Noname manuscript No.
(will be inserted by the editor)

Semi-Inner-Products for Convex Functionals and Their Use in Image Decomposition

Guy Gilboa

Received: date / Accepted: date

Abstract Semi-inner-products in the sense of Lumer are extended to convex functionals. This yields a Hilbert-space like structure to convex functionals in Banach spaces. In particular, a general expression for semi-inner-products with respect to one homogeneous functionals is given. Thus one can use the new operator for the analysis of total variation and higher order functionals like total-generalized-variation (TGV). Having a semi-inner-product, an angle between functions can be defined in a straightforward manner. It is shown that in the one homogeneous case the Bregman distance can be expressed in terms of this newly defined angle. In addition, properties of the semi-inner-product of nonlinear eigenfunctions induced by the functional are derived. We use this construction to state a sufficient condition for a perfect decomposition of two signals and suggest numerical measures which indicate when those conditions are approximately met.

Keywords Semi-inner-product · Total variation · Nonlinear eigenfunctions · Image decomposition

1 Introduction

Formulating image-processing and computer-vision tasks as variational problems, has been used extensively, with great success for denoising, segmentation, optical flow, stereo matching, 3D reconstruction and more [3,18,16]. In those cases regularizing functionals are used to avoid non-physical solutions and to overcome problems related with noisy measurements. For images, depth and optical-flow maps, and many other modalities - the signals have inherent discontinuities. Therefore, an appropriate mathematical modeling should account for that. One-homogeneous functionals, specifically based on the L^1 norm, can cope well with discontinuities. The most classical one is the total variation (TV) functional as first introduced for image processing in [37] and, in recent years, the proposition of total-generalized-variation (TGV) [9] which has increased the applicability of such regularizers from essentially piecewise constant to piecewise smooth solutions.

Recently, there is an emerging branch of studies trying to use functionals in alternative ways, broadening their analytical scope and usability [25,8,13]. In this context solutions of nonlinear eigenvalue problems induced by the regularizer are assumed as the fundamental structuring elements. A nonlinear spectral theory is developed, where operations such as nonlinear low-pass and high-pass filters can be performed.

In this paper we introduce an additional necessary ingredient in nonlinear spectral analysis of functionals, which is a weaker form of the inner-product to Banach spaces. It is referred to as a *semi-inner-product* and was first introduced by Lumer in [32]. We define the properties of a semi-inner-product for functionals and present the formulation for the one-homogeneous case. We then introduce a notion of semi-inner-products of degree q, where for $q = 1/2$ this definition provides a useful construct. Properties of semi-inner-products in the case of nonlinear eigenfunctions are discussed, where things simplify considerably. Finally, we connect these new notions to the problem of image decomposition, see e.g. [34,38,4,5,27,39]. A necessary condition for perfect decomposition is stated and soft indicators of how well two signals can be decomposed using a regularizer and nonlinear spectral filtering are formulated.

Department of Electrical Engineering,
Technion - Israel Institute of Technology,
Haifa 32000, Israel E-mail: guy.gilboa@ee.technion.ac.il

1.1 Main contributions

The main contributions of the paper are:

1. Defining the properties of semi-inner-products for general convex functionals from which angles and orthogonality measures, with respect to a functional, can be derived.
2. Proposing a semi-inner-product formulation for the case of one-homogeneous functionals,
$$[u,v]_J := \langle u, p(v)\rangle J(v),\ p(v)\in\partial J(v).$$
3. Extending semi inner products to be of degree q and showing the applicability for $q=1/2$.
4. Showing that in the case of J being one-homogeneous the Bregman distance [10] can be related to the angle between the functions u and v by
$$D_J(u,v) = J(u)\left(1-\cos(\text{angle}(u,v))\right).$$
5. Connecting the semi-inner-product to image decomposition through the recently proposed variational spectral filtering approach [25] and presenting a sufficient condition for perfect decomposition of functions admitting the nonlinear eigenvalue problem (9), (in Th. 2).
6. Proposing two soft measures to estimate when a good decomposition is expected and validating these through numerical experiments.

2 Preliminaries

We will now summarize four mathematical concepts and notions which are at the basis of this manuscript:

1. The semi-inner-product of Lumer.
2. Convex one-homogeneous functionals and their unique properties.
3. Functions admitting a nonlinear eigenvalue problem induced by a convex regularizer.
4. A recent direction suggested in [25] of analyzing and processing regularization problems using a nonlinear spectral approach.

We will see at the last section how all these components are brought together in the analysis of signal decomposition based on regularizing functionals.

2.1 Semi-inner-product

In [32] Lumer introduced the notion of semi-inner-product (s.i.p.), where Giles [28] refined it by asserting the homogeneity property for both arguments. Semi-inner-products have been used in the analysis of Banach spaces [11,21] and in recent years extending Hilbert-space-like concepts in the context of machine-learning and classification [20,40,31]. In general, a s.i.p. is defined for complex-valued functions. Here we restrict ourselves to real-valued functions and follow the definitions of [20].

Definition 1 (Semi-inner-product) Let $(\mathcal{X},\|\cdot\|)$ be a real Banach space. A semi-inner-product on $\mathcal{X}$ is a real function $[u,v]$ on $\mathcal{X}\times\mathcal{X}$ with the properties:

1. (Linearity in the first argument)
$$[u_1+u_2,v] = [u_1,v]+[u_2,v],$$
2. (Homogeneity in the first argument)
$$[\alpha u, v] = \alpha[u,v],$$
3. (Norm-inducing)
$$[u,u] = \|u\|^2,$$
4. (Cauchy-Schwarz inequality)
$$[u,v]\le\|u\|\|v\|,$$
5. (Homogeneity in the second argument)
$$[u,\alpha v] = \alpha[u,v].$$

Giles [28] has added the fifth property (Homogeneity in the second argument), arguing that in the case of norms this does not impose additional restrictions and increases the structure. In the proposed generalizing to functionals, in some cases this condition will be omitted. In [28] a semi-inner-product for L^p norms $\|u\|_{L^p} = \left(\int_\Omega |u(x)|^p dx\right)^{1/p}$, $1<p<\infty$ was proposed

$$[u,v] := \int_\Omega u(x)v(x)|v(x)|^{p-2}dx\frac{1}{\|v\|_{L^p}^{p-2}}. \tag{1}$$

2.2 One-homogeneous functionals

Let $J(u)$ be a proper, convex, lower semi-continuous regularization functional $J:\mathcal{X}\to\mathbb{R}^+\cup\{\infty\}$ defined on Banach space $\mathcal{X}$. For J which is a one homogeneous functional we have

$$J(\alpha u) = |\alpha| J(u),\ \alpha\in\mathbb{R}. \tag{2}$$

We assume that $J(u)>0$ for $u\in\mathcal{X}\setminus\{0\}$ (as done for instance in [13]). This can be achieved by choosing $\mathcal{X}$ restricted in the right way (note that the null-space of a convex one-homogeneous functional is a linear subspace of $\mathcal{X}$, [8]). E.g. in the case of total variation regularization we would consider the subspace of functions with vanishing mean value. The general case can be reconstructed by adding appropriate nullspace components.

For the spectral representation, defined hereafter, we assume that the gradient descent equation, based on J, is well posed and that the initial condition f admits $J(f) < \infty$ and $\|f\|_{L^2} < \infty$.

Let $p(u) \in \mathcal{X}^*$ (where $\mathcal{X}^*$ is the dual space of $\mathcal{X}$) belong to the *subdifferential* of $J(u)$, defined by:

$$\partial J(u) := \\ \{p(u) \in \mathcal{X}^* \mid J(v) - J(u) \geq \langle v-u, p(u)\rangle, \forall v \in \mathcal{X}\}, \quad (3)$$

where $\langle \cdot, \cdot \rangle$ is the duality product from $X \times X^*$ to $\mathbb{R}$. We denote $p(u) \in \partial J(u)$, where an element $p(u)$ is referred to as a *subgradient*. For convex one homogeneous functionals it is well known [22] that

$$J(u) = \langle u, p(u)\rangle, \forall p(u) \in \partial J(u). \quad (4)$$

And also, for all $p(u) \in \partial J(u)$, $\mathbb{R} \ni \alpha \neq 0$, we have

$$\text{sgn}(\alpha) p(u) \in \partial J(\alpha u), \quad (5)$$

where $\text{sgn}(\cdot)$ is the signum function. From (3) and (4) we have that an element in the subdifferential of one-homogeneous functionals admits the following inequality:

$$J(v) \geq \langle v, p(u)\rangle, \forall p(u) \in \partial J(u),\, v \in \mathcal{X}. \quad (6)$$

In later sections we need a slight extension of this property, where the bound is with respect to the magnitude of the right-hand-side. Since $J(-v) = J(v)$ we can also plug $-v$ in (6) and get the bound $J(v) \geq -\langle v, p(u)\rangle$, hence

$$J(v) \geq |\langle v, p(u)\rangle|, \forall p(u) \in \partial J(u),\, v \in \mathcal{X}. \quad (7)$$

One-homogeneous functionals also admit the triangle inequality:

$$J(u+v) \leq J(u) + J(v). \quad (8)$$

This can be shown by $J(u+v) = \langle u+v, p(u+v)\rangle = \langle u, p(u+v)\rangle + \langle v, p(u+v)\rangle$ and using (6) we have $J(u) \geq \langle u, p(u+v)\rangle$ and $J(v) \geq \langle v, p(u+v)\rangle$.

