Proximity Operators for a Class of Hybrid Sparsity+Entropy Priors. Application to DOSY NMR Signal Reconstruction
Afef Cherni, Emilie Chouzenoux, Marc-André Delsuc

To cite this version:

HAL Id: hal-01421741
https://hal.archives-ouvertes.fr/hal-01421741
Submitted on 22 Dec 2016

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

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
PROXIMITY OPERATORS FOR A CLASS OF HYBRID SPARSITY + ENTROPY PRIORS. APPLICATION TO DOSY NMR SIGNAL RECONSTRUCTION.

Cherni Arefé1,2, Chouzenoux Émilie1, Delsuc Marc-André2

1 Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM, Marne-la-Vallée, France.
2 Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U596, CNRS UMR 7104 Université de Strasbourg, 67404 Illkirch-Graffenstaden, France

ABSTRACT

Inverse problems arising from Laplace transform inversion are ill-posed, and require suitable regularization strategies. Although the maximum entropy regularization approach usually appears as an adequate strategy due to its ability to recover regular positive valued signals, it was observed to lead to poor reconstruction results when the sought signal contains narrow peaks. In that case, a sparsity promoting penalty such as the ℓ_1 norm, combined with a positivity constraint, is more suitable. In order to derive a flexible resolution method, hybrid approaches combining both entropy and sparsity regularization strategies should be envisaged. However, the choice of an efficient optimization algorithm remains a challenging task. Among available optimization techniques, proximal methods have shown their efficiency in solving large scale possibly nonsmooth problems. This paper provides an extensive list of new proximity operators for the sum of entropy possibly nonsmooth problems. This paper provides an extensive list of new proximity operators for the sum of entropy and sparsity penalties. The applicability of these results is illustrated by means of experiments, in the context of DOSY NMR signal reconstruction.

Index Terms—Inverse problems, Laplace inversion, entropy regularization, sparsity prior, proximity operator, spectroscopy, nuclear magnetic resonance.

1. INTRODUCTION

The fast resolution of large size ill-posed linear inverse problems presents a big challenge in the context of biophysical data processing [1]. A difficult inverse problem, arising for instance in the context of diffusion ordered nuclear magnetic resonance spectroscopy (DOSY NMR) [2], consists in the estimation of an original unknown positive-valued signal \( \mathbf{x} = (x_n)_{1 \leq n \leq N} \in [0, +\infty]^N \) from degraded measurements \( \mathbf{y} = (y_m)_{1 \leq m \leq M} \in \mathbb{R}^M \) related to \( \mathbf{x} \) through a linear model

$$ y = K \mathbf{x} + w, \quad (1) $$

where \( K = (K_{m,n})_{1 \leq m \leq M, 1 \leq n \leq N} \in \mathbb{R}^{M \times N} \) is the observation matrix corresponding to a discretized version of a Laplace transform, and \( w = (w_m)_{1 \leq m \leq M} \in \mathbb{R}^M \) represents some additive acquisition noise. An efficient resolution strategy is to employ a penalized approach that defines an estimate \( \hat{\mathbf{x}} \in \mathbb{R}^N \) of \( \mathbf{x} \) as a solution to the following constrained minimization problem [3]:

$$ \minimize_{\mathbf{x} \in \mathbb{R}^N} \Psi(\mathbf{x}) \text{ subject to } ||K \mathbf{x} - \mathbf{y}|| \leq \tau, \quad (2) $$

where \( \tau > 0 \) is a parameter depending on the noise characteristics, and \( \Psi \) is a regularization function that allows to add prior information on the sought signal and to impose the positivity of its entries. The so-called maximum entropy reconstruction approach [4], corresponding to the choice of the Shannon entropy function [5, 6] for the penalization term \( \Psi \), has been at the core of several papers dealing with regularized inverse Laplace transform [7, 8, 9, 10, 11, 12]. A more recent approach consists in adopting for \( \Psi \) a criterion enforcing both sparsity and positivity, with the aim to improve the resolution of narrow peaks possibly present in the sought signal [13, 14, 15, 16], but this strategy may be at the price of losing the smoothness of the solution. Hybrid regularization approaches combining both entropy and sparsity terms in \( \Psi \) should thus be envisaged so as to derive a flexible resolution method. However, the choice of an efficient optimization algorithm to solve the resulting constrained problem (2) (or its Lagrangian formulation) remains a challenging task.

