OPTIMUM ANALOG RECEIVE FILTERS FOR DETECTION AND INFERENCE UNDER A SAMPLING RATE CONSTRAINT

Manuel Stein, Andreas Lenz, Amine Mezghani, and Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Technische Universität München, Munich, Germany

ABSTRACT

The problem of optimum analog receive filtering for digital signal detection and parameter estimation is considered. Here the case of a signal source with bandwidth B_t and a receiver with fixed sampling rate f_s is discussed under the assumption that $2B_t > f_s$. We investigate the impact of adjusting the receive bandwidth B_r of the analog pre-filter, which is applied prior to the sampler, with respect to the *deflection coefficient* or the *Fisher information measure*. This reveals that the design rule $2B_r < f_s$, known as the sampling theorem, does not necessarily lead to optimum system performance. Studying the two analytical information measures under a fix sampling rate f_s and an arbitrary choice of B_r, we provide an example where receive setups with $2B_r > f_s$ achieve higher detection and parameter estimation performance.

Index Terms— analog filtering, detection theory, estimation theory, sub-Nyquist sampling, satellite-based positioning

1. INTRODUCTION

The design of future mobile signal processing systems becomes challenging as it asks for high performance under severe cost, power and complexity constraints. In particular analog-to-digital conversion (ADC) turns out to be a bottleneck for the design of energy efficient receive systems [1]. Especially, when signals of high bandwidth have to be received and processed, the complexity of the ADC has to be taken into consideration. In such cases the well-known sampling theorem [2–4], which provides a sufficient condition for perfect signal reconstruction, guides the engineer's decision to increase the sampling rate along with the bandwidth of the analog pre-filter in order to enhance system performance. While the analog filter is not critical with respect to runtime complexity, the high ADC rate makes this option unattractive. However, many signal processing tasks do not require full signal reconstruction and therefore the sampling theorem is not necessarily the right design rule when heading for an optimum trade-off between system complexity and performance. The works concerned with *compressed sensing* [5–7] show that signal processing is in general possible with random sub-Nyquist sampling, while [8] explores the sampling strategy which maximizes channel capacity. In contrast the discussion here focuses on performance gains for signal detection and estimation problems which can be attained within the classical sampling framework by shaping signal aliasing through the analog pre-filter. Deriving analytical characterizations of detection and estimation theoretic measures under sub-Nyquist sampling, we analyze the optimum bandwidth of the analog pre-filter under a fixed sampling rate. The results show that for example satellite-based navigation receivers (GPS, Galileo, etc.) which violate the sampling theorem have the potential to attain higher performance with the same sampling rate as receivers which follow the traditional filter design.

The authors gratefully acknowledge the German Federal Ministry of Economics and Technology for supporting this work with grant 50NA1110.
Contact authors: {manuel.stein, josef.a.nossek}@tum.de

2. SYSTEM MODEL

2.1. Analog Transmit and Receive Signal

For the discussion, we assume a periodic transmitter $\breve{x}(t) \in \mathbb{R}$

$$\breve{x}(t) = \sum_{m=-\infty}^{\infty} b_{\mathrm{mod}\,(m,M)} \breve{g}(t - mT_c), \tag{1}$$

where $\boldsymbol{b} \in \{-1, 1\}^M$ is a sequence of M binary symbols, each of duration T_c, $\mathrm{mod}\,(\cdot)$ is the modulo-operation and $\breve{g}(t)$ is a pulse with one-sided bandwidth B_t. The receive sensor

$$\breve{y}(t) = \gamma \breve{x}(t - \tau) + \breve{\eta}(t) \tag{2}$$

attains a copy of the transmit signal $\breve{x}(t)$, attenuated by $\gamma \in \mathbb{R}$ and delayed by $\tau \in \mathbb{R}$. Thermal noise and possible interference are taken into account within the receive model by an additive random signal $\breve{\eta}(t)$. During the following pages we will approximate $\breve{\eta}(t)$ by a white and wide-sense stationary Gaussian random process with power spectral density $\frac{N_0}{2}$, i.e.

