SIGNAL RECONSTRUCTION FROM NOISY, ALIASED, AND NONIDEAL SAMPLES: WHAT LINEAR MMSE APPROACHES CAN ACHIEVE

Alvaro Guevara and Rudolf Mester

Visual Sensorics and Information Processing Lab (VSI)
Computer Science Dept., Goethe University, Frankfurt, Germany
Email: {alvaro.guevara, mester}@vsi.cs.uni-frankfurt.de

ABSTRACT

This paper addresses the problem of interpolating a (non-bandlimited) signal from a discrete set of noisy measurements obtained from non-δ sampling kernels. We present a linear estimation approach, assuming the signal is given by a continuous model for which first and second order moments are known. The formula provides a generalization of the well-known discrete-discrete Wiener style estimator, but does not necessarily involve Fourier domain considerations. Finally, some experiments illustrate the flexibility of the method under strong noise and aliasing effects, and shows how the input autocorrelation, the sampling kernel and the noise process shape the form of the optimal interpolating kernels.

1. INTRODUCTION

Sixty years after Shannon's key contribution, we observe a revived interest in sampling theory mainly due to new developments in e.g., spline-based signal processing [1, 2, 3], and estimation theory [4]. The essential message of the present paper is that even for non-ideal sampling, that is: for non-bandlimited signals and/or for sampling kernels that strongly deviate from Dirac δ-pulses, a linear reconstruction which is optimum in the least squares (LS) sense is possible and attractive. This statement should appear obvious and self-evident to anybody trained in statistical signal processing, but there is only very few work which actually goes along that line. This linear MMSE solution can be achieved in a much more straightforward way than this may appear from earlier publications on reconstruction from non-ideal sampling. For this approach, the input signals are assumed to be realizations from a wide-sense stationary (WSS) process, and the first and second order moment functions – the autocorrelation function (ACF) – must be given, or estimated. We do not see many practical situations where these very mild requirements cannot be met. The range of application of this approach includes situations when we do not deal with bandlimited signals, that is: even if the conditions of the Whittaker-Shannon-Kotel'nikov (WSK) theorem are not met. The price to pay for that is (obviously) that a perfect reconstruction is not possible, but a simple, straightforward MSE-optimal solution may be an attractive goal in many situations anyway.

We stress that the essential mathematical principles employed here date back to Gauss and Wiener, and that the decisive point is to use these principles in an unbiased manner. The richness of this theory lies partly on the fact that it is easily adaptable to different tasks in signal and image processing, such as optimal filtering [5]. We focus on this work in the reconstruction from regularly spaced noisy discrete samples, also known as *smoothing* or *approximation* [6]. Statistical approaches to reconstruction from samples have received little attention in the signal and image processing literature during the recent two decades, whereas early extensive work using Wiener filtering (e.g. [7], or [8]) seems to be forgotten or pushed aside, possibly due to the extensive usage of formulations in the Fourier domain, which do not really simplify the exposition and practical application in case of discrete signals. However, more recently, statistical reconstruction methods seem to regain attention. These methods can be classified as discrete or continuous approaches. On the discrete side, the work of Leung et al. [9] on image interpolation shows a comparison on the performance between several ACF image models for both ideal and nonideal sampling. Shi and Reichenbach [10] derive the Wiener filter for 2-D images in the frequency domain, and propose a parametric Markov random field to model the ACF from the sampled (low resolution) data. On the continuous counterpart, Ruiz-Alzola et al. [6] present a comparison between Kriging (a quite popular interpolation method in geostatistics) and Wiener filtering, based on finite sets of noisy samples. A short section in a contribution by Ramani et al. [11] determines the filter in the frequency domain.

