



# Outlier Removal Power of the L1-Norm Super-Resolution

Yann Traonmilin, Saïd Ladjal, Andrés Almansa

## ► To cite this version:

Yann Traonmilin, Saïd Ladjal, Andrés Almansa. Outlier Removal Power of the L1-Norm Super-Resolution. 4th International Conference, SSVM 2013,, Jun 2013, Austria. pp.198-209, 10.1007/978-3-642-38267-3\_17 . hal-00803695

**HAL Id: hal-00803695**

**<https://hal.science/hal-00803695>**

Submitted on 22 Mar 2013

**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.

# Outlier Removal Power of the L1-Norm Super-Resolution

Yann Traonmilin, Saïd Ladjal, and Andrés Almansa

Telecom Paristech LTCI

{yann.traonmilin,ladjal,andres.almansa}@telecom-paristech.fr

**Abstract.** Super-resolution combines several low resolution images having different sampling into a high resolution image. L1-norm data fit minimization has been proposed to solve this problem in a robust way. The outlier rejection capability of this methods has been shown experimentally for super-resolution. However, existing approaches add a regularization term to perform the minimization while it may not be necessary. In this paper, we recall the link between robustness to outliers and the sparse recovery framework. We use a slightly weaker Null Space Property to characterize this capability. Then, we apply these results to super resolution and show both theoretically and experimentally that we can quantify the robustness to outliers with respect to the number of images.

**Keywords:** super-resolution, interpolation, L1-norm

## 1 Introduction

### 1.1 Problem Statement and State of the Art

The objective of super-resolution (SR) is to recover a high resolution (HR) image from several low resolution (LR) images. SR relies on the different sampling caused by motion between LR images acquisition. Several surveys of the subject exist in the literature [1–3]. The variational approach to super-resolution leads to the general form of a regularized minimization of the data-fit functional.

Most of the time, this data fit functional is an  $L^p$ -norm fit to the observed data. The  $L^2$ -norm (least squares data fit) has been the most frequent choice because of the optimality properties of the solution when data is contaminated by random noise [4]. Methods for least squares minimization such as the conjugate gradient are also well-known and efficient. More recently,  $L^1$ -norm minimization has been proposed to remove outliers from images [5] and as a robust way to perform super-resolution. It was shown that this method is robust to outliers in super-resolution [6–8]. Whatever norm is chosen, a regularization term is generally added to the variational problem.

Tychonov [4], bilateral total variation [6,9], total variation [7,8] or non-local regularization [10,11] have been considered. In all these cases, an a priori hypothesis is made on the regularity of the HR image. However, when observation noise is random, it is likely that such regularization is not necessary when many LR images

are available [12]. When there is an unnecessary regularization, high resolution features which could be recovered may be lost instead. In the case of unbounded outliers, results based on the least squares solution of super-resolution are not optimal because they are not well suited to the noise configuration.

In other areas of applied mathematics, it is known that the  $L^1$ -norm minimization has the ability to remove outliers. Candès and Tao showed in [13] that the outliers removal power of  $L^1$ -norm minimization is equivalent to a sparse recovery problem with sparsity having the cardinality of the support of outliers. They also showed that the observation matrix leads to the right result if it fulfills a restricted isometry property (RIP). Since this paper, the Null Space Property (NSP) has been shown to be an equivalent characterization of the capability to recover sparse vector from underdetermined observations [14].

## 1.2 Contributions

To our knowledge characterizations of  $L^1$  norm minimization have not been used in the context of super-resolution. In **Section 2**, we set up the variational super-resolution problem. We then (**Section 3**) formulate the problem of forgiving outliers in the data in a slightly weaker way than in [13]. Vaswani [15] studied partially known support which is a stronger formulation of sparse recovery. Knowledge of the support is also used for structured sparsity where dedicated methods are designed [16]. [17] considered weaker formulation of the robustness of  $L^1$ -norm recovery by considering a fixed sparsity support. We consider arbitrary set of supports for outliers, which will allow an easy application to the super-resolution problem. This leads to an equivalent slightly weaker Null Space Property. In **Section 4**, we apply these results to the super-resolution interpolation problem. We find lower bounds on the number of images ensuring the robustness to a given number of outliers. We also show that allowing for arbitrary sets of supports for outliers can provide better practical results. Finally, we show experiments illustrating these results in **Section 5**.

