Abstract

We show that inverse problems with a truncated quadratic regularization are NP-hard in general to solve, or even approximate up to an additive error. This stands in contrast to the case corresponding to a finite-dimensional approximation to the Mumford-Shah functional, where the operator involved is the identity and for which polynomial-time solutions are known. Consequently, we confirm the infeasibility of any natural extension of the Mumford-Shah functional to general inverse problems. A connection between truncated quadratic minimization and sparsity-constrained minimization is also discussed.

Keywords: inverse problems, Mumford-Shah functional, truncated quadratic regularization, sparse recovery, NP-hard, thresholding, SUBSET-SUM

1 Introduction

Consider a discrete signal $x \in \mathbb{R}^N$ sampled from a piecewise smooth signal and revealed through measurements $y = Ax + e$, where $e \in \mathbb{R}^m$ is observation noise and $A : \mathbb{R}^N \rightarrow \mathbb{R}^m$ is a known linear operator identified with an $m \times N$ real matrix (representing, for instance, a blurring or partial obscuring of the data). Consider the truncated quadratic minimization problem,

$$\hat{x} = \arg\min_{x \in \mathbb{R}^N} J(x),$$

$$J(x) = ||Ax - y||_2^2 + \sum_{j=1}^{N-1} Q(x_{j+1} - x_j),$$

with truncated quadratic penalty term $Q(u) = \alpha \min\{u^2, \beta\}$ parametrized by $\alpha, \beta > 0$. Since its introduction in 1984 by Geman and Geman in the context of image restoration [2,6,8], this problem has been the subject of considerable theoretical and practical interest, finding applications ranging from visual analysis to crack detection in fracture mechanics [9,10]. The choice of regularization is motivated as follows: $Q$ desires to smooth small differences $|x_{j+1} - x_j| \leq \sqrt{\beta}$ where it acts quadratically, but suspends smoothing over larger differences.

*B. Alexeev is with the Department of Mathematics at Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544 USA e-mail: balexeev@math.princeton.edu.
†R. Ward is with the Department of Mathematics at the Courant Institute of Mathematical Sciences, New York University, 251 Mercer St, New York, NY 10012 USA e-mail: rward@cims.nyu.edu.
From a statistical point of view, the quadratic data-fidelity term $\|Ax - y\|^2_2$ can be viewed as a log-likelihood of the data under the hypothesis that $e$ is Gaussian random noise, while the truncated quadratic regularization term corresponds to the energy of a piecewise Gaussian Markov random field \[1\,2\,6\,7].

The truncated quadratic minimization problem is non-smooth and highly non-convex. However, several characterizations of the minimizers have been unveiled \[5\,8\]. It is known for instance that minimizers exist and satisfy a “gap” property \[3\,8\]: the magnitude of successive differences of such solutions are either smaller than a first threshold or larger than a second, strictly larger threshold. These thresholds are independent of the observed data $y = Ax + e$ and depend only on the regularization parameters $\alpha$ and $\beta$. This dependence is explicit, so that a priori information about the thresholds can be incorporated into choice of regularization parameters.

When $A$ is the $N \times N$ identity matrix, the truncated quadratic objective function can be viewed as a discretization of the Mumford-Shah functional \[1\], which motivated the variational approach for edge detection and image segmentation with its introduction in 1988. When $A$ is the identity matrix as such, the truncated quadratic minimization problem can be solved in polynomial-time using dynamic programming \[4\]. However, for general $m \times N$ matrices $A$, existing algorithms for minimizing the functional \[1\] guarantee convergence to local minimizers at best \[5\].

In this paper, we show that the truncated quadratic minimization problem is NP-hard in general, certifying that the present convergence guarantees are the best one could hope for. Consequently, the Mumford-Shah functional \[2\] cannot be tractably extended to general inverse problems.

