



HAL
open science

Nonlinear approximation spaces for inverse problems

Albert Cohen, Matthieu Dolbeault, Olga Mula, Agustin Somacal

► **To cite this version:**

Albert Cohen, Matthieu Dolbeault, Olga Mula, Agustin Somacal. Nonlinear approximation spaces for inverse problems. 2022. ⟨hal-03780952v2⟩

HAL Id: hal-03780952

<https://hal.science/hal-03780952v2>

Preprint submitted on 6 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Nonlinear approximation spaces for inverse problems

Albert Cohen*, Matthieu Dolbeault*, Olga Mula[†], Agustin Somacal*

October 6, 2022

Abstract

This paper is concerned with the ubiquitous inverse problem of recovering an unknown function u from finitely many measurements possibly affected by noise. In recent years, inversion methods based on *linear* approximation spaces were introduced in [7, 27] with certified recovery bounds. It is however known that linear spaces become ineffective for approximating simple and relevant families of functions, such as piecewise smooth functions that typically occur in hyperbolic PDEs (shocks) or images (edges). For such families, *nonlinear* spaces [18] are known to significantly improve the approximation performance. The first contribution of this paper is to provide with certified recovery bounds for inversion procedures based on nonlinear approximation spaces. The second contribution is the application of this framework to the recovery of general bidimensional shapes from cell-average data. We also discuss how the application of our results to n -term approximation relates to classical results in compressed sensing.

1 Introduction

1.1 The recovery problem

In this paper, we treat the following state estimation problem in a general Banach space V . We want to recover an approximation to an unknown function $u \in V$ from data given by m observations

$$z_i := \ell_i(u) + \eta_i, \quad i = 1, \dots, m, \quad (1.1)$$

where $\ell_i : V \mapsto \mathbb{R}$ are known measurement functionals, and η_i is additive noise. The functionals ℓ_i often correspond to the response of a physical measurement device but they can have a different interpretation depending on the application. Their behavior can be linear (in which case the ℓ_i are linear functionals from V' , the dual of V) or nonlinear. This type of recovery problem is clearly ill-posed when the dimension of V exceeds m . It arises ubiquitously in sampling and inverse problem applications where V is infinite dimensional (to name a few, see [1, 3, 22, 26]).

One natural strategy to address this difficulty is to search for a recovery of u by an element of a low-dimensional reconstruction space $V_n \subset V$. The space V_n could be either an n -dimensional linear subspace, or more generally a nonlinear approximation space parametrized by n degrees of freedom, with $n \leq m$.

*Laboratoire Jacques-Louis Lions, Sorbonne Université, 4 place Jussieu, 75005 Paris, France (albert.cohen@sorbonne-universite.fr, matthieu.dolbeault@sorbonne-universite.fr and agustin.somacal@sorbonne-universite.fr)

[†]Technical University of Eindhoven, Netherlands (o.mula@tue.nl)

In order to obtain quantitative results for such recovery procedures, it is necessary to possess additional information about u , usually as an assumption that u belongs to a certain model class \mathcal{K} contained in V . The approximation space V_n is chosen in order to collectively approximate the elements of \mathcal{K} as well as possible, in the sense that

$$\text{dist}(\mathcal{K}, V_n)_V := \max_{u \in \mathcal{K}} \min_{v \in V_n} \|u - v\|_V$$

is as small as possible for moderate values of n .

Numerous theoretical results and numerical algorithms have been proposed in several fields to study and solve the above recovery problem (below we recall some relevant results). However, to the best of our knowledge, they all involve at least one or several of the following assumptions:

- The ℓ_i are *linear* functionals,
- V_n is a *linear* (or affine) subspace of V ,
- V is a *Hilbert* space,
- The model class \mathcal{K} is a *ball in a smoothness space*, e.g., a unit ball in Lipschitz, Sobolev, or Besov spaces. Results involving this type of model classes have been intensively studied in the field of optimal recovery (see [9, 30, 32]).

The goal of this paper is to develop and analyze inversion procedures that do not require any of the above assumptions. Our analysis and numerical algorithms can thus be applied to virtually any recovery problem. The starting point of our development is based on algorithms introduced for inverse state estimation using reduced order models of parametrized Partial Differential Equations (PDEs). We next recall the specific framework. The presentation will also serve to explain more in depth the motivations leading to propose the present generalization.

1.2 State estimation with reduced models for parametrized PDE's

A relevant scenario in inverse state estimation is when the model class \mathcal{K} is given by the set of solutions to some parameter-dependent PDE of the general form

$$\mathcal{P}(u, y) = 0, \tag{1.2}$$

where \mathcal{P} is a differential operator, y a vector of parameters ranging in some domain Y in \mathbb{R}^d , and u is the solution. If well-posedness holds in some Banach space V for each $y \in Y$, we denote by $u(y) \in V$ the corresponding solution for the given parameter value y and by

$$\mathcal{M} := \{u(y) : y \in Y\},$$

the *solution manifold*.

In inverse state estimation, we take $\mathcal{K} = \mathcal{M}$ for the model class so the unknown u to recover belongs to \mathcal{M} . However, the parameter y that satisfies $u = u(y)$ is unknown, so we cannot solve the forward problem (1.2) to approximate u . Instead, we must approximate u from the partial observational data (1.1), and the knowledge of the model class $\mathcal{K} = \mathcal{M}$.

For the manifold \mathcal{M} , efficient approximation spaces V_n are usually obtained by reduced modelling techniques. In their most simple format, reduced models consist into linear spaces $(V_n)_{n \geq 0}$

with $\dim(V_n) = n$. The ideal benchmark in this linear approximation setting is provided by the Kolmogorov n -width

$$d_n(\mathcal{M})_V := \inf_{\dim(V_n) \leq n} \text{dist}(\mathcal{M}, V_n)_V,$$

which describes the optimal approximation performance achievable by an n -dimensional space over the set \mathcal{M} .

Apart from very simplified cases, the space V_n achieving the above infimum is usually out of reach. Practical model reduction techniques such as polynomial approximation in the parametrized domain [15, 16, 39] or reduced bases [19, 25, 28, 37, 41] construct spaces V_n that are “suboptimal yet good”. In particular, the reduced basis method, which generates V_n by a specific selection of particular solution instances $u^1, \dots, u^n \in \mathcal{M}$, has been proved to have approximation error $\text{dist}(\mathcal{M}, V_n)_V$ that decays with the same polynomial or exponential rates as $d_n(\mathcal{M})_V$, and in that sense are close to optimal [17].

1.3 The PBDW method

We take the *Parametrized Background Data Weak* (PBDW) method as a starting point for our analysis. The PBDW method, first introduced in [27], as well as several extensions, has been the object of a series of works [7, 8, 13, 14] on its optimality properties as a recovery algorithm. It has also been used for different practical applications, see [3, 22, 24]. We refer to [31] for an overview of the state of the art on this approach, and its connections with different fields. For our current purposes, it will suffice to recall the first version of the algorithm, which is the goal of this section.

The PBDW method uses a linear approximation space V_n of dimension $n \leq m$. Usually this space is a reduced model in applications. It is assumed that the ℓ_i are continuous linear functionals, that is $\ell_i \in V'$, and that V is a Hilbert space. Then, introducing the Riesz representers $\omega_i \in V$ such that $\ell_i(v) = \langle \omega_i, v \rangle_V$, the data of the noise-free observation

$$\ell(u) := (\ell_1(u), \dots, \ell_m(u)),$$

is equivalent to that of the orthogonal projection $w = P_W u$ on the *Riesz measurement space*

$$W := \text{span}\{\omega_1, \dots, \omega_m\}.$$

Assuming linear independence of the ℓ_i , this space has dimension m . A critical quantity is the number

$$\mu = \mu(V_n, W) := \max_{v \in V_n} \frac{\|v\|_V}{\|P_W v\|_V}, \quad (1.3)$$

that describes the “stability” of the description of an element of V_n by its projection onto W , and may be thought of as the inverse cosine of the angle between W and V_n . In particular, this quantity is finite only when $n \leq m$. It can be explicitly computed as the inverse of the smallest singular value of a cross-grammian matrix between orthonormal bases of V_n and W (see [7, 31]).

The PBDW method consists in solving the minimization problem

$$\min_{v \in V_w} \min_{\tilde{v} \in V_n} \|v - \tilde{v}\|_V,$$

where $V_w := w + W^\perp$ is the set of all states v such that $P_W v = w$. We denote by $(u^*, \tilde{u}) \in V_w \times V_n$ the minimizing pair, which is unique when $\mu < \infty$, and can be computed by solving an $n \times n$ linear

system. The function \tilde{u} may be seen as a particular best-fit estimator of u on V_n , since it is also defined by

$$\tilde{u} := \operatorname{argmin}\{\|P_W v - w\|_V : v \in V_n\}.$$

The function u^* can be derived from \tilde{u} by the correction procedure

$$u^* := \tilde{u} + (w - P_W \tilde{u}),$$

which shows that $u^* \in V_n + W$. It may be thought of as a generalized interpolation estimator, since it agrees with the observed data ($P_W u^* = P_W u$). In the case of noise-free data, it is proved in [7, 27] that these estimators satisfy the recovery bounds

$$\|u - \tilde{u}\|_V \leq \mu \min_{v \in V_n} \|u - v\|_V \quad \text{and} \quad \|u - u^*\|_V \leq \mu \min_{v \in V_n \oplus (V_n^\perp \cap W)} \|u - v\|_V.$$

These bounds reflect a typical trade-off in the choice of the reduced basis space, since making n larger has both effect of decreasing the approximation error $\min_{v \in V_n} \|u - v\|_V$ and increasing the stability constant $\mu = \mu(V_n, W)$.

When the PBDW method is applied to noisy data, amounting in observing a perturbed version \bar{w} of $w = P_W u$, the recovery bounds remain valid up to the additional term $\mu \|w - \bar{w}\|_V$. In summary, one has for both estimators

$$\max\{\|u - \tilde{u}\|_V, \|u - u^*\|_V\} \leq \mu(e_n(u) + \kappa), \quad (1.4)$$

where

$$e_n(u) := \min_{v \in V_n} \|u - v\|_V$$

is the reduced model approximation error and $\kappa := \|w - \bar{w}\|_V$ is the noise error measured in the space W . Note that since the additive perturbations η_i are applied to the data $\ell_i(u)$, a natural model for the measurement noise is to assume a bound of the form

$$\|\eta\|_p \leq \varepsilon, \quad (1.5)$$

for the vector $\eta = (\eta_1, \dots, \eta_m)$, typically in the max norm $p = \infty$ or euclidean norm $p = 2$. Therefore, one has $\kappa \leq \beta\varepsilon$, where

$$\beta := \max_{v \in W} \frac{\|v\|_V}{\|\ell(v)\|_p},$$

resulting in a bound of the form $\mu e_n(u) + \mu\beta\varepsilon$ for both estimators.

1.4 Towards nonlinear approximation spaces

The simplicity of the PBDW method and its variants comes together with a fundamental limitation on its performance: it is by essence a linear reconstruction method with recovery bounds tied to the approximation error $e_n(u)$. When the only prior information is that the unknown function u belongs to a class \mathcal{K} , with $\mathcal{K} = \mathcal{M}$ the solution manifold in the case of parametric PDEs, its best performance over \mathcal{K} is thus limited by the n -width $d_n(\mathcal{K})_V$ and in turn by $d_m(\mathcal{K})_V$ since $n \leq m$.

In several simple yet relevant settings, it is known that n -widths have poor decay with n . One instance is when the class \mathcal{K} contains piecewise smooth states, with a state-dependent location of jump discontinuities. As an elementary example, one can easily check that if $V = L^2([0, 1])$ and \mathcal{K}

is the set all indicator functions $u = \chi_{[a,b]}$ with $a, b \in [0, 1]$, one has $d_n(\mathcal{K})_V \sim n^{-1/2}$. This decay is of course even slower for more general classes of piecewise smooth function in higher dimension, see in particular [5, Chapter 3, equation (3.76)]. Such functions are typical in parametrized hyperbolic PDEs, due to the presence of shocks with positions that differ when parameters entering the velocity vary. We refer to [4, 8, 20, 23, 33, 40] for other examples of parametric PDEs whose solution manifold has slow Kolmogorov n -width decay.

For such classes of functions, nonlinear approximation methods are well known to perform significantly better than their linear counterparts. Typical representatives of such methods include approximation by rational fractions, free knot splines or adaptive finite elements, best n -term approximation in a basis or dictionary, neural network or various tensor formats. In these instances the space V_n still depends on n or $\mathcal{O}(n)$ parameters but is not anymore a linear space. We refer to [18] for a general introduction on the topic of nonlinear approximation.

1.5 Objective and outline

The objective of this paper is to study the natural extensions of the PBDW method to such nonlinear approximation spaces and identify the basic structural properties that lead to near optimal recovery estimates similar to (1.4).

We begin in §2 by considering the most general setting where V is a Banach space, V_n a nonlinear approximation family, and the ℓ_i are functionals defined on V that are not necessarily linear, but Lipschitz continuous, that is

$$\|\ell(v) - \ell(\tilde{v})\|_Z \leq \alpha_Z \|v - \tilde{v}\|_V, \quad v, \tilde{v} \in V. \quad (1.6)$$

Here $\|\cdot\|_Z$ can be any given norm defined over \mathbb{R}^m with the constant α_Z depending on this choice of norm. In this framework, we discuss the best-fit estimation procedure that consists in minimizing the distance to the observed data in a given norm $\|\cdot\|_Z$.