## 2 Super-Resolution Interpolation Model

### 2.1 Low Resolution Image Generation

In a finite dimensional context, LR images are generated by a linear map  $A$ :

$$\begin{aligned} A : \mathbb{R}^{ML \times ML} &\rightarrow (\mathbb{R}^{L \times L})^N \\ u &\rightarrow (A_i u)_{i=1, N} = (SQ_i u)_{i=1, N} \end{aligned} \tag{1}$$

where  $M$  is the super-resolution factor,  $N$  is the number of LR images,  $L \times L$  is the size of LR images,  $u$  is a HR image of size  $ML \times ML$ , the  $A_i$  are linear maps generating LR images,  $S$  is the sub-sampling operator by a factor  $M$  and  $Q_i$  are the deformations associated with each LR image. SR is the process of recovering

$u_0$  from  $w = Au_0 + n$  ( $n$  is the observation noise). In this paper, we suppose that the  $Q_i$  are known. In this setting, the inversion of  $A$  is called super-resolution interpolation.

It has been shown in [12] that  $A$  is almost surely full rank when motions are random compositions of translations and rotations and  $N \geq M^2$ .

## 2.2 Variational Formulation

When  $A$  is full rank and  $M^2 \leq N$ ,  $L^2$ -norm minimization guarantees that the energy of the reconstruction noise is bounded by the energy of observation noise times the operator norm of the pseudo-inverse  $A^\dagger$  of  $A$ . This leads to useful results when observation noise is bounded. In the case of outliers, no assumption is made on the power of the noise and  $L^2$  reconstruction does not guarantee a good result (unbounded reconstruction noise). In this paper, we study the efficiency of the  $L^1$ -norm minimization of the data-fit:

$$\operatorname{argmin}_u \|Au - w\|_1 \quad (2)$$

with  $w = Au_0 + n_0$ . We look for conditions on  $A$  ensuring that  $u_0$  is the unique solution of (2) when  $n_0$  is an outlying noise. Outliers have the form :  $n_0 = n.T$  with  $T$  a vector of 0 and 1 representing the support of the noise (the  $\cdot$  represents the component-by-component vector product). We do not make any hypothesis on  $n$ . In Section 3,  $A$  will be a general full rank matrix of an over-determined system. In other sections,  $A$  will be an over-determined full rank SR operator of size  $NL^2 \times (ML)^2$  with  $N > M^2$ .

## 3 Forgiving Matrices

### 3.1 Definitions

We introduce the concept of a  $\mathcal{T}$ -forgiving matrix  $A$  ( $A : \mathbb{R}^m \rightarrow \mathbb{R}^p$ ) :

**Definition 1. Forgiving Matrix** *Let  $\mathcal{T}$  be a set of supports in  $\mathbb{R}^p$  (subset of  $\{0,1\}^p$ ).  $A$  is called  $\mathcal{T}$ -forgiving if for all  $T \in \mathcal{T}, n \in \mathbb{R}^p, u_0 \in \mathbb{R}^m$ , we have:*

$$u_0 = \operatorname{argmin}_u \|Au - (Au_0 + n.T)\|_1 \quad (3)$$

and  $u_0$  is the unique minimizer.

When a matrix is  $\mathcal{T}$ -forgiving, the  $L^1$  minimization recovers  $u_0$  from any observation  $Au_0$  contaminated by outliers whose support is in  $\mathcal{T}$ .

**Definition 2. Sparse Capable Matrix** *Let  $\mathcal{T}$  be a set of supports in  $\mathbb{R}^p$ .  $B$  ( $\mathbb{R}^p \rightarrow \mathbb{R}^q$ ) is called  $\mathcal{T}$ -sparse capable if for all  $T \in \mathcal{T}, x_0 \in \mathbb{R}^p, y \in \mathbb{R}^q$ , we have:*

$$x_0.T = \operatorname{argmin}_x \|x\|_1 \text{ subject to } Bx = B(x_0.T) \quad (4)$$

and  $x_0.T$  is the unique solution to problem (4).

The Null Space Property found in [14] only depends on the Null-Space of the matrix (and its interaction with supports). It is a non-concentration property which can be stated as follows:

**Definition 3. Non-Concentration Property**

Let  $\mathcal{T}$  be a set of supports in  $\mathbb{R}^p$  and  $V$  a subspace of  $\mathbb{R}^p$ . We say that  $V$  has the  $\mathcal{T}$ -Non-Concentration Property (NCP) if for all  $v \in V \setminus \{0\}$  and all  $T \in \mathcal{T}$

$$\|v.T\|_1 < \|v.T^c\|_1 \quad (5)$$

where  $T^c$  stands for the complement support of  $T$ .

