R.J. Marks II and D. Radbel, "Error of linear estimation of lost samples in an oversampled bandlimited signal", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-32, pp.648-654 (1984).

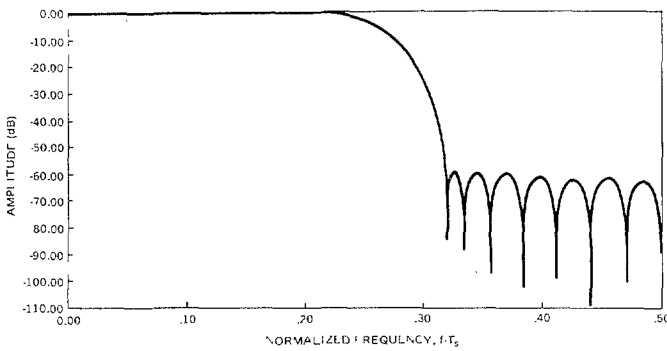

Fig. 4. Frequency response of the 32-tap quadrature mirror filter.

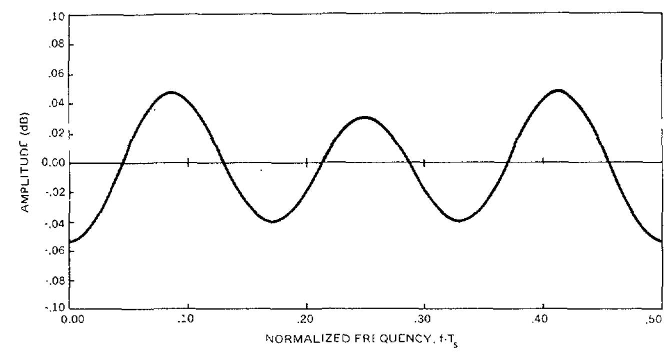

Fig. 5. Reconstruction error between the output and the input of the system of Fig. 1 (codecs not included). Eight taps.

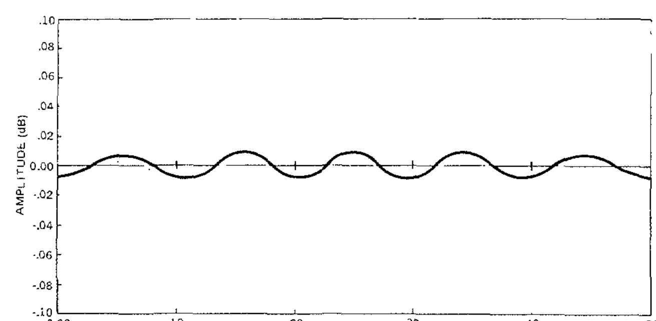

Fig. 6. Reconstruction error between the output and the input of the system of Fig. 1 (codecs not included). 16 taps.

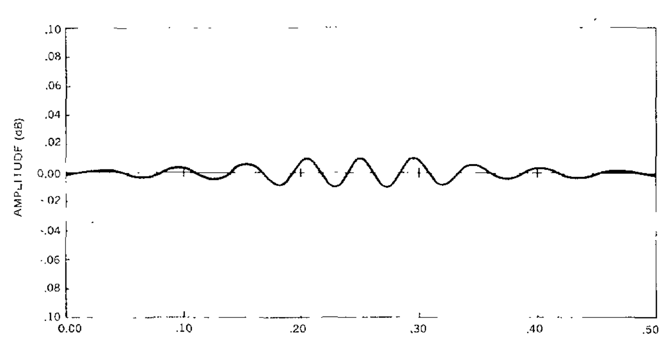

Fig. 7. Reconstruction error between the output and the input of the system of Fig. 1 (codecs not included). 32 taps.

TABLE III
APPROXIMATION ERRORS FOR 8-, 16-, AND 32-TAP FILTERS

N	WINDOWING TECHNIQUE	METHOD OF [4]	PROPOSED METHOD
8	0.969858 E-01	0.297892 E-03	0.254621 E-03
16	0.437511 E-02	0.172692 E-05	0.149813 E-05
32	0.576005 E-01	0.291371 E-05	0.169736 E-05

IV. CONCLUSIONS

An automatic method has been proposed for designing half-band nonrecursive quadrature mirror filters with the desired behavior of the stop and transition bands.

This method is based on an analytical formula which has been derived to represent a duly defined approximation error: in this way the set of coefficients which minimize this error can be computed by means of a well known nonlinear optimization procedure.

REFERENCES

[1] D. Esteban and C. Galand, "32 KBPS CCITT compatible splitband coding scheme," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Tulsa, OK, Apr. 1978, pp. 320–325.

[2] R. Crochiere, "Digital signal processor: Sub-band coding," *Bell Syst. Tech. J.*, vol. 60, pp. 1633–1653, Sept. 1981.

[3] R. S. Cheung and R. L. Winslow, "High quality 16 kb/s voice transmission: The subband coder approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Denver, CO, Apr. 1980, pp. 319–322.

[4] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 291–294, Denver, CO, Apr. 1980.

[5] H. H. Rosenbrock, "An automatic method for finding the greatest or least value of a function," *Comput. J.*, vol. 3, pp. 175–184, 1960.

[6] R. Hooke and T. Jeaves, "Direct search solution of numerical and statistical problems," *J. Ass. Comp. Mach.*, vol. 8, pp. 212–229, Apr. 1961.

[7] L. R. Rabiner, C. A. McGonegal, and D. Paul, "FIR windowed filter design program–WINDOW," in *Programs for Digital Signal Processing*. New York: IEEE Press, 1979.

[8] G. Pirani, F. Rusinà, and V. Zingarelli, "Multiplication-free filters for subband coding of speech," in *Proc. IEEE Int. Symp. Circuits Syst.*, Rome, Italy, May 1982, pp. 848–851.

