\|_X$ for all $x_1, x_2 \in \mathcal{F}$ and for some $\sigma > 0$, $0 < C < \infty$. We then have for the image $A(\mathcal{F})$ that
\[ H(\varepsilon, A(\mathcal{F}), \| \cdot \|_Y) \leq H \left( 2^{-1} C^{-1/\sigma} \varepsilon^{1/\sigma}, \mathcal{F}, \| \cdot \|_X \right) \]
for every $\varepsilon$, $0 < \varepsilon < \infty$. (If $C = 0$, then $H(\varepsilon, A(\mathcal{F}), \| \cdot \|_Y) = 0$ for every $\varepsilon$, $0 < \varepsilon < \infty$.)

2. Let $\mu$ be a Borel-measure on $\Omega$, a (non-empty) Borel subset of $\mathbb{R}^d$, and let $\mathcal{F}$ be a subset of $L^r(\Omega)$ with finite bracketing metric entropy $H_{\| \cdot \|_{r, \mu}}(\varepsilon, \mathcal{F}, \| \cdot \|_{r, \mu})$ for some $1 \leq r \leq \infty$ and all $\varepsilon$, $0 < \varepsilon < \infty$. If $h$ is a nonnegative Borel-measurable function on $\Omega$ satisfying $0 < \|h\|_{\infty, \mu} < \infty$, then for $\mathcal{F}h = \{ fh : f \in \mathcal{F} \}$ we have
\[ H_{\| \cdot \|_{r, \mu}}(\varepsilon, \mathcal{F}h, \| \cdot \|_{r, \mu}) \leq H_{\| \cdot \|_{\infty, \mu}}(\varepsilon, \mathcal{F}, \| \cdot \|_{r, \mu}) \]
for every $\varepsilon$, $0 < \varepsilon < \infty$. (If $\|h\|_{\infty, \mu} = 0$, then $H_{\| \cdot \|_{r, \mu}}(\varepsilon, \mathcal{F}h, \| \cdot \|_{r, \mu}) = 0$ for every $\varepsilon$, $0 < \varepsilon < \infty$.)
Proof. To prove the first part, let $B_i, i = 1, ..., N \left(2^{-1}C^{-1/\sigma} \varepsilon^{1/\sigma} F, \| \cdot \|_X \right)$ be closed balls of radius $2^{-1} C^{-1/\sigma} \varepsilon^{1/\sigma}$ covering $F$. For each $i$ choose $x_i$ from $B_i \cap F$, which obviously is non-empty. Then

$$B_i^* = \left\{ y \in Y : \| y - A(x_i) \|_Y \leq \sup_{x \in B_i \cap F} \| A(x) - A(x_i) \|_Y \right\}$$

is a closed ball in $Y$ containing $A(B_i \cap F)$, hence the union of all balls $B_i^*$ covers $A(F)$. The radius of $B_i^*$ is less than or equal to $\varepsilon$ since

$$\sup_{x \in B_i \cap F} \| A(x) - A(x_i) \|_Y \leq \sup_{x \in B_i \cap F} C \| x - x_i \|_X^\sigma \leq \varepsilon.$$

Thus $H(\varepsilon, A(F), \| \cdot \|_Y) \leq H(\varepsilon^{1/\sigma}, F, \| \cdot \|_X)$. The claim in parentheses is obvious.

To prove the second part, let $[l_i, u_i], i = 1, ..., N$, $N_1 \left(\| h \|_{\infty, \mu}^{-1}, \varepsilon, F, \| \cdot \|_{r, \mu}\right)$ be brackets of $L^r(\mu)$-size less than or equal to $\| h \|_{\infty, \mu}^{-1} \varepsilon > 0$ covering $F$. Since $h$ is nonnegative, the brackets $[l_i h, u_i h]$ cover the set $F h$. The $L^r(\mu)$-size of these brackets can be bounded by

$$\| u_i h - l_i h \|_{r, \mu} \leq \| h \|_{\infty, \mu} \| u_i - l_i \|_{r, \mu} \leq \varepsilon.$$

