New variants of the POCS method using affine subspaces of finite codimension, with applications to irregular sampling

H. G. Feichtinger, C. Cenker, M. Mayer, H. Steier, T. Strohmer

Universität Wien, Institut f. Mathematik
Strudlhofgasse 4, A-1090 WIEN, A U S T R I A

ABSTRACT

The POCS-method (projection onto convex subsets) has been proposed (see Yeh/Stark[21]) as an efficient way of recovering a band-limited signal from irregular sampling values. However, both the ordinary POCS-method (which uses one sampling point at a given time, i.e. consists of a succession of projections onto affine hyperplanes) and the one-step method (which uses all sampling values at the same time) become extremely slow if the number of sampling points gets large. Already for midsize 2D-problems (e.g. 128×128 images) one may easy run into memory problems. Based on the theory of pseudo-inverse matrices[8] new efficient variants of the POCS-method (so to say intermediate versions) are described, which make use of a finite number of sampling points at each step. Depending on the computational environment appropriate strategies of designing those families of sampling points (either many families with few points, or few families with many points, overlapping families or disjoint ones....) have to be found. We also report on numerical results for these algorithms.

1. PROBLEM DESCRIPTION

The irregular sampling problem for band-limited 1-D and 2-D signals [4,5,6,9,11,16,19] has found much interest recently. It concerns the reconstruction of a band-limited signal or image (with known frequency spectrum but unknown amplitudes) from a sufficiently rich family of irregularly spaced sampling values.

Let us start with a description of the problem for the continuous case first. For a closed subset $\Omega \subseteq I\!R^n$ we denote by B^Ω the set of all square integrable functions $f \in L^2(I\!R^n)$, i.e all signals of finite energy with spectrum $spec\ f := supp\ \hat{f} \subseteq \Omega$ (in other words $\hat{f}(s)$ vanishes on the complement of Ω). If Ω is a bounded set the multidimensional version of the well-known Shannon sampling theorem tells us, that for any sufficiently small lattice constant $\alpha > 0$ it is possible to recover the signal f completely from the set of sampling values $(f(\alpha m)_{m \in Z^n})$. In fact, a series representation of the form

$$f(t) = \sum_{m \in Z^n} f(\alpha m)\, g(t - \alpha m) \tag{1}$$

is possible. The series is convergent in the L^2-sense, i.e. the quadratic mean, and uniformly. The building blocks for this series representation are shifted versions of g which can be either a sinc-type function, i.e. the inverse Fourier transform of the indicator function $\mathbf{1}_\Omega$ of Ω, or more generally some $g \in L^2$ with sufficiently small spectrum $spec\, g$, satisfying $\hat{g}(s) \equiv 1$ on Ω. The critical value α_0 (the upper limit of all admissible values α) is usually called the Nyquist rate (for Ω).

In case of irregular sampling no signal reconstruction in such a simple form can be expected. But if the sampling density is high enough, complete reconstruction of the signal is still possible. The reader is referred to the applied literature[3,6,7,10,13,14] for various (iterative and non-iterative) reconstruction methods. In this paper we concentrate on new variants of the POCS method.

For the numerical test of the various methods we have worked with IRSATOL, a toolbox (for research purposes) based on the mathematical software package MATLAB, and developed at the Math.Dept. of the Univ. of Vienna, working on 486PC's and SUN workstations. The setup for the experiments was the following. We start with a complex signal vector or matrix with known spectrum (usually small compared to the signal format) and a family of sampling coordinates. For the reconstruction we assume that only the sampling values of the given signal at those sampling coordinates are available. This approach allows us to measure the mean square error between the given signal and the reconstructed one. Our standard example for this paper is depicted in Fig. 3. It shows a 1-D signal, with real and imaginary parts, and the sampling values marked by crosses. The signal length is 512, the FFT of the sequence has at most 51 non-zero values, and it is known at which coordinates those may appear, and we work with

93 sampling coordinates. This allows some irregularity, but on the other hand some of the larger gaps are about the size of the Nyquist rate.

2. GENERAL POCS METHODS

The POCS method[1,2,12,17,22,23] in general can be described as algorithm which produces, starting with an arbitrary vector which is usually the zero-vector, a sequence of approximations through projections on convex sets in some Hilbert space. In most cases one is searching for the unique element f_0 in the intersection $C := \cap_{i \in I} C_i$ of a family $\mathcal{C} = (C_i)_{i \in I}$ of convex sets. Denoting the projection from the full signal space $\mathcal{H}$ into the convex set C_i by P_i we obtain from an arbitrary starting element f_0 a sequence f_n of approximations through $f_{n+1} := P_{i_n} f_n$. The standard strategy for most applications consists in choosing a finite sequence of convex sets, to use each of those convex sets exactly once to produce a cycle of projections, and to iterate on that cycle, until the changes become marginal, or in many cases the determined number of cycles (or amount of time available) has been used up.

If the convex set happens to be a linear subspace (vector subspace) V_i of $\mathcal{H}$ the projection mapping is just the orthogonal projection from $\mathcal{H}$ onto V_i. It is well known that it can be obtained by taking any orthonormal basis for $(g_j)_{j \in J_i}$ for V_i in order to write P_i as $\sum_{j \in J_i} < f, f_j > f_j$. If V_i is relatively large within $\mathcal{H}$, i.e. if the codimension of V_i with respect to $\mathcal{H}$ is small, it is better to describe the projection as $\mathbf{Id} - P_i$. This is true in particular for hyperspaces $V_{\mathbf{n}} = \{f | < f, \mathbf{n} >== 0\}$, where $\mathbf{n}$ is some normal vector of length one, where we have $P_{V_{\mathbf{n}}}(f) = f - < f, \mathbf{n} > \mathbf{n}$.

