A Time Decoding Realization with a CNN

Aurel A. Lazar*, Tamás Roska†, Ernő K. Simonyi‡ and László T. Tóth§

* Department of Electrical Engineering, Columbia University,
NY 10027, New York, USA, e-mail: aurel@ee.columbia.edu
† Faculty of Information Technology, Pázmány Péter Catholic University,
H-1052 Budapest, Hungary, e-mail: roska@sztaki.hu
‡ National Council of Hungary for Information & Communication Technology,
H-1073 Budapest, Hungary, e-mail: simonyi@nhit.hu
§ Department of Telecommunications and Media Informatics,
Budapest University of Technology and Economics,
H-1117 Budapest, Hungary, e-mail: tothl@tmit.bme.hu

***Abstract*— Time encoding is a novel real-time asynchronous mechanism for encoding amplitude information into a time sequence. The analog bandlimited input is fed into a simple nonlinear neuron-like circuit that generates a strictly increasing time sequence based on which the signal can be reconstructed. The heart of the reconstruction is solving a system of ill-conditioned linear equations. This contribution shows that the equations can be manipulated so that the reconstruction becomes feasible using a Cellular Neural Network (CNN) with a banded system matrix. In particular, the system is first transformed into a well-conditioned smaller system; and then, the Lanczos process is used to lay it out into a set of even smaller systems characterized by a set of tridiagonal matrices. Each of these systems can directly be solved by CNNs, whereas the preprocessing (transformation and Lanczos algorithm) and simple postprocessing phases can be partly or fully implemented by using the digital capabilities of the CNN Universal Machine (CNN-UM). Each step of the proposed formulation is confirmed by numerical (digital) simulations [1].**

I. Introduction

Section I-A gives a short overview of irregular sampling, a topic that is closely related to time encoding and is also used later on. Section I-B summarizes the basic concept of time encoding. A numerical example is presented in Section I-C. Section I-D reviews how related problems are solved by CNN. Section I-E shows how linear equations are solved using neural networks and what extra conditions are needed for a CNN implementation. Section II presents the proposed formulation.

A. Irregular Sampling

Let the analog signal $x(t)$ be bandlimited to $[-\Omega, \Omega]$. The classical sampling theorem ([14], [24]) calls for representing $x(t)$ based on its samples taken uniformly at or above the Nyquist rate. Researchers have long been fascinated of how uniform (traditional) sampling can be generalized. Starting from early achievements [8] substantial results have been accumulated over the years both in theory [26], [12] and efficient numerical solutions [9], [10], just to mention a few. If s_k, is a strictly increasing time sequence, a possible representation is given by

$$x(t) = \sum_{k=-\infty}^{\infty} c_k g(t - s_k) \tag{1}$$

where c_k's are appropriate scalars and $g(t) = \sin(\Omega t)/(\pi t)$ is the impulse response of an ideal lowpass filter with bandwidth Ω. Evaluating both sides of (1) at $t = s_\ell$ gives

$$x(s_\ell) = \sum_{k=-\infty}^{\infty} c_k g(s_\ell - s_k), \tag{2}$$

or in matrix form

$$\mathbf{q} = \mathbf{G}\mathbf{c} \tag{3}$$

with unknown vector $[\mathbf{c}]_k = c_k$ and known vector $\mathbf{q}$ and matrix $\mathbf{G}$ given by:

$$[\mathbf{q}]_\ell = x(s_\ell), \quad \text{and} \quad [\mathbf{G}]_{\ell,k} = g(s_\ell - s_k), \tag{4}$$

If the average density of the s_k's is at or above the Nyquist rate then the unknown $\mathbf{c}$ can in principle be obtained from (3) [12]. The practical solution, however, is challenging because

- $\mathbf{G}$ is typically ill-conditioned,
- $\mathbf{q}$, $\mathbf{c}$, $\mathbf{G}$ should in principle have infinite dimensions.

B. Time encoding

Although a large number of theoretical papers have been published in the area of irregular sampling, its practical applications are limited to a few areas including astronomical measurements, medical imaging, and the lost-data problem in communication theory [1]. The use of irregular sampling in communications is even more limited because of two main reasons. First, as mentioned in Section I-A, the reconstruction for $\mathbf{c}$ based on (3) is difficult. Second, both the times s_k's and the amplitudes $x(s_k)$'s are needed for the reconstruction. Therefore, if the average density of the s_k's are at the Nyquist-rate, then two times as much information is needed for transmission as that in the case of regular sampling where $s_{k+1} - s_k$ is fixed for any k. Thus in the regular sampling

[1] Proceedings of the Seventh Seminar on Neural Network Applications in Electrical Engineering, NEUREL 2004, pp. 97-102, Belgrade, Serbia and Montenegro, September 23-25, 2004.

case, the s_k's carry no information and, therefore, do not need to be transmitted.

A technique related to pulse position modulation (PPM) addressed this problem as follows (see, e.g., [21]). The s_k's are defined as the zero-crossings of the difference signal $x(t) - f(t)$ where $f(t)$ is known. Therefore, $x(s_k)$ no longer carries information since $x(s_k) = f(s_k)$ holds and all the information is carried by the s_k's. A simple implementation can be given by feeding the difference $x(t) - f(t)$ into a comparator that fires at exactly $t = s_k$. By choosing $f(t)$ properly (usually as a periodic ramp or sinusoid signal with appropriate amplitude and frequency) the required density of the s_k's can be guaranteed. Still, the method has not became popular because the high sensitivity to noise and the exact shape of $f(t)$ that VLSI implementations cannot guarantee with high accuracy.