Error of Linear Estimation of Lost Samples in an Oversampled Band-Limited Signal

ROBERT J. MARKS II AND DMITRY RADBEL

Abstract—A finite number of lost samples from an oversampled band-limited signal can be restored from the remaining samples. This paper explores the noise sensitivity of a linear algorithm that performs such restoration. Even though the problem is well posed, restoration noise level can become prohibitively high for a) sampling rates close to the

Manuscript received June 27, 1983; revised November 22, 1983.
The authors are with the University of Washington, Department of Electrical Engineering, Seattle, WA 98195.

0096-3518/84/0600-0648$01.00

Nyquist rate, and b) too many lost samples. Numerical results of restoration noise level are given for the cases of one lost sample, two (not necessarily adjacent) samples and a sequence of M adjacent lost samples. The effects of both truncation and noise are evaluated for the case of a single lost sample in a stochastic signal. The results are compared with the corresponding minimum mean-square error of the lost sample. Although suboptimal, the truncated lost sample sampling estimate is more straightforward computationally and does not require detailed knowledge of the signal or noise second-order statistics.

I. Introduction

In the absence of noise a finite number of lost samples in an oversampled band-limited signal can be regained from the remaining known samples [1]. Such a restoration falls under the title of interpolation—a process that is well posed [2]–[4]; that is, the ratio of restoration error to data noise levels can be bound. Clearly, the performance of such restoration algorithms will generally worsen (and the bound increase) when a) the sampling rate becomes closer to the Nyquist rate, and b) the number of lost samples increases.

In this paper, we investigate the sensitivity of the restoration algorithm in [1] to additive samplewise white noise. The variance of the corresponding restoration uncertainty is used as a measure of the algorithm's performance.

For a single lost sample in a zero-mean wide sense stationary stochastic signal, the effects of both data noise and truncation on the estimation error are considered: first using the truncated lost sample sampling theorem (ST) estimate and then using the minimum linear mean-square error (MSE) estimate. Although the ST approach is suboptimal, its implementation requires only a bound on the signal bandwidth. It does not depend on a detailed knowledge of the second-order statistics of the signal and noise as does the MSE estimate.

II. Preliminaries

Here, without elaboration, we restate the results in [1]. Let $f(x)$ denote a finite energy deterministic band-limited signal with bandwidth $2W$. That is

$$f(x) = \int_{-W}^{W} F(u) \exp(j2\pi ux)\, du$$

where

$$F(u) = \int_{-\infty}^{\infty} f(x) \exp(-j2\pi ux)\, dx.$$

Let $2B$ be a sampling rate equal to or in excess of the Nyquist rate $2W$. Define the sampling rate parameter $r = W/B \leqslant 1$. Let $\mathfrak{M}$ denote a set of indices corresponding to M lost samples. If $r < 1$, we can regain $f(x)$ from the sample set $\{f(m/2B) | m \notin \mathfrak{M}\}$ via

$$f(x) = \sum_{n \notin \mathfrak{M}} f\left(\frac{n}{2B}\right) k(x; n) \tag{1}$$

or

$$f(x) = \sum_{n \notin \mathfrak{M}} f\left(\frac{n}{2B}\right) k_r(x; n) \tag{2}$$

where the interpolation functions are

$$k(x; n) = \operatorname{sinc}(2Bx - n) + r \sum_{p \in \mathfrak{M}} \sum_{q \in \mathfrak{M}} a_{qp} \operatorname{sinc} r(n - q) \operatorname{sinc}(2Bx - p) \tag{3}$$

and

$$k_r(x; n) = k(x; n) * 2W \operatorname{sinc}(2Wx) = r \operatorname{sinc}(2Wx - rn) + r^2 \sum_{p \in \mathfrak{M}} \sum_{q \in \mathfrak{M}} a_{qp} \operatorname{sinc} r(n - q) \operatorname{sinc}(2Wx - rp) \tag{4}$$

where a_{qp} is the qpth element of the matrix $\boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{S})^{-1}$, $\boldsymbol{I}$ is the identity matrix, and $\boldsymbol{S}$ is a Toeplitz matrix with elements $\{s_{mn} = r \operatorname{sinc} r(m - n) | (m, n) \in \mathfrak{M} \times \mathfrak{M}\}$. The asterisk denotes convolution. Note that the r subscript denotes the filtered equivalent of the nonsubscripted case. Correspondingly, we shall refer to (1) as the unfiltered and (2) as the filtered case.

The sampling theorem is also applicable to bandlimited stochastic processes. Let $\mathfrak{f}(x)$ denote a real zero-mean wide sense stationary random signal with autocorrelation

$$R_{\mathfrak{f}}(x - y) = E[\mathfrak{f}(x)\mathfrak{f}(y)]$$

where $E(\cdot)$ denotes the expected value operator. Let $\mathfrak{f}(x)$ be bandlimited with bandwidth $2W$ in the sense that

$$R_{\mathfrak{f}}(x) = \int_{-W}^{W} S_{\mathfrak{f}}(u) \exp(j2\pi ux)\, du$$

where the power spectral density of the process is

$$S_{\mathfrak{f}}(u) = \int_{-\infty}^{\infty} R_{\mathfrak{f}}(x) \exp(-j2\pi ux)\, dx.$$

Motivated by (1) and (2), we define the following processes in terms of the known samples

$$\hat{\mathfrak{f}}(x) = \sum_{n \notin \mathfrak{M}} \mathfrak{f}\left(\frac{n}{2B}\right) k(x; n) \tag{5}$$

and

$$\hat{\mathfrak{f}}_r(x) = \sum_{n \notin \mathfrak{M}} \mathfrak{f}\left(\frac{n}{2B}\right) k_r(x, n). \tag{6}$$