If the convex set is an affine subspace of the signal space $\mathcal{H}$, i.e. some translate of a linear subspace, e.g. $C = u + V$, for some fixed element $u \in \mathcal{H}$, then the convex projection P_C is an affine mapping, which can be desribed through $P_C(f) = P_V(f - u) + u$. In the case of a hyperplane V this means that $P_C(f) = f - (< f, \mathbf{n} > - < u, \mathbf{n} >)\mathbf{n}$.

For the efficiency of any variant of such a POCS method the following three ingredients are relevant:

- The choice of the family of convex sets
- The description of the projection operators (constructive)
- A strategy to run through the family of convex families (in a certain order, possibly with repetitions)

3. VARIANTS OF THE SINGLE STEP POCS METHOD

During the investigation of the POCS-method it turned out that the order at which the sampling values were used may have a strong influence on the rate of convergence of the POCS-method. The following paragraphs explain several strategies to find optimal or suboptimal orders for the sampling points and describe corresponding numerical experiments. Since we are mostly interested in practical applications of the POCS method (and not so much in abstract theory for the infinite case) we only have to discuss cases which involve a finite FT and therefore a finite number of sampling points.

3.1 Standard POCS Method

The standard (single-step) POCS method for 1D-problems as described by Yeh/Stark makes use of the sampling points in their "natural" order, i.e. from left to right. It has been explained in section 9.2 in "Theory and Practice of Irregular Sampling"[5] that this is often not a very good strategy (after the correction at a sampling point has been carried out the sampling value at a close neighbor does not contribute much to improve the approximation in terms of the global signal energy) On the other hand, for 2D situations there is no such natural order of indices for random sampling sets. Therefore alternative strategies (while still keeping the single step approach) have to be investigated.

3.2 POCS method with permutations

For the 1D-case this is only a simple, but efficient modification of the standard POCS method and for the 2D-case it is the only natural one:

Fixing an arbitrary enumeration of the (finite) sampling set the algorithm is applied to each sampling value according to its given index. For the 1D-case we have observed good improvement of the speed of convergence (compared to the standard method) by applying this simple strategy (the extra computational load consists only in

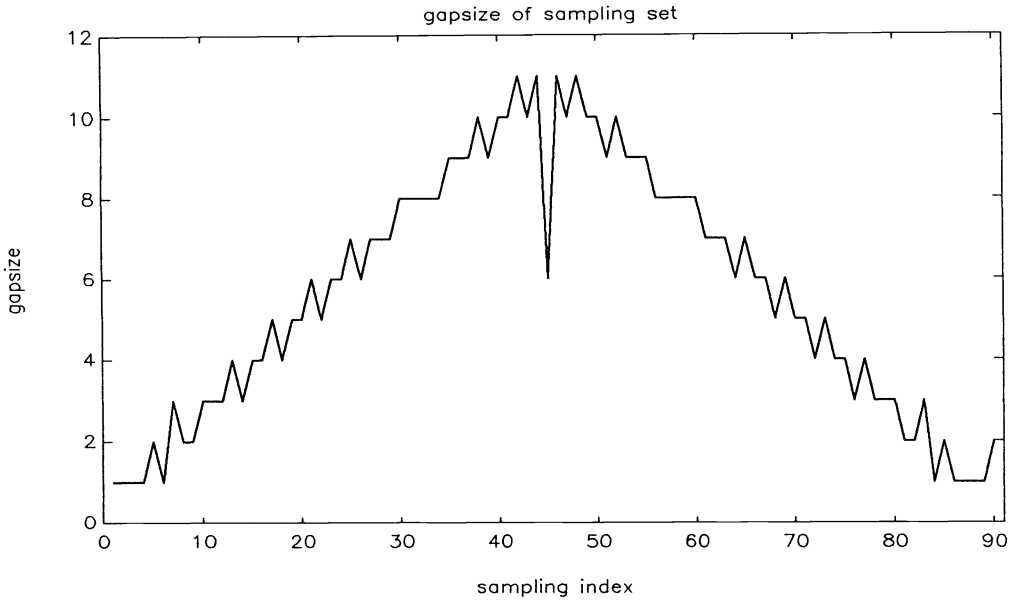

Figure 1:

determining some permutation of the index set). Although the actual rate of convergence depends on the given choice of the index-sequence we have found that "most" index sequences give more or less the same rate of convergence, and only few choices of the order (such as the left-to-right one) result in slower convergence. Only in few cases we have observed index sequences which result in a rate of convergence significantly better than the average rate of convergence. In the 2D-case there is no natural way of going through the sampling set (to compare with), but the fact that "most" choices for the index sequence result in some average rate of convergence is still true. From our experiments we conclude that it does not make sense to look for the "optimal" index sequence for a given sampling set. In the few cases we have found such sequences there was no evident criterion for determining why it was good, nor is there any way of calculating it a priori.

Observe that the structure of this algorithm (one determines first one cycle of single-steps, which is then iterated over and over again) has the consequence that each of the sampling points is used at a regular frequency, i.e. for a large number of steps each point is used almost equally often. Alternative strategies (to be described below) therefore have to involve ideas about the frequency at which the different sampling points are used during the process.

3.3 Maximal Correction Method

The observation that applying a POCS-step at the "next" point x_i may result in a very small improvement of the quality of the approximation led us to the *maximal correction method.* The idea behind the maximal correction method is the following: After each single iteration step the error between the approximation and the given sampling values is calculated, and the next step of this algorithm is a single POCS step using the sampling point at which the error takes its maximum.

Although the computational overhead (at each iteration step two arrays are subtracted and the maximum of the resulting array has to be found) the maximal correction method showed superior performance compared to the standard POCS- method.