$$\breve{r}(t) = \int_{-\infty}^{\infty} \breve{\eta}(\lambda)\breve{\eta}(\lambda - t)\mathrm{d}\lambda = \frac{N_0}{2}\delta(t) \qquad \forall t.$$

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

2.2. Analog Pre-Filtering and Sampling

For further processing the receive signal is band-limited by an ideal analog low-pass filter $h(t; B_r)$ with one-sided bandwidth B_r, i.e.

$$H(\omega; B_r) = \begin{cases} 1 & \text{if } |\omega| \leq 2\pi B_r \\ 0 & \text{else,} \end{cases}$$

where $H(\omega; B_r) = \mathcal{F}\{h(t; B_r)\}$ with $\mathcal{F}\{\cdot\}$ being the Fourier-transform. Consequently, the analog sampler input is

$$y(t) = \breve{y}(t) * h(t; B_r) = \gamma x(t - \tau) + \eta(t). \tag{3}$$

For one signal period $T_o = MT_c$ the analog signal is sampled at a rate of $f_s = \frac{1}{T_s}$, resulting in a digital block signal model

$$\boldsymbol{y} = \gamma \boldsymbol{x}(\tau) + \boldsymbol{\eta},$$

with $N = \frac{T_o}{T_s}$ samples, where

$$\begin{aligned} \boldsymbol{x}(\tau) &= \begin{bmatrix} x(-\tau) & x(T_s - \tau) & \dots & x((N-1)T_s - \tau) \end{bmatrix}^T \in \mathbb{R}^N \\ \boldsymbol{\eta} &= \begin{bmatrix} \eta(0) & \eta(T_s) & \dots & \eta((N-1)T_s) \end{bmatrix}^T \in \mathbb{R}^N \\ \boldsymbol{y} &= \begin{bmatrix} y(0) & y(T_s) & \dots & y((N-1)T_s) \end{bmatrix}^T \in \mathbb{R}^N. \end{aligned}$$

2.3. Noise Properties

The filtered noise process $\eta(t)$ is in general no longer white as the autocorrelation function is given by

$$r(t) = B_r N_0 \operatorname{sinc}(2B_r t),$$

such that the entries of the filtered noise covariance matrix $\boldsymbol{R} = \mathrm{E}\left[\boldsymbol{\eta}\boldsymbol{\eta}^T\right]$ are given by

$$\boldsymbol{R}_{ij} = B_r N_0 \operatorname{sinc}(2B_r T_s |i - j|).$$

3. DETECTION AND ESTIMATION PROBLEMS

With the defined system model we consider two basic problems of statistical signal processing, summarized here briefly.

3.1. Detection Problem

Given the digital receive signal $\boldsymbol{y}$, detection is concerned with distinguishing between two possible receive situations

$$\begin{aligned} \mathcal{H}_0: \quad & \boldsymbol{y} = \boldsymbol{\eta} \\ \mathcal{H}_1: \quad & \boldsymbol{y} = \gamma \boldsymbol{x}(\tau) + \boldsymbol{\eta}, \end{aligned}$$

i.e. to determine if the signal source $\breve{x}(t)$ is active or inactive while the parameters γ and τ are known constants. Under the assumption that both receive situations occur with equal probability, i.e. $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 0.5$, the error-probability

$$P_e = P(\mathcal{H}_0|\mathcal{H}_1)P(\mathcal{H}_1) + P(\mathcal{H}_1|\mathcal{H}_0)P(\mathcal{H}_0)$$

of the *maximum-likelihood detector* which decides for $\mathcal{H}_1$ if

$$p(\boldsymbol{y}|\mathcal{H}_1) > p(\boldsymbol{y}|\mathcal{H}_0),$$

is given by [9]

$$P_e = Q\left(\frac{1}{2}\sqrt{\gamma^2 \boldsymbol{x}^T(\tau)\boldsymbol{R}^{-1}\boldsymbol{x}(\tau)}\right) = Q\left(\frac{1}{2}\sqrt{D(\tau, B_r)}\right),$$