## 2 Truncated quadratic minimization reformulated

It will be helpful to recast the truncated quadratic minimization problem in terms of the discrete differences $u_j = x_{j+1} - x_j$, effectively decoupling the action of the regularization term $Q$. We may express this change of variables in matrix notation

\[\hat{x} = \arg\min_{x \in \text{SBV}[0,1]} \mathcal{F}(x),\]

\[\mathcal{F}(x) = \int_{[0,1]\setminus S_x} \left( (x - y)^2 + \alpha \|\nabla x\|^2_2 \right)dx + \alpha\beta |S_x|,

over the space $\text{SBV}$ of bounded variation functions on $[0,1]$ with vanishing Cantor part. Note that SBV functions have a well-defined discontinuity set $S_x$ of finite cardinality $|S_x|$; see \[5\] for more details.
as \( u = Dx \), with \( D : \mathbb{R}^N \rightarrow \mathbb{R}^{N-1} \) the discrete difference matrix,

\[
D = \begin{pmatrix}
-1 & 1 & 0 & \ldots & 0 \\
0 & -1 & 1 & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & \ldots & \ldots & -1 & 1
\end{pmatrix}.
\]

The null space of \( D \), which we denote in the following by \( \mathcal{N}(D) \), is simply the one-dimensional subspace of constant vectors in \( \mathbb{R}^N \). The orthogonal projection of a vector \( x \in \mathbb{R}^N \) onto this subspace is the constant vector \( c \) whose entries coincide with the mean value \( \frac{1}{N} \sum_{j=1}^{N} x_j \) of \( x \), while its projection onto the orthogonal complement of \( \mathcal{N}(D) \) is given by the least squares solution \( D^\dagger Dx \), where \( D^\dagger \) is the pseudo-inverse matrix of \( D \) in the Moore-Penrose sense. These observations yield the orthogonal decomposition \( x = D^\dagger Dx + c \), or, incorporating the substitution \( u = Dx \), the decomposition \( x = D^\dagger u + c \).

Minimization of \( J \), recast in terms of the variables \( u \) and \( c \), becomes

\[
(\hat{c}, \hat{u}) = \arg\min_{c \in \mathcal{N}(D), u \in \mathbb{R}^{N-1}} J(c, u),
\]

\[
J(c, u) = \|AD^\dagger u + Ac - y\|_2^2 + \sum_{j=1}^{N-1} Q(u_j),
\]

where the primal minimizer \( \hat{x} \) and \((c, u)\)-minimizer \( (\hat{c}, \hat{u}) \) are interchangeable according to \( \hat{x} = D^\dagger \hat{u} + \hat{c} \).

If the null space of \( A \) contains the constant vectors, such as if \( A = TD \) for an \( m \times (N-1) \) matrix \( T \), the minimization problem (2) reduces to a function of \( u \) only,

\[
\hat{u} = \arg\min_{u \in \mathbb{R}^{N-1}} \|AD^\dagger u - y\|_2^2 + \sum_{j=1}^{N-1} Q(u_j).
\]

Making the substitution \( A = TD \) and using that \( DD^\dagger = I \), we see in particular that any optimization problem of the form \( \hat{u} = \arg\min_{u \in \mathbb{R}^{N-1}} \|Tu - y\|_2^2 + \sum_{j=1}^{N-1} Q(u_j) \) can be identified with an instance of a truncated quadratic minimization problem (1).

To summarize,

**Lemma 1.** Let \( T : \mathbb{R}^{N-1} \rightarrow \mathbb{R}^m \) be a linear operator identified with a matrix of \( \mathbb{R}^{m \times (N-1)} \). The minimization problem

\[
\hat{u} = \arg\min_{u \in \mathbb{R}^{N-1}} \|Tu - y\|_2^2 + \sum_{j=1}^{N-1} Q(u_j)
\]

corresponds to a truncated quadratic minimization problem

\[
\hat{x} = \arg\min_{x \in \mathbb{R}^N} \|TDx - y\|_2^2 + \sum_{j=1}^{N-1} Q(x_{j+1} - x_j)
\]

in the sense that \( \hat{x} = D^\dagger \hat{u} \) is a minimizer for (1) if \( \hat{u} \) minimizes (3), while \( \hat{u} = D\hat{x} \) is a minimizer for (3) if \( \hat{x} \) is a minimizer for (1).
3 Reduction to SUBSET-SUM

Recall that the complexity class NP consists of all problems whose solution can be verified in polynomial time given a certificate for the answer. For example, the problem SUBSET-SUM is to determine, given nonzero integers $a_1, \ldots, a_k$ and $C$, whether or not there exists a subset $S$ of $\{1, \ldots, k\}$ such that $\sum_{i \in S} a_i = C$. This problem is in NP because given any particular subset $S$, we can easily check whether or not its corresponding sum is zero.