Investigating the frequency the sampling values are used in the correction method it turns out that some sampling points in regions with a high sampling density are only rarely used or not at all. On th other hand isolated samples as well as samples in sparse regions are used more frequently. To demonstrate this phenomenon we have chosen a sampling set with variable gap sizei, small gaps near the boundary and larger gaps in the center. Fig. 2 describes

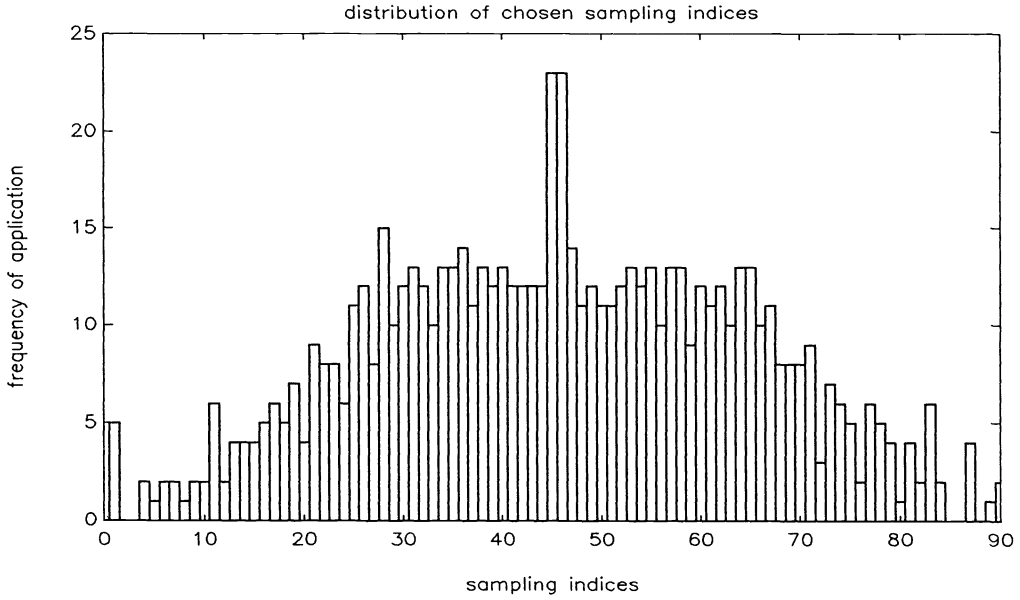

Figure 2:

the size of gaps between the sampling points. The frequency at which the sampling points are used by the maximal correction method is shown in Fig. 1) and indicates a clear correlation betweenthe gap size and the frequency at which points are used.

Nevertheless the following typical problem that may arise with this method: There are two "critical" sampling points at which the given approximation deviates most from the given sampling values. Applying the correction step at the point with the maximal error may increase the error at the second critical sampling point. Therefore that point will be used next. If now that correction causes in turn a large error at the first critical point, that point has to be used next to compensate the error, so that stagnation of the approximation process may occur. To prevent this the correction method with aging, where we have tried to force that each sampling point is used regularely (at least from time to time), was introduced.

3.4 Correction method with aging

The correction method with aging uses an aging-mechanism, which counts how often each sampling point has been ignored. After each step of the POCS algorithm the age of all sampling points that were not used in this step are incremented, whereas the counter for the point that has been used is set to zero.

In the course of the approximation the sampling values are used according to the maximal correction method, till the age of a sampling point exceeds a predefined treshold. If this happens the sampling point is used for the next iteration step independent of the maximum of the approximation error. If the sampling points are very irregular distributed across the domain of the signal to be approximated there are locations (clusters) where the gaps between the sampling points are small compared to the average distance.

To prevent sampling points of a cluster of being used too often, a slight variant of the aging-algorithm was introduced. In the modified algorithm the aging-counter of the left and the right neighbours of each used point were also diminished.

3.5 Red-black strategy

Another strategy for improving the convergence rate of the POCS-method is to use a so called red-black scheme. This means in connection with the POCS method that at the beginning the sampling sequence is divided in two

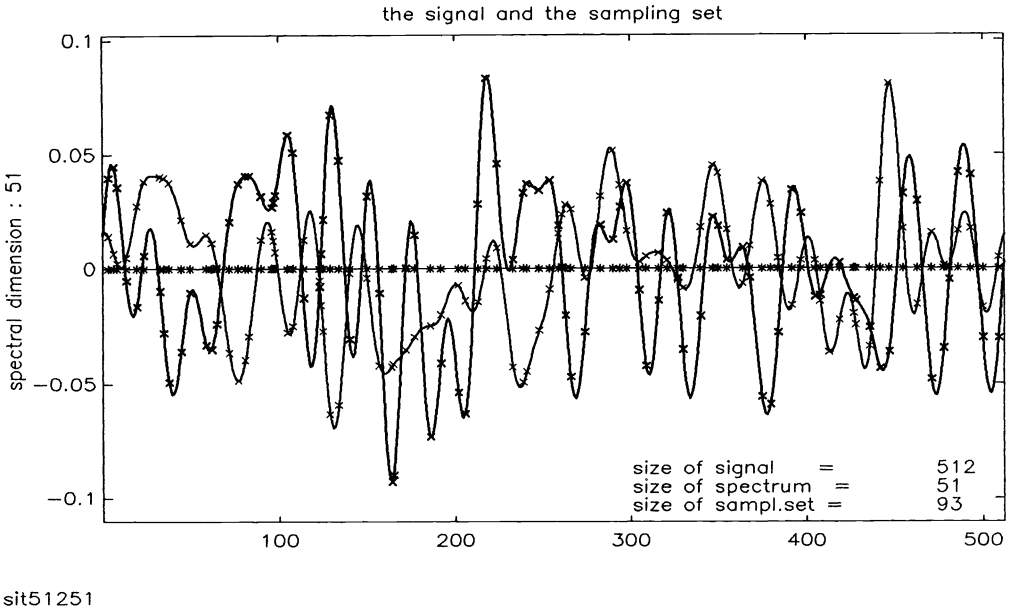

Figure 3:

subsequences, for example the sampling points with even indices ("the red ones") and those with odd indices ("the black ones"). First the algorithm is applied to the "red" and than to the "black" samples. Using a natural generalisation of the red-black scheme first the sampling points with indices $1 + s \cdot k$ ($s \in N$ fixed) then all sampling points with indices $2 + s \cdot k$ and so on are used (s denotes the grid size or offset of the red-black scheme.).