Our main structural assumption on V_n is the following *inverse stability property*: the reduced model is stable with respect to the measurement functionals if there exists a finite constant μ_Z such that

$$\|v - \tilde{v}\|_V \leq \mu_Z \|\ell(v) - \ell(\tilde{v})\|_Z, \quad v, \tilde{v} \in V_n. \quad (1.7)$$

The stability constant μ_Z depends on the Z norm and plays a role similar to that of μ in the linear case. In particular, we show that this constant is finite only if $n \leq m$. The resulting estimator \tilde{u} is then proved to satisfy a general recovery bound of the form

$$\|u - \tilde{u}\|_V \leq C_1 e_n(u) + C_2 \|\eta\|_p,$$

where $e_n(u) := \min_{v \in V_n} \|u - v\|_V$ is the nonlinear reduced model approximation error, $\|\eta\|_p$ the level of measurement noise in ℓ^p norm, and the constants C_1 and C_2 depend on α_Z and μ_Z .

In §3, we consider the more particular setting where the ℓ_i are linear functionals. Then, we show that constants C_1 and C_2 are each minimized by a different choice of norm $\|\cdot\|_Z$, resulting in two different best fit estimators \tilde{u} , as already observed in [6] in the case of linear reduced models. This particular setting also allows us to introduce a generalized interpolation estimator u^* and establish similar recovery estimates for $\|u - u^*\|_V$.

We next apply our framework to the inverse problem that consists in recovering a general shape Ω , identified to its characteristic function χ_Ω , based on cell average data

$$a_T(\Omega) := \frac{1}{|T|} \int_T \chi_\Omega, \quad T \in \mathcal{T},$$

where \mathcal{T} is a fixed cartesian mesh. One motivation for this problem is the design of finite volume schemes for the computation of solutions to transport PDEs on such meshes.

We first discuss in §4 the best estimation rate in terms of the mesh size h that can be achieved by standard linear reconstructions, and which is essentially that of piecewise constant approximations, that is $\mathcal{O}(h^{1/q})$ regardless of the smoothness of the boundary $\partial\Omega$. This intrinsic limitation is due to the presence of the jump discontinuity that is not well resolved by the mesh.

We then discuss in §5 a local recovery strategy based on a nonlinear approximation space V_n that consists of characteristic functions of half-planes which can fit the boundary of Ω at a subcell resolution level, as already proposed in [2, 34, 35, 36]. One main result, whose proof is given in an appendix, is that this approximation space is stable in the sense of (1.7) with respect to cell average measurements on a stencil of 3×3 squares. In turn, if Ω has a C^2 boundary, the recovered shape $\tilde{\Omega}$ is proved to satisfy an estimate of the form

$$\|\chi_\Omega - \chi_{\tilde{\Omega}}\|_{L^q} \leq Ch^{2/q},$$

where h is the mesh size, which cannot be achieved by any linear reconstruction. This paves the way to higher order reconstruction methods for smoother boundaries by using local nonlinear approximation spaces with curved boundaries and larger stencils.

Finally, we discuss in §6 the application of our results to the recovery of large vectors of size N from $m < N$ linear measurements, up to the error of best n -term approximation. This problem is well-known in compressed sensing [10, 21], and was in particular studied in [12] which discusses the importance of the recovery norm $\|\cdot\|_V$ to understand if near-optimal recovery bounds can be achieved with m not much larger than n . We show that the structural assumptions identified in our general setting are naturally related to the so-called *null space property* introduced in [12].

2 Nonlinear reduction of inverse problems

2.1 A general framework

In full generality we are interested in recovering functions u in a general Banach space V with norm $\|\cdot\|_V$, from the measurement vector $z = (z_1, \dots, z_m) \in \mathbb{R}^m$ given by (1.1). A recovery (or inversion) map

$$z \rightarrow R(z),$$

takes this vector to an approximation $R(z)$ of u . We are interested in controlling the recovery error $\|u - R(z)\|_V$.

To build the recovery map R , we use a nonlinear approximation space of dimension n is a family of functions that can be described by n parameters. Loosely speaking, this means that there exists a set $S \subset \mathbb{R}^n$ and a continuous map $\varphi : S \rightarrow V$ such that

$$V_n := \{\varphi(x) : x \in S\}.$$

Note that this definition covers the case of an n dimensional linear subspace since we can choose $S = \mathbb{R}^n$ and φ a linear map.

Our main assumptions are the Lipschitz stability of the functionals ℓ_i over the whole space V and their inverse Lipschitz stability over the nonlinear approximation space V_n , expressed by (1.6) and (1.7), respectively. Note that since \mathbb{R}^m is finite dimensional, the norm $\|\cdot\|_Z$ that is chosen in

\mathbb{R}^m to express these properties could be arbitrary up to a modification of the stability constants α_Z, μ_Z . These constants can be optimally defined as

$$\alpha_Z = \sup_{v_1, v_2 \in V} \frac{\|\ell(v_1) - \ell(v_2)\|_Z}{\|v_1 - v_2\|_V},$$

and

$$\mu_Z = \sup_{v_1, v_2 \in V_n} \frac{\|v_1 - v_2\|_V}{\|\ell(v_1) - \ell(v_2)\|_Z}.$$

Note that one always has $\alpha_Z \mu_Z \geq 1$.

Remark 2.1. Note that when V_n is an n -dimensional space and the ℓ_i are linear functionals, the quantity μ_Z may be rewritten as

$$\mu_Z = \max_{v \in V_n} \frac{\|v\|_V}{\|\ell(v)\|_Z}.$$

As discussed further, the quantity μ defined in (1.3) for the analysis of the PBDW method is an instance of μ_Z corresponding to a particular choice of norm $\|\cdot\|_Z$. Assuming the ℓ_i are independent functionals, one easily checks that finiteness of this quantity imposes that $n \leq m$. Indeed, if $n > m$, there exists a non-trivial $v \in V_n \cap \mathcal{N}$, where

$$\mathcal{N} := \{v : \ell(v) = 0\}$$

is the null space of the measurement map that has codimension m , and therefore μ_Z is infinite.

Remark 2.2. The restriction $n \leq m$ is also needed for nonlinear spaces V_n and measurement ℓ , under assumptions expressing that m and n are local dimensions. More precisely, assume that the map φ defining V_n is differentiable at some x_0 in the interior of S , that ℓ is differentiable at $v_0 = \varphi(x_0)$, and that both tangent maps have full rank at these points, that is,

$$\dim(d\varphi_{x_0}(\mathbb{R}^n)) = n \quad \text{and} \quad \dim(d\ell_{v_0}(V)) = m.$$

Then, by taking $v_1 = v_0$ and $v_2 = \varphi(x_0 + tx)$ in the quotient that defines μ_Z , and letting $t \rightarrow 0$ for arbitrary $x \in \mathbb{R}^n$, one finds that

$$\mu_Z \geq \max_{v \in d\varphi_{x_0}(\mathbb{R}^n)} \frac{\|v\|_V}{\|d\ell_{v_0}(v)\|_Z},$$

and therefore it is infinite if $n \leq m$, by the same argument as in the previous remark.

2.2 The best fit estimator

We define a first recovery map $z \mapsto \tilde{u} = R(z)$ as the best fit estimator in the Z norm

$$\tilde{u} := \operatorname{argmin}\{\|z - \ell(v)\|_Z : v \in V_n\}. \quad (2.8)$$

The existence of such a minimizer is trivial if the space V_n and the measurement map ℓ are linear. It can also be ensured in the nonlinear case under additional assumptions, for example compactness of the set S defining the nonlinear space V_n , which will be the case in the application to shape

recovery discussed in §5. If the minimizer does not exist, we may consider a near minimizer, that is $\tilde{u} \in V_n$ satisfying

$$\|z - \ell(\tilde{u})\|_Z \leq C \|z - \ell(v)\|_Z, \quad v \in V_n,$$

for some fixed $C > 1$. Inspection of the proofs of our main results below reveals that similar recovery bounds can be obtained for such a near minimizer, up to the multiplicative constant C .

Recall that our assumption (1.5) on the noise model is a control on $\|\eta\|_p$ for some $1 \leq p \leq \infty$. For this value of p , we introduce the quantity

$$\beta_Z := \max_{z \in \mathbb{R}^m} \frac{\|z\|_Z}{\|z\|_p}$$

We are now in position to state a recovery bound in this general framework.

Theorem 2.3. *The best fit estimator \tilde{u} from (2.8) satisfies the estimate*

$$\|u - \tilde{u}\|_V \leq C_1 e_n(u) + C_2 \|\eta\|_p, \tag{2.9}$$

where $C_1 := 1 + 2\alpha_Z \mu_Z$ and $C_2 := 2\beta_Z \mu_Z$.

Proof: Consider any $v \in V_n$ and write

$$\|u - \tilde{u}\|_V \leq \|u - v\|_V + \|v - \tilde{u}\|_V \leq \|u - v\|_V + \mu_Z \|\ell(v) - \ell(\tilde{u})\|_Z,$$

where we have used (1.7). On the other hand, the minimizing property of \tilde{u} ensures that

$$\|\ell(v) - \ell(\tilde{u})\|_Z \leq \|z - \ell(v)\|_Z + \|z - \ell(\tilde{u})\|_Z \leq 2\|z - \ell(v)\|_Z.$$

Furthermore, using the stability (1.6) of ℓ and the definition of β_Z , we have

$$\|z - \ell(v)\|_Z \leq \|\ell(v) - \ell(u)\|_Z + \|\eta\|_Z \leq \alpha_Z \|u - v\| + \beta_Z \|\eta\|_p.$$

Combining the three estimates, we reach

$$\|u - \tilde{u}\|_V \leq (1 + 2\alpha_Z \mu_Z) \|u - v\|_V + 2\beta_Z \mu_Z \|\eta\|_p,$$

which gives (2.9) by optimizing over $v \in V_n$. □

The constants C_1 and C_2 in the above recovery estimate depend on the choice of norm $\|\cdot\|_Z$. Note that they are invariant when this norm is scaled by a factor $t > 0$, since this has the effect of multiplying α_Z and β_Z by t and dividing μ_Z by t , which is consistent with the fact that the resulting estimator \tilde{u} is left unchanged by such a scaling. In the next section we show, in the particular setting of linear measurements, that specific choices of $\|\cdot\|_Z$ can be used to minimize C_1 or C_2 . This setting also allows us to introduce and study a generalized interpolation estimator, which is not relevant to the present section since the nonlinear measurement map ℓ is not assumed to be surjective: in the presence of noise, there might exist no $v \in V$ that agrees with the data, in the sense that $z = \ell(u) + \eta$ does not belong to the range of ℓ .

3 Linear observations

In this section, we assume that the $\ell_i \in V'$ are independent linear functionals, still allowing V_n to be a general nonlinear space. In this framework, which contains the example of shape recovery discussed in §5, one has

$$\alpha_Z = \max_{v \in V} \frac{\|\ell(v)\|_Z}{\|v\|_V}$$

and

$$\mu_Z = \max_{v \in V_n^{\text{diff}}} \frac{\|v\|_V}{\|\ell(v)\|_Z},$$

where

$$V_n^{\text{diff}} = V_n - V_n := \{v_1 - v_2 : v_1, v_2 \in V_n\}.$$

In this particular setting, we can identify the norms $\|\cdot\|_Z$ that minimize the constants $C_1 := 1 + 2\alpha_Z\mu_Z$ and $C_2 := 2\beta_Z\mu_Z$, respectively.

3.1 Optimal norms

As $\ell : V \rightarrow \mathbb{R}^m$ is continuous and surjective, we can define a norm on \mathbb{R}^m through

$$\|z\|_W = \min\{\|v\|_V : \ell(v) = z\}. \quad (3.10)$$

Remark 3.1. *If V is a Hilbert space, the minimizer is unique by strict convexity of $\|\cdot\|_V$, and the m -dimensional space*

$$W := \{\operatorname{argmin}_{\ell(v)=z} \|v\|_V, z \in \mathbb{R}^m\}$$

is exactly the span of the Riesz representers of the observation functionals $\ell_i \in V'$. Moreover, denoting P_W the orthogonal projection on W , we have

$$\|\ell(v)\|_W = \|P_W v\|_V, \quad v \in V.$$

For this reason, we sometimes refer to $\|\cdot\|_W$ as the Riesz norm even in the case of a more general Banach space.

The following result shows that the choice $\|\cdot\|_Z := \|\cdot\|_W$ is the one that minimizes the constant C_1 , while C_2 is minimized by simply taking the ℓ^p norm $\|\cdot\|_Z = \|\cdot\|_p$.

Theorem 3.2. *For any norm $\|\cdot\|_Z$, one has*

$$\alpha_W\mu_W = \mu_W \leq \alpha_Z\mu_Z,$$

and

$$\beta_p\mu_p = \mu_p \leq \beta_Z\mu_Z,$$

where $(\alpha_W, \beta_W, \mu_W)$ and $(\alpha_p, \beta_p, \mu_p)$ are the triplets $(\alpha_Z, \beta_Z, \mu_Z)$ when $\|\cdot\|_Z := \|\cdot\|_W$ and $\|\cdot\|_Z = \|\cdot\|_p$, respectively.

