Recovery of Non-Negative Signals from Compressively Sampled Observations Via Non-Negative Quadratic Programming

Paul D. O'Grady, Scott T. Rickard

Recovery of Non-Negative Signals from Compressively Sampled Observations Via Non-Negative Quadratic Programming

Paul D. O'Grady & Scott T. Rickard
Complex & Adaptive Systems Laboratory,
University College Dublin,
Belfield, Dublin 4, Ireland.

***Abstract*— The new emerging theory of Compressive Sampling has demonstrated that by exploiting the structure of a signal, it is possible to sample a signal below the Nyquist rate and achieve perfect reconstruction.**

In this paper, we consider a special case of Compressive Sampling where the uncompressed signal is non-negative, and propose an extension of Non-negative Quadratic Programming—which utilises Iteratively Reweighted Least Squares—for the recovery of non-negative minimum ℓ_p-norm solutions, $0 \leq p \leq 1$. Furthermore, we investigate signal recovery performance where the sampling matrix has entries drawn from a Gaussian distribution with decreasing number of negative values, and demonstrate that—unlike standard Compressive Sampling—the standard Gaussian distribution is unsuitable for this special case.

I. Introduction

The Nyquist-Shannon sampling theorem states that in order for a continuous-time signal to be represented without error from its samples, the signal must be sampled at a rate that is at least twice its bandwidth. In practice, signals are often compressed soon after sampling, trading off perfect recovery for some acceptable level of error. Clearly, this is a waste of valuable sampling resources. In recent years, a new and exciting theory of *Compressive Sampling* (CS) [1], [2] (also known as compressed sensing among other related terms) has emerged, in which a signal is sampled and compressed simultaneously using sparse representations at a greatly reduced sampling rate. The central idea being that the number of samples needed to recover a signal perfectly depends on the structural content of the signal—as captured by a sparse representation that parsimoniously represents the signal—rather than its bandwidth.

More formally, CS is concerned with the solution, $\mathbf{x} \in \mathbb{R}^N$, of an under-determined systems of linear equations of the form $\mathbf{\Phi x} = \mathbf{y}$, where the *sampling matrix* $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ has fewer rows than columns, *i.e.*, $M < N$. Critical to the theory of CS is the assumption that the solution $\mathbf{x}$ is *sparse*, *i.e.*, $\mathbf{y}$ has a parsimonious representation in a known fixed basis. The most natural norm constraint for this assumption is the ℓ_0 (pseudo-)norm, as it indicates the number of nonzero coefficients. However, minimisation of the ℓ_0 norm is a non-convex optimisation, which is *NP-complete* and cannot be computed in polynomial time. For these reasons the ℓ_1 norm is usually specified, as it is computationally tractable and also recovers sparse solutions,

$$\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_1, \text{ subject to } \mathbf{\Phi x} = \mathbf{y}, \tag{1}$$

where the recovered signal, $\mathbf{x}$, is such a solution.

In order to achieve the minimum number of measurements, M, required to perform perfect recovery, $\mathbf{\Phi}$ needs to be maximally *incoherent* with the sparse basis *i.e.*, have a non-parsimonious representation in that basis—a notion which is contrary to sparseness. For real-valued signals, the entries of $\mathbf{\Phi}$ are typically drawn from a random Gaussian distribution, as it is universally incoherent with sparse transformations, and performs exact recovery with the minimum number of measurements with high probability. Furthermore, Candès and Tao [3] present an important result that gives a lower bound on M that reliably achieves perfect recovery for a K-sparse signal ($\|\mathbf{x}\|_0 = K$): $M \geq CK\log(N)$, where C depends on the desired probability of success, which tends to one as $N \to \infty$.

In previous work, we proposed *Non-negative Under-determined Iteratively Reweighted Least Squares* (NUIRlS) [4] for the recovery of compressively sampled non-negative signals for the special case where compressive sampling is entirely non-negative, *i.e.* $\mathbf{\Phi}, \mathbf{x}, \mathbf{y} \geq 0$. Furthermore, we demonstrated that, unlike the compressive sampling of real-valued signals, sparse signals may be recovered from minimum ℓ_2-norm solutions using NUIRLS. Continuing in this direction, we relax the strict non-negative constraint on the CS problem presented in our previous work and turn our attention to the case where only the uncompressed signal, $\mathbf{x}$, is believed to be non-negative. We propose an extension to Non-negative Quadratic Programming—which fits with our new assumption—for the recovery of non-negative minimum ℓ_p-norm solutions, $0 \leq p \leq 1$.