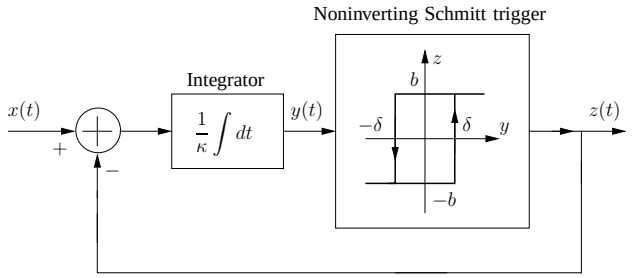

Fig. 1. A possible robust TEM configuration.

An improved time encoding scheme, referred to as Time Encoding Machine (TEM), is depicted in Figure 1 [16]. As shown, the TEM consists of an ideal integrator and a noninverting Schmitt trigger in a feedback arrangement. The output $z(t)$ takes b or $-b$ values at transition times denoted by t_k. It can be shown [16] that this circuit can also be described by (1) and (3) where we have

$$[\mathbf{q}]_\ell = (-1)^k \left[2\kappa\delta - b(t_{k+1} - t_k\right], \tag{5}$$

$$[\mathbf{G}]_{\ell,k} = \int_{t_\ell}^{t_{\ell+1}} g(t - s_k)dt, \quad \text{and} \quad s_k = \frac{t_k + t_{k+1}}{2} \tag{6}$$

with the same condition for the density of the t_k's as that for the s_k's in the case of irregular sampling. This circuit can be shown to be robust in terms of additive noise and circuit imperfections. Intuitively this makes sense, since replacing the Schmitt trigger by a clocked quantizer (a comparator followed or preceded by a sampler) a popular robust circuit, a first order Sigma-Delta modulator is obtained [22]. Although the circuit parameters κ and δ appear in the expression for $\mathbf{q}$ (see (5)), with alternative formulations [18] this dependence can also be eliminated, and $x(t)/b$ can be recovered in terms of the time differences $t_{k+1} - t_k$ only. Due to lack of space and in order to keep the formulation relatively simple, however, subsequently (3) together with (5) and (6) will be used. Following the terminology introduced in [16] the overall reconstruction procedure/implementation will be referred to as the Time Decoding Machine (TDM).

The formulation in (3), (5) and (6) assumes that the dimensionality of the matrices and vectors used is infinite. In practice (and simulations), however, only a finite time window can be used, when the right-hand-side (RHS) of (1) merely approximates $x(t)$ on its left-hand-side (LHS). Therefore, if L denotes the number of the t_k's within some known observation interval T_{obs}, then

$$x(t) \simeq \sum_{k=0}^{L-1} c_k g(t - s_k), \tag{7}$$

where $\mathbf{q}$ and $\mathbf{c}$ in (3) have L components, and $\mathbf{G}$ is an L by L matrix. Let the reconstruction error, the difference between the RHS and the LHS of (7), and its root-mean-square (RMS) value be denoted by $e(t)$ and $\mathcal{E}$, respectively. It can be shown [16], [18] that $e(t)$ can be decreased by increasing L for a given T_{obs} that in turn results in ill-conditioned $\mathbf{G}$ as illustrated in the example of I-C below:

C. Example

Let $x(t)$ be given by its familiar sampling representation

$$x(t) = \sum_{k=-\infty}^{\infty} x(kT) \frac{\sin(\Omega(t - kT)}{\Omega(t - kT)}$$

where $T = \pi/\Omega$ is the Nyquist-period, and the samples $x(T)$ through $x(12T)$, are respectively, -0.1961, 0.186965, 0.207271, 0.0987736, -0.275572, 0.0201665, 0.290247, 0.138374, -0.067588, -0.145661, -0.11133, -0.291498, $x(kT) = 0$, for $k \leq 0$ and $k > 12$. The rest of

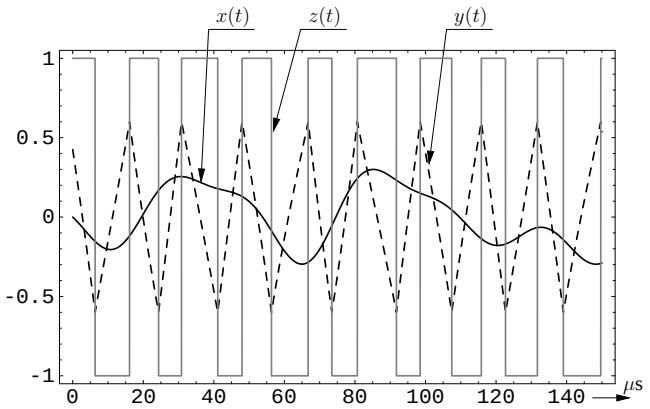

Fig. 2. Simulation results for the signals in Fig. 1.

