On Convex Envelopes and Regularization of Non-Convex Functionals without moving Global Minima.

Marcus Carlsson

Abstract

We provide theory for the computation of convex envelopes of non-convex functionals including an $\ell^2$-term, and use these to suggest a method for regularizing a more general set of problems. The applications are particularly aimed at compressed sensing and low rank recovery problems but the theory relies on results which potentially could be useful also for other types of non-convex problems. For optimization problems where the $\ell^2$-term contains a singular matrix we prove that the regularizations never move the global minima. This result in turn relies on a theorem concerning the structure of convex envelopes which is interesting in its own right. It says that at any point where the convex envelope does not touch the non-convex functional we necessarily have a direction in which the convex envelope is affine.

1 Introduction

This article is a compressed and improved version of [13], which contains more information and potentially more errors. The present work is the extension of a chain of ideas with its roots in compressed sensing. $\ell^1 - \ell^2$-minimization tricks have a long history and got renewed attention with the work of Donoho, Candès and Tao among others [15, 16, 12]. In the same spirit the nuclear norm minimization strategy was investigated by Fazel and coworkers [17, 28] and in both cases it was shown that these methods yield perfect reconstructions in the case of no noise. However, in realistic scenarios these results often do not apply and moreover there is of course noise, in which case the methods come with a (sometimes severe) bias. Moreover they are slow since one needs to find an appropriate value of involved penalty parameters.

Due to such issues there is a wealth of non-convex variations to replace $\ell^1$/nuclear norm in the area of compressed sensing, we refer to [14] for a survey. Two fairly recent contributions in this vein is the work by Carl Olsson and coworkers [20] as well as by Gilles Aubert and coworkers [30]. The former paper deals with non-convex matrix minimization problems with subspace constraints, the latter with sparse reconstructions, and in particular the latter shows that the concrete regularizer considered there has the desirable property of not moving global minima. In this paper we find a unifying framework and show that all these penalties are particular cases of the so called “proximal hull” or “quadratic envelope”. We systematically study this as a regularizer and in particular we lift the result of Aubert et al. to a general context. In order to do so we provide new results on the structure of lower semi-continuous (abbreviated l.s.c.) convex envelopes which are interesting in their own right. More precisely we show that whenever a l.s.c. convex envelope is not in touch with the function that generates it, then it necessarily has a direction in which it is affine linear.

Centre for Mathematical Sciences, Lund University, mc@maths.lth.se
2 Outline and motivation

We develop methods to compute the lower semi-continuous convex envelope of functionals of the form
\[ f(x) + \frac{1}{2} \| x - d \|_2^2, \tag{1} \]
and show that this is of the form \( Q(f)(x) + \frac{1}{2} \| x - d \|_2^2 \), where \( Q(f) \) is the proximal hull or “quadratic envelope”, as we shall call it. Here \( x \) can be in any separable Hilbert space but \( f \) needs to be such that the global minimization of (1) is computable. The practical applications of \( Q(f) \) pertains to optimization of (1) with additional constraints, as well as unconstrained optimization of
\[ f(x) + \frac{1}{2} \| Ax - d \|_2^2, \tag{2} \]
where \( A \) is a linear operator.

To introduce the main ideas behind this work we consider two concrete problems. A multitude of applications can be posed mathematically as finding the lowest rank matrix \( X \) satisfying some equation \( A(X) = d \), where \( A \) is a linear operator and \( d \) is a measurement (see e.g. [28, 32]). Usually the measurement \( d \) is not perfect so in practice one wishes to find the minimum rank given some accepted error; \( \| A(X) - d \| \leq \rho \). The dual formulation of this problem is
\[ \arg\min_X \lambda \text{rank}(X) + \| A(X) - d \|_2^2, \tag{3} \]
where \( \lambda \) is a parameter. However, the functional \( \text{rank}(X) \) is non-convex and highly discontinuous, so the problem can not be solved as stated (in general). It can be solved for the case \( A = I \) but the problem is still hard when combined with additional priors, see e.g. Section 1.1 in [20] for an overview and applications in signal processing and imaging.

Due to the problematic nature of \( \text{rank}(X) \) it has become popular to replace \( \text{rank}(X) \) with the nuclear norm of \( X \). However, \( \text{rank}(X) \) and the nuclear norm are quite far apart and the method leads to a bias in the solution, which led the authors of [20] to suggest working instead with the convex envelope of \( \text{rank}(X) + \frac{1}{2} \| X - D \|_F^2 \) for which they obtained an explicit expression. They also provided the convex envelope when \( \text{rank}(X) \) is replaced by the indicator functional of the set \( \{ X : \text{rank}(X) \leq K \} \), in order to treat problems where a matrix of a fixed rank is sought, and this convex envelope was further studied in [1].

Independently, convex envelopes was used in [30] to suggest a regularizer to functionals of the type
\[ \| x \|_0 + \frac{1}{2} \| Ax - d \|_2^2, \quad x \in \mathbb{R}^n, \tag{4} \]

Figure 1: Regularizing (17) by \( Q(\text{card}) \) and \( Q(\ell^1) \) finds the oracle solution up to noise levels of around \( \| \epsilon \| = 3 \) and 4 (roughly 30\% of \( \| d \| \) whereas \( \ell^1 \)-regularization only finds this solution with no noise.
which is usually dealt with by replacing $\|x\|_0$ by $\lambda\|x\|_1$. The main contribution of their work is to show that their regularizer does not move global minima. A common misconception is that the same holds for $\ell^1$-methods, which is true only if there is no noise \[11\]. In the presence of noise the estimates for $\ell^1$-methods are rather poor and \[30\] is the first framework which allows for regularization without moving minima in a more realistic scenario.

This paper presents a unified approach to this circle of ideas by connecting them with the “quadratic envelope” $Q(f)$. We also extend the findings of \[30\] to any problem of the form \( (2) \) as long as $Q(f)$ is computable. An expanded version of this article is found in \[13\] which contains a long list of instances where $Q(f)$ is computable.

In particular $Q(f)$ is computable for $\ell_K$; the indicator functional for $\{x \in \mathbb{C}^n : \|x\|_0 \leq K \}$. As a proof of concept, we compare performance of \( (4) \) with version of \( (4) \) for $Q(\text{card})$ and with $Q(\ell_K)$ (where $\text{card}(x) = \|x\|_0$). We use a $100 \times 200$ matrix $A$ and minimize the regularized version of \( (4) \) for $d$ of the form $Ax_0 + \epsilon$, where $x_0$ has cardinality 10 and $\epsilon$ takes on various levels of noise. As noted in \[10\] the best one can hope for is then to recover the so-called “oracle solution” $x_S$ (obtained if an oracle a priori revealed the correct support). As Figure \[1\] shows both $Q(\text{card})$ and $Q(\ell_{10})$ outperform $\ell^1$ and finds the oracle solution for fairly large levels of noise. Also $Q(\ell_{10})$ beats $Q(\text{card})$, which is no surprise since it contains additional information about the problem built into it, and demonstrates the versatility of the new $Q$-transform. The article \[13\] contain much more information about this particular case.