2.3 Nonlinear Eigenfunctions

Let us begin by stating the nonlinear eigenvalue problem induced by a convex functional.

Definition 2 (Eigenfunctions and eigenvalues induced by $J(u)$) An eigenfunction u induced by the functional $J(u)$ admits the following equation,

$$\lambda u \in \partial J(u), \quad (9)$$

where $\lambda \in \mathbb{R}$ is the corresponding eigenvalue.

The analysis of eigenfunctions related to non- quadratic convex functionals was mainly concerned with the total variation (TV) regularization. In the analysis of the variational TV denoising, i.e. the ROF model from [37], Meyer [34] has shown an explicit solution for the case of a disk (an eigenfunction of TV), quantifying explicitly the loss of contrast and advocating the use of $TV - G$ regularization. Within the extensive studies of the TV-flow [1, 2, 7, 24] eigenfunctions of TV (referred to as *calibrable sets*) were analyzed and explicit solutions were given for several cases of eigenfunction spatial settings. In [15] an explicit solution of a disk for the inverse-scale-space flow is presented, showing its instantaneous appearance at a precise time point related to its radius and height.

Geometric understanding of TV eigenfunctions

In [1] a connection between the eigenvalue λ and the perimeter to area ratio is established for the total-variation (TV) case. Let us recall this relation. The TV functional is defined by

$$J_{TV}(u) = \sup_{\|\varphi\|_{L^\infty(\Omega)} \leq 1} \int_\Omega u \,\text{div}\, \varphi dx, \quad (10)$$

with $\varphi \in C_0^\infty$. For a convex set $A \subset \mathbb{R}^2$ let f_A be the indicator function of A where $f(x) = 1$ for any $x \in A$ and zero otherwise. If f_A is an eigenfunction (admits Eq. (9)) with respect to the TV functional then

$$\lambda = \frac{P(A)}{|A|}, \quad (11)$$

with $P(A)$ the perimeter of the set A and $|A|$ its area.

2.4 The TV Transform

In [25] a generalization of eigenfunction analysis to the total-variation case was proposed. We would like to decompose and process an input image $f(x) \in BV$ (where BV is the space of bounded variations in which J_{TV} is finite). This is done through TV gradient descent in the following way. Let $u(t;x)$ be the TV-flow solution [1], which stands for the gradient descent of the total variation energy $J_{TV}(u)$, with initial condition $f(x)$:

$$\partial_t u = -p, \qquad p \in \partial J_{TV}(u), \qquad u(t=0) = f(x). \quad (12)$$

The TV spectral representation (referred to also as TV transform) is defined by

$$\phi(t;x) := t\partial_{tt} u(t;x), \quad (13)$$

where $\partial_{tt} u$ is the second time derivative of the solution $u(t;x)$ of the TV flow (12).

We briefly discuss the regularity of ϕ. In [7] a comprehensive analysis is presented for the TV flow in $\mathbb{R}^N$. A strong solution is shown for the case $f \in L^2(\mathbb{R}^N)$ (Th. 2). Moreover, time regularity (Section 7 of [7]) based on semigroup estimates yields that:

$$u_t(t) \in L^2(\mathbb{R}^N) \text{ for } f \in L^2(\mathbb{R}^N), t \geq \varepsilon, \forall \varepsilon > 0,$$

and also

$$u_t(t) \in L^2(\mathbb{R}^N) \text{ for } f \in BV(\mathbb{R}^N), t \geq 0.$$

Note that a finite extinction time T is shown to hold for the TV-flow [1], where for all initial conditions $f \in L^2(\mathbb{R}^N)$ there exists $T \geq 0$ such that $\forall t > T$, $u_t \equiv 0$.

As ϕ can be a measure in the time domain we are mostly concerned with the integral form

$$\Phi_{t_1,t_2}(x) := \int_{t_1}^{t_2} \phi(t;x)dt, \;\; 0 \leq t_1 < t_2 < \infty. \tag{14}$$

This type of integration appears in the reconstruction formula (17) below as well as in all types of filters formulated by equations (18) - (22). Using integration by parts we have

$$\Phi_{t_1,t_2} = u_t(t_2)t_2 - u_t(t_1)t_1 - u(t_2) + u(t_1).$$

Thus we conclude that $\Phi_{t_1,t_2} \in L^2(\mathbb{R}^N)$ for any $t_1 \geq 0$ with $f \in BV(\mathbb{R}^N)$, or for any $t_1 > 0$ with $f \in L^2(\mathbb{R}^N)$. In any spatial discrete setting the input f is naturally of bounded variation and the integrals are well defined.

For a broader study on one-homogeneous spectral representations see [14]. Some regularity results presented in [14], related to the finite dimensional case, are summarized in Section 2.6, where it is shown that $\phi \in \left(W^{1,1}_{loc}(\mathbb{R}^+, \mathbb{R}^n)\right)^*$, in which expressions as in (16) below are admissible.

For $f(x)$ admitting (9), with a corresponding eigenvalue λ, one obtains a gradient flow (12) with the solution

$$u(t,x) = (1 - \lambda t)^+ f(x), \tag{15}$$

where $(q)^+ = q$ if $q > 0$ and 0 otherwise. See e.g. [7] (Th. 4) for a broader discussion and analysis. The spectral response becomes

$$\phi(t;x) = \delta(t - 1/\lambda) f(x), \tag{16}$$

where $\delta(\cdot)$ denotes a Dirac delta distribution. This should be understood as having the spectral representation with a concentrated measure at $t = \frac{1}{\lambda}$, or that $f(x)$ can be recovered by Φ_{t_1,t_2} with an integration over a small time range, $t_2 - t_1 = \Delta t$, where $t_1 < \frac{1}{\lambda} < t_2$.

In the general case, ϕ yields a continuum multiscale representation of the image, generalizing structure-texture decomposition methods like [34,36,5]. For simplicity we assume signals with zero mean $\bar{f} = \frac{1}{\Omega}\int_\Omega f(x)dx = 0$. One can reconstruct the original image $f \in BV$ by:

$$f(x) = \int_0^\infty \phi(t;x)dt. \tag{17}$$

Given a transfer function $H(t) \in \mathbb{R}$, image filtering can be performed by

$$f_H(x) := \int_0^\infty H(t)\phi(t;x)dt. \tag{18}$$

Simple useful filters are ones which either retain or diminish completely scales up to some cutoff scale. The (ideal) low-pass-filter (LPF) can be defined by Eq. (18) with $H(t) = 1$ for $t \geq t_c$ and 0 otherwise, or

$$LPF_{t_c}(f) := \int_{t_c}^\infty \phi(t;x)dt. \tag{19}$$

Its complement, the (ideal) high-pass-filter (HPF), is defined by

$$HPF_{t_c}(f) := \int_0^{t_c} \phi(t;x)dt. \tag{20}$$

Similarly, band-(pass/stop)-filters are filters with low and high cut-off scale parameters ($t_1 < t_2$)

$$BPF_{t_1,t_2}(f) := \int_{t_1}^{t_2} \phi(t;x)dt, \tag{21}$$

$$BSF_{t_1,t_2}(f) := \int_0^{t_1} \phi(t;x)dt + \int_{t_2}^\infty \phi(t;x)dt. \tag{22}$$

The spectrum $S_f(t)$ corresponds to the amplitude of each scale of the input f:

$$S_f(t) := \|\phi(t;x)\|_{L^1(\Omega)} = \int_\Omega |\phi(t;x)|dx. \tag{23}$$

In Fig. 1 an example of spectral TV processing is shown with the response of the four filters defined above in Eqs. (19) through (22).

2.5 Generalized Transform

In [13] the spectral TV framework was generalized in several ways. First the theory was extended to a wider class of one-homogeneous functionals.