We say that a matrix has the  $\mathcal{T}$ -Null Space Property ( $\mathcal{T}$ -NSP) if its null space has the  $\mathcal{T}$ -NCP.

*Remark 1.* Notice that, given the finite-dimensional setting, the NCP property implies the existence of a constant  $\gamma < 1$  such that for all  $v \in V$  and all  $T \in \mathcal{T}$ :

$$\|v.T\|_1 < \gamma \|v.T^c\|_1. \quad (6)$$

This constant is called the NSP constant in the area of sparse recovery.

For the completeness of the paper, we now proceed with the direct proof of equivalence between the forgiveness of  $A$  and the Non-Concentration Property for the image of  $A$  ( $\text{Im}A$ ). This equivalence can be obtained by combining [13] and [14] and slightly modifying the proofs to introduce arbitrary  $\mathcal{T}$  instead of considering families of supports with fixed size. Indeed, [13] proves that forgiveness of a matrix (called linear coding capability) is equivalent to the sparse capability of any matrix whose kernel is the image of the original one and [14] proves that sparse capability is equivalent to the NCP (called there NSP).

### 3.2 Characterization of Forgiveness by the Non-Concentration Property

**Theorem 1.** *The two following propositions are equivalent:*

1.  $A$  is  $\mathcal{T}$ -forgiving
2.  $\text{Im}A$  has the  $\mathcal{T}$ -Non Concentration Property.

*Proof.*  $1 \Rightarrow 2$ : Let  $A$  be  $\mathcal{T}$ -forgiving, and  $T \in \mathcal{T}$ . Let  $w \in \text{Im}A \setminus \{0\}$ , there is  $u_0$  such that  $w = Au_0 \neq 0$ . From the characterization of the  $L^1$  minimizer in (3), we know that the following inequality holds

$$\|n.T\|_1 < \|Au - (w + n.T)\|_1 \quad (7)$$

for all  $n \in \mathbb{R}^p$  and for every sub-optimal  $u \neq u_0$ . The strict inequality is a consequence of the uniqueness. In particular, for  $n = w$  and  $u = 2u_0$  ( $u \neq u_0$  because  $Au_0 \neq 0$ ),  $Au = 2w$  and:

$$\|w.T\|_1 < \|w - w.T\|_1 = \|w.T^c\|_1. \quad (8)$$

This shows that  $\text{Im}A$  satisfies the NCP on  $\mathcal{T}$ .

$2 \Rightarrow 1$ : By hypothesis  $\text{Im}A$  has the NCP on  $\mathcal{T}$ . Let  $u_0 \in \mathbb{R}^m$ ,  $n \in \mathbb{R}^p$  and  $T \in \mathcal{T}$ . We have to show that  $u_0$  is a minimizer of (3). Let  $u \neq u_0$ . The  $L^1$ -norm is the sum of  $L^1$ -norms taken on complementary supports:

$$\begin{aligned} f(u) &= \|Au - (Au_0 + n.T)\|_1 \\ &= \|(Au - (Au_0 + n.T)).T\|_1 + \|(Au - (Au_0 + n.T)).T^c\|_1 \\ &= \|A(u - u_0).T - n.T\|_1 + \|A(u - u_0).T^c\|_1. \end{aligned} \quad (9)$$

We use the triangle inequality followed by the NCP :

$$\begin{aligned} f(u) &\geq \|n.T\|_1 - \|A(u - u_0).T\|_1 + \|A(u - u_0).T^c\|_1 \\ f(u) &> \|n.T\|_1 = f(u_0). \end{aligned} \quad (10)$$

This strict inequality shows that  $u_0$  is the unique minimizer of  $f$ . Consequently,  $A$  is  $\mathcal{T}$ -forgiving.  $\square$

With this slightly different result, the NCP can be checked on particular sets of supports, and not only those having a given cardinal as usually done in the sparse recovery framework. For example, in the context of image super-resolution, it is interesting to consider outliers contaminating a fixed number of LR images. This hypothesis models real situations like new object in the scene, light reflection...

*Remark 2.* The previous result implies the following already known result: the NSP of order  $K$  is equivalent to the  $K$ -sparse recovery capability. We just have to apply the result for  $\mathcal{T}_K$  the set of all supports of cardinal  $K$ .

*Remark 3.* Note that in the context of outlier removal, the NCP could be called “Image Space Property” for  $A$ .

