Interpolation and the Discrete Papoulis-Gerchberg Algorithm

Paulo Jorge S. G. Ferreira

Abstract— **In this paper we analyze the performance of an iterative algorithm, similar to the discrete Papoulis-Gerchberg algorithm, and which can be used to recover missing samples in finite-length records of band-limited data. No assumptions are made regarding the distribution of the missing samples, in contrast with the often studied extrapolation problem, in which the known samples are grouped together. Indeed, it is possible to regard the observed signal as a sampled version of the original one, and to interpret the reconstruction result studied herein as a sampling result. We show that the iterative algorithm converges if the density of the sampling set exceeds a certain minimum value which naturally increases with the bandwidth of the data. We give upper and lower bounds for the error as a function of the number of iterations, together with the signals for which the bounds are attained. Also, we analyze the effect of a relaxation constant present in the algorithm on the spectral radius of the iteration matrix. From this analysis we infer the optimum value of the relaxation constant. We also point out, among all sampling sets with the same density, those for which the convergence rate of the recovery algorithm is maximum or minimum. For low-pass signals it turns out that the best convergence rates result when the distances among the missing samples are a multiple of a certain integer. The worst convergence rates generally occur when the missing samples are contiguous.**

I. Notation and Terminology

We will work in complex n-dimensional space $\mathbb{C}^n$, with the usual inner product and norm. We define a *signal* or *sequence* of length n to be any n-dimensional complex vector $\mathbf{x}$, with components, or samples, $x_0, x_1, \ldots, x_{n-1}$.

The *Fourier matrix* $\mathbf{F}$ is the unitary $n \times n$ matrix with components F_{mk} given by

$$F_{mk} = \frac{1}{\sqrt{n}} e^{j\frac{2\pi}{n}mk}$$

where j denotes the imaginary unit. The discrete Fourier transform (DFT) of $\mathbf{x}$, denoted by $\hat{\mathbf{x}}$, is by definition the sequence $\hat{\mathbf{x}} = \mathbf{F}\mathbf{x}$.

Denote by E_n the set $\{0, 1, \ldots, n-1\}$, and let S_a and S_b be two subsets of E_n with k elements. We say that S_a and S_b are *equivalent* if the elements of S_a can be obtained by addition of an integer constant, modulo n, to the elements of S_b. This simply means that S_a and S_b are related by a circular shift (a cyclic permutation). We say that a subset of E_n of cardinality $k < n$ is *contiguous* if it is equivalent to $E_k = \{0, 1, \ldots, k-1\}$.

We will be interested in two linear operations defined on $\mathbb{C}^n$, which will be called sampling and band-limiting, respectively.

Departamento de Electrónica e Telecomunicações / INESC, Universidade de Aveiro, 3810 Aveiro, Portugal. Tel. +351-34-370525, fax +351-34-370545, e-mail pjf@inesca.pt.

By definition, the *sampling operation* maps a sequence into another by setting to zero a subset of its samples. In matrix form, this corresponds to multiplication by a diagonal matrix $\mathbf{D}$ containing only zeros or ones. The diagonal of $\mathbf{D}$ will be called the sampling set associated with the sampling operation $\mathbf{D}$, and $\mathbf{D}$ itself will be called a *sampling matrix*. The *density* of a sampling set is, by definition, s/n, s being the number of nonzero entries in the sampling set. It is assumed that $s < n$, that is, $\mathbf{D} \neq \mathbf{I}$.

We define a *band-limiting operator* to be any linear operator characterized by a matrix $\mathbf{B}$ of the form $\mathbf{B} = \mathbf{F}^{-1}\mathbf{\Gamma}\mathbf{F}$, where $\mathbf{\Gamma}$ is a sampling matrix other than $\mathbf{I}$. Thus, according to this definition, the band-limiting operation in $\mathbb{C}^n$ is similar to a sampling operation performed in the frequency domain. We define the *bandwidth* of the sequence $\mathbf{y} = \mathbf{B}\mathbf{x}$ to be q/n, where q is the number of nonzero entries in the matrix $\mathbf{\Gamma}$. For real low-pass sequences, a bandwidth of b means that the highest normalized frequency in the sequence is $b/2$. It is natural to call *passband* of $\mathbf{B}$ to the nonzero entries in the diagonal of $\mathbf{\Gamma}$. The low-pass and high-pass signals are fixed-points of band-limiting operators with a contiguous passband.

Clearly, the set of all sampled sequences $\mathbf{y} = \mathbf{D}\mathbf{x}$ with $\mathbf{x} \in \mathbb{C}^n$ is a subspace of $\mathbb{C}^n$. Similarly, the set of all band-limited sequences $\mathbf{y} = \mathbf{B}\mathbf{x}$ with $\mathbf{x} \in \mathbb{C}^n$ is a subspace of $\mathbb{C}^n$. Furthermore, sampling and band-limiting are idempotent operations, and every sampling or band-limiting matrix is a projection matrix.

II. Purpose and Main Results of this Paper

There are four possible models for the band-limited extrapolation problem, which correspond to all possible combinations of discrete and continuous time and frequency. These models and some of their relations were studied in [1], along with a number of approximation results. We will not insist here on the L_2 extrapolation problem nor on its sampled versions. Instead, we will focus on finite-dimensional interpolation and extrapolation problems, characterized by finite (or periodic) time-domain and frequency-domain vectors. In the terminology of [1], this corresponds to the discrete-discrete model.

More precisely, this paper is concerned with the recovery of partially known signals (finite dimensional vectors) under a smoothness constraint (band-limiting) which implies the vanishing of a known subset of the samples of the DFT of the data. The signals are partially known in the sense that only a subset of its samples is supposed to be available for measurement. Our task is to find from the available samples the values of the remaining ones. This

problem presents certain similarities with the familiar sampling problem, which basically consists in restoration of a signal given a subset of its samples.

When the unknown samples are contiguously distributed this reduces to the often discussed *extrapolation problem*. It will be shown that in any case only a restriction of cardinality needs to be imposed upon the set of lost samples, and that, for any given bandwidth, there is a maximum percentage of unknown samples that can be tolerated if errorless restoration is desired, and this percentage naturally increases as the bandwidth of the signals decreases. This can be interpreted as a sampling result, asserting the possibility of errorless recovery of a band-limited signal if a sufficiently dense subset of its samples is known.

The reconstruction algorithm studied herein is an iterative method which reduces to the discrete finite-dimensional version of the well-known Papoulis-Gerchberg extrapolation algorithm [2, 3], if the unknown samples are contiguous, and if a relaxation constant μ is unitary. In the latter case, the algorithm may also be described as an alternating projection algorithm of the type found in [4]. Its convergence will be established under no restriction on the distribution of the missing samples, generalizing results found in [5] and providing a theoretical background for the observations reported in [6].

It is often convenient to have upper and lower bounds for the convergence rate of an iterative algorithm. We will give these bounds and also the particular signals for which the bounds are attained. When the missing samples are contiguously distributed, these signals reduce to the periodic discrete prolate spheroidal sequences (P-DPSS) defined in [7].

We will analyze the effect of the relaxation constant μ upon the convergence rate of the algorithm, and calculate the particular value of μ for which the asymptotic convergence rate of the algorithm is best possible.

We will also shed some light on the nature of the sampling sets for which the rate of convergence of the algorithm is best or worst possible in a certain sense. It turns out that, for the same number of unknown samples, the worst possible sampling set for low-pass or high-pass signals will nearly always be a contiguous one. On the other hand, the optimum sampling sets are those that arise when the distance between missing samples is a multiple of a certain integer, or very nearly so. For signals which are not strictly low-pass or high-pass, or, more precisely, for signals with non-contiguous passbands, this is not necessarily true. Band-pass signals provide an example of an important class of signals for which the worst possible sampling sets might not be contiguous.

We will show that for an optimum sampling set the spectral radius of the iteration matrix, which is an important indicator of the rate of convergence, is independent of the number of missing samples. In fact, a signal with several missing samples may be reconstructed as easily as a signal with a single missing sample, in the sense of asymptotic convergence, if the positions of the missing samples are properly chosen.

III. Related Work

A large percentage of the many published papers concerned with iterative constrained signal restoration is devoted to the band-limited extrapolation problem. Its study goes back to [8], where an extrapolation method is given, based on the double orthogonality property of the prolate spheroidal wave functions. The introduction of the error-reduction algorithm, also known as the Papoulis-Gerchberg iteration [2, 3], brought a new interest to the problem, and, since then, several works dealing with different aspects of the extrapolation problem have appeared. We refer the interested reader to [1, 7, 9–19] among many others. These works address topics that include implementation aspects, numerical stability, convergence acceleration, noniterative extrapolation, the sampled analog of the extrapolation problem, convergence criteria, the effect of stabilizing constraints, etc.

We will briefly review some of the results found in these works and in a few others which are more closely related to the subject that we are about to study. For a numerical comparison of the several existing extrapolation algorithms see [20, 21].

A. Band-limited L_2 and ℓ_2 signals

Many of the mappings often found in iterative constrained restoration are projections, a fact that was used [4, 11] to obtain an alternative framework for a class of restoration problems which admit an interesting geometrical interpretation [4]. In the notation of [4], x satisfies $x = P_b x$, and the observed signal is $y = P_a x$. The algorithm studied therein is based on the iteration

$$x = y + (I - P_a)P_b x.$$

It is shown in [4] that x is completely determined by $y = P_a x$ if and only if P_b and $I - P_a$ have only the zero vector in common. Furthermore, the recovery process is stable if and only if the angle between the subspaces associated with P_b and $I - P_a$ is greater than zero. The solution of the extrapolation problem in L_2 follows from these results if P_a and P_b are identified with truncation in the time and frequency domains, respectively.

The convergence rate of the Papoulis-Gerchberg iteration may be improved by reducing out-of-band noise in an adaptive way. This can be accomplished by thresholding the spectrum, a technique which is effective for signals which are linear combinations of sinusoidal functions [22].