For the general gradient flow of a one-homogeneous functional J, where J admits the condition of Section 2.2, we have

$$\partial_t u(t) = -p(t), \qquad p(t) \in \partial J(u(t)), \; u(0) = f, \tag{24}$$

the spectral transform $\phi(t)$, the eigenfunction response, the reconstruction and the filtering, Eqs. (13), (15),

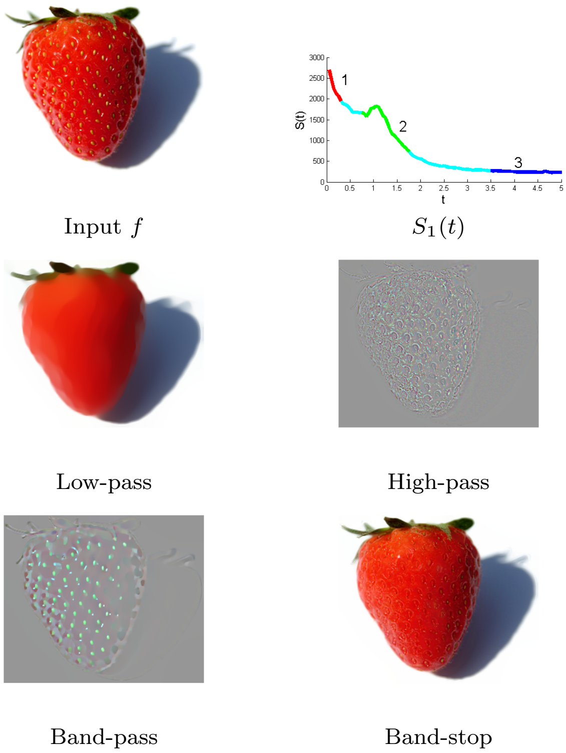

Fig. 1 Total-variation spectral filtering example. The input image (top left) is decomposed into its $\phi(t)$ components, the corresponding spectrum $S_1(t)$ is on the top right. Integration of the ϕ's over the t domains 1, 2 and 3 (top right) yields high-pass, band-pass and low-pass filters, respectively. The band-stop filter (bottom right) is the complement integration domain of region 2. Taken from [13].

(16), (17), (18) all generalize in a straightforward manner, retaining the same expressions.

A new spectrum was defined by

$$S_2(t) = t\sqrt{\frac{d^2}{dt^2}J(u(t))} = \sqrt{\langle \phi(t), 2tp(t)\rangle}, \tag{25}$$

for which an analogue of the Parseval identity can be derived

$$\|f\|^2 = \int_0^\infty S_2(t)^2 \, dt.$$

An orthogonality property was shown

$$\langle \phi(t), u(t)\rangle = 0, \;\forall t > 0. \tag{26}$$

An overview of these ideas with relations to some classical signal processing methods are presented in [26].

2.6 Some Regularity Results

A comprehensive analysis of one-homogeneous transforms is still under way. However, in [14] several results were established for the finite dimensional setting (spatially discrete, time continuous). We summarize them here.

Let J be a proper, convex, lower semi-continuous, absolutely one-homogeneous function on $\mathbb{R}^n$, ($J : \mathbb{R}^n \to \mathbb{R}$). We use the gradient flow as in (24) with arbitrary initial condition $f \in \mathbb{R}^n$. Here we show the more general case, where the null space is not restricted. $\mathcal{N}(J)$ will denote the nullspace of J

$$\mathcal{N}(J) = \{u \in \mathbb{R}^n | J(u) = 0\},$$

and P_0 is the projection operator on $\mathcal{N}(J)$. The spectral representation is

$$\phi(t) = t\partial_{tt}u(t),$$

where $u(t)$ is the solution of (24).

Proposition 1 (Finite extinction time) *There exists a time $T < \infty$ such that $u(T)$ determined via* (24) *meets*

$$u(T) = P_0(f).$$

This can be shown by observing that the time derivative of the square L^2 norm of $u(t)$ is strictly positive as long as u is not in $\mathcal{N}(J)$. A complete proof is in Prop. 3 of [14].

Moreover, based on the theory of gradient flows (cf. [23]), we have that $\partial_t u(t) \in L^\infty$. We can thus state a regularity result for $u(t)$ and $\phi(t)$.

Proposition 2 (Regularity of u and ϕ) *The function $u : \mathbb{R}^+ \to \mathbb{R}^n$ is Lipschitz continuous. The spectral representation ϕ satisfies*

$$\phi \in \left(W^{1,1}_{loc}(\mathbb{R}^+, \mathbb{R}^n)\right)^*.$$

Essentially it is shown that the integral

$$\int_0^\infty v(t)\cdot\phi(t)dt = -\int_0^\infty (t\partial_t v(t) + v(t))\partial_t u(t)dt$$

is well defined for any test function $v \in W^{1,1}_{loc}(\mathbb{R}^+, \mathbb{R}^n)$. More details are in Propositions 2 and 4 in [14]. Similar arguments lead to showing the reconstruction of the input data by

$$f = P_0(f) + \int_0^\infty \phi(t)dt, \tag{27}$$

and for expressing a filtering operation $(w_0, w(t))$, where $w_0 \in \mathbb{R}$ and $w(t) \in W^{1,1}_{loc}$, by

$$f_w = w_0 P_0(f) + \int_0^\infty w(t)\phi(t)dt, \tag{28}$$

which can be expressed through integration by parts also as

$$f_w = w_0 P_0(f) - \int_0^\infty (tw'(t) + w(t))\partial_t u(t)dt.$$

With the preliminary settings and definitions in place we can now continue to the main contributions of the paper concerning generalized s.i.p's.

3 A semi-inner-product for convex functionals

Let us define a semi-inner-product for convex functionals, in a similar manner to Definition 1. As we will show later, a function which admits the properties below may not be unique. Therefore, in a similar manner to the subdifferential, we allow the semi-inner-product to be a set of possibly more than one element. We denote by $[u,v]_J$ an element and by $\{[u,v]_J\}$ the set of admissible s.i.p.'s. We will later see for the one-homogeneous case that when a specific subgradient of the second argument is chosen the s.i.p. is unique.

Definition 3 (Semi-inner-product of a convex functional, partial homogeneity) Let J be a convex functional $J : \mathcal{X} \to \mathbb{R}^+ \cup \{\infty\}$ defined on a Banach space $\mathcal{X}$. A semi-inner-product with partial homogeneity on $\mathcal{X}$ is a real function $[u,v]_J$ on $\mathcal{X} \times \mathcal{X}$ with the properties:

1. (Linearity in the first argument)
$$[u_1 + u_2, v]_J = s + q,\ s \in \{[u_1, v]_J\},\ q \in \{[u_2, v]_J\}.$$
2. (Homogeneity in the first argument)
$$[\alpha u, v]_J \in \{\alpha[u, v]_J\},\ \alpha \in \mathbb{R}.$$
3. (Functional-inducing)
$$[u, u]_J = J^2(u).$$
4. (Cauchy-Schwarz-type inequality)
$$\sqrt{|[u, v]_J[v, u]_J|} \le J(u)J(v). \quad (29)$$

A stricter definition, with homogeneity in both arguments is defined by

Definition 4 (Semi-inner-product of a convex functional, full homogeneity) Following the same notations of Def. 3, $[u,v]_J$ is a semi-inner-product with full homogeneity if it admits all the properties of Def. 3 and in addition:

5. (Homogeneity in the second argument)
$$[u, \alpha v]_J \in \{\alpha[u, v]_J\}.$$

3.1 Semi-inner-product formulations

It can be verified that for functionals of the form

$$J_{\mathcal{H}}(u) = \|u\|^2_{\mathcal{H}}, \quad (30)$$

with $\{\|\cdot\|_{\mathcal{H}}, \langle\cdot,\cdot\rangle_{\mathcal{H}}\}$ a Hilbert-space norm and inner-product, respectively, a semi-inner-product in the sense of Def. 3 is:

$$[u, v]_{J_{\mathcal{H}}} := \langle u, v\rangle_{\mathcal{H}} \|v\|^2_{\mathcal{H}}. \quad (31)$$

However, our main focus of the paper is devoted to functionals not based on a Hilbert-space but on smoothing, discontinuity preserving functionals such as the total-variation or the total-generalized-variation. Those functionals are extremely useful in processing images and many other types of signals with inherent discontinuities, such as depth-maps or optical-flow fields. Those functionals are one-homogeneous and therefore a full homogeneity semi-inner-product can be defined.

Theorem 1 *Let J be a convex one-homogeneous functional, admitting the conditions defined in Section 2.2, and $p(v) \in \partial J(v)$ a subgradient. Then a corresponding semi-inner-product with full homogeneity in the sense of Def. 4 is*

$$[u, v]_J^{p(v)} := \langle u, p(v)\rangle J(v), \quad (32)$$

where $\langle\cdot,\cdot\rangle$ is the duality product of $u \in X$ and $p(v) \in X^*$.