## 4 Application to Super-Resolution

### 4.1 Sufficient Condition for K-Forgiveness

In this section, we suppose that we only have the knowledge of the number of outliers  $K$  for the super-resolution problem.  $A$  is the super-resolution operator and  $\mathcal{T}$  is the set of supports of cardinal  $K$ . We call this special case of  $\mathcal{T}$ -forgiveness the  $K$ -forgiveness. We first give sufficient conditions on the number of observed images for the NCP. Then we use the weaker Restricted Isometry Property (RIP) which is another sufficient condition for sparse capability. For any linear map  $A$  and support  $T$ , we call  $A_T$ , the operator  $u \rightarrow (Au).T$ .

**Sufficient Condition for the NCP** Let  $T$  be a support with cardinal  $K$ . We look for a sufficient condition such that:

$$\frac{\|A_T u\|_1}{\|A_{T^c} u\|_1} < 1 \quad (11)$$

holds for all supports  $T$  of size  $K$ .

We start by bounding the  $L^1$ -operator norm of  $A_T$ . Let  $a_i$  be the lines of  $A$ :

$$\frac{\|A_T u\|_1}{\|u\|_1} = \frac{\sum_{i \in T} |\langle a_i, u \rangle|}{\|u\|_1} \leq \frac{\sum_{i \in T} \sum_j |a_{i,j} u_j|}{\|u\|_1}. \quad (12)$$

Because each coefficient of  $A$  is a sample of a cardinal sine, we have  $|a_{i,j}| \leq 1$ . Therefore, we have

$$\begin{aligned} \frac{\|A_T u\|_1}{\|u\|_1} &\leq \frac{\sum_{i \in T} \sum_j |u_j|}{\|u\|_1} \\ &\leq K. \end{aligned} \quad (13)$$

Now we bound the ratio  $\frac{\|A_T u\|_1}{\|A_{T^c} u\|_1}$ . We use the  $L^1$  conditioning  $\kappa_{A_{T^c},1}$  of  $A_{T^c}$ . The  $L^p$  conditioning of an operator  $A$  is defined by:

$$\kappa_{A,p} = \frac{\sup_{\|u\|_p=1} \|Au\|_p}{\inf_{\|u\|_p=1} \|Au\|_p} \quad (14)$$

This leads to the following inequalities :

$$\begin{aligned} \frac{\|A_T u\|_1}{\|A_{T^c} u\|_1} &\leq \frac{K \|u\|_1}{\|A_{T^c} u\|_1} \\ &\leq K \left( \inf \frac{\|A_{T^c} u\|_1}{\|u\|_1} \right)^{-1} \\ &\leq K \frac{\kappa_{A_{T^c},1}}{\|A_{T^c}\|_1}. \end{aligned} \quad (15)$$

We use the fact that the  $L^1$  operator norm  $\|A_{T^c}\|_1$  can be bounded below the values taken on particular examples. The SR operator transforms constant HR images into constant LR images of same intensity. Consequently,  $\|A_{T^c}\|_1 \geq (NL^2 - K)/(ML)^2$  and:

$$\frac{\|A_T u\|_1}{\|A_{T^c} u\|_1} \leq K(ML)^2 \frac{\kappa_{A_{T^c},1}}{NL^2 - K}. \quad (16)$$

We consider  $\kappa_{A_{T^c},1}^m$  the maximum  $L^1$  condition number of  $A$  restricted to the lines  $T^c$ . A condition for  $K$ -forgiveness is:

$$N > K (M^2 \kappa_{A_{T^c},1}^m + 1). \quad (17)$$

This inferior bound on  $N$  is linear with respect to  $K$  and is tight. Indeed, we can find a case where it is easy to see that  $N$  must be at least greater than a constant times  $K$ : Consider a 1D super-resolution problem with a sub-sampling factor of  $M = 2$  and a number  $N = 2P > 2$  observations with the corresponding translations being  $0, 1, \dots, 0, 1$  respectively (**i.e.** there are  $P$  observation with translation 0 and  $P$  with translation 1). In this case, the reconstruction according

to equation (3) is the following HR signal: each sample is the median of the  $P$  values measured for each sample of the original signal. It is then clear that the  $L^1$  variational setting can not resist to more than  $P/2$  outliers. The worst case being that all outliers contaminate the same pixel (of the original signal) and have the same (unrelated to the signal) value.

**Sufficient Condition for the RIP** A consequence of the equivalence between outlier resistance and sparse recovery is that we can use the Restricted Isometry Property [13] to find a sufficient condition for the  $K$ -forgiveness capability using the more convenient  $L^2$  setting.