We now outline the main contributions of this paper in greater detail. Consider any functional of the form
\[
 f(x) + \frac{\gamma}{2} \|x-d\|^2_V
\]

where $\gamma > 0$ is a parameter, $V$ is an arbitrary separable Hilbert space and $f$ a non-negative functional on $V$. In Section \[3\] we introduce the transform $Q_{\gamma}$ and show that the l.s.c. convex envelope of the functional in \( (5) \) is
\[
 Q_{\gamma}(f)(x) + \frac{\gamma}{2} \|x-d\|^2_V.
\]

In order for $Q_{\gamma}(f)$ to be computable, the global minimization of \( (5) \) needs to be solvable, and hence the problem of minimizing \( (5) \) in itself is not an instance where the $Q_{\gamma}$-transform is useful. However, it is useful for finding global minimizers of \( (5) \) in combination with additional prior restrictions. To illustrate, consider the problem
\[
 \arg \min_{x \in \mathcal{H}} f(x) + \frac{1}{2} \|x-d\|^2,
\]

where $\mathcal{H}$ is a closed convex subset of $V$, and suppose we are unable to find a closed form solution. Upon replacing \( (7) \) with
\[
 \arg \min_{x \in \mathcal{H}} Q_{\gamma}(f)(x) + \frac{1}{2} \|x-d\|^2,
\]

for some fixed $\gamma \leq 1$, we obtain a convex problem which can be solved. However, even for $\gamma = 1$ it is possible that \( (7) \) and \( (8) \) have different solutions, despite the functional in \( (8) \) being the l.s.c. convex envelope of the one in \( (7) \). The rationale behind replacing \( (7) \) by \( (8) \) is pragmatical; since the latter is convex the solution may be found using convex optimization routines. This may seem ad hoc but we remind the reader that replacing e.g. $\|x\|_0$ by $\|x\|_1$ or rank($X$) by the nuclear norm $\|X\|_1$ has had a substantial impact, and that for these concrete cases the modification $Q_{\gamma}(f)$ is much closer to the original functional $f$ (which leads to a better performance as an estimator, see the numerical
sections of \([11, 20]\)). A reason for this is that \(Q_\gamma(f)\) has the desirable feature that \(Q_\gamma(f)(x) = f(x)\) often holds, and since \((8)\) is a convex problem below the original problem \((7)\), it is easy to see that a minimum \(\hat{x}\) to \((3)\) is the solution to \((7)\) whenever \(Q_\gamma(f)(\hat{x}) = f(\hat{x})\). This is highlighted in Figure 2 where the two problems have the same solution. More information and examples on this type of problems is found in Part II of \([13]\).

In Section 5 we consider regularization of functionals like \((3)\) and \((4)\), or more generally

\[
f(x) + \frac{1}{2} \|Ax - d\|_W^2
\]

for arbitrary non-negative \(f\), where \(A : \mathcal{V} \rightarrow \mathcal{W}\) is a linear operator between separable Hilbert spaces. We assume that \(\mathcal{V}\) is such that \(Q_\gamma(f)\) is computable and that the convex envelope of \((9)\) is untractable. We propose to use as regularizer the function \(Q_\gamma(f)\), i.e. we will study the relationship between minimizers of \((9)\) and those of

\[
Q_\gamma(f)(x) + \frac{1}{2} \|Ax - d\|_W^2.
\]

Since it often holds that \(Q_\gamma(f)(x) = f(x)\), we again see that a global minimizer of \((10)\) for which this is the case must also be a global minimizer of \((7)\), in view of the inequality \(Q_\gamma(f) \leq f\) (shown in Section 3). The parameter \(\gamma\) now becomes a useful tool as it tunes the curvature of \(Q_\gamma(f)\) and we pause to illustrate its role by considering a toy problem in one variable; see Figure 3. We let \(\|x\|_0\) be the function equalling 1 on \(\mathbb{R} \setminus \{0\}\) and zero at \(x = 0\). In red we see the functional \(\|x\|_0 + \frac{1}{2} |x - 1|^2\) (which is a particular case of both \((3)\) and \((1)\) in dimension 1, the matrix \(A\) is here the number 1), in blue its convex envelope and in pink the \(\ell^1\) convex relaxation \(\|x\|_1 + \frac{1}{2} |x - 1|^2\). Clearly the global minimum of the red and blue coincide, but the global minimum of the \(\ell^1\)-relaxation is different. For
we have two options, either $|A|^2 > \gamma$ or $|A|^2 < \gamma$. The regularizer \[(10)\] is illustrated in black for these two cases in Figure 3. The circles represent global minima of the respective functions. In the case $|A|^2 > \gamma$ we see that \[(10)\] is a convex minorant of \[(9)\] whose global minima (for this choice of parameters) is equal to that of \[(10)\]. In the case $|A|^2 < \gamma$, \[(10)\] is no longer convex but the local minima of \[(10)\] are also minima of \[(9)\], and \[(10)\] has fewer local minima. In particular the global minima coincide. The main point of the paper is loosely that the general behavior is the same.

In Section 5.1 we generalize the situation in Figure 3 (left) and assume that $\gamma$ satisfies $A^*A \succcurlyeq \gamma I$, \[(11)\] For such choice of $\gamma$ we prove that the functional \[(10)\] is a convex functional below \[(9)\] and hence minimization of \[(10)\] will produce a minimizer which, although not necessarily equal to the minimizer of the original problem, potentially is closer than that obtained by other convex relaxation methods.

For the problem \[(4)\] $A$ is usually a matrix with a large kernel which rules out the above approach. In Section 5.2 we consider the case \[(12)\] generalizing the situation in the right picture of Figure 3. We can then show that \[(10)\] is a continuous (but not convex) functional with the following desirable properties:

i) \[(10)\] lies between \[(9)\] and its l.s.c. convex envelope,

ii) any local minimizer of \[(10)\] is a local minimizer of \[(9)\],

iii) the global minimizers of \[(10)\] and \[(9)\] coincide.

These findings in turn rely on general results about l.s.c. convex envelopes which we provide in Section 4. The computation of the l.s.c. convex envelope of $f(x) + \frac{\gamma}{2}||x||^2$ can be thought of as stretching a plastic foil from below onto the graph of $f(x) + \frac{\gamma}{2}||x||^2$ (see Figure 2). Consider a point $x$ where the plastic foil is not in contact with the graph, i.e. where $Q_\gamma(f)(x) < f(x)$. It is intuitively obvious that the plastic foil, i.e. the graph of $Q_\gamma(f)(x) + \frac{\gamma}{2}||x||^2$, has some direction in
which it is affine linear and thus $Q_\gamma(f)$ should have some direction in which the curvature is $-\gamma$. This is surprisingly difficult to show and despite the wealth of results on l.s.c. convex envelopes it is not found in any standard reference on the topic. The statement is shown in the PhD-thesis [22] for the finite dimensional case. Here we provide a proof is based on an extension of Milman’s theorem due to Arne Brøndsted [9] in a short note from 1966.