It is also known that algorithms which perform successive projections onto convex sets may also outperform the Papoulis-Gerchberg iteration, since they allow any number of *a priori* constraints to be included in the restoration procedure. See, for example, [23–26].

An iterative algorithm for restoring continuously sampled band-limited L_2 functions was proposed in [27]. A continuously sampled signal is obtained by setting to zero a signal except for a neighborhood of a countably infinite set of sampling instants. This algorithm is related to the Papoulis-Gerchberg algorithm and to a general iteration

proposed by Sandberg [28], as discussed in [29–31]. It was shown recently [32] that the same procedure can be used in the case of instantaneous sampling.

Many discrete extrapolation problems lead to ill-conditioned systems of linear equations. Their noniterative solution is attractive from a computational point of view, but the quality of the results is often insufficient due to noise sensitivity or stability problems. Least-squares solutions with minimum norm, which can be found using singular value decomposition [33], are an interesting alternative. An unified approach to noniterative linear signal restoration can be found in [34], using Fredholm's theory of first-kind integral equations. A number of two-step reconstruction algorithms can be understood in terms of this framework. A noniterative algorithm based on a distinct principle is studied in [35].

Reference [36] deals with iterative least-squares solutions of the linear signal restoration problem, and presents an approach to the solution of this problem based on Bialy's iteration, a general procedure to obtain approximate solutions of linear operator equations in Hilbert space. A number of previously known restoration algorithms can be obtained in this way. It is also known [37] that the Papoulis-Gerchberg algorithm can be obtained from Landweber's iteration, an iterative method for the solution of first-kind Fredholm integral equations.

In [38] it is shown that time-limited restoration of shift-invariant blurred ℓ_2 signals can be done iterating an operator equation which defines a contractive mapping. The results are valid for multi-dimensional sequences with finite energy. A similar approach, based on fixed-point theorems, can be outlined under rather general conditions and leads to the constrained iterative restoration framework described in [39].

The extrapolation problem for multidimensional finite energy sequences was studied in [14], and related to the L_2 band-limited extrapolation problem. More precisely, it was shown that the discrete solution tends to the L_2 solution when the sampling rate used in the known segment of the signal tends to infinity.

B. Estimation of L_2 band-limited signals from a finite number of samples

The problem of interpolating a band-limited signal from a finite set of nonuniform samples was formulated as a least-squares optimization problem and solved using Lagrange multipliers, reproducing kernel Hilbert space theory, and other techniques [40–43]. More recently, it was shown [44] that Yen's interpolation formula [40] is a consequence of an optimization result valid in Hilbert space, and that the frequency domain characteristics of this interpolation formula are optimum in a weighted least-squares sense. A block interpolator was also proposed [44] which overcomes some of the computational difficulties found when applying Yen's formula to interpolation problems of large dimensionality.

A mathematically similar problem is addressed in [45], namely, the reconstruction of a time-limited signal from a finite number of unevenly spaced samples of its Fourier transform. It is assumed that there exists a one-to-one differentiable transformation which maps the set of sampling points into a set of equally spaced points, and both one-dimensional and two-dimensional reconstruction problems are studied. In fact, it is possible to arrive at sampling expansions for signals which are not band-limited if such a mapping exists [46]. An alternative, which sometimes leads to good results, is to interpolate a set of uniformly spaced points on the nonuniform grid [47].

The Papoulis-Gerchberg algorithm can also be used [48] to obtain a noniterative algorithm for the recovery of lost samples of oversampled functions belonging to L_2. The effect of noise and out-of-band signal components on a similar noniterative solution was studied in [49]. The algorithm discussed in [48] is also valid for L_2 functions which are integral transforms of compactly supported functions [50].

It is possible to estimate band-limited discrete signals from a finite number of samples, taking into account the length of the interval over which a good extrapolation is desired [51]. Although the result does not have minimum norm, it is optimum for signals concentrated on that interval.

The estimation of two-dimensional signals from a finite number of samples was also studied. In [52] the authors choosed to minimize the maximum of the mean-squared error over a class of signals with respect to the sampling set. They found that the optimum sampling set minimizes the spectral radius of an associated matrix, and that it is evenly spaced for rectangular passbands.

C. Finite-dimensional problems

The band-limited interpolation problem can be studied within the more general framework of constrained iterative restoration [39]. In this case it is natural to take band-limiting as the constraint and time-limiting as the distortion. The convergence of certain of these algorithms, including a class of finite-dimensional problems, was studied in [53].

There are a few papers that report the performance of iterative techniques as interpolators or estimators [6,54], but the analysis normally performed does not take full advantage of the finite dimensionality of the problem. This is an important fact since it allows some of the difficulties that arise when sampling the solution of a continuous problem to be overcome.

More recently [55], a finite-dimensional algorithm for interpolation and extrapolation was proposed which requires less *a priori* information about the bandwidth of the data, and which can be applied to mixed interpolation/extrapolation methods. The new algorithm should be particularly useful whenever precise bandwidth information about the original signals is unavailable.

The analysis of the finite-dimensional version of the Papoulis-Gerchberg algorithm was made in [5], taking advantage of the fact that the known samples are contiguous. This induces a matrix partitioning amenable to analysis. In [56], it is proved that the result of this discrete extrapolation procedure does indeed approximate the solution of

the continuous problem, when the sampling frequency and the number of samples simultaneous increase. For the noisy case see [57]. A similar result is known for multidimensional problems [58].

Finally, we point out that the alternating projection algorithm was explored as a method to compensate for some forms of distortion that result from certain data compression methods [59]. Experimental evidence gathered in [6, 54, 59] shows the usefulness of one-dimensional iterative methods of this kind.

IV. Linear Iterative Algorithms of First Order

The spectral norm of an arbitrary matrix $\mathbf{M}$, which is normally defined as

$$\|\mathbf{M}\| = \sup_{\|\mathbf{x}\|=1} \|\mathbf{M}\mathbf{x}\|,$$

is also given by

$$\|\mathbf{M}\| = \sqrt{\rho(\mathbf{M}^H\mathbf{M})}, \tag{1}$$

where $\rho(\mathbf{X})$ denotes the spectral radius of $\mathbf{X}$, that is, the greatest of its eigenvalues, in absolute value. Note that the spectral norm of a hermitian matrix equals its spectral radius.

Identity (1) can be established in different ways. One that leads to useful conclusions in the kind of problems that we will consider is the following (for an alternative proof see [60]). If we seek he stationary points of the continuous function $\phi(\mathbf{x}) = \|\mathbf{M}\mathbf{x}\|^2$, subject to $\|\mathbf{x}\| = 1$, we are lead to the equation

$$\mathbf{M}^H\mathbf{M}\mathbf{x} = \lambda\mathbf{x}, \tag{2}$$

where λ appears as a Lagrange multiplier. This shows that the unitary vectors $\mathbf{x}$ that render $\|\mathbf{M}\mathbf{x}\|$ stationary are the eigenvectors of $\mathbf{M}^H\mathbf{M}$. If $\mathbf{x}$ is an eigenvector of $\mathbf{M}^H\mathbf{M}$ then

$$\|\mathbf{M}\mathbf{x}\|^2 = \mathbf{x}^H\mathbf{M}^H\mathbf{M}\mathbf{x} = |\lambda|\,\|\mathbf{x}\|^2 = |\lambda|.$$

We are considering the restriction of the continuous function ϕ to a compact set. Thus, the absolute maximum (minimum) of $\phi(\mathbf{x})$ will be attained by the eigenvector of $\mathbf{M}^H\mathbf{M}$ that corresponds to its largest (smallest) eigenvalue in absolute value. These eigenvalues will be denoted by $\lambda_{\max}$ and $\lambda_{\min}$, respectively, and the associated eigenvectors by $\mathbf{v}_{\max}$ and $\mathbf{v}_{\min}$.

Relation (1) follows from this and the definition of norm. Note that $\|\mathbf{M}\mathbf{x}\| \leq \|\mathbf{M}\|\,\|\mathbf{x}\|$. Also, since $\mathbf{M}^H\mathbf{M}$ is hermitian, $\|\mathbf{M}\|^2 = \rho(\mathbf{M}^H\mathbf{M}) = \|\mathbf{M}^H\mathbf{M}\|$.

We recall that a (possibly nonlinear) operator A defined on a metric space M with distance function $d(\cdot,\cdot)$ is *nonexpansive* if $d(Ax, Ay) \leq d(x, y)$ for all $x, y \in M$, and *strictly nonexpansive* if equality holds only for $x = y$. In spaces endowed with a metric d induced by a norm $\|\cdot\|$ we have $d(Ax, Ay) = \|\mathbf{A}\mathbf{x} - \mathbf{A}\mathbf{y}\|$ and $d(x, y) = \|\mathbf{x} - \mathbf{y}\|$. This is the case for linear operators defined on the finite dimensional space $\mathbf{C}^n$, that is, matrices over the complex field, with the usual norm. In these cases nonexpansiveness means that for an arbitrary $\mathbf{v} \in \mathbf{C}^n$ $\|\mathbf{T}\mathbf{v}\| \leq \|\mathbf{v}\|$, and thus the spectral norm of a linear nonexpansive operator $\mathbf{T}$ cannot exceed unity.

It follows from $\rho(\mathbf{T}) \leq \|\mathbf{T}\|$ that the absolute values of the eigenvalues of a nonexpansive matrix $\mathbf{T}$ do not exceed unity. Thus, a strictly nonexpansive matrix $\mathbf{T}$ is convergent to zero (a matrix $\mathbf{T}$ is convergent to zero is if and only if its eigenvalues are all less than 1 in absolute value [60]). Note that strict nonexpansiveness is only a sufficient condition for convergence. In fact, we could easily exhibit examples of expanding linear operators, that is, matrices with norm greater than one, which nevertheless converge to zero. This can never occur for hermitian or normal matrices, whose norms and spectral radii are necessarily equal. A simple example of a matrix whose norm and spectral radius differ is given in [60].

