Detecting stable scales in images via non-smooth K-functionals

Triet M. Le* and Luke Rogers†

UCLA CAM Report 10-63

Abstract

In image decompositions, one is interested in decomposing f into $u+v$ where u and v have different features. In a variational approach, such a decomposition is achieved by solving the following variational problem

$$\inf_{(u,v)\in X_1\times X_2} \{tF_1(u)+F_2(v) : f = u+v\},$$

where F_1 and F_2 are positive functionals defined on some function spaces X_1 and X_2 respectively. Often the scale parameter t is fixed a prior. In this paper, we address the problem of selecting the scale parameter t in a multiscale fashion, and introduce the notion of interpolating scales that are stable with respect to the functional or energy being minimized. The motivation is coming from Peetre's K-functional in interpolation theory.

1 Introduction

In image decompositions ([26], [22], [25], [4], [9], [17], [15], [14], among others), a given image f is decomposed into $u+v$ where u and v belong to some function spaces X_1 and X_2 respectively. The choice of X_1 and X_2 determines the smoothing properties of u and v. In a variational approach, f is decomposed into $u(t)+v(t)$ via the variational problem

$$\inf_{(u,v)\in X_1\times X_2} \{tF_1(u)+F_2(v) : f = u+v\}, \tag{1}$$

for some parameter $t>0$. Here, F_1 and F_2 are positive functionals defined on some function spaces X_1 and X_2 respectively. We are interested in the study of selecting the parameter t in the above minimization problem. For each $t>0$, suppose there exists

*Department of Mathematics, University of Pennsylvania, David Rittenhouse Lab. 209 South 33rd Street Philadelphia, PA 19104-6395 (trietle@math.upenn.edu).

†Department of Mathematics, University of Connecticut, 196 Auditorium Rd., U-3009, Storrs, CT 06269-3009 (rogers@math.uconn.edu).

$(u(t), v(t)) \in X_1 \times X_2$ that is a minimizer of (1). Then $\{u(t)\}_{t>0}$ can be viewed as a multi-scale representation of f and the above functional can be viewed as diffusion (linear or nonlinear) process. The parameter t is often referred to as the global scale parameter. The study of scale parameters (local or global) in images has been shown to be very useful as a tool for feature extraction with applications to object recognition, image segmentation and decompositions, among others.

We first recall a few methods for extracting local features.

Definition 1 (**Dilating Consistency Property).** We say the local scales have the dilating consistency property if the number of local scales of f at x is invariant under dilation. Moreover, denote by $T_f(x)$ the set of local scales of f at x. Then it holds that

$$T_{d_\delta f}(x) = \{\delta^s t : t \in T_f(x)\}, \text{ for some } s < 0. \tag{2}$$

T. Lindeberg [18], [18], [19] represents $u(t)$ as a convolution of the Gaussian kernel K_t with f. Special scale-space features are defined as the set consisting of points $(x, t) \in \mathbb{R}^n \times \mathbb{R}^+$ such that $t\frac{\partial u}{\partial t}(x, t)$ is a local max or local min. In [20], D. Lowe developed the SIFT detector by incorporating the vector consisting of partial derivatives of u at the scale-space feature locations. This detector has shown to be very effective in distinguishing features and objects in images.

Brox-Weickert in [7] study local scales of f at x using the total variation flow, where $\{u(t)\}_{t>0}$ are solutions of the evolution equation

$$u_t = \operatorname{div}(\frac{\nabla u}{|\nabla u|}), \; u(0) = f.$$

For each $x \in \mathbb{R}^n$, they define the local scale $m(x)$ of f at x (single scale) using the average change of $|\partial_t u|$. In other words,

$$m_f(x) = \frac{2T}{\int_0^T |\partial_t u| \; dt},$$

where T is a tuning parameter. They provide texture segmentation as an application with very good segmentation results. However, there are two draw backs to this approach: The parameter T is user defined, and the dilating consistency property is not satisfied. In some way, the local scales depend on the choice of T which is a global parameter.

In [27], Strong-Chan proposed to study scales using the Rudin-Osher-Fatemi diffusion [26]. Fix a $T > 0$, let u be the minimizer of $\mathcal{J}$ given by

$$\mathcal{J}(u) = \int |Du| + \frac{1}{2T}\|f - u\|_{L^2}^2.$$

Suppose $f = \chi_{B_r}$, where B is the ball of radius r centered at the origin. Let the scale $s(x)$ of $x \in B$ defined as

$$s(x) = \frac{|B_r|}{P(B_r)} = \frac{\|\chi_{B_r}\|_{L^1}}{\|\chi_{B_r}\|_{\dot{BV}}} \tag{3}$$

Then it can be shown [27] that for small T (which depends on the G-norm of f [22]) the minimizer u has the same discontinuity set as f and for all $x \in B_r$,

$$|u(x) - f(x)| = \frac{T}{s(x)}.$$

Motivated from this example, the local scale of a general f at x is defined [27] as

$$s(x) = \frac{T}{|u(x) - f(x)|}.$$

This definition of local scales depend on T (and hence u). The minimizer u can be viewed as a "denoised" shape that contains x.

In [21], Luo-et al study local scales by looking at a collection of embedding shapes that contain x and look for a shape that is most contrasted. Then the local scale of f at x is defined as in (3) replacing B_r with the most contrasted shape. In other words, let $\{f_1(x), \cdots, f_{N(x)}(x)\}$ be a collection of shapes (connected components) that contain x with the property that $f_i(x) \subset f_{i+1}(x)$ for all $i = 1, \cdots, N(x)$. $f_i(x)$ is defined as the connected component that contains x such that

$$f_i(x) \subset \{x \in \mathbb{R}^n : f(x) \geq \lambda_i\}, \text{ for some } \lambda_i > 0.$$

Denote by $I(f_i) = \lambda_i$, and define $C(f_i) = |I(f_{i+1}) - I(f_i)|$. Then the most contrasted shape containing x is defined as

$$\hat{f}(x) = \underset{f_i, i=1,\cdots,N(x)}{\arg\max} \; C(f_i).$$

However as commented by the authors in [21], for natural images $C(f_i)$ is typically equal to 1 in the presence of noise and blur. Thus, the most contrasted shape $\hat{f}(x)$ is meaningless. To overcome this problem, they propose to sum up the contrasts of shapes as follows. For $i \geq 1$, the accumulated contrasts are recursively computed as

$$\overline{C}(f_i(x)) = \begin{cases} \overline{C}(f_{i-1}(x)) + C(f_i(x)) & \text{if} \quad |f_i(x)| - |f_{i-1}(x)| \leq TP(f_{i-1}(x)), \\ C(f_i(x)) & \text{otherwise,} \end{cases} \tag{4}$$

where T is some fixed (global) parameter and $P(f_i)$ denotes the perimeter of the shape f_i. The relation

$$|f_i(x)| - |f_{i-1}(x)| \leq TP(f_{i-1}(x)) \tag{5}$$

is not invariant under dilation. In other words, if f is dilated by a factor $\delta > 0$, the equation (5), applying to the dilated shapes, becomes (assuming in $\mathbb{R}^2$)

$$(|f_i(x)| - |f_{i-1}(x)|) \leq (\delta T)P(f_{i-1}(x)).$$

Note that the contrasts in the image is invariant under dilation. Thus the definition of $\overline{C}(f_i(x))$ defined in (4) does not satisfy the dilating consistency property.

In these methods, one obtain a single meaningful scale at each location x. Realizing that x may be embedded in multiple scales, P. Jones and the first author [16] recently propose a new method for extracting a vector of scales at each location x. This notion of local scales is further characterized based on the visibility level of the scales and their separation from other scales. The local scales of f at x is computed by measuring the deviation of f from a linear function near x at different scales t's. This multiscale analysis is intimately related to the theory of wavelets. Let $\phi(x)$ be any approximation to the identity. In particular, consider

$$\phi(x) = e^{-\pi|x|^2}.$$

Here we have $\int_{\mathbb{R}^n} \phi(x)\, dx = 1$. For each $t > 0$, define

$$K_t(x) = t^{-n/2}\phi(x/\sqrt{t}) = (t)^{-n/2} e^{-\pi\frac{|x|^2}{t}}.$$

Let $\psi_t(x) = t\frac{\partial K_t}{\partial t}$, $x \in \mathbb{R}^n$, and define

$$Sf(x,t) = \psi_t * f(x) = \int_{\mathbb{R}^n} \psi_t(x-y) f(y)\, dy. \tag{6}$$

Since ψ_t has zero mean and zero first moments, we see that if f is linear, then $Sf(x,t) = 0$ for all $t > 0$ and $x \in \mathbb{R}^n$. Thus the quantity $|Sf(x,t)|$ measures how f deviates from a linear function of scale t near x. Note that Sf is invariant under addition by a linear function. The local scales of f at x are defined as the set $T_f(x)$ consisting of $t > 0$ such that $|Sf(x,t)|$ is a local maxima. It can be shown that the set of local scales $T_f(x)$ satisfies the dilating consistency property. This notion of local scales is further characterized based on the visibility level of the scales and their separation from other scales.

Passing from local to global, one can consider the global Sf defined by

$$Sf(t) = \|\psi_t * f(x)\|_{L^p} = \left\|\frac{\partial u}{\partial t}\right\|_{L^p}, \text{ for some } 1 \leq p < \infty, \tag{7}$$

or using some weighted L^p spaces. Now the global scales can be defined as the local maxima of $Sf(t)$.

See also Tadmor-Nezzar-Vese [28], Aujol-et al [5], and Vixie-et al [30] for other notions of global scales. In Section 4, we will provide a discussion comparing our method with that of Vixie-et al [30].

In this paper, instead of using (7) to analyze the global scales of f, we approach this problem from an interpolation theory point of view. In Section 4, we will compare the method using (7) with the proposed method for non-smooth functionals.