Hence $N_1 \left(\varepsilon, F h, \| \cdot \|_{r, \mu}\right) \leq N_1 \left(\| h \|_{\infty, \mu}^{-1}, \varepsilon, F, \| \cdot \|_{r, \mu}\right)$. The claim in parentheses follows upon observing that $[\min_i l_i, \max_i u_i]$ is a bracket for $F$, and thus $[\min_i l_i h, \max_i u_i h]$ is a bracket for $F h$ which has $L^r(\mu)$-size zero. ■

3 Applications to Empirical Process Theory

We discuss some applications of the above results to empirical process theory in this section. Here the Borel measure $\mu$ will always be a probability measure $\mathbb{P}$ on a Borel set $\Omega \subseteq \mathbb{R}^d$. Given a collection $\mathcal{F}$ of Borel-measurable functions $f : \Omega \to \mathbb{R}$, the empirical measure $\mathbb{P}_n = 1/n \sum_{i=1}^n \delta_{X_i}$ of $n$ independent random variables $X_1, ..., X_n$ distributed according to $\mathbb{P}$ induces a map from $\mathcal{F} \to \mathbb{R}$ given by $f \mapsto 1/n \sum_{i=1}^n f(X_i) = \mathbb{P}_n f$. With $\mathbb{P} f := \int f(x) \mathbb{P}(x)$, the centered and scaled version of this map is the $\mathcal{F}$–indexed empirical process $\mathcal{G}_n$ given by

$$f \mapsto \mathcal{G}_n f = \sqrt{n} (\mathbb{P}_n - \mathbb{P}) f = \frac{1}{\sqrt{n}} \sum_{i=1}^n (f(X_i) - \mathbb{P} f).$$

3.1 Glivenko-Cantelli Classes

A function class $\mathcal{F} \subseteq L^1(\mathbb{P})$ is a (weak) Glivenko-Cantelli class if

$$\sup_{f \in \mathcal{F}} |(\mathbb{P}_n - \mathbb{P}) f| \xrightarrow{P} 0 \text{ as } n \to \infty, \quad \text{(5)}$$

where the convergence is to be understood w.r.t. outer measure if the supremum is not measurable. By Theorem 7.1.5 in Dudley (1999), $\mathcal{F}$ is Glivenko-Cantelli if $N_1(\varepsilon, \mathcal{F}, \| \cdot \|_{1, \mathbb{P}}) < \infty$ for all $\varepsilon > 0$. The results in Section 2 show that this condition is satisfied for all function classes for which a bound for the bracketing metric entropy was given in this paper. Furthermore, it follows from Section 2.3 that for $\beta > 0$ the sets $D^\beta \mathbb{P}, B^\beta \mathbb{P}, C^\beta$ are uniform (and hence universal) Glivenko-Cantelli classes by using Theorem 6 in Dudley, Giné and Zinn (1991).
While the entropy rates derived in this paper are much stronger than what is needed to obtain Glivenko-Cantelli results, our bracketing-metric entropy bounds also allow one to derive the rate of convergence in (5), by using Corollary 4 in conjunction with either Pollard (1984, Theorem 37) or Kolchinskii (1981). Alternatively, one may use our $L^1(\mathcal{P})$-bracketing metric entropy results from Section 2 together with Yukich (1986, 1987). We note that Yukich’s results are more flexible than the ones in Pollard and Kolchinskii, as the former results allow for non-i.i.d. random variables as well as for unbounded function classes as long as they have a suitable envelope.