Proof Linearity and homogeneity in the first argument are straightforward consequences of using the duality product. We now want to show the property of homogeneity in the second argument. We use Eqs. (2) and (5) to have $p(v) \in \partial J(v)$ and $p(\alpha v) \in \partial J(\alpha v)$ with the relation $p(\alpha v) = \text{sgn}(\alpha)p(v)$ and therefore

$$\begin{aligned}[u, \alpha v]_J^{p(\alpha v)} &= \langle u, p(\alpha v)\rangle J(\alpha v) \\ &= \langle u, \text{sgn}(\alpha)p(v)\rangle |\alpha| J(v) \\ &= \alpha[u, v]_J^{p(v)} \in \{\alpha[u, v]_J\}.\end{aligned}$$

Using (4) we get $[u, u]_J^{p(u)} = \langle u, p(u)\rangle J(u) = J^2(u)$. Finally for the Cauchy-Schwarz property, using (7) we have $\forall p(u) \in \partial J(u)$, $J(v) \ge |\langle v, p(u)\rangle|$ and $\forall p(v) \in \partial J(v)$, $J(u) \ge |\langle u, p(v)\rangle|$, therefore

$$|[u, v]_J^{p(v)}| = |\langle u, p(v)\rangle| J(v) \le J(u)J(v)$$

and also

$$|[v, u]_J^{p(u)}| = |\langle v, p(u)\rangle| J(u) \le J(v)J(u).$$

As noted in the proof, for the one-homogeneous s.i.p. a classical Cauchy-Schwarz inequality holds

$$|[u, v]_J^{p(v)}| \le J(u)J(v). \quad (33)$$

As an example, let us take the L^q norm, $J_{L^q}(u) = \|u\|_{L^q}$, for $1 < q < \infty$. Then $p(u) = |u|^{q-2}u\|u\|^{1-q}_{L^q}$ and Eq. (32) coincides with (1).

3.2 Generalized notions of angle and orthogonality

With the s.i.p. one can define an angle between functions u and v. For brevity, we will omit the superscript $p(v)$ when the context is clear. In the one-homogeneous case, using the above inequality, we can define the angle between u and v (for $u \neq 0$, $v \neq 0$, $J(u) > 0$, $J(v) > 0$) by

$$\text{angle}(u, v) := \cos^{-1}\left(\frac{\lfloor u, v \rfloor_J}{J(u)J(v)}\right). \tag{34}$$

Note that there is no symmetry in the above definition, so in general angle$(u, v) \neq$ angle(v, u).

For a symmetric angle expression, there are two main options, an algebraic mean,

$$\text{angle}_{sym-a}(u, v) := \cos^{-1}\left(\frac{\frac{1}{2}(\lfloor u, v \rfloor_J + \lfloor v, u \rfloor_J)}{J(u)J(v)}\right), \tag{35}$$

and a geometric mean (which also applies for the general convex case, in which the inequality of (29) holds),

$$\text{angle}_{sym-g}(u, v) := \cos^{-1}\left(\frac{\mathcal{S}(\lfloor u, v \rfloor_J, \lfloor v, u \rfloor_J)}{J(u)J(v)}\right), \tag{36}$$

where $\mathcal{S}(a, b) := \text{sgn}(ab)\sqrt{|ab|}$ is a signed square-root.

Orthogonality of two functions can be expressed as having an angle of $\frac{\pi}{2}$ between them. In the case of the nonsymmetric angle of (34) we refer to u as *orthogonal* to v if $0 \in \{[u, v]_J\}$ and to v as *orthogonal* to u if $0 \in \{[v, u]_J\}$.

Definition 5 (Full orthogonality (FO)) We assume $u \neq 0$, $v \neq 0$, $J(u) > 0$, $J(v) > 0$. (u, v) are fully orthogonal if $0 \in \{[u, v]_J\}$ and $0 \in \{[v, u]_J\}$.

3.3 A semi-inner-product of degree q

A slight generalization of the s.i.p. defined above is a *semi inner product of degree* q. Essentially the norm and Cauchy-Schwarz properties are raised to the q's power. The formal definition is as follows.

Definition 6 (Semi-inner-product of degree q of a convex functional) Let J be a convex functional $J : \mathcal{X} \to \mathbb{R}^+ \cup \{\infty\}$ defined on a Banach space $\mathcal{X}$. A semi-inner-product of degree q on $\mathcal{X}$ is a real function $[u, v]_{J,q}$ on $\mathcal{X} \times \mathcal{X}$ with the properties:

1. (Linearity in the first argument)

   $$[u_1+u_2, v]_{J,q} = s+q,\ s \in \{[u_1, v]_{J,q}\},\ q \in \{[u_2, v]_{J,q}\},$$

2. (Homogeneity in the first argument)

   $$[\alpha u, v]_{J,q} \in \{\alpha[u, v]_{J,q}\},\ \alpha \in \mathbb{R},$$

3. (Functional-inducing)

   $$[u, u]_{J,q} = J^{2q}(u),$$

4. (Cauchy-Schwarz-type inequality)

   $$\sqrt{|[u, v]_{J,q} \cdot [v, u]_{J,q}|} \leq J^q(u)J^q(v).$$

We examine more closely the s.i.p. of degree half ($q = \frac{1}{2}$) abbreviated h.s.i.p. For brevity we denote a special symbol for it

$$\lfloor u, v \rfloor_J := [u, v]_{J,1/2}.$$

For the h.s.i.p. property 3 in Def. 6 becomes $\lfloor u, u \rfloor_J = J(u)$, and property 4 becomes $\lfloor u, v \rfloor_J \lfloor v, u \rfloor_J \leq J(u)J(v)$.

In the case of square Hilbert-space functionals, Eq. (30), we get

$$\lfloor u, v \rfloor_{J_\mathcal{H}} = \langle u, v \rangle_\mathcal{H} = \frac{[u, v]_{J_\mathcal{H}}}{J_\mathcal{H}(v)}.$$

We will now examine the one-homogeneous case.

Proposition 3 *Let J be a convex one-homogeneous functional and $p(v) \in \partial J(v)$ a subgradient. Then a corresponding semi-inner-product of degree* $1/2$ *in the sense of Def. 6 is*

$$\lfloor u, v \rfloor_J^{p(v)} := \langle u, p(v) \rangle. \tag{37}$$

Proof The proof is mainly similar to the one of Th. 1. For the third property we use Eq. (4) and for the fourth property, using (7), we have $|\lfloor u, v \rfloor_J^{p(v)}| \leq J(u)$ and $|\lfloor v, u \rfloor_J^{p(v)}| \leq J(v)$.

Note that the s.i.p. of (32) is simply the h.s.i.p. multiplied by $J(v)$,

$$[u, v]_J^{p(v)} = \lfloor u, v \rfloor_J^{p(v)} J(v). \tag{38}$$

Following Eqs. (4), (5), (7) we have for the one-homogeneous h.s.i.p. the following properties:

$$\lfloor u, u \rfloor_J = J(u), \tag{39}$$

$$\lfloor u, \alpha v \rfloor_J \in \{\text{sgn}(\alpha)\lfloor u, v \rfloor_J\}, \qquad \mathbb{R} \ni \alpha \neq 0, \tag{40}$$

$$|\lfloor u, v \rfloor_J| \leq \lfloor u, u \rfloor_J = J(u), \qquad \forall v \in \mathcal{X}. \tag{41}$$

3.4 Relation to Bregman distance

We will now show the close connection between the Bregman distance (also called Bregman divergence) and the s.i.p. in the one-homogeneous case.

Let us first recall the Bregman distance definition [10]. For a convex functional J and a subgradient $p(v) \in \partial J(v)$, the (generalized) Bregman distance is

$$D_J^{p(v)}(u,v) := J(u) - J(v) - \langle u - v, p(v) \rangle. \tag{42}$$

This is not necessarily a distance in the standard sense, as it is not necessarily symmetric and does not admit the triangle inequality, however it is guaranteed to be non-negative and it is identically zero for $u = v$. For J the square L^2 norm we get the Euclidean distance squared,

$$D_{\|\cdot\|^2}(u,v) = \|u - v\|^2.$$

Other known similarity measures, such as the KL- divergence or the Mahalanobis distance, can also be derived from (42) with appropriate functionals [6]. This measure has been widely used in the theoretical analysis of classification, clustering and convex optimization algorithms, see e.g. [6,17,30,19]. Specifically for image processing, a significant branch of studies has presented iterative variational solutions, new evolution formulations and numerical solvers based on the Bregman distance, especially in relation to total-variation and other one-homogeneous regularizing functionals [35,15,29,41,33], see a recent review of the topic in [12].

In the one-homogeneous case we use the relation $J(v) = \langle v, p(v) \rangle$ and the expression in (42) simplifies to

$$D_J^{p(v)}(u,v)|_{(\text{ 1-hom})} = J(u) - \langle u, p(v) \rangle. \tag{43}$$

It is straightforward in this case to infer the relation to the s.i.p. and h.s.i.p.,

$$\begin{aligned} D_J^{p(v)}(u,v)|_{(\text{ 1-hom})} &= J(u) - \tfrac{[u,v]_J^{p(v)}}{J(v)} \\ &= J(u) - \lfloor u,v \rfloor_J^{p(v)}. \end{aligned} \tag{44}$$

An interesting interpretation of the Bregman distance is with respect to the angle between the functions u and v,

$$D_J^{p(v)}(u,v)|_{(\text{ 1-hom})} = J(u)\left(1 - \cos(\text{angle}(u,v))\right), \tag{45}$$

with the angle defined in (34). With this expression we can immediately get the upper and lower bounds

$$0 \le D_J^{p(v)}(u,v) \le 2J(u).$$

Moreover, the interpretation of the Bregman distance is of having direct relation to the angle between the functions; the Bregman distance is zero for zero angle and is monotonically increasing with angle, reaching the maximum at $\text{angle}(u,v) = \pi$.