We recall the following definition of function spaces and their properties from [6] and [29]. Let X_1 and X_2 be two Banach spaces that are continuously embedded in some Hausdorff topological vector space $\mathcal{X}$. If $X_1 \subset X_2$, then one can take $\mathcal{X} = X_2$. We say $f \in X_1 + X_2$, if

$$\|f\|_{X_1+X_2} = \inf_{(u,v)\in X_1\times X_2} \{\|u\|_{X_1} + \|v\|_{X_2} : f = u + v\} < \infty.$$

For each $f \in X_1 + X_2$ and $t > 0$, the K-functional and the L-functional are defined as

$$K(f,t,X_2,X_1) = \inf_{(u,v)\in X_1\times X_2} \{t\|u\|_{X_1} + \|v\|_{X_2} : f = u + v\}, \tag{8}$$

and for any fixed $1 \le p_1, p_2 < \infty$,

$$L(f,t,X_2,X_1) = \inf_{(u,v)\in X_1\times X_2} \left\{ t\|u\|_{X_1}^{p_1} + \|v\|_{X_2}^{p_2} : f = u + v \right\}. \tag{9}$$

Note that L also depends on p_1 and p_2, but for simplicity we remove them from the variables of L. From now on, p_1 and p_2 are assumed to be fixed values in $[1, \infty)$.

It is easy to see that for each positive t, $t\|u\|_{X_1}$ is an equivalent norm in X_1. In other words,

$$\min(t,1)\|u\|_{X_1} \le \|u\|_{X_1} \le \max(t,1)\|u\|_{X_1}.$$

This shows that $K(f,t,X_2,X_1)$ gives an equivalent norm for $X_1 + X_2$, for all $t > 0$. The limiting spaces $\infty X_1 + X_2$ and $X_1 + \infty X_2$ are defined as

$$\infty X_1 + X_2 = \{f \in X_1 + X_2 : \|f\|_{\infty X_1 + X_2} := \lim_{t\to\infty} K(f,t,X_2,X_1) < \infty\}, \text{ and}$$
$$X_1 + \infty X_2 = \{f \in X_1 + X_2 : \|f\|_{X_1 + \infty X_2} := \lim_{t\to\infty} K(f,t,X_1,X_2) < \infty\}.$$

In real interpolation theory, the space $\infty X_1 + X_2$ is known as the Gagliardo completion of X_2 in $X_1 + X_2$ with

$$X_2 \subset \infty X_1 + X_2 \subset X_1 + X_2.$$

Similarly the space $X_1 + \infty X_2$ is the Gagliardo completion of X_1 in $X_1 + X_2$ with

$$X_1 \subset X_1 + \infty X_2 \subset X_1 + X_2.$$

Note that since

$$K(f,t,X_1,X_2) = tK(f,t^{-1},X_2,X_1),$$

the norm in $X_1 + \infty X_2$ can be written as

$$\|f\|_{X_1+\infty X_2} = \lim_{t\to 0} t^{-1} K(f,t,X_2,X_1).$$

Next, we would like to recall the real interpolating spaces which will be crucial in our study of scales. For $\theta \in (0,1)$, the Banach interpolating space $(X_2, X_1)_{\theta,\infty}$ ([6], [29]), is defined as

$$(X_2,X_1)_{\theta,\infty} = \left\{ f \in X_1 + X_2 : \|f\|_{(X_2,X_1)_{\theta,\infty}} = \sup_{0<t<\infty} t^{-\theta} K(f,t,X_2,X_1) < \infty \right\}. \tag{10}$$

Equivalently, these interpolating function spaces can be defined in terms of L-functional defined in (9). Let $\gamma \in (0,1)$ and $1 \le p \le \infty$ be related by

$$p = \gamma p_1 + (1-\gamma)p_2, \text{ and } \theta = \gamma \frac{p_1}{p}. \tag{11}$$

Then the Banach interpolating space $(X_2, X_1)_{\theta,\infty}$ can be defined in terms of the L-functional as

$$(X_2,X_1)_{\theta,\infty} = \left\{ f \in X_1 + X_2 : \|f\|_{(X_2,X_1)_{\theta,\infty}} = \left[\sup_{0<t<\infty} t^{-\gamma} L(f,t,X_2,X_1) \right]^{1/p} < \infty \right\}. \tag{12}$$

The Banach spaces $(X_2, X_1)_{\theta,\infty}$ are intermediate spaces between $X_1 \cap X_2$ and $X_1 + X_2$. In other words, for all $0 < \theta < 1$,

$$X_1 \cap X_2 \subset (X_2, X_1)_{\theta,\infty} \subset X_1 + X_2. \tag{13}$$

If in addition, $X_1 \subset X_2$, then for all $0 < \theta_1 < \theta_2 < 1$,

$$(X_2, X_1)_{\theta_1,\infty} \subset (X_2, X_1)_{\theta_2,\infty}. \tag{14}$$

The study of selecting the parameters t's in (8) and (9) is amount to finding $u(t)$'s which give an efficient multiscale representation of f. For each $f \in X_1 + X_2$ and $\theta, \gamma \in (0, 1)$ and $t > 0$, we define

$$\mathcal{K}f(t, \theta) = t^{-\theta} K(f, t). \tag{15}$$

and

$$\mathcal{L}f(t, \theta) = t^{-\gamma} L(f, t). \tag{16}$$

In this paper, we are interested in analyzing the properties of $\mathcal{K}(t, \theta)$ and $\mathcal{L}(t, \theta)$ which we claim to contain information about scales. We also introduce a notion of stability for scales coming from these diffusions.

2 Interpolating scales

When there is no need to address X_1 and X_2 explicitly, for simplicity, we write

$$K(f, t) = K(f, t, X_2, X_1), \; L(f, t) = L(f, t, X_2, X_1), \text{ etc.}$$

First we would like to show some basic properties for $K(f, t)$ and $L(f, t)$. The following Lemma provides some useful regularity properties and its proof is outlined in [29] and [6]. To make the paper self contained, we provide its proof here.

Lemma 1. *For each $t > 0$, the functional $K(f, t)$ is increasing, continuous, and concave. Moreover, $K(f, t)$ is differentiable almost everywhere, and the points where the derivative exists satisfies*

$$0 \le \frac{dK(f, t)}{dt} \le \frac{K(f, t)}{t}. \tag{17}$$

The same results also hold for the $L(f, t)$.

Proof. Let $0 < t_1 \le t_2 < \infty$, and let $f = u + v$ be any decomposition with $u \in X_1$ and $v \in X_2$. Then

$$K(f, t_1) \le t_1 \|u\|_{X_1} + \|v\|_{X_1} \le t_2 \|u\|_{X_1} + \|v\|_{X_1}.$$

Taking the infimum over all decompositions $f = u + v$, we have

$$K(f, t_1) \le K(f, t_2),$$

which shows that $K(f, t)$ is increasing. In particular,

$$t_1 K(f, t_2) \le t_1 \left[t_2 \|u\|_{X_1} + \|v\|_{X_2} \right] \le t_2 \left[t_1 \|u\|_{X_1} + \|v\|_{X_2} \right].$$

Since the decomposition $f = u + v$ is arbitrary, this implies

$$t_1 K(f, t_2) \leq t_2 K(f, t_1). \tag{18}$$

Hence,

$$0 \leq K(f, t_2) - K(f, t_1) \leq \frac{t_2 - t_1}{t_1} K(f, t_1) \;\Rightarrow\; \frac{K(f, t_2) - K(f, t_1)}{t_2 - t_1} \leq \frac{K(f, t_1)}{t_1},$$

which shows (17) by taking $t_2 \to t_1$ and replacing t_1 with t. To show the concavity of K, let $t_1, t_2 > 0$ and $t = \alpha_1 t_1 + \alpha_2 t_2$ where $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$. Let $f = u + v$ be any decomposition of f with $u \in X_1$ and $v \in X_2$. Then

$$\begin{aligned}\alpha_1 K(f, t_1) + \alpha_2 K(f, t_2) &\leq \alpha_1 (t_1 \|u\|_{X_1} + \|v\|_{X_2}) + \alpha_2 (t_2 \|u\|_{X_1} + \|v\|_{X_2}) \\ &= (\alpha_1 t_1 + \alpha_2 t_2) \|u\|_{X_1} + (\alpha_1 + \alpha_2) \|v\|_{X_2} = t \|u\|_{X_1} + \|v\|_{X_2},\end{aligned}$$

which holds for all decompositions $f = u + v$. Taking the infimum over all decompositions $f = u + v$, we have

$$\alpha_1 K(f, t_1) + \alpha_2 K(f, t_2) \leq K(f, t).$$

This shows that K is concave. A concave and increasing function must be continuous. Moreover, a concave function is both left and right differentiable and hence differentiable almost everywhere. □

For each $f \in X_1 + X_2$, a direct consequence of Lemma 1 is the following property for $\mathcal{K}(t, \theta)$.

Corollary 1. *For each $\theta \in (0, 1)$, $\mathcal{K}f(t, \theta)$ (defined in (15)) is continuous and differentiable almost everywhere. Moreover suppose $f \in (X_2, X_1)_{\theta,\infty}$, then at the points where the derivative exists, we have*

$$\left| \frac{\partial}{\partial t} \mathcal{K}(t, \theta) \right| \leq t^{-\theta-1} (1 + \theta) K(f, t) \leq \frac{1 + \theta}{t} \|f\|_{(X_2, X_1)_{\theta,\infty}}. \tag{19}$$

By a change of variable let $\tau = \log_a t$ for some $a > 1$, then we have

$$\left| \frac{\partial}{\partial \tau} \mathcal{K}(\tau, \theta) \right| \leq C \|f\|_{(X_2, X_1)_{\theta,\infty}}, \text{ with } C = \frac{2}{\ln(a)}. \tag{20}$$

The same results also hold for $\mathcal{L}(t, \theta)$.

Remark 1. Before defining interpolating scales, we would like to make a heuristic remark on the interpolating space $(X_2, X_1)_{\theta,\infty}$. Suppose $X_1 = \dot{BV}$ and $X_2 = L^p$, $1 \leq p < \infty$. For the image f in figures 1 and 2, we expect that $f \in (L_p, \dot{BV})_{\theta_0,\infty}$ for some θ_0 close to 0, which shows that $(L_p, \dot{BV})_{\theta_0,\infty} \approx \dot{BV}$. Here we use the notation "≈" to mean "approximating". On the other hand, for an image f in figures 3 and 4, we expect that $f \in (L_p, \dot{BV})_{\theta_1,\infty}$ for some θ_1 close to 1, which indicates that $(L_p, \dot{BV})_{\theta_1,\infty} \approx L^p$. Thus the regularity of f determines the parameter θ. Recall we have

$$\|f\|_{(L_p, \dot{BV})_{\theta,\infty}} = \sup_{t>0} \left\{ \mathcal{K}(t, \theta) = t^{-\theta} K(f, t, L^p, \dot{BV}) \right\},$$

where $K(f,t,L^p,\dot{BV})$ is concave and a.e. differentiable. The weighted function $t^{-\theta}$, for $\theta \in (0,1)$, determines the prominent scale $t^* = \arg\max_{t>0} \mathcal{K}(t,\theta)$. Small (large) θ penalizes large (small) scales less than large (small) θ. Here, we assume that the readers are familiar with the homogeneous space $\dot{BV}$ of bounded variation. The norm $\|u\|_{\dot{BV}}$ is also known as the total variation of u. We refer to [3] for an in depth discussion of the space BV and its homogeneous version $\dot{BV}$.