3.2 Donsker and Uniform Donsker Classes

A function class $\mathcal{F} \subseteq L^2(\mathcal{P})$ is said to be $\mathcal{P}$-Donsker if it is pregaussian and if $\mathcal{G}_n \rightarrow \mathcal{G}_\mathcal{P}$ in law as $n \rightarrow \infty$ in $l^\infty(\mathcal{F})$ (where $l^\infty(\mathcal{F})$ denotes the set of all bounded real functions $H$ on $\mathcal{F}$ normed by $\|H\|_\infty = \sup_{f \in \mathcal{F}} |H(f)|$) and where $\mathcal{G}_\mathcal{P}$ is a zero-mean Gaussian process over $l^\infty(\mathcal{F})$ with covariance function $\mathbb{E}((f - \mathbb{E}f)(g - \mathbb{E}g))$. Given a (non-empty) family $\mathcal{P}$ of probability measures, a function class $\mathcal{F}$ is said to be $\mathcal{P}$-universal Donsker if $\mathcal{F}$ is $\mathcal{P}$-Donsker for all $\mathcal{P} \in \mathcal{P}$. The class $\mathcal{F}$ is said to be a $\mathcal{P}$-uniform Donsker class if the convergence of $\mathcal{G}_n$ to $\mathcal{G}_\mathcal{P}$ in $l^1(\mathcal{F})$ is uniform in $\mathcal{P}$ in a sense made precise in Giné and Zinn (1991); see also Dudley (1992), Sheehy and Wellner (1992), and Talagrand (2003). The $\mathcal{P}$-uniform Donsker property obviously implies the $\mathcal{P}$-universal Donsker property. We shall omit the prefix $\mathcal{P}$ in both of these concepts if $\mathcal{P}$ is the set of all probability measures (on the underlying set $\Omega$). The uniform Donsker property is in particular useful for establishing bootstrap central limit theorems, see, e.g., Giné and Zinn (1990) and Giné (1997).

The following corollaries provide a comprehensive description of uniform Donsker properties of (weighted) Besov balls $B_{s,p,q}$. Since $D_{s,p}$ and $C^s_d$ are contained in respective sets of the form $B_{s,p,q}$ (for suitable radii), the following results also apply to these sets. Analogous results hold for (weighted) balls in Besov spaces defined over Borel sets $\Omega$ other than $\mathbb{R}^d$. It should be noted that the subsequent corollary immediately delivers $\mathcal{P}$-Donsker classes for a fixed $\mathcal{P}$ upon setting $\mathcal{P} = \{\mathcal{P}\}$.

**Corollary 5** Let $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $\beta \in \mathbb{R}$, and $s - d/p > 0$. Given a family $\mathcal{P}$ of probability measures on $\mathbb{R}^d$ and a $\gamma > 0$ such that

$$\sup_{\mathcal{P} \in \mathcal{P}} \left\| \langle x \rangle^{\gamma - \beta} \right\|_{L^p(\mathcal{P})} < \infty$$

(6) holds, we have the following results:

1. If $\gamma > s - d/p$ and $s > d/2$, then $B_{s,p,q}$ is a $\mathcal{P}$-uniform Donsker class.

2. If $\gamma < s - d/p$ and $\gamma/d + 1/p > 1/2$, then $B_{s,p,q}$ is a $\mathcal{P}$-uniform Donsker class.

**Proof.** Corollary 4 implies that there exists a finite constant $C$ such that

$$\sup_{\mathcal{P} \in \mathcal{P}} \left\| H_{\{1\}}(\epsilon, B_{s,p,q}) \right\|_{L^p(\mathcal{P})} \leq C\varepsilon^{-1/\alpha}$$

where $\alpha$ has been defined in the proof of Theorem 2. Hence, the condition on the convergence of the bracketing integral in Theorem 2.8.4 of van der Vaart and Wellner (1996) is satisfied in the first as well as in the second part of the corollary. Observe that (6) implies for the envelope $K \langle x \rangle^{-\beta}$ (cf. Proposition 3 in the Appendix)

$$\lim_{M \rightarrow \infty} \sup_{\mathcal{P} \in \mathcal{P}} \left\| K \langle x \rangle^{-\beta} \cdot 1_{[K \langle x \rangle^{-\beta} > M]} \right\|_{L^p(\mathcal{P})} = 0,$$
which verifies the envelope condition in Theorem 2.8.4 of van der Vaart and Wellner (1996). ■

We note that the moment condition (6) is trivially satisfied irrespective of the particular choice of $\mathcal{P}$ in case $\gamma \leq \beta$. Therefore, if $\beta > s - d/p > 0$ and $s > d/2$, $\mathcal{B}^2_{s,p,q}$ is a uniform Donsker class; similarly, if $0 < \beta < s - d/p$ and $\beta/d + 1/p > 1/2$, then $\mathcal{B}^2_{s,p,q}$ is again a uniform Donsker class. This follows by setting $\gamma = \beta$ in the above corollary if $\beta \neq s - d/p$ and by setting $\gamma = \beta - \delta$ for some sufficiently small $\delta > 0$ otherwise.