In contrast to that, we proceed as follows: we first find the optimal MMSE estimator in the case where only a finite number of samples are available. This is also done in [6], but in contrast to that paper we also provide the connection to the case of infinitely many discrete samples. The formula obtained is shown to be equivalent to the frequency domain version presented in [12], but both its derivation as well as its final structure are simpler than in [12]. A short section provides the link between the proposed formula and the WSK theorem. For completeness, we review the usual discrete-discrete approach, as found for instance in [13, 8], and illustrate with some experimental results.

2. NOTATION AND FUNDAMENTAL ASSUMPTIONS

In this work, we denote discrete signals with brackets, e.g., $c[k], k \in \mathbb{Z}$ and continuous signals with parenthesis, e.g., $s(x), x \in \mathbb{R}$. The continuous-space Fourier transform of a signal $s(x)$ is expressed as $S(\omega)$ and the discrete-space Fourier transform of a sequence $c[k]$ is expressed as $C(e^{j\omega})$. Discrete convolution is indicated with an asterisk $(*)$ and for its continuous counterpart a star $(\star)$ is employed. We write $\bar{f}(x)$ (resp. $\bar{p}[k]$) for the time-reversed function $\bar{f}(x) = f(-x)$ (resp. $\bar{p}[k] = p[-k]$). Each of the stochastic processes considered here is a zero-mean wide-sense stationary (WSS) process, unless explicitly stated otherwise. For such a process $\{s\}$, we denote by $r_{ss}(d)$ its autocorrelation (or autocovariance)

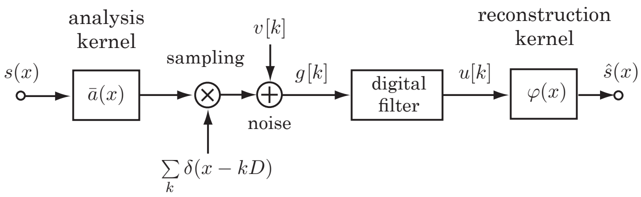

Figure 1: The block diagram represents the sampling-reconstruction problem.

function (ACF), namely,

$$r_{ss}(d) \stackrel{def}{=} \mathsf{E}\left[s(x)\cdot s(x+d)\right],\ x \text{ arbitrary}.$$

For a random vector $\vec{g}$, $\mathbf{C}_g$ denotes its covariance matrix,

$$\mathbf{C}_g \stackrel{def}{=} \mathsf{Cov}\left[\vec{g},\vec{g}\right] = \mathsf{E}\left[\vec{g}\cdot\vec{g}^T\right].$$

For reasons of compact and simple presentation, we address the problem in a single dimension; generalization to higher dimensions is straightforward. The process of sampling and reconstruction can be summarized as follows: the input signal $s(x)$ is sampled with a sampling device characterized by the *analysis kernel*[1] $a(x)$ and the sampling raster width D. Thus, if we define the function $z(x)$ as the convolution between the signal s and the kernel a, i.e.,

$$z(x) = (s \star \bar{a})(x), \tag{1}$$

then the samples $g[k]$ are given by

$$g[k] = z(kD) + v[k], \quad k \in \mathbb{Z},$$

where v represents a zero-mean noise term with known covariance function r_{vv}.

The samples are called *ideal* (or the sampling process is said to be ideal) if the analysis kernel $a(x)$ is equal to the Dirac impulse $\delta(x)$. In this case and in the abscense of noise, the sample $g[k]$ agrees with the signal value $s(kD)$.

The reconstruction aims at finding a linear estimate $\hat{s}(x)$ of $s(x)$ from the samples $g[k]$ of the form

$$\hat{s}(x) = \sum_{k=-\infty}^{+\infty} g[k]\cdot r_{\text{rec}}(x-kD), \tag{2}$$

that is optimum in the least squares sense, that is: the second power Q of the error signal should be minimized. We call the function r_{rec} the reconstruction kernel. In contrast to e.g.[14], the expectation is performed over the error signal process as well as over $\{s(x)\}$, that is: $s(x)$ is considered to be a realization of a random process. In agreement with other authors [1, 12], we show in Section 4.2, that the reconstruction process can be formulated as a two-step process, by first applying a discrete filter on the samples, and secondly convolving the resulting sequence $\{u[k]\}$ with a continuous *'generating function '* $\varphi(x)$. Symbolically,

$$\hat{s}(x) = \sum_{k=-\infty}^{+\infty} u[k]\cdot\varphi(x-kD). \tag{3}$$

The diagram shown in Fig.1 illustrates this process.