From now on, we denote by $Sf(t,\theta)$ either $\mathcal{K}(t,\theta)$ or $\mathcal{L}(t,\theta)$. Suppose $f \in (X_2,X_1)_{\theta,\infty}$ for all $\theta \in (0,1)$. Thus $Sf(t,\theta)$ is bounded for all $\theta \in (0,1)$. Motivated from the previous remark, we have the following definition for interpolating scales.

Definition 2. Let X_1 and X_2 be Banach spaces. Suppose $f \in X_1 + X_2$. Let $Sf(t,\theta)$ be $\mathcal{K}(t,\theta)$ defined in (15) (or $\mathcal{L}(t,\theta)$ defined in (16)). For each fixed θ, denote by $T_{f,\theta}$ the set of $t > 0$ values at which $Sf(t,\theta)$ is a local maximum (as a function of t).

- The *(global) interpolating scales* of f are defined as $T_f = \cup_{\theta\in(0,1)} T_{f,\theta}$.
- For each $t \in T_f$, we say t is a *stable interpolating scale* if there exists $I \subset (0,1)$ such that $|I| > 0$ and $t \in T_{f,\theta}$ for all $\theta \in I$. Here, $|I|$ denotes the Lebesgue measure (length) of the set I in $\mathbb{R}$.
- For each $t \in T_f$, denote by $I_{f,t}$ the maximal open subset in $(0,1)$ such that $t \in T_{f,\theta}$ for all $\theta \in I_{f,t}$. We say $t \in T_f$ is *δ-absolutely-stable* if $|I_{f,t}| > \delta$.
- Let $\delta^* = \sup\{|I_{f,t}| : t \in T_f\}$. Suppose $\delta^* > 0$. Then we say $t \in T_f$ is *δ-relatively-stable* if $|I_{f,t}|/\delta^* > \delta$.

Remark 2. Consider the K-functional given by

$$K(f,t,L^p,\dot{BV}) = \inf_{(u,v)\in \dot{BV}\times L^p} \left\{ t\|u\|_{\dot{BV}} + \|v\|_{L^p} : f = u + v \right\},$$

and for all $\theta \in (0,1)$ we have

$$Sf(t,\theta) = t^{-\theta} K(f,t,L^p,\dot{BV}).$$

We expect that the image f in figure 1 belong to $(L^p,\dot{BV})_{\theta,\infty}$ for θ small (and hence by (14) for all $\theta \in (0,1)$). On the other hand, the image f in figure 3 should only belong to $(L^p,\dot{BV})_{\theta,\infty}$ for θ close to 1. The regularity of f determines the parameter θ which in turn determines the preference of the interpolating scale t. In other words, if θ is close to 1 (or 0), then small (or large) $t > 0$ is weighted more to be chosen as an interpolating scale, respectively. Thus, the parameter θ determines the interpolating scale t. Stable interpolating scales are those that arise from the local maximum of $Sf(t,\theta)$ for more than one θ.

When the image f in consideration is clear, we will drop the subscript f. I.e., we write I_t for $I_{f,t}$, etc. The following examples provide intuitions and meanings to interpolating scales coming from different choices of K and L-functionals.

2.1 Examples

Example 1. In the first example we would like to show a heuristic approximation to (1) given by an L-functional. Let $\{m_i\}$ be an increasing sequence of positive integer, and $f(x) = \sum_{i=0}^{\infty} a_i m_i^{-1/2} \sin(m_i x)$ for $x \in [0, 2\pi]$, where $\sum_{i=0}^{\infty} m_i^{2s} a_i^2 < \infty$ for some large $s > 0$. Consider the L-functional given by

$$L(f,t) = \inf_{u \in \dot{H}^s} \left\{ t\|u\|_{\dot{H}^s}^2 + \|v\|_{L^2}^2 : f = u + v \right\}, \tag{21}$$

where $\|u\|_{\dot{H}^s}^2 = \|(-\Delta)^{s/2}u\|_{L^2}^2 = \|(|\xi|^s \hat{u}(\xi))^\vee(\cdot)\|_{L^2}^2$. Expressing these norms in Fourier domain, $L(f,t)$ can be rewritten as

$$L(f,t) = \inf_{u \in \dot{H}^s} \left\{ \int_{\mathbb{R}} \left[t|\xi|^{2s}|\hat{u}(\xi)|^2 + |\hat{f}(\xi) - \hat{u}(\xi)|^2 \right] \, d\xi \right\}.$$

Note that in this case $X_1 = \dot{H}^s$ and $X_2 = L^2$. For each $t > 0$, there exists a unique $u(t) \in \dot{H}^s$ that minimizes (21) and $u(t)$ is given by

$$\widehat{u(t)}(\xi) = \frac{1}{1 + t|\xi|^{2s}} \cdot \hat{f}(\xi) = \frac{1}{1 + (t^{\frac{1}{2s}}|\xi|)^{2s}} \cdot \hat{f}(\xi).$$

Note that if $s > 0$ is large, then $\frac{1}{1+(t^{\frac{1}{2s}}|\xi|)^{2s}}$ can be approximated ($\approx$) by a characteristic function on the interval $[-t^{-\frac{1}{2s}}, t^{-\frac{1}{2s}}]$.

$$u(t) \approx \begin{cases} 0 & \text{if} \quad m_0^{-s2} < \quad t & \\ a_0 m_0^{-1/2} \sin(m_0 x) & \text{if} \quad m_1^{-2s} < \quad t & \leq m_0^{-2s} \\ \cdots & & \\ \sum_{i=0}^{n} a_i m_i^{-1/2} \sin(m_i x) & \text{if} \quad m_{n+1}^{-2s} < \quad t & \leq m_n^{-2s} \\ \cdots & \cdots \quad \cdots \quad \cdots & \cdots \end{cases}$$

Let $c_n = \sum_{i=0}^{n} m_i^{2s} a_i^2$, $c_\infty = \lim_{n\to\infty} c_n$, and $d_n = \sum_{i=n}^{\infty} a_i^2$. Note that in this case, we have $p = 2$ and $\gamma = \theta$. Therefore,

$$\mathcal{L}f(t,\theta) = t^{-\theta} L(f,t) \approx \begin{cases} t^{-\theta} d_0 & \text{if} \quad m_0^{-s2} < \quad t & \\ t^{-\theta}[tc_0 + d_1] & \text{if} \quad m_1^{-2s} < \quad t & \leq m_0^{-2s} \\ \cdots & & \\ t^{-\theta}[tc_{n-1} + d_n] & \text{if} \quad m_n^{-2s} < \quad t & \leq m_{n-1}^{-2s} \\ \cdots & \cdots \quad \cdots \quad \cdots & \cdots \end{cases}$$

Let $\varphi_n(t) = t^{-\theta}[tc_{n-1} + d_n]$, then φ_n has exactly one global minima at $s_n = \frac{\theta}{1-\theta}\frac{d_n}{c_{n-1}}$. The sequence $\{s_n\}$ is decreasing. In particular,

$$s_n = \frac{\theta}{1-\theta}\frac{d_n}{c_{n-1}} \geq \frac{\theta}{1-\theta}\frac{d_{n+1}}{c_{n-1}} \geq \frac{\theta}{1-\theta}\frac{d_{n+1}}{c_n} = s_{n+1}.$$

Thus, $t_n = m_n^{-2s}$ is a local maxima of $\mathcal{L}f(t,\theta)$ if and only if $s_{n+1} < m_n^{-2s} < s_n$. In other words,

$$\frac{\theta}{1-\theta}\frac{d_{n+1}}{c_n} \leq m_n^{-2s} \leq \frac{\theta}{1-\theta}\frac{d_n}{c_{n-1}}.$$

Equivalently, $\theta \in I_{t_n} \subset (0,1)$ where

$$I_{t_n} = \left(\frac{m_n^{-2s} c_{n-1}}{d_n + m_n^{-2s} c_{n-1}}, \frac{m_n^{-2s} c_n}{d_{n+1} + m_n^{-2s} c_n} \right).$$

We have

$$\begin{aligned} |I_{t_n}| &= m_n^{-2s} \frac{c_n d_n - c_{n-1} d_{n+1}}{(d_{n+1} + m_n^{-2s} c_n)^2} = m_n^{-2s} \frac{(c_{n-1} + m^{2s} a_n^2)(d_{n+1} + a_n^2) - c_{n-1} d_{n+1}}{(d_{n+1} + m_n^{-2s} c_n)^2} \\ &= m_n^{-2s} \frac{m^{2s} a_n^2 d_{n+1} + a_n^2 c_{n-1} + m^{2s} a_n^4}{(d_{n+1} + m_n^{-2s} c_n)^2} = a_n^2 \frac{d_{n+1} + m_n^{-2s} c_n}{(d_{n+1} + m_n^{-2s} c_n)^2} = \frac{a_n^2}{d_{n+1} + m_n^{-2s} c_n} \\ &= \frac{m_n^{2s} a_n^2}{m_n^{2s} d_{n+1} + c_n}, \end{aligned}$$

and

$$\frac{m_n^{2s} a_n^2}{c_\infty} \leq |I_{t_n}| = \frac{m_n^{2s} a_n^2}{m_n^{2s} d_{n+1} + c_n} \leq \frac{m_n^{2s} a_n^2}{m_0^{2s} d_0}.$$

Note that the lower bound $\frac{m_n^{2s} a_n^2}{c_\infty}$ gives the relative size of $m_n^{2s} a_n^2$ with respect to c_∞ which depends on both m_n and a_n. Moreover,

$$1 \leq \sum_{i=0}^{\infty} |I_{t_n}| \leq (m_0^{-2s})(\frac{c_\infty}{d_0}) = (m_0^{-2s}) \frac{\|f\|_{\dot{H}^s}^2}{\|f\|_{L^2}^2}. \tag{22}$$

In this example, all scales $t_n = m_n^{-2s}$ are stable scales. Suppose $a_n = a^n$ and $m_n = b^n$ for some $a \in (0,1)$ and $b > 1$ such that $b^s a < 1$. Then

$$(1 - b^{2s} a^2) b^{2sn} a^{2n} \leq |I_{t_n}| \leq (1 - a^2) b^{2sn} a^{2n}.$$

Example 2. Consider next the L-functional given by the Rudin-Osher-Fatemi model [26]

$$L(f,t) = \inf_{(u,v) \in \dot{BV} \times L^2} \left\{ t\|u\|_{\dot{BV}} + \|v\|_{L^2}^2 \right\}. \tag{23}$$

Let $f = \mathbf{1}_{B_r(0)}$. The calculation from [27] shows that

$$u(t) = \begin{cases} (1 - \frac{t}{r}) \mathbf{1}_{B_r} & \text{if} \quad t < r \\ 0 & \text{if} \quad t > r. \end{cases}$$

For all $\gamma \in (0,1)$, a simple calculation shows that $t^* = 2r\frac{1-\gamma}{2-\gamma} < r$ is the only local maxima for $\mathcal{L}(f,t) = t^{-\gamma} L(f,t)$. Moreover, $T_f = (0,r)$ and none of the interpolating scales in T_f are stable. This example shows that the L-functional (23) fails to see r as a stable interpolating scale.