Further recall that a polynomial-time many-one reduction from a problem $A$ to a problem $B$ is an algorithm that transforms an instance of $A$ to an instance of $B$ with the same answer in time polynomial with respect to the number of bits used to represent the instance of $A$. Intuitively, this captures the notion that $A$ is no harder than $B$, up to polynomial factors, and accordingly one may write $A \leq B$. Finally, a problem $B$ is called NP-hard if every problem in NP is reducible to $B$. (Note that $A \leq B$ and $B \leq C$ imply $A \leq C$, so if an NP-hard problem $B$ reduces to a problem $C$, then $C$ is NP-hard as well.) NP-hard problems can not be solved in polynomial time unless P=NP.

In order to prove our NP-hardness result, we show that the known NP-hard problem SUBSET-SUM admits a polynomial-time reduction to an instance of the truncated quadratic minimization problem. Moreover, we show that any algorithm that could efficiently approximate this minimum (to within an additive error) could solve SUBSET-SUM efficiently as well; that is, the search for even an approximate solution to a truncated quadratic minimization problem is NP-hard as well.

**Theorem 2.** Let $a_1, \ldots, a_k$ and $C$ be given nonzero integers. Then there exists a subset $S$ of $\{1, \ldots, k\}$ such that $\sum_{i \in S} a_i = C$ if and only if $\min_{x \in \mathbb{R}^{2k}} f(x) \leq k$, where

$$f(x) = \left( C - \sum_{i=1}^{k} a_i x_i \right)^2 + P \cdot \sum_{i=1}^{k} (1 - x_i - x_{i+k})^2$$

$$\quad + \sum_{i=1}^{2k} \min \left( 1, \frac{x_i^2}{\varepsilon^2} \right),$$

with $0 < \varepsilon \leq \frac{1}{4(\sum_i |a_i|)}$ and $P \geq \frac{2k}{\varepsilon^2}$. Moreover, this minimum is never strictly between $k$ and $k + \frac{1}{4}$.

**Proof.** Call a subset $S \subseteq \{1, \ldots, k\}$ good if $\sum_{i \in S} a_i = C$. If a good subset $S$ exists, we may set

$$x_i = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{if } i \notin S, \end{cases} \quad \text{and} \quad x_{i+k} = \begin{cases} 0 & \text{if } i \in S, \\ 1 & \text{if } i \notin S, \end{cases}$$

for $1 \leq i \leq k$. Then $f(x) = k$ because the first and second terms vanish, and there are exactly $k$ nonzero $x_i$s. Therefore, if a good subset exists, the minimum is at most $k$. 


Suppose no good subset exists, yet there exists $x$ such that $f(x) < k + \frac{1}{4}$. Consider the $k$ pairs of coordinates $x_i, x_{i+k}$ for $1 \leq i \leq k$. If both $|x_i|$ and $|x_{i+k}|$ were less than $\varepsilon$, a single summand in the second term already exceeds $2k$, as $P \cdot (1 - x_i - x_{i+k})^2 \geq P \cdot (1 - 2\varepsilon)^2 \geq 2k$. Therefore, at least one of $|x_i|$ and $|x_{i+k}|$ exceeds $\varepsilon$, and the third term is already at least $k$. If more than one of $|x_i|$ and $|x_{i+k}|$ exceeded $\varepsilon$, then the third term would be at least $k + 1$, so exactly one of the coordinates in each pair exceeds $\varepsilon$ in absolute value. If $|x_i| \leq \varepsilon$, then $|x_{i+k} - 1| \leq 2\varepsilon$ as otherwise $P \cdot (1 - x_i - x_{i+k})^2 \geq P \cdot (\varepsilon)^2 \geq 2k$; the symmetric holds if $|x_{i+k}| \leq \varepsilon$, so all of the $x_i$ are within $2\varepsilon$ of either 0 or 1.