On the one hand, in the case of entropy regularization, the optimization techniques proposed in the literature are usually ad hoc and difficult to extend to other types of priors [17, 18, 19]. On the other hand, in the case of sparsity regularization, the minimization step is usually handled by proximal optimization methods [20]. These methods, grounded on the use of the proximity operator [21] are highly flexible, and particularly efficient for handling functionals for which the proximity operators have a closed form expression. In this paper, we provide a comprehensive list of formula for the proximity operators of new hybrid regularization functions combining a non necessarily convex sparsity enhancing term and an entropy penalty. This list constitutes a very helpful tool for
deriving fast resolution methods for the regularized inversion of (1). The applicability of our method is illustrated by means of an example in the context of DOSY NMR.

The paper is organized as follows: In section 2, we introduce our notation and recall the definition of the proximity operator. Section 3 presents our main contribution, that is the expression of the proximity operators of several combination of entropy and sparsity promoting penalties. Then, Section 4 presents an application of our results to the resolution of an inverse problem arising in DOSY NMR signal reconstruction. Finally, some conclusions are drawn in Section 5.

2. PROBLEM STATEMENT

2.1. Hybrid regularization

Let us define the following family of hybrid regularization functions of the form:

$$\left(\forall x \in \mathbb{R}^N\right) \Psi(x) = \alpha \Psi_1(x) + \beta \Psi_2(x),$$

where $\Psi_1 : \mathbb{R}^N \rightarrow [-\infty, +\infty]$ and $\Psi_2 : \mathbb{R}^N \rightarrow \mathbb{R}$ are lower semi continuous (lsc) and proper functions and $(\alpha, \beta)$ are some positive weights. In this paper, we focus on the case when $\Psi_1$ is either the Shannon or the Burg entropy [22] and $\Psi_2$ is a sparsity promoting prior. An efficient strategy to promote the sparsity of the sought signal is to choose $\Psi_2$ as the $\ell_0$ penalty that counts the number of nonzeros in $x$. However, the later function is neither differentiable nor continuous at 0, so that its convex approximation, the $\ell_1$ norm, that sums the absolute value of the signal entries is often used instead [23]. Continuous, but non convex, approximations have also been proposed for the $\ell_0$ penalty, namely the log-sum penalty [24] and the Cauchy penalty [25]. Note that all the aforementioned regularization terms are separable, so that (3) can be rewritten as

$$\left(\forall x \in \mathbb{R}^N\right) \Psi(x) = \sum_{n=1}^{N} \psi(x_n),$$

with $\psi = \alpha \psi_1 + \beta \psi_2$, $\psi_1 : \mathbb{R} \rightarrow [-\infty, +\infty]$ and $\psi_2 : \mathbb{R} \rightarrow \mathbb{R}$. Figures 1 and 2 illustrate the considered functions, for the entropy and sparse penalty terms $\psi_1$ and $\psi_2$, respectively.