For example, it follows from the above corollary (with $p = q = \infty$) that a Hölder class $\mathcal{C}^\beta_s$ is $\mathcal{P}$-Donsker, if $\|\langle x \rangle^{-\beta}\|_{2,p} < \infty$ holds for some $\gamma > s > d/2$. Van der Vaart (1994) considers related Hölder-type classes on $\mathbb{R}^d$ and obtains bounds for their $L^r(\mathcal{P})$-bracketing metric entropy for $r < \infty$. His function classes are such that the restrictions of each element to convex bounded subsets $\mathcal{I}$ partitioning $\mathbb{R}^d$ belong to a ball of radius $M_d$ in the Hölder space $C^\alpha(\mathcal{I})$. His envelope conditions for these function classes to be $\mathcal{P}$-Donsker are similar to our result for $\mathcal{C}^\beta_s$. For example, in case $d = 1$ and a constant envelope (i.e. $\beta = 0$), van der Vaart (1994, Example 2) requires the same moment condition, i.e., $\|\langle x \rangle^{\gamma + 1}\|_{2,p} < \infty$ for some $\delta > 0$ and $s > 1/2$.

The limiting case $\gamma = 0$ in Corollary 5 would entail that $\mathcal{B}^2_{s,p,q}$ is a $\mathcal{P}$-uniform Donsker class for $p < 2$ if the moment condition with $\gamma = 0$ is satisfied. In particular, for $\beta = 0$, this would give the uniform Donsker property of $\mathcal{B}^0_{s,p,q}$ for $p < 2$. The bracketing methods used in the present paper, however, prohibit the case $\gamma = 0$ in the above corollary. Nevertheless, at least for $\beta = 0$ and $d = 1$, the result can be established by using the last part of Proposition 3 in the Appendix together with a result in Dudley (1992).

**Corollary 6** Let $1 \leq p < 2$, $1 \leq q \leq \infty$, and $s - 1/p > 0$. Then $\mathcal{B}^0_{s,p,q}$ is a uniform Donsker class.

**Proof.** Note that $\mathcal{B}^0_{s,p,q}$ is contained in a bounded subset of $\mathcal{V}_p(\mathbb{R})$ in view of Proposition 3 in the Appendix. The result now follows from Theorem 2.2 in Dudley (1992). ■

If $p = 2$ and $s - d/2 > 0$, finiteness of $\|\langle x \rangle^{\gamma}\|_{2,p}$ for an arbitrarily small $\gamma > 0$ is required in Corollary 5 for an unweighted Sobolev ball $\mathcal{D}^0_{s,2}$ to be $\mathcal{P}$-Donsker. Marcus (1985), building on Giné (1975), used the fact that $W^{s,2}(\mathbb{R}^d)$ is a Hilbert-space to prove that $\mathcal{D}^0_{s,2}$ is even universally Donsker; however, in contrast to the present paper, he does not provide (bracketing) metric entropy rates for $\mathcal{D}^0_{s,2}$. The proof in Marcus (1985) is based on the Hilbert-Schmidt imbedding

$$W^{s,2}(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d, \mathcal{P})$$

which is compact for any fixed $\mathcal{P}$ as long as $s - d/2 > 0$. In contrast, our results are based on imbedding results in Haroske and Triebel (1994) that apply to more general function spaces on the one hand and are designed to obtain entropy rates on the other hand. Specialized to this particular case, these results establish compactness (and entropy rates) of the above imbedding only for probability measures which have a finite moment of order $2\gamma$ for some $\gamma > 0$. Cf. also Remark 1.] Whether $\mathcal{D}^0_{s,2}$ is a uniform Donsker class for $s - d/2 > 0$ is an open question. In light of Corollaries 5 and 6 we conjecture that the answer is negative.