Proof: One has

$$\alpha_W = \max_{v \in V} \frac{\|\ell(v)\|_W}{\|v\|_V} = \max_{z \in \mathbb{R}^m} \max_{\ell(v)=z} \frac{\|z\|_W}{\|v\|_V} = 1,$$

and so

$$\alpha_W \mu_W = \mu_W = \max_{v \in V_n^{\text{diff}}} \frac{\|v\|_V}{\|\ell(v)\|_W} \leq \max_{v \in V_n^{\text{diff}}} \frac{\|\ell(v)\|_Z}{\|\ell(v)\|_W} \max_{v \in V_n^{\text{diff}}} \frac{\|v\|_V}{\|\ell(v)\|_Z} = \max_{v \in V_n^{\text{diff}}} \frac{\|\ell(v)\|_Z}{\|\ell(v)\|_W} \mu_Z.$$

We now observe that from the definition of W , one has

$$\max_{v \in V_n^{\text{diff}}} \frac{\|\ell(v)\|_Z}{\|\ell(v)\|_W} \leq \max_{z \in \mathbb{R}^m} \frac{\|z\|_Z}{\|z\|_W} = \max_{z \in \mathbb{R}^m} \max_{\ell(v)=z} \frac{\|z\|_Z}{\|v\|_V} = \alpha_Z.$$

We have thus obtained the first claim $\alpha_W \mu_W = \mu_W \leq \alpha_Z \mu_Z$. For the second claim, note that we trivially have $\beta_p = 1$, and so

$$\beta_p \mu_p = \mu_p = \max_{v \in V_n^{\text{diff}}} \frac{\|v\|_V}{\|\ell(v)\|_p} \leq \max_{v \in V_n^{\text{diff}}} \frac{\|\ell(v)\|_Z}{\|\ell(v)\|_p} \max_{v \in V_n^{\text{diff}}} \frac{\|v\|_V}{\|\ell(v)\|_Z} \leq \beta_Z \mu_Z.$$

□

Remark 3.3. *In the particular case where V is a Hilbert space, V_n a linear subspace and $p = 2$, it was already observed in [6] that the reconstruction operators based on the choice $\|\cdot\|_Z = \|\cdot\|_W$ or $\|\cdot\|_Z = \|\cdot\|_2$ are the most stable with respect to the approximation error and the noise error, respectively. The above result may thus be seen as a generalization of this state of affairs to the case of nonlinear subspaces of Banach spaces, and ℓ^p noise.*

3.2 The generalized interpolation estimator

Thanks to the surjectivity of ℓ , we may introduce the space

$$V_z := \{v \in V : \ell(v) = z\},$$

and consider the minimization problem

$$\min_{v \in V_z} \min_{\tilde{v} \in V_n} \|v - \tilde{v}\|_V.$$

If $(u^*, \tilde{u}) \in V_z \times V_n$ is a minimizing pair, the function u^* is given by

$$u^* = u^*(z) \in \operatorname{argmin}\{\operatorname{dist}(v, V_n)_V : \ell(v) = z\},$$

and is called the generalized interpolation estimator, since it exactly matches the data.

Remark 3.4. *The best fit and generalized interpolation estimation may be thought of as the two extreme cases, $t \rightarrow \infty$ and $t \rightarrow 0$, of the penalized estimator*

$$u_t := \operatorname{argmin}\{\|z - \ell(v)\|_Z + t \operatorname{dist}(v, V_n)_V\}.$$

As explained earlier, the generalized interpolation operator may not be well defined in the general case where the ℓ_i are nonlinear. As opposed to the best fit, or the above penalized estimator u_t when $t > 0$, the generalized interpolation estimator does not involve the choice of a particular norm Z .

On the other hand, we see that \tilde{u} is the solution to the problem

$$\min_{\tilde{v} \in V_n} \text{dist}(\tilde{v}, V_z)_V.$$

Observing that

$$\text{dist}(\tilde{v}, V_z)_V = \min_{\ell(v)=z} \|\tilde{v} - v\|_V = \min_{\ell(v')=\ell(\tilde{v})-z} \|v'\|_V = \|\ell(\tilde{v}) - z\|_W,$$

we thus find that \tilde{u} is precisely the best fit estimator for the Riesz norm $\|\cdot\|_Z := \|\cdot\|_W$.

In the Hilbert space setting, the generalized interpolation estimator u^* is therefore the orthogonal projection of this particular best fit estimator \tilde{u} onto the affine space V_z . It may thus also be derived from \tilde{u} by the correction procedure

$$u^* = \tilde{u} + w - P_W \tilde{u},$$

where $w = \text{argmin}_{\ell(v)=z} \|v\|_V \in W$ is the preimage by ℓ of the measurements z . In the noiseless case when $w = P_W u$, this correction can only improve the approximation since it reduces the component of $u - \tilde{u}$ in the W direction while leaving unchanged the orthogonal component, and so, in view of Theorems 2.3 and 3.2, we are ensured that

$$\|u - u^*\|_V \leq C_1 e_n(u),$$

where $C_1 := 1 + 2\mu_W$.

More generally, in the noisy case, and without the assumption that V is a Hilbert space, there is no guarantee that u^* performs better than \tilde{u} , but we still obtain an error estimate on u^* that is similar in nature to that satisfied by \tilde{u} .

Theorem 3.5. *The generalized interpolation estimator u^* satisfies the estimate*

$$\|u - u^*\|_V \leq C_1 e_n(u) + C_2 \|\eta\|_p, \quad (3.11)$$

where $C_1 := 2 + 2\mu_W$ and $C_2 := (1 + 2\mu_W)\beta_W$.

Proof: Take $\delta \in \text{argmin}_{\ell(v)=\eta} \|v\|_V$, so that $\ell(\delta) = \eta$ and $\|\eta\|_W = \|\delta\|_V$. For v and v^* in V_n , decompose

$$\|u - u^*\|_V \leq \|u - v\|_V + \|v - v^*\|_V + \|v^* - u^*\|_V. \quad (3.12)$$

For the middle term, using (1.7), we write

$$\begin{aligned} \|v - v^*\|_V &\leq \mu_W \|\ell(v - v^*)\|_W \\ &\leq \mu_W (\|\ell(v - u)\|_W + \|\ell(u - u^*)\|_W + \|\ell(u^* - v^*)\|_W) \\ &\leq \mu_W (\|v - u\|_V + \|\eta\|_W + \|u^* - v^*\|_V) \end{aligned}$$

since $\alpha_W = 1$, so the decomposition (3.12) becomes

$$\|u - u^*\|_V \leq (1 + \mu_W) \|u - v\|_V + \mu_W \|\eta\|_W + (1 + \mu_W) \|v^* - u^*\|_V.$$

To bound the last term, we optimize over the choice of $v^* \in V_n$ and use the definition of u^* to obtain

$$\inf_{v^* \in V_n} \|v^* - u^*\|_V = \text{dist}(u^*, V_n) \leq \text{dist}(u + \delta, V_n) \leq \text{dist}(u, V_n) + \|\delta\|_V = e_n(u) + \|\eta\|_W$$

since $\ell(u + \delta) = \ell(u) + \eta = z$. Combining the last two estimates and optimizing over $v \in V_n$ gives

$$\|u - u^*\|_V \leq (2 + 2\mu_W)e_n(u) + (1 + 2\mu_W)\|\eta\|_W,$$

and the result follows from the definition of β_W . \square

4 Shape recovery from cell averages

4.1 The shape recovery problem

The problem of reconstructing a function u from its cell averages

$$a_T(u) := \frac{1}{|T|} \int_T u, \quad T \in \mathcal{T},$$

where \mathcal{T} is a partition of the domain $D \subset \mathbb{R}^d$ in which u is defined, appears naturally in two areas:

- In $2d$ or $3d$ image processing, it corresponds to the so-called super-resolution problem, that is, reconstructing a high resolution image from its low resolution version defined on the coarse grid \mathcal{T} of pixels or voxels.
- In numerical simulation of hyperbolic conservation laws, it plays a central role when developing finite volume schemes on the computation mesh \mathcal{T} .

Standard reconstruction methods are challenged when the function u exhibits jump discontinuities which are not well resolved by the partition \mathcal{T} . Such discontinuities correspond to edges in image processing or shocks in conservation laws. Here we may focus on the very simple case of characteristic functions of sets

$$u = \chi_\Omega,$$

that already carry the main difficulty. Therefore we are facing a problem of reconstructing a shape Ω from local averages of χ_Ω .

As a simple example we work in the domain $D = [0, 1]^2$ with a uniform grid based on square cells of sidelength $h = \frac{1}{L}$ for some $L > 1$, therefore of the form

$$\mathcal{T} = \mathcal{T}_h := \{T_{i,j} = [(i-1)h, ih] \times [(j-1)h, jh] : i, j = 1, \dots, L\}.$$

The cardinality of the grid is therefore

$$n := \#(\mathcal{T}) = L^2 = h^{-2}.$$

We consider classes of characteristic functions χ_Ω of sets $\Omega \subset D$ with boundary of a prescribed Hölder smoothness. The definition of these classes requires some precision.

Definition 4.1. For $s \geq 1$, $0 < R < 1/2$ and $M > 0$, we define the class $\mathcal{F}_{s,R,M}$ as consisting of all characteristic functions χ_Ω of domains $\Omega \subset [R, 1-R]^2 \subset D$ with the following property: for all $x \in D$ there exists an orthonormal system (e_1, e_2) and a function $\psi \in \mathcal{C}^s$ with $\|\psi\|_{\mathcal{C}^s} \leq M$, such that

$$y \in \Omega \iff z_2 \leq \psi(z_1),$$

for any $y = x + z_1 e_1 + z_2 e_2$ with $|z_1|, |z_2| \leq R$.

Here, we have used the usual definition

$$\|\psi\|_{\mathcal{C}^s} = \sup_{0 \leq k \leq \lfloor s \rfloor} \|\psi^{(k)}\|_{L^\infty([-R,R])} + \sup_{s,t \in [-R,R]} |s-t|^{\lfloor s \rfloor - s} \left| \psi^{(\lfloor s \rfloor)}(s) - \psi^{(\lfloor s \rfloor)}(t) \right|,$$

for the Hölder norm. In the case of integer smoothness, we use the convention that \mathcal{C}^s denotes functions with Lipschitz derivatives up to order $s-1$, so that in particular the case $s = 1$ corresponds to domains with Lipschitz boundaries.

Remark 4.2. The condition $\Omega \subset [R, 1-R]^2$ imposing that Ω remains away from the boundary ∂D might be quite restrictive in some applications; instead, one can assume that the domains Ω and D are periodic, or symmetrize Ω with respect to ∂D .

4.2 The failure of linear reconstruction methods

The most trivial linear reconstruction method consists in the piecewise constant approximation

$$\tilde{u} = \sum_{T \in \mathcal{T}} a_T(u) \chi_T. \quad (4.13)$$

The approximation rate of this reconstruction over the class $\mathcal{F}_{s,R,M}$ is as follows.

Proposition 4.3. Let $u = \chi_\Omega \in \mathcal{F}_{s,R,M}$, its piecewise constant approximation \tilde{u} by average values on each cell, defined in (4.13), satisfies

$$\|\chi_\Omega - \tilde{u}\|_{L^q} \leq C h^{\frac{1}{q}} = C n^{-\frac{1}{2q}},$$

where the constant C depends on R and M .

Proof: Let $N = \lceil (\sqrt{2}R)^{-1} \rceil$, and partition the domain $D = [0, 1]^2$ into N^2 squares of side $1/N$. Then each subsquare Q is contained in the set $\{x + z_1 e_1 + z_2 e_2, |z_1|, |z_2| \leq R\}$ from Definition 4.1, where x is the center of Q . Thus $\partial\Omega$ is the restriction of the graph of an M -Lipschitz function on Q , so its arc length is bounded by

$$|\partial\Omega \cap Q| \leq \text{diam}(Q) \sqrt{1+M^2} \leq 2R \sqrt{1+M^2}.$$

As any curve of arclength h intersects at most four cells from \mathcal{T} , $\partial\Omega \cap Q$ intersects at most $4 \lceil 2R \sqrt{1+M^2}/h \rceil$ cells, and summing over all subsquares, $\partial\Omega$ intersects at most $4N^2 \lceil 2R \sqrt{1+M^2}/h \rceil$ cells. Denoting $\mathcal{T}_{\partial\Omega}$ the set of these cells, and observing that $u|_T \equiv a_T(u) \in \{0, 1\}$ for $T \notin \mathcal{T}_{\partial\Omega}$, we get

$$\|\chi_\Omega - \tilde{u}\|_{L^q}^q = \sum_{T \in \mathcal{T}} \int_T |u - a_T(u)|^q \leq \sum_{T \in \mathcal{T}_{\partial\Omega}} |T| = h^2 |\mathcal{T}_{\partial\Omega}| \leq 24 \frac{\sqrt{1+M^2}}{R} h$$

for $h \leq R$, and this bound also holds for $h > R$ since $\|\chi_\Omega - \tilde{u}\|_{L^q}^q \leq 1$. \square

The next result shows, for the particular case $q = 2$, that no better rate can actually be achieved by any linear method, regardless of the smoothness s of the boundary. We conjecture that a similar result holds for $1 \leq q \leq \infty$. This motivates the use of nonlinear recovery methods, which are the object of the next section.

We recall that the Kolmogorov n -width of a compact set S from some Banach space V is defined by

$$d_n(S)_V := \inf_{\dim(E) \leq n} \text{dist}(S, E)_V,$$

where $\text{dist}(S, E)_V := \max_{u \in S} \min_{v \in E} \|u - v\|_V$ and the infimum is taken over all finite dimensional spaces E of dimension at most n .