Consider now the sequence of vectors $\mathbf{x}_i \in \mathbb{C}^n$ defined by $\mathbf{x}_{i+1} = \mathbf{M}\mathbf{x}_i + \mathbf{b}$, $\mathbf{M}$ and $\mathbf{b}$ being, respectively, a $n \times n$ matrix and a $n \times 1$ column vector. This is a model of the general linear stationary iterative algorithm of first order with *iteration matrix* $\mathbf{M}$.

Assume that $\mathbf{M}$ is convergent to zero. Then, the sequence $\mathbf{x}_i$ converges to some limit vector $\mathbf{x}$ which satisfies $\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{b}$, since we have

$$\mathbf{x} = \sum_{k=0}^{+\infty} \mathbf{M}^k\mathbf{b} = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{b}.$$

The error or residual at iteration k, $\mathbf{e}_k = \mathbf{x}_k - \mathbf{x}$, is then given by

$$\mathbf{e}_k = \mathbf{M}\mathbf{e}_{k-1},$$

being related to the initial error $\mathbf{e}_0$ by the expression $\mathbf{e}_k = \mathbf{M}^k\mathbf{e}_0$. Thus

$$\|\mathbf{e}_k\| \leq \|\mathbf{M}^k\|\,\|\mathbf{e}_0\|$$

Without imposing further restrictions on $\mathbf{M}$ one can only say that $\|\mathbf{M}^k\|$ tends to zero with k, monotonically for sufficiently high k. For hermitian, or, more generally, for normal matrices, we have

$$\|\mathbf{M}^k\| = \rho(\mathbf{M})^k,$$

and the norm of the error decreases with each iteration, that is, the iterative method has the error-reduction property, independently of k. In this case, knowledge of $\|\mathbf{M}\|$ allows a simple upper-bound for the error to be established.

This is not the case in general, and therefore $\|\mathbf{M}^k\|$ must be studied instead. The asymptotic rate of convergence of the convergent matrix $\mathbf{M}$, $R_\infty(\mathbf{M}) = \ln \rho(\mathbf{M})$, is an often useful quantity [60].

V. The Algorithms

The algorithms which will be studied below are linear iterative algorithms of the first order with non-hermitian iteration matrices. They can be derived using the general framework for constrained iterative restoration established in [39]. In this particular case, we take the sampling and band-limiting operations as, respectively, the distortion and constraint mentioned in [39]. The original signal

satisfies $\mathbf{x} = \mathbf{Bx}$ and the observed signal is $\mathbf{y} = \mathbf{Dx}$, for appropriate matrices $\mathbf{B}$ and $\mathbf{D}$. This suggests the equation

$$\begin{aligned} \mathbf{x} &= \mathbf{Bx} + \mu(\mathbf{y} - \mathbf{Dx}) \\ &= \mu\mathbf{y} + (\mathbf{I} - \mu\mathbf{D})\mathbf{Bx} \end{aligned}$$

where μ is any fixed constant, and $\mathbf{I}$ denotes the identity mapping, as well as the sequence of successive approximations

$$\mathbf{u}_{n+1} = \mu\mathbf{y} + (\mathbf{I} - \mu\mathbf{D})\mathbf{B}\mathbf{u}_n = \mathbf{T}_1\mathbf{u}_n \tag{3}$$

where

$$\mathbf{T}_1(\cdot) = \mu\mathbf{y} + (\mathbf{I} - \mu\mathbf{D})\mathbf{B}(\cdot). \tag{4}$$

It is possible to obtain a band-limited sequence of successive approximations by taking the distortion to be sampling followed by band-limiting. This gives

$$\mathbf{u}_{n+1} = \mu\mathbf{B}\mathbf{y} + (\mathbf{I} - \mu\mathbf{B}\mathbf{D})\mathbf{u}_n = \mathbf{T}_2\mathbf{u}_n \tag{5}$$

Alternatively, we may apply the band-limiting operator to (3). This leads to

$$\mathbf{u}_{n+1} = \mu\mathbf{B}\mathbf{y} + \mathbf{B}(\mathbf{I} - \mu\mathbf{D})\mathbf{u}_n = \mathbf{T}_3\mathbf{u}_n. \tag{6}$$

As we have seen, these algorithms will converge if and only if the spectral radius of the respective iteration matrices is strictly less than unity. Since the iteration matrices of the methods (3) and (6) are the transpose of each other, and (5) is equivalent to (6) if the first approximation is band-limited ($\mathbf{u}_0 = \mathbf{B}\mathbf{u}_0$), it is sufficient to consider, for example, method (3). It will be shown next that this method leads to a sequence of approximations that converges to the required solution, if suitable conditions are imposed upon $\mathbf{D}$, $\mathbf{B}$ and μ. Under the same conditions the other two iterations will also converge to the same solution, provided, of course, that $\mathbf{u}_0 = \mathbf{B}\mathbf{u}_0$ in the case of (5).

VI. Convergence of the Algorithm

A number of alternative approaches to (3) do exist. The one that we will use depends upon well-known results in fixed-point theory. This technique is often used in signal restoration [39,53] and does not assume the linearity of the operators involved. In fact, other possibly non-linear non-expansive constraints can be incorporated in the algorithm without disturbing convergence.

Denote the set of subscripts of the known samples of $\mathbf{x}$ by $I_t = \{i_1, i_2, \ldots, i_s\}$. Let the set of subscripts of the vanishing samples of the DFT of $\mathbf{x}$ be denoted by $\bar{I}_f$, the complement in E_n of the set $I_f = \{k_1, k_2, \ldots, k_q\}$.

Lemma VI.1: Let $\mathbf{D}$ be a sampling operator with density d, and let $\mathbf{B}$ be a band-limiting operator with bandwidth w. If $d \geq w$, $\mathbf{x} = (\mathbf{I} - \mathbf{D})\mathbf{x}$ and $\mathbf{x} = \mathbf{Bx}$, then $\mathbf{x} \equiv 0$.

Proof: Since $\mathbf{x} = (\mathbf{I} - \mathbf{D})\mathbf{x}$, that is, $x_i = 0$ for all $i \in I_t$, s samples of $\mathbf{x}$ vanish and the n equations $\hat{\mathbf{x}} = \mathbf{F}\mathbf{x}$ reduce to

$$\hat{\mathbf{x}} = \sum_{i \notin I_t} \mathbf{F}_i x_i,$$

where $\mathbf{F}_i$ denotes the i-th column of $\mathbf{F}$.

On the other hand, since $\mathbf{x} = \mathbf{Bx}$, that is, $\hat{x}_i = 0$ for all $i \in \bar{I}_f$, $n - q$ samples of $\hat{\mathbf{x}}$ are zero and thus

$$\sum_{i \notin I_t} F_{ki} x_i = 0, \tag{7}$$

for all $k \in \bar{I}_f$. The submatrix of $\mathbf{F}$ that appears in (7) is $(n-q)\times(n-s)$ with $(n-q) \geq (n-s)$ since $d \geq w$. Since $\bar{I}_f$ is contiguous, there is an integer r such that $k_i = i + r \bmod n$ without loss of generality. Since $F_{ki} = (1/\sqrt{n})e^{-j2\pi ki/n}$, we see that the submatrix of $\mathbf{F}$ that appears in (7) can be reduced to a Vandermonde matrix, with linear independent columns. Thus, the only solutions to (7) are given by $x_i = 0$, and therefore $\mathbf{x}$ must be the identically zero vector. ■

Lemma VI.2: Let $0 \leq \mu \leq 2$. Then, the operator $\mathbf{T}_1$ defined by (4) is nonexpansive.

Proof: Let $\mathbf{A}_\mu$ be defined as

$$\mathbf{A}_\mu = (\mathbf{I} - \mu\mathbf{D})\mathbf{B}. \tag{8}$$

It follows from (4) that the nonexpansiveness of $\mathbf{T}_1$ does not depend on $\mathbf{y}$. In fact, $\mathbf{T}_1$ is nonexpansive if and only if $\mathbf{A}_\mu$ is. The nonexpansiveness of the projection operator $\mathbf{B}$ in the Euclidean norm is obvious. Thus, $\mathbf{T}_1$ will be nonexpansive if $\mathbf{I} - \mu\mathbf{D}$ is, which happens if and only if $0 \leq \mu \leq 2$. Otherwise there would be an eigenvalue of $\mathbf{I} - \mu\mathbf{D}$ greater than one in absolute value. ■

Lemma VI.3: If $0 \leq \mu \leq 2$, the equation $\mathbf{x} = \mathbf{T}_1\mathbf{x}$, with $\mathbf{T}_1$ defined by (4), has at least one solution.

Proof: $\mathbf{T}_1$ is continuous and nonexpansive, and maps a non-empty, compact and convex subset of $\mathbb{C}^n$ into itself (consider the set of all sequences with norm less than or equal to a given bound). Therefore, it must have a fixed-point [39,61]. ■

Theorem VI.1: Let $\mathbf{D}$ be a sampling operator with density d, and let $\mathbf{B}$ be a band-limiting operator with bandwidth w. If

$$0 < \mu < 2, \tag{9}$$

$$d \geq w, \tag{10}$$

the operator $\mathbf{T}_1$, defined by (4), is strictly nonexpansive, and the equation $\mathbf{x} = \mathbf{T}_1\mathbf{x}$ has one and only one solution. This solution is given by

$$\mathbf{x} = \lim_{n\to\infty} \mathbf{T}_1^n \mathbf{x}_0,$$

for arbitrary $\mathbf{x}_0$.