**A Appendix**

A normed space $(X, \|\cdot\|_X)$ is said to be imbedded into the normed space $(Y, \|\cdot\|_Y)$ if $X$ is a linear subspace of $Y$ and if the identity map $id : X \rightarrow Y$ is continuous. The imbedding is compact if
the image under the identity map of the unit ball of $X$ is precompact (i.e., totally bounded) in $Y$. We note here the following well-known consequence of the closed graph theorem which we use frequently throughout the paper: Let $X$ and $Y$ be linear subspaces of the vector space of real-valued functions on a (non-empty) set $\Omega \subseteq \mathbb{R}^d$ satisfying $X \subseteq Y$. Let $X$ and $Y$ be equipped with norms $\| \cdot \|_X$ and $\| \cdot \|_Y$, respectively, such that they become Banach spaces. Furthermore, suppose that each of these norm-topologies has the property that norm-convergence of a sequence implies that every subsequence possesses a further subsequence that converges $\lambda \| \Omega \text{-almost everywhere}$ (e.g., $L^p$-norms or Sobolev-norms). Then the map $id : X \to Y$ has a closed graph, and is thus continuous.

We recall the following definition of the entropy numbers of an imbedding:

**Definition 3** Let the normed space $(X, \| \cdot \|_X)$ be imbedded into the normed space $(Y, \| \cdot \|_Y)$. Let $U_X = \{ x \in X : \| x \|_X \leq 1 \}$ and $U_Y = \{ y \in Y : \| y \|_Y \leq 1 \}$ be the closed unit balls in $X$ and $Y$, respectively. Then, for all natural numbers $k$, the $k$-th entropy number of the imbedding operator $id : X \to Y$ is defined as

$$e(k, id(U_X), \| \cdot \|_Y) = \inf \left\{ \varepsilon > 0 : id(U_X) \subseteq \bigcup_{j=1}^{2^{k-1}} (y_j + \varepsilon U_Y) \quad \text{for suitable } y_1, \ldots, y_{2^k-1} \in Y \right\},$$

with the convention that the infimum equals $+\infty$ if the set over which it is taken is empty.

Clearly, $e(k, id(U_X), \| \cdot \|_Y)$ is finite for all $k$ if and only if $X$ is imbedded into $Y$, and the entropy numbers converge to zero as $k \to \infty$ if and only if the imbedding is compact. Obviously, the entropy numbers are closely related to the $\| \cdot \|_Y$-metric entropy of $id(U_X)$. (The definition as given here focuses on the unit ball in $X$, but can easily be extended to bounded subsets of the space $(X, \| \cdot \|_X)$ and to bounded operators other than $id$.)

We shall make use of results for weighted Besov spaces with weighting function $(x)^\beta = (1 + \| x \|^2)^{\beta/2}$, $\beta \in \mathbb{R}$. These spaces are denoted as $B^s_{pq}(\mathbb{R}^d, \langle x \rangle^\beta)$ with the associated norm denoted by $\| \cdot \|_{B^s_{(\beta),sp,q,\lambda}}$ where $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $s > 0$, and $\beta \in \mathbb{R}$. For a definition see Edmunds and Triebel (1996, 4.2.1/6). We note that these spaces with $\beta = 0$ coincide with the “unweighted” spaces $(B^s_{pq}(\mathbb{R}^d), \| \cdot \|_{B,s,p,q,\lambda})$ introduced in Section 2.2. It can be shown that

$$B^s_{pq}(\mathbb{R}^d, \langle x \rangle^\beta) = \left\{ f \in L^p(\mathbb{R}^d) : \| f \cdot \langle x \rangle^\beta \|_{B,s,p,q,\lambda,\lambda} < \infty \right\}$$

holds with equivalent norms

$$\| \cdot \|_{B(\beta),s,p,q,\lambda,\lambda} \sim \| \cdot \|_{B,s,p,q,\lambda};$$

see Edmunds and Triebel (1996), Theorem 4.2.2. [Yet another way of looking at these spaces is to view them as Besov spaces defined relative to the weighting measure $d\nu = \langle x \rangle^{p\beta} dx$ instead of Lebesgue measure; see Schmeisser and Triebel (1987), 5.1.4/5, where one finds a general treatment of weighted function spaces.] Like in the unweighted case, if $s - d/p > 0$ we again view the weighted Besov-spaces as spaces of continuous functions.

