Higher order non-local regularisation

Jan Lellmann[†], Kostas Papafitsoros [†], Carola Schönlieb [†] & Daniel Spector [‡]

[†]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom

[‡]Department of Applied Mathematics, National Chiao Tung University, Hsinchu, Taiwan

EPSRC

Why higher order? Why non-local?

Higher order models: They avoid artifacts – like the staircasing effect in total variation (TV) regularisation – and ideally they also preserve geometric features like edges e.g. total generalised variation (TGV).

Non-local models:

- In image processing, non-local models are capable of exploiting information from far away pixels and repeating patters as well as preserving textures.
- In functional analysis, non-local functionals have been introduced and shown to approximate classical local differential operators as the non-locality vanishes and in certain cases they also provide non-local characterisations of Sobolev and BV spaces.

In this work we study two versions of non-local Hessian functionals, the *explicit* and the *implicit* formulation. For the former one, we prove localisation results and we use it to provide some novel derivative-free characterisations of higher order Sobolev and BV spaces while we apply the latter one in image denoising with very promising results.

Non-local Hessian–The explicit formulation

We define the *explicit* non-local Hessian $H_n u$ of a function $u : \mathbb{R}^N \to \mathbb{R}$ at a point $x \in \mathbb{R}^N$ as follows:

$$H_n u(x) := \frac{N(N+2)}{2} \int_{\mathbb{R}^N} \frac{u(x+h) - 2u(x) + u(x-h)}{|h|^2} \frac{\left(h \otimes h - \frac{|h|^2}{N+2} I_N\right)}{|h|^2} \rho_n(h)dh, \quad (1)$$

where I_N is the $N \times N$ identity matrix and $\rho_n : \mathbb{R}^N \to \mathbb{R}$ are radially symmetric weight functions that satisfy

$$\rho_n \geq 0, \quad \int_{\mathbb{R}^N} \rho_n(x)dx = 1, \quad \lim_{n\to\infty} \int_{|x|>\gamma} \rho_n(x)dx = 0, \quad \forall \gamma > 0. \quad (2)$$

The integral (1) (that can be defined more rigorously as a distribution) exists indeed for functions with at least $\mathrm{BV}^2(\mathbb{R}^N)$ regularity where $\mathrm{BV}^2(\mathbb{R}^N) := \{u \in W^{1,1}(\mathbb{R}^N) : \nabla u \in \mathrm{BV}(\mathbb{R}^N, \mathbb{R}^{N\times N})\}$. Moreover the following localisations hold:

Theorem: Localisation results for $H_n u$

Let ρ_n be a sequence as above. Then the following convergences hold as $n \to \infty$:

(i) If $u \in C_2^2(\mathbb{R}^N)$ then $H_n u \to \nabla^2 u$ uniformly.

(ii) If $u \in W^{2,p}(\mathbb{R}^N)$, $1 \leq p < \infty$ then $H_n u \to \nabla^2 u$ in $L^p(\mathbb{R}^N, \mathbb{R}^{N\times N})$.

(iii) If $u \in \mathrm{BV}^2(\mathbb{R}^N)$ then $H_n u \mathcal{L}^N \to D^2 u$ weakly* in measures (strictly if $N = 1$), where $D^2 u := D(\nabla u)$ is here a $\mathbb{R}^{N\times N}$-valued finite Radon measure.

Using the explicit formulation of non-local Hessian we obtain some novel derivative-free characterisations of the spaces $W^{2,p}(\mathbb{R}^N)$ and $\mathrm{BV}(\mathbb{R}^N)$ as follows:

Theorem: Non-local characterisations of $W^{2,p}(\mathbb{R}^N)$ and $\mathrm{BV}^2(\mathbb{R}^N)$

(i) Let $u \in L^p(\mathbb{R}^N)$, $1 < p < \infty$. Then

$$u \in W^{2,p}(\mathbb{R}^N) \iff \liminf_{n\to\infty} \|H_n u\|_{L^p} < \infty.$$

(ii) Let $u \in L^1(\mathbb{R}^N)$. Then

$$u \in \mathrm{BV}^2(\mathbb{R}^N) \iff \liminf_{n\to\infty} \|H_n u\|_{L^1} < \infty.$$