The final Section 6 is more practical in nature. Critical points of (10) can be found using the forward-backward splitting method (FBS), given that $Q_\gamma(f)$ is “semi-algebraic”, as was shown in [3]. To simplify verification of when $Q_\gamma(f)$ is semi-algebraic we show in Section 6 that this is true as long as $f$ itself is semi-algebraic. Further tools to compute $Q_\gamma(f)$ as well as related proximal operators are found in [13].

3 The quadratic envelope

Let $V$ be a separable Hilbert space over $\mathbb{R}$ or $\mathbb{C}$, such as $\mathbb{C}^n$ with the canonical norm $\|x\|_2^2 = \sum_{j=1}^n |x_j|^2$ or $\mathbb{M}_{m,n}$, equipped with the Frobenius norm which we denote $\|X\|_F$. All Hilbert spaces over $\mathbb{C}$ are also Hilbert spaces over $\mathbb{R}$ with the scalar product $\langle x, y \rangle_\mathbb{R} = \text{Re} \langle x, y \rangle$ and hence it is no restriction to assume that $V$ is a real Hilbert space wherever needed. Even if $V$ is a Hilbert space over $\mathbb{C}$ we will implicitly assume that the scalar product is $\langle x, y \rangle_\mathbb{R}$.

Given any functional $f: V \to \mathbb{R} \cup \{\infty\}$ and parameter $\gamma > 0$ we introduce the “quadratic envelope” $Q_\gamma$ as the supremum of all minimizers of the form $\alpha - \frac{\gamma}{2} \|x - y\|^2$ for $\alpha \in \mathbb{R}$ and $y \in V$;

$$Q_\gamma(f)(x) = \sup_{\alpha \in \mathbb{R}, y \in V} \left\{ \alpha - \frac{\gamma}{2} \|x - y\|^2 : \alpha - \frac{\gamma}{2} \|\cdot - y\|^2 \leq f \right\}. \quad (13)$$

The quadratic envelope has appeared previously e.g. in [29] under the name “proximal hull”, denoted $h_{\gamma^{-1}}$ (Example 1.44), but it seems that the term is not widespread (see the discussion in Section 7) and it seems that its connection with convex envelopes has not been noted or at least not systematically studied. We prefer the term quadratic envelope since it is more illustrative, and prefer the notation $Q_\gamma$ since it would be messy to always have to invert $\gamma$ which in this context has a concrete meaning: The parameter $\gamma$ basically tunes the maximum negative curvature of $Q_\gamma(f)$ as we shall see in Section 4 (Corollary 4.4). When $\gamma = 1$ we simply write $Q$, as opposed to $Q_\gamma$. In this section we first provide some tools to compute $Q_\gamma$, then prove the connection with l.s.c. convex envelopes and end with some auxiliary results and a discussion of connections to previous concepts and terminology.

The Legendre transform (or Fenchel conjugate) is defined as $g^*(y) := \sup_x \langle x, y \rangle - g(x)$. We remind the reader that $g^*$ is l.s.c convex and

Figure 4: Illustration of a non-convex function $f$ (red) and its quadratic envelope $Q_2(f)$ (black). The black graph lies slightly below for illustration only.
that $g^{**}$ equals the l.s.c. convex envelope of $g$ by the Fenchel-Moreau theorem (see e.g. Proposition 13.11 and 13.39 in [4]). We now introduce the transform $S_\gamma$ defined as follows:

$$S_\gamma(f)(y) := \left( f(\cdot) + \frac{\gamma}{2} \| \cdot \|^2 \right)^* (\gamma y) - \frac{\gamma}{2} \| y \|^2 = \sup_x f(x) - \frac{\gamma}{2} \| x - y \|^2.$$  (14)

$S_\gamma$ is simply the negative of the Moreau envelope computed with constant $\gamma^{-1}$. If we set $q_\gamma(x, y) = -\frac{\gamma}{2} \| x - y \|^2$ then, in the terminology of [29] Sec. 11.L, $S_\gamma(f)$ is the $q_\gamma$-conjugate of $f$ and $Q_\gamma(f)$ the $q_\gamma$-envelope of $f$ (reinforcing our choice of terminology “quadratic envelope” for $Q_\gamma$). We introduce the symbol $S_\gamma$ mainly since we believe the notation $-e_{\gamma^{-1}}(f)$ or $q_\gamma f$ (c.f. [29]) or $-\gamma^{-1} f$ (c.f. [4]) would be confusing for our present purposes. Its connection to the quadratic envelope is described by the following proposition;

**Proposition 3.1.** Let $\gamma > 0$ and let $f$ be a $[0, \infty]$-valued l.s.c. functional on a separable Hilbert space $V$. We have $Q_\gamma = S_\gamma \circ S_\gamma := S^{2}_{\gamma}$, i.e.

$$Q_\gamma(f)(x) = \sup_y \left( \inf_w f(w) + \frac{\gamma}{2} \| w - y \|^2 \right) - \frac{\gamma}{2} \| x - y \|^2.$$  (15)

**Proof.** The argument is a replica of Example 1.44 of [29], but is included for completeness. We have $\alpha - \frac{\gamma}{2} \| \cdot \|^2 \leq f$ iff $\alpha \leq f + \frac{\gamma}{2} \| \cdot - y \|^2$ so the maximal $\alpha$ for fixed $y$ is given by $\alpha = -S_\gamma(f)(y)$. Thus $Q_\gamma(f)(x) = \sup_{y \in V} -S_\gamma(f)(y) - \frac{\gamma}{2} \| x - y \|^2 = Q_\gamma(f)(x)$ as desired. $\square$

The next proposition contains some basic observations on the behavior of $S_\gamma$ and $Q_\gamma$.

**Proposition 3.2.** Let $\gamma > 0$ and let $f$ be a $[0, \infty]$-valued l.s.c. functional on a separable Hilbert space $V$. Then $S_\gamma(f)$ takes values in $(-\infty, 0]$ and is continuous whereas $Q_\gamma(f)$ is lower semi-continuous, takes values in $[0, \infty]$ and is continuous in the interior of $\text{dom}(Q_\gamma(f))$.

**Proof.** The statement of the interchanging signs follows easily by the last line of (14) which also shows that $S_\gamma(f)$ avoids $-\infty$. By (14) it follows that $S_\gamma(f)$ (and $Q_\gamma(f)$ by Proposition 3.1) is the difference of an l.s.c. convex functional and a quadratic term. With this in mind the continuity statements follows by standard properties of l.s.c. convex functionals (see e.g. Corollary 8.30 [4]). $\square$

The following result is the key result of this section connecting the $Q_\gamma$-transform with l.s.c. convex envelopes.