**Definition 4.**  $B$  has the restricted isometry property of order  $J$  and constant  $\delta \in ]0, 1[$  if for all  $x \in \mathbb{R}^{N(L \times L)}$ , for all supports  $T$  such that  $|T| = J$

$$(1 - \delta)\|x.T\|_2 \leq \|B(x.T)\|_2 \leq (1 + \delta)\|x.T\|_2. \quad (18)$$

Given a matrix  $A$ , we set  $B$  as the orthogonal projection on  $(\text{Im}A)^\perp$ , that is  $B = P_{(\text{Im}A)^\perp} = I - A(A^H A)^{-1} A^H$ . Showing a RIP of order  $J = K + K'$  with constant  $\delta < \frac{\sqrt{K'} - \sqrt{K}}{\sqrt{K'} + \sqrt{K}}$  for  $B$  gives the  $K$ -sparse capability of  $B$  (See [18]). Consequently,  $\ker B = \text{Im}A$  has the NCP and  $A$  is  $K$ -forgiving. Moreover, if for all  $T$  of cardinal  $J$ :

$$\frac{\|A(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} \leq \sqrt{\delta} \quad (19)$$

then  $B$  has RIP of order  $J$  and constant  $\delta$  (we square equation (18) and use the Pythagorean theorem). We can show using the same reasoning as in equation (13) that  $\|A_T^H\|_2 = \|A_T\|_2 \leq \sqrt{J}$ . Consequently, we bound the ratio:

$$\begin{aligned} \frac{\|A(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} &\leq \sigma_{\max} \frac{\|(A^H A)^{-1} A^H(x.T)\|_2}{\|x.T\|_2} \leq \sigma_{\max} \sigma_{\min}^{-2} \|A_T\|_2 \\ &\leq \frac{\kappa_{A,2}^2 \sqrt{J}}{\sigma_{\max}} \end{aligned} \quad (20)$$

where  $\sigma_?$  are the extremal singular values of  $A$ . Replacing with an admissible value of  $\delta$  gives the condition:

$$\frac{\kappa_{A,2}^4 (K + K')}{\sigma_{\max}^2} \leq \frac{\sqrt{K'} - \sqrt{K}}{\sqrt{K} + \sqrt{K'}}. \quad (21)$$

We take  $K' = 3K$  (which we found is the optimal choice for the resulting constant) and get:

$$\frac{\kappa_{A,2}^4}{\sigma_{\max}^2} \leq \frac{C_1}{\sqrt{K}} \quad (22)$$



where  $C_1 = 0.0670$ .  $\sigma_{max} \geq \frac{\|Au\|_2}{\|u\|_2}$  because  $\sigma_{max}$  is the operator norm of  $A$ . Taking  $u$  as a constant image leads to:  $\sigma_{max} \geq \sqrt{N/M^2}$ . Finally,

$$N > M^2 C_1^{-1} K \kappa_{A,2}^4 \quad (23)$$

is a sufficient condition for  $A$  to be  $K$ -forgiving. This bound uses the  $L^2$  conditioning of the full operator. It has been shown [12, 19] that the conditioning  $\kappa_{A,2}$  converges to 1 for a large number of images and random motions. For a 1D signal and  $M = 2$ , this sufficient condition is roughly  $N > 30K$  asymptotically. This bound has to be compared with the worst case scenario described in the previous section  $N > 4K$  (which is a necessary condition).

#### 4.2 Study of Particular Outlier Configurations

Here, the possibility to choose arbitrary sets of supports shows its benefit. Let  $\mathcal{T}$  be the set of supports contaminating  $N_c$  LR images. In the same way as before, we want to find sufficient conditions for the NCP for  $T \in \mathcal{T}$ . More precisely, we allow for up to  $K = N_c L^2$  outliers as long as they contaminate at most  $N_c$  images. We start by bounding operator norms with a tighter bound. Let  $\mathcal{S}$  be the set of contaminated LR images indices ( $|\mathcal{S}| = N_c$ ):

$$\begin{aligned} \frac{\|A_T u\|_1}{\|u\|_1} &\leq \sum_{i \in \mathcal{S}} \|A_i\|_1 \\ &\leq \sum_{i \in \mathcal{S}} \|S Q_i\|_1 \\ &\leq C_2 N_c \end{aligned} \quad (24)$$

where  $C_2$  is an upper bound of  $\|A_i\|_1$ .  $C_2$  is the maximum  $L^1$ -norm of the *sinc* used for interpolation. For 1D signals, the  $L^1$  norm of the *sinc* is roughly bounded by the logarithm of the size of its support. We plot in Figure 1 a numerical evaluation of this constant for 2D SR. Figure 1 shows the max of the  $L^1$  norms of the *sinc* for translational SR. This bound yields:

$$\begin{aligned} \frac{\|A_T u\|_1}{\|A_{T^c} u\|_1} &\leq \frac{\|A_T u\|_1 \|u\|_1}{\|u\|_1 \|A_{T^c} u\|_1} \\ &\leq \frac{\|u\|_1 C_2 N_c}{\|A_{T^c} u\|_1}. \end{aligned} \quad (25)$$