Using (5) and (6) one can straightforwardly show in a manner paralleling Papoulis [7] that, if $r < 1$, $\hat{\mathfrak{f}}(x)$ and $\hat{\mathfrak{f}}_r(x)$ are equal to $\mathfrak{f}(x)$ in the mean-square sense. That is

$$E[|\mathfrak{f}(x) - \hat{\mathfrak{f}}(x)|^2] = 0 \tag{7}$$

and

$$E[|\mathfrak{f}(x) - \hat{\mathfrak{f}}_r(x)|^2] = 0. \tag{8}$$

III. Noise Sensitivity

Let $\xi(x)$ denote a real zero-mean wide sense stationary stochastic process with autocorrelation

$$R_\xi(x - y) = E[\xi(x)\xi(y)]$$

and noise level

$$\overline{\xi^2} = R_\xi(0)$$

where the overbar denotes the expected value operator. We shall assume that the extent of the autocorrelation is sufficiently small so that when sampled at a rate of $2B$, the samples are white. That is

$$R_\xi\left(\frac{n}{2B}\right) = \overline{\xi^2}\delta_n$$

where δ_n is the Kronecker delta.

If $f(x) + \xi(x)$ is used as input in (1) and (2), the outputs for the unfiltered and filtered cases are $f(x) + \eta(x)$ and $f(x) + \eta_r(x)$, respectively, where

$$\eta(x) = \sum_{n \notin \mathcal{M}} \xi\left(\frac{n}{2B}\right) k(x; n) \tag{9}$$

and

$$\eta_r(x) = \sum_{n \notin \mathcal{M}} \xi\left(\frac{n}{2B}\right) k_r(x; n). \tag{10}$$

The corresponding interpolation noise levels follow, respectively, as

$$\begin{aligned}\overline{\eta^2(x)} &= E[\eta^2(x)] \\ &= \sum_{n \notin \mathcal{M}} \sum_{m \notin \mathcal{M}} R_\xi\left(\frac{n-m}{2B}\right) k(x; n) k(x; m) \\ &= \overline{\xi^2} \sum_{n \notin \mathcal{M}} k^2(x; n)\end{aligned} \tag{11}$$

and

$$\begin{aligned}\overline{\eta_r^2(x)} &= E[\eta_r^2(x)] \\ &= \overline{\xi^2} \sum_{n \notin \mathcal{M}} k_r^2(x; n).\end{aligned} \tag{12}$$

For the unfiltered case, $\overline{\eta^2(k/2B)} = \overline{\xi^2}; k \notin \mathcal{M}$. Upon inspection of the iterative form of the restoration algorithm in [1], one can see that

$$\overline{\eta_r^2(k/2B)} = \overline{\eta^2(k/2B)}; \qquad k \in \mathcal{M}. \tag{13}$$

That is, the noise level at the restored sample points is identical for the filtered and unfiltered cases. For $k \notin \mathcal{M}$, the noise level at the unknown sample locations is generally decreased by filtering.

Consider next the stochastic signal case. We assume $\xi(x)$ is not correlated with the signal $f(y)$ for all (x, y)

$$E[\xi(x) f(y)] = 0. \tag{14}$$

Define the observed signal

$$g(x) = f(x) + \xi(x). \tag{15}$$

If $g(x)$ is sampled and restored using (5), the result is

$$\hat{g}(x) = \hat{f}(x) + \eta(x). \tag{16}$$

In general, $\eta(x)$ is not stationary. Note, from (15), $\eta(x)$ is uncorrelated with both $\hat{f}(y)$ and $\hat{f}_r(y)$.

The restoration error here is

$$\begin{aligned}\epsilon(x) &= E[|f(x) - \hat{g}(x)|^2] \\ &= E[\eta^2(x)] = \overline{\eta^2(x)}.\end{aligned} \tag{17}$$

For the filtered case, define

$$\hat{g}_r(x) = \hat{f}_r(x) + \eta_r(x). \tag{18}$$

The restoration error here is

$$\begin{aligned}\epsilon_r(x) &= E[|f(x) - \hat{g}_r(x)|^2] \\ &= \overline{\eta_r^2(x)}.\end{aligned} \tag{19}$$

Thus, the restoration noise levels in (17) and (19) are the same as for the deterministic signal case in (11) and (12).

No Lost Samples

For purposes of later comparison, consider first the case where $\mathcal{M}$ is empty. For the unfiltered case, (11) becomes

$$\begin{aligned}\overline{\eta^2(x)} &= \overline{\xi^2} \sum_{n \notin \mathcal{M}} k^2(x; n) \\ &= \overline{\xi^2} \sum_{n=-\infty}^{\infty} \operatorname{sinc}^2(2Bx - n).\end{aligned} \tag{20}$$

To evaluate this sum, we rewrite (2) with $f(x) = r \operatorname{sinc} 2W(\alpha - x)$:

$$r \operatorname{sinc} 2W(\alpha - x) = r^2 \sum_{n=-\infty}^{\infty} \operatorname{sinc}(2W\alpha - rn) \operatorname{sinc}(2Wx - rn). \tag{21}$$

Setting $r = 1$ (and thus $W = B$) along with $\alpha = x$, we can evaluate (19):

$$\overline{\eta^2(x)} = \overline{\xi^2}. \tag{22}$$

Thus, the noise level of the interpolation at all points is the same as that of the data samples. Although we have assumed samplewise white noise, one can straightforwardly demonstrate that (22) is true for any wide sense stationary noise.