**Theorem 3.3.** Let $\gamma > 0$ and let $f$ be a $[0, \infty]$-valued functional on a separable Hilbert space $V$. Then $(f + \frac{\gamma}{2} \| \cdot \|^2)^* (y) = S_\gamma(f) \left( \frac{y}{\gamma} + d \right) + \frac{\gamma}{2} \left\| \frac{y}{\gamma} + d \right\|^2 - \frac{\gamma}{2} \| d \|^2$ and

$$\left( f + \frac{\gamma}{2} \| \cdot \|^2 \right)^{*\star} (x) = Q_\gamma(f)(x) + \frac{\gamma}{2} \| x - d \|^2.$$  

In particular, $Q_\gamma(f)(x) + \frac{\gamma}{2} \| x - d \|^2$ is the l.s.c. convex envelope of $f(x) + \frac{\gamma}{2} \| x - d \|^2$ and $0 \leq Q_\gamma(f) \leq f$. 

Proof. We have
\[
(f(\cdot) + \frac{\gamma}{2} \| \cdot - d \|^2)^* (y) = \sup_x \langle x, y \rangle - f(x) - \frac{\gamma}{2} \| x - d \|^2 = \\
= \sup_x -f(x) - \frac{\gamma}{2} \| x - \left( \frac{y}{\gamma} + d \right) \|^2 + \frac{\gamma}{2} \left\| \frac{y}{\gamma} + d \right\|^2 - \frac{\gamma}{2} \| d \|^2
\]
from which the first identity follows. Similarly
\[
(f(\cdot) + \frac{\gamma}{2} \| \cdot - d \|^2)^\ast (x) = \left( S_\gamma(f) \left( \frac{1}{\gamma} + d \right) + \frac{\gamma}{2} \left\| \frac{1}{\gamma} + d \right\|^2 - \frac{\gamma}{2} \| d \|^2 \right)^*(x)
\]
\[
= \sup \langle y, x \rangle - S_\gamma(f) \left( \frac{y}{\gamma} + d \right) - \frac{\gamma}{2} \left\| \frac{y}{\gamma} + d \right\|^2 + \frac{\gamma}{2} \| d \|^2 = \\
= \sup_y -S_\gamma(f) \left( \frac{y}{\gamma} + d \right) - \frac{\gamma}{2} \left\| \frac{y}{\gamma} + d - x \right\|^2 + \frac{\gamma}{2} \| x - d \|^2 = S_\gamma^2(f)(x) + \frac{\gamma}{2} \| x - d \|^2.
\]
The statement about the convex envelope follows by the Fenchel-Moreau theorem and Proposition 3.1 which also gives \( Q_\gamma(f)(x) + \frac{\gamma}{2} \| x - d \|^2 \leq f(x) + \frac{\gamma}{2} \| x - d \|^2 \). This implies the latter part of the inequality \( 0 \leq Q_\gamma(f) \leq f \) whereas the former has already been noticed in Proposition 3.2.

We end this section with some observations about the behavior of \( Q_\gamma(f) \) as a function of \( \gamma \).

**Proposition 3.4.** Let \( f \) be an l.s.c. \([0, \infty]\)-valued functional. Then \( Q_\gamma(f)(x) \) is increasing as a function of \( \gamma \). Moreover
\[
\lim_{\gamma \to \infty} Q_\gamma(f)(x) = f(x) \quad (16)
\]
whereas the limit as \( \gamma \downarrow 0 \) equals a convex minimizer of \( f \) above the l.s.c. convex envelope of \( f \).

We remark that (16) is shown in [29], whereas nothing is said about the case \( \gamma \downarrow 0 \). In fact, \( \lim_{\gamma \to 0} Q_\gamma(f) \) usually equals the l.s.c. convex envelope of \( f \), but this is not necessarily the case in general, which is a surprise at least for the author. To see this, consider \( P = \{ x \in \mathbb{R}^2 : x_1 > 0, x_2 = \sqrt{x_1} \} \), \( Q = \{ x \in \mathbb{R}^2 : x_1 > 0, 0 < x_2 \leq \sqrt{x_1} \} \cup \{ 0 \} \) and \( f = \iota_P \), where \( \iota_P \) is the indicator functional of \( P \). It is easy to see that the l.s.c. convex envelope of \( \iota_P \) equals \( \iota_{cl(Q)} \) (where \( cl \) denotes closure) whereas some thinking reveals that \( \lim_{\gamma \to 0} Q_\gamma(f) = \iota_Q \). However if \( V \) is finite dimensional and \( \lim_{\gamma \to 0} Q_\gamma(f) \) is everywhere finite, then it is automatically continuous (Corollary 8.30 in [4]), and hence it must equal the l.s.c. convex envelope of \( f \).

Proof. If \( \gamma_1 > \gamma_2 \) then \( Q_{\gamma_2}(f)(x) + \frac{\gamma_2}{2} \| x \|^2 \) equals the l.s.c. convex functional \( Q_{\gamma_2}(f)(x) + \frac{\gamma_2}{2} \| x \|^2 \) plus the term \( 2\gamma_2 \| x \|^2 \) so it is l.s.c. and convex. In view of \( Q_{\gamma_2}(f) \leq f \) it also lies below \( f + \frac{\gamma_2}{2} \| x \|^2 \) and so we conclude that
\[
Q_{\gamma_2}(f)(x) + \frac{\gamma_2}{2} \| x \|^2 \leq (f + \frac{\gamma_1}{2} \| x \|^2)^\ast = Q_{\gamma_1}(f)(x) + \frac{\gamma_1}{2} \| x \|^2.
\]
The first claim follows. To see (16) let \( \alpha < f(x) \) be arbitrary. Since \( f \) is l.s.c. the set \( \{ y : f(y) > \alpha \} \) is open and, as \( f \geq 0 \), it follows that for any \( \gamma \) large enough we have \( \alpha - \frac{\gamma}{2} \| \cdot - x \|^2 \leq f \). For such \( \gamma \) we thus have \( \alpha \leq Q_\gamma(f)(x) \leq f(x) \) by (13) and Theorem 3.3 so (16) follows.
Concerning the limit as \( \gamma \downarrow 0 \) set \( g(x) = \lim_{\gamma \downarrow 0} Q_\gamma(f)(x) \) which exist by the first part of this proposition. Since
\[
g(x) = \lim_{\gamma \downarrow 0} Q_\gamma(f)(x) = \lim_{\gamma \downarrow 0} Q_\gamma(f)(x) + \frac{\gamma}{2} \Vert x \Vert^2 = \lim_{\gamma \downarrow 0} \left( f + \frac{\gamma}{2} \Vert \cdot \Vert^2 \right)^**(x) \geq f^{**}
\]
we see that \( g \) is the limit of a decreasing sequence of convex functions, hence it is also convex (Proposition 8.16 [4]), and clearly \( g \leq f \) by Theorem 3.3.

4 Finer Properties of Convex and Quadratic Envelopes

In this section, we prove a result about the structure of l.s.c. convex envelopes which seems relatively unknown. For this we need the concept of weak lower semi-continuity, which is nothing but semi-continuity with respect to the weak topology of the underlying separable Hilbert space \( \mathcal{V} \). We remind the reader that for convex proper functionals there is no difference (Theorem 9.1 [4]) between weakly l.s.c. functionals and standard l.s.c. functionals. Also, if \( \mathcal{V} \) is finite dimensional and the topology is Hausdorff, the two topologies are the same so there is no difference in this case either. However we wish to underline that the difficulty in proving the coming results is present also in the finite-dimensional setting.