Proposition 4.4. *Let $s \geq 1$ be arbitrary. Then for R sufficiently small, and M sufficiently large, there exists $c > 0$ such that the Kolmogorov n -widths of the class $\mathcal{F}_{s,R,M}$ satisfy*

$$d_n(\mathcal{F}_{s,R,M})_{L^2} \geq cn^{-\frac{1}{4}}, \quad n \geq 1.$$

Proof: The proof of this result relies on similar lower bounds for dictionaries of d -dimensional ridge functions

$$\mathbb{P}_k^d := \{x \mapsto \sigma_k(\omega \cdot x + b) : \|\omega\|_2 = 1, c_1 \leq b \leq c_2\}$$

where $\sigma_k(t) := \max\{0, t\}^k$ is the so-called RELU- k function. Here, we work in the space $L^2(B)$ where B is an arbitrary ball of \mathbb{R}^d , and the constants (c_1, c_2) are taken as the inf and sup of $\omega \cdot x$ as $x \in B$ and $\|\omega\|_2 = 1$, respectively, that is we take all b such that the line discontinuity of the k -th derivative of $\sigma_k(\omega \cdot x + b)$ crosses the ball B . Theorem 9 from [38], which improves on earlier results from [29], shows that if

$$B_1(\mathbb{P}_k^d) := \overline{\left\{ \sum_{j=1}^n a_j g_j : n \in \mathbb{N}, g_j \in \mathbb{P}_k^d, \sum_{j=1}^n |a_j| \leq 1 \right\}}$$

denotes the symmetrized convex hull of this dictionary (the closure being taken in $L^2(B)$), then

$$d_n(B_1(\mathbb{P}_k^d))_{L^2(B)} \geq cn^{-\frac{2k+1}{2d}}, \quad n \geq 1,$$

where c depends on k, d , and the diameter of B .

In our case of interest we work with the value $d = 2$ and $k = 0$, so that the ridge functions are simply the characteristic functions of half-planes. By convexity, we have

$$d_n(\mathbb{P}_0^2)_{L^2(B)} = d_n(B_1(\mathbb{P}_0^2))_{L^2(B)} \geq cn^{-\frac{1}{4}}.$$

We take for B the ball of center $(1/2, 1/2)$ and radius $1/4$, which is inside our domain $D = [0, 1]^2$. It is then readily seen that for R small enough and M large enough, we can extend any ridge function $g \in \mathbb{P}_0^2$ into a characteristic function χ_Ω from $\mathcal{F}_{s,R,M}$, as illustrated in Figure 1.

Observing that if E_D is a linear subspace of $L^2(D)$ of dimension at most n , its restriction E_B to B is a linear subspace of $L^2(B)$ of dimension at most n , and one has

$$\text{dist}(\chi_\Omega, E_B)_{L^2(B)} \leq \text{dist}(\chi_\Omega, E_D)_{L^2(D)}.$$

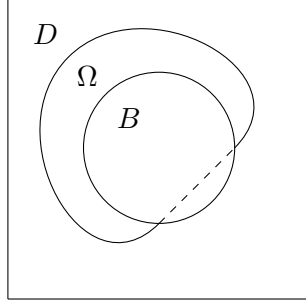


Figure 1: Example of extension of the indicator of a half-plane on B to the indicator of a smooth domain Ω on D

By infimizing, it follows that

$$d_n(\mathcal{F}_{s,R,M})_{L^2(D)} \geq d_n(\mathbb{P}_0^2)_{L^2(B)} \geq cn^{-\frac{1}{4}},$$

which concludes the proof. \square

Remark 4.5. *The fact that we impose conditions on R and M in the above statement is natural since the class $\mathcal{F}_{s,R,M}$ becomes empty if R is not small enough and M not large enough, due to the fact that the sets Ω are assumed to be contained in the interior of D .*

Remark 4.6. *The above results are easily extended to higher dimension $d \geq 2$, with a similar definition for the class $\mathcal{F}_{s,R,M}$. The rate of approximation in L^q norm by piecewise constant functions on uniform partitions is then $n^{-\frac{1}{dq}}$, which in the case $q = 2$ is proved by a similar argument to be the best achievable by any linear reconstruction method. We conjecture that the same holds for more general $1 \leq q \leq \infty$.*

5 Shape recovery by nonlinear least-squares

5.1 Nonlinear reconstruction on a stencil

We now discuss a nonlinear reconstruction method for $u \in \mathcal{F}_{s,R,M}$, whose output \tilde{u} is the indicator of a domain $\tilde{\Omega}$ with polygonal boundary : on each cell T , the domain $\tilde{\Omega}$ coincides with a certain half plane. In order to define the delimiting line we only use the average values of u on a 3×3 stencil of cells centered at T .

We assume that $h < R$, so that Ω does not intersect the boundary cells $T_{i,j}$ with i or j in $\{1, L\}$, and fix indices $1 < i, j < L$. For the cell $T = T_{i,j}$, denote $\bar{x} = ((i - \frac{1}{2})h, (j - \frac{1}{2})h)$ its center, and

$$S = [(i - 2)h, (i + 1)h] \times [(j - 2)h, (j + 1)h] = \bigcup_{i-1 \leq i' \leq i+1, j-1 \leq j' \leq j+1} T_{i'j'}$$

the stencil composed of T and its 8 neighboring cells. We define the nonlinear approximation space

$$V_2 := \{ \chi_{\bar{n} \cdot (x - \bar{x}) \geq c} : \bar{n} \in \mathbb{S}^1, c \in \mathbb{R} \}, \quad (5.14)$$

which is a two-parameter family as each function is determined by $(\arg \vec{n}, c) \in [0, 2\pi) \times \mathbb{R}$, where $\arg \vec{n}$ is the angle of \vec{n} with respect to the horizontal axis.

Here, our measurements are the average values of u on the cells contained in S

$$\ell(u) = (a_{T'}(u))_{T' \subset S} \in \mathbb{R}^9.$$

In order to find a reconstruction of u in V_2 based on these measurements, we need an inverse stability property of the form (1.7). This is not possible here, since ℓ cancels on all functions $\chi_\Omega \in V_2$ with $\Omega \cap S = \emptyset$. We therefore restrict the nonlinear family V_2 , and consider only indicators of half-planes whose boundary passes through the central cell T :

$$V_{2,T} := \left\{ \chi_\Omega \in V_2, \partial\Omega \cap T \neq \emptyset \right\} = \left\{ \chi_{\vec{n} \cdot (x - \bar{x}) \geq c}, \vec{n} \in \mathbb{S}^1, |c| \leq \frac{h}{2} |\vec{n}|_1 \right\}. \quad (5.15)$$

In this setting, we prove the existence of the following stability constants for $V = L^1(S)$ and $Z = \ell^1$, which is the best norm on \mathbb{R}^m in view of Theorem 3.2. For notational simplicity, we omit the reference to Z in these constants.

Proposition 5.1. *One has*

$$\|\ell(u)\|_1 \leq \alpha \|u\|_{L^1(S)}, \quad u \in L^1(D), \quad (5.16)$$

and

$$\|u - v\|_{L^1(S)} \leq \mu \|\ell(u - v)\|_1, \quad u, v \in V_{2,T}, \quad (5.17)$$

where $\alpha = h^{-2}$ and $\mu = \frac{3}{2}h^2$ are the optimal constants.

The proof of the stability property (5.16) is trivial since on each cell

$$|a_{T'}(u)| \leq |T'|^{-1} \|u\|_{L^1(T')} = h^{-2} \|u\|_{L^1(T')},$$

with equality in case u does not change sign. The proof of the inverse stability (5.17) is quite technical and left to the appendix.

Given the noisy observation

$$z = \ell(u) + \eta \in \mathbb{R}^9,$$

we define the estimator of u on the cell T by

$$\tilde{u}_T \in \operatorname{argmin}_{v \in V_2} \|z - \ell(v)\|_1. \quad (5.18)$$

Here we minimize over all V_2 , that is on all indicators of half planes, but we note that we may restrict to half-planes whose boundary passes through the stencil S .

The following result, which uses Proposition 5.1, shows that its distance to u in $L^1(T)$ is comparable to the error between u and its best approximation in the $L^1(S)$ norm

$$\bar{u}_S := \operatorname{argmin}_{v \in V_2} \|u - v\|_{L^1(S)}.$$

Lemma 5.2. *For all $u \in \mathcal{F}_{s,R,M}$, one has*

$$\|u - \tilde{u}_T\|_{L^1(T)} \leq C_1 \|u - \bar{u}_S\|_{L^1(S)} + 2\beta\mu \|\eta\|_p,$$

where $C_1 = 1 + 2\alpha\mu = 4$ and $C_2 = 2\beta\mu = 3^{3-\frac{2}{p}}h^2$, with α, μ as in Proposition 5.1, and $\beta = 9^{1-\frac{1}{p}}$ the maximal ratio between ℓ^p and ℓ^1 norm in \mathbb{R}^9 .

Proof. We distinguish two cases:

- If $\tilde{u}_T \in V_{2,T}$ and $\bar{u}_S \in V_{2,T}$, that is, both boundaries pass through the central cell T , we apply Theorem 2.9 together with Proposition 5.1

$$\begin{aligned} \|u - \tilde{u}_T\|_{L^1(T)} &\leq \|u - \tilde{u}_T\|_{L^1(S)} \leq C_1 \min_{v \in V_{2,T}} \|u - v\|_{L^1(S)} + C_2 \|\eta\|_p \\ &= C_1 \|u - \bar{u}_S\|_{L^1(S)} + C_2 \|\eta\|_p. \end{aligned}$$

with $C_1 = 1 + 2\alpha\mu$, $C_2 = 2\beta\mu$.

- Otherwise, either \tilde{u}_T or \bar{u}_S has constant value 0 or 1 on T , so $\tilde{u}_T - \bar{u}_S$ has constant sign on T , and thus

$$\begin{aligned} \|\bar{u}_S - \tilde{u}_T\|_{L^1(T)} &= h^2 |a_T(\tilde{u}_T - \bar{u}_S)| \leq h^2 \|\ell(\tilde{u}_T - \bar{u}_S)\|_1 \leq h^2 (\|\ell(\bar{u}_S) - z\|_1 + \|\ell(\tilde{u}_T) - z\|_1) \\ &\leq 2h^2 \|\ell(\bar{u}_S) - z\|_1 \leq 2h^2 \|\ell(\bar{u}_S - u)\|_1 + 2h^2 \|\eta\|_1 \leq 2\|u - \bar{u}_S\|_{L^1(S)} + 2h^2 \beta \|\eta\|_p. \end{aligned}$$

By triangle inequality, it follows that

$$\|u - \tilde{u}_T\|_{L^1(T)} \leq 3\|u - \bar{u}_S\|_{L^1(S)} + 2h^2 \beta \|\eta\|_p,$$

which has better constants than in the estimate obtained in the first case, since the constant C_0 is larger than 1. \square

The order of the best local approximation error $\|u - \bar{u}_S\|_{L^1(S)}$ that appears as a bound for the reconstruction error $\|u - \tilde{u}_T\|_{L^1(T)}$ depends on the smoothness of the boundary, as expressed in the following lemma.

Lemma 5.3. *For all $u \in \mathcal{F}_{s,R,M}$, with $R \geq \frac{3}{\sqrt{2}}h$, one has*

$$\|u - \bar{u}_S\|_{L^1(S)} \leq M(3\sqrt{2}h)^{\min(s,2)+1}.$$

Proof. We apply the definition of $\mathcal{F}_{s,R,M}$ at point \bar{x} : as $R \geq \frac{3}{\sqrt{2}}h$, the stencil S is contained in the domain

$$\{\bar{x} + z_1 e_1 + z_2 e_2, |z_1|, |z_2| \leq R\},$$

so $u|_S$ is the indicator of a domain delimited by a \mathcal{C}^s function ψ , with $\|\psi\|_{\mathcal{C}^s} \leq M$. From the definition of \mathcal{C}^s , there exists an affine function ξ such that

$$|\psi(z_1) - \xi(z_1)| \leq M(3\sqrt{2}h)^{\min(s,2)}, \quad |z_1| \leq \frac{3}{\sqrt{2}}h.$$

Then the function $v : \bar{x} + z_1 e_1 + z_2 e_2 \mapsto \chi_{z_2 \leq \xi(z_1)}$ belongs to V_2 , and we have

$$\|u - \bar{u}_S\|_{L^1(S)} \leq \|u - v\|_{L^1(S)} \leq M(3\sqrt{2}h)^{\min(s,2)+1}.$$

\square

5.2 Global nonlinear reconstruction

We now consider the process of recovering $u \in \mathcal{F}_{s,R,M}$ globally from its data

$$z = \ell(u) + \eta,$$

where now $\ell(u) := (a_T(u))_{T \in \mathcal{T}} \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^n$ is the noise vector. Applying to each inner cell $T \in \mathcal{T}$ the previous reconstruction procedure based on the 3×3 stencil S centered at T , we obtain a global recovery $\tilde{u} = \tilde{u}(z)$ such that

$$\tilde{u}|_T = \tilde{u}_T|_T, \quad T = T_{i,j} \in \mathcal{T}, \quad 1 < i, j < L,$$

where \tilde{u}_T is the local estimator from (5.18). On the boundary cells $T = T_{i,j}$ with i or j in $\{1, L\}$, $u|_T$ is zero by Definition 4.1 so we simply set $\tilde{u}|_T = 0$. Note that \tilde{u} is of the form

$$\tilde{u} = \chi_{\tilde{\Omega}},$$

where $\tilde{\Omega}$ has piecewise linear boundary with respect to the mesh \mathcal{T} . The following result gives a global approximation bound, which confirms the improvement over linear methods when $s > 1$.

Theorem 5.4. *For all $u \in \mathcal{F}_{s,R,M}$, one has*

$$\|u - \tilde{u}\|_{L^q(D)} \leq C_1 n^{-\frac{\min(1,s/2)}{q}} + C_2 n^{-\frac{1}{pq}} \|\eta\|_p^{\frac{1}{q}}.$$

Proof. First notice that if the result is proved for $p = q = 1$, as $u - v$ has values in $\{-1, 0, 1\}$,

$$\|u - v\|_{L^q(D)}^q = \|u - v\|_{L^1(D)} \leq C_1 n^{-1} + C_2 n^{-1} \|\eta\|_1 \leq \left(C_1^{\frac{1}{q}} n^{-\frac{1}{q}} + C_2^{\frac{1}{q}} n^{-\frac{1}{pq}} \|\eta\|_1^{\frac{1}{q}} \right)^q,$$

so it suffices treat the case $p = q = 1$.

