RECONSTRUCTION OF SIGNALS FROM IRREGULARLY SAMPLED DATA

Vamsidhar Juvvigunta, Balakrishna Nakshatrala, Nagamohan Kompella

Engineering Research Center for Computational Field Simulation
Mississippi State University, Mississippi State, MS 39762

Correspondence

Vamsidhar Juvvigunta

Engineering Research Center for Computational Field Simulation, PO Box 9627
Mississippi State University, Mississippi State, MS 39762

Email: vamsi@erc.msstate.edu Phone: (601) 325 5135

ABSTRACT

This paper aims to describe the implementation of some of the available iterative algorithms for the reconstruction of irregularly sampled, band limited signals. These algorithms include the Adaptive Weights Method (ADPW), the Projection onto Convex Sets (POCS) method, and the Wiley/Marvasti (WILMAR) method. This work deals with one dimensional signals. The algorithms were tested on synthetic data with sizes up to 1024 samples. Then, they were tested on real signals (audio signals were considered) with sample sizes up to 131072 samples which corresponds to nearly 3 seconds of audio data (upsampled to 48kHz). The algorithms were compared based on their speed of convergence and accuracy of reconstruction. The influence of the auxiliary parameters specific to the WILMAR method was also examined.

1. INTRODUCTION

The sampling problem is one of the standard problems in signal analysis. Since a signal $f(x)$ cannot be stored in its entirety, it is sampled in some fashion, giving rise to a sample sequence. The important question now is how best f can be reconstructed or at least approximated from its sampling values $f(x_n)$. Now, if the samples are equally spaced, the well known sampling theorem by Shannon provides an explicit reconstruction. However, the problem of reconstruction is more difficult when the original signal f is irregularly sampled. Such a scenario arises whenever a signal is transmitted across a physical medium and there is a loss of information at the receiving end. Also, noise in the channel might contribute to the loss of samples in the signal introducing gaps in the signal, all of which contribute to making this problem of reconstructing signals from its irregularly sampled values an important one.

Another field where non-uniform sampling and reconstruction is of use is in (adaptive) compression of signals, as in the field of Computer Tomography where, given the sheer size of the (3-D) data sets involved, considerable savings in storage space can be obtained by sampling with large gaps in regions where the signal changes by small amounts and vice versa.

The primary motivation for this work, however, is the problem of feature detection in Computational Fluid Dynamics (CFD) solutions. Due to the complexity of the flow domain, CFD problems are usually solved on unstructured grids. The solution is a pressure function evaluated at the vertices of the elements of the grid. Feature detection on an irregular collection of points would be simplified greatly if the solution could be described by a collection of uniform grid points amenable to 2-D wavelet transforms [1], [2]. This transformation of the solution from an irregular collection of points to a regular collection of points is a problem in signal reconstruction from irregular samples. The implementation of reconstruction algorithms for irregularly sampled one dimensional signals is an effort at analyzing this larger problem.

Many algorithms have been proposed [3], most of which can be classified as either iterative methods or as direct methods (in which the problem is formulated as a linear problem and then described in terms of a matrix). However, it has been found that direct methods are either impractical or, in some cases, outright impossible to implement for real life problems, more so when large data sets are involved as in audio and video applications. Therefore, in the scope of the present work, we sought to compare and contrast three different algorithms that fall under the former category, namely, that of iterative methods. These methods, namely, the adaptive weights method (ADPW), the Wiley/Marvasti method (WILMAR) [4] and the projection onto convex sets method (POCS) [5], [6], [7] can be considered as alternating mapping methods, using the given information about the unknown signal (its spectral support and the sampling values) in a repetitive way.

2. ITERATIVE ALGORITHMS FOR SIGNAL RECONSTRUCTION

The three algorithms considered here share the same basic approach in reconstructing a signal from its irregularly sampled values. The following iterative two-step-algorithm is typical of most of these iterative methods.

(a)   As a first step, an auxilary signal is constructed from the given sampling values of f by an approximation operator A (A is a linear operator, except in the POCS method). The resulting function Af can be either a step function, a piece-wise linear interpolation or a well chosen discrete measure depending on the particular algorithm being made use of. It is of significance that only the sampling values of f are required to construct Af.

