On certain variational non-additive noise models and beyond

J. Darbon[1], T. Meng[1] and E. Resmerita[2]

[1]Brown University, [2] Alpen-Adria University of Klagenfurt

Optimization and Operator Theory Workshop, Technion

Dedicated to Professor Lev Bregman

November 15, 2021

Motivation

- Image denoising (Gaussian noise)

noisy image

restoration

Motivation

- Image denoising (Poisson noise)

noisy image

restoration

Motivation

- Image denoising as optimization problem (MAP in a Bayesian framework)
- Variational model for approximating u:

$$\min_{u \in \mathbb{R}^n} \underbrace{D(u,x)}_{\text{Data Fidelity}} + t \underbrace{J(u)}_{\textit{a priori}/\text{Regularisation}}$$

    - Observed image: $x \in \mathbb{R}^n$ (a matrix of dimension $n_1 n_2 := n$)
    - Regularization parameter $t \in (0, +\infty)$
- Examples:
    - Gaussian noise: $D(u,x) = \frac{1}{2}\|u - x\|_2^2$
    - Promote sparsity:
        - on the signal: $J = \|\cdot\|_1$ (e.g. Compressive Sensing)
        - on the variation of the signal: $J(u) = \sum_{i,j} |u_j - u_i|$ (Total Variation denoising: Rudin, Osher, Fatemi 92)
- Many available algorithms to compute a minimizer.

HJ equations and additive noise models $x = u + \eta$

- Hamilton-Jacobi equation (HJ)

$$\begin{cases} \frac{\partial S}{\partial t}(x,t) + H(\nabla_x S(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,+\infty), \\ S(x,0) = J(x) & \forall x \in \mathbb{R}^n, \end{cases}$$

- Solution is given by the Lax-Oleinik representation formula

$$S(x,t) = \inf_{u \in \mathbb{R}^n} \left\{ J(u) \; + \; tH^* \left(\frac{x-u}{t} \right) \right\}$$

- Initial data J correspond to the regularization J.
- Hamiltonian H defines the data fidelity for additive noise model via H^*
- The spatial variable x corresponds to the observed image, the time t is the regularization parameter.
- Details in [Darbon 15], extension to multitime [Darbon, Meng 20], Bayesian posterior mean [Darbon, Provencher 20]

Variational non-additive noise models

$$\min_{v \in \mathbb{R}^n} \{f(v) + g(t, x, v)\},$$

- f is the penalizing term
- g is the data fidelity
- t is a positive parameter
- $\frac{x}{t}$ denotes the observed image.

In this work, we focus on

$$\min_{v \in \mathrm{int\ dom}\ H^*} \underbrace{J^*(\nabla H^*(v))}_{f(v)} + \underbrace{tD_{H^*}\left(\frac{x}{t}, v\right)}_{g(t,x,v)},$$

- H is a Legendre function
- J is a convex function.

Goals

- Establish connections between (possibly non-convex) non-additive noise models

$$(NA) \qquad \min_{v \in \mathrm{int\ dom}\ H^*} \underbrace{J^*(\nabla H^*(v))}_{f(v)} + \underbrace{tD_{H^*}\left(\frac{x}{t}, v\right)}_{g(t,x,v)},$$

  and additive noise models:

$$(A) \qquad \min_{v \in \mathbb{R}^n} \{J(x - tv) + tH^*(v)\}.$$

- Establish connections between the above noise models and some HJ PDEs.

Main contributions

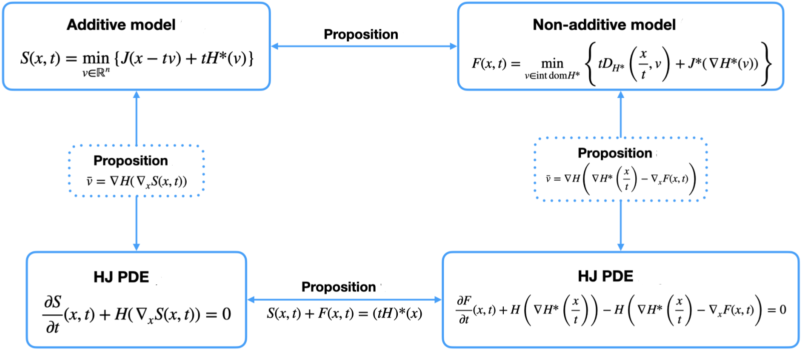


Figure: An illustration of the relations among the non-additive model, the additive one and the corresponding PDEs

Main assumptions

- J and H are proper, convex, lower semicontinuous functions on $\mathbb{R}^n$.
- J is 1-coercive, and H is a Legendre function.