parameters are $b = 1$, $\delta = 0.6$, $\kappa = 6.667$ μs, $\Omega = 2\pi \cdot 40$ kHz hence $T = 12.5$ μs. The evaluation of the t_k's was carried out in the interval $-2T \leq t \leq 15T$ ($T_{\text{obs}} = 17T$) based on the numerical simulation of the TEM in Figure 1. Figure 2 shows the input $x(t)$, the integrator output $y(t)$, and the TEM output $z(t)$. In this example we have

$$L = 26 \tag{8}$$

number of t_k's (only 18 are shown in Fig. 2) determined with high accuracy, when (using the spectral norm, $\|\cdot\|_2$) the

condition number of $\mathbf{G}$ turned out to be:

$$\text{cond}(\mathbf{G}) = 1.476 \times 10^{11} = 233 \text{ dB} \tag{9}$$

Since such an ill-conditioned matrix cannot be inverted directly, $\mathbf{c}$ in (3) was calculated by [25]

$$\mathbf{c} = \mathbf{G}^{+}\mathbf{q} \tag{10}$$

where $\mathbf{G}^{+}$ stands for the pseudo (Moore-Penrose) [2] inverse of $\mathbf{G}$. Using "perfect" t_k's the simulation result is shown by the solid trace of Figure 3.

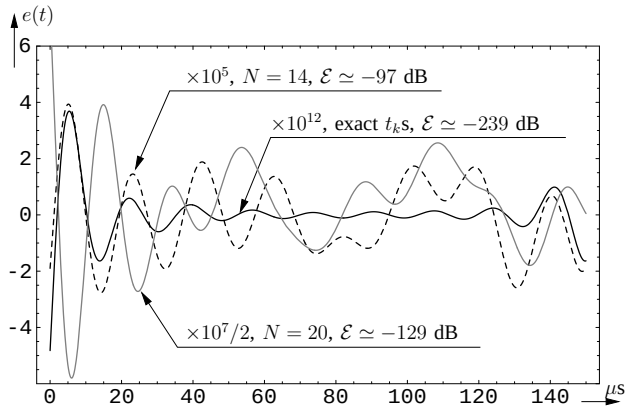

Fig. 3. Simulation results for the error function, $e(t)$, with perfect t_k's and t_k's quantized to $N = 14$ and $N = 20$ bits.

The extremely small RMS error, $\mathcal{E}$, cannot be achieved if the t_k's are inaccurate. In particular, if time encoding is used in telecommunications, then only quantized t_k's can be transmitted. If $x(t) < c$ holds for some given c, then the difference $t_{k+1} - t_k$ turns out to be bounded [16] as:

$$\frac{2\kappa\delta}{b+c} \le t_{k+1} - t_k \le \frac{2\kappa\delta}{b-c} \tag{11}$$

Therefore, if $2\kappa\delta/(b-c)$ is known, only the difference of the upper and lower bound in (11) needs to be quantized. In this way, if N bits are used for time quantization, then the corresponding error of $t_{k+1} - t_k - 2\kappa\delta/(b+c)$ is in the range:

$$\Delta = \frac{1}{2^N}\left(\frac{2\kappa\delta}{b-c} - \frac{2\kappa\delta}{b+c}\right) = \frac{2\kappa\delta}{2^N}\frac{2c}{b^2-c^2}$$

The quantization of $t_{k+1} - t_k - 2\kappa\delta/(b+c)$ can be implemented by a single clock with frequency $1/\Delta$ started at t_k and stopped after the appearance of t_{k+1}. From these values the (erroneous) t_k's, that are needed for generating the matrix $\mathbf{G}$ (see (6)), can easily be calculated. Figure 3 shows the corresponding simulation results with $c = 0.3$ (see $x(t)$ in Figure 2) for $N = 14$ and $N = 20$. As seen, with time quantization the error signals are substantially increased as opposed to using perfect t_k's (machine-precision with 64 bits).

[2]If $\mathbf{G}$ is nonsingular (appropriately well-conditioned), then $\mathbf{G}^{+} = \mathbf{G}^{-1}$.

D. Neural networks for time encoding

After the appearance of the TEM in Section I-B, similar circuits have been disclosed. The TEM investigated in [17] consists of an integrate-and-fire neuron with an absolute refractory period. Although the refractory period seems intuitively to lead to information loss, it is shown in [17] that under simple conditions, bandlimited signals encoded with an integrate and fire neuron can be perfectly recovered from the neural spike train at its output.

Many time-coding related problems can be solved via the CNN Universal Machine (CNN-UM) [5]. One particularly interesting problem is the hyperacuity in time mechanisms [20] with sparse global line in each row. This is a key feature of the interaural time difference computing mechanism in the barn owl [13]. In addition, the CNN dynamics is described by a system of nonlinear differential equations characterized by a band matrix. Its solution with programmable parameters can be made, for a 128×128 system, in a few microseconds. This means, that for a properly formulated TDM problem the CNN-UM based visual microprocessor chip [23] might serve as an ideal physical implementation. If some additional terms are needed in the feedback matrix [4], a few special, additional terms could be added, to make it a special purpose TDM chip.