For the filtered case and $\mathcal{M}$ empty, (12) becomes

$$\begin{aligned}\overline{\eta_r^2(x)}/\overline{\xi^2} &= \sum_{n \notin M} k_r^2(x; n) \\ &= r^2 \sum_{n=-\infty}^{\infty} \operatorname{sinc}^2(2Wx - rn) \\ &= r\end{aligned} \tag{23}$$

where we have used (21) with $\alpha = x$. Thus, increasing the sampling rate reduces the noise level. Contrary to appearance, the noise level cannot be made arbitrarily small by a corresponding increase in sampling rate. Eventually, adjacent samples will become correlated and the white noise assumption violated.

One Lost Sample

Let $M = 1$. With no loss in generality, let that sample be at the origin. For the unfiltered case, (11) becomes

$$\begin{aligned}\overline{\eta^2(x)}/\overline{\xi^2} &= \sum_{n \neq 0} \left[\operatorname{sinc}(2Bx - n) + \frac{r}{1-r} \operatorname{sinc}(rn) \operatorname{sinc}(2Bx)\right]^2 \\ &= \sum_{n=-\infty}^{\infty} \Bigg[\operatorname{sinc}^2(2Bx - n) \\ &\quad + \left(\frac{r}{1-r}\right)^2 \operatorname{sinc}^2(rn) \operatorname{sinc}^2(2Bx) \\ &\quad + \frac{2r}{1-r} \operatorname{sinc}(2Bx - n) \operatorname{sinc}(rn) \operatorname{sinc}(2Bx)\Bigg] \\ &\quad - \frac{1}{1-r} \operatorname{sinc}^2(2Bx).\end{aligned}$$

With appropriate choice of r and α, (21) can be used to evaluate each of the three n sums. After simplifying, we obtain

$$\overline{\eta^2(x)}/\overline{\xi^2} = 1 + \frac{2r}{1-r} \operatorname{sinc}(2Wx) \operatorname{sinc}(2Bx) - \frac{1}{1-r} \operatorname{sinc}^2(2Bx) \tag{24}$$

Note the normalized noise level approaches unity for large x – equal to the no lost sample case in (22). The noise level of the interpolated point at the origin follows from (24) as

$$\begin{aligned}\overline{\eta^2(0)} &= \frac{r}{1-r} \overline{\xi^2} \\ &= \overline{\eta_r^2(0)}.\end{aligned} \tag{25}$$

The result is monotonically increasing on $0 \leqslant r \leqslant 1$. Interestingly, for $r < \frac{1}{2}$, the normalized interpolation noise level in (25) is less than unity which is less than the noise level of the known sample data. Note, however, that we have yet to filter the high-frequency components of the samplewise white noise.

For the filtered case for one lost sample, (12) becomes

$$\overline{\eta_r^2(x)}/\overline{\xi^2} = r^2 \sum_{n \neq 0} \left[\operatorname{sinc}(2Wx - rn) + \frac{r}{1-r} \operatorname{sinc}(rn) \operatorname{sinc}(2Wx)\right]^2.$$

Proceeding in a manner similar to that for the unfiltered case above, we obtain

$$\overline{\eta_r^2(x)}/\overline{\xi^2} = \frac{r}{1-r}[1 - r\{1 - \operatorname{sinc}^2(2Wx)\}]. \tag{26}$$

For large x, the noise level goes to the no lost sample filtered equivalent in (23).

Two Lost Samples

Let $M = 2$ and let the lost samples be located at the origin and at $x = k/2B$ for some specified k. The 2×2 $\boldsymbol{A}$ matrix then has elements

$$a_{11} = a_{22} = \frac{1-r}{\Delta}$$

$$a_{12} = a_{21} = \frac{r \operatorname{sinc}(rk)}{\Delta}$$

where

$$\Delta = (1-r)^2 - r^2 \operatorname{sinc}^2(rk).$$

After straightforward yet tedious calculations, (11) and (12) become

$$\begin{aligned}\overline{\eta^2(x)}/\overline{\xi^2} = {} & 1 - (\alpha^2 + \beta^2) + 2r[a_{11}(\alpha\tau + \beta\rho) + a_{12}(\alpha\rho + \beta\tau)] \\ & + r^2[(a_{11}^2 + a_{12}^2)(\alpha^2\lambda + 2\alpha\beta\gamma + \beta^2\lambda) \\ & + 2a_{11}a_{12}(\alpha^2\gamma + 2\alpha\beta\lambda + \beta^2\gamma)]\end{aligned} \tag{27}$$

and

$$\begin{aligned}\overline{\eta_r^2(x)}/\overline{\xi^2} = {} & r - r^2(\alpha_1^2 + \beta_1^2) + 2r^3[a_{11}(\alpha_1\tau_1 + \beta_1\rho_1) \\ & + a_{12}(\alpha_1\rho_1 + \beta_1\tau_1)] \\ & + r^4[(a_{11}^2 + a_{12}^2)(\alpha_1^2\lambda + 2\alpha_1\beta_1\gamma + \beta_1^2\lambda) \\ & + 2a_{11}a_{12}(\alpha_1^2\gamma + 2\alpha_1\beta_1\lambda + \beta_1^2\gamma)]\end{aligned} \tag{28}$$

where

$$\alpha = \operatorname{sinc}(2Bx - k) \qquad \alpha_1 = \operatorname{sinc}(2Wx - rk)$$

$$\beta = \operatorname{sinc}(2Bx) \qquad \beta_1 = \operatorname{sinc}(2Wx)$$

and

$$\rho = \sum_{p \neq 0, k} \operatorname{sinc}(rp) \operatorname{sinc}(2Bx - p) = \beta_1 - \beta - \alpha \operatorname{sinc}(rk)$$

$$\rho_1 = \sum_{p \neq 0, k} \operatorname{sinc}(rp) \operatorname{sinc}(2Wx - rp) = \frac{1-r}{r}\beta_1 - \alpha_1 \operatorname{sinc}(rk)$$