(b)   Now, the in general not band-limited signal Af is projected onto the space of all square integrable functions $f \in L^2(IR)$ by an orthogonal projection P_Ω in order to kill the high frequencies of Af. This projection can also be described as a low-pass filtering phenomenon, using the indicator function of the set Ω as transfer function. Alternately, this projection of f onto the space of all square integrable functions can be described as a convolution of f with a sinc-type function.

After these two steps, the first iteration is finished and we have obtained an approximation signal, $f_a = Af \bullet \sin c_\Omega$. The next iteration starts with the construction of a new auxilary signal $Af^{(1)}$ using the operator A, where $f^{(1)} = f - Af \bullet \sin c_\Omega$. So, the values of $f^{(1)}$ are known at every point and equal exactly the difference between the given irregularly sampled values and the values of the previous approximation signal. The low-pass filtered difference signal is then added to the approximated signal to obtain $f^{(2)}$ and so on. Hence, the (n+1)th iteration can be described as

$$f^{(n+1)} = f^{(n)} + A(f - f^{(n)}) \bullet \sin c_\Omega \quad \text{or,}$$

$$f^{(n+1)} = f^{(n)} - Af^{(n)} \bullet \sin c_\Omega + Af \bullet \sin c_\Omega$$

Rewriting it in the form

$$f^{(n+1)} = Tf^{(n)} + b,$$

we have,

$$\left\| f^{(n+1)} - f^{(n)} \right\| = \left\| Tf^{(n)} + b - Tf^{(n-1)} - b \right\|$$

$$= \left\| T(f^{(n)} - f^{(n-1)}) \right\|$$

$$< \gamma \left\| f^{(n)} - f^{(n-1)} \right\|$$

Then, convergence of the expression $f^{(n+1)} - f^{(n)}$ depends on the choice of the approximation operator, A, (as detailed in [6], [7], [8]) and also on the choice of γ, since as can be seen, smaller its value, faster the convergence, though its value increases with increasing irregularity in the sample sets [6].

3. DETAILS OF ALGORITHMS CONSIDERED

The details of the three methods considered in this work are now presented.

3.1 The adaptive weights method (ADPW)

In the ADPW method, the approximation operator, A, is given by the discrete measure,

$$Af = \sum_{1=1}^{p} \omega_i f(x_i) \cdot \delta_{xi}.$$

The weights ω_i are established in an adaptive way, depending on the sampling geometry. A good way to determine the ω_i's is given by the "nearest neighborhood method" as:

$$\omega_i = \frac{x_{(i+1)} - x_{(i-1)}}{2}$$

The essential effect of the adaptive weights algorithm is to attach lower importance to those sampling points which are in the regions of higher sampling density, and vice versa. We get the best results by using voronoi weights because they directly reflect the sampling geometry. The theory says that convergence for Adaptive weights algorithm is guaranteed, if the sampling sequences satisfy the Nyquist criterion, i.e. maximal gaps between the samples are smaller than the Nyquist gap.

3.2 The projection onto convex sets method (POCS)

This is a recursive method for finding a point in the intersection of p given closed convex sets. (a set C is said to be convex if for any two points $x_1, x_2 \in C$, the whole line between x_1 and x_2 belongs to C).

Consider the Hilbert space $L^2(\Re)$ with norm $\|.\|$ of all square integrable functions, and a convex set $C \in L^2(\Re)$. For any $f \in L^2(\Re)$, the projection Pf of f onto C is by definition the closest neighbor to f in C. If C is closed and convex, Pf exists and is uniquely determined by f and C from the minimality criterion.

$$\|f - Pf\| = \min_{g \in C}\|f - g\|$$

This rule defines the (in general) non-linear operator $P|(L^2(\Re) \to C)$.

The central idea of the POCS methods is as follows. Nearly most data related to an unknown signal $f \in L^2(\Re)$ can be viewed as a constraint that restricts f to lie in a closed convex set C_i in $L^2(\Re)$ (in our case such a property is given by the sample values). Thus for p known properties there are p closed convex sets $C_i, i = 1, 2, ..., p$ and f must lie in the intersection

$$C_o : = \bigcap_{i = 1}^{p} C_i$$

The problem is then to find a point in C_o given the sets C_i and projection operators P_i projecting onto C_i for $i = 1, \ldots, p$. The convergence properties of the sequence $(f^{(k)})$ generated by the recursion relation

$$f^{(k + 1)} = P_p P_{p-1} \ldots P_1 f^{(k)} \quad k = 0, 1.....$$

are based on fundamental proofs in functional analysis.