E. Neural networks for solving linear equations

Based on (3) the heart of the reconstruction in time encoding and irregular sampling is to solve a set of ill-conditioned linear equations. This and related problems have been addressed by using neural networks. In [6] neuron-like processors, in essence classical analog computers, were proposed for the real-time solution of ill-conditioned linear equations. Further generalizations and simplifications were presented for solving linear least squares problems in [7] including new on-chip adaptive learning algorithms. These methods, however, are not directly applicable for CNN solutions where one cell is allowed to be connected mainly to its neighbors. In terms of linear equations this means that the system matrix has to have a banded structure: the matrix elements are zero apart from its main and a few neighboring diagonals. For time encoding the matrix $\mathbf{G}$ is dense. As it is shown below, recent results for CNNs [2] can be also used for implementing time decoding. Thus, a complete CNN or a CNN-convertible approach can be designed to perform a full time encoding and decoding procedure between analog bandlimited input and output points [2].

II. A CNN-COMPATIBLE TIME DECODING

Section II-A gives a general overview of the proposed reconstruction techniques while Section II-B discusses the details.

A. General scheme and CNN considerations

The original time encoding and decoding scheme [16] is depicted in Figure 4, where the TEM's output is represented by irregularly spaced L pulses with common (unity) weights. In agreement with (10) and (7), the reconstruction is based on calculating the pseudo-inverse of the dense and ill-conditioned

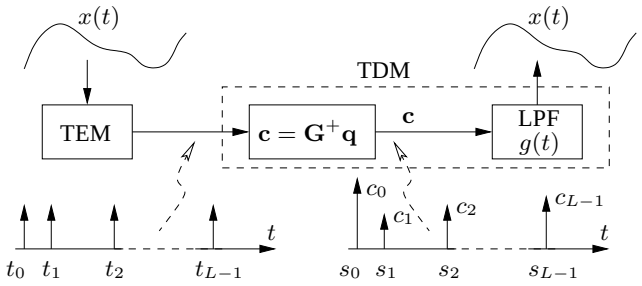

Fig. 4. Original reconstruction [16].

L by L matrix $\mathbf{G}$; and the output signal is obtained by passing *irregularly* spaced pulses with the weights c_k through an ideal lowpass filter (LPF) with impulse response $g(t)$.

Figure 5 depicts the general scheme of the proposed formulation. The Parallel Algorithm block first reduces the original L equations to M, ($M < L$) well-conditioned equations. Then, using the Lanczos algorithm the K by K matrices $\mathbf{T}_1, \ldots, \mathbf{T}_M$ are generated so that each has a tridiagonal form (see also (25)) and reduced size ($K < M$). In terms of these matrices M linear equations $\mathbf{T}_n\mathbf{x}_n = \mathbf{e}_1$ have to be solved for $\mathbf{x}_n$ where $n = 0, \ldots, M-1$ and $\mathbf{e}_1$ is the first column of a K by K identity matrix, as seen. These computations are well-suited for parallel processing, typically for the structural needs of CNNs. Also, as opposed to the original reconstruction in Figure 4, $x(t)$ is now recovered by passing *regularly* spaced pulses through the LPF. This clearly has practical advantages. The computations of the Parallel Algorithm and the Postprocessing blocks can be implemented digitally using dedicated circuitry and/or the digital capabilities of CNNs by placing them into the Global Analog Control Unit (GACU).

In particular, the CNN $\mathbf{A}$ matrix in (8) of [4] and the equivalent matrix $\mathbf{T}$ in (25) of this paper have the same structure. If we are using a spatially variant form of the simple CNN with a template $[\rho, \alpha, \beta]$, then ρ, α, and β are locally variant. In that case a special GACU for the additional computation, i. e. a special-purpose CNN-UM chip can be built. Details of the required accuracy and algorithmic control are to be developed. The use of the CNN-UM chip with a combined DSP to be integrated into a single chip is particularly challenging.

B. Parallel Algorithm

As illustrated in I-C the L by L matrix $\mathbf{G}$ in (3) is ill-conditioned. The first idea of the proposed formulation is to express the functions $g(t - s_k)$ by using oversampled and regularly spaced sampling times as

$$g(t - s_k) \simeq \sum_{n=0}^{M-1} F_{n,k}\, g(t - nT_s) \tag{12}$$

with given M and appropriate coefficients $F_{n,k}$. Since $g(t-s_k)$ is approximated by similar functions, $g(t-nT_s)$, intuitively it makes sense to expect a good approximation for $M < L$. In terms of M and T_{obs} the sampling period T_s is set to $T_s = T_{\text{obs}}/M$. Substituting (12) into (7) gives

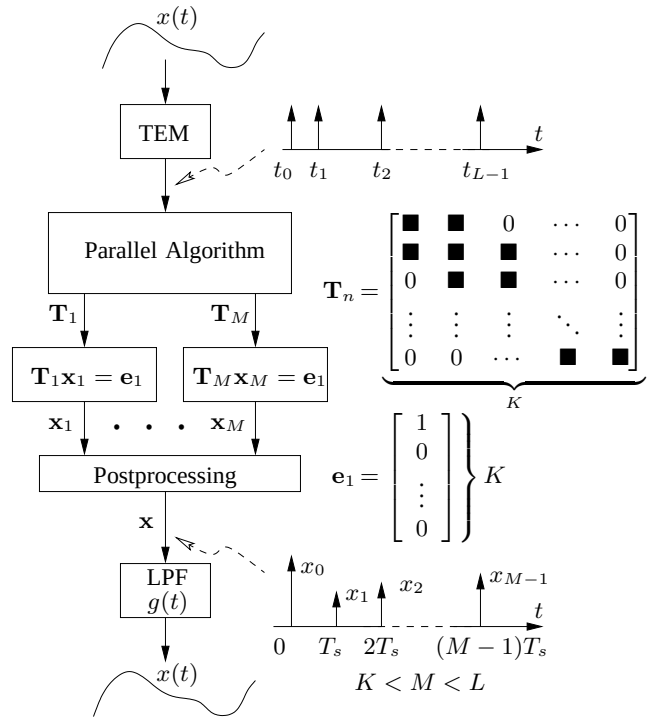

Fig. 5. Proposed reconstruction.