An extension of this relation which applies to the general convex case is not known at this point. We now define the final notions needed for the decomposition theorem.

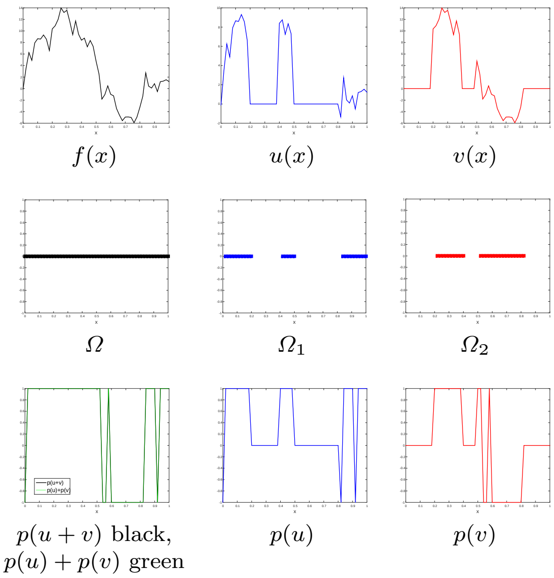

Fig. 2 LIS example for the L^1 norm.

Definition 7 (Linearity in the subdifferential (LIS)) (u,v) are linear in the subdifferential if for any $\mathbb{R} \ni \{\alpha_1, \alpha_2\} \neq 0$ there exist $p(\alpha_1 u + \alpha_2 v) \in \partial J(\alpha_1 u + \alpha_2 v)$, $p(\alpha_1 u) \in \partial J(\alpha_1 u)$, $p(\alpha_2 v) \in \partial J(\alpha_2 v)$, such that

$$p(\alpha_1 u + \alpha_2 v) = p(\alpha_1 u) + p(\alpha_2 v). \tag{46}$$

(LIS) implies the h.s.i.p. is linear in the second argument. If the pair (v_1, v_2) admit the (LIS) condition then there exist 3 subgradient elements $p(\alpha_1 v_1 + \alpha_2 v_2)$, $p(\alpha_1 v_1)$, $p(\alpha_2 v_2)$ such that for all $u \in \mathcal{X}$ we have

$$\begin{aligned} &\lfloor u, \alpha_1 v_1 + \alpha_2 v_2 \rfloor_J^{p(\alpha_1 v_1 + \alpha_2 v_2)} \\ &= \lfloor u, \alpha_1 v_1 \rfloor_J^{p(\alpha_1 v_1)} + \lfloor u, \alpha_2 v_2 \rfloor_J^{p(\alpha_2 v_2)}. \end{aligned} \tag{47}$$

This is shown by writing the left-hand-side, according to (37), as $\langle u, p(\alpha_1 v_1 + \alpha_2 v_2) \rangle$ and using (46).

We give a simple example of two signals admitting (LIS) in the case of J being the L^1 norm, for the 1D case within the unit interval $\Omega = [0,1]$. Let $f(x)$ be a real function in Ω, $f : \Omega \to \mathbb{R}$. We define the following two functions: $u(x) = f(x)$ if $x \in [0, 0.5)$ and 0 otherwise, $v(x) = f(x)$ if $x \in [0.5, 1]$ and 0 otherwise. Then it can be verified that u and v are (LIS). Any other partition $\Omega_1 \subset \Omega$, $\Omega_2 = \Omega \setminus \Omega_1$ for u and v will produce similar results, see Fig. 2.

Definition 8 (Independent functions) (u,v) are independent functions if they are fully orthogonal (FO) and linear in the subdifferential (LIS), according to Def. 5 and Def. 7, respectively.

We shall now show that for one-homogeneous functionals, all functions which are (LIS) are also (FO) and are therefore independent.

Proposition 4 *Let J be a convex one-homogeneous functional. If the pair (u,v) are (LIS) according to Def. 7, then (u,v) are (FO) and therefore are independent (Def. 8).*

Proof From (41) we have

$$J(u) \geq \lfloor u, u+v \rfloor_J,$$

using (LIS), for some fixed subgradients $p(u+v)$, $p(u)$, $p(v)$, we have

$$\begin{aligned} J(u) \geq \lfloor u, u+v \rfloor_J^{p(u+v)} &= \lfloor u, u \rfloor_J^{p(u)} + \lfloor u, v \rfloor_J^{p(v)} \\ &= J(u) + \lfloor u, v \rfloor_J^{p(v)}. \end{aligned}$$

We therefore have $\lfloor u, v \rfloor_J^{p(v)} \leq 0$. On the other hand, taking $\alpha_1 = 1, \alpha_2 = -1$ in Def. 7 we get that also $u, -v$ are (LIS). In this case, using (40), we reach $\lfloor u, v \rfloor_J^{p(v)} \geq 0$. We can conclude that $\lfloor u, v \rfloor_J^{p(v)} = 0$ hence

$$\{[u,v]_J\} \ni \lfloor u, v \rfloor_J^{p(v)} J(v) = 0.$$

The same arguments hold for the pair (v,u).

An interesting characteristic of independent functions is that they reach the upper bound of the triangle inequality (Eq. (8)).

Proposition 5 *Let J be a convex one-homogeneous functional. If (u,v) are independent (Def. 8) then $J(u+v) = J(u) + J(v)$.*

Proof

$$\begin{aligned} J(u+v) &= \lfloor u+v, u+v \rfloor_J^{p(u+v)} \\ &\overset{\text{LIS}}{=} \lfloor u, u \rfloor_J^{p(u)} + \lfloor u, v \rfloor_J^{p(v)} + \lfloor v, u \rfloor_J^{p(u)} + \lfloor v, v \rfloor_J^{p(v)} \\ &\overset{\text{FO}}{=} \lfloor u, u \rfloor_J^{p(u)} + \lfloor v, v \rfloor_J^{p(v)} \\ &= J(u) + J(v). \end{aligned}$$

3.5 S.I.P. for eigenfunctions

We will analyze now the case of nonlinear eigenfunctions (functions admitting (9)). Here we restrict ourselves to the simple case of L^2 embedding, where the duality product is the L^2 inner product (or ℓ^2 in finite dimensions), denoted by $\langle \cdot, \cdot \rangle_2$. Under this setting, things simplify considerably. For $\lambda u \in \partial J(u)$ we get

$$J(u) = \langle u, p(u) \rangle_2 = \langle u, \lambda u \rangle_2 = \lambda \|u\|_2^2,$$

where $\|\cdot\|_2$ is the 2 norm. For semi-inner-products we will often use the subgradient element corresponding to the eigenfunction, this will be denote by a superscript $\lambda_v v$. We therefore have the following relations for the s.i.p and h.s.i.p: For the s.i.p., for any $u \in \mathcal{X}$, $\lambda_v v \in \partial J(v)$,

$$\lambda_v^2 \langle u, v \rangle_2 \|v\|_2^2 = [u,v]_J^{\lambda_v v} \in \{[u,v]_J\}, \tag{48}$$

and for the h.s.i.p.,

$$\lambda_v \langle u, v \rangle_2 = \lfloor u, v \rfloor_J^{\lambda_v v} \in \{\lfloor u, v \rfloor_J\}. \tag{49}$$

Another consequence is related to orthogonality.

Proposition 6 *1. For any $u \in \mathcal{X}$, $\lambda_v v \in \partial J(v)$, $\lambda_v > 0$, $\|v\|_2 > 0$,*

$$[u,v]_J^{\lambda_v v} = 0 \quad \textit{iff} \quad \langle u, v \rangle_2 = 0.$$

2. For $p(u) = \lambda_u u \in \partial J(u)$, $p(v) = \lambda_v v \in \partial J(v)$, $\lambda_u, \lambda_v > 0$, $\|u\|_2, \|v\|_2 > 0$, the following statements are identical:
(a) $[u,v]_J^{\lambda_v v} = 0$,
(b) $[v,u]_J^{\lambda_u u} = 0$,
(c) $\lfloor u, v \rfloor_J^{\lambda_v v} = 0$,
(d) $\lfloor v, u \rfloor_J^{\lambda_u u} = 0$,
(e) $\lfloor u, v \rfloor_J^{\lambda_v v} + \lfloor v, u \rfloor_J^{\lambda_u u} = 0$,
(f) $\langle u, v \rangle_2 = 0$.