We begin with a neat fact concerning weakly l.s.c. convex envelopes which does not seem to have made its way into the modern literature on the subject. It is a reformulation of Arne Brøndsted’s extension of Milman’s theorem [9]. To state it we remind the reader that a functional \( g \) is coercive if and only if its (lower) level sets are bounded (see e.g. Proposition 11.11 [4]). Note that l.s.c. convex envelopes of the type \( Q_\gamma(f)(x) + \frac{\gamma}{2} \Vert x - d \Vert^2 \) (for positive \( f \) always are coercive, by virtue of Proposition 3.2 and the quadratic term. A function \( f \) on \( \mathbb{R} \) is called affine if it is of the form \( f(t) = at + b \) with \( a, b \in \mathbb{R} \).

**Theorem 4.1.** Let \( g \) be a weakly l.s.c. functional on a separable Hilbert space \( \mathcal{V} \) such that \( g^{**} \) is coercive. Given any \( x \in \mathcal{V} \) such that \( g(x) \neq g^{**}(x) \) there exists a unit vector \( \nu \) and \( t_0 > 0 \) such that the function \( h(t) = g^{**}(x_0 + t\nu) \) is affine on \((-t_0, t_0)\).

To prove Theorem 4.1 we recall some concepts from [9]. Given a convex function \( f \) a point \( x \) is called extremal if and only if \( (x, f(x)) \) is extremal for the epigraph of \( f \), denoted \( [f] \). Equivalently, \( x \) is extremal if and only if \( x \in \text{dom} f \) and \( f \) is not affine on any relatively open segment containing \( x \). Moreover \( f_{\text{ext}} \) denotes the functional which equals \( f(x) \) for all extremal points \( x \) and \( \infty \) else. As a consequence of Theorem 1 in [9] we have:

**Theorem 4.2.** Let \( g \) be a weakly l.s.c. functional on a separable Hilbert space \( \mathcal{V} \) such that \( g^{**} \) is coercive, then
\[
[(g^{**})_{\text{ext}}] \subset [g].
\]

**Proof.** In the setting of [9] we let \( E \) be the separable Hilbert space \( \mathcal{V} \) with the weak topology. Since convex functionals are l.s.c. with respect to the weak topology if and only if they are with respect to the norm topology it follows that the l.s.c convex envelope of \( g \) equals the weakly l.s.c. convex envelope. In the notation of Theorem 1 of [9] we can then take \( f = g^{**} \) and the theorem states that \( [f_{\text{ext}}] \subset [g_{cl}] \) where \( g_{cl} \) is the greatest l.s.c. minorant of \( g \). Since \( g \) is assumed to be l.s.c. we have \( g = g_{cl} \) and the desired inclusion follows. It remains to check that the conditions of Theorem 1 are fulfilled, which is that “\( g \) is inf-compact in some direction” (with respect to the weak topology,
referring to the terminology of [39]). For this it suffices to check that $g^{**}$ is inf-compact i.e. that all level sets are compact. The level sets of $g^{**}$ are closed and convex and since $g^{**}$ is assumed coercive they are also bounded. It follows that such level sets are compact in the weak topology and the proof is complete.

Based on this we can now easily prove Theorem 4.1.

**Proof of Theorem 4.1.** Since $g \geq g^{**}$, Theorem 4.2 clearly implies that $g(x) = g^{**}(x)$ for all extremal points $x$ for $g^{**}$. Consequently, if $g(x) = g^{**}(x)$ does not hold, then $x$ is not extremal for $g^{**}$ and the existence of $\nu$ follows by the definition of an extremal point for $g^{**}$.

Next we discuss what the theorem implies about minimizers of $g$ versus $g^{**}$. Denote by $G$ the set of global minimizers of $g$ and by $G^{**}$ the set of global minimizers of $g^{**}$.

**Corollary 4.3.** Let $g$ be a weakly l.s.c. functional on a separable Hilbert space $\mathcal{V}$ such that $g^{**}$ is coercive. Then $G^{**}$ is a closed bounded convex set containing $G$. Letting $G^{**}_{ext}$ denote the extremal points of $G^{**}$ we also have that $G^{**}_{ext} \subset G$. Finally the closed convex hull of $G^{**}_{ext}$ equals $G^{**}$.

**Proof.** The convexity of $G^{**}$ and the inclusion $G \subset G^{**}$ are immediate. The boundedness of $G^{**}$ follows since $g^{**}$ is coercive. Let $x$ be in the closure of $G^{**}$ and let $c$ be the value of the global minimum. Then $g^{**}(x) \leq c$ follows by l.s.c. and the reverse inequality is obvious from the fact that $c$ is a global minimum. It follows that $x \in G^{**}$ and hence $G^{**}$ is closed.

The existence of points in $G^{**}_{ext}$ and the statement concerning the closed convex hull are now immediate consequences of the Krein-Milman theorem and the fact that bounded closed convex sets are weakly compact in separable Hilbert spaces (Theorem 3.33, [39]). It remains to prove that $G^{**}_{ext} \subset G$. Let $x_0 \in G^{**}_{ext}$ suppose $x_0 \notin G$. Then Theorem 4.1 implies the existence of a direction $\nu$ on which $g^{**}$ is constant near $x_0$ contradicting that $x_0$ is an extremal point.

We end by noting that Theorem 4.1 implies that $\gamma$ tunes the maximum negative curvature in the $\mathcal{Q}_\gamma$-transform as discussed in the introduction.

**Corollary 4.4.** Let $f$ be a weakly l.s.c. $[0, \infty]$-valued functional on a separable Hilbert space $\mathcal{V}$. For each $x_0 \in \mathcal{V}$ with $f(x_0) > Q_\gamma(f)(x_0)$ there exists a unit vector $\nu$ such that $Q_\gamma(f)(x_0 + t\nu) = a + bt - \frac{1}{2}t^2$ for $t$ near 0 and some $a, b \in \mathbb{R}$.

**Proof.** Set $g(x) = f(x) + \frac{1}{2}\|x\|^2$. By Theorem 3.3 we have $Q_\gamma(f)(x) + \frac{1}{2}\|x\|^2 = g^{**}(x)$ by which it is immediate that $g^{**}$ is coercive (since $Q_\gamma(f) \geq 0$ by Proposition 3.2). It also follows that $g(x_0) > g^{**}(x_0)$ and hence Theorem 4.1 implies that a unit vector $\nu$ exists such that $t \mapsto Q_\gamma(f)(x_0 + t\nu)$ equals an affine function minus $\frac{1}{2}\|x + t\nu\|^2$ in a neighborhood of $t = 0$.