We introduce the pseudo-inverse  $A_{T^c}^\dagger = (A_{T^c}^H A_{T^c})^{-1} A_{T^c}^H$  (recall that  $A_{T^c}$  has full column rank if  $N - N_c < M^2$ ):

$$\begin{aligned} \sup_u \frac{\|u\|_1}{\|A_{T^c} u\|_1} &= \sup_{v \in \text{Im } A_{T^c}} \frac{\|A_{T^c}^\dagger v\|_1}{\|v\|_1} \\ &\leq \|(A_{T^c}^H A_{T^c})^{-1}\|_1 \sup_{v \in \text{Im } A_{T^c}} \frac{\|A_{T^c}^H v\|_1}{\|v\|_1} \\ &\leq \|(A_{T^c}^H A_{T^c})^{-1}\|_1 C_3 \end{aligned} \quad (26)$$

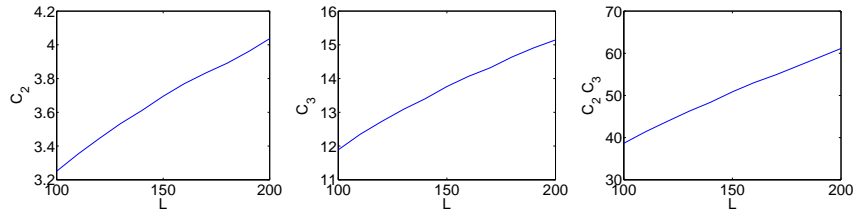
where  $C_3$  is the maximum  $L^1$  norm of the columns of  $Q_i^H S^H$ . This leads to the following sufficient condition :

**Proposition 1.** *If  $N_c$  images are contaminated, having  $N$  images with :*

$$N_c C_2 C_3 \| (A_{T^c}^H A_{T^c})^{-1} \|_1 < 1 \quad (27)$$

*guarantees a perfect reconstruction by  $L^1$  minimization*

We evaluate in Figure 1, the constant  $C_2$  and the product  $C_2 C_3$ . We cannot bound  $\|(A_{T^c}^H A_{T^c})^{-1}\|_1$  without knowledge of the motions because the its  $L^2$  operator norm cannot be bounded (LR grids could be arbitrarily close). However, with random motions, we know that  $(A_{T^c}^H A_{T^c})^{-1} \sim \frac{1}{N} I$  (see [12]) when  $T$  is fixed (on the first images for example) and  $N \rightarrow \infty$ . Asymptotically, the constraint is  $N_c < C_4 N$  (for  $L = 200$ ,  $C_4 = 60$ ). This is much better than the previous result without hypothesis on the support, where the equivalent constant would have been  $L^2 C_1^{-1} = 597000$  for  $L = 200$ . To have an idea of how robust the  $L^1$  SR problem is, we can compare this result asymptotically with the case of random matrices [20] which have been studied in the context of sparse recovery. The equivalent condition would be: for outliers with sparsity  $K$ , with  $N L^2 - M^2 L^2$  observations and a signal of size  $N L^2$ , the constraint would be  $K < (N - M^2) L^2 / \log(N / (N - M^2))$ . We see that asymptotically, this constraint is much better because  $\log(N / (N - M^2)) \rightarrow 0$  when  $N$  grows



**Fig. 1.** Constants for translational SR (a) evaluation of  $C_2$  with respect to  $L$  (b) evaluation of  $C_3$  with respect to  $L$  (c) evaluation of  $C_2 C_3$  with respect to  $L$ .