Example 3. Consider the K-functional given by [9]

$$K(f,t) = \inf_{(u,v) \in \dot{BV} \times L^1} \left\{ t\|u\|_{\dot{BV}} + \|v\|_{L^1} : f = u + v \right\}. \tag{24}$$

Let $f = \mathbf{1}_{B_r(0)}$, where $B_r(p)$ denotes the ball of radius r centered at $p \in \mathbb{R}^2$. Let $(u(t), v(t) = f - u(t)) \in \dot{BV} \times L^1$ such that

$$K(f,t) = t\|u(t)\|_{\dot{BV}} + \|v\|_{L^1}.$$

Then the calculation from [9] gives

$$u(t) = \begin{cases} f & \text{if} \quad t < \frac{r}{2} \\ 0 & \text{if} \quad t > \frac{r}{2} \end{cases} \Rightarrow K(f,t) = \begin{cases} 2\pi r t & \text{if} \quad t < \frac{r}{2} \\ \pi r^2 & \text{if} \quad t > \frac{r}{2}. \end{cases}$$

For $\theta \in (0,1)$, we have

$$\mathcal{K}f(t,\theta) = \begin{cases} 2\pi r t^{1-\theta} & \text{if} \quad t < \frac{r}{2} \\ \pi r^2 t^{-\theta} & \text{if} \quad t > \frac{r}{2}. \end{cases} \Rightarrow \frac{\partial \mathcal{K}f(t,\theta)}{\partial t} = \begin{cases} (1-\theta)2\pi r t^{-\theta} > 0 & \text{if} \quad t < \frac{r}{2} \\ -\theta\pi r^2 t^{-\theta-1} < 0 & \text{if} \quad t > \frac{r}{2}. \end{cases}$$

This implies that, for all $\theta \in (0,1)$, $t^* = \frac{r}{2}$ is the only local maxima for $\mathcal{K}f(\cdot,\theta)$, and $I_{t^*} = (0,1)$. In this case t^* is the only stable scale.

Suppose next $f = \mathbf{1}_{Q_r(p)}$ where $Q_r(p)$ is the square centered at $p \in \mathbb{R}^2$ with side lengths r. Let $\theta^* = \frac{8+6\pi}{20+3\pi} \approx 0.9125$. A simple calculation shows that $t^* = \frac{r}{2}$ is the only local maxima for $\mathcal{K}f(t,\theta)$ for all $\theta \in I_{t^*} = (0, \theta^*)$, and for each $\theta \in (\theta^*, 1)$, $\bar{t} = \frac{16r(1-\theta)}{3(4-\pi)(2-\theta)}$ is the only local maxima for $\mathcal{K}f(t,\theta)$. Note that $t^* = \frac{r}{2}$ is the only stable scale.

Example 4. In this example, we consider again the K-functional given by (24). Let $\{r_k\}$ be a decreasing sequence of nonnegative real numbers such that $\sum_{k=0}^{\infty} r_k < \infty$, and let $f = \sum_{k=0}^{\infty} \mathbf{1}_{B_{r_k}(p_k)}$. We assume that $|p_k - p_j| \gg 1$ if $j \neq k$. Let

$$c_n = 2\pi \sum_{k=0}^{n} r_k, \; c_\infty = \lim_{n\to\infty} c_n, \text{ and } d_n = \pi \sum_{k=n}^{\infty} r_k^2.$$

We have

$$\mathcal{K}f(t,\theta) = \begin{cases} t^{-\theta} d_0 & \text{if} \quad \frac{r_0}{2} \le t < \infty \\ t^{-\theta}\left[tc_0 + d_1\right] & \text{if} \quad \frac{r_1}{2} \le t < \frac{r_0}{2} \\ \cdots & \\ t^{-\theta}\left[tc_{n-1} + d_n\right] & \text{if} \quad \frac{r_n}{2} \le t < \frac{r_{n-1}}{2} \\ \cdots & \end{cases} \tag{25}$$

Following the same arguments as in example 1, we obtain that $t_n = \frac{r_n}{2}$ is a local maxima of $\mathcal{K}f(t,\theta)$ for all $\theta \in I_{t_n}$ where

$$I_{t_n} = \left(\frac{(r_n/2)c_{n-1}}{d_n + (r_n/2)c_{n-1}}, \frac{(r_n/2)c_n}{d_{n+1} + (r_n/2)c_n} \right)$$

We have

$$\begin{aligned} |I_{t_n}| &= \frac{r_n}{2} \frac{2\pi^2 r_n^3 + \pi r_n^2 c_{n-1} + 2\pi r_n d_{n+1}}{(d_{n+1} + (r_n/2)c_n)^2} = \pi r_n^2 \frac{d_{n+1} + (r_n/2)c_n}{(d_{n+1} + (r_n/2)c_n)^2} \\ &= \frac{\pi r_n^2}{d_{n+1} + (r_n/2)c_n} \ge \frac{2\pi r_n}{c_\infty}. \end{aligned} \tag{26}$$

Again, the lower bound $2\pi r_n/c_\infty$ measures the relative size of $2\pi r_n$ with respect to c_∞. Also,

$$|I_{t_n}| = \frac{\pi r_n^2}{d_{n+1} + (r_n/2)c_n} = \frac{2\pi r_n}{2d_{n+1}/r_n + c_n} \leq \frac{2\pi r_n}{(2/r_0)d_0}.$$

Moreover,

$$1 \leq \sum_{n=0}^{\infty} |I_{t_n}| \leq (\frac{r_0}{2})(\frac{c_\infty}{d_0}) = (\frac{r_0}{2})\frac{\|f\|_{\dot{BV}}}{\|f\|_{L^1}}, \tag{27}$$

which agrees with (22). Note that all scales $t_n = \frac{r_n}{2}$ are stable scales. Denote by $K'_-(f,t)$ and $K'_+(f,t)$ the left and right derivative of $K(f,t)$ at t. Then at $t_n = \frac{r_n}{2}$, we have

$$K'_-(f,t_n) = c_n > c_{n-1} = K'_+(f,t_n).$$

This implies

$$K'_-(f,t_n) - K'_+(f,t_n) = c_n - c_{n-1} = 2\pi r_n.$$

We see that in this case

$$I_{t_n} = \frac{t_n}{K(t_n)}[K'_-(f,t_n) - K'_+(f,t_n)], \tag{28}$$

where $K(t_n) = t_n c_n + d_{n+1} = t_n c_{n-1} + d_n$. In this example, suppose $r_n = a^n$, $0 < a < 1$. Then

$$(1-a)a^n \leq |I_{t_n}| \leq (1-a^2)a^n$$

which shows that $|I_{t_n}|$ decays exponentially.

Example 5. In this example, we extend example 4 by considering $f = \sum_{k=0}^{\infty} a_k \mathbf{1}_{B_{r_k}(p_k)}$, where $a_k > 0$ and the sequence $\{r_k\}$ is decreasing such that $\sum_{k=0}^{\infty} a_k r_k < \infty$. Let

$$c_n = 2\pi \sum_{k=0}^{n} a_k r_k,\ c_\infty = \lim_{n\to\infty} c_n,\ \text{and}\ d_n = \pi \sum_{k=n}^{\infty} a_k r_k^2.$$

Follow the same calculation as in example 4, one obtains that $t_n = \frac{r_n}{2} \in T_f$ for all n and

$$I_{t_n} = \left(\frac{(r_n/2)c_{n-1}}{d_n + (r_n/2)c_{n-1}}, \frac{(r_n/2)c_n}{d_{n+1} + (r_n/2)c_n}\right),$$

and

$$|I_{t_n}| = \frac{\pi a_n r_n^2}{d_{n+1} + (r_n/2)c_n}.$$

Remark 3. Following example 4, suppose $f = \chi_{B_{r_0}}(p_0) + \chi_{B_{r_1}}(p_1)$ with $r_0 > r_1$ and $|p_0 - p_1| >> 1$. With the notation that $c_{-1} = 0$ and $d_2 = 0$, we have

$$I_{t_0} = \left(0, \frac{r_0^2}{r_0^2 + r_1^2}\right) \text{ and } I_{t_1} = \left(\frac{r_0 r_1}{r_0 r_1 + r_1^2}, 1\right).$$

Suppose r_0 is fixed.

1. Letting $r_1 \to 0$ implies $|I_{t_1}| \to 0$ and $|I_{t_0}| \to 1$. This shows that the scale t_0 is more prominent than t_1.
2. Letting $r_1 \to r_0$ implies $|I_{t_1}| \to 1/2$ and $|I_{t_0}| \to 1/2$. This shows that both scales r_0 and r_1 are equally prominent.

