A GENERAL WAVELET-BASED PROFILE DECOMPOSITION IN THE CRITICAL EMBEDDING OF FUNCTION SPACES

HAJER BAHOURI

Centre de Mathématiques Faculté de Sciences et Technologie
Université Paris XII - Val de Marne
61, avenue du Général de Gaulle
94010 Creteil Cedex, France
hbahouri@math.cnrs.fr

ALBERT COHEN*

Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
175 Rue du Chevaleret, 75013 Paris, France
cohen@ann.jussieu.fr

GABRIEL KOCH

Mathematical Institute
24-29 St Giles', Oxford, OX1 3LB, England
koch@maths.ox.ac.uk

We characterize the lack of compactness in the critical embedding of functions spaces $X \subset Y$ having similar scaling properties in the following terms: a sequence $(u_n)_{n \geq 0}$ bounded in X has a subsequence that can be expressed as a finite sum of translations and dilations of functions $(\phi_l)_{l > 0}$ such that the remainder converges to zero in Y as the number of functions in the sum and n tend to $+\infty$. Such a decomposition was established by Gérard in [13] for the embedding of the homogeneous Sobolev space $X = \dot{H}^s$ into the $Y = L^p$ in d dimensions with $0 < s = d/2 - d/p$, and then generalized by Jaffard in [15] to the case where X is a Riesz potential space, using wavelet expansions. In this paper, we revisit the wavelet-based profile decomposition, in order to treat a larger range of examples of critical embedding in a hopefully simplified way. In particular, we identify two generic properties on the spaces X and Y that are of key use in building the profile decomposition. These properties may then easily be checked for typical choices of X and Y satisfying critical embedding properties. These includes Sobolev, Besov, Triebel-Lizorkin, Lorentz, Hölder and BMO spaces.

Keywords: Sobolev embedding; defect of compactness; profile decomposition.

*Corresponding author

1. Introduction

The critical embedding of homogeneous Sobolev spaces in dimension d states that for $0 \leq t < s$ and $1 \leq p < q < \infty$ such that $d/p - d/q = s - t$, one has

$$\dot{W}^{s,p}(\mathbb{R}^d) \subset \dot{W}^{t,q}(\mathbb{R}^d). \tag{1.1}$$

The lack of compactness in this embedding can be described in terms of an asymptotic decomposition following Gérard [13] who considered the case $p = 2$ and $t = 0$, and Jaffard [15] who considered general values $p > 1$ with the Riesz potential spaces $\dot{H}^{s,p}$ in replacement of Sobolev spaces $\dot{W}^{s,p}$, again with $t = 0$. Their results can be formulated in the following terms: a sequence $(u_n)_{n \geq 0}$ bounded in $\dot{H}^{s,p}(\mathbb{R}^d)$ can be decomposed up to a subsequence extraction according to

$$u_n = \sum_{l=1}^{L} h_{l,n}^{s-d/p} \phi^l \left(\frac{\cdot - x_{l,n}}{h_{l,n}} \right) + r_{n,L}, \tag{1.2}$$

where $(\phi^l)_{l>0}$ is a family of functions in $\dot{H}^{s,p}(\mathbb{R}^d)$ and where

$$\lim_{L \to +\infty} \left(\limsup_{n \to +\infty} \|r_{n,L}\|_{L^q} \right) = 0.$$

This decomposition is "asymptotically orthogonal" in the sense that for $k \neq l$

$$|\log(h_{l,n}/h_{k,n})| \to +\infty \quad \text{or} \quad |x_{l,n} - x_{k,n}|/h_{l,n} \to +\infty, \quad \text{as } n \to +\infty.$$

This type of decomposition was also obtained earlier in [5] for a bounded sequence in $H_0^1(D, \mathbb{R}^3)$ of solutions of an elliptic problem, with D the open unit disk of $\mathbb{R}^2$ and in [27] and [26] for the critical injections of $W^{1,2}(\Omega)$ in Lebesgue space and of $W^{1,p}(\Omega)$ in Lorentz spaces respectively, with Ω a bounded domain of $\mathbb{R}^d$. They were also studied in [25] in an abstract Hilbert space framework and in [4] in the Heisenberg group context.

The above-mentioned references treat different types of examples of critical embedding by different methods. One of the motivations of this paper is to identify some fundamental mechanisms that lead to such results for a general critical embedding

$$X \subset Y,$$

in a unified way. Here X and Y are generic homogeneous function spaces which, similar to the above particular cases, have the same scaling properties in the sense that for any function f and $h > 0$

$$h^r \|f(h\cdot)\|_X = \|f\|_X \quad \text{and} \quad h^r \|f(h\cdot)\|_Y = \|f\|_Y, \tag{1.3}$$

for the same value of r. In a similar way to Jaffard, we use wavelet bases in order to construct the functions ϕ_l, yet in a somehow different and hopefully simpler way. Our construction is based on two basic key properties of wavelet expansions

in the spaces X and Y, which may then be easily checked on particular pairs of spaces of interest. In particular, any critical embedding involving Sobolev, Besov, Triebel-Lizorkin, Lorentz, Hölder or BMO spaces is covered by our approach.

The study of the lack of compactness in the critical embedding of Sobolev spaces supplies us with a large amount of information about solutions of nonlinear partial differential equations, both in the elliptic frame or the evolution frame. One has, for example,

- the pioneering works of P.-L. Lions [21] and [22] for the sake of geometric problems,
- the description of bounded energy sequences of solutions to the defocusing semilinear quintic wave equation, up to remainder terms small in energy norm in [2],
- the characterization of the defect of compactness for Strichartz estimates for the Schrödinger equation in [18],
- the understanding of features of solutions of nonlinear wave equations with exponential growth in [3],
- the sharp estimate of the time life-span of the focusing critical semilinear wave equation by means of the size of energy of the Cauchy data in [17],
- the study of the bilinear Strichartz estimates for the wave equation in [28].

For further applications, we refer to [10, 12, 14, 23, 20] and the references therein.

Our results which cover a broad spectrum of spaces could be at the origin of several prospectus of similar types of regularity results for Navier–Stokes systems (as in [16, 11]), qualitative study of nonlinear evolution equations or estimates of the span life of focusing semilinear dispersive evolution equations.

1.1. *Wavelet expansions*

Wavelet decompositions of a function have the form

$$f = \sum_{\lambda \in \nabla} d_\lambda \psi_\lambda, \tag{1.4}$$

where $\lambda = (j, k)$ concatenates the scale index $j = j(\lambda)$ and space index $k = k(\lambda)$: for $d = 1$, we have with the L^2 normalization,

$$\psi_{j,k} = \psi_\lambda = 2^{j/2}\psi(2^j \cdot -k), \quad j \in \mathbb{Z}, \ k \in \mathbb{Z},$$

where ψ is the so-called "mother wavelet". In higher dimension $d > 1$, one needs several generating functions ψ^e for $e \in E$ a finite set, so that setting $\psi_\lambda := (\psi_\lambda^e)_{e \in E}^T$ and $d_\lambda = (d_\lambda^e)_{e \in E}$, we can again write (1.4) with $d_\lambda \psi_\lambda$ a finite-dimensional inner product and

$$\psi_{j,k} = \psi_\lambda = 2^{dj/2}\psi(2^j \cdot -k), \quad j \in \mathbb{Z}, \ k \in \mathbb{Z}^d.$$

The index set ∇ in (1.4) is thus always defined as

$$\nabla := \mathbb{Z} \times \mathbb{Z}^d.$$

Note that λ may also be identified to a dyadic cube

$$\lambda \sim 2^{-j}(k + [0,1]^d).$$

We shall sometimes use the notation

$$|\lambda| := j(\lambda),$$

for the scale level of λ. In all the sequel, we systematically normalize our wavelets in X which is equivalent to normalizing them in Y in view of (1.3):

$$\psi_{j,k} = \psi_\lambda = 2^{rj}\psi(2^j \cdot - k). \tag{1.5}$$

It is known that, in addition to be Schauder bases, wavelet bases are unconditional bases for "most" classical function spaces, including in particular the family of Besov and Triebel-Lizorkin spaces: for such spaces X there exists a constant D such that for any finite subset $E \subset \nabla$ and coefficients vectors $(c_\lambda)_{\lambda \in E}$ and $(d_\lambda)_{\lambda \in E}$ such that $|c_\lambda| \leq |d_\lambda|$ for all λ, one has

$$\left\| \sum_{\lambda \in E} c_\lambda \psi_\lambda \right\|_X \leq D \left\| \sum_{\lambda \in E} d_\lambda \psi_\lambda \right\|_X. \tag{1.6}$$

We refer to [6, 7, 24] for more details on the construction of wavelet bases and on the characterization of classical function spaces by expansions in such bases.