This paper is organised as follows: We overview Iteratively Reweighted Least Squares in Section II and Non-negative Quadratic Programming in Section III. We propose an extension of Non-negative Quadratic Programming, which utilises under-determined IRLS, and perform signal recovery experiments in Section IV. We finish with a discussion and conclusion in Section V & VI respectively.

II. Iteratively Reweighted Least Squares

For our purposes, we desire the minimum ℓ_1-norm solution (Eq. 1) and require an objective function that recovers such solutions. However, the ℓ_1-norm objective has a discontinuity at the origin, and is therefore non-differentiable and cannot be minimised using standard gradient methods. Typically, the ℓ_1-norm objective is approximated by a function such as the *Huber* M-estimator [5], where the function penalises reconstruction error linearly for large error and behaves quadratically when the error falls beneath some small threshold close to the discontinuity. Another approach is to use *Iteratively Reweighted Least Squares* (IRLS), which approximates the ℓ_1-norm objective by reweighting the differentiable least-squares objective, where the residual error e, as specified by the ℓ_p norm, is computed by reweighting the minimum ℓ_2-norm solution: $|e|^p \equiv |e|^{p-2}e^2$, where $p = 1$.

In the context of CS, an IRLS algorithm specific to under-determined systems of equations is required,

$$\min_{\mathbf{x}\in\mathbb{R}^N} \|\mathbf{Q}^{-1}\mathbf{x}\|_2, \text{ subject to } \mathbf{\Phi x} = \mathbf{y}. \tag{2}$$

IRLS algorithms, unlike the pseudo-inverse, have no closed-form solution, as $\mathbf{Q}$ is dependent on the previous ℓ_p-norm solution estimate, $\mathbf{x}$. Therefore, in order to improve the estimate of the ℓ_p-norm solution, the procedure is repeated for a number of iterations.

A popular algorithm for performing under-determined IRLS is the FOCUSS algorithm [6], which performs non-convex ℓ_p-norm minimisations, *i.e.*, $0 \le p \le 1$, and recovers sparse solutions. Furthermore, the FOCUSS algorithm is used in the recovery of compressively sampled real-valued signals [7]. In Section I we mention the intractability of non-convex optimisations for the case where $0 \le p < 1$. However, local optimisation methods may be used to compute a global optimum for a non-convex optimisation if initialised sufficiently close to the global optimum. The FOCUSS algorithm follows such an approach and uses the least-squares solution as an initial estimate, which achieves good results in practice. The update equations for FOCUSS are:

$$\mathbf{x}_{k+1} = \mathbf{Q}_k\mathbf{\Phi}^\mathsf{T}(\mathbf{\Phi}\mathbf{Q}_k\mathbf{\Phi}^\mathsf{T})^{-1}\mathbf{y}, \tag{3a}$$

$$\mathbf{Q}_k = \operatorname{diag}(|\mathbf{x}_k|^{(2-p)}), \tag{3b}$$

where $\mathbf{Q}_k$ is initialised with the identity matrix, $\mathbf{Q}_0 = \mathbf{I}$, resulting in the initial solution estimate, $\mathbf{x}_1$, being the minimum ℓ_2-norm solution (least squares). On first inspection of Eq. 3a it may be tempting to suggest that the update equation is derived using normal equations, as is the case with the pseudo-inverse. However, this is not the case, the FOCUSS algorithm is derived by solving for the Lagrange multipliers of Eq. 2, then substituting the result into the solution, which gives a fixed point iteration for recovering minimum ℓ_p-norm solutions [6].