**Proposition 2** (Haroske and Triebel) Suppose $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $s - d/p > 0$, $\beta \in \mathbb{R}$, and $\gamma > 0$ hold. Then $B^s_{pq}(\mathbb{R}^d, \langle x \rangle^\beta)$ is compactly imbedded into $B^0_{\infty 1}(\mathbb{R}^d, \langle x \rangle^{\beta - \gamma})$. Furthermore, the entropy numbers of this imbedding satisfy:
1. If \( \gamma > s - d/p \) then
\[
e \left( k, \id \left( U_{B^p_{pq}(\mathbb{R}^d, (x)^s)} \right), \| \cdot \|_{B(\beta - \gamma), 0, \infty, 1, 1} \right) \sim k^{-s/d}
\]
for all \( k \in \mathbb{N} \).

2. If \( \gamma < s - d/p \) then
\[
e \left( k, \id \left( U_{B^p_{pq}(\mathbb{R}^d, (x)^s)} \right), \| \cdot \|_{B(\beta - \gamma), 0, \infty, 1, 1} \right) \sim k^{-(\gamma/d + 1/p)}
\]
for all \( k \in \mathbb{N} \).

**Proof.** This follows as a special case from Theorem 4.1 in Haroske and Triebel (2004); see also Haroske and Triebel (1994).

We next summarize some properties of bounded subsets in (weighted) Besov spaces which have been used in the paper. Since \( W^{s,p}(\mathbb{R}^d) \) is imbedded into \( B_{s,p}^0(\mathbb{R}^d) \) for either \( q = p \) (s non-integer) or \( q = 1 \) (s integer), the subsequent proposition also delivers corresponding results for the set \( D_{s,p}^0 \). For a measurable function \( f : \mathbb{R} \rightarrow \mathbb{R} \) its \( \pi \)-variation \( (1 < \pi < \infty) \) is defined as
\[
v_\pi (f) = \sup \left\{ \sum_{i=1}^n \left| f(x_i) - f(x_{i-1}) \right| : -\infty < x_0 < x_1 < \ldots < x_n < \infty, \ n \in \mathbb{N} \right\}
\]
and \( V_\pi (\mathbb{R}) \) denotes the linear space of functions for which \( v_\pi (f) < \infty \). Endowed with the norm \( \| f \|_{V_\pi} = \| f \|_\infty + (v_\pi (f))^{1/\pi} \), this space is a Banach space.

**Proposition 3** Suppose \( 1 \leq p \leq \infty, 1 \leq q \leq \infty, \beta \in \mathbb{R}, \) and \( s - d/p > 0 \).

1. If \( \beta \geq 0 \) and \( p < \infty \), then \( \lim_{\| x \| \rightarrow \infty} f(x) = 0 \) for all \( f \in B_{s,p}^\beta \).
2. \( \sup_{f \in B_{s,p}^\beta} |f(x)| \leq K \langle x \rangle^{-\beta} \) holds for some real number \( K > 0 \).
3. If \( \beta \geq 0 \), each \( f \in B_{s,p}^\beta \) is an element of the Hölder space \( C^{s-d/p} (\mathbb{R}^d) \).
4. For \( d = 1, s > 1/p, \) and \( p < \infty \), the Besov ball \( B_{s,p}^0 \) is contained in a bounded subset of \( V_p (\mathbb{R}) \); in fact, \( B_{s,p}^0 (\mathbb{R}) \) is imbedded into \( V_p (\mathbb{R}) \).

**Proof.** Since obviously \( B_{s,p}^\beta = B_{s,p}^0 \langle x \rangle^{-\beta} \) holds, it is sufficient to prove parts 1 and 3 only for the case \( \beta = 0 \). The first part then follows from Lemma 2 below. Since \( s - d/p > 0 \), we have the following chain of imbeddings
\[
B_{pq}^s (\mathbb{R}^d) \hookrightarrow C^{s-d/p} (\mathbb{R}^d) \hookrightarrow UC (\mathbb{R}^d)
\]
(Triebel (1983), 2.7.1/12-13). This immediately implies the third part of the proposition. To prove the second part, observe that the imbedding (9) implies for some real number \( c \) (depending only on \( s, p, q, d \)) and all \( f \in B_{s,p}^\beta \),
\[
\sup_{x \in \mathbb{R}^d} \left| f(x) \cdot \langle x \rangle^\beta \right| = \left\| f \cdot \langle x \rangle^\beta \right\|_\infty \leq c \left\| f \cdot \langle x \rangle^\beta \right\|_{B,s,p,q,\lambda} \leq cb.
\]
This then gives $|f| \leq K (x)^{-\beta}$ for $K = \epsilon b > 0$ and all $f \in B_{s,p,q}^2$.