1.2. *Main results*

Our profile decomposition relies on two key assumptions concerning wavelet decompositions and the spaces X and Y.

In addition we always work under the general assumption that our wavelet basis $(\psi_\lambda)_{\lambda \in \Lambda}$ is an unconditional basis for both spaces X and Y. We therefore assume that (1.6) *holds with some constant D for both norms.*

Our first assumption involves the *nonlinear projector* that we define for each $M > 0$ as follows: if $f \in X$ has the expansion in the wavelet basis given by (1.4), then

$$Q_M f := \sum_{\lambda \in E_M} d_\lambda \psi_\lambda, \tag{1.7}$$

where $E_M = E_M(f)$ is the subset of ∇ of cardinality M that corresponds to the M largest values of $|d_\lambda|$.

Such a set always exists due to the fact that $(\psi_\lambda)_{\lambda \in \Lambda}$ is a Schauder basis for X, since this implies that for any $\eta > 0$ only finitely many coefficients d_λ are larger than η in modulus. This set may however not be unique when some $|d_\lambda|$ are equal,

in which case we may choose an arbitrary realization of such a set. Recall that we have assumed the normalization (1.5) making $\|\psi_\lambda\|_X$ or $\|\psi_\lambda\|_Y$ independent of λ, therefore E_M also corresponds to the M largest $\|d_\lambda\psi_\lambda\|_X$ or $\|d_\lambda\psi_\lambda\|_Y$.

Assumption 1.1. The nonlinear projection satisfies

$$\lim_{M\to+\infty} \max_{\|f\|_X\leq 1} \|f - Q_M f\|_Y = 0. \tag{1.8}$$

The fact that the convergence of $Q_M f$ towards f in Y holds uniformly on the unit ball of X is tied to the nonlinear nature of the operator Q_M: if instead we took Q_M to be the projection onto a *fixed* M-dimensional space, then (1.8) would be in contradiction with the fact that the critical embedding of X into Y is not compact. As will be recalled further, nonlinear approximation theory actually allows for a more precise quantification of the above property in most cases of interest, through an estimate of the form

$$\max_{\|f\|_X\leq 1} \|f - Q_M f\|_Y \leq CM^{-s}, \quad M > 0$$

for some $s > 0$ and C only depending on the choice of X and Y. However, Assumption 1.1 alone will be sufficient for our purpose.

Our second assumption only concerns the behavior of wavelet expansions with respect to the X norm. It reflects the fact that this norm is stable with respect to certain operations such as "shifting" the indices of wavelet coefficients, as well as perturbating the value of these coefficients. This is expressed as follows.

Assumption 1.2. Consider a sequence of functions $(f_n)_{n>0}$ which are uniformly bounded in X and may be written as

$$f_n = \sum_{\lambda\in\nabla} c_{\lambda,n}\psi_\lambda, \tag{1.9}$$

and such that for all λ, the sequence $c_{\lambda,n}$ converges towards a finite limit c_λ as $n \to +\infty$. Then, the series $\sum_{\lambda\in\nabla} c_\lambda\psi_\lambda$ converges in X with

$$\left\|\sum_{\lambda\in\nabla} c_\lambda\psi_\lambda\right\|_X \leq C \liminf_{n\to+\infty} \|f_n\|_X, \tag{1.10}$$

where C is a constant only depending on the space X and on the choice of the wavelet basis.

As will be recalled further, for practical choices of X such as Besov or Triebel-Lizorkin spaces, the X norm of a function is equivalent to the norm of its wavelet coefficients in a certain sequence space. This allows us to establish (1.10) essentially by invoking Fatou's lemma.

We are now in a position to state the main theorem of this paper. For any function ϕ, not necessarily a wavelet, and any scale-space index $\lambda = (j, k)$ we use the notation

$$\phi_\lambda := 2^{rj} \phi(2^j \cdot -k), \tag{1.11}$$

for the version of ϕ scaled and translated according to λ.

Theorem 1.1. *Assume that X and Y are two function spaces with the same scaling (1.3) and continuous embedding $X \subset Y$, and assume that there exists a wavelet basis $(\psi_\lambda)_{\lambda \in \nabla}$ which is unconditional for both X and Y, and such that Assumptions 1.1 and 1.2 hold. Let $(u_n)_{n>0}$ be a bounded sequence in X. Then, up to subsequence extraction, there exists a family of functions $(\phi^l)_{l>0}$ in X and sequences of scale-space indices $(\lambda_l(n))_{n>0}$ for each $l > 0$ such that*

$$u_n = \sum_{l=1}^{L} \phi^l_{\lambda_l(n)} + r_{n,L}, \tag{1.12}$$

where

$$\lim_{L \to +\infty} \left(\limsup_{n \to +\infty} \|r_{n,L}\|_Y \right) = 0.$$

The decomposition (1.12) is asymptotically orthogonal in the sense that for any $k \neq l$,

$$|j(\lambda_k(n)) - j(\lambda_l(n))| \to +\infty \quad or$$
$$|k(\lambda_k(n)) - 2^{j(\lambda_k(n)) - j(\lambda_l(n))} k(\lambda_l(n))| \to +\infty, \quad as \quad n \to +\infty. \tag{1.13}$$

Moreover, we have the following for the specific case where X is a Besov or Triebel-Lizorkin space:

Theorem 1.2. *The decomposition in Theorem 1.1 is stable in the sense that, for some $\tau = \tau(X)$ we have*

$$\|(\|\phi^l\|_X)_{l>0}\|_{\ell^\tau} \leq CK, \tag{1.14}$$

where C is a constant that only depends on X and on the choice of the wavelet basis and where $K := \sup_{n \geq 0} \|u_n\|_X$.

Remark 1.1. For certain sequences $(u_n)_{n>0}$, it is possible that for any $L > 0$ the decomposition (1.12) only involves a finite number of profiles ϕ^l for $l = 1, \dots, L_0$, which means that $\phi^l = 0$ for $l > L_0$. Inspection of our proof shows that the theorem

remains valid in such a case, in the sense that

$$u_n = \sum_{l=1}^{L_0} \phi^l_{\lambda_l(n)} + r_n,$$ (1.15)

where

$$\lim_{n \to +\infty} \|r_n\|_Y = 0.$$

In particular, the sequence $(u_n)_{n>0}$ is compact in Y if and only if $\phi^l = 0$ for all $l > 0$.

Remark 1.2. Inspection of our proof also shows that in Assumption 1.1, we may use for Q_M a more general nonlinear projector than the one obtained by taking the M largest values of $|d_\lambda|$. Generally speaking, we may consider a nonlinear projector Q_M that has the general form (1.7), where the sets $E_M = E_M(f)$ of cardinality M depend on f and satisfy

$$E_M(f) \subset E_{M+1}(f).$$

Such a generalization appears to be useful when treating certain types of embedding, see Sec. 3.

1.3. *Layout*

The effective construction of the decomposition is addressed in Sec. 2, together with the proof of Theorem 1.1.

In Sec. 3, we discuss examples of X and Y with critical embedding for which Assumptions 1.1 and 1.2 can be proved. This includes all previously treated cases, and many others such as the embedding of Sobolev, Besov and Triebel-Lizorkin spaces into Lebesgue, Lorentz, BMO and Hölder spaces, or into other Sobolev, Besov and Triebel-Lizorkin spaces.

Finally, in Sec. 4, we prove the stability Theorem 1.2 for both setting of Besov and Triebel-Lizorkin spaces.

2. Construction of the Decomposition and Proof of Theorem 1.1

In this section, we place ourselves under the assumptions of Theorem 1.1. Let $(u_n)_{n>0}$ be a bounded sequence in the space X and define

$$K := \sup_{n>0} \|u_n\|_X < +\infty.$$

The decomposition construction and the proof of Theorem 1.1 proceed in several steps.