An important example for C is the set of all square integrable functions denoted by B_Ω^{2}, where the fourier transform of f vanishes outside Ω which is a closed linear subspace of $C \in L^2(\Re)$.

For a given sequence of samples $f(x_i), i = 1, 2, ..., p$ of an unknown function

$f \in B_\Omega{}^2$ we form the p sets.

$$C_i = \{g \mid g \in B_\Omega^2 \ \ and \ \ g(x_i) = f(x_i)\}$$

In words, C_i is the set of band limited functions whose values at the sampling points x_i coincide with the values of the sampled function. C_i can be described as $f(x_i) + C_o$. The projection of an arbitrary function h onto C_i is given by

$$P_i h = h - (h(x_i) - f(x_i))T_{x_i}(sinc_\Omega).$$

This clearly satisfies $P_i h(x_i) = f(x_i)$.

3.3 The Wiley/Marvasti method (WILMAR)

The Marvasti approach has its origin in Wiley's Natural sampling method and makes use of the formula (f = f * sinc) and the interpretation of the shift operator T_{xi} as convolution with Dirac measures. Thus the Marvasti approximation operator *Af* is nothing but a discrete measure of the form

$$Af = \sum_{1=1}^{p} f(x_i) \cdot \delta_{x_x}$$

and the first approximation signal can be obtained by convolving *Af* with the Sinc function. Hence

$$f_a = \sum_{i=1}^{p} f(x_i) T_{x_i} sinc_\Omega$$

Often, the speed of the convergence of the algorithm increases rapidly if one multiplies (6) by a **global relaxation parameter** λ and we obtain

$$A_\lambda f = \lambda \sum_{i=1}^{p} f(x_i) \cdot \delta_{x_x}$$

The speed of convergence depends on the choice of λ. If the relaxation parameter is too large, the iterations will diverge. On the other hand, a small value of λ brings slow convergence but ensures stability. A good choice of λ is

$$\lambda = \frac{n}{\gamma \cdot p}$$

where n is the length of the signal f and p is the number of sampling points. γ is a safety factor, which helps to prevent divergence. For highly irregular sampled signals one can enforce convergence by choosing a larger value for γ. But that would have a detrimental effect on the rate of convergence.

The Marvasti method is a special case of the adaptive weights method, if one determines all the weights ω_i to be equal (to λ).

4. Implementation

The algorithms were tested on synthetic signals generated on the computer as well as real signals (audio signals). The audio signals were recorded (at 8 kHz) with 16-bit precision from CD using the program *soundeditor* on an SGI machine. The audio files thus recorded in the AIFF format were then converted to a raw data stream (upsampled to

48 kHz) using the program *sox* on a Sun Sparc machine. This data was then given as input to our program (henceforth referred to as 'BMV'). To start with, BMV uses a random number generator function to randomly remove samples from the regularly sampled signal. BMV gives the user the option of having the maximal gap in the irregularly sampled signal to be either more or less than the Nyquist gap, the specification of which has important ramifications on the reconstruction of the signal [8]. The essential input to the program (made use of, by all three algorithms) is the irregularly sampled data and the spectral support of the signal that we wish to reconstruct. The various operating parameters like the sampling frequency used to discretize the signal, the error threshold and the maximum number of iterations allowed are also required as input. BMV also provides the user the option of selecting one or more of the three methods at a time.

ADPW and WILMAR ran successfully on both large and small data sets within reasonable time and exhibited $O(N)$ behavior. However POCS, as implemented, exhibited an $O(N^2)$ behavior and hence, could not successfully operate on the large audio data sets examined.

5. CONCLUSIONS

Accurate reconstruction of signals is a key issue when transmitting data over noisy channels as well as in (adaptive) data compression in fields such as Computer Tomography. All three algorithms

worked successfully on the synthetic and real signals considered.