Proof The first part is an immediate consequence of Eq. (48). For the second part, let us write the equivalent of statements (a) through (e):

(A) $[u,v]_J^{\lambda_v v} = \lambda_v^2 \langle u, v \rangle_2 \|v\|_2^2$,
(B) $[v,u]_J^{\lambda_u u} = \lambda_u^2 \langle v, u \rangle_2 \|u\|_2^2$,
(C) $\lfloor u, v \rfloor_J^{\lambda_v v} = \lambda_v \langle u, v \rangle_2$,
(D) $\lfloor v, u \rfloor_J^{\lambda_u u} = \lambda_u \langle v, u \rangle_2$,
(E) $\lfloor u, v \rfloor_J^{\lambda_v v} + \lfloor v, u \rfloor_J^{\lambda_u u} = (\lambda_v + \lambda_u) \langle u, v \rangle_2$.

We observe that in the case where both u and v are eigenfunctions all expressions reduce to the L^2 inner product up to a strictly positive multiplicative factor and are therefore identical when $\langle u, v \rangle_2 = 0$.

4 Decomposition

Let f_1, f_2 be two functions in $\mathcal{X}$ and $f = f_1 + f_2$. Naturally a decomposition from a single measurement f into two signals f_1 and f_2 is not possible in general. One should use some a priori knowledge and assumptions on the signals (depicted in the choice of the regularizer J). A classical decomposition problem is how and under what conditions we can decompose f into f_1 and f_2. This issue is significant in signal processing, for instance when f_1 is the signal and f_2 is noise or for structure-texture decomposition, where f_1 is structure and f_2 is texture (assumed to be additive). We will try to give an answer to this using the spectral filtering technique and conditions from the above framework.

We can now state a sufficient condition for spectral filtering to perfectly decompose f into f_1 and f_2.

Theorem 2 *Let J be a one-homogeneous functional as defined in Section 2.2. If f_1, f_2 are eigenfunctions with corresponding eigenvalues λ_1, λ_2, with $\lambda_1 < \lambda_2$, independent in the sense of Def. 8, then $f = f_1 + f_2$ can be perfectly decomposed into f_1 and f_2 using the following spectral decomposition: $f_1 = LPF_{\frac{1}{\lambda_c}}(f)$, $f_2 = HPF_{\frac{1}{\lambda_c}}(f)$ with $\lambda_1 < \lambda_c < \lambda_2$.*

Proof The theme of the proof is to show that we get an additive spectral response

$$\phi(t,x) = \delta(t - 1/\lambda_1)f_1(x) + \delta(t - 1/\lambda_2)f_2(x)$$

and therefore the spectral filtering proposed above (Eqs. (16),(19), (20) ,which hold for the general one-homogeneous case) decomposes f correctly.

We examine the gradient flow (24) with initial conditions $f = f_1 + f_2$. Let us show that given the above assumptions the solution is

$$u(t,x) = (1 - \lambda_1 t)^+ f_1(x) + (1 - \lambda_2 t)^+ f_2(x). \tag{50}$$

It is easy to see that for (50) the first time derivative is

$$\partial_t u(t,x) = \begin{cases} -\lambda_1 f_1(x) - \lambda_2 f_2(x), & 0 \le t < 1/\lambda_2 \\ -\lambda_1 f_1(x) - \lambda_2 f_2(x), & 1/\lambda_2 \le t < 1/\lambda_1 \\ 0, & 1/\lambda_1 \le t \end{cases}$$

We now need to check the subdifferential. We do this for $0 \le t < 1/\lambda_2$, similar results can be shown for the other time intervals. We denote by $p(\cdot)$ an element in $\partial J(\cdot)$.

$$\begin{aligned} \partial J(u(t)) = \; & \partial J\left((1 - \lambda_1 t)f_1(x) + (1 - \lambda_2 t)f_2(x)\right) \\ \overset{\text{LIS}}{\ni} \; & p((1 - \lambda_1 t)f_1(x)) + p((1 - \lambda_2 t)f_2(x)) \\ \overset{\text{Eq.(5)}}{=} \; & p(f_1(x)) + p(f_2(x)) \\ \overset{\text{Eq.(9)}}{=} \; & \lambda_1 f_1(x) + \lambda_2 f_2(x) \\ = \; & -\partial_t u(t,x). \end{aligned}$$

We can conclude that two eigenfunctions with different eigenvalues which are independent, with respect to the regularizer J, can be perfectly decomposed using spectral decomposition based on J.

4.1 Decomposition measures

The conditions stated in the above theorem are somewhat strict. We would like to have a soft measure for the independence of two signals which attains the value 1 for completely independent signals (in the sense of Def. 8) and 0 for completely correlated signals. It is expected that this measure will indicate how well two signals can be decomposed.

4.2 Orthogonality measure

Let an orthogonality indicator be defined by

$$\mathcal{O}(u,v) = 1 - \frac{\sqrt{|[u,v]_J [v,u]_J|}}{J(u)J(v)}. \tag{51}$$

We have that $0 \le \mathcal{O}(u,v) \le 1$ and $\mathcal{O} = 1$ in the orthogonal case, if either $[u,v]_J = 0$ or $[v,u]_J = 0$. For the fully correlated case $v = au$, $a > 0$, we get $\mathcal{O}(u, au) = 0$.

4.3 LIS measure

Here a more direct relation to the (LIS) property is defined. We measure how different is $p(u+v)$ from $p(u) + p(v)$. This is done in terms of h.s.i.p.,

$$\begin{aligned} E(u,v) &:= \lfloor u+v, u \rfloor_J + \lfloor u+v, v \rfloor_J - \lfloor u+v, u+v \rfloor_J \\ &= \langle u+v, p(u) + p(v) - p(u+v) \rangle. \end{aligned} \tag{52}$$

We show below that $E \le J(u+v)$. Also we have that $E \to 0$ as $p(u) + p(v) \to p(u+v)$. A possible indicator L for the (LIS) property can therefore be

$$L(u,v) := 1 - \frac{|E(u,v)|}{J(u+v)}. \tag{53}$$

Let us show that

$$E(u,v) \le J(u+v).$$

From (41) we have $\lfloor u+v, u \rfloor_J \le \lfloor u+v, u+v \rfloor_J$ and $\lfloor u+v, v \rfloor_J \le \lfloor u+v, u+v \rfloor_J$, where $\lfloor u+v, u+v \rfloor_J = J(u+v)$. Note also that for the fully correlated case, $v = au$, $a > 0$, we get $L(u, au) = 0$.

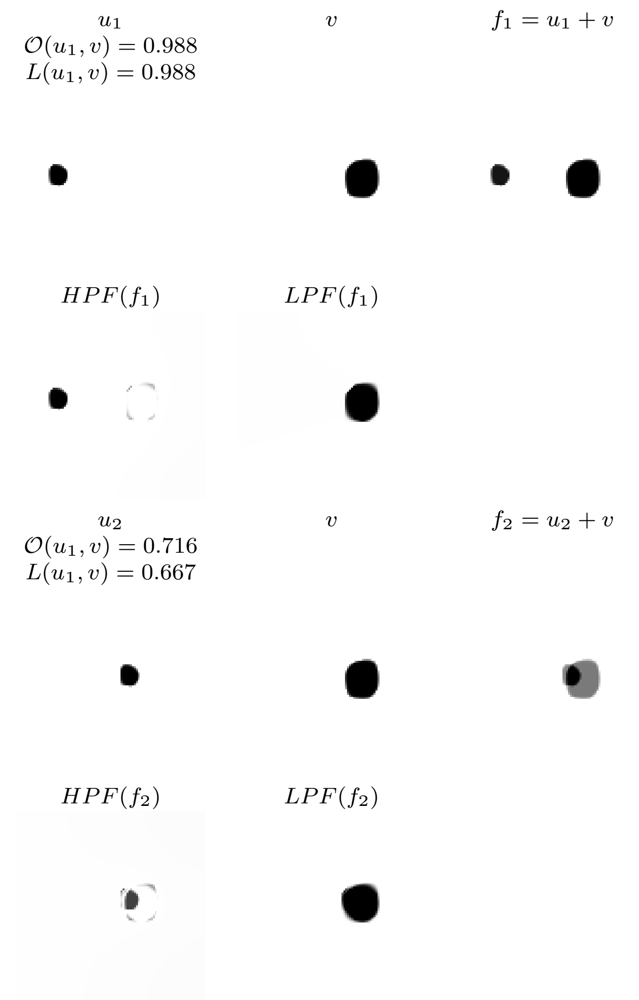

Fig. 3 Separating blobs of different scale using spectral filtering.