Step 1. Rearrangements. We first introduce the wavelet decompositions

$$u_n = \sum_{\lambda \in \nabla} d_{\lambda,n} \psi_\lambda. \tag{2.1}$$

For each $n > 0$, we consider the non-increasing rearrangement $(d_{m,n})_{m>0}$ of $(d_{\lambda,n})_{\lambda \in \nabla}$ according to their moduli. We may therefore write

$$u_n = \sum_{m>0} d_{m,n} \psi_{\lambda(m,n)}. \tag{2.2}$$

Using the nonlinear projector Q_M defined by (1.7), we further split this expansion into

$$u_n = \sum_{m=1}^{M} d_{m,n} \psi_{\lambda(m,n)} + R_M u_n, \tag{2.3}$$

with $R_M u_n = u_n - Q_M u_n$. Combining Assumption 1.1 with the boundedness of $(u_n)_{n>0}$ in X, we infer that

$$\lim_{M \to +\infty} \sup_{n>0} \|R_M u_n\|_Y = 0. \tag{2.4}$$

Our next observation is that if $(\psi_\lambda)_{\lambda \in \nabla}$ is an unconditional basis of X, then the coefficients $d_{m,n}$ are uniformly bounded: indeed, (1.6) implies that the rank one projectors

$$P_\mu : f = \sum_{\lambda \in \Lambda} d_\lambda \psi_\lambda \mapsto P_\mu f := d_\mu \psi_\mu,$$

satisfy the uniform bound

$$\|P_\mu\|_{X \to X} \leq D, \quad \mu \in \nabla.$$

Since we have assumed that our wavelets are normalized in X, for example according to $\|\psi_\mu\|_X = 1$ for all $\mu \in \nabla$, we thus have

$$\sup_{\lambda,n} |d_{\lambda,n}| = \sup_{m,n} |d_{m,n}| \leq DK.$$

Up to a diagonal subsequence extraction procedure in n, we may therefore assume that for all $m > 0$, the sequence $(d_{m,n})_{n>0}$ converges towards a finite limit that depends on m,

$$d_m = \lim_{n \to +\infty} d_{m,n}.$$

Note that $(|d_m|)_{m>0}$ is a non-increasing sequence since all sequences $(|d_{m,n}|)_{m>0}$ are non-increasing. We may thus write

$$u_n = \sum_{m=1}^{M} d_m \psi_{\lambda(m,n)} + t_{n,M},$$

where

$$t_{n,M} := \sum_{m=1}^{M} (d_{m,n} - d_m)\psi_{\lambda(m,n)} + R_M u_n.$$

Step 2. Construction of approximate profiles. We construct the profiles ϕ^l as limit of sequences $\phi^{l,i}$ obtained by the following algorithm. At the first iteration $i = 1$, we set

$$\phi^{1,1} = d_1\psi, \quad \lambda_1(n) := \lambda(1,n), \quad \varphi_1(n) := n. \tag{2.5}$$

Assume that after iteration $i - 1$, we have constructed $L - 1$ functions $(\phi^{1,i}, \ldots, \phi^{L-1,i})$ and scale-space index sequences $(\lambda_1(n), \ldots, \lambda_{L-1}(n))$ with $L \leq i$, as well as an increasing sequence of positive integers $\varphi_{i-1}(n)$ such that

$$\sum_{m=1}^{i-1} d_m\psi_{\lambda(m,\varphi_{i-1}(n))} = \sum_{l=1}^{L-1} \phi^{l,i}_{\lambda_l(\varphi_{i-1}(n))}.$$

At iteration i we shall use the ith component $d_i\psi_{\lambda(i,\varphi_{i-1}(n))}$ to either modify one of these functions or build a new one according to the following dichotomy.

(i) First case: assume that we can extract $\varphi_i(n)$ from $\varphi_{i-1}(n)$ such that for $l = 1, \ldots, L - 1$ at least one of the following holds:

$$\lim_{n \to +\infty} |j(\lambda_l(\varphi_i(n))) - j(\lambda(i, \varphi_i(n)))| = +\infty \tag{2.6}$$

or

$$\lim_{n \to +\infty} |k(\lambda(i, \varphi_i(n))) - 2^{j(\lambda(i,\varphi_i(n)))-j(\lambda_l(\varphi_i(n)))}k(\lambda_l(\varphi_i(n)))| = +\infty. \tag{2.7}$$

In such a case, we create a new profile and scale-space index sequence by defining

$$\phi^{L,i} := d_i\psi, \quad \lambda_L(n) := \lambda(i,n)$$

and we set $\phi^{l,i} = \phi^{l,i-1}$ for $l = 1, \ldots, L - 1$.

(ii) Second case: assume that for some subsequence $\varphi_i(n)$ of $\varphi_{i-1}(n)$ and some $l \in \{1, \ldots, L - 1\}$ both (2.6) and (2.7) do not hold. Then it is easily checked that $j(\lambda_l(\varphi_i(n))) - j(\lambda(i, \varphi_i(n)))$ and $k(\lambda(i, \varphi_i(n))) - 2^{j(\lambda(i,\varphi_i(n)))-j(\lambda_l(\varphi_i(n)))}k(\lambda_l(\varphi_i(n)))$ only take a finite number of values as n varies. Therefore, up to an additional subsequence extraction, we may assume that there exist numbers a and b such that for all $n > 0$,

$$j(\lambda(i, \varphi_i(n))) - j(\lambda_l(\varphi_i(n))) = a \tag{2.8}$$

and

$$k(\lambda(i, \varphi_i(n))) - 2^{j(\lambda(i,\varphi_i(n)))-j(\lambda_l(\varphi_i(n)))}k(\lambda_l(\varphi_i(n))) = b. \tag{2.9}$$

We then update the function $\phi^{l,i-1}$ according to

$$\phi^{l,i} = \phi^{l,i-1} + d_i 2^{ar}\psi(2^a \cdot -b) \tag{2.10}$$

and $\phi^{l',i} = \phi^{l',i-1}$ for $l' \in \{1,\ldots,L-1\}$ and $l' \neq l$.

From this construction, and after extracting a diagonal subsequence which eventually coincides with a subsequence of $\varphi_i(n)$ for each i, we see that for each value of M there exists $L = L(M) \leq M$ such that

$$\sum_{m=1}^{M} d_m \psi_{\lambda(m,n)} = \sum_{l=1}^{L} \phi_{\lambda_l(n)}^{l,M}.$$

More precisely, for each $l = 1, \ldots, L$, we have

$$\phi_{\lambda_l(n)}^{l,M} = \sum_{m \in E(l,M)} d_m \psi_{\lambda(m,n)},$$

where the sets $E(l,M)$ for $i = 1, \ldots, L$ constitute a disjoint partition of $\{1, \ldots, M\}$. Note that $E(l,M) \subset E(l,M+1)$ with $\#(E(l,M+1)) \leq \#(E(l,M))+1$. Similarly, the number of profiles $L(M)$ grows at most by 1 as we move from M to $M+1$. As explained in Remark 1.2, it is possible that $L(M)$ terminates at some maximal value L_0. Finally, note that for any $m,m' \in E_{l,M}$, we have that

$$j(\lambda(m,n)) - j(\lambda(m',n)) = a(m,m') \tag{2.11}$$

and

$$k(\lambda(m,n)) - 2^{j(\lambda(m,n))-j(\lambda(m',n))}k(\lambda(m',n)) = b(m,m'), \tag{2.12}$$

where $a(m,m')$ and $b(m,m')$ do not depend on n.

Step 3. Construction of the exact profiles. We now want to define the functions ϕ^l as the limits in X of $\phi^{l,M}$ as $M \to +\infty$. For this purpose, we shall make use of Assumption 1.2, combined with the scaling property (1.3) of the X norm and the fact that $(\psi_\lambda)_{\lambda \in \nabla}$ is an unconditional basis. For some fixed l and M such that $l \leq L(M)$, let us define the functions

$$g^{l,M} := \sum_{m \in E(l,M)} d_m \psi_{\lambda(m)},$$

$$f^{l,M,n} := \sum_{m \in E(l,M)} d_{m,n} \psi_{\lambda(m)},$$

with $\lambda(m) := \lambda(m,1)$. From the scaling property (1.3) and the properties (2.11) and (2.12), we find that

$$\|f^{l,M,n}\|_X = \left\| \sum_{m \in E(l,M)} d_{m,n} \psi_{\lambda(m,n)} \right\|_X.$$

Since $\sum_{m \in E(l,M)} d_{m,n} \psi_{\lambda(m,n)}$ is a part of the expansion of u_n, we thus find that

$$\|f^{l,M,n}\|_X \leq DK,$$

where D is the constant in (1.6) and $K := \sup_{n>0} \|u_n\|_X$. Invoking Assumption 1.2, we therefore find that $g^{l,M}$ converges in X towards a limit g^l as $M \to +\infty$. We finally notice that, by construction, the $g^{l,M}$ are rescaled versions of the $\phi^{l,M}$: there exists $A > 0$ and $B \in \mathbb{R}^d$ such that

$$\phi^{l,M} = 2^{Ar} g^{l,M} (2^A \cdot -B).$$

By (1.3), we therefore conclude that $\phi^{l,M}$ converges in X towards a limit $\phi^l := 2^{Ar} g^l (2^A \cdot -B)$ as $M \to +\infty$.