[1] see [2, p. 571]

3. OBTAINING THE COVARIANCE FUNCTIONS

For readers familiar with linear estimation theory, it is not at all surprising that the optimal reconstruction $\hat{s}$ depends only on the covariances $\mathsf{Cov}\left[s(x),s(\tilde{x})\right] = r_{ss}(x-\tilde{x})$, $\mathsf{Cov}\left[s(x),g[k]\right]$, and $\mathsf{Cov}\left[g[k],g[\ell]\right]$. In this section, we express the last two covariances in terms of the signal autocorrelation function r_{ss}, the sampling kernel a, and the noise autocorrelation r_{vv}, by adapting results from the theory of linear systems with stochastic inputs.

Eq.1 is the defining expression for the linear shift-invariant (LSI) system with stochastic input $s(x)$, impulse response $\bar{a}$, and output $z(x)$. According to the well-known correlations formulas for LSI systems (cf. [15, p.272]), an straightforward adaption to our current setting shows that

$$\begin{aligned} r_{sz}(\tau) &= (r_{ss}\star a)(\tau), &(4)\\ r_{zz}(\tau) &= (r_{ss}\star a\star\bar{a})(\tau). &(5)\end{aligned}$$

Assume furthermore that the noise and the signal are uncorrelated. Then, we obtain

$$\mathsf{Cov}\left[g[k],s(x)\right] = \mathsf{Cov}\left[z[kD],s(x)\right] = (r_{ss}\star a)(x-kD), \tag{6}$$

as a consequence of eq.4. Similarly,

$$\begin{aligned} \mathsf{Cov}\left[g[i],g[\ell]\right] = &\ \mathsf{Cov}\left[z[iD],z[\ell D]\right] + \mathsf{Cov}\left[v[i],v[\ell]\right]\\ = &\ (r_{ss}\star a\star\bar{a})((\ell-i)D) + r_{vv}[\ell-i], \quad (7)\end{aligned}$$

following eq.5.

4. MMSE ESTIMATION IN THE MIXED CONTINUOUS-DISCRETE CASE

4.1 Estimation from finite noisy samples

Let us assume that we observe K samples from the signal s, which are assembled in the vector $\vec{g} = (g[N_1],\ldots,g[N_K])^T$. There are no special requirements on the choice of the sampling locations N_i, but in our case we assume that they are equally spaced. For each point x, we design the estimate $\hat{s}(x)$ to explicitly depend linearly on the sample vector $\vec{g}$ as follows:

$$\hat{s}(x) = \vec{w}^T(x)\cdot\vec{g}. \tag{8}$$

where the weighting vector $\vec{w}$ is to be chosen in order to minimize the mean-square error $Q = \mathsf{E}\left[(s(x)-\hat{s}(x))^2\right]$.

Introducing the vector

$$\vec{f}(x) \stackrel{def}{=} (f_1(x),\ldots,f_K(x))^T$$

where

$$\begin{aligned} f_i(x) &= \mathsf{E}\left[s(x)\cdot g[N_i]\right]\\ &= (r_{ss}\star a)(x-N_iD),\end{aligned}$$