## 5 Experiments

### 5.1 Algorithm

The equivalence of the  $L^1$  minimization with sparse recovery shown in Section 3 allows for the use of existing algorithms. Daubechies *et al.* [18] showed that iteratively reweighted least squares (IRWLS) convergence to the  $L^1$   $K$ -sparse solution is guaranteed when  $A$  is  $K + \frac{2\gamma}{1-\gamma}$  sparse capable (with  $\gamma$  the NSP constant, see the remark in section 3.1) and when weights are carefully chosen

(and the regularization of the weights  $\epsilon_n \rightarrow 0$ ). We use this algorithm with the super-resolution  $L^2$  data-fit functional. We construct iterations equivalent to [18]:

$$\begin{aligned}
u_{n+1} &= \operatorname{argmin}_u \|\Omega_n(Au - w)\|_2^2 \\
z_{n+1} &= Au_{n+1} - w \\
r_{n+1} &= \text{decreasing sort of } \operatorname{abs}(z_{n+1}) \\
\epsilon_{n+1} &= \min(\epsilon_n, r_{n+1}(K + 1)) \\
\Omega_{n+1} &= \operatorname{diag} \left( [z_{n+1}^2 + \epsilon_{n+1}^2]^{-1/4} \right)
\end{aligned} \tag{28}$$

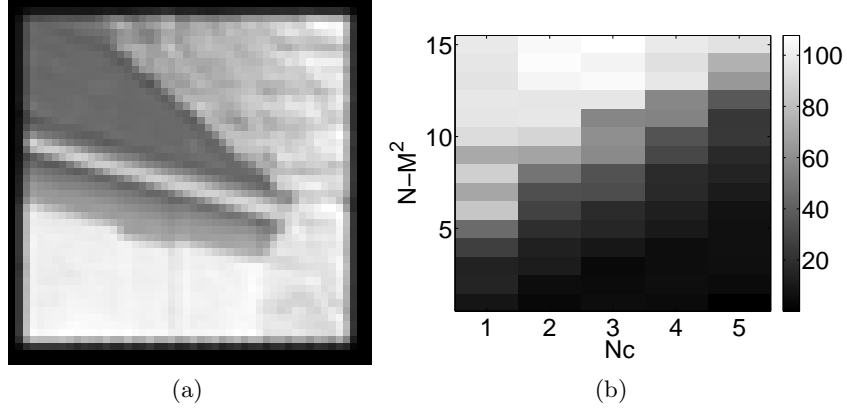
We chose this algorithm because it converges quickly (a few iterations in practice) and convergence can be checked by looking at the variations of  $\epsilon$ . Our aim is to give practical cases when outliers can be rejected.

## 5.2 Results

We show examples of outlier rejection using IRWLS. These practical results are better than our theoretical bounds which match the experience from compressed sensing. In Figure 2, we show an experimental evaluation of the number of images needed when  $N_c$  images are fully contaminated by outliers. For each  $N_c, N$ , the PSNR of the result of IRWLS is calculated for 30 experiments with different motion parameters. We plot the value of 10<sup>th</sup> percentile (90% of the reconstructions have a better PSNR). Each line of this matrix can be interpreted as a phase transition diagram. In Figure 3, we contaminate one LR image with the absolute value of Gaussian random noise of variance 125 (pixels take values in  $[0, 255]$ ). In this case, 6 clean images give a perfect reconstruction of the HR image. In Figure 4, even with more contaminated images (4 noisy LR images on 8 LR images), if the location of outliers is different between LR images,  $L^1$  minimization is still robust.

## 6 Conclusion

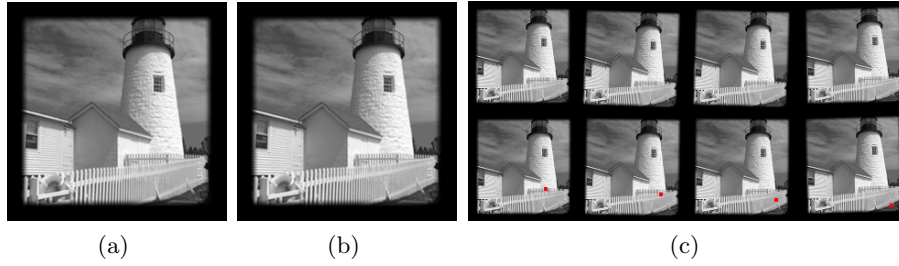
We have studied the outlier rejection capability of  $L^1$  super-resolution in a quantitative way. The link between the outliers resistance problem and sparse recovery allows for the direct translation of the results of the literature of sparse recovery to over-determined super-resolution. We showed that if enough images are available, outlying noise can be completely removed from the observations. We gave theoretical bounds on the ratio between the number of images and outliers to ensure a perfect reconstruction without regularization. We showed that some conditions on the support of outliers allows for a robustness to more outliers. This result takes the form of much better theoretical bounds derived using these particular supports. Experiments show that fewer images are necessary to resist outliers in practice.