$$\tau = \sum_{p \neq 0, k} \operatorname{sinc} r(p-k) \operatorname{sinc}(2Bx - p) = \alpha_1 - \alpha - \beta \operatorname{sinc}(rk)$$

$$\tau_1 = \sum_{p \neq 0, k} \operatorname{sinc} r(p-k) \operatorname{sinc}(2Wx - rp) = \frac{1-r}{r}\alpha_1 - \beta_1 \operatorname{sinc}(rk)$$

$$\lambda = \sum_{p \neq 0, k} \operatorname{sinc}^2(rp) = \sum_{p \neq 0, k} \operatorname{sinc}^2 r(p-k) = \frac{1}{r} - 1 - \operatorname{sinc}^2(rk)$$

$$\gamma = \sum_{p \neq 0, k} \operatorname{sinc}(rp) \operatorname{sinc} r(p-k) = \left(\frac{1}{r} - 2\right) \operatorname{sinc}(rk).$$

All of the sums above are evaluated as special cases of (21). Numerical examples of (27) and (28) are shown in Fig. 1 for $k = 1$ and 5 with $r = 0.2$. The lost sample locations here are at the minima of the unfiltered noise level curves. Note the consistency with (12). A second example for $r = 0.8$ and $k = 1$ is shown in Fig. 2. The lost samples are at zero and unity. The filtered and unfiltered curves are indistinguishable near those points.

At the lost sample point locations

$$\begin{aligned}\overline{\eta^2(0)} = \overline{\eta^2\left(\frac{k}{2B}\right)} &= r^2\overline{\xi^2}[(a_{11}^2 + a_{12}^2)\lambda + 2a_{11}a_{12}\gamma] \\ &= \overline{\eta_r^2(0)} = \overline{\eta_r^2\left(\frac{k}{2B}\right)}.\end{aligned} \tag{29}$$

For large k, the noise level at the origin approaches that for a single lost sample. If kr is an integer, the noise levels for one and two lost samples are equal at the lost sample locations. A plot of (29) is shown in Fig. 3 for $k = 1$, 2, and 5. The single lost sample noise level in (25) is nearly graphically indistinguishable from the $k = 5$ curve.

Sequence of Lost Samples

Clearly, closed form expressions for (11) and (12) become intractible for larger M. Numerical results for three lost samples in a row are shown in Fig. 4. The noise level at the lost sample locations is shown in Fig. 5 for M lost samples in a row. The noise level increases drastically with respect to the number of adjacent lost samples and sampling rate parameter. Correspondingly, the condition number [6] of the $\boldsymbol{A} = [\boldsymbol{I} - \boldsymbol{S}]^{-1}$ matrix increases greatly with larger M and r.

IV. Truncation Effects for a Lost Sample in a Stochastic Signal

Consider the case where a single lost sample is estimated by N samples on each side of the missing point. In the limit, we know the sampling estimates in (5) and (6) are optimal in the sense of minimum linear mean-square error. Truncation and noise destroy this optimality. In this section, the effects of these sources of error are considered. The corresponding minimum mean-square error is then evaluated for purposes of comparison.

Let $M = 1$ and, with no loss in generality, let the lost sample be at the origin. The estimate of the lost sample follows from (16) and (18) as

$$\hat{g}(0) = \hat{g}_r(0) = \frac{r}{1-r} \sum_{n \neq 0} g\left(\frac{n}{2B}\right) \operatorname{sinc}(rn). \tag{30}$$

Using only N samples from each side, the corresponding truncated estimate is

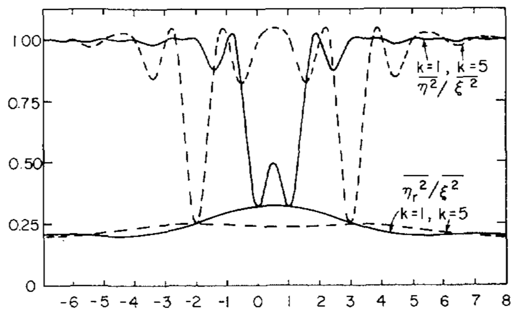

Fig. 1. Restoration noise level for two lost samples when sampling at five times the Nyquist rate ($r = 0.2$). The solid curves are for $k = 1$. The lost samples are at zero and one. The broken line graphs are for $k = 5$ with lost samples at -2 and 3. In both cases, the lower curve represents the filtered case and the upper curve the nonfiltered case.

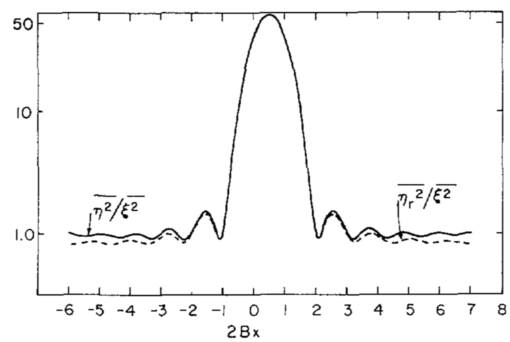

Fig. 2. Restoration noise level for two lost samples when $r = 0.8$ and $k = 1$. The lost samples are at zero and one. The solid curve is for the unfiltered case and broken line plot for the filtered case. The two plots are graphically indistinguishable in the region of the lost samples.