Proof: Assume, to the contrary, that (9) and (10) hold without $\mathbf{T}_1$ being strictly nonexpansive. This means that

$$\|\mathbf{A}_\mu\mathbf{v}\| = \|(\mathbf{I} - \mu\mathbf{D})\mathbf{B}\mathbf{v}\| = \|\mathbf{B}\mathbf{v}\| = \|\mathbf{v}\|$$

for some $\mathbf{v} \neq 0$. We will show that this assumption leads to a contradiction. The last equality implies the vanishing of $\hat{v}_i$ for $i \in \bar{I}_f$, and thus $\mathbf{B}\mathbf{v} = \mathbf{v}$. This reduces the second equality to

$$\|(\mathbf{I} - \mu\mathbf{D})\mathbf{v}\| = \|\mathbf{v}\|.$$

But

$$\begin{aligned}\|(\mathbf{I}-\mu\mathbf{D})\mathbf{v}\|^2 &= (1-\mu)^2\sum_{i\in I_t} v_i^2 + \sum_{i\in \bar{I}_t} v_i^2 \\ &= \|\mathbf{v}\|^2 - \mu(2-\mu)\sum_{i\in I_t} v_i^2\end{aligned}$$

which implies the vanishing of the v_i for $i \in I_t$, since, by hypothesis, $0 < \mu < 2$. Applying lemma VI.1, we see that $\mathbf{v}$ must be the zero vector, a contradiction, and thus $\mathbf{T}_1$ must be strictly nonexpansive. The theorem then follows since a strictly nonexpansive mapping cannot have more than one fixed point. ■

VII. Upper and Lower Error Bounds

We will now give upper and lower bounds for the reconstruction error at any given iteration, assuming that $\mu = 1$. The bounds are best-possible and we give the initial error vectors for which they are attained.

Recalling (8), let $\mathbf{v}$ denote an eigenvector of $\mathbf{A}_1^H\mathbf{A}_1$ pertaining to the eigenvalue λ, that is,

$$\mathbf{A}_1^H\mathbf{A}_1\mathbf{v} = \mathbf{B}(\mathbf{I}-\mathbf{D})\mathbf{B}\mathbf{v} = \lambda\mathbf{v}. \tag{11}$$

We claim that $\mathbf{A}_1\mathbf{v}$ will be an eigenvector of $\mathbf{A}_1$ pertaining to the same eigenvalue, and thus

$$\mathbf{A}_1{}^2\mathbf{v} = \lambda\mathbf{A}_1\mathbf{v}.$$

To check that this is true left-multiply (11) by $\mathbf{A}_1 = (\mathbf{I}-\mathbf{D})\mathbf{B}$ and use the idempotency of $\mathbf{B}$.

It is now easy to derive bounds for the error $\mathbf{e}_k$ at iteration k of the algorithm. From the discussion in section IV it is clear that $\mathbf{e}_k = \mathbf{A}_1\mathbf{e}_{k-1}$ and thus

$$\sup_{\|\mathbf{e}_0\|=1}\|\mathbf{e}_1\|^2 = \sup_{\|\mathbf{e}_0\|=1}\|\mathbf{A}_1\mathbf{e}_0\|^2 = \lambda_{\max}$$

and

$$\inf_{\|\mathbf{e}_0\|=1}\|\mathbf{e}_1\|^2 = \inf_{\|\mathbf{e}_0\|=1}\|\mathbf{A}_1\mathbf{e}_0\|^2 = \lambda_{\min},$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are, as we have seen, the smallest and largest of the eigenvalues of $\mathbf{A}_1^H\mathbf{A}_1$. The supremum and the infimum are attained when $\mathbf{e}_0$ equals the $\mathbf{v}_{\max}$ or $\mathbf{v}_{\min}$, the eigenvectors of $\mathbf{A}_1^H\mathbf{A}_1$ that correspond to the eigenvalues $\lambda_{\max}$ and $\lambda_{\min}$, respectively.

Now, we have just seen that $\mathbf{A}_1\mathbf{v}_{\max}$ and $\mathbf{A}_1\mathbf{v}_{\min}$ will also be eigenvectors of $\mathbf{A}_1$, and therefore

$$\mathbf{e}_2 = \mathbf{A}_1{}^2\mathbf{v}_{\max} = \lambda_{\max}\mathbf{A}_1\mathbf{v}_{\max},$$

if $\mathbf{e}_0 = \mathbf{v}_{\max}$, and

$$\mathbf{e}_2 = \mathbf{A}_1{}^2\mathbf{v}_{\min} = \lambda_{\min}\mathbf{A}_1\mathbf{v}_{\min},$$

if $\mathbf{e}_0 = \mathbf{v}_{\min}$. By induction it is now clear that the inequalities

$$\|\mathbf{e}_{k+1}\| \le \lambda_{\max}^k\|\mathbf{e}_1\| \tag{12}$$

$$\|\mathbf{e}_{k+1}\| \ge \lambda_{\min}^k\|\mathbf{e}_1\|, \tag{13}$$

hold. Equality is attained if $\mathbf{e}_1 = \mathbf{v}_{\max}$, in case of (12), or $\mathbf{e}_1 = \mathbf{v}_{\min}$, in case of (13).

In many cases one takes $\mathbf{x}_0$ to be the zero vector, which gives $\mathbf{x}_1 = \mathbf{y}$, and $\mathbf{e}_k = \mathbf{A}_1^k\mathbf{x}$. Taking $\mathbf{x} = \mathbf{v}_{\max}$ then results in the lowest possible asymptotic convergence rate. In this sense, the $\mathbf{v}_{\max}$ are the signals which result in the worst possible algorithm performance. A similar argument, regarding best possible performance, holds for $\mathbf{v}_{\min}$.

For pure extrapolation of low-pass signals and $\mu = 1$, this analysis is related to the singular value analysis of the operator $\mathbf{DB}$ done in [7], and to the periodic discrete prolate spheroidal sequences (P-DPSS) discussed therein. Note that setting $\mathbf{M} = \mathbf{A}_1$ in equation (2) gives

$$\mathbf{B}(\mathbf{I}-\mathbf{D})\mathbf{B}\mathbf{v} = \lambda\mathbf{v}, \tag{14}$$

which turns out to be equivalent to

$$\mathbf{B}\mathbf{D}\mathbf{B}\mathbf{v} = (1-\lambda)\mathbf{v} \tag{15}$$

since any solution of (14) must be band-limited (left multiply (14) by $\mathbf{B}$). Equation (15) is equivalent to the equation used to define the P-DPSS in [7]. The eigenvectors of (14) $\mathbf{v}_{\min}$ and $\mathbf{v}_{\max}$, which correspond to the smallest and largest eigenvalues $\lambda_{\min}$ and $\lambda_{\max}$ in (14), are therefore examples of P-DPSS.

VIII. The Optimum Relaxation Constant

Let us now study the effect of the parameter μ upon the convergence rate of the algorithm.

Let $\mathbf{S}_\mu = \mathbf{B}(\mathbf{I}-\mu\mathbf{D})\mathbf{B}$. We claim that if $\mathbf{v}$ is an eigenvector of $\mathbf{S}_\mu$ pertaining to the eigenvalue λ, then $\mathbf{A}_\mu\mathbf{v}$ is an eigenvector of $\mathbf{A}_\mu$ pertaining to the same eigenvalue. This follows from the idempotency of $\mathbf{B}$, which implies

$$\mathbf{A}_\mu\mathbf{S}_\mu = (\mathbf{I}-\mu\mathbf{D})\mathbf{B}^2(\mathbf{I}-\mu\mathbf{D})\mathbf{B} = \mathbf{A}_\mu^2.$$

Therefore if $\mathbf{S}_\mu\mathbf{v} = \lambda\mathbf{v}$ then $\mathbf{A}_\mu{}^2\mathbf{v} = \lambda\mathbf{A}_\mu\mathbf{v}$. This shows that every eigenvalue of $\mathbf{S}_\mu$ is also an eigenvalue of $\mathbf{A}_\mu$, and thus $\rho(\mathbf{S}_\mu) \le \rho(\mathbf{A}_\mu)$.

On the other hand, if $\mathbf{v}$ is an eigenvector of $\mathbf{A}_\mu$ pertaining to the eigenvalue λ, then $\mathbf{B}\mathbf{v}$ is an eigenvector of $\mathbf{S}_\mu$ pertaining to the same eigenvalue. To check this, we left multiply $\mathbf{A}_\mu\mathbf{v} = \lambda\mathbf{v}$ by $\mathbf{B}$, obtaining

$$\mathbf{B}(\mathbf{I}-\mu\mathbf{D})\mathbf{B}\mathbf{v} = \lambda\mathbf{B}\mathbf{v},$$

which is equivalent to

$$\mathbf{S}_\mu\mathbf{B}\mathbf{v} = \lambda\mathbf{B}\mathbf{v},$$

again using the idempotency of $\mathbf{B}$. This shows that $\rho(\mathbf{A}_\mu) \le \rho(\mathbf{S}_\mu)$. Since we have already shown the converse, it follows that $\rho(\mathbf{A}_\mu) = \rho(\mathbf{S}_\mu)$.

The effect of the parameter μ upon the convergence rate of the algorithm may be judged from its effect upon the spectral radius $\rho(\mathbf{A}_\mu)$ of the iteration matrix. Since $\rho(\mathbf{A}_\mu) = \rho(\mathbf{S}_\mu)$, it is sufficient to compare the spectral radii of the matrices $\mathbf{S}_\mu$ and $\mathbf{S}_1$.

With this in mind, we will show that if $\mathbf{v}$ and λ are, respectively, an eigenvector/eigenvalue pair of $\mathbf{S}_1$, then $\mathbf{v}$ and $1-\mu(1-\lambda)$ are an eigenvector/eigenvalue pair of $\mathbf{S}_\mu$.