Non-local Hessian–The implicit formulation

We define an *implicit* formulation of a non-local Hessian which is more versatile than (1) and allows for non-symmetric weights, thus making it more suitable for imaging applications. We define the explicit non-local gradient and non-local Hessian, $G'_u(x)$ and $H'_u(x)$ respectively as

$$(G'_u(x), H'_u(x)) := \underset{G_u\in\mathbb{R}^N, H_u\in \mathrm{Sym}(\mathbb{R}^{N\times N})}{\operatorname{argmin}} \frac{1}{2} \int_{\Omega-\{x\}} \left(u(x+h) - u(x) - G_u^\top h - \frac{1}{2} h^\top H_u h\right)^2 \rho_x(h)dh. \quad (3)$$

Since the optimality conditions for (3) are linear we can use H'_u as a regulariser as follows:

$$\min_u \frac{1}{s}\|u - f\|_s^s + \alpha\|H'u\|_1, \quad s = 1, 2, \quad \text{subject to (3)}.$$

Convex objective + linear constraints $\Longrightarrow$ Solvable with standard convex solvers.

We choose the weight function ρ_x in a way that it is small for points that are separated from x by a strong edge and large otherwise. This is done by solving the *Eikonal* equation

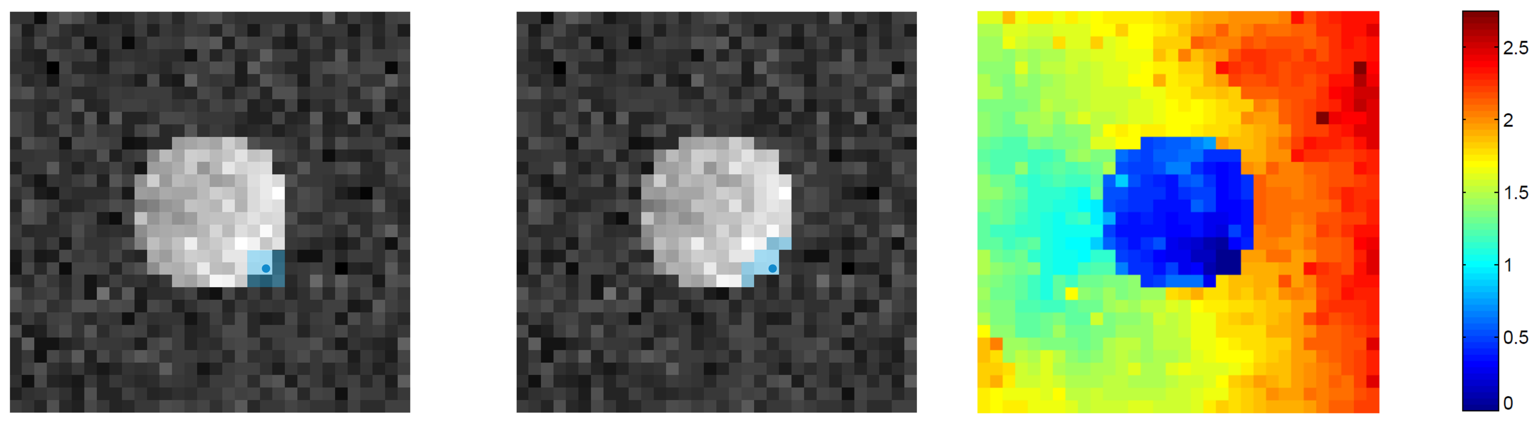
$$\|\nabla c_x\|_2 = \phi(\nabla f), \quad c_x(x) = 0, \quad \phi(\nabla f) = \|\nabla f\|_2^2 + \epsilon,$$

and setting

$$\rho_x(y_i) = \begin{cases} \frac{1}{(c_x(y_i))^2} & \text{if } i \leq K, \\ 0 & \text{if } i > K, \end{cases}$$

for the K closest neighbours of x (in the "c_x-distance" sense).

$\Longrightarrow$ we obtain a regulariser that preserves **edges** and **affine** structures as well.



Left: Standard discretisation of 2[nd] order derivatives uses all points in a 3×3 neighbourhood. *Middle*: The proposed discretisation involves only the points that are likely to belong to the same affine region. *Right*: The values of c_x for all the points. Points that are separated from x by a strong edge have a large value of c_x and are therefore not included in the set of points that is used to define the non-local Hessian at x.