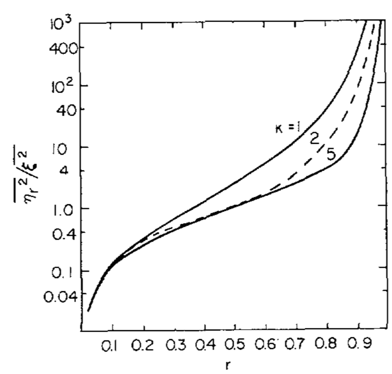

Fig. 3. Restoration noise level for two lost samples at the lost sample location as a function of the sampling rate parameter. The noise level for a single lost sample is almost graphically indistinguishable from the $k = 5$ plot.

$$\hat{g}_N(0) = \frac{r}{1-r} \sum_{\substack{|n| \le N \\ n \ne 0}} g\left(\frac{n}{2B}\right) \operatorname{sinc}(rn).$$

The mean-square error of this estimate is

$$\hat{\epsilon}_N = E[|f(0) - \hat{g}_N(0)|^2]$$
$$= E\left[\left| f(0) - \frac{r}{1-r} \sum_{\substack{|n| \le N \\ n \ne 0}} \left\{ f\left(\frac{n}{2B}\right) + \xi\left(\frac{n}{2B}\right)\right\} \operatorname{sinc}(rn)\right|^2\right]. \tag{31}$$

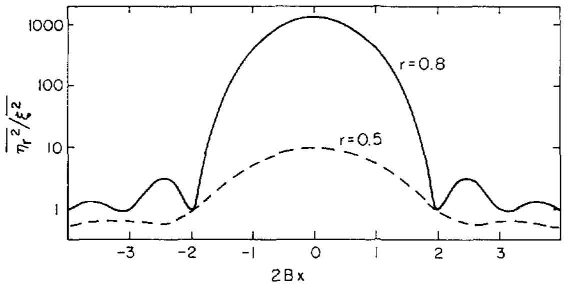

Fig. 4. Filtered restoration noise level for three lost samples at 0, +1, and −1.

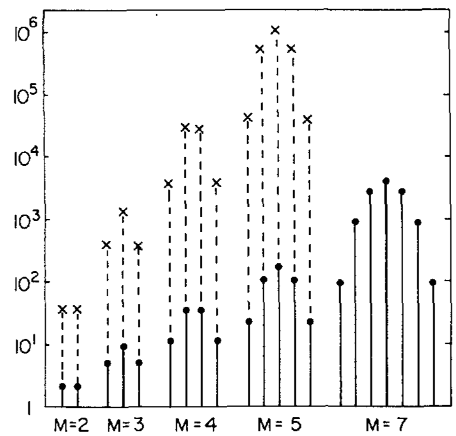

Fig. 5. Restoration noise level for M lost samples in a row at the lost sample locations. The lower dot values in each case correspond to $r = 0.5$ and the upper x's to $r = 0.8$.

For samplewise white noise (31) becomes

$$\epsilon_N = R_f(0) - \frac{2r}{1-r} \sum_{\substack{|n| \le N \\ n \ne 0}} R_f\left(\frac{n}{2B}\right) \operatorname{sinc}(rn)$$
$$+ \left(\frac{r}{1-r}\right)^2 \left[\overline{\xi^2} \sum_{\substack{|n| \le N \\ n \ne 0}} \operatorname{sinc}^2(rn) + \sum_{\substack{|n| \le N \\ n \ne 0}} \sum_{\substack{|m| \le N \\ m \ne 0}} R_f\left(\frac{n-m}{2B}\right) \operatorname{sinc}(rn) \operatorname{sinc}(rm)\right].$$

We normalize the error to the noise variance

$$\tilde{\epsilon}_N = \epsilon_N / \overline{\xi^2}.$$

Note that the corresponding normalization

$$\tilde{R}_f(x) \equiv R_f(x)/\overline{\xi^2}$$

is, for $x = 0$, the signal to noise ratio. Then (31) becomes

$$\tilde{\epsilon}_N = \tilde{R}_f(0) - \frac{4r}{1-r} \sum_{n=1}^{N} \tilde{R}_f\left(\frac{n}{2B}\right) \operatorname{sinc}(rn)$$
$$+ 2\left(\frac{r}{1-r}\right)^2 \left[\sum_{n=1}^{N} \operatorname{sinc}^2(rn) + \sum_{n=1}^{N} \sum_{m=1}^{N} \left\{ \tilde{R}_f\left(\frac{n-m}{2B}\right) + \tilde{R}_f\left(\frac{n+m}{2B}\right)\right\} \cdot \operatorname{sinc}(rn) \operatorname{sinc}(rm)\right]. \tag{32}$$

After some manipulation, the following computationally more attractive iterative form of (32) results in

$$\tilde{\epsilon}_{N+1} = \tilde{\epsilon}_N - \frac{4r}{1-r}\tilde{R}_f\left(\frac{N+1}{2B}\right)\operatorname{sinc} r(N+1) + 2\left(\frac{r}{1-r}\right)^2\left[\left\{1+\tilde{R}_f(0)+\tilde{R}_f\left(\frac{2N+2}{2B}\right)\right\}\cdot \operatorname{sinc}^2 r(n+1) + 2\operatorname{sinc} r(N+1)\cdot\sum_{n=1}^{N}\left\{\tilde{R}_f\left(\frac{n-N-1}{2B}\right)+\tilde{R}_f\left(\frac{n+N+1}{2B}\right)\right\}\operatorname{sinc}(rn)\right]$$

with initialization

$$\tilde{\epsilon}_1 = \tilde{R}_f(0) - \frac{4r}{1-r}\tilde{R}_f\left(\frac{1}{2B}\right)\operatorname{sinc}(r) + 2\left(\frac{r}{1-r}\right)^2\left[1+\tilde{R}_f(0)+\tilde{R}_f\left(\frac{2}{2B}\right)\right]\operatorname{sinc}^2(r)\Bigg]. \tag{33}$$

The error for the ST estimate in (33) will later be compared to the minimum linear mean-square estimate of a single lost sample. We note from (25) that

$$\lim_{N\to\infty} \tilde{\epsilon}_N = \frac{r}{1-r}$$

a result independent of both the signal and noise structure.