5 The Quadratic Envelope as a Regularizer

We now let $A : \mathcal{V} \to \mathcal{W}$ be a bounded linear operator, where $\mathcal{V}, \mathcal{W}$ are possibly different (separable) Hilbert spaces, and consider functionals of the type

$$J(x) = f(x) + \frac{1}{2}\|Ax - d\|^2_W, \quad x \in \mathcal{V},$$

(17)

Our aim is to develop strategies to deal with the general problem (17), in the case when $f$ is an $[0, \infty]$-valued functional such that $Q_\gamma(f)$ is computable, and focus on computing (explicit) approximations of the l.s.c convex envelope of $J$. The theory is split in two cases, either we approximate the convex
envelope from below by a convex functional, or we approximate it from above with a non-convex functional having a number of desirable properties, most notably the fact that local minimizers do not change. More precisely, we will study the relationship between the original problem \( \text{(17)} \) and the modified problem

\[
\mathcal{J}_\gamma(x) = Q_\gamma(f)(x) + \frac{1}{2} \|Ax - d\|_W^2, \quad x \in V
\]

under the assumption that \( \gamma I \preceq A^*A \) or \( \gamma I \succ A^*A \) (c.f. [11]-[12] and recall Figure [3]). Note that \( \gamma I \succ A^*A \) if and only if \( \gamma \geq \|A\|^2 \).

### 5.1 Case \( A^*A \succ \gamma I \)

Let \( f \) be a \([0, \infty] \)-valued functional and \( A : V \rightarrow W \) a bounded linear operator. The main result of this section states that \( \mathcal{J}_\gamma \) is a convex minorant of the l.s.c. convex envelope \( \mathcal{J}^{**} \).

**Theorem 5.1.** For \( \gamma > 0 \) such that \( A^*A \succ \gamma I \), \( \mathcal{J}_\gamma \) is convex and \( \mathcal{J}_\gamma \leq \mathcal{J}^{**} \). Moreover, if \( A^*A \succ \gamma I \) then it is strongly convex, in which case it has a unique minimizer. Finally, a minimizer \( \hat{x} \) of \( \mathcal{J}_\gamma \) is a minimizer of \( \mathcal{J} \) whenever \( f(\hat{x}) = Q_\gamma(f)(\hat{x}) \).

**Proof.** Upon expanding \( \|Ax - d\|^2 = \|Ax\|^2 - 2 \langle Ax, d \rangle + \|d\|^2 \) and noting that the latter two terms are affine linear, it is easily seen that it suffices to prove the first part of the statement for \( d = 0 \). That \( \mathcal{J}_\gamma \) is l.s.c. and that \( \mathcal{J}_\gamma \leq \mathcal{J} \) follows immediately by Theorem [3.3] and thus \( \mathcal{J}_\gamma \leq \mathcal{J}^{**} \) follows immediately upon showing that \( \mathcal{J}_\gamma \) is convex. Define \( \langle x, y \rangle_W = \langle Ax, Ay \rangle_W - \gamma \langle x, y \rangle_V \) and note that this is a semi-inner product, as long as \( A^*A \succ \gamma I \), which is an inner product if the inequality is strict. In either case \( \|x\|^2_W := \langle x, x \rangle_W \) is convex. It follows that

\[
Q_\gamma(f)(x) + \frac{1}{2} \|Ax\|^2_W = \left( Q_\gamma(f)(x) + \frac{\gamma}{2} \|x\|^2_V \right) + \frac{1}{2} \|x\|^2_W
\]

which by Theorem [3.3] implies that \( \mathcal{J}_\gamma \) equals the l.s.c. convex envelope of \( f(x) + \frac{\gamma}{2} \|x\|^2_V \) plus the term \( \frac{1}{2} \|x\|^2_W \). We conclude that \( \mathcal{J}_\gamma \) is a convex functional which is strongly convex when \( A^*A \succ \gamma I \). In the latter case the existence of a unique minimizer follows by Corollary 11.15 in [H] (supercoercivity of \( \mathcal{J}_\gamma \) is obvious by the term \( \frac{1}{2} \|x\|^2_W \)). Finally let \( d \) be fixed and let \( \hat{x} \) be a minimizer of \( \mathcal{J}_\gamma \). Suppose that \( f(\hat{x}) = Q_\gamma(f)(\hat{x}) \) and let \( y \in V \) be arbitrary. Then \( \mathcal{J}(y) \geq \mathcal{J}_\gamma(y) \geq \mathcal{J}_\gamma(\hat{x}) = \mathcal{J}(\hat{x}) \) showing that \( \hat{x} \) is a global minimizer of \( \mathcal{J} \).

### 5.2 Case \( A^*A \preceq \gamma I \)

Let \( f \) be a \([0, \infty] \)-valued functional and \( A : V \rightarrow W \) a bounded linear operator. Again we are interested in the relationship between \( \mathcal{J} \) and \( \mathcal{J}_\gamma \) defined in \([17]\) and \([18]\) respectively. The main result of this section is that \( \mathcal{J}_\gamma \) does not move minima for \( \gamma \) in the stated range, but we begin by noting the following inequalities, the first one being reverse of the one proved in Theorem [5.1].

**Proposition 5.2.** For \( \gamma \) such that \( \|A\|^2 \leq \gamma \) we have \( \mathcal{J}^{**} \leq \mathcal{J}_\gamma \leq \mathcal{J} \).

**Proof.** The right inequality is immediate since \( Q_\gamma(f) \leq f \) by Theorem [3.3] As in Theorem [5.1] we moreover see that it suffices to prove the left inequality for \( d = 0 \). To this end set \( h(x) = \mathcal{J}^{**}(x) - \frac{\gamma}{2} \|Ax\|^2 \). Since \( \mathcal{J}^{**} \leq f + \frac{1}{2} \|Ax\|^2 \) we have \( h \leq f \) and moreover

\[
h(x) + \frac{\gamma}{2} \|x\|^2 = \mathcal{J}^{**} + \left( \frac{\gamma}{2} \|x\|^2 - \frac{1}{2} \|Ax\|^2 \right).
\]
Theorem 5.3. Let $\gamma > 0$ and $x$ be a minimizer of $J(x)$. Then $\gamma^2 \|x\|^2 \leq (f + \frac{\gamma}{2} \|\cdot\|^2)''(x) = Q_\gamma(f)(x) + \frac{\gamma}{2} \|x\|^2$.

The right hand side is convex and l.s.c. by which we conclude that

$$h(x) + \frac{\gamma}{2} \|x\|^2 \leq (f + \frac{\gamma}{2} \|\cdot\|^2)''(x) = Q_\gamma(f)(x) + \frac{\gamma}{2} \|x\|^2$$

(19) that there exists a unit vector $\nu$ such that $J(x) \leq Q_\gamma(f)(x) + \frac{\gamma}{2} \|Ax\|^2$ which is the desired inequality (for $d = 0$).

We now come to the main theorem of this section, inspired by Theorems 4.5 and 4.8 in [30].

We say that $x$ is a local minimizer (resp. strict local minimizer) of $J$, then it is also a local minimizer (resp. strict local minimizer) of $J$, and $J(x) = J(x)$. In addition the global minimizers coincide.