By an argument similar to the proof of Proposition 4.3, $\partial\Omega$ intersects at most $16N^2[2R\sqrt{1+M^2}/h]$ stencils of 9 cells. Using the fact that $u = \bar{u}_S$ is a constant on any other stencil, we get

$$\begin{aligned} \|u - \tilde{u}\|_{L^1(D)} &= \sum_{T \text{ inner cell}} \|u - \tilde{u}\|_{L^1(T)} \leq \sum_{T \text{ inner cell}} (1 + 2\alpha\mu) \|u - \bar{u}\|_{L^1(S)} + 2\beta\mu \|\eta\|_{\ell^1(S)} \\ &\leq 16N^2 \left\lceil \frac{2R\sqrt{1+M^2}}{h} \right\rceil M(3\sqrt{2}h)^{\min(s,2)+1} + 18\beta\mu \|\eta\|_1 \leq C_1 h^{\min(s,2)} + C_2 h^2 \|\eta\|_1. \end{aligned}$$

We conclude by recalling that $n = h^{-2}$. □

Remark 5.5. *Here the convergence rate for the noiseless term $n^{-\frac{\min(1,s/2)}{q}}$ is limited due to the use of polygonal domains in the reconstruction. So the best approximation rate $h^{\frac{2}{q}} = n^{-\frac{1}{q}}$ is already attained for C^2 boundaries. When the smoothness parameter s is larger than 2, better rates $n^{-\frac{s}{2q}}$ should be reachable if we use non-linear approximation spaces that are richer than the space V_2 , for example indicator functions of domains with boundary that have a higher order polynomial description rather than straight lines. Of course, the stable identification of these approximants in the sense of (1.7) might require stencils that are of larger size than 3×3 .*

Remark 5.6. *If $\|\eta\|_\infty \leq \frac{1}{9}$, then \tilde{u} is exactly equal to u on any cell whose corresponding stencil does not intersect $\partial\Omega$, so the error is concentrated on $\mathcal{O}(\sqrt{n})$ cells, leading to an improved rate $n^{-\frac{p+1}{2pq}}$ instead of $n^{-\frac{1}{pq}}$ for the noise term.*

5.3 Numerical illustration

We study the behavior of the above discussed linear and non-linear recovery methods from cell averages for the particular target function $u = \chi_\Omega$, with Ω a slightly decentered disk of radius $r = 0.325$.

The linear method consists of the piecewise constant approximation (4.13), referred to as *PiecewiseConstant*. As to the nonlinear method, for the local best fit problem, we use the ℓ^2 norm on \mathbb{R}^9 instead of the ℓ^1 norm. By norm equivalence on \mathbb{R}^9 , the same convergence results can be proved to hold with different constants. This method, which we refer to as *LinearInterface*, does not ensure consistency of the reconstruction in the sense that $a_T(\tilde{u}) = a_T(u)$. One way to approach this consistency property is to modify the ℓ^2 norm by putting a large weight on the central cell. We refer to this variant as *LinearInterfaceCC*, here taking the weight 100.

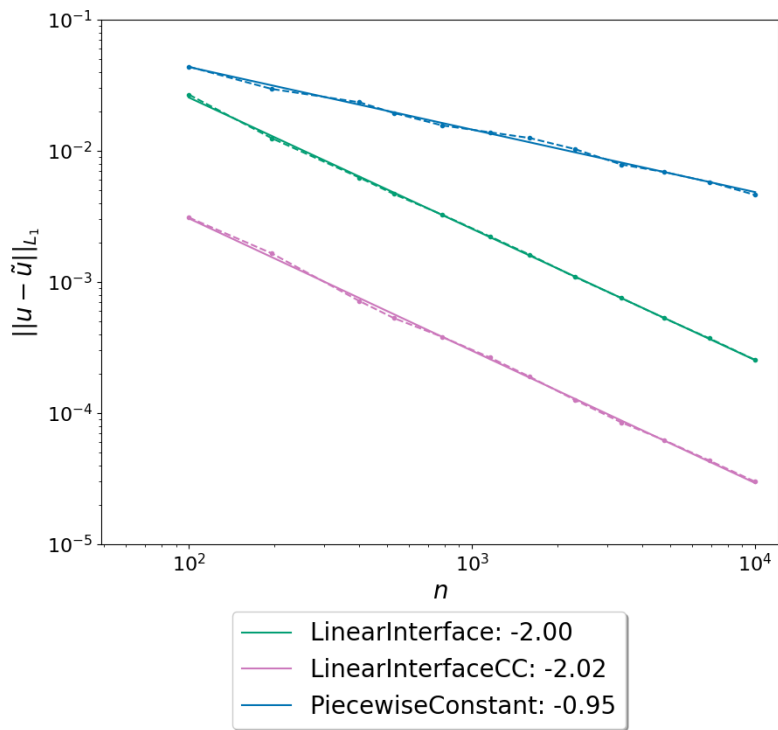
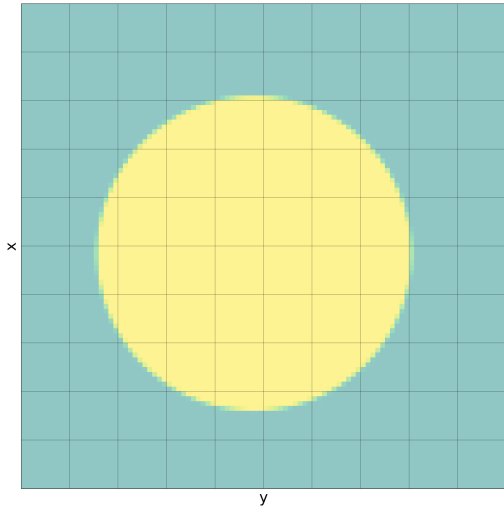
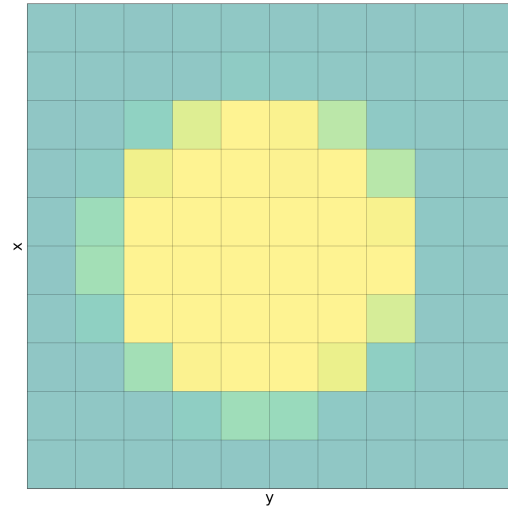


Figure 2: Convergence curves for the linear and nonlinear recovery methods

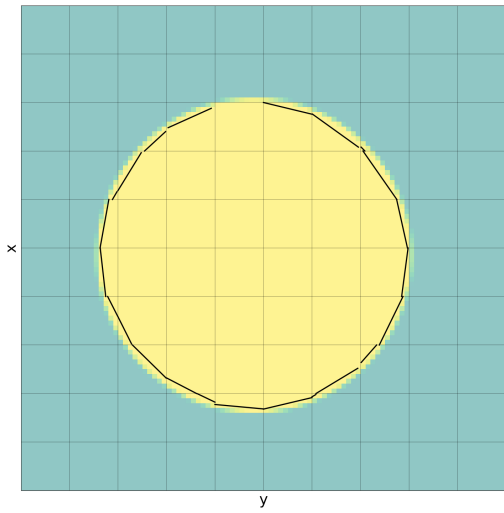
Figure 2 shows the convergence rates of the three methods in the L^1 norm. The expected h^{-2} decay is observed in both non-linear methods while the linear method lags behind with a decay rate of h^{-1} . It is relevant to note that although both non-linear methods benefit from the same rate, the associated constants differ by an order of magnitude, showing the practical improvement gained by imposing consistency. This improvement is also visible on Figure 3 which shows that in the *LinearInterface* method, the interfaces that minimize the l_2 error on the 9 surrounding cells lay always inside the circle as the curvature of the boundary pushes them towards the center. On the contrary, *LinearInterfaceCC* seems to find the right compromise between sticking to the cell average



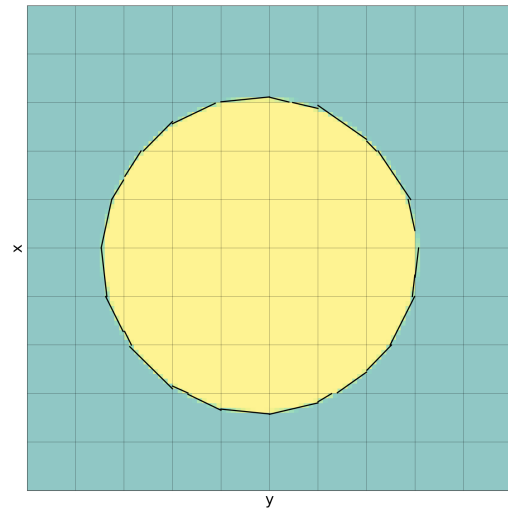
(a)



(b)



(c)



(d)

Figure 3: (a) The target function, (b) its recovery by PiecewiseConstant showing the cell-average data, and the recovered boundaries by (c) LinearInterface and (d) LinearInterfaceCC methods.

while capturing at the same time the curvature trend hinted by the surrounding cell averages.

6 Relation to compressed sensing

6.1 Compressed sensing and best n -term approximation

In this section we discuss the application of our setting to the sparse recovery of large vectors from a few linear observations. We thus take

$$V = \mathbb{R}^N,$$

equipped with some given norm $\|\cdot\|_V$ of interest. The linear measurements of $u = (u_1, \dots, u_N)^\top \in \mathbb{R}^N$ are given by

$$(\ell_1(u), \dots, \ell_m(u))^\top = \Phi u,$$

where Φ is an $m \times N$ measurement matrix, with typically $m \ll N$.

The topic of compressed sensing deals with sparse recovery of u from such measurements, that is, searching to recover an accurate approximation to u by a vector with only a few non-zero components. We refer to [10] for some first highly celebrated breakthrough results and to [21] for a general treatment.

We define the nonlinear space of n -sparse vectors as

$$V_n := \left\{ u \in \mathbb{R}^N : \|u\|_0 := \#\{i : u_i \neq 0\} \leq n \right\},$$

and the best n -term approximation error in the V norm as

$$e_n(u)_V := \min_{v \in V_n} \|u - v\|_V.$$

One natural question is to understand for which type of measurement matrices Φ does the noise-free measurement $y = \Phi u$ contain enough information, in order to recover any u up to an error $e_n(u)_V$. In other words, one asks if there exists a recovery map $R : \mathbb{R}^m \rightarrow \mathbb{R}^N$ such that one has the *instance optimality property* at order n

$$\|u - R(\Phi u)\|_V \leq C_0 e_n(u)_V, \quad u \in \mathbb{R}^N, \quad (6.19)$$

with C_0 a fixed constant, which we denote by $IOP(n, C_0)$. This question has been answered in [12] in terms of the null space $\mathcal{N} := \{v \in \mathbb{R}^N : \Phi v = 0\}$. We say that Φ satisfies the *null space property* at order k with constant C_1 , denoted by $NSP(k, C_1)$ if and only if

$$\|v\|_V \leq C_1 e_k(v)_V, \quad v \in \mathcal{N}. \quad (6.20)$$

This property quantifies how much vectors from the null space can be concentrated on a few coordinates. One main result of [12] is the equivalence between IOP at order n and NSP at order $2n$ in the following sense.

Theorem 6.1. *One has $IOP(n, C_0) \Rightarrow NSP(2n, C_0)$ and conversely $NSP(2n, C_1) \Rightarrow IOP(n, 2C_1)$.*

One natural question is whether matrices Φ with such properties can be constructed with a number of rows/measurements m barely larger than n . As we recall further the answer to this question is strongly tied to the norm V used on \mathbb{R}^N .

6.2 Stability and the null space property

The nonlinear estimation results that we have obtained in §2 and §3 can be applied to the setting of sparse recovery, offering us a different vehicle than the null space property to establish instance optimality.

In the present setting, for a given norm $\|\cdot\|_Z$, the stability property (1.6) takes the form

$$\|\Phi u\|_Z \leq \alpha_Z \|u\|_V, \quad u \in \mathbb{R}^N \quad (6.21)$$

and the inverse stability property (1.7) takes the form

$$\|v\|_V \leq \mu_Z \|\Phi v\|_Z, \quad v \in V_{2n}, \quad (6.22)$$

since for sparse vectors we have $V_n^{\text{diff}} = V_n - V_n = V_{2n}$. We refer to these properties as $S(\alpha_Z)$ and $IS(2n, \mu_Z)$, respectively.

Application of Theorem 2.3 in the noiseless case immediately gives us that the nonlinear best fit recovery $R(\Phi u) = \tilde{u}$ satisfies the instance optimality bound (6.19) with constant $C_0 = 1 + 2\alpha_Z \mu_Z$. In other words

$$S(\alpha_Z) \text{ and } IS(2n, \mu_Z) \Rightarrow IOP(n, C_0), \quad C_0 = 1 + 2\alpha_Z \mu_Z. \quad (6.23)$$

The following result shows that (S, IS) is actually equivalent to NSP , and thus to IOS , in the sense that a converse result holds when $\|\cdot\|_Z$ is chosen to be the Riesz norm(3.10).