Minimum Mean-Square Error Estimate

Given N samples on each side of a lost sample, we can form the more general linear estimate

$$\check{g}_N(0) = \sum_{\substack{|n|\leq N\\ n\neq 0}} \alpha_n g\left(\frac{n}{2B}\right) \tag{34}$$

and choose the coefficients so that the mean-square error

$$\check{\epsilon}_N = E[|f(0) - \check{g}_N(0)|^2] \tag{35}$$

is minimized. Clearly, the resulting error is lower here than in the truncated sampling theorem in (30) which is a (nonoptimal) special case of (34).

From the orthogonality principle [7] the error in (34) is minimum when $f(0) - \check{g}_N(0)$ is orthogonal to $\check{g}_N(0)$ and thus to each element in $\{g(n/2B) \| |n| \leq N, n \neq 0\}$. Thus, we require

$$\left[E\ \{f(0) - g(0)\}\ g\left(\frac{n}{2B}\right)\right] = 0; \qquad |n| \leq N,\ n\neq 0.$$

Expanding and recognizing by symmetry that $\alpha_n = \alpha_{-n}$ yields the optimal α_n's as solutions to the set of equations

$$R_f\left(\frac{m}{2B}\right) = \sum_{n=1}^{N}\alpha_n\left[R_g\left(\frac{n-m}{2B}\right) + R_g\left(\frac{n+m}{2B}\right)\right] \tag{36}$$

where, since the signal and noise are uncorrelated, the observed signal's autocorrelation is

$$R_g(x) = R_f(x) + R_\xi(x)$$

the α_n's can clearly be determined. The result, when substituted into (34), yields the minimum linear mean-square error MSE estimate of the lost sample, $f(0)$. The corresponding minimum error is

$$(\check{\epsilon}_N)_{\min} = R_f(0) - 2\sum_{n=1}^{N}\alpha_n R_f\left(\frac{n}{2B}\right). \tag{37}$$

Comparison of Lost Sample Estimates

Numerical examples for ST minimum and MSE estimates of a lost sample are shown in Figs. 6 and 7 for the case where the signal's power spectral density is $\overline{f^2}/2W$ over the interval $|u| \leq W$ and zero elsewhere. The noise is samplewise white. The signal-to-noise (power) ratio follows as

$$\text{SNR} = \overline{f^2}/\overline{\xi^2}.$$

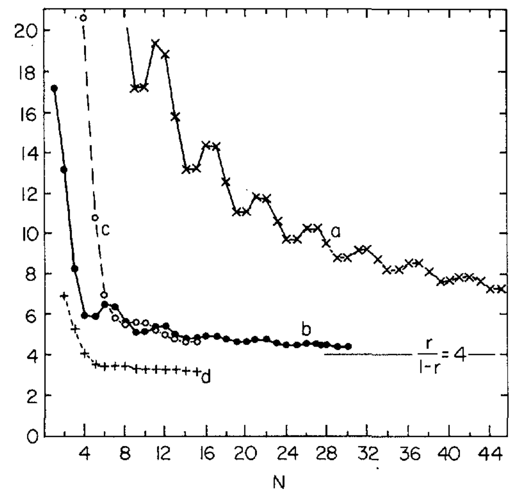

Fig. 6. Mean-square error of lost sample restoration when N samples on both sides are used in the estimate. Here, $r = 0.8$. ST: (a) SNR = 100, (b) SNR = 10. MSE: (c) SNR = 100, (d) SNR = 10.

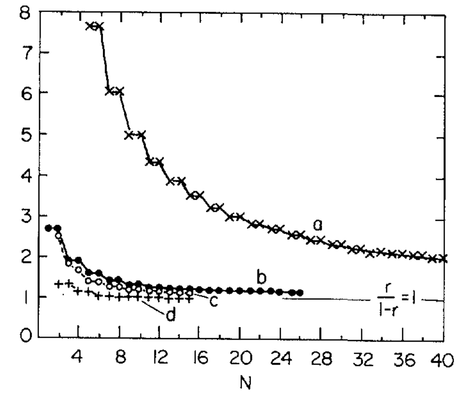

Fig. 7. Same as Fig. 6, but $r = 0.5$.

For the ST estimate, the normalized error approaches $r/(1 - r)$ independent of the SNR as $N \to \infty$.

Interestingly, both the ST and MSE estimates do not necessarily improve as we increase the number of samples. For the sampling theorem, this is graphically evident in Fig. 6.

Computation Comparison

By design, the optimal MSE estimate of the lost sample yields superior results than the ST estimate. From (34) and (36), however, the MSE estimate requires detailed knowledge of the autocorrelation of $g(x)$ at specified sample points. The ST estimate requires only bandwidth information. Clearly, this is also the case when restoring a larger number of lost samples.

In general, let K be the number of known samples used to estimate M lost samples. For good results, we require $K \gg M$. To apply the ST, we need to invert an $M \times M$ matrix to find the a_{qp}'s used in (3) and (4). For the MSE estimate, we are required to solve K linear equations with K unknowns in (36) to find the needed α_n's.

REFERENCES

[1] R. J. Marks II, "Restoring lost samples from an oversampled band-limited signal," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 752–755, June 1983.

[2] D. C. Youla, "Generalized image restoration by method of alternating orthogonal projections," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 694-702, 1978.

[3] H. Stark, D. Cahana, and H. Webb, "Restoration of arbitrary finite-energy optical objects from limited spatial and spectral information," *J. Opt. Soc. Amer.*, vol. 71, p. 635, 1981.

[4] D. Kaplan and R. J. Marks II, "Noise sensitivity of interpolation and extrapolation matrices," *Appl. Opt.*, vol. 21, pp. 4489-4492, 1982.