The synthetic signals considered were superpositions of sine waves of different frequencies in the range 400 Hz to 1000Hz and were of lengths 512 and 1024 samples. The real signals were of lengths 2^{17} and 2^{18} samples. The performance of each of the three algorithms on the two synthetic data sets (of lengths 512 and 1024 samples respectively) is shown in Fig. 1 and Fig. 2 respectively. The performance graphs for the ADPW and the WILMAR methods on the audio signal are shown in Fig 3. The POCS method was found to be impractical vis a vis the time consumed for large data sets like the audio signal considered. Hence, it was not evaluated for the audio signal.

The results cannot be used to decisively conclude about the effectiveness of the algorithms for an arbitrary irregular data set. The POCS method converged to a lower error as compared to the other two methods for the smaller synthetic data set considered, but, exhibited a higher error when the larger synthetic data set was used. The reason for this, we believe, is the strong local influence of samples with large amplitudes when high frequency signals are considered. Finally, because of the high computational cost involved in using the POCS method for reconstructing of large signals, the ADPW and the WILMAR methods are concluded to be the methods of choice, with the ADPW method performing better than the

WILMAR method. However, for reconstructing small signals (with signal lengths of the order of a few thousand samples), the POCS method sometimes outperforms the ADPW method in terms of accuracy.

A graphical user interface has been developed that displays the results of each of the algorithms at different iterations on the chosen data set. The user can also, at any point of time, compare the results of each algorithm with the original signal, in addition to being able to view the error at each iteration. We expect this to be a valuable tool for academic applications as well as a simple tool for the reconstruction of audio data when parts of the data is lost due to wear and tear of the media used for storing the file(s), as might happen in the case of magnetic media.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] J.P. Antoine, P. Carrette, R. Murenzi, B. Piette. "Image analysis with 2-D continuous wavelet transform", *Signal Processing*, Vol 31, pp. 241-272, 1991.

[2] R. Machiraju, A. Gaddipatti, R. Yagel, "Detection and enhancement of scale-coherent structures using wavelet transform products". *Proc. of the Technical Conference on Wavelets in Signal and Image Processing*, SPIE Annual Meeting, San Diego, CA, July 21 - August 2, 1997.

[3] H.G. Feichtinger, K. Gröchenig. "Theory and practice of irregular sampling". In Benedetto J. and Frazier M., *Wavelets: Mathematics and Applications,* CRC Press, pp 305-363, 1993.

[4] F. A. Marvasti, M. Analoui. "Recovery of signals from non-uniform samples using iterative methods". *Proc. of IEEE; Int. Conf. on Circuits and Systems (CAS)*, April 1989

[5] S. Yeh, H. Stark. "Iterative and one step reconstruction from non-uniform samples by convex projections". *J. Opt. Soc. Am.* A 7/3, pp. 491-499, 1990

[6] T. Strohmer. *Irregular sampling, frames and pseudo-inverse.* M.S. Thesis, University of Vienna, 1991.

[7] H.G. Feichtinger, C. Cenker, M. Herrmann. *"*Iterative algorithms in irregular sampling: A first comparison of methods". *Proc. ICCCP`91*, March 1991, Phoenix, AZ pp 483-489, 1991.

[8] H. G. Feichtinger C. Cenker.