where $D(\tau, B_r)$ is referred to as deflection coefficient. With the *Q-function* $Q(\cdot)$ being a monotonically decreasing function, the optimum receiver setup $B_r^{\mathrm{D}}(\tau)$ minimizing the error-probability P_e is

$$B_r^{\mathrm{D}}(\tau) = \arg \max_{B_r \in \mathbb{R}^+} D(\tau, B_r).$$

3.2. Estimation Problem

Estimation is concerned with the inference of an unknown signal parameter from a noisy receive signal, like for example the time-delay τ. Under the assumption that γ is a constant and τ is unknown but deterministic, the optimum unbiased estimator is the *maximum likelihood estimator* [10]

$$\hat{\tau}_{\mathrm{ML}}(\boldsymbol{y}) = \arg \max_{\tau \in \mathbb{R}} p(\boldsymbol{y}; \tau).$$

As this estimator is asymptotically efficient for the considered system model, its performance with sufficiently large N is equal to the Cramér-Rao lower bound (CRLB) [10]

$$\begin{aligned} \mathrm{Var}\left(\hat{\tau}_{\mathrm{ML}}(\boldsymbol{y})\right) &= \mathrm{CRLB}(\tau, B_r) \\ &= \frac{1}{F(\tau, B_r)}, \end{aligned}$$

where the Fisher information measure is defined by

$$\begin{aligned} F(\tau, B_r) &= \int_{\mathcal{Y}} p(\boldsymbol{y}; \tau) \left(\frac{\partial \ln p(\boldsymbol{y}; \tau)}{\partial \tau}\right)^2 \mathrm{d}\tau \\ &= \gamma^2 \frac{\partial \boldsymbol{x}^T(\tau)}{\partial \tau} \boldsymbol{R}^{-1} \frac{\partial \boldsymbol{x}(\tau)}{\partial \tau}. \end{aligned}$$

The receiver setup $B_r^{\mathrm{E}}(\tau)$ minimizing $\mathrm{Var}\left(\hat{\tau}_{\mathrm{ML}}(\boldsymbol{y})\right)$ is

$$B_r^{\mathrm{E}}(\tau) = \arg \max_{B_r \in \mathbb{R}^+} F(\tau, B_r).$$

4. COMPACT ANALYTIC PERFORMANCE CHARACTERIZATION

To attain compact analytical expressions without inversion of large matrices, frequency domain characterizations of the information measures $D(\tau, B_r)$ and $F(\tau, B_r)$ are derived. Using the fact that $x(t)$ is periodic with $\omega_0 = \frac{2\pi}{T_o}$,

$$x(t) = \frac{1}{T_o} \sum_{k=-\infty}^{\infty} X(k\omega_0)\, \mathrm{e}^{\mathrm{j}\, k\omega_0 t} = \sum_{k=-\infty}^{\infty} \tilde{x}_k\, \mathrm{e}^{\mathrm{j}\, k\omega_0 t},$$

where

$$X(\omega) = \int_{T_o} x(t)\,\mathrm{e}^{-\mathrm{j}\,\omega t}\,\mathrm{d}t \tag{4}$$

and

$$\tilde{x}_k = \frac{1}{T_o} X(k\omega_0),$$

with $X(k\omega_0)$ being the Fourier transform (4) of the signal $x(t)$ evaluated at the discrete frequency point $k\omega_0$. Note that due to the signal model (1), (2) and (3)

$$X(k\omega_0) = B(k\omega_0)\breve{G}(k\omega_0)H(k\omega_0; B_r),$$

with $B(k\omega_0)$ being the coefficients of the discrete Fourier transform of the binary sequence $\boldsymbol{b}$. Defining the spectrum of the sampled signal

$$X'(k\omega_0;\tau) = \sum_{l=-\infty}^{\infty} X(k\omega_0 - l\omega_s)\,\mathrm{e}^{-\mathrm{j}(k\omega_0 - l\omega_s)\tau},$$

