Appl. Comput. Harmon. Anal. 36 (2014) 239–255

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

# Hyperbolic wavelet thresholding methods and the curse of dimensionality through the maxiset approach

F. Autin [a], G. Claeskens [b], J.-M. Freyermuth [b,*]

[a] Aix-Marseille Université – L.A.T.P., 39, rue F. Joliot Curie, 13453 Marseille Cedex 13, France
[b] KU Leuven, ORSTAT and Leuven Statistics Research centre, Naamsestraat 69, 3000 Leuven, Belgium

ARTICLE INFO

*Article history:*
Received 4 September 2012
Received in revised form 27 March 2013
Accepted 29 April 2013
Available online 9 May 2013
Communicated by Dominique Picard

*Keywords:*
Anisotropy
Hyperbolic wavelet basis
Information pooling
Maxiset
Wavelet thresholding
Atomic dimension

ABSTRACT

In this paper we compute the maxisets of some denoising methods (estimators) for multidimensional signals based on thresholding coefficients in hyperbolic wavelet bases. That is, we determine the largest functional space over which the risk of these estimators converges at a chosen rate. In the unidimensional setting, refining the choice of the coefficients that are subject to thresholding by pooling information from geometric structures in the coefficient domain (e.g., vertical blocks) is known to provide 'large maxisets'. In the multidimensional setting, the situation is less straightforward. In a sense these estimators are much more exposed to the curse of dimensionality. However we identify cases where information pooling has a clear benefit. In particular, we identify some general structural constraints that can be related to compound functional models and to a minimal level of anisotropy.



## 1. Introduction

Multivariate wavelet bases for $L_2([0,1)^d)$ can be constructed by taking the tensor product of univariate wavelet functions. The formed wavelet bases, so-called hyperbolic, have been proved to be of particular interest for many applications and are studied since a long time for their approximation theoretic properties [18,28,27], in data compression [47,23,12,16], in statistics [36] and more recently in very active research areas such as compressive sensing [20] and multifractal analysis [1]. In this paper, we give new theoretical results about their abilities for denoising multidimensional signals by different thresholding methods of the noisy hyperbolic wavelet coefficients.

Wavelet thresholding for nonparametric function estimation has been widely studied from both theoretical and practical points of view. Most of the results for multidimensional function estimation follow as 'direct' extensions of unidimensional results whenever considering the simple to handle multidimensional wavelet bases generated by functions that are product of unidimensional scaling/wavelet functions with the same parameter of scale (here so-called isotropic wavelet bases). These methods have been proved useful for example in [45] for estimating functions with isotropic regularity. Nevertheless, when going to the estimation of multivariate objects, a key concept to integrate is the one of *anisotropy*. That is, to allow these objects to have different smoothness properties along the different coordinate axes. This concept is also essential when the different directions have completely different meaning as in [37] for time-varying spectral density estimation. By comparing the isotropic and hyperbolic wavelet bases, Neumann and von Sachs [37] first show that optimal estimation of two-dimensional functions in anisotropic Sobolev classes can be achieved using the hyperbolic wavelet basis (sometimes equally referred to as anisotropic wavelets). Neumann [36] continues that study by proving that thresholding in these

* Corresponding author.
*E-mail addresses:* autin@cmi.univ-mrs.fr (F. Autin), Gerda.Claeskens@kuleuven.be (G. Claeskens), Jean-Marc.Freyermuth@kuleuven.be (J.-M. Freyermuth).

1063-5203/$ – see front matter 
http://dx.doi.org/10.1016/j.acha.2013.04.003

hyperbolic bases allows for adaptive estimation to both spatially inhomogeneous smoothness and anisotropy captured by anisotropic Besov scales.

In the field of anisotropic function estimation, the seminal work of Donoho [19], who constructs a dyadic classification and regression tree estimator and proves its quasi-minimax optimality for estimating anisotropic function classes, had a considerable impact as evidenced in the recent papers of Klemela [35] and Akakpo [2]. The former generalizes the approach of [19] for estimating multivariate densities with histograms and the latter uses adapted partitioning with piecewise polynomial fits. Other approaches rely on kernel methods with adaptive bandwidths [30,10]. Goldenshluger and Lepski [24] notably discuss the application of a pointwise adaptive procedure based on the selection from a large collection of kernels to estimate a function that belongs to the union of anisotropic Hölder classes. More sophisticated frameworks than the Gaussian white noise model are also considered. For example, Comte and Lacour [15] study adaptive anisotropic kernel estimators in a multidimensional convolution model and Ingster and Stepanova [29] consider the problem of detecting signals with particular attention being paid to the case of infinite dimension.