$$x(t) \simeq \sum_{n=0}^{M-1} x_n g(t - nT_s), \quad x_n = [\mathbf{x}]_n, \quad \mathbf{x} = \mathbf{F}\mathbf{c} \tag{13}$$

where $[\mathbf{F}]_{n,k} = F_{n,k}$ and $\mathbf{c}$ was introduced in (3). Evaluating both sides of (12) at $t = mT_s$ gives $\mathbf{D} \simeq \mathbf{B}\mathbf{F}$ with

$$[\mathbf{D}]_{m,k} = g(mT_s - s_k) \quad \text{and} \quad [\mathbf{B}]_{m,n} = g((m-n)T_s). \tag{14}$$

Therefore, with appropriate $M < L$ we have:

$$\mathbf{F} \simeq \mathbf{B}^{-1}\mathbf{D} \tag{15}$$

Substituting now (12) into (6) gives

$$\mathbf{G} \simeq \hat{\mathbf{D}}\mathbf{F} \quad \text{with} \quad [\hat{\mathbf{D}}]_{k,n} = \int_{t_k}^{t_{k+1}} g(t - nT_s)dt. \tag{16}$$

Therefore, (3) and the third relationship in (13) imply $\mathbf{q} = \mathbf{G}\mathbf{c} \simeq \hat{\mathbf{D}}\mathbf{F}\mathbf{c} = \hat{\mathbf{D}}\mathbf{x}$. Multiplying both sides by $\mathbf{D}$ gives $\mathbf{D}\mathbf{q} \simeq \mathbf{D}\hat{\mathbf{D}}\mathbf{x}$, from which we have $\mathbf{x} \simeq (\mathbf{D}\hat{\mathbf{D}})^{-1}\mathbf{D}\mathbf{q}$, since $\mathbf{D}\hat{\mathbf{D}}$ is a quadratic (M by M) matrix with smaller size and improved condition number as opposed to $\mathbf{G}$. The condition can further be improved by defining

$$\hat{\mathbf{F}} \simeq \hat{\mathbf{D}}\mathbf{B}^{-1} \tag{17}$$

when using (15) gives

$$\mathbf{x} \simeq \mathbf{B}^{-1}(\mathbf{F}\hat{\mathbf{F}})^{-1}\mathbf{F}\mathbf{q} \tag{18}$$

as the solution for the transformed system in (13).

Based on numerical simulations we found that cond($\mathbf{F}\hat{\mathbf{F}}$) is substantially smaller than cond($\mathbf{D}\hat{\mathbf{D}}$). Intuitively, this is

because $\mathbf{B}$ is structurally similar not only to $\mathbf{D}$ but also to $\hat{\mathbf{D}}$ as well due to the mean value theorem, $[\hat{\mathbf{D}}]_{k,n} = (t_{k+1}-t_k)g(\xi_k - nT_s)$ with appropriate $\xi_k \in [t_k, t_{k+1}]$. In this way, since $\mathbf{D}\hat{\mathbf{D}} = \mathbf{B}\mathbf{F}\hat{\mathbf{F}}\mathbf{B}$, the multiplication of the inverse of $\mathbf{B}$ acts as a preconditioner on both sides of $\mathbf{D}\hat{\mathbf{D}}$.

Since $\mathbf{B}$ is a fixed symmetric Toeplitz matrix [3], [19], its inverse can be precalculated with arbitrary accuracy and stored; and only the multiplications in (15) and (17) are to be carried out on-line. In addition, not only the inverse but other parametrized forms of $\mathbf{B}$ can be calculated beforehand. For example, using its eigenvalue decomposition $\mathbf{B} = \mathbf{U}\text{diag}(\lambda_k)\mathbf{U}^{-1}$ with modal matrix $\mathbf{U}$ and eigenvalues λ_k a possible generalization of $\mathbf{B}^{-1}$ is

$$\mathbf{X} = \mathbf{U}\text{diag}(\lambda_k^{-1})^{\alpha}\mathbf{U}^{-1}, \tag{19}$$

and therefore

$$\mathbf{x} \simeq \mathbf{X}(\mathbf{F}\hat{\mathbf{F}})^{-1}\mathbf{F}\mathbf{q} \tag{20}$$

where

$$\mathbf{F} = \mathbf{X}\mathbf{D} \quad \text{and} \quad \hat{\mathbf{F}} = \hat{\mathbf{D}}\mathbf{X}. \tag{21}$$

Choosing $\alpha = 1$ gives $\mathbf{X} = \mathbf{B}^{-1}$, hence the solution for $\mathbf{x}$ in (20) becomes the same as that in (18). The example in II-B.1 below demonstrates that cond$(\mathbf{F}\hat{\mathbf{F}})$ can further be improved by appropriately choosing the parameter α without affecting the RMS error as opposed to the case obtained by $\alpha = 1$.