Again, $\mathbf{S}_1\mathbf{v} = \lambda\mathbf{v}$ implies $\mathbf{B}\mathbf{v} = \mathbf{v}$ (left multiply by $\mathbf{B}$ and use idempotency), and thus we have

$$\mathbf{B}(\mathbf{I}-\mathbf{D})\mathbf{B}\mathbf{v} = \mathbf{v} - \mathbf{B}\mathbf{D}\mathbf{B}\mathbf{v} = \lambda\mathbf{v}.$$

Consequently, $\mathbf{v}$ is an eigenvector of $\mathbf{BDB}$ pertaining to the eigenvalue $1-\lambda$. This implies

$$\mathbf{B}(\mathbf{I}-\mu\mathbf{D})\mathbf{B}\mathbf{v} = \mathbf{v} - \mu\mathbf{B}\mathbf{D}\mathbf{B}\mathbf{v} = \mathbf{v} - \mu(1-\lambda)\mathbf{v}$$

as we wanted to show.

The graphic $\{\lambda, f(\lambda)\}$ of $f(\lambda) = 1-\mu(1-\lambda)$ $(0 \le \lambda \le 1)$ is a straight line of positive slope connecting the points $\{0, 1-\mu\}$ and $\{1, 1\}$. This mapping allows the determination of the optimal value of μ. In the following discussion, we take $\mu = 1$ as the reference value. Note that decreasing μ below 1 is uninteresting since it increases the eigenvalues of the iteration matrix. On the contrary, increasing μ towards 2 may reduce the spectral radius and therefore lead to better convergence rates.

Assume that $\mu > 1$, and let $\lambda_{\min}$ and $\lambda_{\max}$ be the smallest and largest eigenvalues of $\mathbf{S}_1$, respectively. Note that $\lambda_{\min} = 0$ and $\lambda_{\max} < 1$ under the conditions of theorem VI.1. The eigenvalues of $\mathbf{S}_\mu$ will therefore lie in the interval $[1-\mu, 1-\mu(1-\lambda_{\max})]$, the spectral radius of $\mathbf{A}_\mu$ being given by

$$\rho(\mathbf{A}_\mu) = \max_{1<\mu<2}\{\mu - 1, 1-\mu(1-\lambda_{\max})\}.$$

This expression is minimized when $\mu - 1 = 1-\mu(1-\lambda_{\max})$. Solving we find the optimal value of μ,

$$\mu_{\text{opt}} = \frac{2}{2-\lambda_{\max}}.$$

For $\mu = \mu_{\text{opt}}$ the spectral radius of the iteration matrix is

$$\rho(\mathbf{A}_\mu) = \frac{\lambda_{\max}}{2-\lambda_{\max}} = \frac{\rho(\mathbf{A}_1)}{2-\rho(\mathbf{A}_1)},$$

which may result in considerably better convergence rates.

To summarize, reducing μ below 1 should be avoided since it decreases the convergence rate. When μ increases between 1 and μ_{opt}, the spectral radius of the iteration matrix decreases, and therefore the asymptotic convergence rate is improved. Increasing μ above μ_{opt} should be avoided since it again reduces the convergence rate and will ultimately lead to divergence. Note, however, that $0 < \lambda_{\max} < 1$ implies $1 < \mu_{\text{opt}} < 2$ and $0 < \rho(\mathbf{A}_\mu) < 1$, which means that for μ just greater than μ_{opt} the method will still be convergent.

IX. Sampling Sets and Convergence Rate

We will now study the norm and spectral radius of the matrix $\mathbf{A}_1$ defined by (8). We would like to know which sampling sets minimize or maximize the asymptotic convergence rate of the algorithm, for a given density. We will see that the answer depends on the $\mathbf{B}$ matrix. For low-pass data, for example, the missing sample positions that give best-possible sampling sets define a grid with a certain spacing. Thus, the optimum sampling strategy is, in a sense, uniform sampling. On the other hand, the worst possible sampling sets are contiguous. Again, this may not be the case for data neither low-pass nor high-pass.

We start by noting that the general inequality $\|\mathbf{M}\| \ge \rho(\mathbf{M})$ can be made precise if $\mathbf{M} = \mathbf{A}_1$. First, recall from section VIII that $\rho(\mathbf{A}_\mu) = \rho(\mathbf{S}_\mu)$. When $\mu = 1$ we see that $\mathbf{A}_1^H\mathbf{A}_1 = \mathbf{S}_1$, and therefore, using (1), we conclude that

$$\|\mathbf{A}_1\|^2 = \rho(\mathbf{A}_1^H\mathbf{A}_1) = \rho(\mathbf{A}_1). \tag{16}$$

In the following we will study the norm of $\mathbf{A}_1$ instead of its eigenvalues. However, because of (16), knowledge of the norm of $\mathbf{A}_1$ is equivalent to knowledge of its spectral radius, and therefore of its asymptotic rate of convergence. This can be of help in understanding the effect of the sampling sets upon the convergence properties of the algorithm.

As before, let s be the number of known samples, and let their subscripts be $I_t = \{i_1, i_2, \ldots, i_s\}$.

We will not impose any contiguity constraints upon the band-limiting matrix $\mathbf{B}$, which is only assumed to have bandwidth q/n (see section I for the meaning of q). We denote the columns of $\mathbf{B}$ by $\mathbf{b}_i$, $i = 0, 1, \ldots, n-1$.

Since the lines of $\mathbf{A}_\mu$ are given by

$$(1-\mu D_{ii})\mathbf{b}_i^H,$$

the components of the vector $\mathbf{A}_\mu\mathbf{x}$ are

$$(1-\mu D_{ii})\langle \mathbf{b}_i, \mathbf{x}\rangle.$$

It follows that

$$\begin{aligned}
\mathbf{x}^H{\mathbf{A}_\mu}^H\mathbf{A}_\mu\mathbf{x} &= \|\mathbf{A}_\mu\mathbf{x}\|^2 \\
&= \sum_{i=0}^{n-1}(1-\mu D_{ii})^2\langle \mathbf{b}_i, \mathbf{x}\rangle^2 \\
&= \sum_{i\notin I_t}\langle \mathbf{b}_i, \mathbf{x}\rangle^2 + (1-\mu)^2\sum_{i\in I_t}\langle \mathbf{b}_i, \mathbf{x}\rangle^2.
\end{aligned} \tag{17}$$

From this equation several conclusions can be drawn. We start by examining a few simple consequences.

A. The Minimum Norm of the Iteration Matrix

Recall from section IV that $\|\mathbf{A}_\mu\|^2$ equals the spectral radius of the hermitian matrix $\mathbf{A}_\mu^H\mathbf{A}_\mu$, which, in turn, can be obtained radius of the hermitian matrix $\mathbf{A}_1^H\mathbf{A}_1$, which, in turn, can be obtained by taking the supremum of (17) over all $\mathbf{x}$ of unit norm. From (17) it follows that this quantity attains its least possible value when $\mu = 1$. Thus, the norm of $\mathbf{A}_\mu$ is minimized when the relaxation constant μ is unitary.

A second proof could be given based on the results of section VIII.

B. One Missing Sample

When only one sample is missing and $\mu = 1$, $\bar{I}_t$ is reduced to one element. Denoting it by i, we have

$$\|\mathbf{A}_1\mathbf{x}\|^2 = \langle \mathbf{b}_i, \mathbf{x}\rangle^2.$$

The supremum of this expression over all $\mathbf{x}$ with unit norm will be attained for $\mathbf{x} = \mathbf{b}_i/\|\mathbf{b}_i\|$. Thus

$$\|\mathbf{A}_1\|^2 = \|\mathbf{b}_i\|^2 = q/n,$$

where q/n is just the normalized bandwidth of the band-limiting matrix. Thus $\mathbf{A}_1$ will be a linear contractive operator for any $q < n$.

It follows from (16) that the spectral radius of $\mathbf{A}_1$ equals the normalized bandwidth of $\mathbf{B}$. This can be checked easily, noting that the iteration matrix $\mathbf{A}_1$ will have, in this case, only one nonzero row. Therefore it may have one nonzero eigenvalue at most. Since the trace of $\mathbf{A}_1 = (\mathbf{I} - \mathbf{D})\mathbf{B}$ is given, in this case, by $B_{00} = q/n$, we may immediately conclude that the nonzero eigenvalue of $\mathbf{A}_1$, and consequently its spectral radius, is also given by q/n and equals the squared norm of $\mathbf{A}_1$.

Note that each iteration of the reconstruction algorithm will reduce the squared error by a factor of at least q/n. Using the optimum relaxation constant $\mu = \mu_{\text{opt}} = 2n/(2n - q)$ this can be improved to $q/(2n - q)$.

C. Two Missing Samples

Suppose now that μ still equals 1 but that two samples are missing, that is, $\bar{I}_t = \{i, j\}$. In this case we have

$$\|\mathbf{A}_1\mathbf{x}\|^2 = \langle \mathbf{b}_i, \mathbf{x}\rangle^2 + \langle \mathbf{b}_j, \mathbf{x}\rangle^2.$$

The stationary points of this quadratic form are the eigenvectors of the matrix $\mathbf{A}_1$, which in this case has only two nonzero rows, rows i and j, given by $\mathbf{b}_i^H$ and $\mathbf{b}_j^H$, respectively. In fact

$$\mathbf{A}_1^H\mathbf{A}_1 = \mathbf{b}_i\mathbf{b}_i^H + \mathbf{b}_j\mathbf{b}_j^H.$$

It is easy to find the eigenvectors and eigenvalues of such a matrix. It can be directly verified that, in the real case, the quantities

$$\|\mathbf{b}_i\|^2 \pm |\langle \mathbf{b}_i, \mathbf{b}_j\rangle| = \frac{q}{n} \pm |\langle \mathbf{b}_i, \mathbf{b}_j\rangle|$$

are eigenvalues of $\mathbf{A}_1^H\mathbf{A}_1$ pertaining to the orthogonal eigenvectors

$$\mathbf{b}_i \pm \mathbf{b}_j$$

(consider $\mathbf{A}_1^H\mathbf{A}_1\mathbf{b}_i$ and $\mathbf{A}_1^H\mathbf{A}_1\mathbf{b}_j$). Since the trace of $\mathbf{A}_1^H\mathbf{A}_1$ is $2\|\mathbf{b}_i\|^2 = 2q/n$, and since $\mathbf{A}_1^H\mathbf{A}_1$ cannot have any negative eigenvalues, we conclude that

$$\|\mathbf{A}_1\|^2 = \frac{q}{n} + |\langle \mathbf{b}_i, \mathbf{b}_j\rangle|.$$

The inner product $\langle \mathbf{b}_i, \mathbf{b}_j\rangle$ can be put into a simpler form. Considering that each $\mathbf{b}_i$ is a column of $\mathbf{B}$, and that $\mathbf{B}$ is idempotent, that is, $\mathbf{B}^2 = \mathbf{B}$, we see that

$$\langle \mathbf{b}_i, \mathbf{b}_j\rangle = \mathbf{b}_p(q) = \mathbf{b}_q(p),$$

where $\mathbf{b}_i(k)$ denotes the k-th element of the i-th column of $\mathbf{B}$. This result may also be easily obtained by direct calculation.