2.2 Properties of interpolating scales

Let the dilating operator be defined as $d_\delta f(x) = f(\delta x)$ for $\delta > 0$. Suppose $\|\cdot\|_{X_1}$ and $\|\cdot\|_{X_2}$ are homogenous of degree s_1 and s_2 respectively. In other words,

$$\|\partial_\delta g\|_{X_1} = \delta^{s_1}\|g\|_{X_1}, \text{ and } \|d_\delta g\|_{X_2} = \delta^{s_2}\|g\|_{X_2}. \tag{29}$$

Suppose also that $Sf(t,\theta) = t^{-\theta}K(f,t)$, where $K(f,t)$ is defined in (8). Then we have

$$K(d_\delta f, t) = \delta^{s_2}K(f, t\delta^{s_1-s_2}),$$

which implies

$$S(d_\delta f)(t,\theta) = \delta^{\theta s_1+(1-\theta)s_2}Sf(t\delta^{s_1-s_2},\theta).$$

In particular,

$$S(d_\delta f)(t',\theta) = \delta^{\theta s_1+(1-\theta)s_2}Sf(t,\theta),$$

where $t' = t\delta^{s_2-s_1}$. This shows that the interpolating scales satisfies the dilating consistency property, and we can relate the scales of $d_\delta f$ with the scales of f. Moreover, the stability of scales are invariant under dilation. In particular, suppose $Sf(t,\theta)$ is given by a K-functional satisfying the homogeneous property (29). Then the interpolating scales of $d_\delta f$ is given by

$$T_{d_\delta f} = \left\{t' = t\delta^{s_2-s_1} : t \in T_f\right\}.$$

Moreover,

$$I_{d_\delta f,t'} = I_{f,t}, \ \forall t \in T_f \text{ and } t' = t\delta^{s_2-s_1} \in T_{d_\delta f}.$$

Moreover, if $s_1 = s_2$, then $t' = t$ which shows that interpolating scales are invariant under dilations.

The same result also holds if Sf is given by an L-functional.

Remark 4. Suppose we are in $\mathbb{R}^n$:

- For the functional $L(f,t,L^2,\dot{H}^s)$ given by (21) with $p_1 = p_2 = 2$, $s_1 = -n+2s$ and $s_2 = -n$.
- For the functional $L(f,t,L^2,\dot{BV})$ with $p_1 = 1$ and $p_2 = 2$, $s_1 = -n+1$ and $s_2 = -n$.
- For the functional $K(f,t,L^1,\dot{BV})$ given by (24), $s_1 = -n+1$ and $s_2 = -n$.
- For the K-functional $K(f,t,L^2,\dot{BV})$, $s_1 = -n+1$ and $s_2 = -n/2$. If $n = 2$ (i.e. in $\mathbb{R}^2$), then $s_1 = s_2 = -1$. This shows that the interpolating scales of f with respect to $K(f,t,L^2,\dot{BV})$ are invariant under dilation.

Examples 1, 3 and 4 show that the set $I_{f,t}$ for each $t \in T_f$ is a subinterval in $(0,1)$. The following result indeed shows that in general $I_{f,t}$ is a subinterval in $(0,1)$.

Proposition 1. *Let $\bar{t} \in T_f(\theta_1) \cap T_f(\theta_2)$, for some $0 < \theta_1 < \theta_2 < 1$. Then $\bar{t} \in T_f(\theta)$ for all $\theta \in (\theta_1,\theta_2)$. In other words, $I_{f,\bar{t}}$ is a subinterval in $(0,1)$.*

Proof. $\bar{t}$ being a local maxima for both $Sf(t,\theta_1)$ and $Sf(t,\theta_2)$ implies there exists a small $\epsilon > 0$ such that $Sf(\bar{t},\theta_i) \geq Sf(t,\theta_i)$, for all $t \in (\bar{t}-\epsilon, \bar{t}+\epsilon)$, for $i = 1, 2$. We have for all $t \in (\bar{t}, \bar{t}+\epsilon)$,

$$Sf(\bar{t},\theta) = \bar{t}^{-\theta+\theta_1} Sf(\bar{t},\theta_1) > t^{-\theta+\theta_1} Sf(t,\theta_1) = Sf(t,\theta),$$

since $t^{-\theta+\theta_1}$ is strictly decreasing. Similarly, for all $t \in (\bar{t}-\epsilon, \bar{t})$,

$$Sf(\bar{t},\theta) = \bar{t}^{-\theta+\theta_2} Sf(\bar{t},\theta_2) > t^{-\theta+\theta_2} Sf(t,\theta_2) = Sf(t,\theta),$$

since $t^{-\theta+\theta_2}$ is strictly increasing. Hence, $\bar{t}$ is a local maxima of $Sf(t,\theta)$. □

Suppose $\bar{t} \in T_f$, and $K(f,\cdot)$ is differentiable at $\bar{t}$. Then for some $\bar{\theta} \in (0,1)$, we have

$$0 = \frac{\partial}{\partial t} Sf(\bar{t},\bar{\theta}) = -\bar{\theta} t^{-\bar{\theta}-1} K(f,\bar{t}) + t^{-\bar{\theta}} K'(f,\bar{t}), \;\Rightarrow\; \bar{\theta} = \bar{t}\frac{K'(f,\bar{t})}{K(f,\bar{t})},$$

which is uniquely determined by $\bar{t}$. In other words, if $K(f,\cdot)$ is differentiable at $\bar{t}$, then $I_{f,\bar{t}} = \{\bar{\theta}\}$. Since for a fixed θ, $Sf(\cdot,\theta)$ is differentiable almost everywhere, we have that $|I_{f,t}| = 0$ for almost every $t \in T_f$. Thus the following result holds.

Proposition 2. *For each $t \in T_f$, either $I_{f,t}$ is an interval in $(0,1)$ or $I_{f,t}$ consists of a single point. Moreover, for almost every $t \in T_f$, $|I_{f,t}| = 0$.*

In examples 1 and 2 we see that each θ, $Sf(t,\theta)$ is differentiable for all t. Thus, $I_{f,t} = 0$ for all $t > 0$. From (28), we see that the length of the interval $I_{f,t}$ is directly related to the left and the right derivative of K at t. Denote by $K'_-(f,t)$ and $K'_+(f,t)$ the left and right derivatives of $K(f,t)$ at t respectively. Then the following result holds.

Proposition 3. *For all $t \in T_f$, we have*

$$|I_{f,t}| = \frac{t}{K(f,t)}\left[K'_-(f,t) - K'_+(f,t)\right]. \tag{30}$$

Proof. By Proposition 2, we may assume $K'_-(f,t) > K'_+(f,t)$. We have

$$Sf'_-(t,\theta) = -\theta t^{-\theta-1} K(f,t) + t^{-\theta} K'_-(f,t) > 0 \iff \theta < \frac{t}{K(f,t)} K'_-(f,t).$$

Similarly,

$$Sf'_+(t,\theta) = -\theta t^{-\theta-1} K(f,t) + t^{-\theta} K'_+(f,t) < 0 \iff \theta > \frac{t}{K(f,t)} K'_+(f,t).$$

Thus, $\theta \in I_{f,t}$ if and only if

$$\frac{t}{K(f,t)} K'_+(f,t) < \theta < \frac{t}{K(f,t)} K'_-(f,t),$$

which shows (30). □

Remark 5. The previous Proposition shows that the stability of the scale t is given by the jump in the derivatives of $K(f,t)$ at t weighted by $t/K(f,t)$. Suppose $t_1 < t_2$ are two interpolating scales such that

$$K'_-(f,t_1) - K'_+(f,t_1) = K'_-(f,t_2) - K'_+(f,t_2). \tag{31}$$

By (18), we have

$$\frac{t_1}{K(f,t_1)} \leq \frac{t_2}{K(f,t_2)}.$$

This implies that the scale t_2 is more stable than t_1 even though (31) holds.

Recall that we have

$$K(f,t) = K(f,t,X_2,X_1) = \inf_{(u,v)\in X_1\times X_2} \{t\|u\|_{X_1} + \|v\|_{X_2} : f = u+v\}.$$

Suppose $f \in X_2$. Take $u = 0$ and $v = f$, then

$$K(f,t) \leq \|f\|_{X_2}, \text{ for all } t > 0.$$

On the other hand suppose $f \in X_1$. Let $u = f$ and $v = 0$, then for all $t > 0$,

$$t^{-1}K(f,t) \leq \|f\|_{X_1}.$$

In particular,

$$\lim_{t\to\infty} K(f,t) \leq \|f\|_{X_2} \text{ and } \lim_{t\to 0^+} t^{-1}K(f,t) \leq \|f\|_{X_1}.$$

For example if we take $X_1 = L^\infty$ and $X_2 = L^1$, then it has been shown (see [6], Chapter 5, Theorem 1.6) that $\infty L^\infty + L^1 = L^1$ and $L^\infty + \infty L^1 = L^\infty$ with the same norms. In other words,

$$\|f\|_{\infty L^\infty + L^1} = \lim_{t\to\infty} K(f,t) = \|f\|_{L^1}, \text{ and } \|f\|_{L^\infty+\infty L^1} = \lim_{t\to 0^+} t^{-1}K(f,t) = \|f\|_{L^\infty}.$$

Thus it is possible that $K(f,t)$ is unbounded, i.e. $\lim_{t\to\infty} K(f,t) = \infty$ in which case $f \notin L^1$. Similarly, if $f \notin L^\infty$, then $\lim_{t\to 0} t^{-1}K(f,t) = \infty$. In this case, we have $f \in L^\infty + L^1$, but $f \notin L^\infty$ and $f \notin L^1$.

Next, we would like to know what structure there is on the sets $I_{f,t}$, in particular if there is a bound on how many of them there are of length greater than some fixed $\delta > 0$. The following example shows that if $K(f,t)$ can be an arbitrary increasing concave function then no such bound exists. We take $K(f,t)$ to be a piecewise linear approximation to a power function with exponent less than 1. Specifically, take $A > B > 1$ and by Proposition 3, let $K(f,t)$ be piecewise linear with vertices (t_j, K_j) for all $j \in \mathbb{Z}$ such that

$$|I_{f,t_j}| = \frac{t_j}{K_j}\left(\frac{K_j - K_{j-1}}{t_j - t_{j-1}} - \frac{K_{j+1} - K_j}{t_{j+1} - t_j}\right) = \frac{B-1}{A-1}\left(\frac{A}{B} - 1\right).$$

Hence it is evident that for any fixed $\delta > 0$, we can choose $A > B > 1$ such that there are infinitely many $I_{f,t}$ of length greater than δ. In this example, $K(f,t)$ is unbounded and has unbounded right derivative at 0.