Alternatively, Eq. 2 can be restated as

$$\min_{\mathbf{g}\in\mathbb{R}^N} \|\mathbf{g}\|_2, \text{ subject to } \mathbf{\Phi Q g} = \mathbf{y}, \tag{4}$$

where the new problem is to find $\mathbf{x} = \mathbf{Qg}$, which results in the following algorithm:

$$\mathbf{g}_{k+1} = \mathbf{Q}_k\mathbf{\Phi}^\mathsf{T}(\mathbf{\Phi}\mathbf{Q}_k\mathbf{\Phi}^\mathsf{T})^{-1}\mathbf{y}, \tag{5a}$$

$$\mathbf{x}_{k+1} = \mathbf{Q}_k\mathbf{g}_{k+1}, \tag{5b}$$

$$\mathbf{Q}_k = \operatorname{diag}(|\mathbf{x}_k|^{(1-(p/2))}), \tag{5c}$$

where Eq. 5b computes the reweighting of the minimum norm solution Eq. 5a.

To explain how iterative reweighting results in a minimum ℓ_p-norm solution consider the objective minimised at each iteration for Eq. 2 & Eq. 3b,

$$\|\mathbf{Q}_k^{-1}\mathbf{x}_{k+1}\|^2 = \sum_{i=1}^{N}\Big(\frac{x_{k+1}^i}{|x_k^i|^{(2-p)}}\Big)^2. \tag{6}$$

The relatively large entries in $\mathbf{Q}$ deemphasise the contribution of the corresponding entries of $\mathbf{x}$ to the objective (Eq. 6), and vice versa. Therefore, large entries in $\mathbf{x}_k$ result in larger corresponding entries in $\mathbf{x}_{k+1}$, if the respective columns in $\mathbf{\Phi}$ are significant in fitting $\mathbf{y}$, implying that once a favourable weighting is obtained, the weighting at the next iteration continues to be favourable resulting in convergence to a minimum ℓ_p-norm solution. For $p = 1$, reweighting by $\mathbf{Q}$ deemphasises signal outliers in the much the same way as they are by ℓ_1-norm regression methods. Furthermore, the solutions recovered by the FOCUSS algorithm correspond to those recovered by ℓ_1-norm minimisation using the simplex algorithm, where each solution has at most M non-zero entries, ensuring that the recovered signal is sparse.

III. Non-negative Quadratic Programming

A multiplicative update for the solution of quadratic programs with non-negativity constraints on the solution has been recently proposed by Sha *et al.* [8]. The algorithm is referred to as *Non-negative Quadratic Programming* (NQP) and follows the reasoning behind *Non-negative Matrix Factorisation* (NMF) [9], where a non-negativity constraint is motivated by the underlying characteristics of the problem. The general problem of non-negative quadratic programming involves minimisation of the following quadratic function:

$$F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \mathbf{b}^\mathsf{T}\mathbf{x}, \tag{7}$$

subject to a non-negativity constraint on the solution, $\mathbf{x} \ge 0$. Furthermore, the matrix $\mathbf{A}$ is assumed to be symmetric and strictly positive definite, which guarantees that the function is bounded from below, *i.e.*, has a lower bound, and that its optimisation is convex. The NQP algorithm provides solutions that are expressed in terms of the positive and negative entries of the matrix $\mathbf{A}$ (Eq. 7). The elements of which are segregated by sign, giving two non-negative matrices $\mathbf{A}^+$ & $\mathbf{A}^-$:

$$A_{ij}^+ = \begin{cases} A_{ij} & \text{if } A_{ij} > 0, \\ 0 & \text{otherwise.} \end{cases};\ A_{ij}^- = \begin{cases} |A_{ij}| & \text{if } A_{ij} < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

from which $\mathbf{A}$ can be trivially constructed, $\mathbf{A} = \mathbf{A}^{+} - \mathbf{A}^{-}$. The NQP update is expressed in terms of $\mathbf{A}^{+}$ & $\mathbf{A}^{-}$, resulting in the following element-wise update:

$$x_i \leftarrow x_i \left[\frac{-b_i + \sqrt{b_i^2 + 4(\mathbf{A}^{+}\mathbf{x})_i(\mathbf{A}^{-}\mathbf{x})_i}}{2(\mathbf{A}^{+}\mathbf{x})_i} \right]. \tag{9}$$

The update rule—like NMF—is parameter independent, *i.e.*, requires no parameter selection, and is easy to implement. Furthermore, since $\mathbf{x}, \mathbf{A}^{+}$ & $\mathbf{A}^{-}$ are all non-negative, the factor on the right hand side of Eq. 9 will also be non-negative, which ensures the non-negativity of x_i as the optimisation is restricted to the non-negative orthant. For the case where $\mathbf{A}$ itself is non-negative, Eq. 9 reduces to the standard NMF update, and in this sense NQP can be considered to be a generalisation of NMF. Furthermore, it has been shown [8] that the NQP update has fixed points wherever the objective function, $F(\mathbf{x})$, achieves its minimum value, and converges monotonically to the global minimum.