  Short recap: H is Legendre if it satisfies

  1. $\mathrm{int}\ \mathrm{dom}\ H$ is non-empty.
  2. H is differentiable on $\mathrm{int}\ \mathrm{dom}\ H$.
  3. $$\partial H(x) = \begin{cases} \emptyset, & \text{if } x \in (\mathrm{dom}\ H) \setminus (\mathrm{int}\ \mathrm{dom}\ H), \\ \{\nabla H(x)\}, & \text{if } x \in \mathrm{int}\ \mathrm{dom}\ H. \end{cases}$$
  4. H is strictly convex on $\mathrm{int}\ \mathrm{dom}\ H$.

Main results

Proposition

Let $x \in \mathbb{R}^n$ *and* $t > 0$ *satisfy* $x \in (\mathrm{dom}\ J + t\,\mathrm{int}\ \mathrm{dom}\ H^*) \cap (t\,\mathrm{dom}\ H^*)$. *Then, the minimizers in* (NA) *and* (A) *exist and are unique. Also,* $\bar{v}$ *is the minimizer in* (NA) *if and only if it is the minimizer in* (A)*:*

$$\arg\min_{v \in \mathrm{int}\ \mathrm{dom}\ H^*} \left\{ tD_{H^*}\left(\frac{x}{t}, v\right) + J^*(\nabla H^*(v)) \right\} = \arg\min_{v \in \mathbb{R}^n} \{J(x - tv) + tH^*(v)\}$$

Moreover, the minimal values in the two problems satisfy

$$\begin{aligned} \min_{v} \left\{ tD_{H^*}\left(\frac{x}{t}, v\right) + J^*(\nabla H^*(v)) \right\} \quad &+ \quad \min_{v}\{J(x - tv) + tH^*(v)\} \\ &= \quad (tH)^*(x). \end{aligned}$$

"Generalized Moreau decomposition"
[Moreau 65] for $t = 1$ and $H = \frac{1}{2}\|\cdot\|_2^2$, [Combettes, Reyes 11] for more general cases.

Useful tools for the analysis

Consider

$$(P1) \quad \min_{u\in\mathbb{R}^n} \left\{ tH^* \left(\frac{x-u}{t} \right) + J(u) \right\},$$

$$(P2) \quad \min_{p\in\mathbb{R}^n} \{\langle p, x\rangle - tH(p) - J^*(p)\},$$

which are equivalent to (A) and (NA), respectively,

$$(A) \quad \min_{v\in\mathbb{R}^n} \{J(x-tv) + tH^*(v)\}.$$

$$(NA) \quad \min_{v\in \mathrm{int\ dom}\ H^*} J^*(\nabla H^*(v)) + tD_{H^*}\left(\frac{x}{t}, v\right).$$

Main results

Proposition

Let $x \in \mathrm{dom}\ J + t\,\mathrm{int}\ \mathrm{dom}\ H^*$. Then, there exist unique vectors $\bar{u}$ and $\bar{p}$ in $\mathbb{R}^n$ solving (P1) and (P2), respectively. Moreover, the following two statements are equivalent:

(a) *The vectors $\bar{u}$ and $\bar{p}$ solve (P1) and (P2), respectively.*

(b) *There hold*

$$x = \bar{u} + t\nabla H(\bar{p}), \quad \text{and} \quad J(\bar{u}) + J^*(\bar{p}) - \langle \bar{p}, \bar{u} \rangle = 0.$$

Additionally, it holds

$$\min_{u \in \mathbb{R}^n} \left\{ tH^* \left(\frac{x-u}{t} \right) + J(u) \right\} = \max_{p \in \mathbb{R}^n} \{ \langle p, x \rangle - tH(p) - J^*(p) \}.$$

Relations to some HJ PDEs

- The minimal value $S(x,t)$ in (A) solves a HJ PDE,

$$S(x,t) := \begin{cases} \inf_{v\in\mathbb{R}^n} \{J(x-tv) + tH^*(v)\}, & x\in\mathbb{R}^n, t>0, \\ (J^* + I_{\text{dom } H})^*(x), & x\in\mathbb{R}^n, t=0, \\ +\infty, & x\in\mathbb{R}^n, t<0. \end{cases}$$