(using eq.6), and noticing that the element $c_g(i,\ell)$ of the covariance matrix $\mathbf{C}_g$ equals

$$c_g(i,\ell) = (r_{ss}\star a\star\bar{a})((N_\ell-N_i)D) + r_{vv}[N_\ell-N_i], \tag{9}$$

in virtue of eq.7, then a short algebraic manipulation shows that the optimal $\vec{w}$, given by the critical point of Q, satisfies the normal equation

$$\mathbf{C}_g\cdot\vec{w} = \vec{f}(x). \tag{10}$$

Substituting in eq.8, we conclude that the optimal estimate $\hat{s}(x)$ is given by

$$\hat{s}(x) = \vec{f}^T(x) \cdot \mathbf{C}_g^{-1} \cdot \vec{g}, \tag{11}$$

and the variance of the error term $e(x) = s(x) - \hat{s}(x)$ equals

$$\mathsf{Var}[e(x)] = r_{ss}(0) - \vec{f}^T(x) \cdot \mathbf{C}_g^{-1} \cdot \vec{f}(x).$$

4.2 The case of infinitely many noisy samples

Let us address the case where the sample sequence $\vec{g}$ is infinitely long, i.e., $\vec{g} = (g[k])_{k\in\mathbb{Z}}$. Then we need to reinterpret eq.10 and eq.11 accordingly. In particular, $\mathbf{C}_g$ can be interpreted as a linear operator that maps an (infinitely long) vector $\vec{w}$ to

$$\mathbf{C}_g \cdot \vec{w} = (\vec{c}_i \cdot \vec{w})_{i\in\mathbb{Z}} \tag{12}$$

where, for a fixed i, $\vec{c}_i$ is the 'i-th row" of $\mathbf{C}_g$, that is, $\vec{c}_i$ given by eq.9, and the product in 12 should be understood in some appropriate sequence space, ℓ^2 being the usual choice. Another interpretation is to characterize $\mathbf{C}_g$ as a convolution operator: in fact, eq.12 can be written as

$$\mathbf{C}_g \cdot \vec{w} = \vec{t} * \vec{w},$$

where

$$t[i] = (r_{ss} \star a \star \bar{a})(x)\big|_{x=iD} + r_{vv}[i].$$

Now, assume that $\mathbf{C}_g$ is an invertible operator and let $\vec{u}$ be the unique vector such that

$$\mathbf{C}_g \cdot \vec{u} = \vec{g}.$$

In terms of convolution, $\vec{t} * \vec{u} = \vec{g}$. Therefore, the formula for the optimal estimate $\hat{s}(x)$, derived in eq.11, reduces to

$$\hat{s}(x) = \vec{f}(x)^T \cdot \vec{u}.$$

Expanding this product we obtain

$$\begin{aligned} \hat{s}(x) &= \sum_{i=-\infty}^{\infty} u_i \cdot f_i(x) \\ &= \left(\sum_{i=-\infty}^{\infty} u_i \cdot \delta(x - iD) \right) \star (r_{ss} \star a)(x). \end{aligned} \tag{13}$$

From this formula, we can fully describe the reconstruction process as a two-step process:

(i) The samples $\vec{g}$ are linearly filtered by a discrete-discrete filter characterized by the matrix $\mathbf{C}_g^{-1}$, whose impulse response is given by the inverse under convolution of the vector $\vec{t}$.

(ii) The filtered samples $\vec{u}$ are linearly filtered with a discrete-continuous filter whose impulse response is given by $\varphi(x) = (r_{ss} \star a)(x)$.

This completes eq.3 and the reconstruction diagram in Fig.1.

4.3 Connection with Fourier domain formulation

In [11], the authors express the reconstruction process just described (for the case $D = 1$) by rewriting the estimation problem in the frequency domain and expressing the discrete-discrete filter in terms of its Fourier transform. We proceed to verify that the result obtained in [11] is (as expected) equivalent to the one contained in eq.13.