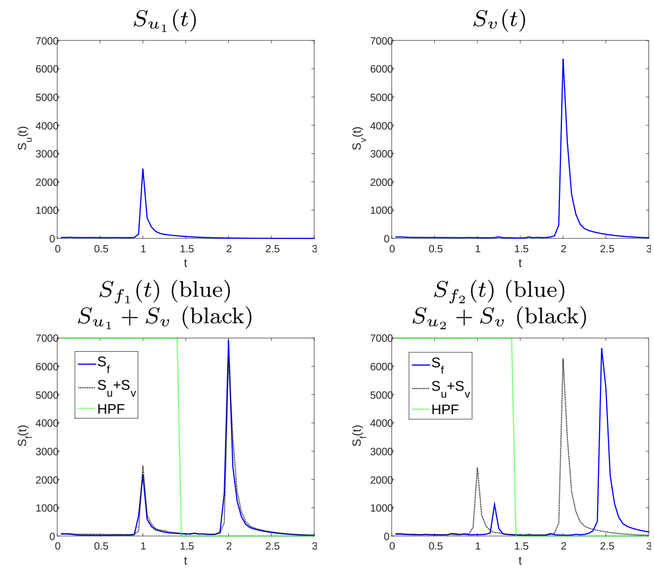

Fig. 4 Spectra of the different blob signals.

5 Experiments

The following experiments are performed to show the behavior of the soft measures described in the previous section for the TV functional. We compare the orthogonality measure $\mathcal{O}(u, v)$, Eq. (51), and the LIS measure $L(u, v)$, Eq. (53).

In Figs. 3 and 4 two cases are shown. In the first one (Fig. 3 top 2 rows) $f_1 = u_1 + v$, where u_1 and v are two blobs which are spatially well separated. The decomposition indicators are close to 1 ($\mathcal{O}(u_1, v) = L(u_1, v) = 0.988$). A high-pass-filter, as defined in (20), was used to separate u_1 with a cutoff between the peaks, see the green line in Fig. 4, bottom left, which visualizes the filter transfer function. One can observe a relatively good separation (with some residual of v as it is not a precise eigenfunction). In Fig. 4 the spectrum of u_1 and v are shown and the spectrum of their sum superimposed on the spectrum of f_1 (bottom left), which are close to identical.

The case of overlapping signals u_2 and v, $f_2 = u_2 + v$, is shown as well with significantly lower $\mathcal{O}$ and L indicators and lower quality decomposition (the spectra are also not additive).

In Fig. 5 u and v are constructed to be precise discrete eigenfunctions. One can see numerically the pointwise ratio at the top (black, dashed) $\frac{p(u)}{u}$ which is practically constant (for all x). This means that u is indeed an eigenfunction of TV and admits $p(u) = \lambda u$. The same goes for v on the top right side. We denote by d the distance between the centers of the peak parts of u and v (shown on the second row on the left). The eigenfunction u is displaced from being at $d = 0$ to $d = 25$, where for each d the measures $\mathcal{O}(u, v)$ and $L(u, v)$ are computed. Both indicators are well correlated, with L yielding slightly sharper results. As can be expected, as the peaks of the functions u and v are farther apart, decomposition is easier and both indicator approach 1. Several instances of the composition $f = u + v$ are shown on the bottom row on the right.

In Figs. 6 and 7 a 2D experiment is shown. Here d is the distance between the centers of two discs of identical size (radius). In the continuous case, in an unbounded domain $\mathbb{R}^2$ a disc is an eigenfunction of TV. Here we have a bounded domain and cannot produce discretely real discs, so this is an approximation. As those discs are identical (radius and height), in principle they cannot be decomposed through spectral filtering since they have the same eigenvalue. However we can compare this case to a theoretical analysis done by Bellettini et al [7].

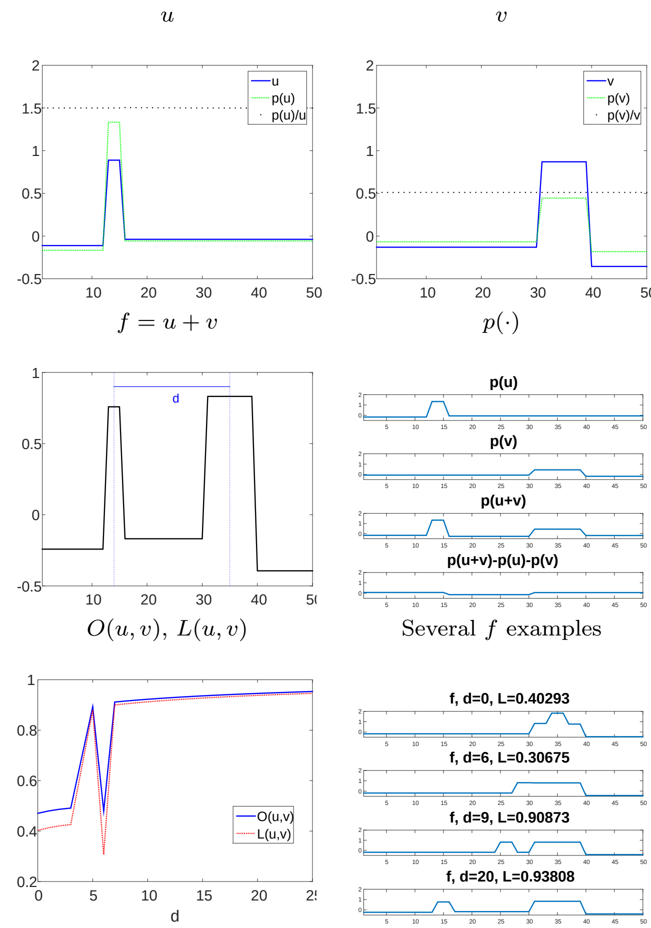

Fig. 5 Comparison of $\mathcal{O}(u, v)$ and $L(u, v)$ for the 1D case of u and v being 2 precise TV eigenfunctions in a finite domain. Top row: u (blue), $p(u)$ (green) and $\frac{p(u)}{u}$ (black, dashed) are shown on the left, v, $p(v)$ and $\frac{p(v)}{v}$ (right). Middle row (from left), $f = u + v$ and d is shown which is the distance between the centers of u and v. An example ($d = 21$) of $p(u)$, $p(v)$, $p(u + v)$ and the difference $p(u + v) - p(u) - p(v)$ (from top subplot, respectively). As the difference vanishes the LIS measure $L(u, v)$ approaches 1. Bottom row, $O(u, v)$ (blue) and $L(u, v)$ (red, dashed) as a function of the distance d. On the right, several cases of f for different values of d.

It was shown in [7] that for two identical discs of radius r the sum of the two discs is also an eigenfunction (meaning they admit (LIS)) if $d \geq \pi r$. Therefore the values of $\mathcal{O}(u, v)$ and $L(u, v)$ are plotted as a function of $\frac{d}{r}$, with critical points at $\frac{d}{r} = 2$, that is the discs are just separated but touch each other at a single point, and at $\frac{d}{r} = \pi$, the theoretical critical distacne. As can be seen, $\mathcal{O}$ and L are almost identical here, however the critical point may not be that significant and as soon as the discs do not touch each other, $\frac{d}{r} > 2$, the values approach 1 fast. One can notice in the numerical examples for several d values on the right of Fig. 7 that for

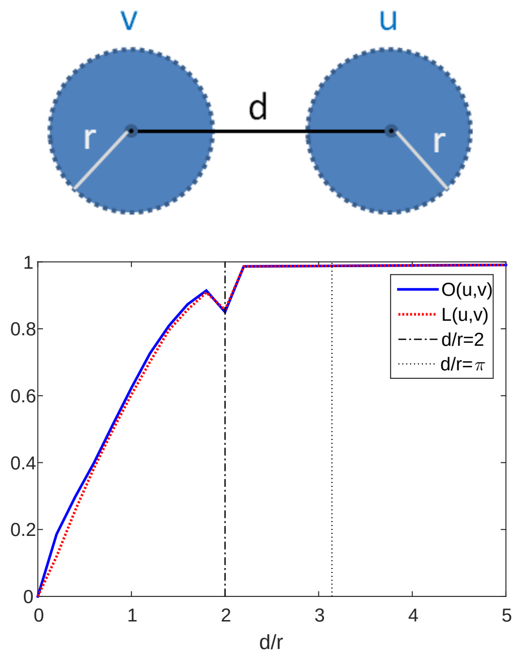

Fig. 6 Experiment of 2 identical discs. Top - illustration of the discs u and v of radius r and the distance d between the centers of the discs. Bottom $\mathcal{O}(u, v)$, Eq. (51), and $L(u, v)$, Eq. (53), as a function of d/r.

$\frac{d}{r} = 4 > \pi$ indeed we get that $p(u + v) - p(u) - p(v)$ almost vanishes numerically.