We finally prove the last part. If $p = 1$ we have the following chain of imbeddings

$$B_{1,q}^s(\mathbb{R}) \hookrightarrow B_{1,1}^{s'}(\mathbb{R}) = W^{s',1}(\mathbb{R}) \hookrightarrow W^{1,1}(\mathbb{R}) \hookrightarrow C(\mathbb{R})$$

(10)

for some non-integer $s'$ satisfying $s > s' > 1$ (Triebel (1983, 2.3.2/7, 2.5.7/9) and Adams (1975, 5.4/7)). The space $W^{1,1}(\mathbb{R})$ consists of continuous functions $f$ with a Lebesgue-integrable generalized derivative. By Theorem 5.3.5 in Ziemer (1989) the essential total variation of every such $f$ is then finite. Since $f$ is continuous this implies that $v_p(f)$ is finite, and hence so is $\|f\|_{v_p} < \infty$.

This establishes the set-inclusion $B_{1,q}^s(\mathbb{R}) \subseteq \mathcal{V}_1(\mathbb{R})$. In case $1 < p < \infty$, consider the following chain of imbeddings

$$B_{p,q}^s(\mathbb{R}) \hookrightarrow B_{1,p}^{1/s}(\mathbb{R}) \rightarrow \dot{B}_{p,1}^{1/s}(\mathbb{R}) \cap C_0(\mathbb{R}) \hookrightarrow BV_p(\mathbb{R})$$

for $s > 1/p$ and $p < \infty$, where the first and second space are endowed with the respective Besov-norms. Furthermore, $\dot{B}_{p,1}^{1/s}(\mathbb{R})$ is the homogeneous Besov space (for a definition see 5.1.3/4 in Triebel (1983)), and the third space in the above display is endowed with the restriction of the usual (semi)norm on $\dot{B}_{p,1}^{1/s}(\mathbb{R})$. The fourth space is defined in Section 2.2 of Bourdaud, de Cristoforis and Sickel (2004). The first imbedding follows from Triebel (1983, 2.3.2/7). Since $B_{1,p}^{1/s}(\mathbb{R}) = L^p(\mathbb{R}, \lambda) \cap \dot{B}_{p,1}^{1/s}(\mathbb{R})$ (Triebel (1983), Remark 5.2.3/3) and since $B_{1,p}^{1/s}(\mathbb{R})$ is imbedded into $UC_0(\mathbb{R})$, and hence into $C_0(\mathbb{R})$, by Lemma 2 below, the set-inclusion

$$B_{1,p}^{1/s}(\mathbb{R}) \subseteq \dot{B}_{p,1}^{1/s}(\mathbb{R}) \cap C_0(\mathbb{R})$$

follows. Both spaces are Banach spaces. This is obvious for the first space, and follows from Proposition 10 of Bourdaud, de Cristoforis and Sickel (2004) for the second one. In view of this proposition and Lemma 2 below, it is obvious that norm convergence in both spaces implies convergence almost everywhere. The second imbedding now follows from the closed graph theorem as discussed in the first paragraph of the Appendix. The third imbedding is to be interpreted as a continuous quotient map (i.e., it associates to each function $f \in \dot{B}_{p,1}^{1/s}(\mathbb{R}) \cap C_0(\mathbb{R})$ the equivalence class $[f]$ modulo equality almost everywhere such that $f \in [f] \in BV_p(\mathbb{R})$). The existence of this imbedding follows from a modification of a theorem of Petree given as Theorem 5 in Bourdaud, de Cristoforis and Sickel (2004). By definition of the space $BV_p(\mathbb{R})$ it follows that every $f \in B_{p,q}^s(\mathbb{R})$ coincides with a function $g$ in $V_p(\mathbb{R})$ outside of a Lebesgue null-set $N$. Fix $\epsilon > 0$ and a grid of points such that $-\infty < x_0 < x_1 < \ldots < x_n < \infty$. Since $f$ is uniformly continuous by (9) and the null-set $N$ is nowhere dense, we can find $x_i^* \in \mathbb{R} \setminus N$ such that $|f(x_i) - f(x_i^*)| < (\epsilon/n)^{1/p}$. Then observing that $f$ and $g$ agree outside of $N$ we obtain