NQP has been applied to a number of problems including acoustic time delay estimation [10] and parameter estimation for maximum margin hyperplanes in support vector machines [11].

A. Non-negative Least Squares using NQP

The NPQ algorithm finds solutions for general non-negative quadratic programming problems. For our purposes, we require a non-negative least squares algorithm:

$$\min_{\mathbf{x}} \ \frac{1}{2}\|\mathbf{y} - \mathbf{\Phi x}\|_{\mathrm{F}}^2 \quad \text{subject to } \mathbf{x} \geq 0, \tag{10}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius matrix norm, *i.e.*, $\|\mathbf{A}\|_{\mathrm{F}} = \mathrm{tr}(\mathbf{A}^{\mathsf{T}}\mathbf{A})$, $\mathbf{y} \in \mathbb{R}^{M\times 1}$ is the vector to be factorised, $\mathbf{\Phi} \in \mathbb{R}^{\geq 0, M\times R}$ is a known sampling matrix and $\mathbf{x} \in \mathbb{R}^{\geq 0, R\times 1}$ is the minimum non-negative ℓ_2-norm solution. The least-squares objective function is quadratic and therefore a solution can be found using quadratic programming. It is necessary to arrange Eq. 10 into a form compatible with Eq. 9; since $\mathbf{\Phi}$ is known, we fix $\mathbf{\Phi}$ and expand the reconstruction objective (Eq. 10):

$$\begin{aligned} \|\mathbf{y} - \mathbf{\Phi x}\|_{\mathrm{F}}^2 &= \mathrm{tr}((\mathbf{y} - \mathbf{\Phi x})^{\mathsf{T}}(\mathbf{y} - \mathbf{\Phi x})) \\ &= \mathrm{tr}(\mathbf{x}^{\mathsf{T}}\mathbf{\Phi}^{\mathsf{T}}\mathbf{\Phi x}) - 2\mathrm{tr}(\mathbf{y}^{\mathsf{T}}\mathbf{\Phi x}) + \mathrm{tr}(\mathbf{y}^{\mathsf{T}}\mathbf{y}), \end{aligned} \tag{11}$$

the parameters for the NQP algorithm are therefore $\mathbf{A} = \mathbf{\Phi}^{\mathsf{T}}\mathbf{\Phi}$ & $\mathbf{b} = -(\mathbf{\Phi}^{\mathsf{T}}\mathbf{y})$. In contrast to the standard least-squares solution provided by the moore-penrose pseudoinverse, there exists no closed-form solution due to the added non-negativity constraint, which necessitates the use of an iterative approximative algorithm such as NQP. For the case where $\mathbf{A}$ itself is non-negative, Eq. 9 reduces to the least squares NMF update.

IV. Under-Determined IRLS and NQP

We propose an algorithm for the recovery of sparse signals from compressively sampled non-negative data. The algorithm performs under-determined IRLS, as stated in Eq. 4, with an additional non-negativity constraint on the solution, *i.e.*, $\mathbf{g} \geq 0$. We refer to the algorithm as *Iteratively Reweighted Non-negative Quadratic Programming* (IRNQP), which performs non-convex minimisations recovering non-negative ℓ_p-norm, $0 \leq p \leq 1$, solutions. Furthermore, the algorithm is derived within the framework of NQP, resulting in a multiplicative update equation.