3.7 Comparison of POCS-methods

Fig. 3 shows a typical test signal of length 512 consisting of 51 pure frequences and 93 sampling values. Each method stopped the reconstruction after 8 cycles (i.e. $8 \cdot 93 = 744$ single steps) and the reconstructed signals were compared to the original one (cf. Fig. 4).

As can be seen from Fig. 4 the maximal correction method is close to machine precision after about 8 cycles. The POCS method with permutations ("Perm. POCS") and the Red/Black variant also show an acceptable behavior of convergence, at least a much better rate of convergence than using the standard algorithm. For the sake of comparison we have also included a random POCS method, where the sampling points are used in a complete random order without making use of cycles ("Random POCS"). This behavior of convergence for the different methods is typical for many 1-D situations.

3.8 Conclusion

Using the sampling values as they occur in the index set (standard method) has the advantage that no lengthly search is required to determine which point shall be used for the next iteration step. But the experiments showed that even with a small data set the convergence rate of this algorithm can be insufficient for practical applications.

At least in a noise free environment the correction method showed the best results, but growing data sets might lead to the conclusion that even the small computational overhead necessary for this algorithm is not tolerable on a given machine and for a given implementation.

The advantage of the improved rate of convergence (compared to the standard method) of the permutation strategy and the red-black scheme is the simplicity of the algorithms. It costs only little time to generate the corresponding index sequences.

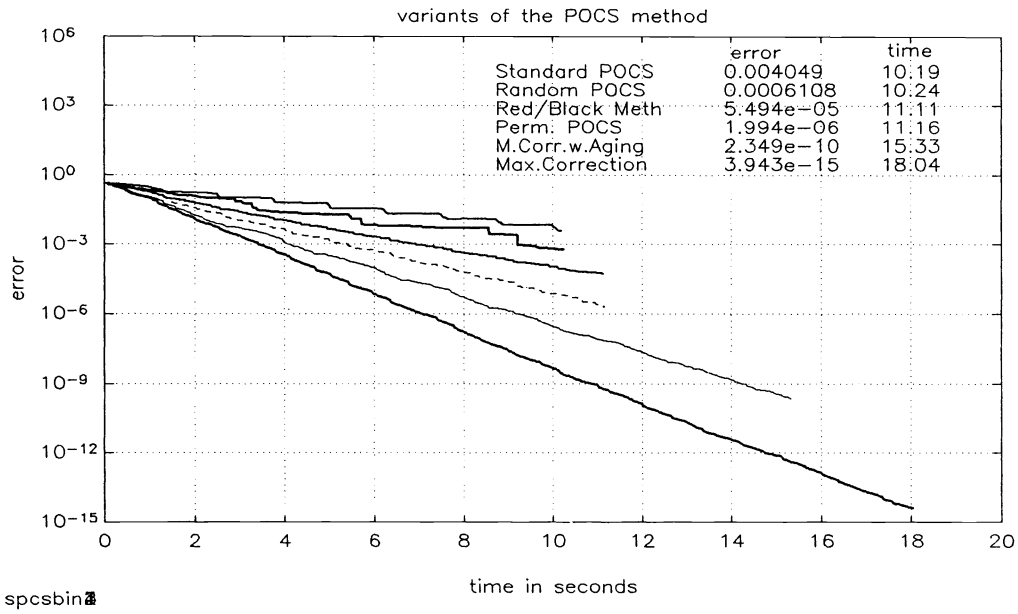

Figure 4:

4. PSEUDO-INVERSE MATRICES AND IRREGULAR SAMPLING

4.1 Geometric description

From an abstract point of view we have the problem of determining a vector in some finite-dimensional, let us say n-dimensional Euclidean space (real or complex) with inner product (i.e. real/complex Hilbert space $\mathcal{H}$ denoted by $\mathbb{C}^n$ or $\mathbb{R}^n$ to our best knowledge from a sequence of scalar products $(\langle f, g_i \rangle)_{i=1}^r$. Since these scalar products do not contain any information about directions orthogonal to the linear subspace space $\mathcal{H}_0$ generated by the system of functions $(g_i)_{i=1}^r$ the best we can achieve is the best approximation $f_a \in \mathcal{H}_0$ of the given function f, i.e. the orthogonal projection of f onto $\mathcal{H}_0$. Let M be the $r \times n$ matrix whose i-th row is the vector g_i and let s be the set of scalar products $(\langle f, g_i \rangle)_{i=1}^r$. Obviously s satisfies $s = f \cdot M'$. It is clear that $dim\mathcal{H}_0 \leq r$, but of course one has strict inequality if the sequence is not linear independent. In such a situation there are many ways to represent f_a as a linear combination of the "basis" vectors and the question is which one should be determined. Since all coefficient vectors $\Lambda = (\lambda_i)_{i=1}^r$ which represent the same vector $f = \sum_{i=1}^r \lambda_i g_i = \Lambda \cdot M$ form an affine subspace of $\mathbb{C}^r$ (or $\mathbb{R}^r$) there is a uniquely determined sequence Λ_0 (for each f in $\mathcal{H}$) with minimal ℓ^2-norm in this affine space. If M is invertible, the inverse matrix M^{-1} describes the linear mapping $f \mapsto \Lambda_0$ (by means of matrix multiplication from the right). If M is not invertible, the pseudo-inverse[18] of M will describe the required linear mapping. We shall write $pinv(M)$ or M^+ for the pseudo-inverse also called **Moore-Penrose inverse** of a matrix M. In the general case the matrix $pinv(M)$ can be used to solve the minimal norm least square problem (MNLSQ-problem) described above [8].