Let p_W be the convolution inverse of $\vec{t}$, and let $q_W = \varphi$ (the notation is chosen in order to match that of [11]). By definition of the convolution inverse and the formula of $\vec{t}$, we have that

$$p_W[k] * \left(r_{vv}[k] + (q_W \star \bar{a})(x)\big|_{x=k}\right) = \delta[k].$$

Taking Fourier transform on both sides, we obtain

$$P_W(e^{j\omega}) \cdot \left(R_{vv}(e^{j\omega}) + \sum_{k=-\infty}^{\infty} Q_W(\omega + 2\pi k) \cdot A(\omega + 2\pi k) \right) = 1.$$

Therefore, the impulse response of the discrete-discrete filtering step in Fourier domain equals

$$P_W(e^{j\omega}) = \frac{1}{\left(R_{vv}(e^{j\omega}) + \sum_{k=-\infty}^{\infty} Q_W(\omega + 2\pi k) \cdot A(\omega + 2\pi k) \right)},$$

which coincides with formulas (49) and (50) given by Ramani, Van de Ville, Blu, and Unser in [11].

4.4 Connection to WSK theorem

As a simple illustration, we show that the WSK theorem is a particular case of eq.13. The WSK conditions assume that the analysis kernel is ideal, namely, $a(x) = \bar{a}(x) = \delta(x)$, and the process s to be bandlimited. For simplicity, let the process $\{s\}$ be *uniformly bandlimited* to the frequency $f_{\max} = \frac{1}{2D}$. Thus, its power spectrum density is given by $R_{ss}(f) = \mathsf{rect}(D \cdot f)$, and consequently, its autocorrelation function is given by $r_{ss}(d) = \frac{1}{D}\mathsf{sinc}\left(\frac{d}{D}\right)$. Then, we see that

$$\begin{aligned} f_i(x) &= \mathsf{sinc}\left(\frac{x}{D} - i\right), \text{ and} \\ \mathbf{C}_g \cdot w &= \vec{t} * \vec{w} = \frac{1}{D}\vec{w} \end{aligned}$$

Substituting these results in eq.10 and simplifying, we obtain $\vec{w} = \left(\mathsf{sinc}\left(\frac{x}{D} - i\right)\right)_{i\in\mathbb{Z}}$ and the optimal estimate $\hat{s}(x)$ is given by

$$\hat{s}(x) = \vec{w} \cdot \vec{g} = \sum_i s(iD) \cdot \mathsf{sinc}\left(\frac{x}{D} - i\right),$$

which coincides with the WSK interpolation formula. It is also possible to check that the error is actually zero.

5. MMSE ESTIMATION IN THE DISCRETE-DISCRETE CASE

In this section, we provide a short description of the reconstruction process in the case where the original process is assumed to be discrete, in order to show how it follows as a special case of the mixed continuous-discrete setting previously studied. The discrete-discrete setting is probably the most popular among researchers since first, DSP implementations of the principles shown so far are to be implemented in terms of discrete signals anyway, and second, the necessary mathematical tools are significantly simpler to handle. As should be expected, the reconstruction has the same basic form as the one found in Section 4.

So, assume that the signal $\vec{s} = (s[1], s[2], \ldots s[M])^T$ is a realization of a WSS discrete process $\{s[n]\}$ and the samples $g[k]$, arranged in the vector $\vec{g} = (g[1], g[2], \ldots g[K])^T$ are obtained by performing the scalar product of the sequence $\{s[n]\}$ with an analysis kernel sequence $a_k[n]$ according to $g[k] = \vec{a}_k^T \vec{s} + v[k]$. In matrix form, $\vec{g} = \mathbf{H} \cdot \vec{s} + \vec{v}$, where $\mathbf{H}$ is the matrix whose ith row equals $\vec{a}_i^T$, for $i = 1, \ldots, M$. It is assumed that there exists a positive integer D such that $M = KD$ (subsampling factor).