In contrast to standard NQP, for CS signal recovery we typically require minimum ℓ_p-norm solutions, and achieve this by combining NQP with IRLS, where least squares (minimum ℓ_2-norm solution) is expressed as a quadratic program, which is then reweighted to recover the minimum ℓ_p-norm solution:

$$\min_{\mathbf{g}} \ \frac{1}{2}\|\mathbf{y} - \mathbf{\Phi Q}_k\mathbf{g}\|_{\mathrm{F}}^2 \quad \text{subject to } \mathbf{g} \geq 0, \tag{12}$$

which is expanded as follows,

$$\begin{aligned} \|\mathbf{y} - \mathbf{\Phi Q g}\|_{\mathrm{F}}^2 &= \mathrm{tr}((\mathbf{y} - \mathbf{\Phi Q}_k\mathbf{g})^{\mathsf{T}}(\mathbf{y} - \mathbf{\Phi Q}_k\mathbf{g})) \\ &= \mathrm{tr}(\mathbf{g}^{\mathsf{T}}\mathbf{Q}_k\mathbf{\Phi}^{\mathsf{T}}\mathbf{\Phi Q}_k\mathbf{g}) \\ &\quad - 2\mathrm{tr}(\mathbf{y}^{\mathsf{T}}\mathbf{\Phi Q}_k\mathbf{g}) + \mathrm{tr}(\mathbf{y}^{\mathsf{T}}\mathbf{y}), \end{aligned} \tag{13}$$

therefore $\mathbf{A}_k = \mathbf{Q}_k\mathbf{\Phi}^{\mathsf{T}}\mathbf{\Phi Q}_k$ & $\mathbf{b}_k = -(\mathbf{Q}_k\mathbf{\Phi}^{\mathsf{T}}\mathbf{y})$ where the reweighted NQP update is now

$$\mathbf{g}_{k+1} \leftarrow \mathbf{g}_k \left[\frac{-\mathbf{b}_k + \sqrt{\mathbf{b}_k^2 + 4(\mathbf{A}_k^{+}\mathbf{g}_k)(\mathbf{A}_k^{-}\mathbf{g}_k)}}{2(\mathbf{A}_k^{+}\mathbf{g}_k)} \right], \tag{14a}$$

$$\mathbf{x}_{k+1} = \mathbf{Q}_k\mathbf{g}_k, \tag{14b}$$

$$\mathbf{Q}_k = \mathrm{diag}((\mathbf{g}_k)^{(1-(p/2))}), \qquad \mathbf{Q}_0 = \mathbf{I}. \tag{14c}$$

For the case where $\mathbf{A}$ is non-negative the algorithm becomes the NUIRLS algorithm.

As discussed in Section III, NQP is an iterative algorithm, IRLS is also an iterative algorithm, combining both results in a two-step iterative algorithm, where the minimum ℓ_2-norm solution at each NQP iteration is iteratively reweighted to recover a minimum ℓ_p-norm solution, which is used in the next NQP iteration and so on. As this process is repeated, the algorithm converges to a local optimum of Eq. 12.

A. Numerical Experiments

We use IRNQP to recover sparse signals from compressively sampled non-negative data, and compare the recovered signals to those recovered by non-negative least squares as implemented by NQP (Section III-A). We perform compressive sampling where $\mathbf{\Phi}$ and $\mathbf{x}$ are of fixed dimension, $M = 80$ & $N = 128$, and test for a number of signals with increasing K-sparseness, with $K = 60$ being the maximum. Since NQP does not have non-negative constraint on $\mathbf{\Phi}$, we perform two sets of experiments where the entries of $\mathbf{\Phi}$ are drawn from a folded (or rectified) Gaussian Distribution (where the absolute value of negative values is used) and a standard zero-mean Gaussian Distribution, where the former corresponds to least squares NMF when $p = 2$ and NUIRLS when $0 \leq p \leq 1$.We run the algorithm for 1500 NQP iterations, each having 150 IRLS iterations, and specify $p = \{0, 0.5, 1\}$. In order to keep both the proposed algorithm and non-negative least squares NQP in an even setting, the latter is run for 225000 (1500×150) iterations. The experiment is repeated for 20 Monte Carlo runs,