Using the fact that $\mathbf{B}$ is circulant we have

$$\mathbf{b}_p(q) = \mathbf{b}_q(p) = \mathbf{b}_0(p \ominus q),$$

where $\ominus$ denotes subtraction modulo n. Therefore

$$\|\mathbf{A}_1\|^2 = \frac{q}{n} + |\mathbf{b}_0(i \ominus j)|.$$

When the set of vanishing samples of the DFT of the signal is contiguous, and in particular for low-pass or high-pass real band-limiting matrices, we have

$$|\langle \mathbf{b}_i, \mathbf{b}_k\rangle| = |\mathbf{b}_0(i \ominus k)| = \left|\frac{\sin[\pi q(i \ominus k)/n]}{n \sin[\pi(i \ominus k)/n]}\right|. \qquad (18)$$

This provides a convenient way of evaluating the squared norm and spectral radius of the iteration matrix as a function of the lag between missing samples.

D. Missing Sample Distributions that Minimize or Maximize the Norm of the Iteration Matrix

When only two samples are missing it is possible to obtain the missing sample distributions that minimize (maximize) the norm and spectral radius of the iteration matrix $\mathbf{A}_1$ simply by choosing the lag $i \ominus j$ between samples in such a way that $|\mathbf{b}_0(i \ominus j)|$ is minimum (maximum).

For signals with q contiguous nonzero harmonics, an hypothesis satisfied by low-pass or high-pass signals, for example, the first row of $\mathbf{B}$ will consist of samples of the modulus of the Dirichlet kernel, as seen in (18). In this case it can be verified numerically that

$$|\mathbf{b}_0(k)| \le |\mathbf{b}_0(1)| \qquad k = 1, 2, \ldots, \left\lfloor \frac{n}{2} \right\rfloor,$$

for $1 < q < n - 1$ for realistic values of n. The worst case situation corresponds, in these cases, to contiguous missing samples, that is, $i \ominus j = 1$. See figure 1, where the norm of the iteration matrix is depicted as a function of the lag between missing samples, for low-pass (or high-pass) real signals (it follows from the symmetry of $\mathbf{b}_0(x)$ that only the first $\lfloor n/2 \rfloor$ possible lags need to be considered).

For band-pass signals, for example, the worst case situation may no longer correspond to a contiguous sampling set (see figure 2).

Let us now consider a number of missing samples greater than 2. Again, let s/n be the density of the sampling matrix, and $n - s$ the number of unknown samples. What can we say about the best and worst possible missing sample distributions?

As we have seen, the squared norm of $\mathbf{A}_\mu$ is given by

$$\sup_{\|x\|=1} \left(\sum_{i \notin I_t} \langle \mathbf{b}_i, \mathbf{x}\rangle^2 + (1 - \mu)^2 \sum_{i \in I_t} \langle \mathbf{b}_i, \mathbf{x}\rangle^2 \right). \qquad (19)$$

There are two extreme situations. Roughly speaking, the first and more favorable arises when all of vectors $\mathbf{b}_i$ with

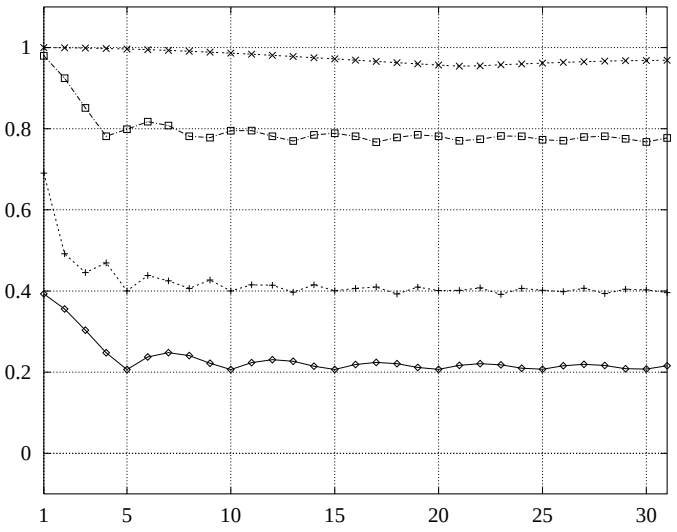


Fig. 1. Norm of the iteration matrix for real low-pass signals ($n = 64$) with $2m + 1$ nonzero harmonics ($m = 30, 24, 12$ and 6, from top to bottom) as a function of the lag between unknown samples.

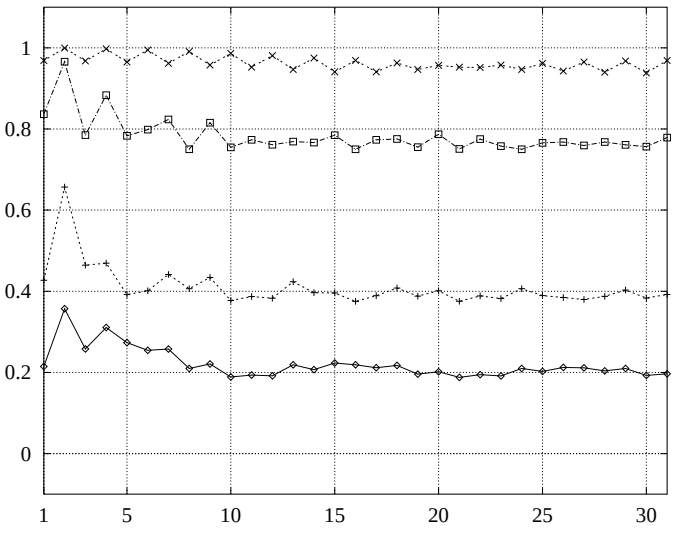


Fig. 2. Norm of the iteration matrix for real band-pass signals ($n = 64$) with $2m$ nonzero harmonics ($m = 30, 24, 12$ and 6, from top to bottom) as a function of the lag between unknown samples.

$i \notin I_t$ are orthogonal, or very nearly so. The worst possible situation occurs when the $\mathbf{b}_i$ with $i \notin I_t$ are very nearly linearly dependent.

Assume that all vectors $\mathbf{b}_i$ with $i \notin I_t$ are orthogonal, and let X be the subspace spanned by them. Since they all have squared norm equal to q/n, we may write

$$\mathbf{x} = \frac{n}{q} \sum_{i \notin I_t} \langle \mathbf{b}_i, \mathbf{x} \rangle \mathbf{b}_i$$

for any $\mathbf{x}$ belonging to X. Thus

$$\|\mathbf{x}\|^2 = \frac{n}{q} \sum_{i \notin I_t} \langle \mathbf{b}_i, \mathbf{x} \rangle^2,$$

or, equivalently,

$$\sum_{i \notin I_t} \langle \mathbf{b}_i, \mathbf{x} \rangle^2 = \|\mathbf{x}\|^2 \frac{q}{n}.$$

It follows from (19) with $\mu = 1$ that

$$\|\mathbf{A}_1\|^2 = \sup_{\|x\|=1} \frac{q}{n} \|\mathbf{x}\|^2 = \frac{q}{n},$$

that is, $\mathbf{A}_1$ will be a convergent matrix with squared norm q/n. Note that in this case the norm is independent of the number of missing samples, and equals the value obtained for a single missing sample. Clearly, if the missing samples are distributed in this way, the recovery of one, two, or more missing samples can be performed with a matrix whose norm is determined only by the bandwidth of the data.

Also, the asymptotic convergence rate of the reconstruction algorithm will be independent of the number of missing samples, provided that the lag between missing samples is adequately chosen. We assume, of course, that the conditions of theorem VI.1 are satisfied, that is, the sampling set has a sufficiently high density.

For signals with a contiguous set of nonzero harmonics, and in particular for low-pass signals, it follows from (18) that $|\mathbf{b}_0(k)|$ can be zero only for the following values of k,

$$\frac{n}{q}, \frac{2n}{q}, \frac{3n}{q}, \ldots$$

assuming that q divides n. These quantities determine the possible positions of the missing samples which correspond to optimum sampling sets. If q does not divide n, they may still originate sampling sets with near-optimum convergence properties.

X. Examples

In this section we present the results of computer simulations that illustrate and confirm the results presented so far.

The reconstruction algorithm described in section V was applied to two reconstruction problems. The input signal for the first example is depicted in figure 3. It is a randomly generated vector of 64 samples, band-limited with a low-pass filter of normalized bandwidth 0.52. The (randomly generated) sampler vector is depicted in figure 4. It has a normalized density of 0.69. Since its normalized density exceeds the normalized bandwidth of the input data, we may expect the reconstruction procedure to converge.

The optimum value for the relaxation constant was found from the results of section VIII, and the algorithm was run twice, once with $\mu = 1$ and another with $\mu = \mu_{\text{opt}}$. In both cases we took the observed signal as the first approximation. The resulting error curves are shown in figure 5. As expected, the asymptotic rate of convergence is best in the latter case.