2.2. Proximity Operator

When $\Psi$ is convex on $\mathbb{R}^N$, its proximity operator at $x \in \mathbb{R}^N$ is defined as the unique minimizer of $\Psi + \frac{1}{2}\|y-x\|^2$ [26]. This operator has been generalized for lsc proper functions that are not necessarily convex in [27, Sec.XV-4], as the multi-valued operator:

$$\text{prox}_\psi : x \mapsto \text{Argmin}_{y \in \mathbb{R}^N} \left( \frac{1}{2}\|y-x\|^2 + \Psi(y) \right).$$

Since $\Psi$ in (4) takes a separable form, its proximity operator (5) is given by ([21]):

$$\text{prox}_\psi(x) = (p_n(x_n))_{1 \leq n \leq N},$$

where, for every $n \in \{1, \ldots, N\}$,

$$p_n(x_n) = \text{prox}_{\alpha \psi_1 + \beta \psi_2}(x_n).$$

Fig. 1. Examples of entropy based penalties, in the scalar case: Shannon (continuous red line) and Burg (dashed blue line) entropy priors. Note that $\psi_1 = +\infty$ for negative values of $x$.

Fig. 2. Examples of sparsity promoting penalties, in the scalar case: $\ell_0$ (continuous black line), $\ell_1$ (dashed-dotted green line), log-sum (dashed red line) and Cauchy (dotted blue line) priors.

In the sequel, we will focus our attention on the resolution of the scalar optimization problem (7), when $\alpha \in [0, +\infty[$ and $\beta \in [0, +\infty[$.
3. PROXIMITY OPERATORS FOR HYBRID SPARSE + ENTROPY PRIORS

3.1. Shannon entropy

The Shannon entropy regularization is defined, for all \( x \in \mathbb{R} \), as
\[
\psi_1(x) = \begin{cases} 
\log x & \text{if } x > 0 \\
0 & \text{if } x = 0 \\
\infty & \text{elsewhere.}
\end{cases}
\]
(8)

According to [20], the proximity operator of \( \alpha \psi_1 \) at \( x \in \mathbb{R}^N \), with \( \alpha \in [0, +\infty[ \), reads:
\[
\text{prox}_{\alpha \psi_1}(x) = \alpha W \left( \frac{1}{\alpha} \exp \left( \frac{x}{\alpha} - 1 \right) \right),
\]
(9)
where \( W \) states for the \( W \) Lambert function [28]. The remainder of this section shows how to generalize this expression in the case when \( \beta \in [0, +\infty[ \).

3.1.1. Shannon entropy + \( \ell_1 \)

Set \( (\alpha, \beta) \in [0, +\infty[^2 \). The hybrid Shannon entropy + \( \ell_1 \) penalty is defined, for every \( x \in \mathbb{R} \), as
\[
\psi(x) = \begin{cases} 
\alpha x \log x + \beta x & \text{if } x > 0 \\
0 & \text{if } x = 0 \\
\infty & \text{elsewhere.}
\end{cases}
\]
(10)

Its proximity operator reads, for every \( x \in \mathbb{R} \):
\[
\text{prox}_{\psi}(x) = \alpha W \left( \frac{1}{\alpha} \exp \left( \frac{x - \beta}{\alpha} - 1 \right) \right),
\]
(11)

3.1.2. Shannon entropy + \( \ell_0 \)

Set \( (\alpha, \beta) \in [0, +\infty[^2 \). The hybrid Shannon entropy + \( \ell_0 \) penalty is defined, for every \( x \in \mathbb{R} \), as
\[
\psi(x) = \begin{cases} 
\alpha x \log x + \beta x & \text{if } x > 0 \\
0 & \text{if } x = 0 \\
\infty & \text{elsewhere.}
\end{cases}
\]
(12)

Its proximity operator reads, for every \( x \in \mathbb{R} \):
\[
\text{prox}_{\psi}(x) = \begin{cases} 
p & \text{if } \beta < \beta \text{ (13)} \\
\{0, p\} & \text{if } \beta = \beta \\
0 & \text{elsewhere,}
\end{cases}
\]
where \( \beta = \frac{1}{2} p^2 + \alpha p \in [0, +\infty[ \) and
\[
p = \alpha W \left( \frac{1}{\alpha} \exp \left( \frac{x}{\alpha} - 1 \right) \right).
\]
(14)

Note that, due to the discontinuity of the \( \ell_0 \) penalty in 0, \( \text{prox}_{\psi} \) is multi-valued in the limit case when \( \beta = \beta \).