Theorem 6.2. *For any norm $\|\cdot\|_Z$, one has*

$$S(\alpha_Z) \text{ and } IS(2n, \mu_Z) \Rightarrow NSP(2n, C_1), \quad C_1 = 1 + \alpha_Z \mu_Z. \quad (6.24)$$

Conversely, let $\|\cdot\|_W$ be the Riesz norm so that $\|\Phi u\|_W = \min_{\Phi v = \Phi u} \|v\|_V$, then

$$NSP(2n, C_1) \Rightarrow S(\alpha_W) \text{ and } IS(2n, \mu_W), \quad \alpha_W = 1 \text{ and } \mu_W = 1 + C_1. \quad (6.25)$$

Proof. Assume that $S(\alpha_Z)$ and $IS(2n, \mu_Z)$ hold. Let $v \in \mathcal{N}$ and \tilde{v} its best approximation in V_{2n} , then

$$\begin{aligned} \|v\|_V &\leq \|v - \tilde{v}\|_V + \|\tilde{v}\|_V \\ &\leq e_{2n}(v)_V + \mu_Z \|\Phi \tilde{v}\|_W \\ &= e_{2n}(v)_V + \mu_Z \|\Phi(v - \tilde{v})\|_W \leq (1 + \alpha_Z \mu_Z) e_{2n}(v)_V. \end{aligned}$$

This shows that $NSP(2n, C_1)$ holds with $C_1 = 1 + \alpha_Z \mu_Z$.

Conversely, assume that $NSP(2n, C_1)$ holds. From the definition of the Riesz norm, it is immediate that $S(\alpha_W)$ holds with $\alpha_W = 1$. For $v \in V_{2n}$, let \tilde{v} be the minimizer of $\min_{\Phi \tilde{v} = \Phi v} \|\tilde{v}\|_V$. Then, one has

$$\|v\|_V \leq \|\tilde{v}\|_V + \|v - \tilde{v}\|_V \leq \|\tilde{v}\|_V + C_1 \sigma_{2n}(v - \tilde{v})_V \leq (1 + C_1) \|\tilde{v}\|_V,$$

by using v as a sparse approximation to $v - \tilde{v}$. Since $\|\tilde{v}\|_V = \|\Phi v\|_W$, this shows that $IS(2n, \mu_W)$ holds with $\mu_W = 1 + C_1$. \square

6.3 The case of ℓ^p norms

The range of m allowing the properties to be fulfilled is best understood in the case of the ℓ^p norms, that is $\|\cdot\|_V = \|\cdot\|_p$, as discussed in [12] which points out a striking difference between the case $p = 2$ and $p = 1$:

1. In the case $p = 2$, it is proved that $NSP(2, C_1)$ cannot hold unless $N \leq C_1^2 m$. In other words, instance optimality in ℓ^2 even at order $n = 1$ requires a number of measurements that is proportional to the full space dimension.
2. In the more favorable case $p = 1$, it is proved that for matrices which satisfy the ℓ^2 -RIP property of order $3n$

$$(1 - \delta)\|v\|_2^2 \leq \|\Phi v\|_2^2 \leq (1 + \delta)\|v\|_2^2, \quad v \in V_{3n},$$

with parameter $0 < \delta < \frac{(\sqrt{2}-1)^2}{3}$, the $NSP(2n, C_1)$ holds with C_1 depending on δ . Such matrices are known to exist with $m \sim n \log(N/n)$ rows.

Our setting based on the stability properties S and IS applies more naturally to a different class of matrices built from graphs, which is also known to be well adapted for sparse recovery in the ℓ^1 norm. A bipartite graph with (N, m) left and right vertices, and of left degree d , is an (l, ε) -graph expander if

$$|X| \leq l \Rightarrow |N(X)| \geq d(1 - \varepsilon)|X|, \quad X \subset \{1, \dots, N\},$$

where $N(X) \subset \{1, \dots, m\}$ is the set of vertices connected to X . We necessarily have $|N(X)| \leq d|X|$, and $(1 - \varepsilon)d|X| \geq m$. From [11], it is known that there exists a $(2n, \frac{1}{2})$ -graph expander with $d \sim \log \frac{N}{n}$ and $m \sim nd \sim n \log(N/n)$.

Now denote $\Phi \in \{0, 1\}^{m \times N}$ the adjacency matrix of this graph, so that each column of Φ has d nonzero entries. Then

$$\|\Phi x\|_1 \leq d\|x\|_1, \quad x \in \mathbb{R}^N,$$

and

$$\|\Phi x\|_1 \geq d(1 - \varepsilon)\|x\|_1, \quad x \in V_{2n}.$$

Therefore $S(\alpha_1)$ and $IS(2n, \mu_1)$, hold with $\alpha_1 = d$ and $\mu_1 = \frac{1}{d(1-\varepsilon)} = \frac{2}{d}$, which by (6.24) and (6.23) gives $NSP(2n, C_1)$ with $C_1 = 3$ and $IOP(n, C_0)$ with $C_0 = 5$.

7 Appendix: Proof of Proposition 5.1

The proof contains 15 cases, represented on a tree in Figure 4. These cases correspond to different geometric situations, up to certain symmetries that leave the final relevant quantities $\|\ell(w)\|_1$ and $\|w\|_{L^1(S)}$ unchanged.

Node a: Take $w = u - v \in V_{2,T}^{\text{diff}}$, with $u, v \in V_{2,T}$, and denote \vec{n}_u, \vec{n}_v and c_u, c_v the corresponding unit vectors and offsets from the definition 5.15 of $V_{2,T}$. Recalling that $\bar{x} = (\bar{x}_1, \bar{x}_2)$ is the center of S , we also denote

$$\Delta_u = \{x \in \mathbb{R}^2, (x - \bar{x}) \cdot \vec{n}_u = c_u\}$$

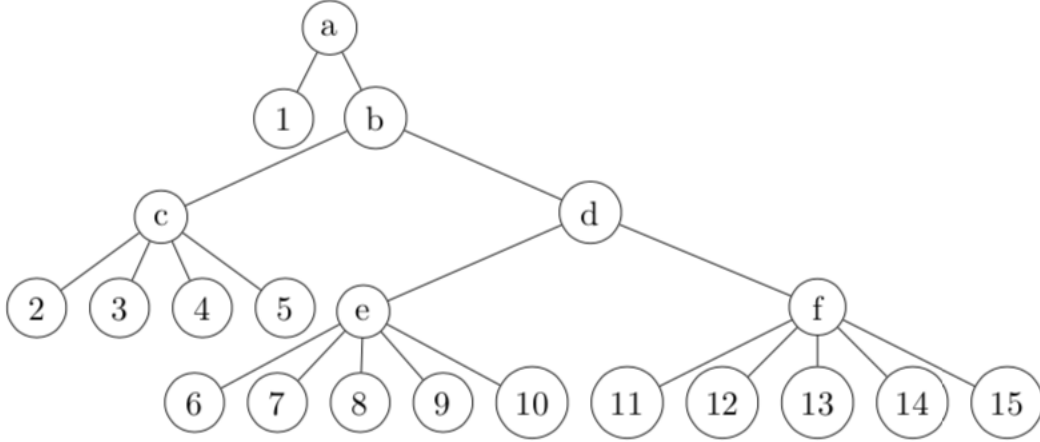


Figure 4: Structure of the proof, each leaf corresponds to a different case, and each node contains a general treatment valid for all its sons

the delimiting line between $\{u = 0\}$ and $\{u = 1\}$, and define Δ_v in a similar way.

Case 1: If $\vec{n}_u = \vec{n}_v = \vec{n}$, we have

$$w = \begin{cases} \chi_{c_u \leq \vec{n} \cdot (x - \bar{x}) < c_v} & \text{if } c_u \leq c_v \\ -\chi_{c_v \leq \vec{n} \cdot (x - \bar{x}) < c_u} & \text{otherwise} \end{cases}$$

so w has constant sign, which implies $\|w\|_{L^1(S)} = h^2 \|\ell(w)\|_1$.

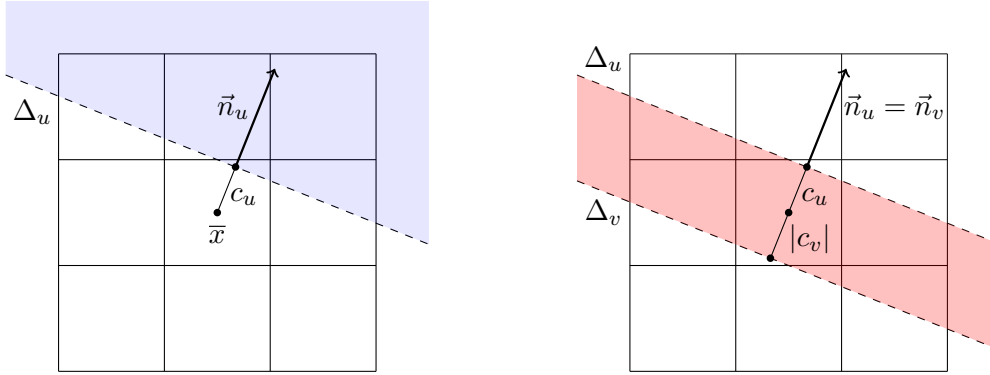


Figure 5: Left: 3×3 stencil S , with \bar{x} its center, and an example of function $u \in V_{2,T}$ with directing vector \vec{n}_u and offset $c_u > 0$. Here the dotted line corresponds to Δ_u , and the shaded region to $u = 1$, while $u = 0$ elsewhere. Right: Representation of Case 1 ($\vec{n}_u = \vec{n}_v$), here $c_v < 0 < c_u$ so $w = -1$ on the shaded region and $w = 0$ elsewhere

Node b: In all other cases, the cones

$$\mathcal{C}_+ = \{x \in \mathbb{R}^2, w(x) = 1\} \quad \text{and} \quad \mathcal{C}_- = \{x \in \mathbb{R}^2, w(x) = -1\}$$

are non-empty, and we can define the external bisector

$$\Delta = \{x \in \mathbb{R}^2, (\vec{n}_u - \vec{n}_v) \cdot (x - \bar{x}) = c_u - c_v\},$$

which is the line of symmetry between \mathcal{C}_+ and \mathcal{C}_- . We also denote

$$\mathcal{C} = \mathcal{C}_+ \cup \mathcal{C}_- = \{x \in \mathbb{R}^2, |w(x)| = 1\}.$$

Observing that

$$\|w\|_{L^1(S)} = |S \cap \mathcal{C}| \tag{7.26}$$

and

$$\|\ell(w)\|_1 = h^{-2} \sum_{T \subset S} \left| |T \cap \mathcal{C}_+| - |T \cap \mathcal{C}_-| \right|, \tag{7.27}$$

the stability property (5.17) can be rewritten as

$$|S \cap \mathcal{C}| \leq \frac{3}{2} \sum_{T \subset S} \left| |T \cap \mathcal{C}_+| - |T \cap \mathcal{C}_-| \right| = \frac{3}{2} \left(|S \cap \mathcal{C}| - 2 \sum_{T \subset S} \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|) \right),$$

or equivalently

$$|S \cap \mathcal{C}| \geq 6 \sum_{T \subset S} \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|). \tag{7.28}$$

Up to a rotation of S by a multiple of $\frac{\pi}{2}$, we may assume without loss of generality that

$$\arg(\vec{n}_u - \vec{n}_v) \in \left[\frac{\pi}{4}, \frac{3\pi}{4} \right],$$

that is, Δ is at an angle of at most $\frac{\pi}{4}$ with the horizontal axis, and \mathcal{C}_+ lies above Δ . Take (\vec{e}_1, \vec{e}_2) the canonical basis of \mathbb{R}^2 .

Node c: Consider the situation where $(\vec{n}_u \cdot \vec{e}_2)(\vec{n}_v \cdot \vec{e}_2) > 0$. As $\vec{n}_u \neq \vec{n}_v$ and $\vec{n}_u \neq -\vec{n}_v$, the lines Δ_u and Δ_v intersect at one point $X \in \mathbb{R}^2$. Moreover, the above condition implies $X + \vec{e}_2 \notin \mathcal{C}$. Using the fact that $|\arg(\Delta)| \leq \frac{\pi}{4}$, we also get $X + \vec{e}_1 \notin \mathcal{C}$.

Up to a symmetry with respect to the vertical axis, we can assume that \mathcal{C}_+ is included in the quadrant $X + \mathbb{R}_+^2$. Now consider a cell $T \subset S$ such that $\min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|) \neq 0$, then there exist points $x \in T \cap \mathcal{C}_-$ and $y \in T \cap \mathcal{C}_+$. As $x_1 \leq X_1 \leq y_1$ and $x_2 \leq X_2 \leq y_2$, we get $X \in T$, so there is at most one such cell T , and inequality (7.28) reduces to

$$|S \cap \mathcal{C}| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

Case 2: If $X \notin S$, then w has constant sign on S , so $\|w\|_{L^1(S)} = h^2 \|\ell(w)\|_1$.