**Fig. 2.** Experimental outlier rejection (a) HR image used for all experiments (b) 10% percentile of the PSNR (in dB) with respect to the number of outliers  $N_c$  and number of images  $N - M^2$



**Fig. 3.**  $L^1$  SR interpolation outlier removal for  $M = 2$  and  $N = 7$  (a) Ideal HR image (b) Reconstructed image (c) LR images (outliers on the last image)



**Fig. 4.**  $L^1$  SR interpolation outlier removal for  $M = 2$  and  $N = 8$  (a) Ideal HR image (b) Reconstructed image (c) LR images (outliers simulating saturated pixels (red squares) on the last 4 ones)

## References

1. Farsiu, S., Robinson, D., Elad, M., Milanfar, P.: Advances and challenges in super-resolution. Int. J. Imaging Syst. Technol. **14**(2) (2004) 47–57

2. Milanfar, P.: Super-resolution imaging. Volume 1. CRC Press (2010)
3. Tian, J., Ma, K.K.: A survey on super-resolution imaging. *Signal, Image and Video Processing* **5**(3) (September 2011) 329–342
4. Hardie, R.C., Barnard, K.J., Armstrong, E.E.: Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *Image Processing, IEEE Transactions on* **6**(12) (December 1997) 1621–1633
5. Nikolova, M.: A Variational Approach to Remove Outliers and Impulse Noise. **20**(1-2) (2004) 99–120
6. Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multiframe super resolution. *Image Processing, IEEE Transactions on* **13**(10) (October 2004) 1327–1344
7. He, Y., Yap, K.H., Chen, L., Chau, L.P.: A Nonlinear Least Square Technique for Simultaneous Image Registration and Super-Resolution. *Image Processing, IEEE Transactions on* **16**(11) (November 2007) 2830–2841
8. Yap, K.H., He, Y., Tian, Y., Chau, L.P.: A Nonlinear  $\ell_1$ -Norm Approach for Joint Image Registration and Super-Resolution. *Signal Processing Letters, IEEE* **16**(11) (November 2009) 981–984
9. Robinson, M.D., Toth, C.A., Lo, J.Y., Farsiu, S.: Efficient Fourier-Wavelet Super-Resolution. *Image Processing, IEEE Transactions on* **19**(10) (October 2010) 2669–2681
10. Protter, M., Elad, M., Takeda, H., Milanfar, P.: Generalizing the nonlocal-means to super-resolution reconstruction. *Image Processing, IEEE Transactions on* **18**(1) (jan. 2009) 36–51
11. Peyré, G., Bougleux, S., Cohen, L.: Non-local Regularization of Inverse Problems *Computer Vision - ECCV 2008*. In Forsyth, D., Torr, P., Zisserman, A., eds.: *Computer Vision - ECCV 2008*. Volume 5304 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, Berlin, Heidelberg (2008) 57–68
12. Traonmilin, Y., Ladjal, S., Almansa, A.: On the amount of regularization for Super-Resolution interpolation. In: *20th European Signal Processing Conference 2012 (EUSIPCO 2012)*, Bucharest, Romania (August 2012)
13. Candes, E.J., Tao, T.: Decoding by linear programming. *Information Theory, IEEE Transactions on* **51**(12) (December 2005) 4203–4215
14. Cohen, A., Dahmen, W., DeVore, R.: Compressed sensing and best k-term approximation. *J. Amer. Math. Soc* **22**(1) (2009)
15. Vaswani, N., Lu, W.: Modified-CS: Modifying Compressive Sensing for Problems With Partially Known Support. *Signal Processing, IEEE Transactions on* **58**(9) (2010) 4595–4607
16. Bach, F., Jenatton, R., Mairal, J., Obozinski, G.: Structured Sparsity through Convex Optimization. *Statistical Science* **27** (2012) 450–468
17. Xu, W., Hassibi, B.: On sharp performance bounds for robust sparse signal recoveries. In: *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, IEEE (June 2009) 493–497
18. Daubechies, I., DeVore, R., Fornasier, M., Güntürk, C.S.: Iteratively reweighted least squares minimization for sparse recovery. *Comm. Pure Appl. Math.* **63**(1) (2010) 1–38
19. Champagnat, F., Le Besnerais, G., Kulcsár, C.: Statistical performance modeling for superresolution: a discrete data-continuous reconstruction framework. *J. Opt. Soc. Am. A* **26**(7) (July 2009) 1730–1746
20. Dossal, C., Peyré, G., Fadili, J.: A numerical exploration of compressed sampling recovery. *Linear Algebra and its Applications* **432**(7) (March 2010) 1663–1679