Step 4. Conclusion of the proof. For any given $L > 0$, we may write

$$u_n = \sum_{l=1}^{L} \phi^l_{\lambda_l(n)} + r_{n,L},$$

where, for any value of M such that $L \leq L(M)$, the remainder $r_{n,L}$ may be decomposed into

$$\sum_{l=1}^{L} \left(\phi^{l,M}_{\lambda_l(n)} - \phi^l_{\lambda_l(n)} \right) + \sum_{l=1}^{L} \sum_{m \in E(l,M)} (d_{m,n} - d_m) \psi_{\lambda(m,n)}$$

$$+ \sum_{l=L+1}^{L(M)} \sum_{m \in E(l,M)} d_{m,n} \psi_{\lambda(m,n)} + R_M u_n. \tag{2.13}$$

Note that each of these terms depends on the chosen value of M but their sum $r_{n,L}$ is actually independent of M. We rewrite this decomposition as

$$r_{n,L} = r_1(n, L, M) + r_2(n, L, M),$$

where r_1 and r_2 stand for the first and last terms in (2.13), respectively. By construction, all values of m which appear in the third term of (2.13) are between $L+1$ and M. Therefore the last two terms in (2.13) may be viewed as a partial sum of

$$R_L u_n = \sum_{m>L} d_{m,n} \psi_{\lambda(m,n)}.$$

Since we have assumed that $(\psi_\lambda)_{\lambda \in \nabla}$ is an unconditional basis for Y, we may therefore write

$$\|r_2(n, L, M)\|_Y \leq D \|R_L u_n\|_Y.$$

According to Assumption 1.1, which is expressed by (2.4), the right-hand side converges to 0 as $L \to +\infty$ uniformly in n and therefore

$$\lim_{L \to +\infty} \sup_{n,M} \|r_2(n, L, M)\|_Y = 0.$$

We now consider the first two terms in (2.13). For the first term, we have

$$\left\| \sum_{l=1}^{L} (\phi_{\lambda_l(n)}^{l,M} - \phi_{\lambda_l(n)}^l) \right\|_X \leq \sum_{l=1}^{L} \left\| \phi_{\lambda_l(n)}^{l,M} - \phi_{\lambda_l(n)}^l \right\|_X = \sum_{l=1}^{L} \left\| \phi^{l,M} - \phi^l \right\|_X.$$

Therefore, for any fixed L, this term goes to 0 in X as $M \to +\infty$. For the second term, we first notice that for any fixed L and M such that $L \leq L(M)$, all values of m which appear in this term are less than or equal to M. Since we have assumed that $(\psi_\lambda)_{\lambda \in \nabla}$ is an unconditional basis for X, it follows that

$$\left\| \sum_{l=1}^{L} \sum_{m \in E(l,M)} (d_{m,n} - d_m) \psi_{\lambda(m,n)} \right\|_X \leq D \left\| \sum_{m=1}^{M} (d_{m,n} - d_m) \psi_{\lambda(m,n)} \right\|_X$$

$$\leq CD \sum_{m=1}^{M} |d_{m,n} - d_m|,$$

where $C = \|\psi\|_X = \|\psi_\lambda\|_X$ for all $\lambda \in \nabla$. Therefore for any fixed L and M such that $L \leq L(M)$, this term goes to 0 in X as $n \to +\infty$. Combining these observations, we find that for any fixed L and any $\varepsilon > 0$, there exist M and n_0 such that for all $n \geq n_0$,

$$\|r_1(n, L, M)\|_X \leq \varepsilon.$$

By continuous embedding, the same holds for $\|r_1(n, L, M)\|_Y$. Since M was arbitrary in the decomposition (2.13) of $r_{n,L}$, we obtain that

$$\lim_{L \to +\infty} \left(\limsup_{n \to +\infty} \|r_{n,L}\|_Y \right) = 0,$$

which concludes the proof of the theorem.

3. Examples

Our main result applies to a large range of critical embedding. Specifically, we consider

(i) For the space X: spaces of Besov type $\dot{B}_{p,a}^s$ or Triebel-Lizorkin type $\dot{F}_{p,a}^s$ with $1 \leq p < \infty$ and $1 \leq a \leq \infty$.

(ii) For the space Y: spaces of Besov type $\dot{B}_{q,b}^t$, Triebel-Lizorkin type $\dot{F}_{q,b}^t$, Lebesgue type L^q, Lorentz type $L^{q,b}$, and the space BMO, with $1 \leq q \leq \infty$ and $1 \leq b \leq \infty$.

Note that Lebesgue spaces may be thought of as a particular case of Triebel-Lizorkin spaces since $L^q = \dot{F}^0_{q,2}$, yet we treat them separately since several results that we invoke further have been proved in an isolated manner for the specific case of Lebesgue spaces.

The critical embedding for such spaces imposes that $t < s$ together with the scaling

$$r = \frac{d}{p} - s = \frac{d}{q} - t, \tag{3.1}$$

where $t = 0$ if Y is of Lebesgue or Lorentz type, and $t = 0$ and $q = \infty$ if $Y = BMO$. It also imposes some relations between the fine-tuning indices a and b. For example for $s > 0$ and p, q such that $\frac{d}{p} - s = \frac{d}{q}$, the space $\dot{B}^s_{p,a}$ embeds continuously into $L^{q,b}$ if $b \geq a$.

Note that for non-integer $t > 0$ the Hölder space $\dot{C}^t$ coincides with the Besov space $\dot{B}^t_{\infty,\infty}$, and that for all integer $m \geq 0$, the Sobolev space $\dot{W}^{m,p}$ coincides with the Triebel-Lizorkin space $\dot{F}^m_{p,2}$ when $1 < p < \infty$. In particular, $L^p = \dot{F}^0_{p,2}$ for $1 < p < \infty$. For $p = 1$, it is known that $\dot{F}^0_{1,2}$ coincides with the Hardy space $\mathcal{H}^1$ which is a closed subspace of L^1. We refer to [1] and [29] for an introduction to all such spaces.

It is known that properly constructed wavelet bases are unconditional for all such spaces, see in particular [24]. In addition, Besov and Triebel-Lizorkin spaces, as well as BMO, may be characterized by simple properties on wavelet coefficients. More precisely, for $f = \sum_{\lambda \in \nabla} d_\lambda \psi_\lambda$ and wavelets normalized according to (1.5) with r given by (3.1), we have the following norm equivalences (see [6, 7, 24]):

(i) For Besov spaces,

$$\|f\|_{\dot{B}^t_{q,b}} \sim \left(\sum_{j \in \mathbb{Z}} \left(\sum_{|\lambda|=j} |d_\lambda|^q \right)^{b/q} \right)^{1/b}, \tag{3.2}$$

with the standard modification when $q = \infty$ or $b = \infty$.

(ii) For Triebel-Lizorkin spaces,

$$\|f\|_{\dot{F}^t_{q,b}} \sim \left\| \left(\sum_{\lambda \in \nabla} |d_\lambda \chi_\lambda|^b \right)^{1/b} \right\|_{L^q}, \tag{3.3}$$

where $\chi_\lambda = 2^{dj/q} \chi(2^j \cdot -k)$ with $\chi = \chi_{[0,1]^d}$ for $\lambda \sim (j,k)$. When $b = \infty$, $(\sum_{\lambda \in \nabla} |d_\lambda \chi_\lambda|^b)^{1/b}$ should be replaced by $\sup_{j \in \mathbb{Z}} \sum_{|\lambda|=j} |d_\lambda \chi_\lambda|$.

(iii) For BMO,

$$\|f\|_{BMO} \sim \max_{\lambda \in \nabla} \left(2^{d|\lambda|} \sum_{\mu \subset \lambda} |d_\mu|^2 2^{-d|\mu|} \right)^{1/2}, \tag{3.4}$$

where by definition $\mu \subset \lambda$ means that $2^{-j(\mu)}([0,1]^d + k(\mu)) \subset 2^{-j(\lambda)}([0,1]^d + k(\lambda))$.

Note that due to the discretization of the scale-space index d_λ, the above equivalent norms do not exactly satisfy the scaling relation (1.3). These norm equivalences readily imply that $(\psi_\lambda)_{\lambda \in \Lambda}$ is an unconditional basis for such spaces. Note that there exists no simple wavelet characterization of Lorentz spaces $L^{q,b}$ when $b \neq q$. However, the unconditionality of $(\psi_\lambda)_{\lambda \in \Lambda}$ in such spaces follows by interpolation of Lebesgue spaces for any $1 < b, q < \infty$.