- The minimal value $F(x,t)$ in (NA) satisfies another HJ PDE,

$$F(x,t) := \begin{cases} \inf_{v} \left\{ tD_{H^*}\left(\frac{x}{t}, v\right) + J^*(\nabla H^*(v)) \right\}, & t>0,\ (x,t)\in D_F \\ I^*_{\text{dom } H}(x) - (J^* + I_{\text{dom } H})^*(x), & t=0,\ (x,t)\in D_F \end{cases}$$

More details

Proposition

The following statements hold:

(a) *S is a convex and lower semi-continuous function with respect to the joint variable.*

(b) *S is differentiable at any $(x, t) \in \mathrm{int}\ \mathrm{dom}\ S$, with*

$$\nabla S(x, t) = (\bar{p}, -H(\bar{p})),$$

where $\bar{p}$ is the unique maximizer of

$$(P2) \quad \min_{p \in \mathbb{R}^n} \{\langle p, x\rangle - tH(p) - J^*(p)\}.$$

(c) *For all $x \in \mathbb{R}^n$, we have $S(x, 0) \leq J(x)$. If $x \in \mathrm{dom}\ J$ satisfies $\partial J(x) \cap \mathrm{cl}\ \mathrm{dom}\ H \neq \emptyset$, then $S(x, 0) = J(x)$ and $\partial_x S(x, 0) = \partial J(x) \cap \mathrm{cl}\ \mathrm{dom}\ H$, where $\partial_x S(x, 0)$ is the subdifferential of the function $y \mapsto S(y, 0)$.*

Relations to some HJ PDEs

Proposition

The function S solves the following HJ PDE

$$\begin{cases} \frac{\partial S}{\partial t}(x,t) + H(\nabla_x S(x,t)) = 0, & \text{if } (x,t) \in \mathrm{int\ dom}\ S, \\ S(x,0) = J_S(x), & \text{if } x \in \mathbb{R}^n, \end{cases}$$

where J_S is a proper, convex, lsc function with $\mathrm{dom}\ J_S^* = \mathrm{cl\ dom}\ H$ *and* $J_S^* = J^*$ *in the domain of* J_S^**. Moreover, the minimizer* $\bar{v}$ *in* (A) *satisfies*

$$\bar{v} = \nabla H(\nabla_x S(x,t)),$$

for all $t > 0$ *and* $x \in \mathrm{dom}\ J + t\,\mathrm{int\ dom}\ H^*$*.*

Relations to some HJ PDEs

Proposition

Assume that $0 \in \mathrm{dom}\ J$*. Then* F *is continuously differentiable on the interior of its domain and satisfies the following differential equation for all* $(x,t) \in \mathrm{int\ dom}\ F$

$$\frac{\partial F}{\partial t}(x,t) + H\left(\nabla H^*\left(\frac{x}{t}\right)\right) - H\left(\nabla H^*\left(\frac{x}{t}\right) - \nabla_x F(x,t)\right) = 0.$$

Also, F *satisfies*

$$\lim_{t \to 0^+} F(x+td,t) = F(x,0), \quad \forall d \in \mathrm{int\ dom}\ H^*,\ x \in \mathrm{dom}\left(I^*_{\mathrm{dom}\ H}\right).$$

Moreover, the minimizer $\bar{v}$ *in* (NA) *satisfies*

$$\bar{v} = \nabla H\left(\nabla H^*\left(\frac{x}{t}\right) - \nabla_x F(x,t)\right),$$

for all $t > 0$ *and* $x \in t\,\mathrm{int\ dom}\ H^*$.

Poisson noise model

Define

$$H(p) = \sum_{i=1}^{n} e^{p_i}, \qquad p = (p_1, \dots, p_n) \in \mathbb{R}^n.$$

The function H is a Legendre function with H^* given by

$$H^*(y) = \sup_{p \in \mathbb{R}^n} \sum_{i=1}^{n} (p_i y_i - e^{p_i}) = \begin{cases} \sum_{i=1}^{n} (y_i \log y_i - y_i), & \text{if } y \in [0, +\infty)^n, \\ +\infty, & \text{otherwise.} \end{cases}$$