"Reconstruction algorithms for discrete non-uniform sampled band-limited signals". *Proc. 15.ÖAGM*, ÖCG, volume 56, pages 51 61, May 1991.
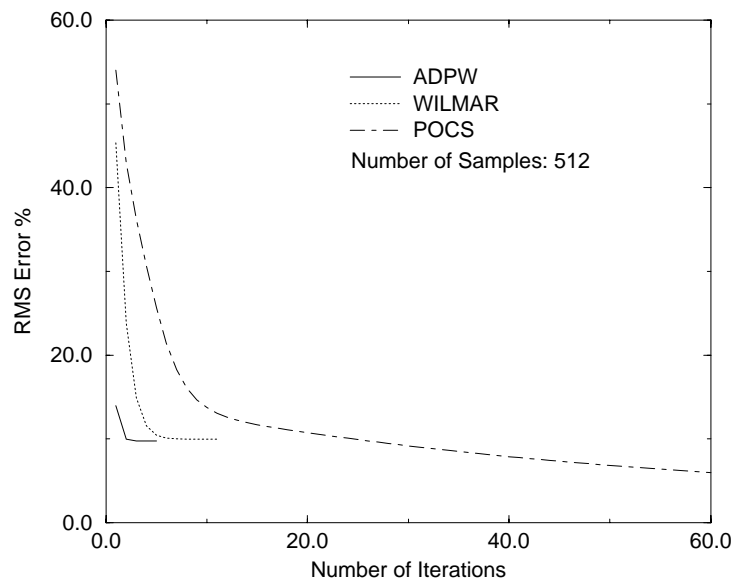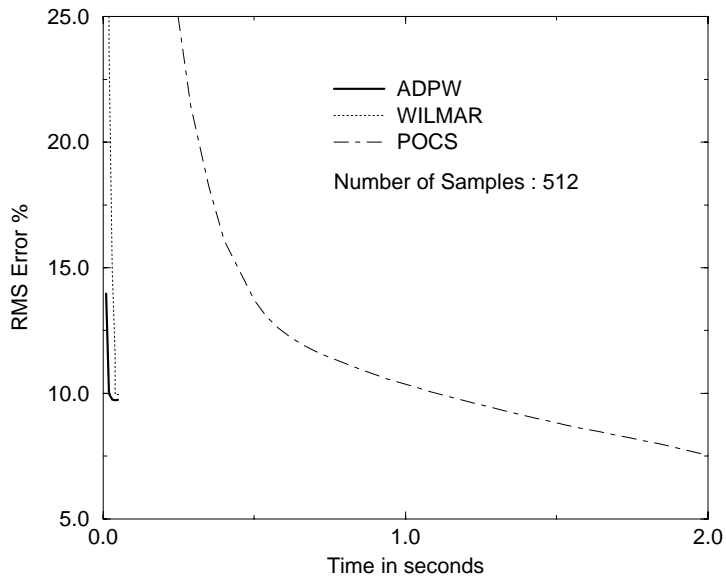
Figure 1. Performance on a synthetic data set with 512 samples