If the estimate $\hat{s}[m]$ of $s[m]$ is modeled as a linear function of $\vec{g}$ according to $\hat{s}[m] = \vec{w}_m^T \vec{g}$, then following the same considerations for the mixed continuous-discrete approach, one can show that the operator $\mathbf{C}_g$ and the vector $\vec{f}$ are given by

$$\mathbf{C}_g = \mathbf{H}\mathbf{C}_s\mathbf{H}^T + \mathbf{C}_v, \qquad \vec{f} = \mathbf{H}\mathbf{C}_s^T,$$

and thus the optimal estimate $\hat{\vec{s}}$ (in the MMSE sense) of the vector $\vec{s}$ is given by

$$\hat{\vec{s}} = \mathbf{C}_s\mathbf{H}^T \left(\mathbf{H}\mathbf{C}_s\mathbf{H}^T + \mathbf{C}_v\right)^{-1} \vec{g}.$$

This is of course the well-known discrete-discrete Wiener/Bayes estimator formula available in [13, p.364] and in [8, p.292].

6. EXPERIMENTS

In this section, we illustrate the method presented before using two types of process. The first one is a third order autoregressive process AR(3) formed by applying three times the impulse response $h(s) = 1/(1 - \alpha s)$ to a white noise process, with $\alpha = 0.95$. We will label this process as AR(3,α).

For this value of α, the process AR(3,α), although non-bandlimited, is quite smooth and is thus a study case for which the sampling-reconstruction process does not strongly deviate from the WSK conditions.

The second type of process is a first order normalized AR(1) or Markov process, whose ACF is given by $r_{ss}(d) = \alpha^{|d|}, 0 < \alpha < 1$. We denote this process by AR(1,α).

The sampling process for both type of processes was performed using the 'averaging' analysis function $a(x) = \mathrm{rect}(x/\tau)$. The values of τ used were $\tau = 1$ (maximum width), $\tau = 3/5$ (intermediate width) and $\tau \to 0$ (ideal sampling). The SNR (in db) was computed according to $10\log_{10}(\sigma_s^2/\sigma_v^2)$.

Reconstruction results from nonideal noisy samples for AR(3,α) are shown in Fig.2. We remark that, due to the presence of noise, the optimal reconstruction does not necessarily pass through the samples, as they are just partially reliable.

In Fig.3 we show reconstruction kernels, that is, the function $r_{\mathrm{rec}}(x)$ in eq.2, for the process AR(3,α), with nonideal sampling in Fig.3 a) and with ideal sampling in Fig.3 b). In this case, the non-δ kernel $a(x)$ has an stretching effect on the shape of the kernel, due to the averaging operation performed according to eq.13. The effect is more noticeable under favorable noise conditions. Moreover, for both plots presented in Fig.3, the noise has a clearly visible damping effect on the reconstruction kernel, in an attempt to reduce the influence of high frequencies. Nevertheless, notice that the two dashed kernels (SNR=10 db) in Fig.3(a) and (b) are indeed very similar to each other. Thus, the magnitude of the damping effect caused by the noise term on the interpolating kernels depends on the signal statistics and the analysis kernel. This agrees with what was observed in [9].

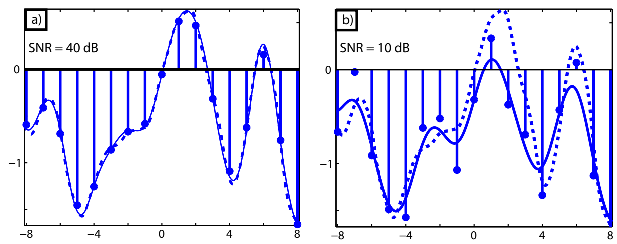

Figure 2: Examples of optimal reconstructions for the third order process AR(3,α) with $\alpha = 0.95$. Dashed line: input signal. Solid line: reconstructed signal. Filled circles: nonideal noisy samples. Both process were sampled using $\tau = 1$.