The following theorem reveals the relationship between the explicit, and the implicit formulation of non-local Hessian:

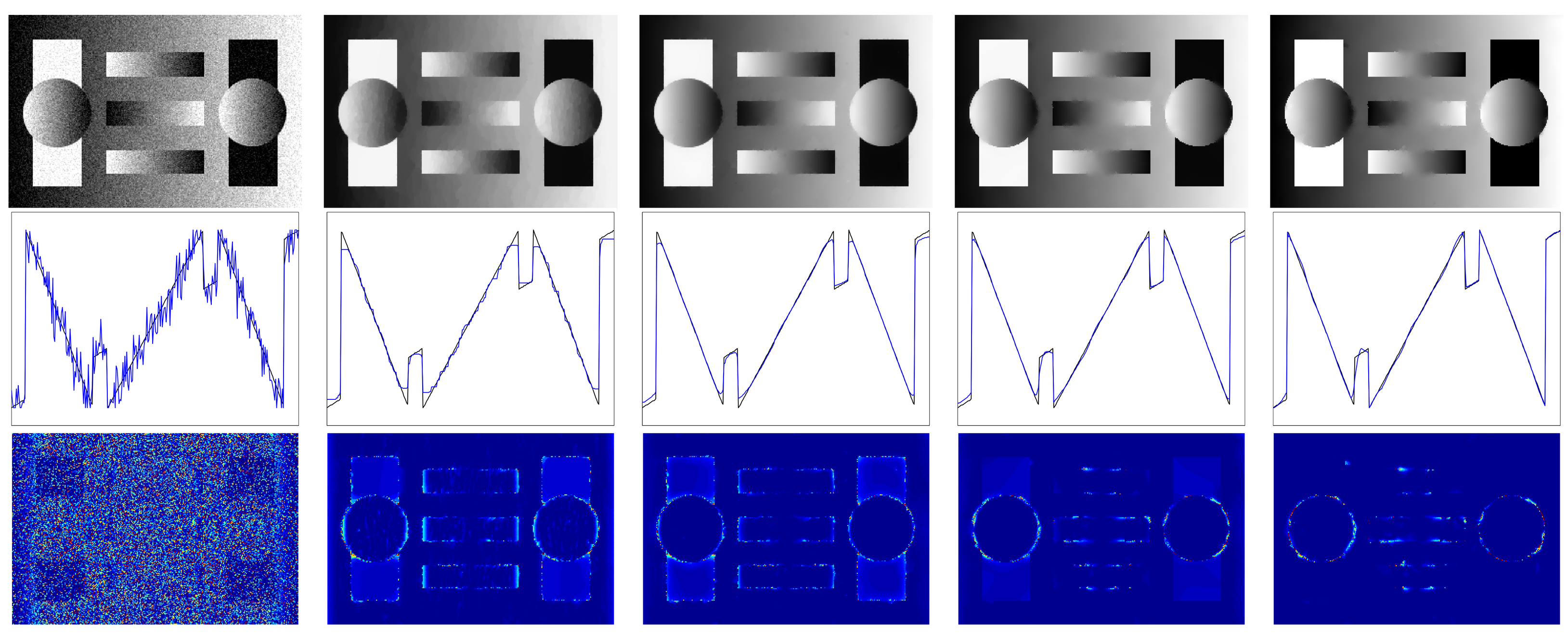
Theorem: Connection between the explicit and implicit non-local Hessian

Assume that $\rho_x = \rho$ is radially symmetric. Then the explicit non-local Hessian Hu can be written as

$$Hu = \int_0^\infty H'_u(h) \int_{hS^{N-1}} \rho(z) d\mathcal{H}^{N-1}(z) dh.$$

Some denoising examples

We are able to obtain true piecewise affine reconstructions with minimal loss of contrast:



Input, PSNR=22.82	TV, PSNR=31.83	L^2–TGV, PSNR=33.28	L^2–H'_u, PSNR=35.39	L^1–H'_u, PSNR=36.06

First row: Denoising results. *Second row*: Middle row slices. *Third row*: Distribution of the L^2 error to the ground truth. The parameters are chosen for optimal PSNR.

References

(1) J. Lellmann, K. Papafitsoros, C.B. Schönlieb and D. Spector, *Analysis and application of a non-local Hessian*, in preparation, (2014).

(2) K. Bredies, K. Kunisch and T. Pock, *Total generalised variation*, SIAM Journal on Imaging Sciences, 3 (2009), 1–42.

(3) T. Mengesha and D. Spector, *Localization of nonlocal gradients in various topologies*, Calculus of Variations and Partial Differential Equations (2014).

This research has been funded by the Cambridge Centre for Analysis, EPSRC and the King Abdullah University of Science and Technology (KAUST).