with $\omega_s = \frac{2\pi}{T_s}$, the signal samples with delay can be written

$$\begin{aligned}
x[n] &= \frac{1}{T_o} \sum_{k=-\infty}^{\infty} X(k\omega_0)\,\mathrm{e}^{\mathrm{j}\,k\omega_0(nT_s - \tau)} \\
&= \frac{1}{T_o} \sum_{k=-\infty}^{\infty} X(k\omega_0)\,\mathrm{e}^{-\mathrm{j}\,k\omega_0\tau}\,\mathrm{e}^{\mathrm{j}\,kn\frac{2\pi}{N}} \\
&= \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \left(\frac{1}{T_o} \sum_{l=-\infty}^{\infty} X(k\omega_0 - l\omega_s)\,\mathrm{e}^{-\mathrm{j}(k\omega_0 - l\omega_s)\tau} \right) \mathrm{e}^{\mathrm{j}\,kn\frac{2\pi}{N}} \\
&= \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \tilde{x}'_k(\tau)\,\mathrm{e}^{\mathrm{j}\,kn\frac{2\pi}{N}},
\end{aligned} \tag{5}$$

where it was used that $\omega_0 T_s = \frac{2\pi T_s}{T_0} = \frac{2\pi\omega_0}{\omega_s} = \frac{2\pi}{N}$ and $\mathrm{e}^{\mathrm{j}\,kn\frac{2\pi}{N}} = \mathrm{e}^{\mathrm{j}(k-lN)n\frac{2\pi}{N}}, \forall l \in \mathbb{Z}$. A modified version of the inverse discrete Fourier transformation matrix

$$W_{nk} = \frac{1}{\sqrt{N}}\,\mathrm{e}^{\mathrm{j}\,2\pi\frac{\left(-\frac{N}{2}-1+k\right)\left(n-1\right)}{N}},$$

with $k, n = 1, \ldots, N$ and (5) allows to write

$$\boldsymbol{x}(\tau) = \sqrt{N}\boldsymbol{W}\tilde{\boldsymbol{x}}'(\tau),$$

where the Fourier coefficient vector is defined

$$\tilde{\boldsymbol{x}}'(\tau) = \left[\tilde{x}'_{-\frac{N}{2}}(\tau) \quad \tilde{x}'_{-\frac{N}{2}+1}(\tau) \quad \ldots \quad \tilde{x}'_{\frac{N}{2}-1}(\tau)\right]^T \in \mathbb{C}^N.$$

Accordingly, with

$$\tilde{\psi}_k(\tau) = \frac{1}{T_o} \sum_{l=-\infty}^{\infty} \mathrm{j}\,l\omega_s X(k\omega_0 - l\omega_s)\,\mathrm{e}^{-\mathrm{j}(k\omega_0 - l\omega_s)\tau},$$

the sampled signal derivative is given by

$$\begin{aligned}
\frac{\partial x[n]}{\partial \tau} &= \frac{1}{T_o} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \Bigg(\sum_{l=-\infty}^{\infty} -\mathrm{j}(k\omega_0 - l\omega_s) \cdot \\
&\qquad \cdot X(k\omega_0 - l\omega_s)\,\mathrm{e}^{-\mathrm{j}(k\omega_0 - l\omega_s)\tau} \Bigg) \mathrm{e}^{\mathrm{j}\,kn\frac{2\pi}{N}} \\
&= \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \left(-\mathrm{j}\,k\omega_0 \tilde{x}'_k(\tau) + \tilde{\psi}_k(\tau) \right) \mathrm{e}^{\mathrm{j}\,kn\frac{2\pi}{N}},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\psi}_k(\tau) &= \frac{1}{T_o} \Psi(k\omega_0;\tau) \\
&= \frac{1}{T_o} \sum_{l=-\infty}^{\infty} \mathrm{j}\,l\omega_s X(k\omega_0 - l\omega_s)\,\mathrm{e}^{-\mathrm{j}(k\omega_0 - l\omega_s)\tau}.
\end{aligned}$$