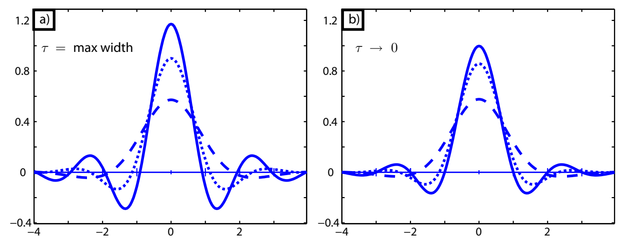

Figure 3: Examples of optimal reconstruction kernels for the third order process AR(3,α) with $\alpha = 0.95$, under different noise levels. For figure a), the process was sampled using $\tau = 1$ (maximum width), while for figure b), the sampling process was ideal. Solid line: SNR=40 db. Pointed line: SNR=20 db. Dashed line: SNR=10 db.

For the AR(1,α) process, Fig.4 shows the reconstruction results for $\alpha = 0.98$. This process is then not bandlimited but the area of the nonbandlimited part of the signal accounts only for about 10% of the total integral.

In Fig. 4 (c) the interpolation looks almost piecewise linear, but is actually of exponential type. This family of kernels are known in the literature as exponential splines [3]. Considering that Markov models are in many situations adequate for both image and geostatistics data, then it is not surprising that, in certain cases, linear interpolation outperforms other more more sophisticated polynomial splines interpolating functions. Fig. 4 (a) and Fig. 4 (b) are included in order to stress the smoothing effect of the analysis kernel on the reconstructions.

For positive values of α away from 1, AR(1,α) processes exhibit very high frequencies, thus violating drastically the WSK conditions. Therefore, any attempt to reconstruct the signal from discrete samples as performed in the previous cases renders a very rough estimate of the signal. In any case, it is interesting to regard these extreme cases for which the optimal interpolation scheme is far from being linear. Fig.5 shows reconstruction kernels for AR(1,α) with $\alpha = 0.70$ and $\alpha = 0.50$. Decreasing the value of α produces an increase on the concavity of the kernel in the ideal sampling case. This effect is in agreement with the examples shown in [3].

In summary, we see that the linear MMSE estimator is able to handle both mild and strong aliasing, as well as noise present on the samples, and the resulting interpolating function include the ubiquitous sinc and polynomial interpolation-type schemes, as well as other more exotic examples.

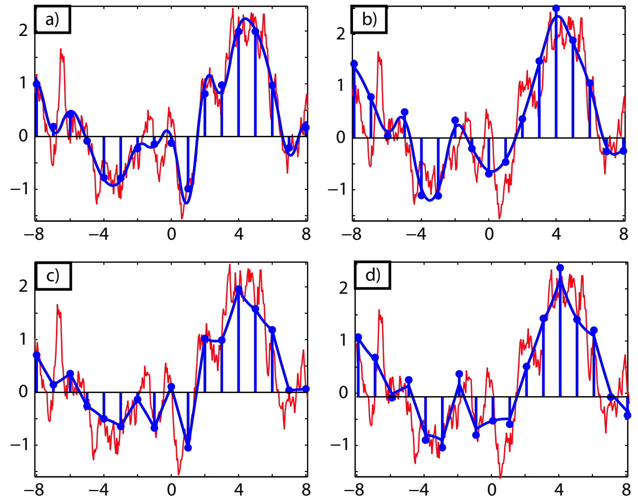

Figure 4: Examples of optimal reconstructions for the first order process AR(1,α) with $\alpha = 0.98$. The thin red line is the original signal, while the thicker blue line is the optimal reconstruction from the samples shown. Figure a): nonideal sampling with $\tau = 1$, noiseless conditions. Figure b): nonideal sampling with $\tau = 1$, SNR=10 db. Figure c): ideal sampling, noiseless conditions. Figure d): ideal sampling, SNR=10 db.