The standard method of calculating $pinv(M)$ involves the determination of a singular value composition (SVD) for the matrix M. The following identity[20] may signifincantly reduce the computational load for determining $pinv(M)$.

$$pinv(M) = M' \cdot pinv(M \cdot M') = pinv(M' \cdot M) \cdot M'. \tag{2}$$

Observe that $G_M := M \cdot M'$ is exactly the **Gram matrix** of the sequence $(g_i)_{i=1}^r$ which appears at various places in the applied literature[19] (e.g. the one-step method of Yeh and Stark[21] makes use of G_M, although they only discuss the case where G_M is invertible.)

Note also that in certain cases the application of formula 2 in general might be numerically less stable than the direct determination of $pinv(M)$, but we consider this as a minor problem for our situations compared to the memory problems with a huge SVD for M itself. In fact, the dimension of $G_M = M \cdot M'$ is exactly $r \times r$, i.e. the number of vectors involved, which is usually very small compared to the signal length. But for images or long signals with many samples even using formula 2 will not avoid that this pseudo-inverse method reaches its bounds (requiring too much memory or too much time).

4.2 Application in irregular sampling

In this section we discuss shortly how pseudo-inverse matrices can be used in connection with the irregular sampling problem. Pseudo-inverse matrix methods work efficient if the number of sampling points is small. While large gaps between sampling points may lead to divergence in some iterative algorithms, pseudo-inverse methods can still provide a good reconstruction of the signal.

Recall that $f \in B^\Omega$ is a band-limited discrete signal of length n with sampling values $(f(t_i))_{i=1}^r = (\langle f, g_i \rangle)_{i=1}^r$ where g_i is a *sinc*-type function, obtained by cyclic shift of $sinc_\Omega$ along the sampling coordinate t_i (since we have the reproducing convolution equation[7,20] $f = f * sinc_\Omega$). B^Ω, the space of band-limited signals can be spanned by the collection $(g_i)_{i=1}^r$, i.e. any element f in B^Ω can be represented as a finite linear combination of these shifted *sinc*-functions : $f = \sum_{i=1}^r \lambda_i g_i$, if there are sufficiently many sampling points available, in particular, if the maximal gap between sampling points is sufficiently small (depending on the diameter of the spectrum Ω).

If the matrix M contains the shifted *sinc*-function g_i as i-th row then the optimal approximation f_a of f through linear combinations of these functions, i.e. the orthogonal projection of f by elements of B^Ω (their linear span) can be obtained by means of the pseudo-inverse $pinv(M)$, even if there are *linear dependencies* between the rows of M (which happens, if $r > |\Omega|$, i.e. there are more samples than spectral points):

$$f_a = (f \cdot pinv(M)) \cdot M = \Lambda_0 \cdot M. \tag{3}$$

Here $\Lambda_0 := (\lambda_i)_{i=1}^r = f \cdot pinv(M)$ is the unique sequence with minimal ℓ^2-norm among those sequences Λ which satisfy $\Lambda \cdot M = f_a$.

Applying formula 2 we can determine f_a as:

$$f_a = (f \cdot M') \cdot pinv(M \cdot M') \cdot M. \tag{4}$$

Equation 4 shows that it is possible to determine f from the sampling sequence $(f(t_i))_{i=1}^r = f \cdot M'$ alone even without knowing f itself. Finally observe that it is not even necessary to establish the full matrix M in order to calculate the coefficient sequence $(f(t_i))_{i=1}^r \cdot pinv(M \cdot M')$, because the entries of the Gram matrix G_M are just values of the *sinc*-function at differences of sampling coordinates, due to the fact that the inner product of two shifted *sinc*-functions is just the value of that *sinc*-function at the difference of the shift parameters (remember that the *sinc*-function enjoys the reproducing convolution equation[7,20] $sinc * sinc = sinc$). Furthermore, it is easy to obtain therefrom the linear combination of shifted *sinc*-function using a double FFT. In fact, it suffices to establish first a sequence with the coefficients Λ_0 inserted at the sampling coordinates and zero elsewhere. The FFT of this sequence is the multiplied by $\mathbf{1}_\Omega$ and the inverse of the resulting sequence is exactly $f_a = \Lambda_0 \cdot M$. The limitation of this approach is therefore not the signal length, but rather the number of sampling points. Thus even for signals which would not allow to take correspondingly huge matrices M into the computer memory can be applied if these tricks are used for the implementation, and if there are not too many sampling points.

5. FINITE SPLITTINGS OF SAMPLING SETS

In Section 3. we have been concentrating on sequences of projections on hyperplanes, i.e. subspaces of the signal space of codimension one. The correction at each step was just a multiple of one of the functions g_i. In the preceding paragraph we have seen how we can obtain projections for an arbitrary finite dimensional subspace $\mathcal{H}_0$ via the pseudo-inverse of the corresponding matrix M. This method enables us to build correction terms, which are finite linear combinations of (a small number of) vectors which give the best possible improvement of the approximation by terms of this form. We shall now use this knowledge to provide alternative strategies, still following the general idea of POCS.

5.1 Strategies of splitting the sampling set

One obvious variant of the POCS method for irregular sampling is to choose an arbitrary covering $(I_j)_{j=1}^l$ of the sampling index sequence $I = \{1, \ldots r\}$. We will mainly concentrate on partitions, although it is possible to use some of the indices repeatedly. Let then $\mathcal{H}_j$ be the linear span of $\{g_i | i \in I_j\}$. Furthermore let M_j be the matrix containing the signals $\{g_i | i \in I_j\}$ as rows. The mapping $P_j : h \mapsto (h \cdot M_j') \cdot pinv(M_j)$ describes therefore the projection from $\mathcal{H}$ onto $\mathcal{H}_j$. It is used to describe the projection of a given approximate signal f_a onto the affine space $A_j := \{h : h(t_i) = s_i \; \forall i \in I_j\}$. Recall that $s = (\langle f, g_i \rangle)_{i=1}^r$ is the given sequence of sampling values.