6 Conclusion

In this work several new concepts were presented, which can be helpful in future theoretical understanding and better employment of convex regularizers. The properties of semi-inner-products for convex functionals were stated, following the s.i.p. of Lumer for normed spaces. Essentially, linearity and homogeneity are kept in the first argument, the functional is induced by the s.i.p. and a Cauchy-Schwartz-type property holds. The s.i.p., however, does not behave linearly with respect to the second argument. For non-smooth functionals s.i.p.'s are similar to the subdifferential and may contain several elements (in this case it is unique when a subgradient element is chosen).

For the one-homogeneous case a general formulation of the s.i.p. was given. This yields natural definitions of orthogonality and angles between 2 functions, with respect to regularizing functionals like TV or TGV. The relation to the Bregman distance was shown, where in the one-homogeneous case the Bregman distance between two functions can be expressed in terms of the angle between those functions.

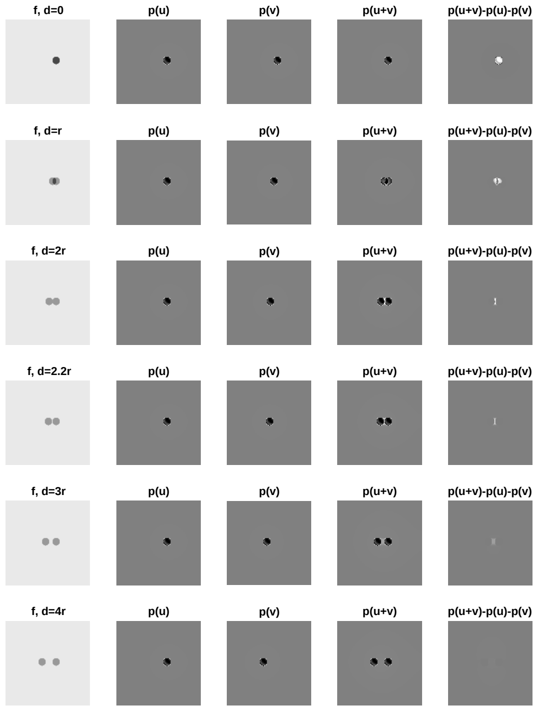

Fig. 7 A few examples of the 2 discs experiment. Each column (from left): $f = u + v$, $p(u)$, $p(v)$, $p(u + v)$, $p(u + v) - p(u) - p(v)$. The rows are results of different distances between the discs, $d/r = 0, 1, 2, 2.2, 3, 4$.

An extension of s.i.p.'s to general degrees was suggested, where the case of half-semi-inner-products (h.s.i.p.) was further developed. Finally, it was shown that when the h.s.i.p. is linear in the second argument one can decompose two eigenfunctions (with different eigenvalues) perfectly, using the spectral filters proposed in [25, 13]. As the conditions for perfect decomposition are quite strict, two soft indicators based on s.i.p.'s and h.s.i.p.'s were suggested. Their goal is to measure how close we are to fulfilling those conditions. Initial experiments indicate both measures are useful in assessing the separability of signals with a dominant scale (where the one based on the (LIS) property yields slightly sharper results).

Acknowledgements

The author would like to thank the support of the Israel Science Foundation (ISF), grant 2097/15.

References

1. F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential and Integral Equations*, 14(3):321–360, 2001.
2. F. Andreu, V. Caselles, JI Dıaz, and JM Mazón. Some qualitative properties for the total variation flow. *Journal of Functional Analysis*, 188(2):516–547, 2002.
3. G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2002.
4. J.F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition into a bounded variation component and an oscillating component. *JMIV*, 22(1), January 2005.
5. J.F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition – modeling, algorithms, and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, 2006.
6. A. Banerjee, S. Merugu, I. Dhillon, and J. Ghosh. Clustering with bregman divergences. *The Journal of Machine Learning Research*, 6:1705–1749, 2005.
7. G. Bellettini, V. Caselles, and M. Novaga. The total variation flow in R^N. *Journal of Differential Equations*, 184(2):475–525, 2002.
8. M. Benning and M. Burger. Ground states and singular vectors of convex variational regularization methods. *Methods and Applications of Analysis*, 20(4):295–334, 2013.
9. K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imaging Sciences*, 3(3):492–526, 2010.
10. L.M. Bregman. The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Comp. Math. and Math. Phys.*, 7:200–217, 1967.
11. R. Bruck. Nonexpansive projections on subsets of banach spaces. *Pacific Journal of Mathematics*, 47(2):341–355, 1973.
12. M. Burger. Bregman distances in inverse problems and partial differential equation. *arXiv preprint arXiv:1505.05191*, 2015.
13. M. Burger, L. Eckardt, G. Gilboa, and M. Moeller. Spectral representations of one-homogeneous functionals. In *Scale Space and Variational Methods in Computer Vision*, pages 16–27. Springer, 2015.
14. M. Burger, G. Gilboa, M. Moeller, L. Eckardt, and D. Cremers. Spectral decompositions using one-homogeneous functionals, 2016. Submitted. Online at http://arxiv.org/pdf/1601.02912v1.pdf.
15. M. Burger, G. Gilboa, S. Osher, and J. Xu. Nonlinear inverse scale space methods. *Comm. in Math. Sci.*, 4(1):179–212, 2006.
16. M. Burger and S. Osher. A guide to the tv zoo. In *Level Set and PDE Based Reconstruction Methods in Imaging*, pages 1–70. Springer, 2013.
17. Y. Censor and T. Elfving. A multiprojection algorithm using bregman projections in a product space. *Numerical Algorithms*, 8(2):221–239, 1994.
18. A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock. An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9:263–340, 2010.
19. B. Cox, A. Juditsky, and A. Nemirovski. Dual subgradient algorithms for large-scale nonsmooth learning problems. *Mathematical Programming*, 148(1-2):143–180, 2014.
20. R. Der and D. Lee. Large-margin classification in banach spaces. In *International Conference on Artificial Intelligence and Statistics*, pages 91–98, 2007.
21. S. S. Dragomir. *Semi-inner products and applications.* Nova Science Publishers New York, 2004.

22. I Ekeland and R Témam. Convex analysis and variational problems. *Classics in Applied Mathematics. Society for Industrial and Applied Mathematics*, 1999.
23. L.C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1991.
24. Y. Giga and R.V. Kohn. Scale-invariant extinction time estimates for some singular diffusion equations. *Hokkaido University Preprint Series in Mathematics*, (963), 2010.
25. G. Gilboa. A total variation spectral framework for scale and texture analysis. *SIAM J. Imaging Sciences*, 7(4):1937–1961, 2014.
26. G. Gilboa, M. Moeller, and M. Burger. Nonlinear spectral analysis via one-homogeneous functionals - overview and future prospects. Submitted. Preprint at http://arxiv.org/abs/1510.01077.
27. G. Gilboa, N. Sochen, and Y.Y. Zeevi. Variational denoising of partly-textured images by spatially varying constraints. *IEEE Trans. on Image Processing*, 15(8):2280–2289, 2006.
28. J.R. Giles. Classes of semi-inner-product spaces. *Transactions of the American Mathematical Society*, pages 436–446, 1967.
29. T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.
30. P. Jain, B. Kulis, J.V. Davis, and I.S. Dhillon. Metric and kernel learning using a linear transformation. *The Journal of Machine Learning Research*, 13(1):519–547, 2012.
31. M. Lange, M. Biehl, and T. Villmann. Non-euclidean principal component analysis by hebbian learning. *Neurocomputing*, 147:107–119, 2015.
32. G. Lumer. Semi-inner-product spaces. *Transactions of the American Mathematical Society*, pages 29–43, 1961.
33. S. Ma, D. Goldfarb, and L. Chen. Fixed point and bregman iterative methods for matrix rank minimization. *Mathematical Programming*, 128(1-2):321–353, 2011.
34. Y. Meyer. Oscillating patterns in image processing and in some nonlinear evolution equations, March 2001. The 15th Dean Jacquelines B. Lewis Memorial Lectures.
35. S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation based image restoration. *SIAM Journal on Multiscale Modeling and Simulation*, 4:460–489, 2005.
36. S. Osher, A. Sole, and L. Vese. Image decomposition and restoration using total variation minimization and the H^{-1} norm. *SIAM Multiscale Modeling and Simulation*, 1(3):349–370, 2003.
37. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
38. L. Vese and S. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19:553–572, 2003.
39. Y. Yang, S. Han, T. Wang, W. Tao, and X.-C. Tai. Multilayer graph cuts based unsupervised color–texture image segmentation using multivariate mixed student's t-distribution and regional credibility merging. *Pattern Recognition*, 46(4):1101–1124, 2013.
40. H. Zhang, Y. Xu, and J. Zhang. Reproducing kernel banach spaces for machine learning. *The Journal of Machine Learning Research*, 10:2741–2775, 2009.
41. X. Zhang, M. Burger, X. Bresson, and S. Osher. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM Journal on Imaging Sciences*, 3(3):253–276, 2010.