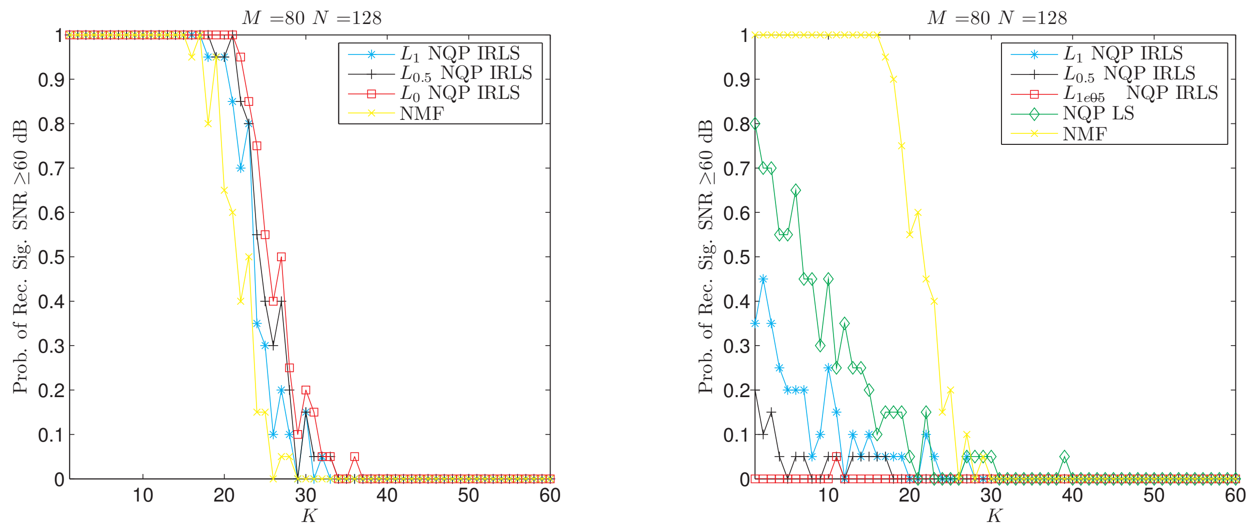

Fig. 1. Probability of recovering a signal with SNR $\geq$ 60 dB as a function of signal K-sparseness, where $\mathbf{\Phi}$ has entries drawn from a folded Gaussian (left) and standard zero-mean Gaussian distribution (right).

where a new $\mathbf{\Phi}$ is constructed for each run. The Signal-to-Noise Ratio (SNR) of the recovered signals are averaged over all Monte Carlo runs, and the probability of perfect recovery, SNR $\geq$ 60 dB, versus K-sparseness is plotted in Figure 1.

For the case where $\mathbf{\Phi}$ is non-negative (left plot in Figure 1), the plot indicates that for recovered signals with a required SNR of 60 dB, standard NMF successfully achieves the desired SNR for $K \leq 17$, while NUIRLS achieves the same SNR for $K \leq 21$, which demonstrates that if the compressively sampled signal is sufficiently sparse, in this case $K \leq 17$, a non-negativity constraint is enough to recover the signal [12]. Therefore, standard NMF can be employed in the recovery of sufficiently sparse signals from compressively sampled non-negative data. Unlike, signal recovery from compressively sampled real-valued data, where the ability to recover signals is dependent on the selection of a norm constraint that matches the sparseness assumption of the signal. These results correspond to those previously reported for the NMF implementation of the NUIRLS algorithm [4].

For the case where $\mathbf{\Phi}$ has standard Gaussian entries (right plot in Figure 1), the plot indicates a degradation in the quality of the recovered signals, which degrades further as $p \to 0$. Furthermore, it is evident that the sparse norms, $0 \leq p \leq 1$, perform worse than least squares under an assumption of non-negativity on the solution and a standard Gaussian $\mathbf{\Phi}$, while they perform better when $\mathbf{\Phi}$ is non-negative. Taking both plots together, it is easy to imagine a variation of recovery performance between the case where the entries of $\mathbf{\Phi}$ are drawn from a folded Gaussian and standard Gaussian distribution—the difference between both being the proportion of negative entries, which is 0% and 50% respectively.