$$\sum_{i=1}^{n} |f(x_i) - f(x_i - 1)|^p$$

$$\leq \sum_{i=1}^{n} \left( |f(x_i) - f(x_i^*)| + |f(x_i^*) - f(x_i - 1)| + |f(x_i - 1) - f(x_i)| \right)^p$$

$$\leq 2^{2p-2} \left[ \sum_{i=1}^{n} |f(x_i) - f(x_i^*)|^p + |f(x_i^*) - f(x_i - 1)|^p + |f(x_i - 1) - f(x_i)|^p \right]$$

$$\leq 2^{2p-2} \varepsilon + 2^{2p-2} \sum_{i=1}^{n} |g(x_i^*) - g(x_i - 1)|^p \leq 2^{2p-2} \varepsilon + 2^{2p-2} v_p(g).$$

19
This shows that $\nu_p(f)$ is finite. Since also $\|f\|_\infty < \infty$ holds by (9) we have $\|f\|_{\nu_p} < \infty$. Summarizing, we have established the set-inclusion $B_{\nu_p}^p(\mathbb{R}) \subseteq V_p(\mathbb{R})$ for $1 \leq p < \infty$. That this inclusion is in fact an imbedding now follows from the closed graph theorem. 


**Lemma 2** Suppose $1 \leq p < \infty$, $1 \leq q \leq \infty$, and $s - d/p \geq 0$. If $s - d/p > 0$ or if $q = 1$ then $B_{\nu_p}^p(\mathbb{R}^d)$ is imbedded into $UC_0(\mathbb{R}^d)$.

**Proof.** We have the imbedding $B_{\nu_p}^p(\mathbb{R}^d) \hookrightarrow UC(\mathbb{R}^d)$ by (9) above in case $s - d/p > 0$, and by (Triebel (1983), 2.7.1/13) in case $s - d/p = 0$ and $q = 1$. Note that this implies $\|\cdot\|_\infty \lesssim \|\cdot\|_{B_{s,p,q,\lambda}}$. Therefore, since the Schwartz space $S$ of rapidly decreasing infinitely differentiable functions on $\mathbb{R}^d$ is dense in the Banach space $(B_{\nu_p}^p(\mathbb{R}^d), \|\cdot\|_{B_{s,p,q,\lambda}})$ for $p < \infty$ and $q < \infty$ (see Theorem 2.3.3 in Triebel (1983)) and since $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$ is the $\|\cdot\|_\infty$-completion of $S$, the conclusion of the lemma follows for $q < \infty$. If $q = \infty$, under the assumptions of the lemma $s - d/p > 0$ follows, and hence $B_{\nu_p}^p(\mathbb{R}^d)$ is imbedded into $B_{p,\infty}^{d/p}(\mathbb{R}^d)$ in view of Triebel (1983, 2.3.2/7), thus reducing this case to the case $q = 1 < \infty$ established before.


**Lemma 3** Suppose $s > 0$, $\beta \in \mathbb{R}$. Then there exists a finite constant $b'$ (depending on the radius $b$ of $C_0^{s,\lambda}$) such that $C_0^{s,\lambda} \subseteq \left\{ f \in L^0(\mathbb{R}^d) : \|f \cdot (x)^{\beta}\|_{B_{s,\infty,\infty,\lambda}} \leq b' \right\}$.

**Proof.** This follows from the fact that $\|f \cdot (x)^{\beta}\|_{B_{s,\infty,\infty,\lambda}} \lesssim \|f \cdot (x)^{\beta}\|_{s,\infty}$ holds by either 2.5.7/9, 2.5.7/6 (s non-integer) or 2.5.7/11 (s integer) of Triebel (1983).

**Acknowledgement.** We would like to acknowledge helpful comments on general issues related to the subject from H.G. Feichtinger, M. Kunzinger, V. Losert, H. Rindler, H. Triebel, and W.P. Ziemer.

**B References**


Department of Statistics
University of Vienna
Universitätsstrasse 5/3
1010 Vienna, Austria
E-mail: richard.nickl@univie.ac.at, benedikt.poetscher@univie.ac.at