One step in any of the new variants of the POCS method is performed as follows: A particular index set I_j is choosen. The "subproblem" is to find an approximation $f_{a,j} \in A_j$ for f. Now $f_{a,j}$ can be obtained from the previous approximation by the *pinv*-solution of $(s_i)_{i \in I_j} = f \cdot M_j'$. As described in section 4.2 this requires only to establish the pseudo-inverse of the matrix $G_j := M_j \cdot M_j'$ which contains only sampling values of the *sinc*-function. Since each of those projections will be used very often we start the algorithms by establishing all those matrices for any of the occurring index families $(I_j)_{j=1}^l$ and keep them in storage.

One cycle of POCS-steps consists of a sequence of projections, using all of the index sets, in some order. Once more one cycle is counted as one iteration, and successive approximations to f are obtained by running a couple of iterations of the given algorithm.

Three "splittting strategies" to obtain the partition $(I_j)_{j=1}^l$ have been implemented:

1. **Arithmetical strategy:** Each I_j simply contains k successive sampling indices for some fixed k.
2. **Tiling strategy:** Put a regular grid over the signal domain and collect the indices of the sampling values in each "tile" to form I_j.
3. **Grid strategy:** start as in 2., but then form "grids" by taking one sampling index out of each tile to get the set I_j.

5.2 Numerical results

The grid strategy prefers sampling values in "sparse" regions, because as soon as all points of a block are used up, it begins to use the previous samples once more. In the one dimensional case numerical experiments showed that the numerical stability of the matrices M_j obtained by the grid strategy are usually better than for the other strategies. In this case the sampling values are better distributed over the signal domain, which appears to keep the linear dependencies among the rows of M_j "lower" Also all matrices M_j of the grid strategy are of the same size which is usually not true for the tiling method.

It is clear that the behavior of the three splitting variants depends on the sampling set and on the number of blocks, in which the sampling set is split. We use again the signal shown in Fig. 3 for our experiments. To minimize the computational effort, it is obvious that one should store the various appearing pseudo-inverse matrices, because they do not change during iterations. To make the comparison ot the three strategies fair, we have splitted the sampling set for each method in such a way, that one obtains the same number l of submatrices M_j for all methods. In the first experiment (the result is shown in Fig 5) l is 5. That means for the arithmetic strategy, that we take the first $\frac{n}{r}$ successive sampling points, then the next $\frac{n}{r}$, doing this 5 times (so all matrices $(M_j)_{j=1}^l$ are of size $\frac{n}{r} \times \frac{n}{r}$). The sizes of the 5 matrices of the tiling strategy depend on the number of sampling points in each "tile". For the grid strategy the size of all matrices M_j is equal to the number of sampling points of that block with most sampling points. As mentioned above samples in sparse regions may be used multiple. This kind of weigthing results in higher computational effort, but the advantage is better stability for the reconstruction. Fig. 6 shows the convergence behavior of the three splitting methods using 20 submatrices.

The influence of the chosen number and size of the matrices is documented in Fig 7. The test image is of size 64×64 with 137 spectral points and 196 samples. The method under use is the grid strategy. The name *pocs1* says that the matrices M_j are of size 1×1, which is nothing than the single step method. Analogous *pocs16* means that the matrices are of size $16^2 \times 16^2$, (16^2 since we are dealing with a 2-D situation). One can see that using few large matrices gives a good reconstruction but the pseudo-inversion of large matrices costs much time. The single step