Case 3: If X is in the central cell T , the dilation of T with respect to X by a factor 2 is a subset of S , and the image of $\mathcal{C} \cap T$ is in $\mathcal{C} \cap S$, so

$$|S \cap \mathcal{C}| \geq 4|T \cap \mathcal{C}| \geq 8 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

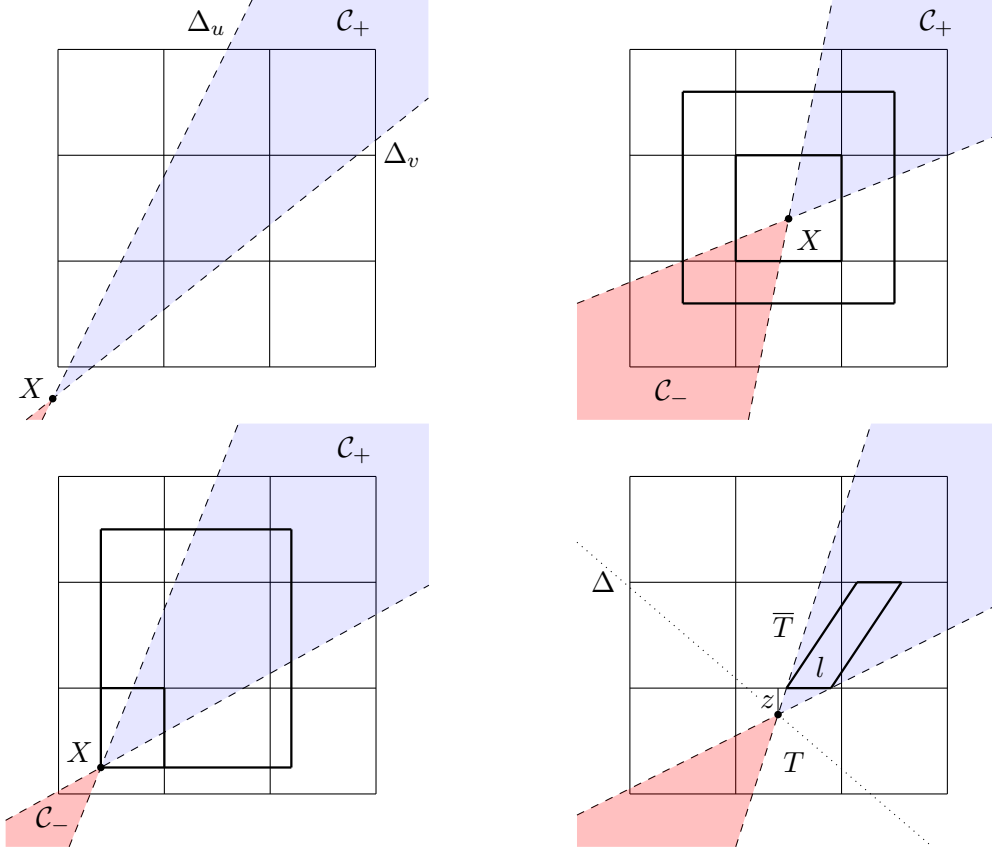


Figure 6: Cases 2, 3, 4, and 5

Case 4: If X is in the lower left cell T , the dilation of $T \cap \mathcal{C}_+$ with respect to X by a factor 3 is in $S \cap \mathcal{C}_+$, so

$$|S \cap \mathcal{C}| \geq |S \cap \mathcal{C}_+| \geq 9|T \cap \mathcal{C}_+| \geq 9 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

The same argument holds with \mathcal{C}_- instead of \mathcal{C}_+ when X is in the upper right cell. Moreover, as Δ_u and Δ_v go through the central cell, X may not be in the upper left or lower right cells.

Case 5: If X is in the lower central cell T , denote $l = |\partial T \cap \mathcal{C}_+| \in (0, h)$ the distance between Δ_u and Δ_v when they pass from T to the central cell \bar{T} , and $z = \text{dist}(X, \bar{T}) \in (0, h)$ the depth of the point of intersection. Then

$$|T \cap \mathcal{C}_+| = \frac{zl}{2} \quad \text{and} \quad |T \cap \mathcal{C}_-| \leq \frac{zl}{2} \left(\frac{h-z}{z} \right)^2,$$

so $\min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|) \leq \frac{hl}{4}$. On the other hand, the parallelogram of base $\partial T \cap \mathcal{C}_+$, of height h , and with sides orthogonal to Δ belongs to $(S \setminus T) \cap \mathcal{C}_+$ (it does not escape to the right of S because Δ is close to the horizontal axis, so the sides of the parallelogram are at an angle at most $\frac{\pi}{4}$ with the vertical axis), and has an area hl , which proves that

$$|\mathcal{C} \cap S| \geq hl + |\mathcal{C}_+ \cap T| + |\mathcal{C}_- \cap T| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

A similar construction can be applied to the remaining cases where X is in the upper central, central left or central right cell, which concludes the proof for Node c.

Node d: If now $(\vec{n}_u \cdot \vec{e}_2)(\vec{n}_v \cdot \vec{e}_2) \leq 0$, as $\arg(\vec{n}_u - \vec{n}_v) \in [\frac{\pi}{4}, \frac{3\pi}{4}]$, we get $\vec{n}_u \cdot \vec{e}_2 \geq 0 \geq \vec{n}_v \cdot \vec{e}_2$. Observe that $\mathcal{C}_+ + \vec{e}_2 \subset \mathcal{C}_+$ since for all $x \in \mathcal{C}_+$,

$$(x + \vec{e}_2 - \bar{x}) \cdot \vec{n}_u \geq (x - \bar{x}) \cdot \vec{n}_u \geq c_u \quad \text{and} \quad (x + \vec{e}_2 - \bar{x}) \cdot \vec{n}_v \leq (x - \bar{x}) \cdot \vec{n}_v < c_v.$$

In the same way, $\mathcal{C}_- - \vec{e}_2 \subset \mathcal{C}_-$. We now divide S into columns separated by the vertical boundaries between cells, and in addition by vertical lines where Δ intersects the two horizontal lines separating cells of S , as illustrated in Figure 7.

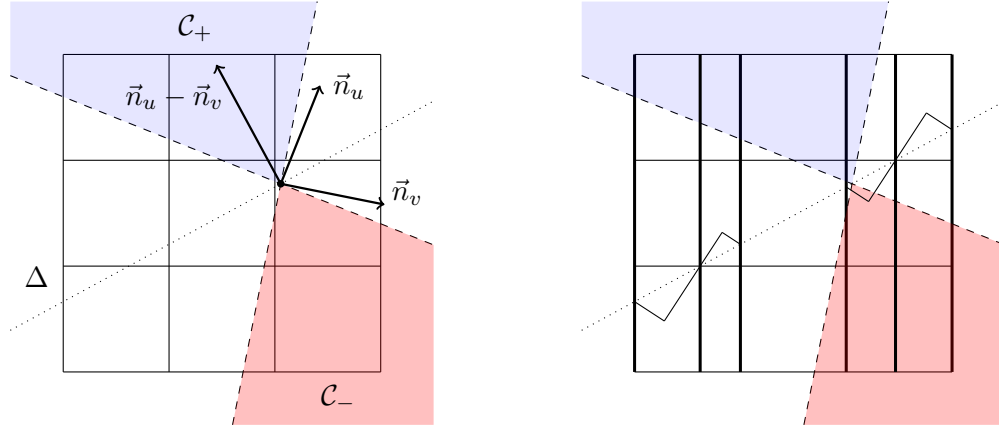


Figure 7: Generic situation for Node d, and partition of S into 5 columns: here, in addition to the 4 vertical lines delimiting the cells of S , we added 2 vertical lines passing through the intersections of Δ with the 2 horizontal cell delimiters

Let U be such a column, and T a cell intersecting U . If $T \cap U \neq T$, Δ intersects either the upper or lower boundary of T , but not both since Δ is at an angle of at most $\frac{\pi}{4}$ with the horizontal axis. If it is the upper boundary, the symmetric of the part of $T \cap U$ above Δ with respect to Δ is in $T \cap U$. If it is the lower boundary, the symmetric of the part of $T \cap U$ below Δ with respect to Δ is in $T \cap U$. Using the fact that \mathcal{C}_+ and \mathcal{C}_- are symmetric with respect to Δ , we obtain

$$\min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|) = \min(|T \cap U \cap \mathcal{C}_+|, |T \cap U \cap \mathcal{C}_-|) + \min(|T \cap U^c \cap \mathcal{C}_+|, |T \cap U^c \cap \mathcal{C}_-|).$$

Thanks to this observation, instead of (7.28) we only have to prove the inequality

$$|U \cap \mathcal{C}| \geq 6 \sum_{T \subset U} \min(|T \cap U \cap \mathcal{C}_+|, |T \cap U \cap \mathcal{C}_-|) \quad (7.29)$$

on each column U separately. We thus consider only one column U in the sequel, and assume up to a horizontal dilation (which preserves the condition $|\arg(\Delta)| \leq \frac{\pi}{4}$) that U has width h and is composed of three full cells.

According to the definition of the columns, there is at most one cell $T \subset U$ such that $T \cap \Delta \neq \emptyset$, and as Δ separates \mathcal{C}_+ and \mathcal{C}_- , it is only for this cell that we may have $\min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|) \neq 0$.

If there is no such cell, (7.29) trivially holds. Otherwise, similar to Node c, we only need to prove

$$|U \cap \mathcal{C}| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|),$$

where $T \subset U$ is the cell containing $\Delta \cap U$. Denoting P_1, P_2, P_3 and P_4 the upper left, upper right, lower left and lower right corner points of T , we observe that the assumptions on Δ and U imply $P_1, P_2 \notin \mathring{\mathcal{C}}_-$ and $P_3, P_4 \notin \mathring{\mathcal{C}}_+$.

Node e: If $U \cap \Delta_u \cap \Delta_v = \emptyset$, that is, if U contains no intersection point between Δ_u and Δ_v , we match 5 cases depending on the position of T in U , and of its corners with respect to \mathcal{C} . They are illustrated in Figure 8.

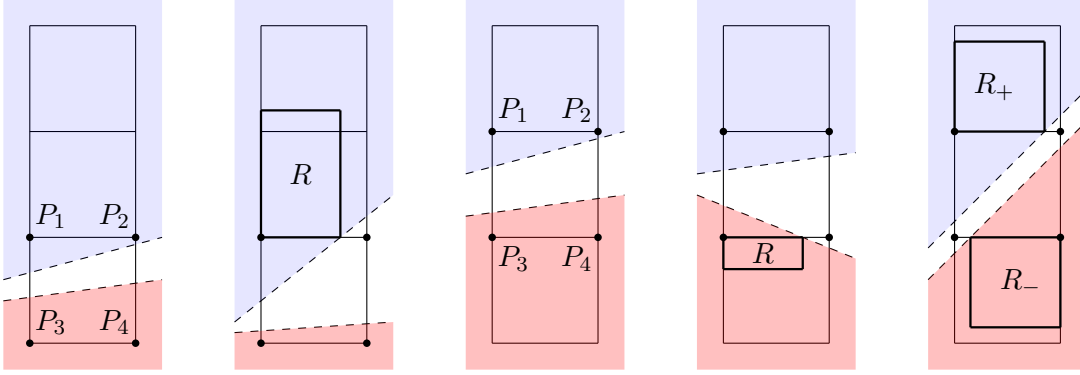


Figure 8: Cases 6, 7, 8, 9 and 10

Case 6: If T is the bottom cell and $P_1, P_2 \in \mathcal{C}_+$, then the two other cells are included in \mathcal{C}_+ , so

$$|U \cap \mathcal{C}| \geq 2h^2 + |T \cap \mathcal{C}| \geq 3|T \cap \mathcal{C}| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

Case 7: If T is the bottom cell and $P_1 \in \mathcal{C}_+$ but $P_2 \notin \mathcal{C}_+$, $T \cap \mathcal{C}_+$ is a triangle of width and height at most h , so there is a rectangle $R \subset (U \setminus T) \cap \mathcal{C}_+$ of same width and twice as high, and thus

$$|U \cap \mathcal{C}| \geq |R| + |T \cap \mathcal{C}| = 4|T \cap \mathcal{C}_+| + |T \cap \mathcal{C}| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

The same argument holds when $P_2 \in \mathcal{C}_+$ but $P_1 \notin \mathcal{C}_+$, and we necessarily have P_1 or P_2 in \mathcal{C}_+ since $T \cap \mathcal{C}_+ \neq \emptyset$. If T is the top cell, applying a symmetry with respect to the horizontal axis and exchanging \mathcal{C}_+ with \mathcal{C}_- brings us back to Cases 6 and 7.

Case 8: If T is the central cell, $P_1, P_2 \in \mathcal{C}_+$ and $P_3, P_4 \in \mathcal{C}_-$ the two other cells are included in \mathcal{C}_+ and \mathcal{C}_- , and we conclude as in Case 6.

Case 9: If T is the central cell, $P_1, P_2 \in \mathcal{C}_+$, $P_3 \in \mathcal{C}_-$ but $P_4 \notin \mathcal{C}_-$, the top cell is included in \mathcal{C}_+ , and there is a rectangle $R \subset \mathcal{C}_-$ of same width and height as $T \cap \mathcal{C}_-$ in the bottom cell, so

$$|U \cap \mathcal{C}| \geq h^2 + |T \cap \mathcal{C}| + |R| \geq 2|T \cap \mathcal{C}| + 2|T \cap \mathcal{C}_-| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

The same situation occurs when only three points among P_1, \dots, P_4 are in \mathcal{C} .

Case 10: If T is the central cell, only one vertex among P_1, P_2 is in \mathcal{C}_+ , and only one among P_3, P_4 is in \mathcal{C}_- , both $T \cap \mathcal{C}_+$ and $T \cap \mathcal{C}_-$ are triangles, and there exist rectangles R_+ and R_- of same widths and heights, so

$$|U \cap \mathcal{C}| \geq |R_+| + |T \cap \mathcal{C}| + |R_-| \geq 3|T \cap \mathcal{C}_+| + 3|T \cap \mathcal{C}_-| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

As \mathcal{C}_+ and \mathcal{C}_- each contain at least one corner of T , we treated all cases for Node e.

Node f: Finally, we consider the situation where there is an intersection point $X \in \Delta_u \cap \Delta_v$ in U , and therefore in T . We again match 5 cases, illustrated in Figure 9, depending on the position of T in U , and of its corners with respect to \mathcal{C} .