Hence, $D_{H^*}(\frac{x}{t}, v)$ is the Kullback-Leibler distance

$$D_{H^*}\left(\frac{x}{t}, v\right) = \sum_{i=1}^{n} \left(\frac{x_i}{t} \log \frac{x_i}{t} - \frac{x_i}{t} + v_i - \frac{x_i}{t} \log v_i \right).$$

The variational model with Poisson noise:

$$\min_{v \in (0, +\infty)^n} \left\{ t D_{H^*}\left(\frac{x}{t}, v\right) + J^*(\log v_1, \dots, \log v_n) \right\}.$$

Poisson noise model

- The related additive noise model is

$$\min_{v\in\mathbb{R}^n}\left\{J(x-tv)+t\sum_{i=1}^{n}(v_i\log v_i-v_i)\right\}.$$

- The related HJPDEs are

$$\frac{\partial S}{\partial t}(x,t)+\sum_{i=1}^{n}\exp\left(\frac{\partial S}{\partial x_i}(x,t)\right)=0,$$

$$\frac{\partial F}{\partial t}(x,t)-\frac{1}{t}\sum_{i=1}^{n}x_i\left(\exp\left(-\frac{\partial F}{\partial x_i}(x,t)\right)-1\right)=0.$$

Poisson noise model: Remarks

- A widely used regularization term is the total variation TV [Le, Chartrand, Asaki 07].
  However, there is no convex lsc function J satisfying $J^*(\nabla H^*(v)) = J^*(\log v_1, \ldots, \log v_n) = \mathrm{TV}(v)$.
- An example of appropriate penalizing function $f = J^*(\nabla H^*)$ is

$$f(v) = \mathrm{TV}(\log v_1, \ldots, \log v_n),$$

where J is the indicator ball of Meyer's norm.
The corresponding variational denoising model:

$$\min_{v_i > 0} \left\{ \sum_{i=1}^{n} \left(tv_i - x_i \log v_i + x_i \log\left(\frac{x_i}{t}\right) - x_i \right) + \mathrm{TV}(\log v_1, \ldots, \log v_n$$

[Oh, Harmani, Willet 13]

Poisson noise model: Remarks

The parameter t is related to the exposure time of the sensor:

- Let v be the gray level array of the original image, which does not change over time.
- The observed image is a sample from a Poisson distribution whose rate equals tv, where t is the exposure time of the sensor [Tendero, Osher 16].
- The probability mass function of the Poisson distribution at $x \in \mathbb{Z}^n$ equals

$$P(x|v) = \prod_{i=1}^{n} \frac{(tv_i)^{x_i} e^{-tv_i}}{x_i!}.$$

Then, the corresponding MAP estimator for the denoising problem with Poisson noise reads

$$\bar{v} = \operatorname*{arg\,min}_{v \in (0,+\infty)^n} \left\{ \sum_{i=1}^{n} (tv_i - x_i \log v_i) + f(v) \right\}.$$

Multiplicative noise model

$$H(p) := \begin{cases} \sum_{i=1}^{n}(-1 - \log(-p_i)), & \text{if } p = (p_1, \ldots, p_n) \in (-\infty, 0)^n, \\ +\infty, & \text{otherwise,} \end{cases}$$

$$H^*(y) = \begin{cases} -\sum_{i=1}^{n} \log y_i, & \text{if } y = (y_1, \ldots, y_n) \in (0, +\infty)^n, \\ +\infty, & \text{otherwise,} \end{cases}$$

which is the Burg entropy, thus yielding the Itakura-Saito distance

$$D_{H^*}\left(\frac{x}{t}, v\right) = \sum_{i=1}^{n} \left(-\log \frac{x_i}{t} + \log v_i + \frac{x_i/t}{v_i} - 1 \right).$$

The multiplicative noise variational model is

$$\min_{v \in (0,+\infty)^n} \left\{ t \sum_{i=1}^{n} \left(-1 + \log v_i + \frac{x_i/t}{v_i} - \log \frac{x_i}{t} \right) + J^*\left(-\frac{1}{v_1}, \ldots, -\frac{1}{v_n} \right) \right\}$$

Multiplicative noise model

- The related additive noise model is

$$\min_{v\in(0,+\infty)^n}\left\{J(x-tv)-t\sum_{i=1}^{n}\log v_i\right\}.$$

- The related HJPDEs become

$$\frac{\partial S}{\partial t}(x,t)-\sum_{i=1}^{n}\left(1+\log\left(-\frac{\partial S}{\partial x_i}(x,t)\right)\right)=0,$$

$$\frac{\partial F}{\partial t}(x,t)+\sum_{i=1}^{n}\log\left(1+\frac{x_i}{t}\frac{\partial F}{\partial x_i}(x,t)\right)=0.$$