3.1.3. Shannon + log-sum

Set \( (\alpha, \beta, \delta) \in [0, +\infty[^3 \). The hybrid Shannon entropy + log-sum penalty is defined, for every \( x \in \mathbb{R} \), as
\[
\psi(x) = \begin{cases} 
\alpha x \log x + \beta \log(\delta + x) & \text{if } x > 0 \\
\beta \log(\delta) & \text{if } x = 0 \\
+\infty & \text{elsewhere.}
\end{cases}
\]
(15)

For every \( x \in \mathbb{R} \), its proximity operator is given by
\[
\text{prox}_{\psi}(x) = \text{Argmin}_{p \in [0, +\infty[} \{ \frac{1}{2} (x - p)^2 + \psi(p) \},
\]
with \( \varphi(p) = p^2 + (\delta - x + \alpha)p + \alpha(\delta + p) \log p + \delta(\alpha - x) + \beta \).

3.1.4. Shannon + Cauchy

Set \( (\alpha, \beta, \delta) \in [0, +\infty[^3 \). The hybrid Shannon entropy + Cauchy penalty is defined, for every \( x \in \mathbb{R} \), as
\[
\psi(x) = \begin{cases} 
\alpha x \log x + \beta \log(\delta + x^2) & \text{if } x > 0 \\
\beta \log(\delta) & \text{if } x = 0 \\
+\infty & \text{elsewhere.}
\end{cases}
\]
(16)

For every \( x \in \mathbb{R} \), its proximity operator is given by
\[
\text{prox}_{\psi}(x) = \text{Argmin}_{p \in [0, +\infty[} \{ \frac{1}{2} (x - p)^2 + \psi(p) \},
\]
with \( \varphi(p) = p^3 + (\alpha - x)p^2 + (\delta + 2\beta)p + \alpha(\delta + p^2) \log p + \delta(\alpha - x) \).

3.2. Case of Burg entropy

The Burg entropy regularization is defined, for all \( x \in \mathbb{R} \), as
\[
\psi_1(x) = \begin{cases} 
-\log x & \text{if } x > 0 \\
+\infty & \text{elsewhere.}
\end{cases}
\]
(17)

According to [29], the proximity operator of the Burg entropy function reads, for every \( x \in \mathbb{R}^N \), for every \( \alpha \in [0, +\infty[ \):
\[
\text{prox}_{\alpha \psi_1}(x) = \frac{x + \sqrt{x^2 + 4\alpha}}{2}.
\]
(18)

Let us now present the extension of this result to the case when \( \beta \in [0, +\infty[ \).

3.2.1. Burg entropy + \( \ell_1 \)

Set \( (\alpha, \beta) \in [0, +\infty[^2 \). The hybrid Burg entropy + \( \ell_1 \) penalty is defined, for every \( x \in \mathbb{R} \), as
\[
\psi(x) = \begin{cases} 
-\alpha \log x + \beta x & \text{if } x > 0 \\
+\infty & \text{elsewhere.}
\end{cases}
\]
(19)

Its proximity operator reads, for every \( x \in \mathbb{R} \) ([29]):
\[
\text{prox}_{\psi}(x) = \frac{x - \beta + \sqrt{(\beta - x)^2 + 4\alpha}}{2}.
\]
(20)
3.2.2. Burg + $\ell_0$

Set $(\alpha, \beta) \in [0, +\infty]^2$. The hybrid Burg entropy + $\ell_0$ penalty is defined, for every $x \in \mathbb{R}$, as

$$
\psi(x) = \begin{cases} 
-\alpha \log x + \beta & \text{if } x > 0 \\
+\infty & \text{elsewhere.}
\end{cases}
$$

(21)