[5] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, p. 1565, 1977.

[6] T. K. Sarkar, D. D. Weiner, and V. K. Jain, "Some mathematical considerations in dealing with the inverse problem," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 373-379, 1981.

[7] A. Papoulis, *Probability, Random Variables and Stochastic Processes.* New York: McGraw-Hill, 1965, pp. 370, 390.

A Discussion of Various Approaches to the Linear System Identification Problem

TAPAN K. SARKAR, SOHEIL A. DIANAT, AND DONALD D. WEINER

Abstract—**This paper deals with the pole zero identification of a linear system from a measured input-output record. One objective is to show that the pencil-of-function method minimizes a weighted version of the Kalman equation error. It follows that the pencil-of-function method is capable of yielding robust estimates for poles located in a given region of the complex s plane. The second objective of this paper is to illustrate that identical sets of equations arise in three supposedly different analytical techniques for obtaining the impulse response of a system. The techniques investigated are 1) the least squares technique based on the discrete Wiener-Hopf equation, 2) Pisarenko's eigenvalue method, and 3) Jain's pencil-of-function method. The proof of equivalence is valid only for the noise-free case when the system order is known. Instead of using the conventional differential equation formulation, equivalence is shown with the integral form utilized in the pencil-of-function method.**

I. Introduction

In linear system identification, one is often interested in obtaining a pole-zero model of an unknown system from measured records of the input and output. If $x(t)$ and $y(t)$ are the respective time domain input and output to the system, then we are interested in characterizing the impulse response $h(t)$ by a sum of complex exponentials, i.e.,

$$h(t) \simeq \sum_{i=1}^{n} A_i \exp(s_i t). \tag{1}$$

Here n is referred to as the order of the system. s_i and A_i are the poles and the residues at the poles, respectively. In the Laplace domain, the problem is to model the transfer function $H(s)$ [which is the Laplace transform of $h(t)$] by a ratio of two polynomials as

$$\frac{Y(s)}{X(s)} = H(s) \simeq \frac{b_0 + b_1 s + \cdots + b_m s^m}{a_0 + a_1 s + \cdots + a_n s^n} \triangleq \frac{B(s)}{A(s)} \tag{2}$$

where $Y(s)$ and $X(s)$ are the Laplace transforms of the input and output, respectively. Zero initial conditions have been assumed in (2). Equality in (2) is attained when $y(t)$ and $x(t)$ are noise free and the system order n is exactly chosen.

Three basic approaches to solving the identification problem are a) the least squares approach (based on the Wiener-Hopf technique [1]-[3]), b) the eigenvalue method (based on Koopman's results [12] which were later applied by Levine [11] and Pisarenko [4]-[5]), and c) the pencil-of-function method (based on the linear dependence/independence of a set of functions) [6]-[8].

In this paper, we show that the three techniques yield analytically equivalent equations when there is no noise in the measured waveforms $x(t)$ and $y(t)$ and the system order n is correctly chosen. However, in the presence of noise, performance differs from one technique to another [8].

II. The Concept of Error in the Various Techniques

Given a specified input $x(t)$, one would like to minimize the mean-squared error between the actual output $y(t)$ and the predicted output from the system model. In the Laplace domain, this is mathematically equivalent to minimization of $|E^1(s)|^2$ where

$$E^1(s) = Y(s) - \frac{B(s)}{A(s)} X(s) = \frac{Y(s)A(s) - B(s)X(s)}{A(s)} \triangleq \frac{E(s)}{A(s)} \tag{3}$$

and the unknowns a_i and b_j appear in A and B [as defined in (2)].

However, even though minimization of $|E^1(s)|^2$ with respect to b_j is a linear problem, the minimization of the squared error with respect to a_i is a nonlinear problem [9]. Hence, we tend to minimize $|E(s)|^2$ (where $E(s)$ is popularly known as the equation error, after Kalman [10]) rather than $|E^1(s)|^2$. This is because minimization of $|E(s)|^2$ with respect to a_i and b_j is a linear problem. In fact, beginning with Kalman [10] in 1958, almost all pole zero modeling techniques utilize this error criterion. The first two techniques—the least-squares and the eigenvalue methods—as implemented by the present researchers, utilize the minimization of $|E(s)|^2$. On the other hand, the third technique—the pencil-of-function method—minimizes a weighted $|E(s)|^2$. This weighting is particularly useful when one is interested in very accurate locations of poles and zeros in a specified region of the complex s plane.

The obvious question now raised is "what guarantee does one have of obtaining a 'good' solution if $|E(s)|^2$ is minimized?" It is clear when the data is noise free and the system order n is correctly chosen that minimization of $|E(s)|^2$ is equivalent to minimization of $|E^1(s)|^2$. This is because when $E(s)$ is zero, $E^1(s)$ is of course zero because the latter is the result of passing the former through the linear filter $1/A(s)$.

However, if $y(t)$ is contaminated with noise, such that the noise contaminated output $Y_N(s)$ is

$$Y_N(s) = Y(s) + N(s) \tag{4}$$

then minimization of the error results in

$$E^1(s_i) = N(s_i). \tag{5}$$

Thus, the output noise plays a crucial role in computation of the system poles by minimizing $|E(s)|^2$. This is a well observed fact for Prony's method [8], which is similar to the least-squares technique [3].

Manuscript received July 29, 1982; revised August 19, 1983 and January 12, 1984. This work was supported by the Office of Naval Research under Contract N00014-79-C-0598.

T. K. Sarkar and S. A. Dianat are with the Department of Electrical Engineering, Rochester Institute of Technology, Rochester, NY 14623.

D. D. Weiner is with the Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY 13210.

0096-3518/84/0600-0654$01.00