Proof. Let $x$ be a local minimizer of $J$. If $Q_\gamma(f)(x) = f(x)$ does not hold then Corollary 4.4 implies that there exists a unit vector $\nu$ such that

$$\frac{d^2}{dt^2} J_\gamma(x + t\nu)(0) = \frac{d^2}{dt^2} \left( Q_\gamma(f)(x + t\nu) + \frac{\gamma}{2} \|Ax + t\nu\|^2 \right)(0) = \|A\nu\|^2 - \gamma < 0. \quad (19)$$

We thus conclude that $Q_\gamma(f)(x) = f(x)$ holds which immediately gives that $J_\gamma(x) = J(x)$.

In view of Proposition 5.2 it follows that $x$ is a local minimizer also for $J_\gamma$. The same argument applies to strict local minimizers.

We now prove that the global minimizers coincide. Note that global minimizers of $J_\gamma$ are global minimizers of $J$ in view of Proposition 5.2 and the fact that $J(x) = J(x)$ for all global minimizers $x$. From this we also see that the global minimum of $J_\gamma$ and $J_\gamma$ coincide, let us denote this value by $c$. Conversely suppose that $x$ is a global minimizer of $J_\gamma$ (i.e. $J_\gamma(x) = c$). Then it is a local minimizer of $J$ by the first part, which automatically is global for $J$ since we otherwise would have $J(y) < c$ for some other value $y$. The proof is complete.

The situation when $\gamma = \|A\|^2$ is a bit more involved so we content ourselves with the following statement concerning the global minimizers.

Theorem 5.4. Set $\gamma = \|A\|^2$, let $G$ be the global minimizers of $J$ and $G_\gamma$ the global minimizers of $J_\gamma$. Then $G \subset G_\gamma$ and each connected component of $G_\gamma$ contains points of $G$.

Proof. The statement $G \subset G_\gamma$ follows as in the above proof, as well as the fact that the global minimum of $J$ and $J_\gamma$ coincide; we denote it by $c$. If $x \in G_\gamma$ and $J(x) > c$ then it follows by (19) that there exists a unit vector $\nu$ such that $\frac{d^2}{dt^2} J_\gamma(x + t\nu) \leq 0$ in a neighborhood of $t = 0$. Strict inequality contradicts the assumption of global minima, so we deduce that $\gamma \|\nu\|^2 = \|A\nu\|^2$. Introducing the semi-norm $\|x\|_U = \gamma \|x\|_V^2 - \|Ax\|^2_W$, this means that $\|\nu\|_U = 0$, i.e. that $\nu$ lies in the kernel of the semi-norm $\|\cdot\|_U$ (which is a linear subspace by convexity of the semi-norm). Let $P$ be the affine hyperplane $P = x + \ker \|\cdot\|_U$ and set $S = P \cap G_\gamma$. For $y \in \ker \|\cdot\|_U$ we have

$$J(x + y) = \left( Q_\gamma(f)(x + y) + \frac{\gamma}{2} \|x + y\|_V^2 \right) - \frac{1}{2} \|x\|_U^2 - (Ax + y, d)_W + \frac{1}{2} \|d\|_W^2, \quad (20)$$

and

$$J(x) = \left( Q_\gamma(f)(x) + \frac{\gamma}{2} \|x\|_V^2 \right) - \frac{1}{2} \|x\|_U^2 - (Ax, d)_W + \frac{1}{2} \|d\|_W^2.$$
so Theorem 3.3 implies that $\mathcal{J}_\gamma$ is convex on $P$. In particular $S$ is convex. Since $\mathcal{J}_\gamma$ is l.s.c. it is also closed. Moreover $S$ is bounded due to the quadratic term $\|x + y\|_V^2$ in (20). $S$ is therefore weakly closed and hence it equals the closed convex hull of its extremal points by the Krein-Milman theorem. If $x$ now is one of these extremal points then we can argue as in the beginning of this proof and conclude that $\mathcal{J}_\gamma(x) = \mathcal{J}(x)$, since the existence of a $\nu$ with the properties stated initially would contradict that $x$ is an extremal point of $S$.  

6 The $S$-Transform and Semi-Algebraicity

We briefly treat semi-algebraicity of $Q_\gamma(f)$ since it was shown in [3] that this is a necessary condition for the forward backward splitting method to converge in the non-convex setting. We remind the reader that a function on a finite dimensional space is semi-algebraic if its graph is a semi-algebraic set [6].

Theorem 6.1. If $\mathcal{V}$ is finite dimensional and $f$ is semi-algebraic then so is $S_\gamma(f)$ and $Q_\gamma(f)$.

Proof. We assume for simplicity that $\gamma = 1$. It is a consequence of the Tarski-Seidenberg theorem that the set of semi-algebraic functions is closed under addition (see e.g. Prop. 2.2.6 in [6]) and similarly one can prove that the epigraph of a semi-algebraic function is a semi-algebraic set. If $f$ is semi-algebraic on $\mathbb{R}^n$ it follows that $g(x, y) = \langle x, y \rangle - (f(x) + \frac{1}{2} \|x\|^2)$ is semi-algebraic on $\mathbb{R}^{2n}$ and by the argument following Theorem 2.2 in [3] it follows that the Legendre transform of $f + \frac{1}{2} \|x\|^2$ is semi-algebraic. The first result now follows since this function minus $\frac{1}{2} \|y\|^2$ equals $S_\gamma(f)(y)$ by (14), and the second is immediate by Proposition 3.1.

7 Related Works

The operations $S_\gamma(f)$ and $Q_\gamma(f)$ were introduced around 1970 in greater generality by J-J. Moreau [26] and (seemingly independently) E-A. Weiss [33], and were further studied around 1990 by R. Poliquin [27] with a focus on smoothness properties. Variations of Propositions 3.1 and 3.2 date back to these early articles, and are also found e.g. in Rockafellar-Wets [29] Section 11.L. The transforms $S_\gamma$ and $Q_\gamma$ go under names like “Φ-conjugate”/”proximal transform” and “Φ-biconjugate”/”Φ-convex envelope”, and arise by the concrete choice $\Phi(x, y) = q_\gamma(x, y) = -\gamma \frac{1}{2} \|x - y\|^2$. Following Rockafellar-Wets [29] $Q_\gamma(f)$ should be called “proximal hull” or “$q_\gamma$-envelope”. We believe that the “quadratic envelope”, which is closer to the latter, is more suggestive. Functions that satisfy $Q_\gamma(f) = f$ has been called e.g. $\gamma^{-1}$-proximal or quadratically convex.