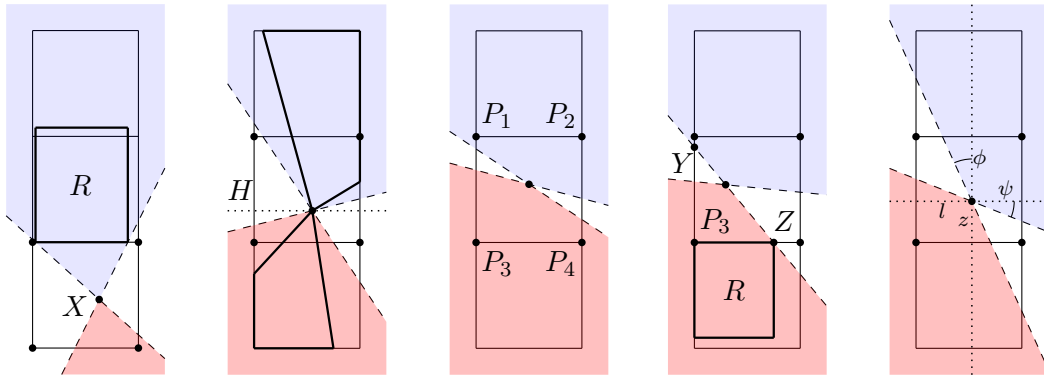


Figure 9: Cases 11, 12, 13, 14 and 15

Case 11: If T is the bottom cell, as Δ_u and Δ_v pass through the central cell of S , U is included in the central column of S , and no corner of T can be in \mathcal{C}_+ , since otherwise Δ would have to pass through that corner, according to the definition of the columns. As a consequence, Δ_u and Δ_v necessarily pass through the central cell of U , so $T \cap \mathcal{C}_+$ is a triangle, and we proceed as in Case 7. The same happens if T is the top cell, so in the rest of the proof we only consider situations where T is the central cell.

Case 12: If the horizontal line H passing through X does not intersect \mathcal{C} at any other point, \mathcal{C}_+ is entirely above H and \mathcal{C}_- entirely below. Denoting $z = X_2 - \bar{x}_2 + \frac{h}{2} \in (0, h)$, the vertical dilation with respect to H by a factor $\frac{2h-z}{h-z}$ sends $T \cap \mathcal{C}_+$ in $U \cap \mathcal{C}_+$, and the vertical dilation with respect to H by a factor $\frac{h+z}{z}$ sends $T \cap \mathcal{C}_-$ in $U \cap \mathcal{C}_-$, so

$$|U \cap \mathcal{C}| \geq \frac{2h-z}{h-z} |T \cap \mathcal{C}_+| + \frac{h+z}{z} |T \cap \mathcal{C}_-| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|)$$

because $\frac{2h-z}{h-z} + \frac{h+z}{z} = 2 + \frac{h^2}{z(h-z)} \geq 6$ for $z \in (0, h)$.

In the remaining cases, up to a symmetry with respect to the vertical axis, we can assume that $X + \mathbb{R}_+^2 \subset \mathcal{C}_+$ and $X + \mathbb{R}_-^2 \subset \mathcal{C}_-$, and in particular $P_2 \in \mathcal{C}_+$ and $P_3 \in \mathcal{C}_-$.

Case 13: If $P_1 \in \mathcal{C}_+$ and $P_4 \in \mathcal{C}_-$, the situation is similar to Case 8.

Case 14: If $P_1 \in \mathcal{C}_+$ and $P_4 \notin \mathcal{C}_-$, the top cell is included in \mathcal{C}_+ , and one of the lines Δ_u or Δ_v intersects the line segments $[P_1, P_3]$ and $[P_3, P_4]$ at points Y and Z . Then the triangle YP_3Z is included in T and contains $T \cap \mathcal{C}_-$, so there is a rectangle R of same width and height in $(U \setminus T) \cap \mathcal{C}_-$. In the end

$$|U \cap \mathcal{C}| \geq h^2 + |T \cap \mathcal{C}| + |R| \geq 2|T \cap \mathcal{C}| + 2|T \cap \mathcal{C}_-| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|).$$

The same approach treats the symmetric case $P_1 \notin \mathcal{C}_+$ and $P_4 \in \mathcal{C}_-$,

Case 15: Finally, if $P_1 \notin \mathcal{C}_+$ and $P_4 \notin \mathcal{C}_-$, denote $l = X_1 - \bar{x}_1 + \frac{h}{2} \in (0, h)$, $z = X_2 - \bar{x}_2 + \frac{h}{2} \in (0, h)$, $\phi \in (0, \frac{\pi}{4})$ the angle between the vertical axis and the line among Δ_u and Δ_v that intersects $[P_1, P_2]$, and $\psi \in (0, \frac{\pi}{4})$ the angle between the line among Δ_u and Δ_v that intersects $[P_1, P_3]$ and the horizontal axis. As $|\arg(\Delta)| \leq \frac{\pi}{4}$, $\phi \geq \psi$ so $\tan(\psi) \leq \tan(\phi) =: t \leq 1$.

We can now compute

$$\begin{aligned} |T \cap \mathcal{C}_+| &= (h-l)(h-z) + \frac{1}{2}(h-l)^2 \tan \psi + \frac{1}{2}(h-z)^2 \tan \phi, \\ |T \cap \mathcal{C}_-| &= lz + \frac{1}{2}l^2 \tan \psi + \frac{1}{2}z^2 \tan \phi, \end{aligned}$$

and

$$|(U \setminus T) \cap \mathcal{C}| \geq (h-l)h + (h-z)th + lh + zth = (1+t)h^2.$$

If $l+z \leq h$, we get

$$|(U \setminus T) \cap \mathcal{C}| \geq (1+t)(l+z)^2 - (1-t)(l-z)^2 = 4lz + 2t(l^2 + z^2) \geq 4|T \cap \mathcal{C}_-|.$$

Similarly, $l+z \geq h$ implies $|(U \setminus T) \cap \mathcal{C}| \geq 4|T \cap \mathcal{C}_+|$. In any case, we found

$$|U \cap \mathcal{C}| = |T \cap \mathcal{C}| + |(U \setminus T) \cap \mathcal{C}| \geq 6 \min(|T \cap \mathcal{C}_+|, |T \cap \mathcal{C}_-|),$$

which concludes the proof.

As a last remark, note that the constants $\alpha = h^{-2}$ and $\mu = \frac{3}{2}h^2$ in Proposition 5.1 are sharp, since equality is attained by functions of constant sign on each cell for α , and by $w = u - v$ with $\arg(\vec{n}_u) \in \frac{\pi}{4}\mathbb{Z}$, $c_u = 0$ and $v = u - 1$ for μ .

References

- [1] B. Adcock, A. C. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM Journal on Mathematical Analysis*, 45(5):3132–3167, 2013.
- [2] F. Arandiga, A. Cohen, R. Donat, and N. Dyn. Interpolation and approximation of piecewise smooth functions. *SIAM Journal on Numerical Analysis*, 43(1):41–57, 2005.
- [3] J.-P. Argaud, B. Bouriquet, F. de Caso, H. Gong, Y. Maday, and O. Mula. Sensor placement in nuclear reactors based on the generalized empirical interpolation method. *Journal of Computational Physics*, 363:354 – 370, 2018.

- [4] B. Battisti, T. Blickhan, G. Enchery, V. Ehrlacher, D. Lombardi, and O. Mula. Wasserstein model reduction approach for parametrized flow problems in porous media. working paper or preprint, May 2022.
- [5] P. Benner, A. Cohen, M. Ohlberger, and K. Willcox. *Model Reduction and Approximation: Theory and Algorithms*, volume 15. SIAM, 2017.
- [6] P. Berger, K. Gröchenig, and G. Matz. Sampling and reconstruction in distinct subspaces using oblique projections. *arXiv preprint arXiv:1706.06444*, 2017.
- [7] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Data assimilation in reduced modeling. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):1–29, 2017.
- [8] P. Binev, A. Cohen, O. Mula, and J. Nichols. Greedy algorithms for optimal measurements selection in state estimation using reduced models. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1101–1126, 2018.
- [9] B. Bojanov. Optimal recovery of functions and integrals. In *First European Congress of Mathematics*, pages 371–390. Springer, 1994.
- [10] Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.*, 59(8):1207–1223, 2006.
- [11] Michael Capalbo, Omer Reingold, Salil Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders. In *Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*, pages 659–668, 2002.
- [12] A. Cohen, W. Dahmen, and R. DeVore. Compressed sensing and best k-term approximation. *Journal of the American mathematical society*, 22(1):211–231, 2009.
- [13] A. Cohen, W. Dahmen, R. DeVore, J. Fadili, O. Mula, and J. Nichols. Optimal reduced model algorithms for data-based state estimation. *SIAM Journal on Numerical Analysis*, 58(6):3355–3381, 2020.
- [14] A. Cohen, W. Dahmen, O. Mula, and J. Nichols. Nonlinear reduced models for state and parameter estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):227–267, 2022.
- [15] A. Cohen and R. DeVore. Approximation of high-dimensional parametric pdes. *Acta Numerica*, 24:1–159, 2015.
- [16] A. Cohen, R. DeVore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE’s. *Analysis and Applications*, 09(01):11–47, 2011.
- [17] R. DeVore, G. Petrova, and P. Wojtaszczyk. Greedy algorithms for reduced bases in Banach spaces. *Constructive Approximation*, 37(3):455–466, 2013.
- [18] R.A. DeVore. Nonlinear approximation. *Acta numerica*, 7:51–150, 1998.
- [19] J. L. Eftang, A. T. Patera, and E. M. Rönquist. An ”hp” certified reduced basis method for parametrized elliptic partial differential equations. *SIAM Journal on Scientific Computing*, 32(6):3170–3200, 2010.

- [20] V. Ehrlacher, D. Lombardi, O. Mula, and F.-X. Vialard. Nonlinear model reduction on metric spaces. application to one-dimensional conservative pdes in wasserstein spaces. *ESAIM M2AN*, 54(6):2159–2197, 2020.
- [21] Simon Foucart and Holger Rauhut. An invitation to compressive sensing. In *A mathematical introduction to compressive sensing*, pages 1–39. Springer, 2013.
- [22] F. Galarce, D. Lombardi, and O. Mula. State estimation with model reduction and shape variability. application to biomedical problems. *SIAM Journal on Scientific Computing*, 44(3):B805–B833, 2022.
- [23] C. Greif and K. Urban. Decay of the kolmogorov n-width for wave problems. *Applied Mathematics Letters*, 96:216–222, 2019.
- [24] J. K. Hammond, R. Chakir, F. Bourquin, and Y. Maday. PBDW: A non-intrusive reduced basis data assimilation method and its application to an urban dispersion modeling framework. *Applied Mathematical Modelling*, 76:1–25, 2019.
- [25] J. S. Hesthaven, G. Rozza, and B. Stamm. Certified reduced basis methods for parametrized partial differential equations. *SpringerBriefs in Mathematics*, 2015.
- [26] T. Hrycak and K. Gröchenig. Pseudospectral fourier reconstruction with the modified inverse polynomial reconstruction method. *Journal of Computational Physics*, 229(3):933–946, 2010.
- [27] Y. Maday, A. T. Patera, J. D. Penn, and M. Yano. A parameterized-background data-weak approach to variational data assimilation: formulation, analysis, and application to acoustics. *International Journal for Numerical Methods in Engineering*, 102(5):933–965, 2015.
- [28] Y. Maday and B. Stamm. Locally adaptive greedy approximations for anisotropic parameter reduced basis spaces. *SIAM J. Scientific Computing*, 35(6):A2417–A2441, 2013.
- [29] Yuly Makovoz. Random approximants and neural networks. *Journal of Approximation Theory*, 85(1):98–109, 1996.
- [30] C. A. Micchelli and Th. J. Rivlin. *A survey of optimal recovery*. Springer, 1977.
- [31] O. Mula. Inverse problems: A deterministic approach using physics-based reduced models. *Submitted*, 2022.
- [32] E. Novak and H. Wozniakowski. Tractability of multivariate problems, volume i: Linear information, european math. *Soc., Zürich*, 2(3), 2008.
- [33] M. Ohlberger and S. Rave. Reduced basis methods: Success, limitations and future challenges. In *Proceedings of the Conference Algorithm*, pages 1–12, 2016.
- [34] James Edward Pilliod. *An analysis of piecewise linear interface reconstruction algorithms for volume-of-fluid methods*. U. of Calif., Davis, 1992.
- [35] James Edward Pilliod Jr and Elbridge Gerry Puckett. Second-order accurate volume-of-fluid algorithms for tracking material interfaces. *Journal of Computational Physics*, 199(2):465–502, 2004.

- [36] Elbridge Gerry Puckett. A volume-of-fluid interface tracking algorithm with applications to computing shock wave refraction. In *proceedings of the fourth international symposium on Computational Fluid Dynamics*, pages 933–938, 1991.
- [37] G. Rozza, D. B. P. Huynh, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):1, 9 2007.
- [38] Jonathan W Siegel and Jinchao Xu. Sharp bounds on the approximation rates, metric entropy, and n -widths of shallow neural networks. *arXiv preprint arXiv:2101.12365*, 2021.
- [39] H. Tran, C. G. Webster, and G. Zhang. Analysis of quasi-optimal polynomial approximations for parameterized pdes with deterministic and stochastic coefficients. *Numerische Mathematik*, 137(2):451–493, 2017.
- [40] G. Welper. Transformed snapshot interpolation. *arXiv preprint arXiv:1505.01227*, 2015.
- [41] Z. Zou, D. Kouri, and W. Aquino. An adaptive local reduced basis method for solving pdes with uncertain inputs and evaluating risk. *Computer Methods in Applied Mechanics and Engineering*, 345:302–322, 2019.