Let $\bar{x}_i$ be the closer of 0 and 1 to $x_i$. Then because no good subset exists and $C$ and the $a_i$ are integers, $|C - \sum_{i=1}^{k} a_i \bar{x}_i| \geq 1$. It follows that the first term
\[
\left(C - \sum_{i=1}^{k} a_i x_i\right)^2 \geq (1 - \sum 2\varepsilon a_i)^2 \geq \left(\frac{1}{4}\right)^2 = \frac{1}{4}.
\]
But the third term was already at least $k$, so this is a contradiction. Therefore, if no good subset exists, we must have $\min_x f(x) \geq k + \frac{1}{4}$. \hfill \Box

**Corollary 3.** Solving the truncated quadratic regularization problem, even to within an additive error, is NP-hard.

**Proof.** In light of Lemma 1, minimization of the function $f$ is a truncated quadratic minimization problem, with $m = k + 1$ and $N = 2k$. Therefore, we have reduced the known NP-hard problem SUBSET-SUM to a truncated quadratic minimization problem. It remains to verify that this reduction is polynomial-time. To see this, note that Theorem 2 ensures that the minimum of $f$ is either at most $k$ or at least $k + \frac{1}{4}$; thus, we only need to approximate each of the polynomially-many entries in the matrices and vectors in $f$ to within a number of bits that is polynomial compared to the number of bits needed to represent $\sum_i |a_i|$. \hfill \Box

### 4 Connection to sparse recovery

The only properties of the quadratic regularization term $\min\{1, x_i^2/\varepsilon^2\}$ needed for Theorem 2 were that it be bounded between 0 and 1, equal to 0 if $x_i = 0$, and equal to 1 if $|x_i| \geq \varepsilon$. Indeed, Theorem 2 holds for any regularization term satisfying these properties; for example, one could consider hard thresholding,

\[
|x|_0 = \begin{cases} 
0 & x = 0, \\
1 & x \neq 0,
\end{cases}
\]

which generates the $\ell_0$ “counting norm” $\|x\|_0 = \sum_{j=1}^{N} |x_j|_0$. We then reprove the known result that the $\ell_0$-regularized optimization problem,

\[
\hat{u} = \arg\min_{u \in \mathbb{R}^N} \|Tu - y\|_2^2 + \gamma \|u\|_0,
\]

is NP-hard in general. This functional is of considerable interest in the emerging area of sparse recovery, as it is guaranteed to produce sparse solutions for sufficiently large $\gamma$ and over a certain class of matrices $\mathcal{E}$. In this light, the truncated quadratic

\[\text{2See [3] for matrix constructions that admit polynomial-time recovery algorithms for the $\ell_0$-regularized optimization problem. All constructions at present involve an element of randomness,}\]
minimization problem may be interpreted as a relaxation of the $\ell_0$-regularized optimization problem, and our main result as showing that even such relaxations of the $\ell_0$-regularized functional are NP-hard.

References

[1] J. E. Besag. Digital image processing: Toward Bayesian image analysis. *J. Appl. Statist.*, 16(3):395–407, 1989.

[2] A. Blake and A. Zisserman. *Visual reconstruction*. MIT Press Series in Artificial Intelligence. MIT Press, Cambridge, MA, 1987.

[3] E. J. Candes and M. B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[4] A. Chambolle. Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations. *SIAM J. Appl. Math.*, 55(3):827–863, 1995.

[5] M. Fornasier and R. Ward. Free discontinuity problems meet iterative thresholding. Submitted, Foundations of Computational Mathematics.

[6] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(6):721–741, November 1984.

[7] F.-C. Jeng and J. W. Woods. Simulated annealing in compound Gaussian random fields. *IEEE Trans. Inform. Theory*, 36(1):94–107, 1990.

[8] M. Nikolova. Thresholding implied by truncated quadratic regularization. *IEEE Trans. Signal Process.*, 48(12):3437–3450, 2000.

[9] L. Rondi. A variational approach to the reconstruction of cracks by boundary measurements. *J. Math. Pures Appl. (9)*, 87(3):324–342, 2007.

[10] L. Rondi. Reconstruction in the inverse crack problem by variational methods. *European J. Appl. Math.*, 19(6):635–660, 2008.

and a complete characterization of such matrices forms the core of the area known as compressed sensing.