The rest of the formulation is the direct application of the method developed by Feldmann and Freund [11] for implementing an efficient algorithm for computing the Padé approximation of Laplace-domain transfer functions via the Lanczos process. As shown in the Appendix, the coefficients x_n in (13) are given by

$$x_n = \mathbf{l}_n^T\mathbf{r}\mathbf{e}_1^T\mathbf{T}_n^{-1}\mathbf{e}_1 \tag{22}$$

where $\mathbf{e}_1$ was defined earlier (see also Figure 5). Similarly, denoting the n-th column of an M by M identity matrix by $\hat{\mathbf{e}}_n$ the parameters in (22) are

$$\mathbf{l}_n^T = \hat{\mathbf{e}}_n^T\mathbf{X}, \quad \mathbf{A} = \mathbf{F}\hat{\mathbf{F}}, \quad \mathbf{r} = \mathbf{F}\mathbf{q}, \tag{23}$$

where $\mathbf{T}_n$ are K by K ($K \leq M$) tridiagonal matrices introduced in Figure 5. The Appendix contains the relationships (Lanczos Algorithm) for calculating the elements of $\mathbf{T}_n$ in (25).

As a result, the Postprocessing block of Figure 5 calculates $x_n = \mathbf{l}_n^T\mathbf{r}\mathbf{e}_1^T\mathbf{x}_n$ where $\mathbf{x}_n$ is the solution of $\mathbf{T}_n\mathbf{x}_n = \mathbf{e}_1$. The example in II-B.1 below shows that K can be substantially less than M without noticeably increasing the RMS error, $\mathcal{E}$.

1) Example: Using the parameters and t_k's of the example in I-C numerical simulations were carried out to determine the RMS recovery error, $\mathcal{E}$, and cond$(\mathbf{F}\hat{\mathbf{F}})$. Figure 6 shows the results obtained for different values by M and $\alpha = 1$. For $M = 22$ and different values of α the simulation results for cond$(\mathbf{F}\hat{\mathbf{F}})$ are shown in Figure 7. As shown, in agreement with these figures, for $M = 22$ and $\alpha = 1$ we have cond$(\mathbf{F}\hat{\mathbf{F}}) = 739.4 \simeq 57.4$ dB, whereas with $M = 22$ and $\alpha = 0.8$ cond$(\mathbf{F}\hat{\mathbf{F}}) = 35.36 \simeq 31$ dB.

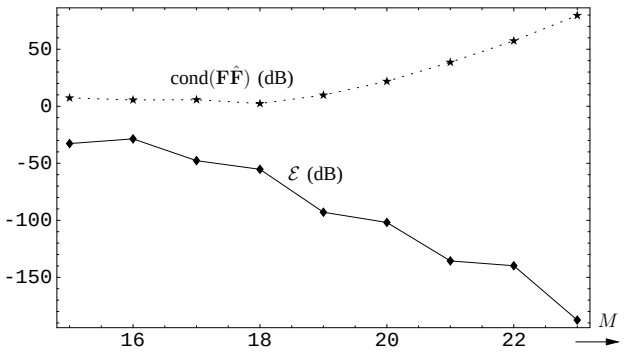

Fig. 6. Simulation results for $\mathcal{E}$ and cond$(\mathbf{F}\hat{\mathbf{F}})$ with $\mathbf{X} = \mathbf{B}^{-1}$ ($\alpha = 1$) in terms of M.

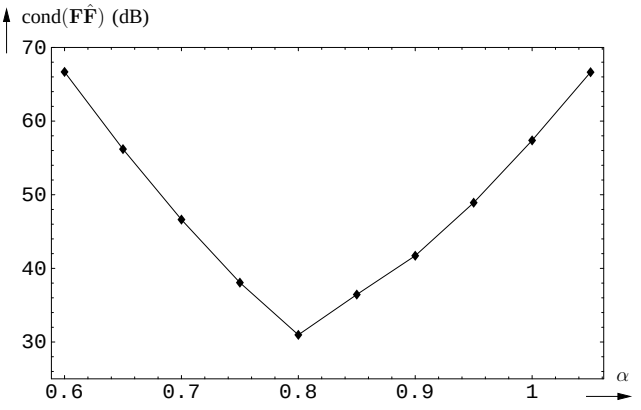

Fig. 7. Simulation results for cond$(\mathbf{F}\hat{\mathbf{F}})$ with $M = 22$ and different values of α in (19).

Finally, for $M = 22$ and $\alpha = 0.8$ (see Figure 7), the matrices $\mathbf{T}_1, \ldots, \mathbf{T}_M$ were generated for several K, and the corresponding RMS error was calculated. As shown in Figure 8, the size of $\mathbf{T}_1, \ldots, \mathbf{T}_M$ can be reduced from 22 to 12 resulting in practically the same $\mathcal{E}$. The RMS error obtained for $K \geq 12$ matches with that shown in Figure 6 for $M = 22$.