However they are called, it seems that the connection with convex envelopes a la Theorem 3.3 has not been investigated, which is the main novelty of this publication along with the structural result Corollary 4.4 and its applications to regularization in Section 5. Apart from the already mentioned works by Aubert, Blanc-Feraud, Soubies and Larsson, Olsson we have not found any similar result in the literature. The fairly recent survey paper [23] is about the closely related concept of computing Fenchel conjugates, and also mentions proximal hulls, yet it has no overlap with the present paper despite citing 262 other papers. It primarily deals with numeric computation of convex envelopes in cases when symbolic formulas are not available, and as such it is an interesting alternative to the methods developed here. The same goes for the papers [24] and [7]. The importance of computing convex envelopes is stressed in [25] where techniques for computing convex envelopes of so called
“convex polyhedral” functions are developed. Convex approximations from below are considered in [8] which should be compared with the results in Section 5. An alternative to approximating the convex envelope is to numerically try to compute the proximal operator of the original functional directly, which is pursued in [19]. The papers [2, 31] deal with Lasry-Lions approximants in Hilbert space but do not make the connection with the convex envelopes. For parameters $s < t$ the Lasry-Lions approximation of $f$ is defined by
\[
S_{1/s}S_{1/t}(f)(x) = -\left(\inf_y \left(\inf_w f(w) + \frac{1}{2t} \|w - y\|^2\right) + \frac{1}{2s} \|x - y\|^2\right) = \sup_y \left(\inf_w f(w) + \frac{1}{2t} \|w - y\|^2\right) - \frac{1}{2s} \|x - y\|^2
\]
which for $s = t$ gives $Q_{s-1}$. This regularization is also studied in Section 6 of the more recent publication [31] (with the notation $C(1)f$), mainly with focus on differentiability-results. It is also closely connected to the more general “proximal average”, see e.g. [5, 18]. However the proximal average has been used mainly for modification of convex functions whereas $Q_{\gamma}(f) = f$ for any l.s.c. convex function.

8 Conclusions

We have provided theory for computing l.s.c. convex envelopes of certain functionals and shown a connection with quadratic envelopes (a.k.a. proximal hulls), which was then used to regularize more intricate problems. We showed that for sufficiently small values of the parameter $\gamma$, this yields convex functionals below the original functional, which coincide with the original functional on a large part of the underlying Hilbert space. For $\gamma$ sufficiently large on the other hand we lose convexity but gain the desirable feature that the modified functional has the same global minimizers as the original one, and fewer local ones. This in turn was based on results regarding the structure of l.s.c. convex envelopes. The results are inspired from prior work by Carl Olsson and Viktor Larsson as well as Emmanuel Soubies, Laure Blanc-Féraud and Gilles Aubert.

Particular cases of these ideas have already been applied to compressed sensing, imaging, signal processing and frequency estimation. Currently we are working on more concrete results regarding low rank approximation, improvements of frequency estimation techniques, as well as an application to the classical phase retrieval problem. We hope that other researchers will try these methods on their problems and find that the method is a valuable tool. To aid with this task an expanded version of this article is available on arXiv [13] with many more examples and useful details.

References

[1] Fredrik Andersson, Marcus Carlsson, and Carl Olsson. Convex envelopes for fixed rank approximation. Optimization Letters, 11(8):1783–1795, 2017.

[2] H Attouch and D Azé. Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry-Lions method. In Annales de l'IHP Analyse non linéaire, volume 10, pages 289–312, 1993.
[3] Hedy Attouch, Jérôme Bolte, and Benar Fux Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. *Mathematical Programming*, 137(1-2):91–129, 2013.

[4] Heinz H Bauschke and Patrick L Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer Science & Business Media, 2011.

[5] Heinz H Bauschke, Rafal Goebel, Yves Lucet, and Xianfu Wang. The proximal average: basic theory. *SIAM Journal on Optimization*, 19(2):766–785, 2008.

[6] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.

[7] Jonathan M Borwein and Chris H Hamilton. Symbolic Fenchel conjugation. *Mathematical Programming*, 116(1-2):17–35, 2009.

[8] Bernard Brighi and Michel Chipot. Approximated convex envelope of a function. *SIAM journal on numerical analysis*, 31(1):128–148, 1994.

[9] Arne Brøndsted. Milman’s theorem for convex functions. *Mathematica Scandinavica*, 19:5–10, 1966.

[10] E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

[11] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.

[12] Emmanuel J Candes and Terence Tao. Decoding by linear programming. *IEEE transactions on information theory*, 51(12):4203–4215, 2005.

[13] Marcus Carlsson. On convexification/optimization of functionals including an l2-misfit term. *arXiv preprint arXiv:1609.09378*, 2016.

[14] Marcus Carlsson, Daniele Gerosa, and Carl Olsson. A un-biased approach to compressed sensing. *arXiv preprint*, 2018.

[15] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.

[16] David L Donoho. For most large underdetermined systems of linear equations the minimal $\ell^1$-norm solution is also the sparsest solution. *Communications on pure and applied mathematics*, 59(6):797–829, 2006.

[17] Maryam Fazel. *Matrix rank minimization with applications*. PhD thesis, Stanford University, 2002.

[18] Waren L Hare. A proximal average for nonconvex functions: a proximal stability perspective. *SIAM Journal on Optimization*, 20(2):650–666, 2009.
[19] Warren Hare and Claudia Sagastizábal. Computing proximal points of nonconvex functions. Mathematical Programming, 116(1-2):221–258, 2009.

[20] Viktor Larsson and Carl Olsson. Convex low rank approximation. International Journal of Computer Vision, pages 1–21, 2016.

[21] Jean-Michel Lasry and Pierre-Louis Lions. A remark on regularization in Hilbert spaces. Israel Journal of Mathematics, 55(3):257–266, 1986.

[22] Yves Lucet. The Legendre-Fenchel Transform and the Convex Hull of a Function: Fast Computational Algorithms, Second-Order Smoothness and Analysis. PhD-thesis; https://people.ok.ubc.ca/ylucet/thesis/1997-PhD Yves lucet.pdf, 1997.

[23] Yves Lucet. What shape is your conjugate? A survey of computational convex analysis and its applications. SIAM review, 52(3):505–542, 2010.

[24] Garth P McCormick. Computability of global solutions to factorable nonconvex programs: Part I Convex underestimating problems. Mathematical programming, 10(1):147–175, 1976.

[25] Clifford A Meyer and Christodoulos A Floudas. Convex envelopes for edge-concave functions. Mathematical programming, 103(2):207–224, 2005.

[26] Jean-Jacques Moreau. Inf-convolutions, sous-additive, convexite des fonctions numeriques. Journal de mathématiques pures et appliquées, 49:109–154, 1970.

[27] René Poliquin. Subgradient monotonicity and convex functions. Nonlinear Analysis: Theory, Methods & Applications, 14(4):305–317, 1990.

[28] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev., 52(3):471–501, August 2010.

[29] R Tyrrell Rockafellar and Roger J-B Wets. Variational analysis, volume 317. Springer, 2009.

[30] Emmanuel Soubies, Laure Blanc-Féraud, and Gilles Aubert. A continuous exact $\ell_0$ penalty (CEL0) for least squares regularized problem. SIAM Journal on Imaging Sciences, 8(3):1607–1639, 2015.

[31] Thomas Strömberg. On regularization in banach spaces. Arkiv för Matematik, 34(2):383–406, 1996.

[32] Paul Tseng. Approximation accuracy, gradient methods, and error bound for structured convex optimization. Mathematical Programming, 125(2):263–295, 2010.

[33] Ernst-August Weiss. Konjugierte funktionen. Archiv der Mathematik, 20(5):538–545, 1969.