Next, let us note that $\frac{K(f,t)}{t}$ is decreasing in t and bounded below by $K(f,1)$ when $t \in (0,1)$ and thus on this interval

$$\frac{t}{K(f,t)} \le K(f,1)^{-1}, \ \forall t \in (0,1).$$

This gives

$$|I_{f,t}| \le \frac{1}{K(f,1)}(K'_-(f,t) - K'_+(f,t)). \tag{32}$$

If $\lim_{t\to 0^+} K'(f,t)$ is bounded (which is the case when $f \in X_1$), then equation (32) implies the summability of the lengths of the $I_{f,t}$ over the scales $t \in (0,1)$ since the sum is telescoping. For the scales $t \ge 1$, we instead use the fact that $K(f,t)$ is increasing and the fact that

$$|I_{f,t}| \le \frac{t}{K(f,1)}(K'_-(f,t) - K'_+(f,t)).$$

However, when examining the graph of $K(f,t)$, it is easy to see that $t(K'_-(f,t) - K'_+(f,t))$ is the interval subtended on the vertical axis by the left and right tangent lines at t. Concavity implies that these intervals are disjoint for distinct values of t, and their union is contained in the range of $K(f,t)$. In particular if $K(f,t)$ is bounded (which is the case when $f \in X_2$) then the lengths of $I_{f,t}$ over the scales $t \ge 1$ are summable. Combining this with the previous result, we see that if $K(f,t)$ is bounded and has bounded right derivative at 0, then the lengths of $I_{f,t}$ are summable over all scales. Thus we have the following Lemma.

Lemma 2. *Suppose* $\lim_{t\to\infty} K(f,t) < \infty$ *and* $\lim_{t\to 0^+} t^{-1}K(f,t) < \infty$. *In other words, suppose* $f \in (\infty X_1 + X_2) \cap (X_1 + \infty X_2)$. *Then*

$$\sum_{t \in T_f} |I_{f,t}| < \infty.$$

A direct consequence of Lemma 2 is the following result.

Corollary 2. *Suppose* $f \in X_1 \cap X_2$. *Then*

$$\sum_{t \in T_f} |I_{f,t}| < \infty.$$

A stronger estimate for $K(f,t)$ as $t \to \infty$ is as follows. Suppose X_1 is continuously embedded in X_2. In other words, there exists $C > 0$ such that $\|u\|_{X_2} \le C\|u\|_{X_1}$ for all $u \in X_1$. Then for all $t \ge C$, we have $K(f,t) = \|f\|_{X_2}$, assuming $f \in X_2$. Indeed, let $f = u + v$ be any decomposition, then

$$t\|u\|_{X_1} + \|v\|_{X_2} \ge \frac{t}{C}\|u\|_{X_2} + \|v\|_{X_2} \ge \|u\|_{X_2} + \|v\|_{X_2} \ge \|f\|_{X_2}.$$

Taking the infimum over all decompositions, we have $K(f,t) \ge \|f\|_{X_2}$, but $u = 0$ and $v = f$ is one valid decomposition. This implies,

$$K(f,t) = \|f\|_{X_2}, \text{ for all } t \ge C.$$

Thus, we have the following result.

Corollary 3. *Suppose X_1 is continuously embedded in X_2, and let C be the smallest constant such that*

$$\|u\|_{X_2} \le C\|u\|_{X_1}, \text{ for all } u \in X_1.$$

Let $f \in X_2$. Then for all $t \ge C$, $K(f,t) = \|f\|_{X_2}$. The decomposition $u = 0$ and $v = f$ minimizes

$$J(u,t) = t\|u\|_{X_1} + \|v\|_{X_2}.$$

Remark 6. In $\mathbb{R}^n$. Define

$$p^* = \begin{cases} \frac{np}{n-p} & \text{if } p < n \\ \infty & \text{if } p = n \end{cases}$$

with $1^* = \infty$ if $n = 1$ and $\frac{n}{n-1}$ otherwise. Let $\Omega \subset \mathbb{R}^n$ be a bounded with Lipschitz boundary. Poincaré Inequalities imply that $\dot{BV}(\Omega)$ is continuously embedded in $L^p(\Omega)$, for all $1 \le p \le 1^*$. In other words, there exists a constant $C = C(p,\Omega)$ such that

$$\|u - u_\Omega\|_{L^p(\Omega)} \le C\|u\|_{\dot{BV}(\Omega)}, \text{ for all } u \in \dot{BV}(\Omega),\ 1 \le p \le 1^*.$$

Let $f \in L^p(\Omega)$ and consider the K-functional $K(f,t) = K(f,t,\dot{BV}(\Omega),L^p)$, $1 \le p \le 1^*$. Give any decomposition $f = u + v \in \dot{BV}(\Omega) + L^p(\Omega)$ we have

$$\begin{aligned} t\|u\|_{\dot{BV}(\Omega)} + \|v\|_{L^p(\Omega)} &= t\|u - u_\Omega\|_{\dot{BV}(\Omega)} + \|v\|_{L^p(\Omega)} \ge \frac{t}{C}\|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)} \\ &\ge \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)} \ge \|f\|_{L^p}. \end{aligned}$$

Thus for all $t \ge C$, $K(f,t) = \|f\|_{L^p}$.

We have seen that if $\lim_{t\to\infty} K(f,t) = K(f,\infty) < \infty$ (i.e if $f \in \infty X_1 + X_2$), then for any fixed $t_0 > 0$, the lengths of the intervals $I_{f,t}$, $t \in T_f$ and $t \ge t_0$, are always summable. In particular,

$$\sum_{t\in T_t, t\ge t_0} |I_{f,t}| \le K(f,\infty) - K(f,t_0).$$

Also, if $\lim_{t\to 0^+} t^{-1}K(f,t) < \infty$ (i.e. $f \in X_1 + \infty X_2$) , then the summability of the lengths of the intervals $I_{f,t}$, for $t \in I_f$ and $0 < t < t_0$, are ensured. However, the boundedness of $\lim_{t\to 0^+} t^{-1}K(f,t)$ is not necessary. In other words, there are cases where $\lim_{t\to 0^+} t^{-1}K(f,t) = \infty$ and

$$\sum_{t\in T_t, t\le t_0} |I_{f,t}| < \infty. \tag{33}$$

Indeed, suppose $f \in X_2 \setminus X_1$ and $X_1 \subset X_2$ such that

$$\inf_{u\in X_1} \|f - u\|_{X_2} = \gamma > 0. \tag{34}$$

For instance, X_2 is a finite dimensional vector space of dimension m and X_1 is any subspace of X_2 with dimension $k < m$. For each $t > 0$, let $f = u + v$ be any decomposition of f. Then we have

$$t\|u\|_{X_1} + \|v\|_{X_2} \ge \|v\|_{X_2} = \|f - u\|_{X_2} \ge \gamma.$$

Taking the infimum over all decompositions of f, we have $K(f,t) \geq \gamma$. Using the fact that $t(K'_-(f,t) - K'_+(f,t))$ is the interval subtended on the vertical axis by the left and right tangent lines at t, we have that (33) holds. In particular,

$$\sum_{t \in T_t, t \leq t_0} |I_{f,t}| \leq K(f,t_0) - \gamma. \tag{35}$$

The thresholding parameter δ as an interpolating scale: Fix f and define $s : \mathbb{R}^+ \to \mathbb{R}^+$ by $s(t) = |I_{f,t}|$. For each $\lambda \geq 0$ denote

$$D_\lambda = \{t > 0 : s(t) > \lambda\}.$$

From (30) and the fact that $K(f,t)$ is concave, we have D_λ is a finite set, i.e. $\#D_\lambda < \infty$. Consider μ_s and s^* defined as

$$\mu_s(\lambda) = \#D_\lambda, \text{ for all } \lambda > 0,$$

and

$$s^*(\delta) = \inf_{\mu_s(\lambda) \leq \delta} \lambda.$$

Let $d_\delta(t)$ be defined as

$$d_\delta(t) = \begin{cases} s^*(\delta) & \text{if} \quad s(t) > s^*(\delta), \\ s(t) & \text{otherwise.} \end{cases}$$

Equipped l^1 and l^∞ on $\mathbb{R}^+$ with the discrete measure, then $d_\delta(t)$ is the minimizer for the variational problem

$$K(s,\delta,l^1,l^\infty) = \inf_{d \in l^\infty} \left\{ \delta \|d\|_{l^\infty} + \|s-d\|_{l^1} \right\}. \tag{36}$$

Thus, the thresholding parameter δ can be viewed as an interpolating scale with respect to the K-functional $K(s,\delta,l^1,l^\infty)$. Thus, δ can also be automated.

3 Numerical results

In this section, we provide some numerical results on the interpolating scales with respect to the following functional:

$$K(f,t) = \inf_{(u,v) \in \dot{BV} \times L^1} \left\{ t\|u\|_{\dot{BV}} + \|v\|_{L^1} \right\}. \tag{37}$$

As seen from Example 4, $K(f,t)$ is localized in the spacial domain. This functional $K(f,t)$ was proposed by Nikolova in [24] as a discrete model to remove outliners and impulse noise. This functional was also proposed in [10] as an REU Summer project. Chan-Esedoglu [9] further analyzed $K(f,t)$ as a continuous model and studied the case where $f = \chi_\Omega$, for some $\Omega \subset \mathbb{R}^n$. They also showed that for each $t > 0$, there exists a minimizer $u(t)$ that is also a characteristic function. Allard in [1] and [2] provided further analysis of $K(f,t)$ and computed exact minimizers for a certain class

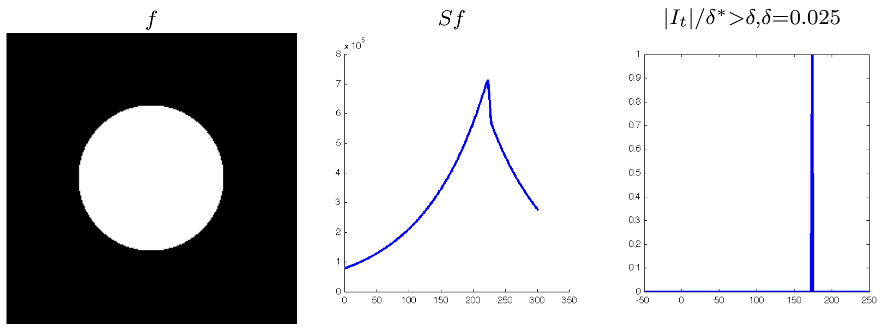

Figure 1: In this figure, f is a characteristic function of a circle. $Sf(t,\theta) = t^{-\theta}K(f,t)$ with $\theta = 0.5$. $|I_t|/\delta^*$ shows the length of the intervals for the δ-relatively-stable scales with $\delta = 0.025$. In this case, there is only one δ-relatively-stable scale.

of characteristic functions. See also Morgan-Vixie [23], Duval-Aujol [11] and references there in for further analysis of the functional $K(f,t)$. Recently there are fast numerical methods for computing functional involving the total variation [12], [8], [13]. For the numerical results shown below, we use the algorithm from [12] for computing minimizers in (37). The software can be downloaded through http://www.wotaoyin.com/

Based on the result from Corollary 1, the Sf function is discretize uniformly with respect to (τ,θ), where $\tau = \log_a t$ with $a = 1.02$. More specifically, let

$$D_\tau = \left\{ t_k = a^k : k = -50, \cdots, 250, a = 1.02 \right\}, \text{ and}$$

$$D_\theta = \left\{ \theta_j = j\Delta\theta : \Delta\theta = 0.005, j = 1, \cdots, \frac{1}{\Delta\theta} - \Delta\theta \right\}.$$

Then Sf is computed discrete on $D_t \times D_\theta$.