The number of samples in the figures refer to the number of samples in the reconstructed grid. The size of the irregular sampling set taken is around 70 % of the grid size.
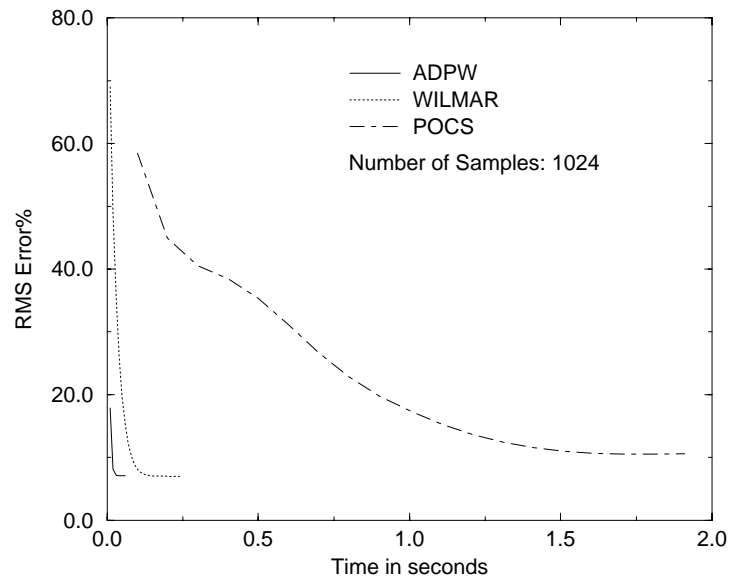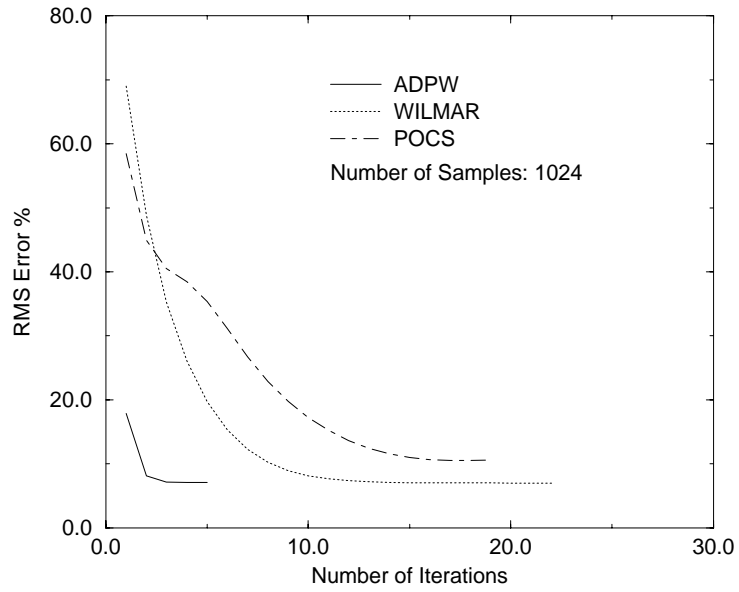
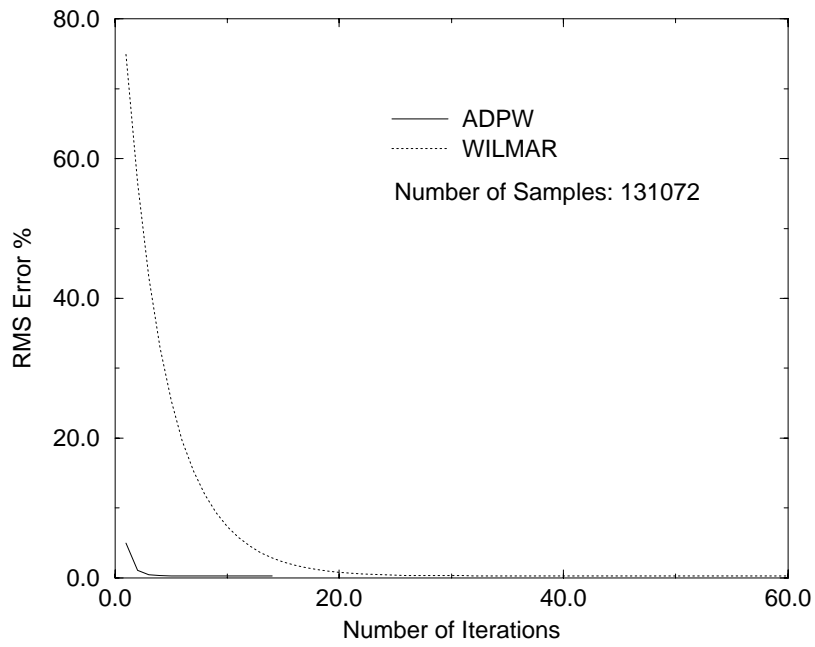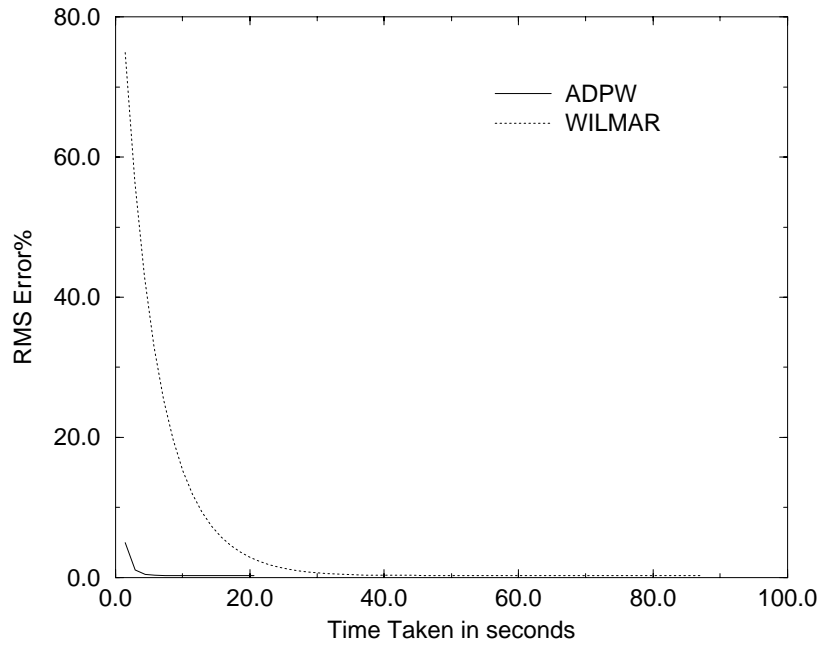Figure 2. Performance on a synthetic signal with 1024 samples