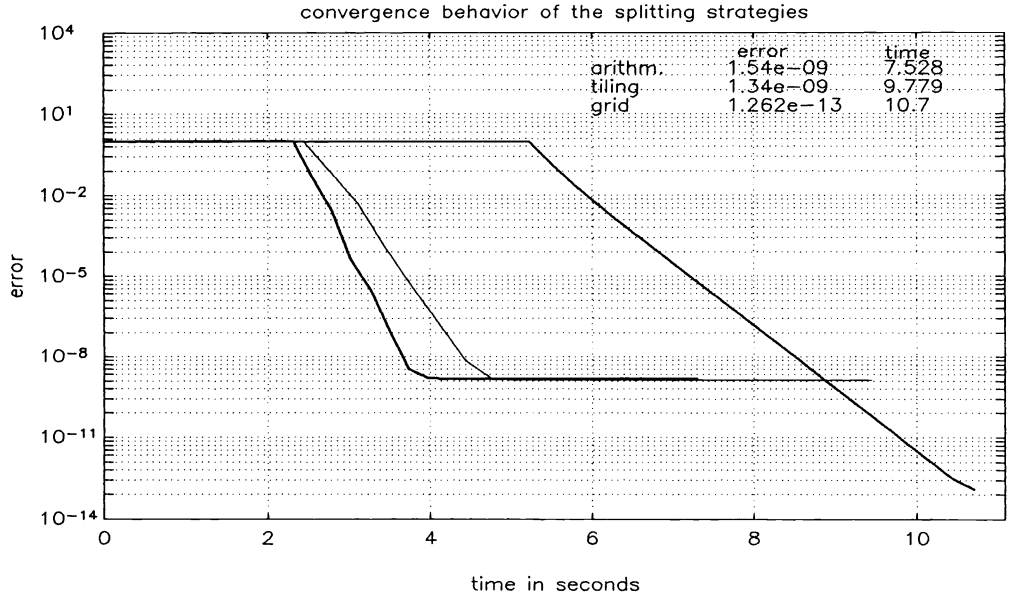

Figure 5: number of matrices = 10

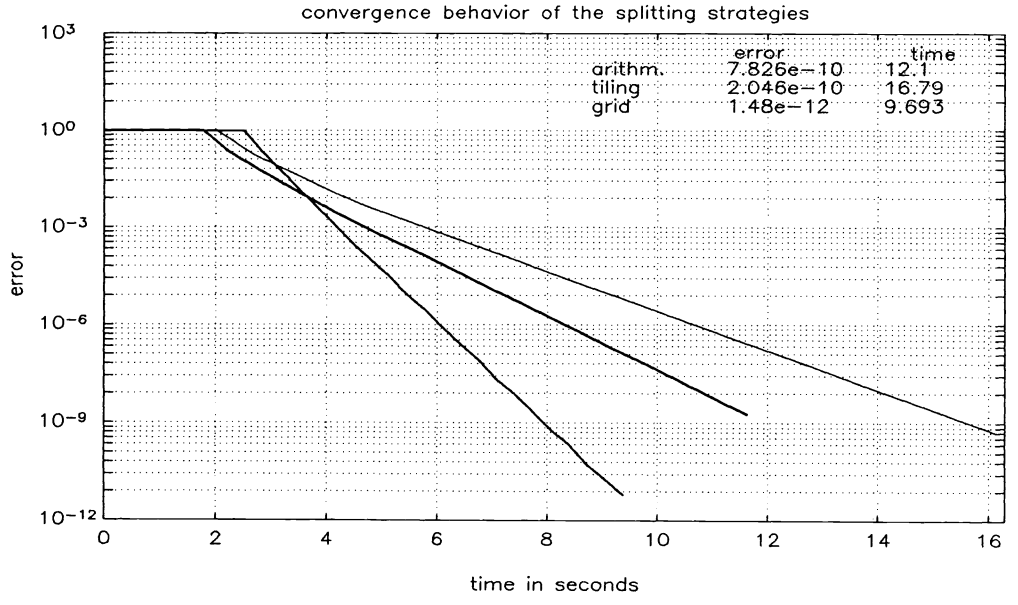

Figure 6: number of matrices = 20

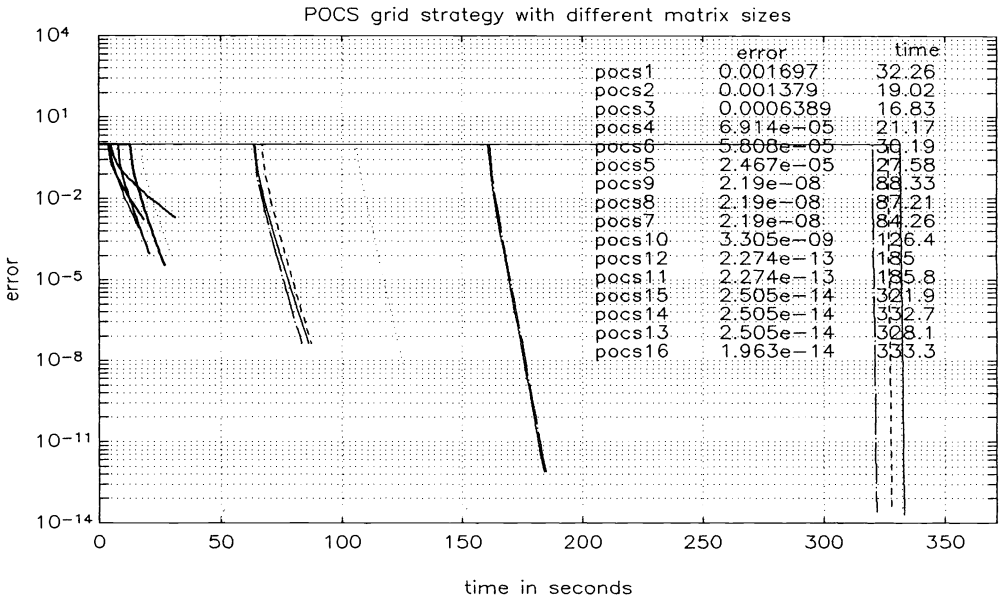

Figure 7:

method is fast but the approximation not very good. The best performance arises for an "intermediate" strategy, i.e. splitting the sampling set in such a way, that one obtains not too many midsize matrices.

It appears as a promising and easy task to implement these variants of the POCS method on parallel computers, especially for larger images (for example 256 × 256 and beyond).

5.3 Bunched sampling

Actually a special case of the general approach taken here is the case of "bunched" sampling. We call a discrete sampling sequence a *bunched sampling set*, if it is the finite union of (possibly different) lattices. To be more precise let us give a detailed description for the 2D-case: We call a family $X = (x_i)_{i \in I}$ a bunched sampling set for the plane, if there are offset parameters x_i and y_i as well as lattice constants α_i and β_i , such that $X = \cup_{i \in I}(x_i, y_i) + L_i$, where the lattices L_i are given by $L_i := (\alpha_i Z) \times (\beta_i Z)$. In that particular case we can split the irregular but bunched sampling set into several regular subsequences and the convex sets to be used for the reconstruction of f_0 are the families $C_i := \{f \in B^{\Omega} \mid f(l_m^i) = f_0(l_m^i)\ \forall m \in Z^2\}$. The corresponding pseudo-inverse of such a family C_i has a very simple strucure. Actually the pseudo-inverse matrices do not have to be established, but the necessary calculations can be carried out very efficient just making use of few FFT's. This has been explained in detail in our contribution for the SPIE91[8] and goes back to the paper by Oakley/Cunningham/Little[15].

The fact that in case of bunched sampling pseudo-inversion can be replaced by FFT's allows to go for signal lengths and image dimensions, which could never be treated with ordinary matrix methods (using SVD , for example).