Figure 1 shows indeed there is only one stable scale, which agrees with example 3.

Figure 2 shows the interpolating scales for the beetle image. The plot of I_t/δ^* shows that the most stable and localized scale in this image is t_k with $k = 153$ which corresponds to the body of the beetle. The scale t_k with $k = 72$, which corresponds to beetle's legs, is not very well localized. There seems to be a mixture of scales ranging from $k = 62$ to $k = 82$ for the beetle's legs. See the caption in figure 2 for further explanations.

Figures 3-4 show the case for texture decompositions. In this case, the prominent stable scale most occur at small t_k's.

4 Discussion

From Proposition 3, we see that t^* is a stable scale if and only if there is a jump in the derivative of $K(f,t)$ at t^*. This shows that the $\frac{d}{\partial t}K(f,t)$ is non-smooth in t. The set of stable interpolating scales can be used as a tool to determine the non-smoothness of K.

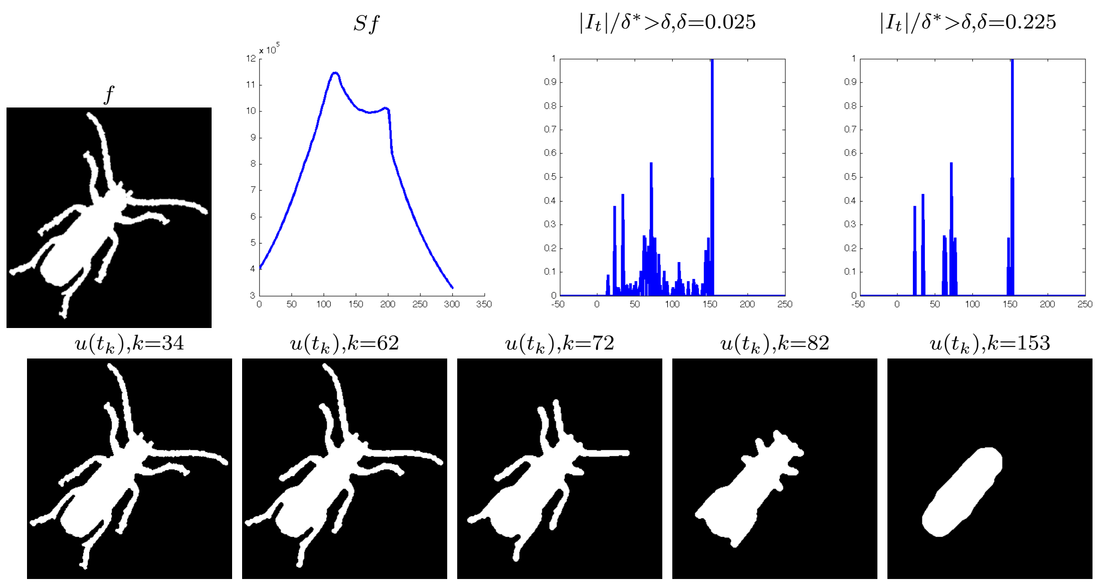

Figure 2: In this figure, f is a characteristic function of the beetle. $Sf(t, \theta) = t^{-\theta} K(f, t)$ with $\theta = 0.5$ which shows there are two local maximas. $|I_t|/\delta^*$ shows the length of the intervals for the δ-relatively-stable scales with $\delta = 0.025, 0.225$. In this case, we see that there are three cluster of scales, which occur at t_k's with $k = 34, 72, 153$. The cluster of scales at $k = 72$ contains a mixture of scales ranging from $k = 62$ to $k = 82$.

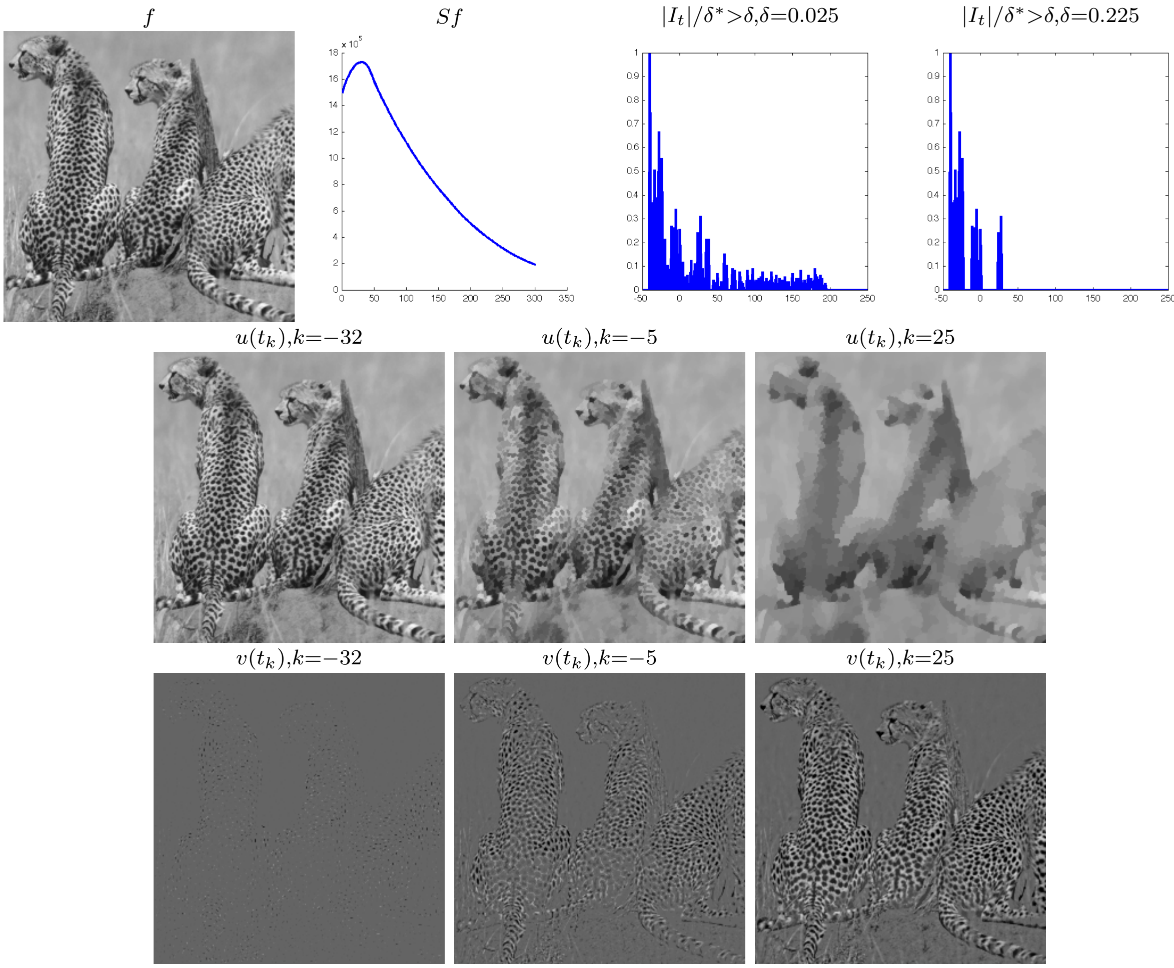

Figure 3: In this figure, $Sf(t,\theta) = t^{-\theta}K(f,t)$ with $\theta = 0.5$ which shows there are two local maxima. $|I_t|/\delta^*$ shows the length of the intervals for the δ-relatively-stable scales. In this case, we see that there are a few clusters of scales, which occur at t_k's with $k = -32, -5, 25$.