We now need to discuss the validity of Assumptions 1.1 and 1.2, for such choices of spaces. We first discuss Assumption 1.2 which is only concerned with the space X. Since we assumed here that X is of Besov type $\dot{B}^s_{p,a}$ or Triebel-Lizorkin type $\dot{F}^s_{p,a}$, we may use equivalent norms given by (3.2) and (3.3). Therefore, the X norm of $f_n = \sum_{\lambda \in \nabla} c_{\lambda,n} \psi_\lambda$ is either equivalent to

$$\left(\sum_{j \in \mathbb{Z}} \left(\sum_{|\lambda|=j} |c_{\lambda,n}|^p \right)^{a/p} \right)^{1/a}$$

or

$$\left\| \left(\sum_{\lambda \in \nabla} |c_{\lambda,n} \chi_\lambda|^a \right)^{1/a} \right\|_{L^p}.$$

In both cases, we may invoke Fatou's lemma to conclude that for the limit sequence (c_λ), we have

$$\left(\sum_{j \in \mathbb{Z}} \left(\sum_{|\lambda|=j} |c_\lambda|^p \right)^{a/p} \right)^{1/a} \leq \liminf_{n \to +\infty} \left(\sum_{j \in \mathbb{Z}} \left(\sum_{|\lambda|=j} |c_{\lambda,n}|^p \right)^{a/p} \right)^{1/a}$$

and

$$\left\| \left(\sum_{\lambda \in \nabla} |c_\lambda \chi_\lambda|^a \right)^{1/a} \right\|_{L^p} \leq \liminf_{n \to +\infty} \left\| \left(\sum_{\lambda \in \nabla} |c_{\lambda,n} \chi_\lambda|^a \right)^{1/a} \right\|_{L^p}.$$

Therefore, Assumption 1.2 holds for all Besov and Triebel-Lizorkin spaces.

We next discuss Assumption 1.1, for some specific examples of pairs X and Y which satisfy the critical embedding property. The study of the nonlinear projector Q_M is an important chapter of approximation theory. The process of approximating a function

$$f = \sum_{\lambda \in \nabla} d_\lambda \psi_\lambda,$$

by a function of the form

$$\sum_{\lambda \in E_M} c_\lambda \psi_\lambda,$$

with $\#(E_M) \leq M$ is sometimes called *best M-term approximation*, and has been studied extensively. The most natural choice is to take for E_M the indices corresponding to the largest coefficients $|d_\lambda|$ and to set $c_\lambda = d_\lambda$, which corresponds to our definition of Q_M. However as already mentioned in Remark 1.2, other more relevant choices could be used if necessary for proving the validity of Assumption 1.1 for certain pairs (X, Y) and a specific instance will be mentioned below.

The study of the convergence of $Q_M f$ towards f is particularly elementary in the case where $X = \dot{B}^s_{p,p}$ and $Y = \dot{B}^t_{q,q}$, with $\frac{1}{p} - \frac{1}{q} = \frac{s-t}{d}$. Indeed, according to (3.2), we have for such spaces

$$\|f\|_{\dot{B}^s_{p,p}} \sim \|(d_\lambda)_{\lambda \in \nabla}\|_{\ell^p} \quad \text{and} \quad \|f\|_{\dot{B}^t_{q,q}} \sim \|(d_\lambda)_{\lambda \in \nabla}\|_{\ell^q},$$

and therefore for any $f \in X$, using the decreasing rearrangement $(d_m)_{m>0}$ of the $|d_\lambda|$, we obtain

$$
\begin{aligned}
\|f - Q_M f\|_{\dot{B}^t_{q,q}} &\sim \left(\sum_{\lambda \notin E_M} |d_\lambda|^q \right)^{1/q} \\
&= \left(\sum_{m>M} |d_m|^q \right)^{1/q} \\
&\leq |d_M|^{1-p/q} \left(\sum_{m>M} |d_m|^p \right)^{1/q} \\
&\leq \left(M^{-1} \sum_{m=1}^{M} |d_m|^p \right)^{1/p-1/q} \left(\sum_{m>M} |d_m|^p \right)^{1/q} \\
&\leq M^{-(1/p-1/q)} \left(\sum_{m>0} |d_m|^p \right)^{1/p} \\
&\leq M^{-\frac{s-t}{d}} \|(d_\lambda)_{\lambda \in \nabla}\|_{\ell^p} \sim M^{-\frac{s-t}{d}} \|f\|_{\dot{B}^s_{p,p}}.
\end{aligned}
$$

We have thus proved that

$$\sup_{\|f\|_{\dot{B}^s_{p,p}} \leq 1} \|f - Q_M f\|_{\dot{B}^t_{q,q}} \leq C M^{-\sigma}, \quad \sigma := \frac{s-t}{d} > 0, \tag{3.5}$$

which shows that Assumption 1.1 holds in such a case.

For other choices of X and Y, the study of best M-term approximation is more involved and we just describe the available results without proof.

The case of the embedding of the Besov space $X = \dot{B}^s_{p,p}$ into the Lebesgue space $Y = L^q$, with $q < \infty$ and $\frac{1}{p} - \frac{1}{q} = \frac{s}{d} > 0$ has first been treated in [9] — see also [6] and [8] — where it was proved that

$$\sup_{\|f\|_{\dot{B}^s_{p,p}} \leq 1} \|f - Q_M f\|_{L^q} \leq CM^{-\sigma}, \quad \sigma := \frac{s}{d} > 0. \tag{3.6}$$

Therefore Assumption 1.1 also holds in such a case. Note that when $q \leq 2$, one has continuous embedding of $\dot{B}^0_{q,q}$ in L^q and therefore (3.6) may be viewed as a consequence of (3.5), however this is no longer the case when $2 \leq q < \infty$, yet (3.6) still holds.

A finer result, that may be obtained by interpolation techniques, states that, with the same relations between p and q, the Besov space $\dot{B}^s_{p,q}$ — which is strictly larger than $\dot{B}^s_{p,p}$ is continuously embedded in L^q, and one may therefore ask if Assumption 1.1 is still valid in such a case. A positive answer was given in [19] for the more general embedding of $X = \dot{B}^s_{p,q}$ into $Y = \dot{F}^t_{q,b}$ with $\frac{1}{p} - \frac{1}{q} = \frac{s-t}{d}$, where $b \in]0, \infty]$ is arbitrary: we have the convergence estimate

$$\sup_{\|f\|_{\dot{B}^s_{p,q}} \leq 1} \|f - Q_M f\|_{\dot{F}^t_{q,b}} \leq CM^{-\sigma}, \quad \sigma := \frac{s-t}{d} > 0, \tag{3.7}$$

and therefore Assumption 1.1 is again valid. Note that $L^q = \dot{F}^0_{q,2}$ is a particular case.

For Besov spaces, the critical embedding of $X = \dot{B}^s_{p,a}$ into $Y = \dot{B}^t_{q,b}$ with $\frac{1}{p} - \frac{1}{q} = \frac{s-t}{d}$ is known to hold whenever $a \leq b$ (it is an immediate consequence of the norm equivalence (3.2)). The study of best M-term approximation in this context was done in [19], where the following result was proved: there exists a nonlinear projector Q_M of the form (1.7), such that when $\frac{1}{a} - \frac{1}{b} \geq \frac{s-t}{d}$, one has

$$\sup_{\|f\|_{\dot{B}^s_{p,a}} \leq 1} \|f - Q_M f\|_{\dot{B}^t_{q,b}} \leq CM^{-\sigma}, \quad \sigma := \frac{s-t}{d} > 0. \tag{3.8}$$

The set $E_M(f)$ used in the definition of Q_M is however not generally based on picking the M largest $|d_\lambda|$, which is not a problem for our purposes as already mentioned in Remark 1.2. Therefore, Assumption 1.1 is valid for such pairs.

In this last example, the restriction $\frac{1}{a} - \frac{1}{b} \geq \frac{s-t}{d}$ is stronger than $a \leq b$ which is sufficient for the critical embedding. However, we may still obtain the validity of Assumption 1.1 when $a < b$ by a general trick which we shall re-use further: introduce an auxiliary space Z with continuous embedding

$$X \subset Z \subset Y, \tag{3.9}$$

such that Assumption 1.1 either holds for the embedding between X and Z, or between Z and Y, which immediately implies the validity of Assumption 1.1

between X and Y. In the present case we take

$$Z = \dot{B}_{\tilde{p},a}^{\tilde{s}} \quad \text{with } t < \tilde{s} < s, \quad \frac{1}{p} - \frac{1}{\tilde{p}} = \frac{s - \tilde{s}}{d}.$$

The continuous embeddings (3.9) clearly hold. In addition, when $\tilde{s}$ is sufficiently close to t, we have that $\frac{1}{a} - \frac{1}{b} \geq \frac{\tilde{s}-t}{d}$, so that Assumption 1.1 is valid for the pair (Z, Y) according to (3.8), and thus also for (X, Y).