We investigate the recovery performance of least squares NQP where the proportion of entries of $\mathbf{\Phi}$ range from 0% (folded Gaussian Dist.) to 50% (standard Gaussian Dist.). We use the same sampling and algorithm parameters that we specified in the previous experiment, and perform signal recovery for a number of signals with different K-sparseness, $K = \{10, 20, 30, 40, 50, 60\}$. The experiment is repeated for 50 Monte Carlo runs, where a new $\mathbf{\Phi}$ and set of K-sparse signals are generated for each run. The SNR of the recovered signals are averaged over all runs, and are plotted in Figure 2.

We first note that the recovered signal SNRs at 0% correspond to NMF in Figure 1, while the recovered SNRs at 50% correspond to least squares NQP. It is evident that for the sparsest signals, $K = \{10, 20, 30\}$, the recovery performance peaks around 30% resulting in improvements over NMF of around 15dB, while providing 40dB improvements over least squares NQP. Moreover, for the sparsest signals, there is steep decline from the peak to 50%, which gives the worst performance and indicates that a sampling matrix with entries drawn from a standard Gaussian distribution impedes the recovery of compressively sampled non-negative signals. In contrast, for standard CS where the recovered signal may be real valued, a Gaussian sampling matrix is frequently employed and achieves good results.

We repeat the above experiment (20 Monte Carlo runs) for the proposed algorithm where $K = \{5, 30\}$ and $p = 0.5$, the results are treated as before and plotted in Figure 3. For K=5, it is evident that recovery performance peaks around 10% resulting in improvements over NUIRLS of around 20dB. Furthermore, perfect recovery is achieved up to around 45% negative entries. For K=30, there is no sharp peak in performance and the performance plot closely follows the trajectory presented for the same class of signals in Figure 2. It is evident that as the proportion of negative entries increases, the difference in recovery performance between both plots decreases dramatically, which contrasts with Figure 2, where the difference in performance between the sparsest signals is generally constant.

V. Discussion

For the case where $\mathbf{\Phi}$ has entries drawn from a folded Gaussian, Figure 1 indicates that sparse signal recovery per-

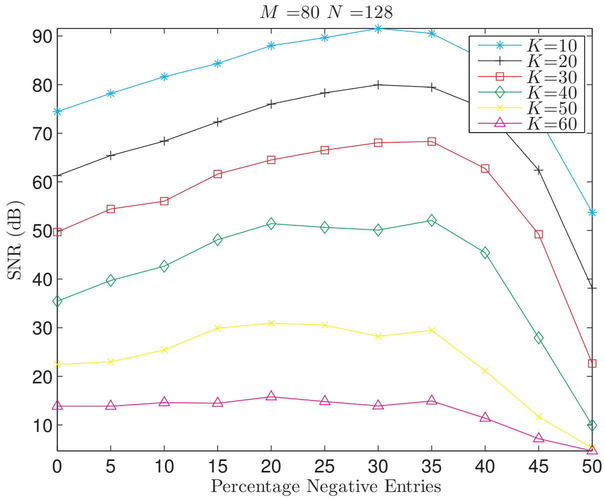

Fig. 2. Plot of recovery performance for least squares NQP versus the percentage of negative entries in $\mathbf{\Phi}$, where the recovered signals have varying K-sparseness.

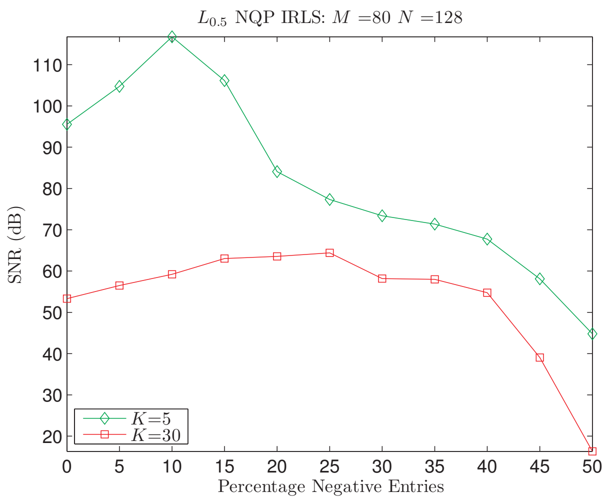

Fig. 3. Plot of recovery performance for the proposed algorithm versus the percentage of negative entries in $\mathbf{\Phi}$, where the recovered signals have varying K-sparseness.

forms as expected, with the ℓ_0 norm achieving the best results. In contrast, for the case where $\mathbf{\Phi}$ has entries drawn from a standard Gaussian, the recovery performance of sparse norms is reversed with the ℓ_2 norm performing best. One suggestion for this behaviour may be that initialising the reweighting procedure using the least-squares solution may not be sufficiently close to the global optimum for ℓ_p-norm minimisations, resulting in a non-convex-like degradation in recovery performance.