Then, with the diagonal matrix

$$\Sigma_{kk} = -\mathrm{j}\,\omega_0\Big(-\frac{N}{2} - 1 + k\Big)$$

and the vector

$$\tilde{\boldsymbol{\psi}}(\tau) = \left[\tilde{\psi}_{-\frac{N}{2}}(\tau) \quad \tilde{\psi}_{-\frac{N}{2}+1}(\tau) \quad \ldots \quad \tilde{\psi}_{\frac{N}{2}-1}(\tau)\right]^T \in \mathbb{C}^N,$$

the signal derivative in vector notation can be written

$$\frac{\partial \boldsymbol{x}(\tau)}{\partial \tau} = \sqrt{N}\boldsymbol{W}\Big(\boldsymbol{\Sigma}\tilde{\boldsymbol{x}}'(\tau) + \tilde{\boldsymbol{\psi}}(\tau)\Big).$$

Note that for large N the noise covariance matrix [11]

$$\boldsymbol{R} \approx \boldsymbol{W}\boldsymbol{\Omega}\boldsymbol{W}^H,$$

where the entries of the diagonal eigenvalue matrix are

$$\Omega_{kk} = \frac{1}{T_s} R'\left(\Big(-\frac{N}{2} - 1 + k\Big)\omega_0\right),$$

with the sampled noise power spectral density

$$R'(k\omega_0) = \sum_{l=-\infty}^{\infty} R(k\omega_0 - l\omega_s).$$

This allows to write the deflection coefficient

$$\begin{aligned}
D(\tau, B_r) &= \gamma^2 N \big(\boldsymbol{W}\tilde{\boldsymbol{x}}'(\tau)\big)^H \boldsymbol{W}\boldsymbol{\Omega}^{-1}\boldsymbol{W}^H\boldsymbol{W}\tilde{\boldsymbol{x}}'(\tau) \\
&= \frac{\gamma^2}{T_0} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \frac{|X'(k\omega_0;\tau)|^2}{R'(k\omega_0)}
\end{aligned} \tag{6}$$

and the Fisher information measure

$$\begin{aligned}
F(\tau, B_r) &= \gamma^2 N \big(\boldsymbol{\Sigma}\tilde{\boldsymbol{x}}'(\tau) + \tilde{\boldsymbol{\psi}}(\tau)\big)^H \boldsymbol{\Omega}^{-1} \big(\boldsymbol{\Sigma}\tilde{\boldsymbol{x}}'(\tau) + \tilde{\boldsymbol{\psi}}(\tau)\big) \\
&= \frac{\gamma^2}{T_0} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \frac{(k\omega_0)^2 |X'(k\omega_0;\tau)|^2}{R'(k\omega_0)} + \frac{|\Psi(k\omega_0;\tau)|^2}{R'(k\omega_0)} - \\
&\qquad - \frac{2k\omega_0\,\mathrm{Im}\left\{X'^*(k\omega_0;\tau)\Psi(k\omega_0;\tau)\right\}}{R'(k\omega_0)}.
\end{aligned} \tag{7}$$

4.1. Expected Performance Measures

The information measures (6) and (7) are dependent on the specific choice of the binary sequence $\boldsymbol{b}$. Since $\boldsymbol{b}$ is in general a white random sequence, i.e.

$$\mathrm{E}_{\boldsymbol{b}}\left[b_i b_j\right] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else,} \end{cases} \tag{8}$$

an average performance characterization, independent of $\boldsymbol{b}$, is needed. To this end, the expected informations measures

$$\begin{aligned} \bar{D}(\tau, B_r) &= \mathrm{E}_{\boldsymbol{b}}\left[D(\tau, B_r)\right] \\ &= \frac{\gamma^2}{T_0} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \frac{\mathrm{E}_{\boldsymbol{b}}\left[|X'(k\omega_0;\tau)|^2\right]}{R'(k\omega_0)} \end{aligned} \tag{9}$$