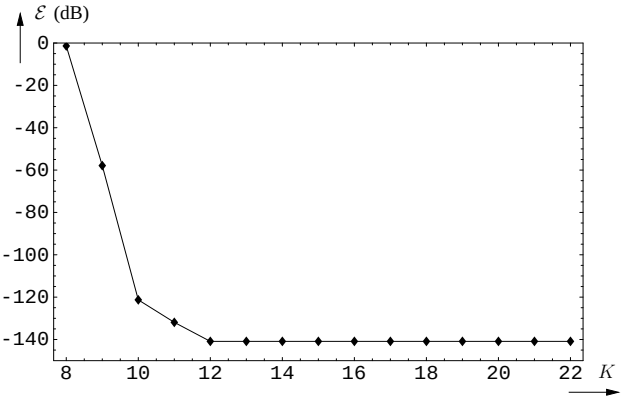

Fig. 8. Simulation results for $\mathcal{E}$ in terms of different values of K.

APPENDIX

A simplified form of the result in [11] can be written as

$$\mathbf{l}^T(p\mathbf{I}-\mathbf{A})^{-1}\mathbf{r} \simeq \mathbf{l}^T\mathbf{r}\mathbf{e}_1^T(p\mathbf{I}-\mathbf{T})^{-1}\mathbf{e}_1 \qquad (24)$$

where the RHS is the Padé approximant of the LHS. Here p is a scalar, $\mathbf{I}$ denotes the identity matrix (with appropriate dimensions), superscript T denotes transposition, $\mathbf{l}$ and $\mathbf{r}$ are given vectors with M elements, and $\mathbf{A}$ is a given M by M matrix. On the RHS $\mathbf{T}$ is a K by K matrix and $\mathbf{e}_1$ was defined earlier (see also Figure 5). If $K = M$, then (24) holds as an equality, whereas for $K < M$ an approximation takes place in a certain range for p. Matrix $\mathbf{T}$ has the tridiagonal form

$$\mathbf{T} = \begin{bmatrix} \alpha_1 & \beta_2 & 0 & 0 & \cdots & 0 & 0 \\ \rho_2 & \alpha_2 & \beta_3 & 0 & \cdots & 0 & 0 \\ 0 & \rho_3 & \alpha_3 & \beta_4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \rho_{K-1} & \alpha_{K-1} & \beta_K \\ 0 & 0 & 0 & \cdots & 0 & \rho_K & \alpha_K \end{bmatrix} \qquad (25)$$

where in terms of $\mathbf{l}$, $\mathbf{r}$, and $\mathbf{A}$ the parameters ρ_k, α_k, and β_k are calculated by the following well-known iterative procedure:

The classical Lanczos Algorithm: [15]
In the initialization part set $\rho_1 = \| \mathbf{r} \|_2$, $\eta_1 = \| \mathbf{l} \|_2$. In the n-th step of the algorithm we need to calculate the scalar d_n and the vectors $\mathbf{v}_n$ and $\mathbf{w}_n$ with setting the initial values as $d_0 = 1$, $\mathbf{v}_0 = \mathbf{w}_0 = \mathbf{0}$, $\mathbf{v}_1 = \mathbf{r}/\rho_1$, and $\mathbf{w}_1 = \mathbf{l}/\eta_1$. Then, for $n = 1, 2, \ldots, K$ do:

1) Compute $d_n = \mathbf{w}_n^T\mathbf{v}_n$.
2) Set

$$\alpha_n = \frac{\mathbf{w}_n^T\mathbf{A}\mathbf{v}_n}{d_n}, \quad \beta_n = \frac{d_n}{d_{n-1}}\eta_n, \quad \gamma_n = \frac{d_n}{d_{n-1}}\rho_n,$$

3) Calculate the auxiliary parameters $\mathbf{v}$ and $\mathbf{w}$ as

$$\begin{aligned} \mathbf{v} &= \mathbf{A}\mathbf{v}_n - \mathbf{v}_n\alpha_n - \mathbf{v}_{n-1}\beta_n \\ \mathbf{w} &= \mathbf{A}^T\mathbf{w}_n - \mathbf{w}_n\alpha_n - \mathbf{w}_{n-1}\gamma_n \end{aligned}$$

4) Set $\rho_{n+1} = \| \mathbf{v} \|_2$, $\eta_{n+1} = \| \mathbf{w} \|_2$, and

$$\mathbf{v}_{n+1} = \frac{\mathbf{v}}{\rho_{n+1}}, \quad \mathbf{w}_{n+1} = \frac{\mathbf{w}}{\eta_{n+1}},$$

The algorithm can be implemented very efficiently, but for ill-conditioned $(p\mathbf{I} - \mathbf{A})$ a breakdown (triggered by division by zero) typically occurs.

Using $\hat{\mathbf{e}}_n$ defined before (23), and using (20), the coefficients in (13) can be calculated by:

$$x_n = \hat{\mathbf{e}}_n^T\mathbf{x} \simeq \hat{\mathbf{e}}_n^T\mathbf{X}(\mathbf{F}\hat{\mathbf{F}})^{-1}\mathbf{F}\mathbf{q}$$

Therefore, using (24) with $p = 0$ gives (22).

REFERENCES

[1] A. Aldroubi and K. Gröchenig, "Non-uniform sampling and reconstruction in shift-invariant spaces", *SIAM Review*, Vol. 43, pp. 585-620, 2001.