Remark 3.1. It is not difficult to check that Assumption 1.1 does not hold for the embedding of $X = \dot{B}_{p,a}^s$ into $Y = \dot{B}_{q,a}^t$, and we also conjecture that the profile decomposition does not generally exist for such an embedding. As an example, consider $a = \infty$, and a sequence $(u_n)_{n>0}$ obtained by piling up one wavelet at each scale $j = 0, \ldots, n$ at position $k = 0$:

$$u_n = \sum_{j=0}^{n} 2^{rj} \psi(2^j \cdot).$$

All wavelets in u_n contribute equally to the X and Y norm (which is equivalent to the supremum of the coefficients, equal to 1) and the extraction of profiles with asymptotically orthogonal scale-space localization seems impossible.

The above trick based on the intermediate space Z may be used to prove Assumption 1.1 for other types of critical embeddings:

- Embedding of Besov spaces into BMO:

$$\dot{B}_{p,p}^s \subset BMO, \quad s = \frac{d}{p} > 0,$$

  which includes as a particular case the well-known embedding $\dot{H}^{d/2} \subset BMO$, and may be easily proved from the wavelet characterization (3.2) and (3.4). Choosing $Z = \dot{B}_{\tilde{p},\tilde{p}}^{\tilde{s}}$ for any $0 < \tilde{s} < s$ and $\tilde{p}$ such that $\tilde{s} = \frac{d}{\tilde{p}}$, we clearly have the continuous embeddings (3.9). In addition, Assumption 1.1 is valid for the pair (X, Z) according to (3.5), and thus also for (X, Y).
- Embedding of Besov spaces into Lorentz spaces:

$$\dot{B}_{p,a}^s \subset L^{q,b}, \quad \frac{1}{p} - \frac{1}{q} = \frac{s}{d} > 0,$$

  which is valid for any $a \leq b$. If $a < b$, we may introduce for any $0 < \tilde{s} < s$

$$Z = \dot{B}_{\tilde{p},b}^{\tilde{s}}, \quad \frac{1}{\tilde{p}} - \frac{1}{q} = \frac{\tilde{s}}{d} > 0,$$

  so that we have the continuous embeddings (3.9). In addition, we have already proved that Assumption 1.1 holds for the pair (X, Z). It therefore holds for the pair (X, Y). One may easily check that Assumption 1.1 does not hold for the embedding of $X = \dot{B}_{p,a}^s$ into $Y = L^{q,a}$, and conjecture that the profile

decomposition does not generally exist for such an embedding, by an argument analogous to the one in Remark 3.1.

- Embedding of Triebel-Lizorkin spaces into Triebel-Lizorkin or Besov spaces: for any $a, b > 0$, consider $X = \dot{F}^s_{p,a}$ and $Y = \dot{F}^t_{q,b}$ with $\frac{1}{p} - \frac{1}{q} = \frac{s-t}{d}$. It is known, see [29], that X is continuously embedded into

$$Z = \dot{B}^{\tilde{s}}_{\tilde{p},\tilde{p}}, \quad \tilde{s} < s \quad \text{and} \quad \frac{1}{p} - \frac{1}{\tilde{p}} = \frac{s - \tilde{s}}{d}.$$

If we assume $t < \tilde{s} < s$, we have the continuous embeddings (3.9). Moreover, we have already proved that Assumption 1.1 holds for the pair (Z, Y). It therefore holds for the pair (X, Y). The same type of reasoning allows one to prove Assumption 1.1 for the embedding of $X = \dot{F}^s_{p,a}$ into $Y = \dot{B}^t_{q,b}$ with $b > p$.

Remark 3.2. It is easily seen that if X and Y are a pair of spaces such that both Assumptions 1.1 and 1.2 hold for a certain wavelet basis (ψ_λ), then the corresponding vector fields spaces $(X)^d$ and $(Y)^d$ also satisfy the same assumptions for the vector-valued wavelet basis

$$\psi_{\lambda,i} := \psi_\lambda e_i, \quad \lambda \in \nabla, \quad i = 1, \ldots, d,$$

where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the canonical basis vector, and the wavelet coefficients d_λ are defined accordingly as vectors.

4. Stability of the Decomposition

Finally we want to show that the decomposition is stable in the sense that the sum of the $\|\phi^l\|_X$ raised to an appropriate power remains bounded. In our discussion, we distinguish between the cases where X is a Besov or Triebel-Lizorkin space. We first address the Besov case.

Theorem 4.1. *Assume that $X = \dot{B}^s_{p,a}$ with $1 \leq p < \infty$ and $1 \leq a \leq \infty$. We then have*

$$\|(\|\phi^l\|_X)_{l>0}\|_{\ell^\tau} \leq CK, \quad \tau := \max\{p, a\}, \tag{4.1}$$

where C is a constant that only depends on X and on the choice of the wavelet basis and where $K := \sup_{n \geq 0} \|u_n\|_X$.

Proof. Fix an arbitrary $L > 0$ and let M be such that $L \leq L(M)$ as in Step 4 of Sec. 2. For $l = 1, \ldots, L$, we recall the approximate profiles

$$\phi^{l,M}_{\lambda_l(n)} = \sum_{m \in E(l,M)} d_m \psi_{\lambda(m,n)}$$

and we also define

$$\phi^{l,M,n} := \sum_{m \in E(l,M)} d_{m,n} \psi_{\lambda(m,n)},$$

which are disjoint part of the wavelet expansion of u_n. More precisely, we have

$$u_n = \sum_{l=1}^{L} \phi^{l,M,n} + \sum_{m>M} d_{m,n} \psi_{\lambda(m,n)}.$$

We next claim that if $E_1, \ldots, E_L$ are disjoint finite sets in ∇, then for any coefficient sequence (d_λ), one has

$$\left(\sum_{l=1}^{L} \left\| \sum_{\lambda \in E_l} d_\lambda \psi_\lambda \right\|_X^\tau \right)^{1/\tau} \le C \left\| \sum_{l=1}^{L} \sum_{\lambda \in E_l} d_\lambda \psi_\lambda \right\|_X, \tag{4.2}$$

where C is a constant that only depends on X and on the choice of the wavelet basis, and with the standard modification of the sum to the power $1/\tau$ by a supremum on the left-hand side when $\tau = \infty$.

Before proving this claim, we first show that it leads to the conclusion of the proof. Indeed, for $l = 1, \ldots, L$, the functions $\phi^{l,M,n}$ are linear combinations of wavelets with indices in disjoint finite sets $E_1, \ldots, E_L$ (that vary with n), and therefore according to (4.2), when $\tau < \infty$,

$$\left(\sum_{l=1}^{L} \|\phi^{l,M,n}\|_X^\tau \right)^{1/\tau} \le C \left\| \sum_{l=1}^{L} \phi^{l,M,n} \right\|_X.$$

Using the unconditionality inequality (1.6), we thus find that for all $n > 0$

$$\left(\sum_{l=1}^{L} \|\phi^{l,M,n}\|_X^\tau \right)^{1/\tau} \le CK,$$

up to a multiplication of the constant C by D. Since $\|\phi_{\lambda_l(n)}^{l,M} - \phi^{l,M,n}\|_X \to 0$ as $n \to \infty$, it follows that for any $\varepsilon > 0$ we have

$$\left(\sum_{l=1}^{L} \|\phi_{\lambda_l(n)}^{l,M}\|_X^\tau \right)^{1/\tau} \le CK + \varepsilon,$$

for n sufficiently large. By the scaling invariance (1.3) we thus find that

$$\left(\sum_{l=1}^{L} \|\phi^{l,M}\|_X^\tau \right)^{1/\tau} \le CK.$$

Letting M go to $+\infty$, we obtain the same inequality for the exact profiles

$$\left(\sum_{l=1}^{L} \|\phi^l\|_X^\tau \right)^{1/\tau} \le CK$$

and we thus conclude that (4.1) holds, by letting $L \to +\infty$. The case $\tau = \infty$ is treated in an exact similar way, replacing the sum to the power $1/\tau$ by a supremum.