Although we do not explore the possibility here, IRNQP may be employed as a reweighted quadratic program algorithm for more general quadratic problems beyond least squares, where $\mathbf{x}$ is replaced with $\mathbf{Q}_k\mathbf{g}$, which results in an optimisation for $\mathbf{g}$ as in Eq. 13.

Finally, Figures 2 & 3 suggest an optimal form for $\mathbf{\Phi}$, which is based on the dimensionality of the sampling matrix, K-sparseness of the signals and the ℓ_p norm used for signal recovery. For future work, we endeavour to provide an analytical exposition of our observations for more general sampling matrices.

VI. Conclusion

In this paper, we proposed an extension to the non-negative quadratic programming algorithm for the purposes of signal recovery from compressively sampled observations, where the uncompressed signal is non-negative and the sampling matrix is real-valued. The algorithm is referred to as Iteratively Reweighted Non-negative Quadratic Programming and recovers non-negative minimum ℓ_p-norm solutions, $0 \leq p \leq 1$.

We investigated signal recovery performance where the sampling matrix has entries drawn from a Gaussian distribution with a decreasing proportion of negative values, and demonstrate that—unlike standard Compressive Sampling—the standard Gaussian distribution is unsuitable for this special case. Moreover, sparse norms perform particularly badly for the standard Gaussian sampling matrix.

Finally, this paper complements our previously reported work on the NUIRLS algorithm and demonstrated that, unlike the compressive sampling of real-valued signals, sparse non-negative signals may be recovered from minimum ℓ_2-norm solutions using not only non-negative sampling matrices but real-valued matrices also.

Acknowledgment

This material is based upon works supported by the Science Foundation Ireland under Grant No. 05/YI2/I677.

References

[1] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] David L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–306, 2006.

[3] Emmanuel Candès and Terence Tao, “Near optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–25, 2006.

[4] Paul D. O’Grady and Scott T. Rickard, “Compressive sampling of non-negative signals,” in *International Workshop on Machine Learning for Signal Processing*, Canćun, Mexico, Oct. 16–19 2008, IEEE Press.

[5] P. J. Huber, *Robust Statistics*, Wiley, 1981.

[6] Bhaskar D. Rao and Kenneth Kreutz-Delgado, “An affine scaling methodology for best basis selection,” *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 187–200, January 1999.

[7] Rick Chartrand and Wotao Yin, “Iteratively reweighted algorithms for compressive sensing,” Submitted, 2008.

[8] Fei Sha, Yuanqing Lin, Lawrence K. Saul, and Daniel D. Lee, “Multiplicative updates for nonnegative quadratic programming,” *Neu. Comp.*, vol. 19, no. 8, pp. 2004–2031, 2007.

[9] Daniel D. Lee and H. Sebastian Seung, “Algorithms for non-negative matrix factorization,” in *Adv. in Neu. Info. Proc. Sys. 13*. 2001, pp. 556–62, MIT Press.

[10] Yuanqing Lin and Daniel D. Lee, “Bayesian regularization and non-negative deconvolution for time delay estimation,” in *Adv. in Neu. Info. Proc. Sys. 17*. 2005, pp. 809–816, MIT Press.

[11] Fei Sha, Lawrence K. Saul, and Daniel D. Lee, “Multiplicative updates for nonnegative quadratic programming in support vector machines,” in *Adv. in Neu. Info. Proc. Sys. 15*. 2003, pp. 1041–1048, MIT Press.

[12] Alfred M. Bruckstein, Michael Elad, and Michael Zibulevsky, “A non-negative and sparse enough solution of an underdetermined linear system of equations is unique,” Submitted to IEEE Transactions on Information Theory, 2008.