[2] D. Castro, R. Espejo, A. Rodrígues-Vázques, A. Carmona, P. Földesy, Á. Zarándy, P. Szolgay, T. Szirányi, and T. Roska, "A 0.8-μm CMOS two-dimensional programmable mixed-signal focal-plane array processor with on-chip binary imaging and instructions storage", *IEEE Journal on Solid-State Circuits*, Vol. 32 No. 7 : 1013-1026. July 1997, pp. 1013-1026.

[3] R. H. Chan, A. M. Yip, and M. K. Ng, "The best circulant preconditioners for Hermitian Toeplitz systems", *SIAM Journal on Numerical Analysis*, Vol. 38, No. 3, pp. 876-896.

[4] L. O. Chua and T. Roska, "Stability of a class of nonreciprocal cellular neural network, *IEEE Transactions on Circuits and Systems I*, Vol. 37, No. 12, December 1990, pp. 1520-1527.

[5] L. Chua and T. Roska, *Cellular Neural Networks and Visual Computing*, Cambridge University Press, 2002.

[6] A. Cichocki and R. Unbehauen, "Neural networks for solving systems of linear equations and related problems", *IEEE Transactions on Circuits and Systems I*, Vol. 39, No. 2, February 1992, pp. 124-138.

[7] A. Cichocki and R. Unbehauen, "Simplified neural networks for solving linear least squares and total least squares problems in real time", *IEEE Transactions on Neural Networks*, Vol. 5, No. 6, November 1994, pp. 910-923.

[8] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series", *Transactions of the American Mathematical Society*, Vol. 72, pp. 341-366, 1952.

[9] H. G. Feichtinger and K. Gröchenig, *Theory and Practice of Irregular Sampling*. In J.J. Benedetto and M.W. Frazier, editors, Wavelets: Mathematics and Applications, pp. 305-363, CRC Press, Boca Raton, FL, 1994.

[10] H. G. Feichtinger, K. Gröchenig, and T. Strohmer, "Efficient numerical methods in non-uniform sampling theory", *Numerische Mathematik*, 69:423-440, 1995.

[11] P. Feldmann and R. W. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 14, No. 5, May 1995, pp. 639-649.

[12] S. Jaffard, "A density criterion for frames of complex exponentials", *Michigan Mathematical Journal*, Vol. 38, No. 3, pp. 339-348, 1991.

[13] M. Konishi, "Sound localization in the barn owl: tuning neuronal hardware with microsecond precision", *Scientific American* Vol. 268/4, 1993.

[14] V. A. Kotel'nikov, *On the Transmission Capacity of the Ether and Wire in Electrocommunications*, In J.J. Benedetto and P.J.S.G. Ferreira, editors, Modern Sampling Theory, Mathematics and Applications, pp. 27-45. Birkhauser, Boston, MA, 2001. Translated by V.E. Katsnelson from the Russian original published in Izd. Red. Upr. Suyazi RKKA, Moscow 1933.

[15] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators", *Journal of Research of the National Bureau of Standards*, Vol. 45, October 1950, pp. 255-282.

[16] A. A. Lazar and L. T. Tóth, "Time encoding and perfect recovery of bandlimited signals", *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. VI, pp. 709-712, April 6-10, 2003, Hong Kong.

[17] A. A. Lazar, " Time encoding with an integrate-and-fire neuron with a refractory period", *Neurocomputing*, Vol. 58-50, pp. 53-58, 2004. Presented at The Computational NeuroScience Meeting, July 5-9, 2003, Alicante, Spain.

[18] A. A. Lazar and L. T. Tóth, "Sensitivity analysis of time encoded bandlimited signals", *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. II, pp. 901-904, May 17-21, 2004, Montreal.

[19] A. A. Lazar, , E. K. Simonyi, and L. T. Tóth, "Fast recovery algorithms for time encoded bandlimited signals" , *BNET Technical Report*, #1-04, Department of Electrical Engineering, Columbia University, New York, NY 10027, June 2004.

[20] K. Lotz, L. Bölöni, T. Roska, and J. Hámori, " Hyperacuity in time: a CNN model of a time-coding pathway of sound localization", *IEEE Transactions on Circuits and Systems I*, Vol. 46, No. 8, August 1999, pp. 994-1002.

[21] F. Marvasti and M. Sandler, *Applications of Nonuniform Sampling to Nonlinear Modulation, A/D and D/A Techniques*. In F. Marvasti, editor, Nonuniform Sampling, Theory and Practice, pp. 647-687, Kluwer Academic/Plenum Publishers, New York, 2001.

[22] S. R. Norsworthy, R. Schreier, and G. C. Temes (editors), *Delta-Sigma Data Converters*, IEEE Press, New York, 1997.

[23] T. Roska and Rodriguez-Vazquez, "Towards Visual Microprocessors", *Proceedings of the IEEE*, July, 2002.

[24] C. E. Shannon, "Communications in the presence of noise", *Proceedings of the IRE*, Vol. 37, pp. 10-21, January 1949.

[25] T. Strohmer, *Irregular Sampling, Frames and Pseudoinverse*, *MS Thesis*, 1993, Department of Mathematics, University of Vienna, Vienna, Austria.

[26] R. M. Young, *Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.