Figure 6 depicts the input data vector which minimizes the asymptotic convergence rate of the algorithm, for $\mu = 1$, the sampler depicted in figure 4, and frequency domain

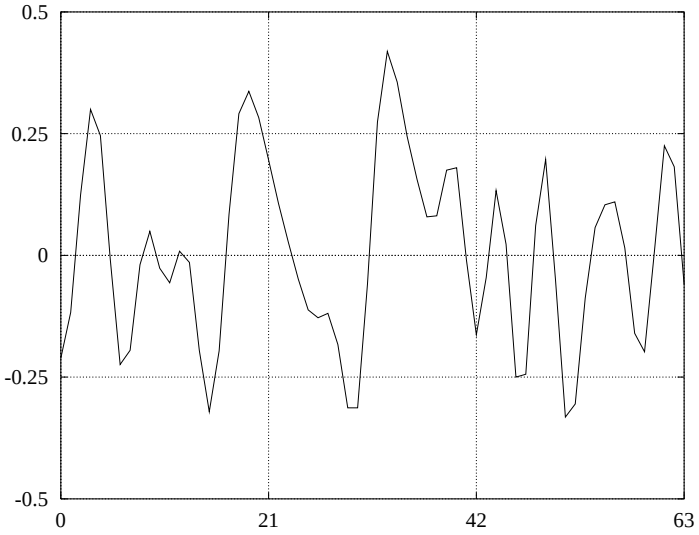


Fig. 3. Input signal. Randomly generated low-pass signal ($n = 64$) with 33 nonzero harmonics (normalized bandwidth of approximately 0.52).

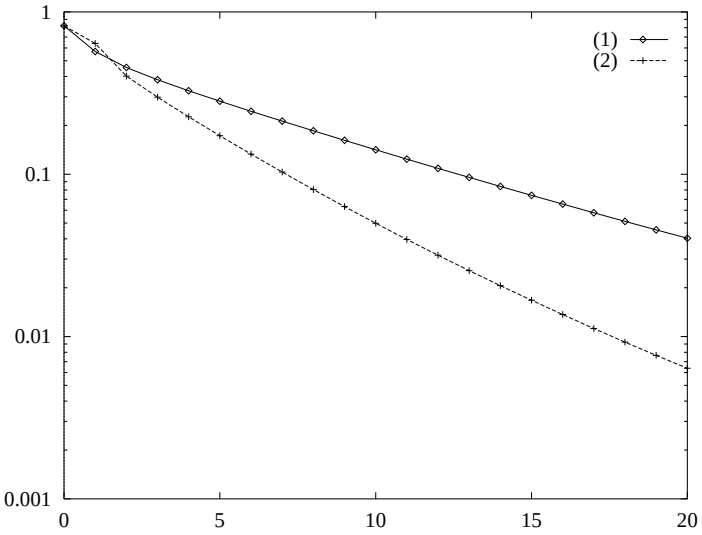


Fig. 5. Euclidean norm of the error as a function of the number of iterations. (1) $\mu = 1$, (2) $\mu = \mu_{\text{opt}}$.

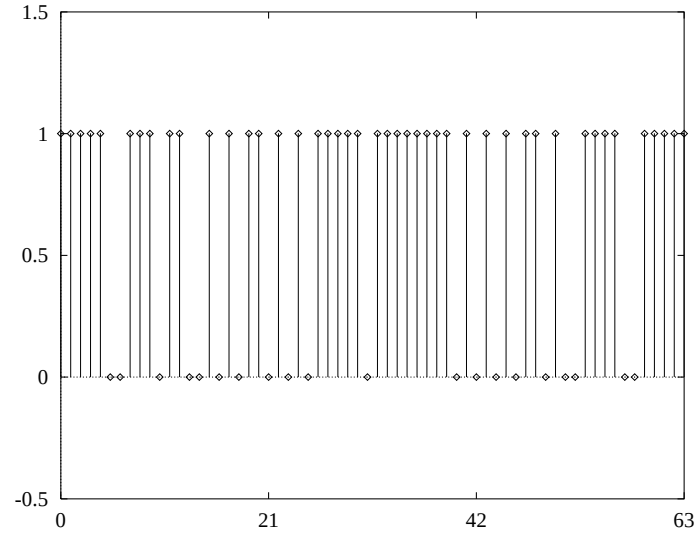


Fig. 4. Sampler. Randomly generated, $n = 64$, 20 zeros (normalized density of approximately 0.69).

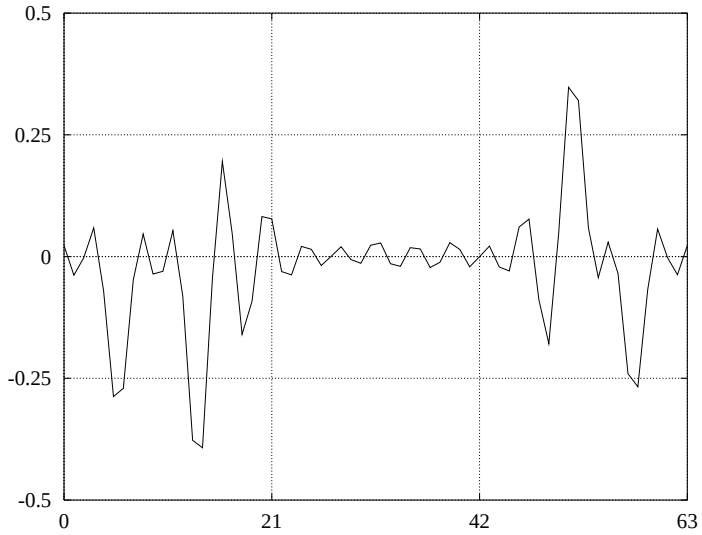


Fig. 6. Eigenvector $\mathbf{v}_{\max}$ that corresponds to the sampler depicted in figure 4.

characteristics equal to those of the signal depicted in figure 3. It is the eigenvector $\mathbf{v}_{\max}$ of $\mathbf{S}_1$ which pertains to its largest eigenvalue.

We ran the algorithm twice using this vector as input, once with $\mu = 1$ and another with $\mu = \mu_{\text{opt}}$. The error evolution may be found in figure 7. Again, the algorithm performed as expected. Note the improvement in convergence rate in the case $\mu = \mu_{\text{opt}}$.

In all cases, the reconstructed signals were visually indistinguishable from the originals, on the scale of the figure.

XI. Conclusions

We studied the convergence properties of a finite-dimensional iterative reconstruction algorithm which can be useful for the solution of certain interpolation and extrapolation problems. We gave best-possible lower and upper bounds for the error as a function of the number of iterations, and the signals for which the bounds are exact. We analyzed the effect of the relaxation constant μ and established its optimum value. Starting with the simplest possible cases, we elaborated on the best and worst possible sampling sets, in the sense of best and worst possible asymptotic convergence rates. Finally, we illustrated the theoretical results with the results of a computer simulation.

This work complements existing works where the same or similar algorithms are used and shown to perform well, but which do not include a detailed theoretical analysis.

Although the algorithm is capable of interesting performance under certain circumstances, there are classes of problems for which it does not seem very well suited. We point out, for example, the finite-dimensional low-pass extrapolation problem as an example. In this case, and, more

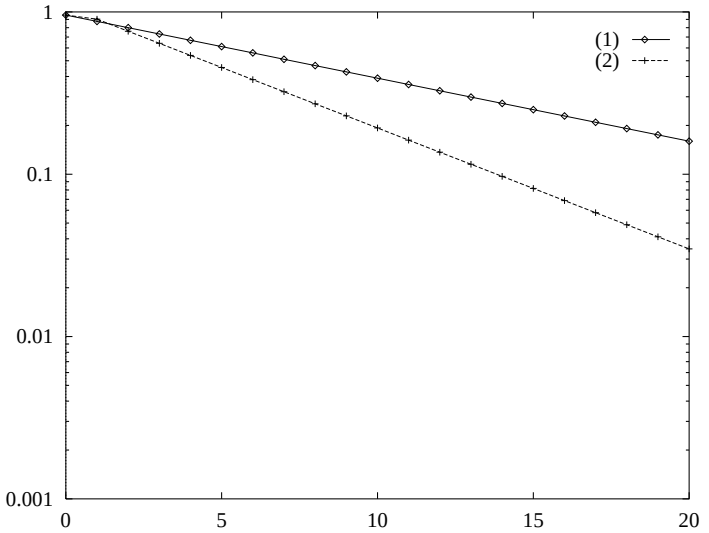


Fig. 7. Euclidean norm of the error as a function of the number of iterations. (1) $\mu = 1$, (2) $\mu = \mu_{\text{opt}}$.

generally, whenever there are long contiguous gaps of missing samples, the convergence rate may fall to quite low levels. We hope to address this problem in future works.

References

[1] J. L. C. Sanz and T. S. Huang, "Some aspects of band-limited signal extrapolation: Models, discrete approximations, and noise", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 6, pp. 1492–1501, Dec. 1983.

[2] A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation", *IEEE Transactions on Circuits and Systems*, vol. 22, no. 9, pp. 735–742, Sept. 1975.

[3] R. W. Gerchberg, "Super resolution through error energy reduction", *Opt. Acta*, vol. 21, no. 9, pp. 709–720, 1974.

[4] D. C. Youla, "Generalized image restoration by the method of alternating orthogonal projections", *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp. 694–702, Sept. 1978.

[5] M. C. Jones, "The discrete Gerchberg algorithm", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 3, pp. 624–626, June 1986.

[6] P. S. Naidu and B. Paramasivaiah, "Estimation of sinusoids from incomplete time series", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 3, pp. 559–562, June 1984.

[7] A. K. Jain and S. Ranganath, "Extrapolation algorithms for discrete signals with application in spectral estimation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 4, pp. 830–845, Aug. 1981.

[8] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty — I", *The Bell System Technical Journal*, vol. 40, no. 1, pp. 43–63, Jan. 1961.

[9] G. A. Viano, "On the extrapolation of optical image data", *Journal of Mathematical Physics*, vol. 17, no. 7, pp. 1160–1165, July 1976.