6. SUMMARY

In this note we have presented "intermediate" versions of the POCS method (between the single-step and the one-step version), as well as improved strategies for the single-step POCS method. For each of those methods one can say that the sampling set has to be grouped into subfamilies, which are used in a certain order. From our experiments so far it appears that the optimal layout (which of course depends on the specific problem size and implementation) will be obtained by choosing not too many families, but try to keep on the other hand the number of matrices (pseudo-inverse matrices) to be kept in storage comparatively small. For the overall effort to obtain a certain degree of approximation one has to take into account the time to determine the sampling subfamilies,

to establish the corresponding pseudo-inverse matrices, and finally the number of cycles required. At least for 2D problems those strategies can be much more efficient than the simple version of POCS for the irregular sampling problem. More detailed investigations, also concerning the question of parallel implementations of those algorithms, have to be carried out in the future.

7. ACKNOWLEDGEMENTS

The authors would like to acknowledge support from the "Jubiläumsfond of the Austrian National Bank" as well as the "Universitätspreis der Kammer der gewerblichen Wirtschaft für Wien" , which allowed extensive numerical experiments on the algorithms as described above. These experiments were carried out on different machines, including AT-PC's, 386-PC's and SUN Workstations, using the mathematical software package MATLAB^{TM}. The project is continued as a project supported by the Austrian Science Foundation, "FWF", project number "PH8784". First results concerning section 3 have been presented at the IMACS conference 1991 in Dublin.

8. REFERENCES

1. L.M Bregman, "Finding the Common Point of Convex Sets by the Method of Successive Projection". *Dokl.Akad.Nauk. SSSR*, 162 (3), pp.487-490, 1965

2. G. Crombez, "Image Recovery by Convex Combinations of Projections". *J.Math.Anal.Appl* 155, pp. 413-419 , 1991

3. H.G.Feichtinger and K.H.Gröchenig, "Iterative reconstruction of multivariate band-limited functions from irregular sampling values". *SIAM J.Math.Anal.* 231 (1992), pp. 244–261.

4. H.G.Feichtinger and K.H.Gröchenig, "Error analysis in regular and irregular sampling theory", *Applicable Analysis*, to appear 1992.

5. H.G.Feichtinger and K.Gröchenig, "Theory and Practice of Irregular Sampling", contribution to another book about "wavelets", J.Benedetto and M.Frazier (Eds.), to appear 1993.

6. H.G.Feichtinger, C.Cenker and H.Steier, "Fast iterative and non-iterative reconstruction methods in irregular sampling", *Proc. ICASSP'1991*, Toronto,IEEE Signal Processing Soc., 1991, Vol. 3, pp 1773/1776.

7. H.G.Feichtinger, C. Cenker and H. Hermann, "Iterative algorithms in irregular sampling: A first comparison of methods", *ICCCP-91*, March 91, Phoenix, Arizona, USA. pp. 483-489

8. H.G.Feichtinger, "Pseudo-inverse matrix methods for signal reconstruction from partial data", *Proc.Conf. SPIE*, Boston , Nov. 1991.,Vis.Comm., 483-489.

9. K. Gröchenig, "Reconstruction algorithms in irregular sampling theory", *Comput.Math.*, to appear.

10. S. Kaczmarz, "Angenäherte Auflösung von Systemen linearer Gleichungen". *Bull.Int.Acad.Pol.Sci.Lett. A*, pp.355-357.

11.. S.P.Kim and N.K.Bose, "Reconstruction of $2-D$ bandlimited discrete signals from nonuniform samples". *IEEE Proc.*, Vol. 137, Pt.F,No.3, pp. 197–204, June 1990.

12. R.J.Mammone and R.J.Rothacker, "General iterative method of restoring linearly degraded images". *J.Opt.Soc.Am.* A/Vol.4, No.1, pp.208-215, Jan.1987.

13. F.A.Marvasti, "An iterative method to compensate for interpolation distortion", *IEEE Trans. ASSP* 37/10 (1989), 1617-1621.

14. F.A. Marvasti, M.Analoui and M.Gamshadzahi, " Recovery of signals from nonuniform samples using iterative methods", *IEEE Trans. Signal Processing*, Vol. 39/4 , pp. 872–878, 1991.

15. J.P.Oakley, M.J.Cunningham, and G.Little, "A Fourier-Domain Formula for the least-squares projection of a function onto a repetitive basis in N-dimensional space". *IEEE Trans.ASSP*, Vol. 38/1, pp.114–120, 1990.

16. K.D.Sauer and J.P.Allebach, "Iterative reconstruction of band-limited images from nonuniformly spaced samples". *IEEE Trans.* CAS-34/12 (1987).

17. M.I.Sezan and H. Stark, "Applications of Convex Projection Theory to Image Recovery in Tomography and Related Areas". *Image Recovery: Theory and Application*, pp. 29-77, Acad. Press, New York, 1987.

18. G.Strang, *Linear Algebra and Its Applications*, Sec.Ed., Academic Press, New York, 1980.

19. H. Stark (Ed.), "Image Recovery: Theory and Application", Academic Press, New York 1987

20. T.Strohmer, *Irregular Sampling, Frames and Pseudo-inverse*, Thesis, Univ. Vienna, 1991.

21. S.Yeh and H.Stark, "Iterative and one-step reconstruction from nonuniform samples by convex projections". *J.Opt.Soc. Amer. A*, Vol. 7/3, pp. 491–499 , 1990.

22. D.C Youla, "Mathematical Theory of Image Restoration by the Method of Convex Projections". *Image Recovery: Theory and Application*, pp. 29-77, Acad. Press, New York, 1987.

23. D.C.Youla and H. Webb, "Image Restoration by the Method of Convex Projections, Part 1, Theory", *IEEE Trans. Med. Imaging 1* (1982), pp. 81-94 .