It remains to prove (4.2). We actually claim that this property holds with constant $C = 1$ if we take for $\| \cdot \|_X$ the equivalent norm given by (3.2). This is obvious when $p = a = \tau$ since this equivalent norm is then simply the ℓ^τ norm of the wavelet coefficients. When $p \neq a$, we distinguish between the cases $p < a$ and $p > a$. We denote by

$$E_{j,l} := \{\lambda \in E_l; \ |\lambda| = j\},$$

so that

$$E_l = \bigcup_{j \in \mathbb{Z}} E_{l,j}.$$

First consider the case $\tau = a > p$. We then have, when $a < \infty$,

$$\left\| \sum_{l=1}^{L} \sum_{\lambda \in E_l} d_\lambda \psi_\lambda \right\|_X^a = \sum_{j \in \mathbb{Z}} \left(\sum_{l=1}^{L} \sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{a/p}$$

$$\geq \sum_{j \in \mathbb{Z}} \sum_{l=1}^{L} \left(\sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{a/p}$$

$$= \sum_{l=1}^{L} \sum_{j \in \mathbb{Z}} \left(\sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{a/p}$$

$$= \sum_{l=1}^{L} \left\| \sum_{\lambda \in E_l} d_\lambda \psi_\lambda \right\|_X^a ,$$

where for the inequality we have simply used the fact that $a/p > 1$. Therefore (4.2) holds. When $a = \infty$, we obtain the same result by writing

$$\left\| \sum_{l=1}^{L} \sum_{\lambda \in E_l} d_\lambda \psi_\lambda \right\|_X = \sup_{j \in \mathbb{Z}} \left(\sum_{l=1}^{L} \sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{1/p}$$

$$\geq \sup_{j \in \mathbb{Z}} \left(\sup_{l \leq L} \sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{1/p}$$

$$= \sup_{l \leq L} \sup_{j \in \mathbb{Z}} \left(\sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{1/p}$$

$$= \sup_{l \leq L} \left\| \sum_{\lambda \in E_l} d_\lambda \psi_\lambda \right\|_X .$$

We next consider the case $\tau = p > a$. To treat this case where $p < \infty$ by hypothesis, we introduce the notation

$$b_{j,l} := \left(\sum_{\lambda \in E_{l,j}} |d_\lambda|^p \right)^{a/p},$$

and we remark that (4.2) is then equivalent to

$$\left(\sum_l \left(\sum_j b_{j,l} \right)^{p/a} \right)^{a/p} \leq \sum_j \left(\sum_l |b_{j,l}|^{p/a} \right)^{a/p},$$

which trivially holds by applying the triangle inequality in $\ell^{p/a}$. $\qquad \square$

Our last result addresses the Triebel-Lizorkin case.

Theorem 4.2. *Assume that* $X = \dot{F}^s_{p,a}$ *with* $1 \leq p < \infty$ *and* $1 \leq a \leq \infty$. *We then have*

$$\| (\| \phi^l \|_X)_{l>0} \|_{\ell^p} \leq CK, \tag{4.3}$$

where C *is a constant that only depends on* X *and on the choice of the wavelet basis and where* $K := \sup_{n \geq 0} \| u_n \|_X$.

Proof. We only give the proof in the case $a < \infty$, the case $a = \infty$ being treated by the same type of arguments up to notational changes. Fix an arbitrary $L > 0$ and let M be such that $L \leq L(M)$ as in Step 4 of Sec. 2. By the unconditionality (1.6) of the wavelet basis with respect to X, we first observe that

$$\left\| \sum_{m=1}^M d_{m,n} \psi_{\lambda(m,n)} \right\|_X \leq DK.$$

It follows that for any $\varepsilon > 0$, we have

$$\left\| \sum_{m=1}^M d_m \psi_{\lambda(m,n)} \right\|_X \leq DK + \varepsilon$$

for n sufficiently large. Recall that the sum inside the norm may be rewritten in terms of the approximate profiles:

$$\sum_{m=1}^M d_m \psi_{\lambda(m,n)} = \sum_{l=1}^L \sum_{m \in E(l,M)} d_m \psi_{\lambda(m,n)} = \sum_{l=1}^L \phi^{l,M}_{\lambda_l(n)}.$$

We associate to the approximate profile $\phi^{l,M}$ a piecewise constant function $\chi^{l,M}$ defined by

$$\chi^{l,M}_{\lambda_l(n)} := \sum_{m \in E(l,M)} |d_m \chi_{\lambda(m,n)}|^a,$$

where $\chi_\lambda = 2^{dj/p}\chi(2^j \cdot -k)$ with $\chi = \chi_{[0,1]^d}$ for $\lambda \sim (j,k)$. Thus according to the wavelet characterization (3.3) of Triebel-Lizorkin spaces, we have

$$c \int_{\mathbb{R}^d} |\chi^{l,M}_{\lambda_l(n)}(x)|^{p/a} dx \leq \|\phi^{l,M}_{\lambda_l(n)}\|^p_X \leq C \int_{\mathbb{R}^d} |\chi^{l,M}_{\lambda_l(n)}(x)|^{p/a} dx,$$

as well as

$$c \int_{\mathbb{R}^d} \left| \sum_{l=1}^L \chi^{l,M}_{\lambda_l(n)}(x) \right|^{p/a} dx \leq \left\| \sum_{m=1}^M d_{m,n} \psi_{\lambda(m,n)} \right\|^p_X \leq C \int_{\mathbb{R}^d} \left| \sum_{l=1}^L \chi^{l,M}_{\lambda_l(n)}(x) \right|^{p/a} dx,$$

where $0 < c \leq C$ only depends on the choice of the wavelet basis. In the case where $a \leq p$, we obviously have

$$\sum_{l=1}^L \int_{\mathbb{R}^d} |\chi^{l,M}_{\lambda_l(n)}(x)|^{p/a} dx \leq \int_{\mathbb{R}^d} \left| \sum_{l=1}^L \chi^{l,M}_{\lambda_l(n)}(x) \right|^{p/a} dx.$$

It therefore follows that

$$\sum_{l=1}^L \|\phi^{l,M}\|^p_X = \sum_{l=1}^L \|\phi^{l,M}_{\lambda_l(n)}\|^p_X \leq \frac{C}{c} \left\| \sum_{m=1}^M d_{m,n} \psi_{\lambda(m,n)} \right\|^p_X \leq \frac{C}{c}(DK + \varepsilon)^p.$$

Since this holds for any $\varepsilon > 0$, and $L > 0$ and M such that $L \leq L(M)$, we therefore obtain (4.3) by a limiting argument, up to renaming $(C/c)^{1/p}D$ into C.

In order to reach the same conclusion in the case $p < a$, we need to exploit the "asymptotic orthogonality" of the scales $\lambda_l(n)$ as expressed by (1.13) in the statement of Theorem 1.1. For this purpose, let us define

$$\Omega_n := \mathrm{Supp}\left(\sum_{l=1}^L \chi^{l,M}_{\lambda_l(n)} \right) = \bigcup_{l=1}^L \mathrm{Supp}(\chi^{l,M}_{\lambda_l(n)}).$$

For any $x \in \Omega_n$, we denote by l^* the number in $\{1, \dots, L\}$ such that

$$\chi^{l^*,M}_{\lambda_{l^*}(n)}(x) = \max_{l=1,\dots,L} \chi^{l,M}_{\lambda_l(n)}(x).$$

Note that l^* depends on both x and n. We claim that a consequence of (1.13) is that the function $\chi^{l^*,M}_{\lambda_{l^*}(n)}$ tends to dominate all other $\chi^{l,M}_{\lambda_l(n)}$ at the point x as $n \to +\infty$ in the following uniform sense:

$$\lim_{n \to +\infty} \min_{x \in \Omega_n} \frac{\chi^{l^*,M}_{\lambda_{l^*}(n)}(x)}{\sum_{l=1}^L \chi^{l,M}_{\lambda_l(n)}(x) - \chi^{l^*,M}_{\lambda_{l^*}(n)}(x)} = +\infty. \tag{4.4}$$

Before proving this claim, let us show how it leads us to the conclusion of the theorem. We observe that (4.4) also means that $|\chi^{l^*,M}_{\lambda_{l*}(n)}|^{p/a}$ tends to dominate all other $|\chi^{l,M}_{\lambda_l(n)}|^{p/a}$ at the point x as $n \to +\infty$. Therefore, for any $\varepsilon > 0$, we have for n large enough

$$\sum_{l=1}^{L} |\chi^{l,M}_{\lambda_l(n)}(x)|^{p/a} \leq (1+\varepsilon)|\chi^{l^*,M}_{\lambda_{l*}(n)}(x)|^{p/a} \leq (1+\varepsilon) \left| \sum_{l=1}^{L} \chi^{l,M}_{\lambda_l(n)}(x) \right|^{p/a}$$

for all $x \in \Omega_n$, and thus

$$\sum_{l=1}^{L} \int_{\mathbb{R}^d} |\chi^{l,M}_{\lambda_l(n)}(x)|^{p/a} dx \leq (1+\varepsilon) \int_{\mathbb{R}^d} \left| \sum_{l=1}^{L} \chi^{l,M}_{\lambda_l(n)}(x) \right|^{p/a} dx.$$

We may then conclude the proof as in the case $a \leq p$.