Its proximity operator reads, for every $x \in \mathbb{R}$:

$$
\text{prox}_\psi(x) = \frac{x + \sqrt{x^2 + 4\alpha}}{2}.
$$

(22)

3.2.3. Burg + log-sum

Set $(\alpha, \beta, \delta) \in [0, +\infty]^3$. The hybrid Burg entropy + log-sum penalty is defined, for every $x \in \mathbb{R}$, as

$$
\psi(x) = \begin{cases} 
-\alpha \log x + \beta \log(\delta + x) & \text{if } x > 0 \\
+\infty & \text{elsewhere.}
\end{cases}
$$

(23)

For every $x \in \mathbb{R}$, its proximity operator is given by

$$
\text{prox}_\psi(x) = \text{Argmin}_{p \in [0, +\infty] \text{ s.t. } \varphi(p) = 0} \left( \frac{1}{2} (x - p)^2 + \varphi(p) \right),
$$

with

$$
\varphi(p) = p^3 + (\delta - x)p^2 + p(\beta - \delta x - \alpha) - \delta \alpha.
$$

3.2.4. Burg + Cauchy

Set $(\alpha, \beta, \delta) \in [0, +\infty]^3$. The hybrid Burg entropy + Cauchy penalty is defined, for every $x \in \mathbb{R}$, as

$$
\psi(x) = \begin{cases} 
-\alpha \log x + \beta \log(\delta + x^2) & \text{if } x > 0 \\
+\infty & \text{elsewhere.}
\end{cases}
$$

(24)

For every $x \in \mathbb{R}$, its proximity operator is given by

$$
\text{prox}_\psi(x) = \text{Argmin}_{p \in [0, +\infty] \text{ s.t. } \varphi(p) = 0} \left( \frac{1}{2} (x - p)^2 + \varphi(p) \right),
$$

with

$$
\varphi(p) = p^4 - x^3 \beta + (\delta + 2 \beta - \alpha)p^2 - \delta x p - \delta \alpha.
$$

4. APPLICATION TO NMR SIGNAL RECONSTRUCTION

4.1. Problem formulation

During the DOSY (Diffusion Order SpectroscopY) experiment, proposed by [30] to analyze the properties of complex chemical mixtures, a series of measurements are acquired for different pulsed field gradient strengths, the data are then processed with the aim to separate different species according to their diffusion coefficient. The DOSY NMR data $y \in \mathbb{R}^M$ gathers the results of $M \geq 1$ experiments corresponding to different acquisition settings characterized by the set of parameters $t = (t_m)_{1 \leq m \leq M}$. The problem is then to estimate, from these measurements, the values of the true diffusion distribution $\mathcal{V}(T)$ at given positions $T = (T_m)_{1 \leq m \leq N}$. The relation between $y$ and the sought signal $\mathcal{V}$ is given, for every $m \in \{1, \ldots, M\}$ and every $n \in \{1, \ldots, N\}$ by

$$
K_{m,n} = \exp(-T_m t_n),
$$

(25)

where $w \in \mathbb{R}^M$ is a perturbation noise. In practice, the noise is assumed to be zero-mean Gaussian, i.i.d, with known standard deviation $\sigma > 0$ so that we propose to find an estimate $\hat{x} \in \mathbb{R}^N$ of $\mathcal{V}$ by solving (2) with $\tau = \eta \sqrt{M} \sigma$, where $\eta > 0$ is a weight closed to 1 [31].