Multiplicative noise model: Remarks

- The non-convex regularization term $f(v) = \mathrm{TV}(\log v_1, \dots, \log v_n)$ has been often employed in the literature [Shi, Osher 08].
- However, there is no convex function J such that

$$J^* \left(-\frac{1}{v_1}, \dots, -\frac{1}{v_n} \right) = \mathrm{TV}(\log v_1, \dots, \log v_n), \quad v \in (0, +\infty)^n.$$

- We will use $f(v) = TV(v)$ [Aubert, Aujol 08].

Multiplicative noise model: Remarks

- The observation J_i on the i-th pixel in the model is the average of L observations $I_1, \ldots, I_L$, which are i.i.d. sampled from the exponential distribution with rate $\frac{1}{v_i}$.
- The distribution of J_i is the Gamma distribution with parameters L and $\frac{L}{v_i}$, with density function

$$f_{J_i}(J_i = z_i | v_i) = \left(\frac{L}{v_i}\right)^L \frac{1}{\Gamma(L)} z_i^{L-1} e^{-Lz_i/v_i}, \quad \forall z_i \in [0, +\infty).$$

- Since the pixels J_i are independent from each other, the density function of the whole image $J = (J_1, \ldots, J_n)$ is

$$f_J(J = z | v) = \prod_{i=1}^{n} \left(\frac{L}{v_i}\right)^L \frac{1}{\Gamma(L)} z_i^{L-1} e^{-Lz_i/v_i}, \quad z \in [0, +\infty)^n.$$

for all $z \in [0, +\infty)^n$.

Multiplicative noise model: Remarks

- The corresponding MAP estimator for the denoising problem with multiplicative noise is

$$\arg\min_{v\in(0,+\infty)^n} \left\{ \sum_{i=1}^{n} \left(L \log v_i + \frac{Lz_i}{v_i} \right) + f(v) \right\},$$

which is equivalent to the variational model

$$\arg\min_{v\in(0,+\infty)^n} \left\{ \sum_{i=1}^{n} \left(-t + t \log v_i + t\frac{x_i/t}{v_i} - t \log \frac{x_i}{t} \right) + f(v) \right\}.$$

with $t = L$ and $x = Lz$.

- The time variable t is the number of the observed images, and the spatial variable x is the summation of the t observed images.

Numerical experiments for a Poisson model

For the non-convex variational model

$$\min_{v\in(0,+\infty)^n}\left\{\sum_{i=1}^{n}(tv_i - x_i\log v_i) + \alpha\mathrm{TV}(\log v_1,\ldots,\log v_n)\right\},$$

which is equivalent to the convex one

$$\min_{v\in(0,+\infty)^n}\left\{\mathrm{TV}^*\left(\frac{x-tv}{\alpha}\right) + t\sum_{i=1}^{n}(v_i\log v_i - v_i)\right\},$$

ADMM is employed, that is

$$v^{(k+1)} = \arg\min_{v\in(0,+\infty)^n}\left\{t\sum_{i=1}^{n}(v_i\log v_i - v_i) + \frac{\lambda}{2}\left\|w^{(k)} + tv - x + y^{(k)}\right\|^2\right\},$$

$$w^{(k+1)} = \arg\min_{w\in\mathbb{R}^n}\left\{J(w) + \frac{\lambda}{2}\|w + tv^{(k+1)} - x + y^{(k)}\|^2\right\}$$

$$y^{(k+1)} = y^{(k)} + w^{(k+1)} + tv^{(k+1)} - x.$$

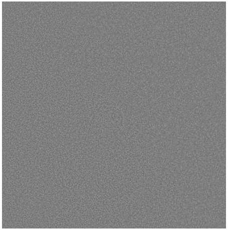
Numerical experiments for a Poisson model

(a)

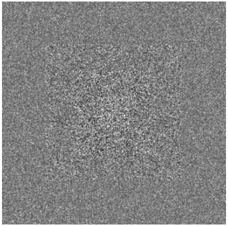
Numerical experiments for a Poisson model