[10] M. S. Sabri and W. Steenaart, "An approach to band-limited signal extrapolation: The extrapolation matrix", *IEEE Transactions on Circuits and Systems*, vol. 25, no. 2, pp. 74–78, Feb. 1978.

[11] J. A. Cadzow, "An extrapolation procedure for band-limited signals", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no. 1, pp. 4–12, Feb. 1979.

[12] W. Y. Xu and C. Chamzas, "On the extrapolation of band-limited functions with energy constraints", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 5, pp. 1222–1234, Oct. 1983.

[13] H. J. Trussel, "Convergence criteria for iterative restoration methods", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 1, pp. 129–136, Feb. 1983.

[14] J. L. C. Sanz and T. S. Huang, "Discrete and continuous band-limited signal extrapolation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 5, pp. 1276–1285, Oct. 1983.

[15] C. C. Chamzas and W. Y. Xu, "An improved version of Papoulis-Gerchberg algorithm on band-limited extrapolation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 2, pp. 437–440, Apr. 1984.

[16] C. Yeh and R. T. Chin, "Error analysis of a class of constrained iterative restoration algorithms", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 6, pp. 1593–1598, Dec. 1985.

[17] S. Singh, S. N. Tandon, and H. M. Gupta, "An iterative restoration technique", *Signal Processing*, vol. 11, pp. 1–11, 1986.

[18] C. Cenker, H. G. Feichtinger, and H. Steier, "Fast iterative and non-iterative reconstruction of band-limited functions from irregular sampling values", in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Toronto, May 1991.

[19] P. J. S. G. Ferreira, "Noniterative and faster iterative methods for interpolation and extrapolation", submitted to *IEEE Transactions on Signal Processing*.

[20] T. S. Huang, J. L. C. Sanz, H. Fan, J. Shafii, and B.-M. Tsai, "Numerical comparison of several algorithms for band-limited signal extrapolation", *Applied Optics*, vol. 23, no. 2, pp. 307–317, Jan. 1984.

[21] C. Cenker, H. G. Feichtinger, and M. Herrmann, "Iterative algorithms in irregular sampling: a first comparison of methods", in *Proceedings of the 10th IEEE IPCCC*, Scottsdale, Mar. 1991, pp. 483–489.

[22] A. Papoulis and C. Chamzas, "Detection of hidden periodicities by adaptive extrapolation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 27, no. 5, pp. 492–500, Oct. 1979.

[23] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1 — theory", *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 81–94, Oct. 1982.

[24] M. I. Sezan and H. Stark, "Image restoration by the method of convex projections: Part 2 — applications and numerical results", *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 95–101, Oct. 1982.

[25] H. J. Trussel and M. R. Civanlar, "The feasible solution in signal restoration", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 2, pp. 201–212, Apr. 1984.

[26] H. J. Trussel and M. R. Civanlar, "The Landweber iteration and projection onto convex sets", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 6, pp. 1632–1634, Dec. 1985.

[27] R. G. Wiley, "Recovery of bandlimited signals from unequally spaced samples", *IEEE Transactions on Communications*, vol. 26, no. 1, pp. 135–137, Jan. 1978.

[28] I. W. Sandberg, "On the properties of some systems that distort signals — I", *The Bell System Technical Journal*, vol. 42, pp. 2033–2046, Sept. 1963.

[29] R. G. Wiley, "On an iterative technique for recovery of bandlimited signals", *Proceedings of the IEEE*, vol. 66, no. 4, pp. 522–523, Apr. 1978.

[30] R. G. Wiley, "Concerning the recovery of a bandlimited signal or its spectrum from a finite segment", *IEEE Transactions on Communications*, vol. 27, no. 1, pp. 251–252, Jan. 1979.

[31] I. W. Sandberg, "A note on the *PQ* theorem and the extrapolation of signals", *IEEE Transactions on Signal Processing*, vol. 41, no. 6, pp. 2259–2261, June 1993.

[32] F. A. Marvasti, M. Analoui, and M. Gamshadzahi, "Recovery of signals from nonuniform samples using iterative methods", *IEEE Transactions on Signal Processing*, vol. 39, no. 4, pp. 872–878, Apr. 1991.

[33] B. J. Sullivan and B. Liu, "On the use of singular value decomposition and decimation in discrete-time band-limited signal extrapolation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 6, pp. 1201–1212, Dec. 1984.

[34] J. L. C. Sanz and T. S. Huang, "A unified approach to noniterative linear signal restoration", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 2, pp. 403–409, Apr. 1984.

[35] F. M. Marvasti, "Reconstruction of speech signals with lost samples", *IEEE Transactions on Signal Processing*, vol. 40, no. 12, pp. 2897–2903, Dec. 1992.

[36] J. L. C. Sanz and T. S. Huang, "Unified Hilbert space approach to iterative least-squares linear signal restoration", *Journal of the Optical Society of America*, vol. 73, no. 11, pp. 1455–1465, Nov. 1983.

[37] J. L. Sanz and T. S. Huang, "On the Gerchberg-Papoulis algorithm", *IEEE Transactions on Circuits and Systems*, vol. 30, no. 12, pp. 907–908, Dec. 1983.

[38] J. L. Sanz and T. S. Huang, "Iterative time-limited signal reconstruction", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 3, pp. 643–649, June 1983.

[39] R. W. Schafer, R. M. Mersereau, and M. A. Richards, "Constrained iterative restoration algorithms", *Proceedings of the IEEE*, vol. 69, no. 4, pp. 432–450, Apr. 1981.

[40] J. L. Yen, "On nonuniform sampling of bandwidth-limited signals", *IRE Transactions on Circuit Theory*, vol. 3, pp. 251–257, Dec. 1956.

[41] L. Levi, "Fitting a bandlimited signal to given points", *IEEE Transactions on Information Theory*, pp. 372–376, July 1965.

[42] K. Yao, "Applications of reproducing kernel Hilbert spaces — bandlimited signal models", *Information and Control*, vol. 11, pp. 429–444, 1967.

[43] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty — V: The discrete case", *The Bell System Technical Journal*, vol. 57, no. 5, pp. 1371–1429, May 1978.

[44] G. Calvagno and D. C. Munson, Jr., "New results on Yen's approach to interpolation from nonuniformly spaced samples", in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Apr. 1990, pp. 1535–1538.

[45] M. Soumekh, "Band-limited interpolation from unevenly spaced sampled data", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 1, pp. 110–122, Jan. 1988.

[46] J. J. Clark, R. P. Palmer, and P. D. Lawrence, "A transformation method for the reconstruction of functions from nonuniformly spaced samples", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 5, pp. 1151–1165, Oct. 1985.

[47] H. J. Trussell, L. L. Arnder, P. R. Moran, and R. C. Williams, "Corrections for nonuniform sampling distortions in magnetic resonance imagery", *IEEE Transactions on Medical Imaging*, vol. 7, no. 1, pp. 32–44, Mar. 1988.

[48] R. J. Marks II, "Restoring lost samples from an oversampled band-limited signal", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 3, pp. 752–755, June 1983.

[49] P. Delsarte, A. J. E. M. Janssen, and L. B. Vries, "Discrete prolate spheroidal wave functions and interpolation", *SIAM Journal of Applied Mathematics*, vol. 45, no. 4, pp. 641–650, Aug. 1985.

[50] P. J. S. G. Ferreira, "Incomplete sampling series and the recovery of missing samples from oversampled band-limited signals", *IEEE Transactions on Signal Processing*, vol. 40, no. 1, pp. 225–227, Jan. 1992.

[51] D. P. Kolba and T. W. Parks, "Optimal estimation for band-limited signals including time domain considerations", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 1, pp. 113–122, Feb. 1983.

[52] D. S. Chen and J. P. Allebach, "Analysis of error in reconstruction of two-dimensional signals from irregularly spaced samples", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 2, pp. 173–180, Feb. 1987.

[53] V. T. Tom, T. F. Quatieri, M. H. Hayes, and J. H. McClellan, "Convergence of iterative nonexpansive signal reconstruction algorithms", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 5, pp. 1052–1058, Oct. 1981.

[54] F. A. Marvasti, "An iterative method to compensate for the interpolation distortion", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 10, pp. 1617–1621, Oct. 1989.

[55] S. D. Cabrera and T. W. Parks, "Extrapolation and spectral estimation with iterative weighted norm modification", *IEEE Transactions on Signal Processing*, vol. 39, no. 4, pp. 842–851, Apr. 1991.

[56] H.-J. Schlebusch and W. Splettstösser, "On a conjecture of J. L. C. Sanz and T. S. Huang", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 6, pp. 1628–1630, Dec. 1985.

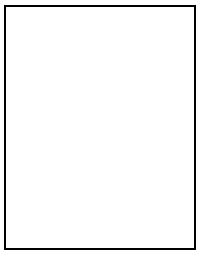
[57] X. Zhou and X. Xia, "A Sanz-Huang conjecture on band-limiting extrapolation with noise", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 9, pp. 1468–1472, Sept. 1989.

[58] X. Zhou and X. Xia, "The extrapolation of high-dimensional band-limited functions", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 10, pp. 1576–1580, Oct. 1989.

[59] P. J. S. G. Ferreira, "Reconstrução de sinais a partir de informação parcial", Provas de aptidão pedagógica e capacidade científica, Departamento de Electrónica e Telecomunicações, Universidade de Aveiro, Apr. 1988.

[60] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[61] D. R. Smart, *Fixed Point Theorems*, Cambridge University Press, Cambridge, 1980.



Paulo J. S. G. Ferreira was born in Torres Novas, Portugal, and received the Ph.D. degree in Electrical Engineering from Universidade de Aveiro, Portugal. He is presently with the Departamento de Electrónica e Telecomunicações, Universidade de Aveiro. His research interests are in the area of signal processing, with emphasis on sampling theory and constrained iterative and non-iterative restoration. He is especially interested in the theory and applications of linear algebra as well as in applications of numerical and functional analysis to signal and systems theory.