4.2. Experimental results

We now present the experimental results obtained when, for every $n \in \{1, \ldots, N\}$, for every $m \in \{1, \ldots, M\}$,

$$
\begin{align*}
\left\{ t_m = t_{\min} + \frac{m-1}{M} t_{\max}, \\
T_n = T_{\min} \exp \left( -\frac{(m-1)}{N} \log \left( \frac{T_{\max}}{T_{\min}} \right) \right) \right. 
\end{align*}
$$

(26)

with $M = 50$, $N = 200$, $T_{\min} = 1$, $T_{\max} = 10^3$, $t_{\min} = 0$, $t_{\max} = 1.5$. We consider two synthetic signals A and B. The synthetic signal A models a monodisperse distribution with symmetric log-normal shape located at $T = 16$ while signal B corresponds to a polydisperse distribution that is the sum of two log-normal patterns located at $T_1 = 4$ and $T_2 = 32$. Table 1 presents the resulting signal to noise ratio (SNR) defined as $10 \log_{10} \left( \|\mathcal{V}\|^2 / \|\hat{x} - \mathcal{V}\|^2 \right)$ which was obtained when solving Problem (2) for both datasets, for different values of $\sigma$, and several choices for the regularization term $\psi$, namely (8), (10), (17) and (19). Here, we choose to focus only on convex priors so that the optimization problem can be efficiently solved with the PPXA+ algorithm [32]. Parameter $\eta$ is set equals to 1.2, since it was observed to lead to the best reconstruction results. Moreover, when hybrid penalties are considered, we fix $\beta = 1 - \alpha$, and we optimize $\alpha \in [0, 1]$ manually so as to obtain a minimal reconstruction error.

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>Shannon prior</th>
<th>Shannon $+\ell_1$ prior</th>
<th>Burg prior</th>
<th>Burg $+\ell_1$ prior</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^{-2}</td>
<td>12.45</td>
<td><strong>13.16</strong></td>
<td>12.92</td>
<td><strong>12.92</strong></td>
</tr>
<tr>
<td>10^{-3}</td>
<td>18.16</td>
<td><strong>20.86</strong></td>
<td>12.11</td>
<td><strong>13.44</strong></td>
</tr>
<tr>
<td>10^{-4}</td>
<td>20.87</td>
<td><strong>25.95</strong></td>
<td>12.03</td>
<td><strong>15.53</strong></td>
</tr>
<tr>
<td>10^{-5}</td>
<td>11.14</td>
<td><strong>18.52</strong></td>
<td>6.20</td>
<td><strong>9.54</strong></td>
</tr>
<tr>
<td>Dataset A</td>
<td>18.11</td>
<td><strong>20.23</strong></td>
<td>7.41</td>
<td><strong>10.95</strong></td>
</tr>
<tr>
<td>10^{-5}</td>
<td>19.05</td>
<td><strong>26.30</strong></td>
<td>7.54</td>
<td><strong>10.98</strong></td>
</tr>
</tbody>
</table>

Table 1. SNR in dB of the restored signals A and B for various choices of the penalization function $\psi$. 
One can observe that the addition of the $\ell_1$ norm in the penalty functions improves significantly the quality of restoration results, especially in the case of the polydisperse signal B. The best results are obtained with the combination of Shannon entropy and $\ell_1$ penalties. This is also confirmed by visual inspection, as it can be noticed in Figure 3 displaying an example of reconstruction result we obtained in the case of dataset B, with $\sigma = 10^{-5}$.

![Original signal, Shannon, Shannon + $\ell_1$, Burg, Burg + $\ell_1$](image)

**Fig. 3.** Reconstruction results for dataset B and $\sigma = 10^{-5}$.

### 5. CONCLUSION

This work expands the extensive list of proximity operators available in the literature [20, 29, 33] by considering the case of separable functions combining entropy and sparsity promoting terms. Through numerical experiments, we show that these functions may serve as efficient hybrid penalties for solving ill-posed inverse problems in the context of DOSY NMR spectroscopy. Since only convex penalties were tested in our experiments, the PPXA+ algorithm was retained. It should however be emphasized that the novel proximity operators we derive could be applied in a variety of proximal algorithms, in the convex [34] or the non-convex case [35, 36]. The latter case should be of particular interest in the context of blind signal restoration problems such as those encountered in [37, 38] where the proposed hybrid penalties could be beneficial.

### 6. REFERENCES