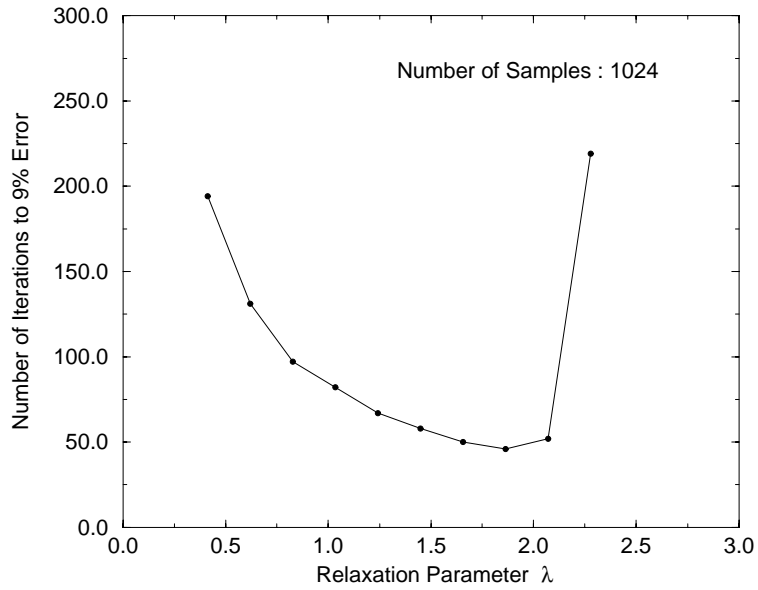Figure 3. Performance for an 8KHz audio signal with 131072 samples.

Figure 4. Performance variation of the WILMAR method with changes in the relaxation parameter
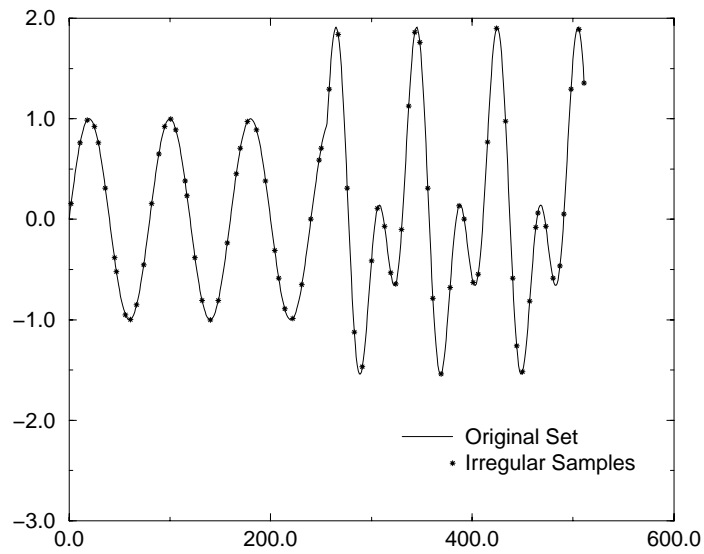
Figure 5. A typical re-sampled set

The datasets referred to are as follows.

1. Signal 1: the audio signal with 131072 samples

2. Signal 2: The synthetic signal with 512 samples

3. Signal 3: The synthetic signal with 1024 samples

DATA	POCS METHOD		
	Number of Iterations	Time Taken (in seconds)	RMS Error (as%)
Signal 1	--------	--------	--------
Signal 2	61	1.91	10.5902
Signal 3	200	9.72	2.8416

TABLE 1.

DATA	WILMAR METHOD		
	Number of Iterations	Time Taken (in seconds)	RMS Error (as%)
Signal 1	61	87.18	0.2842
Signal 2	22	0.24	6.9973
Signal 3	11	0.06	9.9482

TABLE 2.

DATA	ADPW METHOD		
	Number of Iterations	Time Taken (in seconds)	RMS Error (as%)
Signal 1	14	20.61	0.2859
Signal 2	5	0.06	7.0728
Signal 3	5	0.05	9.7346

TABLE 3.