(a) The restored image using our model with $TV(\log)$

(b) The restored image using TV



(c) The residual image (+0.5) using our model with $TV(\log)$



(d) The residual image (+0.5) using TV

Numerical experiments for a multiplicative model

For the non-convex problem

$$\min_{v\in(0,+\infty)^n}\left\{t\sum_{i=1}^{n}\left(-1+\log v_i+\frac{x_i/t}{v_i}-\log\frac{x_i}{t}\right)+\alpha f(v)\right\},$$

with $f(v)=\mathrm{TV}\left(-\frac{1}{v_1},\ldots,-\frac{1}{v_n}\right)$, consider the convex one with $J=(\alpha TV)^*$

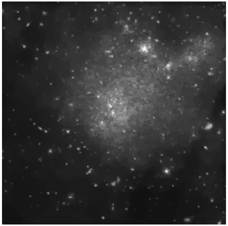
$$\min_v\left\{J(x-tv)-t\sum_{i=1}^{n}\log v_i\right\}=\min_v\left\{\mathrm{TV}^*\left(\frac{x-tv}{\alpha}\right)-t\sum_{i=1}^{n}\log v_i\right\}$$

and apply also an ADMM.

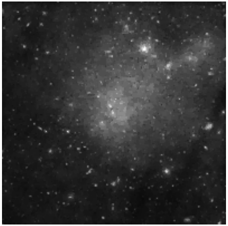
Numerical experiments for a multiplicative model

(a)

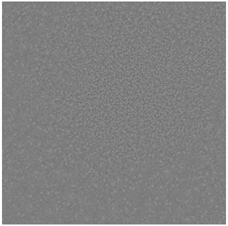
Numerical experiments for a multiplicative model



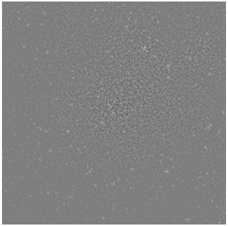
(a) The restored image using our model with TV



(b) The restored image using $TV(\log)$



(c) The residual image $(+0.5)$ using our model with TV



(d) The residual image $(+0.5)$ using $TV(\log)$

Summary

- We have proposed a variational model for denoising images with non-additive noise

$$\min_{v \in \mathrm{int\ dom}\ H^*} \left\{ J^*(\nabla H^*(v)) + tD_{H^*}\left(\frac{x}{t}, v\right) \right\}.$$

- We have established its connections with additive noise models and with HJ PDEs. This could be used for a knowledge transfer between the corresponding fields.
- We have discussed the Poisson noise and the multiplicative noise cases.
- We have shown numerical experiments based on ADMM.

On Hamilton-Jacobi PDEs and image denoising models with certain non-additive noise, J. Darbon, T. Meng, E. Resmerita, JMIV, in revision.

References

- Aubert, G., Aujol, J.F.: A variational approach to removing multiplicative noise. SIAM Journal on Applied Mathematics 68(4), 925–946 (2008).
- Combettes, P.L., Reyes, N.N.: Moreau's decomposition in Banach spaces. Math. Program. 139(1-2, Ser. B), 103–114 (2013).
- Darbon, J.: On convex finite-dimensional variational methods in imaging sciences and hamilton-jacobi equations. SIAM J. Imaging Sci. 8(4), 2268–2293 (2015).
- Darbon, J., Meng, T.: On decomposition models in imaging sciences and multi-time Hamilton–Jacobi partial differential equations. SIAM Journal on Imaging Sciences 13(2), 971–1014 (2020).
- Le, T., Chartrand, R., Asaki, T.J.: A variational approach to reconstructing images corrupted by poisson noise. J. Math. Imaging Vis. 27(3), 257–263 (2007).
- Moreau, J.J.: Proximite et dualite dans un espace hilbertien. Bulletin de la Societe Mathematique de France 93, 273–299 (1965).
- Oh, A.K., Harmany, Z.T., Willett, R.M.: Logarithmic total variation regularization for cross-validation in photon-limited imaging. In: 2013 IEEE International Conference on Image Processing, pp. 484–488 (2013).
- Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D 60(1-4), 259–268 (1992).
- Shi, J., Osher, S.: A nonlinear inverse scale space method for a convex multiplicative noise model. SIAM Journal on Imaging Sciences 1(3), 294–321 (2008).
- Tendero, Y., Osher, S.: On a mathematical theory of coded exposure. Res. Math. Sci. 3, Paper No. 4, 39 (2016).