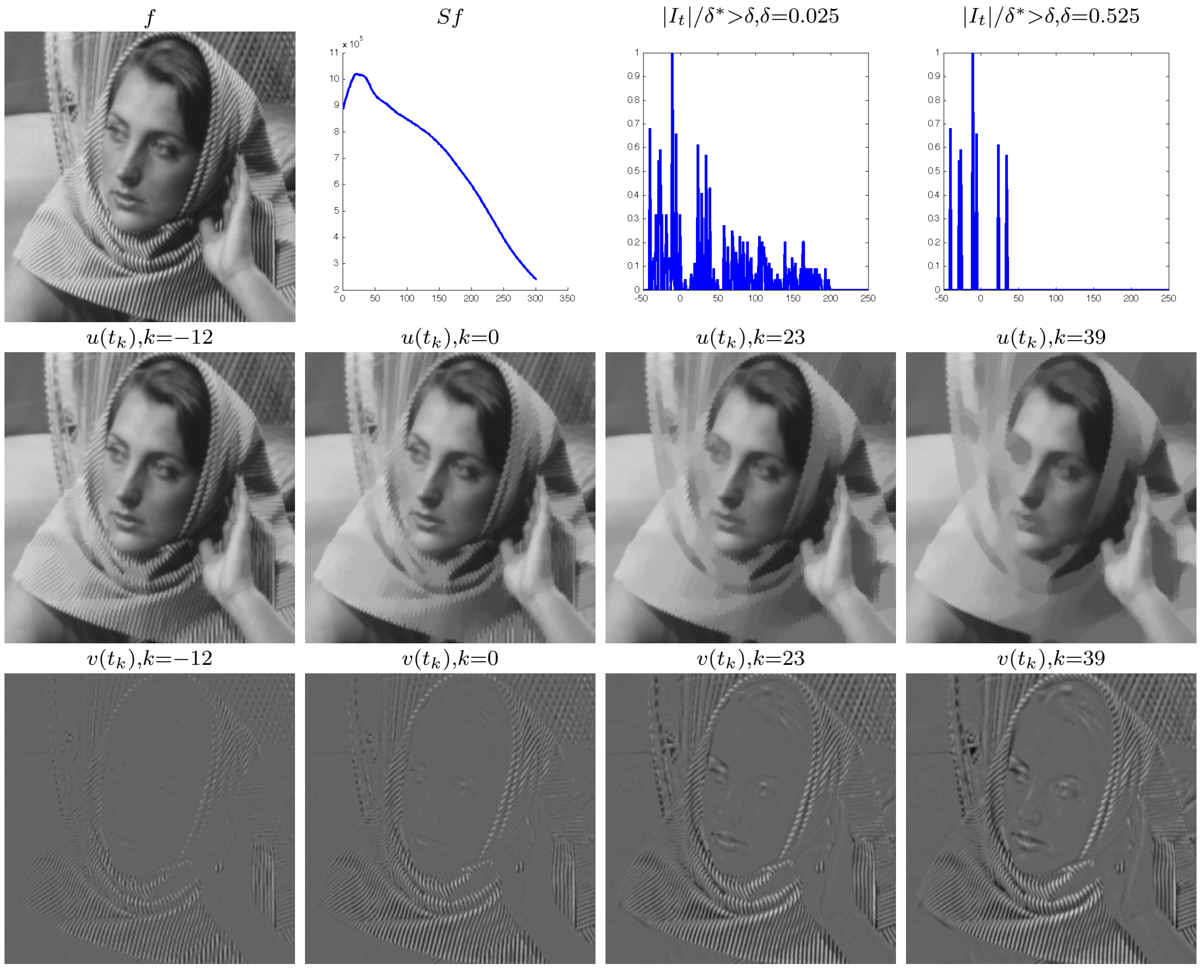

Figure 4: In this figure, $Sf(t,\theta) = t^{-\theta}K(f,t)$ with $\theta = 0.5$ which shows there is one local maxima. $|I_t|/\delta^*$ shows the length of the intervals for the δ-relatively-stable scales. The most stable cluster of scales ranging from t_k's with $k = -12$ to $k = 0$. Another cluster of stable scales ranging from $k = 23$ to $k = 39$.

Next, we would like to discuss another approach for selecting the (global) scale parameter t in (8) (or (9)), which is related to the local scale method described in [16]. For each $t > 0$, let $u(t) \in X_1$ such that

$$K(f,t) = t\|u(t)\|_{X_1} + \|f - u(t)\|_{X_2}.$$

To study local scales of f, we consider [16],

$$Sf(t,x) = t\frac{\partial u}{\partial t}(t,x).$$

The local scales of f at x are then defined as local maxima of $|Sf(\cdot,x)|$. As for the global scale, one can consider $Sf(t) = t\left\|\frac{\partial u}{\partial t}(t)\right\|_{X_1}$. The global scales are defined as local maxima of $|Sf(\cdot)|$. If we go back to example 3, then $Sf(t) = -t2\pi r\delta_{r/2}(t)$, which shows that $t = r/2$ is the only local maxima and $|Sf(r/2)| = \pi r^2$ provides a notion of visibility (stability) of the scale $t = r/2$. In example 4, we have

$$u(t) = \begin{cases} 0 & \text{if } \frac{r_0}{2} \le t < \infty, \\ \mathbf{1}_{B_{r_0}(p_0)} & \text{if } \frac{r_1}{2} \le t < \frac{r_0}{2}, \\ \cdots & \\ \sum_{k=0}^{n-1}\mathbf{1}_{B_{r_k}(p_k)} & \text{if } \frac{r_n}{2} \le t < \frac{r_{n-1}}{2}, \\ \cdots & \end{cases}$$

Thus, $Sf(t) = -t\sum_{n=0}^{\infty} 2\pi r_n \delta_{r_n/2}(t)$, which shows that the only local maxima for $|Sf|$ are $r_n/2$, and

$$|Sf(r_n/2)| = \pi r_n^2 = \frac{r_n}{2}[K'_-(f, r_n/2) - K'_+(f, r_n/2)].$$

A related discussion of global scales is proposed by Vixie-et al [30], where the authors consider $Sf(t) = \frac{\partial}{\partial t}\left[\|u(t)\|_{\dot{BV}}\right]$. In This case,

$$|Sf(r_n/2)| = 2\pi r_n = K'_-(f, r_n/2) - K'_+(f, r_n/2).$$

For instance, if one considers

$$J_{t_n} = \frac{2\pi r_n}{\|f\|_{\dot{BV}}} = \frac{r_n}{\sum_{k=0}^{\infty} r_k}$$

as a measure of stability for the scale $t_n = r_n/2$, then as shown in example 4, $|I_{t_n}| > J_{t_n}$.

An example of a K-functional: Suppose $f \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. For each $\lambda \ge 0$ and $t \ge 0$, define

$$\mu_f(\lambda) = |\{x \in \mathbb{R}^n : |f(x)| > \lambda\}|,$$

and

$$f^*(t) = \inf_{\mu_f(\lambda) \le t} \lambda.$$

For each $t > 0$ and $x \in \mathbb{R}^n$, define

$$u(t,x) = \min\{|f(x)|, f^*(t)\} \cdot \operatorname{sgn}(f(x)), \text{ and}$$

$$v(t,x) = f(x) - u(t,x) = \max\{|f(x)| - f^*(t), 0\} \cdot \text{sgn}(f(x)).$$

Then one has ([6])

$$t\|u(t,\cdot)\|_{L^\infty} + \|v(t,\cdot)\|_{L^1} = K(f,t,L^1,L^\infty).$$

Moreover,

$$K(f,t,L^1,L^\infty) = \int_0^t f^*(s)\, ds.$$

The same also holds if the Lesbegue spaces L^1 and L^∞ are replaced by l^1 and l^∞, respectively, with the discrete measure.

Acknowledgement: T.M.L. is very grateful for the supports from NSF DMS 0809270 and ONR N000140910108.

References

[1] W.K. Allard. Total variation regularization for image denoising, I. Geometric theory. *SIAM Journal on Mathematical Analysis*, 39(4):1150–1190, 2008.

[2] W.K. Allard. Total variation regularization for image denoising, II. Examples. *SIAM Journal on Imaging Sciences*, 1(4):400–417, 2008.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford University Press, USA, 2000.

[4] J.F. Aujol and A. Chambolle. Dual norms and image decomposition models. *International Journal of Computer Vision*, 63(1):104, 2005.

[5] J.F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decompositionmodeling, algorithms, and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, 2006.

[6] C. Bennett and R.C. Sharpley. *Interpolation of operators*. Academic Pr, 1987.

[7] T. Brox and J. Weickert. A TV flow based local scale measure for texture discrimination. *European Conf. on Computer Vision*, 3022:578–590, 2004.

[8] A. Chambolle and J. Darbon. On total variation minimization and surface evolution using parametric maximum flows. *International Journal of Computer Vision*, 84(3):288–307, 2009.

[9] T.F. Chan and Esedoglu S. Aspects of total variation regularized L^1 function approximation. *Siam Journal of Applied Mathematics*, 65(5):1817–1837, 2005.

[10] E. Cheon, A. Paranjpye, L. Vese, and S. Osher. Noise removal by total variation minimization. *UCLA MATH*, Math 199 REU Project, 2002.

[11] V. Duval, J.F. Aujol, and Y. Gousseau. The TVL1 model: a geometric point of view. 2009.

[12] D. Goldfarb and W. Yin. Parametric maximum flow algorithms for fast total variation minimization. *Rice University CAAM Technical Report TR07-09*, 2007.

[13] T. Goldstein and S. Osher. The split Bregman method for L1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[14] J.B. Garnett and P.W. Jones and T.M. Le and L.A. Vese. Modeling oscillatory components with the homogeneous spaces $B\dot{M}O^{-\alpha}$ and $\dot{W}^{-\alpha,p}$. Technical report, UCLA CAM Report 07-21, July 2007, to appear in PAMQ.

[15] J.B. Garnett and T.M. Le and Y. Meyer and L.A. Vese. Image decomposition using total variation and homogeneous besov spaces. *Applied and Computational Harmonic Analysis*, 23(1):25–56, July 2007.

[16] P.W. Jones and T.M. Le. Local scales and multiscale image decompositions. *Applied and Computational Harmonic Analysis*, 26(3):371–394, 2009.

[17] L.H. Lieu and L.A. Vese. Image restoration and decomposition via bounded total variation and negative Hilbert-Sobolev spaces. *Applied Mathematics and Optimization*, 58(2):167–193, 2008.

[18] T. Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. *International Journal of Computer Vision*, 11(3):283–318, 1993.

[19] T. Lindeberg. Feature Detection with Automatic Scale Selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.

[20] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[21] B. Luo, J.F. Aujol, Y. Gousseau, S. Ladjal, and H. Maıtre. Characteristic scale in satellite images. In *ICASSP 2006*, volume 2, pages 809–812, 2006.

[22] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*. American Mathematical Society Providence, RI, 2001.

[23] S.P. Morgan and K.R. Vixie. L1TV computes the flat norm for boundaries. In *Abstract and Applied Analysis*, volume 2007, page 15, 2007.

[24] M. Nikolova. A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision*, 20(1):99–120, 2004.

[25] S.J. Osher, A. Sole, and L.A. Vese. Image decomposition and restoration using total variation minimization and the h^{-1} norm. *Multiscale Modeling and Simulation*, 1(3):349–370, 2003.

[26] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, 1992.

[27] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse problems*, 19:S165–S187, 2003.

[28] E. Tadmor, S. Nezzar, and L. Vese. A multiscale image representation using hierarchical (bv, l^2) decompositions. *Multiscale Modeling and Simulation*, 2:554–579, 2004.

[29] H. Triebel. *Interpolation Theory Function Spaces Differential Operators*. johann Ambrosius barth, 2nd edition, 1995.

[30] K.R. Vixie, K. Clawson, T.J. Asaki, G. Sandine, S.P. Morgan, and B. Price. Multiscale Flat Norm Signatures for Shapes and Images. *Applied Mathematical Sciences*, 4(14):667–680, 2010.