It remains to prove (4.4). Our first observation is that the asymptotic orthogonality of the scales $\lambda_l(n)$ expressed by (1.13), shows that for a given x, the profile scales $|\lambda_l(n)|$ for those $l \in \{1, \ldots, L\}$ such that $x \in \text{Supp}(\chi^{l,M}_{\lambda_l(n)})$ tend to get far apart as n grows. Indeed, these $\lambda_l(n)$'s do not get far apart in space since the supports of $\chi^{l,M}_{\lambda_l(n)}$ all contain the same point x.

We introduce l_* the number that maximizes $|\lambda_l(n)|$ among all those $l \in \{1, \ldots, L\}$ such that $x \in \text{Supp}(\chi^{l,M}_{\lambda_l(n)})$. Similar to l^*, the number l_* depends both of x and n. From the previous observation, we know that for any arbitrarily large $B > 0$, there exists n_0 such that for all $n \geq n_0$, we have

$$|\lambda_{l_*}(n)| \geq |\lambda_l(n)| + B, \tag{4.5}$$

for all $l \in \{1, \ldots, L\}$ such that $x \in \text{Supp}(\chi^{l,M}_{\lambda_l(n)})$ and $l \neq l_*$. Moreover, we may choose this n_0 independent of the selected point x for the same $B > 0$.

We claim that as n grows $\chi^{l_*,M}_{\lambda_{l_*}(n)}$ tends to dominate all other $\chi^{l,M}_{\lambda_l(n)}$ at the point x as $n \to +\infty$, in the sense that

$$\lim_{n \to +\infty} \min_{x \in \Omega_n} \frac{\chi^{l_*,M}_{\lambda_{l_*}(n)}(x)}{\sum_{l=1}^{L} \chi^{l,M}_{\lambda_l(n)}(x) - \chi^{l_*,M}_{\lambda_{l_*}(n)}(x)} = +\infty. \tag{4.6}$$

This clearly implies (4.4) (and shows that $l^* = l_*$ for n large enough).

In order to prove (4.6), we observe that if $x \in \text{Supp}(\chi^{l,M}_{\lambda_l(n)})$ for some $l \in \{1, \ldots, L\}$, we may then frame $\chi^{l,M}_{\lambda_l(n)}(x)$ according to

$$c2^{\frac{a dj_l(n)}{p}} \leq \chi^{l,M}_{\lambda_l(n)}(x) \leq C2^{\frac{a dj_l(n)}{p}}$$

where

$$j_l(n) := \min_{m \in E(l,M)} |\lambda(m,n)| \quad \text{and} \quad J_l(n) := \max_{m \in E(l,M)} |\lambda(m,n)|,$$

and where

$$c := |d_M|^a \quad \text{and} \quad C := \sum_{m=1}^{M} |d_m|^a.$$

The constants c and C of course depend on L and M which are fixed at that stage. Note that from the construction of the profile there exists $A > 0$ (that also depends on L and M) such that for all $l \in \{1, \dots, L\}$

$$|\lambda_l(n)| - A \le j_l(n) \le J_l(n) \le |\lambda_l(n)| + A,$$

and therefore, up to a modification in the constants c and C we may write

$$c2^{\frac{ad|\lambda_l(n)|}{p}} \le \chi_{\lambda_l(n)}^{l,M}(x) \le C2^{\frac{ad|\lambda_l(n)|}{p}}.$$

Combining this observation with (4.5), we easily obtain (4.6).   □

Acknowledgment

References

1. R. Adams, *Sobolev Spaces*, Pure and Applied Mathematics, Vol. 65 (Academic Press, 1975).
2. H. Bahouri and P. Gérard, High frequency approximation of solutions to critical nonlinear wave equations, *Amer. J. Math.* **121** (1999) 131–175.
3. H. Bahouri, M. Majdoub and N. Masmoudi, On the lack of compactness in the 2D critical Sobolev embedding, *J. Funct. Anal.* **260** (2011) 208–252.
4. J. Ben Ameur, Description du défaut de compacité de l'injection de Sobolev sur le groupe de Heisenberg, *Bull. Soc. Math. Belgique* **15** (2008) 599–624.
5. H. Brezis and J. M. Coron, Convergence of solutions of H-Systems or how to blow bubbles, *Arch. Rational Mech. Anal.* **89** (1985) 21–86.
6. A. Cohen, *Numerical Analysis of Wavelet Methods* (Elsevier, 2003).
7. I. Daubechies, *Ten Lectures on Wavelets* (SIAM, 1992).
8. R. DeVore, Nonlinear Approximation, *Acta Nume.* **7** (1998) 51–150.
9. R. DeVore, B. Jawerth and V. Popov, Compression of wavelet decompositions, *Amer. J. Math.* **114** (1992) 737–785.
10. I. Gallagher and P. Gérard, Profile decomposition for the wave equation outside convex obstacles, *J. Math. Pures Appl.* **80** (2001) 1–49.
11. I. Gallagher, G. S. Koch and F. Planchon, A profile decomposition approach to the $L_t^\infty(L_x^3)$ Navier–Stokes regularity criterion, arXiv:1012.0145.
12. I. Gallagher, Profile decomposition for solutions of the Navier–Stokes equations, *Bull. Soc. Math. France* **129** (2001) 285–316.
13. P. Gérard, Description du défaut de compacité de l'injection de Sobolev, *ESAIM Control Optim. Calc. Var.* **3** (1998) 213–233.

14. S. Ibrahim, Comparaison des ondes linéaires et non linéaires à coefficients variables, *Bull. Soc. Math. Belgique* **10** (2003) 299–312.
15. S. Jaffard, Analysis of the lack of compactness in the critical Sobolev embeddings, *J. Funct. Anal.* **161** (1999) 384–396.
16. C. E. Kenig and G. S. Koch, An alternative approach to the Navier–Stokes equations in critical spaces, *Ann. l'Inst. Henri Poincaré*, *Anal. Non Linéaire*, DOI:10.1016/j.anihpc.2010.10.004.
17. C. E. Kenig and F. Merle, Global well-posedness, scattering and blow-up for the energy critical focusing nonlinear wave equation, *Acta Math.* **201** (2008) 147–212.
18. S. Keraani, On the defect of compactness for the Strichartz estimates of the Schrödinger equation, *J. Differential Equations* **175** (2001) 353–392.
19. G. Kyriasis, Nonlinear approximation and interpolation spaces, *J. Approx. Theory* **113** (2001) 110–126.
20. C. Laurent, On stabilization and control for the critical Klein-Gordon equation on a 3-D compact manifold, to appear in *J. Funct. Anal.*
21. P.-L. Lions, The concentration-compactness principle in the calculus of variations. The limit case. I, *Rev. Mat. Iber.* **1** (1985) 145–201.
22. P.-L. Lions, The concentration-compactness principle in the calculus of variations. The limit case. II, *Rev. Mat. Iber.* **1** (1985) 45–121.
23. M. Majdoub, Qualitative study of the critical wave equation with a subcritical perturbation, *J. Math. Anal. Appl.* **301** (2005) 354–365.
24. Y. Meyer, *Ondelettes et Opérateurs* (Hermann, 1990).
25. I. Schindler and K. Tintarev, An abstract version of the concentration compactness principle, *Rev. Math. Complut.* **15** (2002) 417–436.
26. S. Solimini, A note on compactness-type properties with respect to Lorentz norms of bounded subset of a Sobolev space, *Ann. l'IHP Anal. Non linéaire* **12** (1995) 319–337.
27. M. Struwe, A global compactness result for boundary value problems involving limiting nonlinearities, *Math. Z.* **187** (1984) 511–517.
28. T. Tao, An inverse theorem for the bilinear L^2 Strichartz estimate for the wave equation, arXiv:0904.2880.
29. H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, 2nd edn. (Johann Ambrosius Barth, 1995).