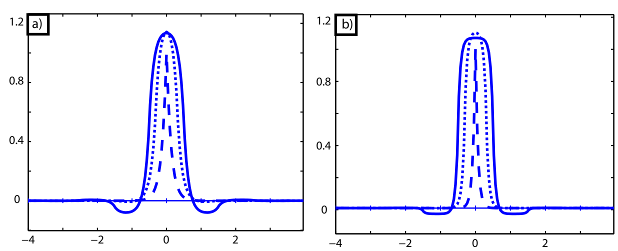

Figure 5: Examples of optimal reconstruction kernels for the first order process AR(1,α) under noiseless conditions for three different values of the width τ. Figure a) shows results for $\alpha = 0.70$, while for figure b), $\alpha = 0.50$ was used. Solid line: $\tau = 1$. Pointed line: $\tau = 3/5$. Dashed line: $\tau \to 0$.

7. CONCLUSION

We have shown how to obtain a mixed continuous-discrete version of the Wiener/Gauss-Markov estimation technique commonly known for discrete data. It is applicable to the problem of interpolating a signal from nonideal samples, under the criterion of minimizing the MSE. For this approach, it is not necessary to assume a bandlimited input process; the derivation is straightforward and does not need a frequency domain formulation. There is a strong relation of our results shown here both with [4, 11] as well as with the classical 'smoothing splines' approach by Wahba [16]. The most important difference to all these works is that the interpolation kernel is determined here directly from the measurable (!) statistical properties of the regarded signals, and does not require to set up *a priori* a class of signals (polynomial splines, exponential splines) in which the solution has to reside. Furthermore, with given statistical models for the signal and the noise there is no need to reason about the determination of a *'regularization parameter'* λ, since the correct balance between the measurement model and the prior follows directly from the theoretical framework, thanks to Mr. Wiener.

Acknowledgment

We gratefully acknowledge funding by the German Federal Ministry of Education and Research (BMBF) in the project Bernstein Fokus Neurotechnologie – Frankfurt Vision Initiative 01GQ0841.

REFERENCES

[1] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 2915–2925, 1994.

[2] M. Unser, "Sampling-50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, 2000.

[3] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 2146–2159, 2005.

[4] Y. Eldar and M. Unser, "Nonideal sampling and interpolation from noisy observations in shift-invariant spaces," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2636–2651, 2006.

[5] M. Mühlich and R. Mester, "A statistical unification of image interpolation, error concealment, and source-adapted filter design," in *6th SSIAI*, 2004, pp. 128–132.

[6] J. Ruiz-Alzola, C. Alberola-López, and C. Westin, "Kriging filters for multidimensional signal processing," *Signal Processing*, vol. 85, no. 2, pp. 413–439, 2005.

[7] H. C. Andrews and B. R. Hunt, "Digital image restoration." Englewood Cliffs, N.J.: Prentice-Hall, 1977.

[8] A. Jain, *Fundamentals of digital image processing*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1989.

[9] W. Leung, P. Bones, and R. Lane, "Statistical interpolation of sampled images," *Optical Engineering*, vol. 40, pp. 547–553, 2001.

[10] J. Shi and S. Reichenbach, "Spatially Constrained Wiener Filter with Markov Autocorrelation Modeling for Image Resolution Enhancement," in *ICIP*, 2006, pp. 2681–2684.

[11] S. Ramani, D. Van De Ville, T. Blu, and M. Unser, "Nonideal sampling and regularization theory," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1055–1070, 2008.

[12] S. Ramani, D. Van De Ville, and M. Unser, "Sampling in practice: Is the best reconstruction space bandlimited?" in *ICIP 2005*, vol. 2, 2005.

[13] S. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall Upper Saddle River, New Jersey, 1993.

[14] Y. Eldar, "Mean-squared error sampling and reconstruction in the presence of noise," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4619–4633, 2006.

[15] A. Papoulis, *Probability, random variables and stochastic processes*. McGraw-Hill, 1984.

[16] G. Wahba, *Spline Models for Observational Data*. SIAM, 1990.