In this paper, we study several hyperbolic wavelet thresholding estimators. In contrast to Neumann [36] who gives some minimax properties under the quadratic risk, we determine, under a more general loss function, the maximal functional space (maxiset) for which the risk of these estimators reaches a given rate of convergence. We are particularly interested in thresholding estimators which pool information from geometric structures in the coefficient domain. In the unidimensional setting such estimators have been proven powerful from both theoretical and practical points of view (see among numerous references [26,11,3]). In this paper, we restrict our study to hierarchically structured wavelet estimators that impose to the coefficients that have survived the threshold to be arranged over a hierarchical structure [21,22]. More particularly, we study a generalization of the *hard tree* estimator [4], namely the *hyperbolic hard tree* estimator. The hard tree estimator has been proven to be the best element of a large family of vertical block thresholding estimators [5] and to outperform the hard thresholding estimator under the $L_2$-risk [4]. In a multidimensional context, we show that these estimators have a complex behavior according to the risk function, the dimension and the rate. Despite that the hyperbolic hard tree estimator is much more exposed to the curse of dimensionality than the hyperbolic hard thresholding estimator, it still outperforms the latter in several identified cases. Within these cases, we identify a general structural constraint that is interpreted as a minimal level of anisotropy and that is related to compound functional models.

The paper is structured as follows: we describe in Section 2 the construction of the $d$-dimensional hyperbolic wavelet basis and the concept of heredity. Section 3 introduces the maxiset approach and Section 4 presents the estimators that will be studied. The main maxiset results are given in Section 5. In Section 6 we interpret our maxiset results and discuss the impact of the curse of dimensionality on information pooling. A conclusion is given in Section 7. The proofs are deferred to Appendix A.

## 2. Multidimensional wavelet bases

There are several ways to construct wavelet bases of $L_2([0,1)^d)$ from a unidimensional wavelet basis of $L_2([0,1))$ which is built from the dilations and translations of a scaling function, say $\phi$, and a wavelet, say $\psi$. To deal with anisotropic functions, Triebel [46] introduced a specific anisotropy adapted wavelet basis (a specific multidimensional wavelet basis associated with an anisotropic parameter that matches the anisotropic smoothness of the signal). Those anisotropic wavelet bases are useful tools in functional analysis and in theory of approximation since they give a benchmark (known level of anisotropy); but they are obviously not well suited to design denoising methods that are required to be adaptive to a signal with unknown smoothness and anisotropy. Hyperbolic wavelet bases have been proved useful for approximating functions from anisotropic smoothness classes (see [18,44,28]) and are known to be more appropriate than the $d$-dimensional isotropic wavelet basis (see among others [37,36,44,28]). This mainly motivated our choice to consider them.

### 2.1. d-dimensional hyperbolic wavelet basis

We detail the construction of the hyperbolic wavelet basis from the following unidimensional periodized wavelet basis with $V$ (for some $V \geqslant 1$) vanishing moment(s)

$$\mathcal{B}_1 = \{\phi(.), \psi_{j,k}(.) = 2^{j/2}\psi(2^j . - k);\ j \in \mathbb{N},\ k \in \{0, \ldots, 2^j - 1\}\}.$$

We construct the $d$-dimensional hyperbolic wavelet basis, denoted as $\mathcal{B}_d$, as follows:

$$\mathcal{B}_d = \{\phi_{\underline{0},\underline{0}}, \psi^{\underline{i}}_{\underline{j},\underline{k}};\ \underline{i} \in \{0,1\}^d \setminus \underline{0},\ \underline{j} \in \mathbb{J}^{\underline{i}},\ \underline{k} \in \mathbb{K}_{\underline{j}}\}$$

where $\underline{0} = (0, \ldots, 0)$ and

$$\phi_{\underline{0},\underline{0}}(.) = \phi(.) \times \cdots \times \phi(.), \quad \text{and} \quad \psi^{\underline{i}}_{\underline{j},\underline{k}}(.) = \psi^{i_1}_{j_1,k_1}(.) \times \cdots \times \psi^{i_d}_{j_d,k_d}(.),$$

with the following notations:

$$\psi^{i_u}_{j_u,k_u}(.) := \begin{cases} 2^{j_u/2}\phi(2^{j_u} . - k_u) & \text{if } i_u = 0, \\ 2^{j_u/2}\psi(2^{j_u} . - k_u) & \text{if } i_u = 1 \end{cases} \tag{1}$$