and

$$\begin{aligned} \bar{F}(\tau, B_r) &= \mathrm{E}_{\boldsymbol{b}}\left[F(\tau, B_r)\right] \\ &= \frac{\gamma^2}{T_0} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \frac{(k\omega_0)^2\,\mathrm{E}_{\boldsymbol{b}}\left[|X'(k\omega_0;\tau)|^2\right]}{R'(k\omega_0)} + \\ &+ \frac{\mathrm{E}_{\boldsymbol{b}}\left[|\Psi(k\omega_0;\tau)|^2\right]}{R'(k\omega_0)} - \\ &- \frac{2k\omega_0\,\mathrm{Im}\left\{\mathrm{E}_{\boldsymbol{b}}\left[X'^*(k\omega_0;\tau)\Psi(k\omega_0;\tau)\right]\right\}}{R'(k\omega_0)} \end{aligned} \tag{10}$$

are introduced. From (8) it follows that

$$\mathrm{E}_{\boldsymbol{b}}\left[B(k\omega_0 - l\omega_s)B^*(k\omega_0 - l\omega_s)\right] = M \cdot \mathcal{I}(l,m),$$

where the indicator function is defined

$$\mathcal{I}(l,m) = \begin{cases} 1 & \text{if } \frac{l-m}{M}\frac{\omega_s}{\omega_0} \in \mathbb{Z} \\ 0 & \text{else.} \end{cases}$$

Therefore, the expected values in (9) and (10) can be written

$$\begin{aligned} \mathrm{E}_{\boldsymbol{b}}\left[|X'(k\omega_0;\tau)|^2\right] &= M \cdot \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \zeta(l,m;\tau) \\ \mathrm{E}_{\boldsymbol{b}}\left[|\Psi(k\omega_0;\tau)|^2\right] &= M \cdot \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} lm\omega_s^2\zeta(l,m;\tau) \\ \mathrm{E}_{\boldsymbol{b}}\left[X'^*(k\omega_0;\tau)\Psi(k\omega_0;\tau)\right] &= M \cdot \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} \mathrm{j}\, l\omega_s\zeta(l,m;\tau) \end{aligned}$$

with

$$\zeta(l,m;\tau) = \mathcal{I}(l,m)\,\mathrm{e}^{\mathrm{j}(l-m)\omega_s\tau}\, G(k\omega_0 - l\omega_s)G^*(k\omega_0 - m\omega_s),$$

where $G(\omega)$ is the Fourier transform of the band-limited pulse $g(t) = \breve{g}(t) * h(t;B_r)$.

5. OPTIMUM ANALOG FILTER BANDWIDTH

To visualize the gains with optimized bandwidth B_r, we assume $M = 1023$, a rectangular pulse $\breve{g}(t)$, $f_c = \frac{1}{T_c} = 1.023$ Mhz (GPS C/A L1), $T_o = MT_c$ and $\tau = \frac{T_c}{2}$. The sampling rate is $f_s = 2.5 f_c$ and the filter bandwidth $B_r = \alpha\frac{f_s}{2}$ with $\alpha \in \mathbb{R}^+$. For each α, the performance is evaluated and normalized to a reference system operating at $B_r = \frac{f_s}{2}$, i.e. $\alpha = 1$. Fig. 1 shows the detection performance with maximum at $\alpha = 1.32$ and Fig. 2 shows the estimation performance which attains its maximum at $\alpha = 1.40$. We notice that here the performance hardly depends on the sampling time τ.

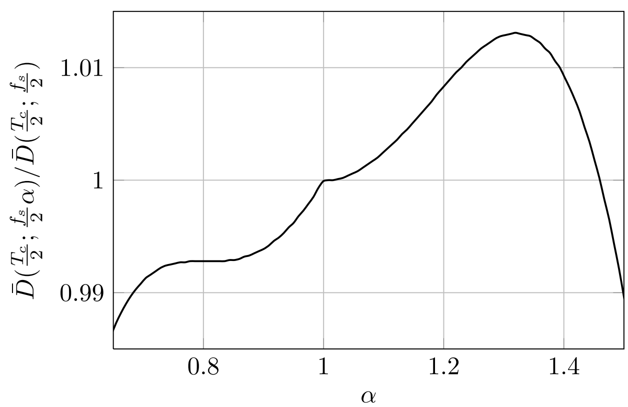

Fig. 1. Detection performance versus filter bandwidth B_r

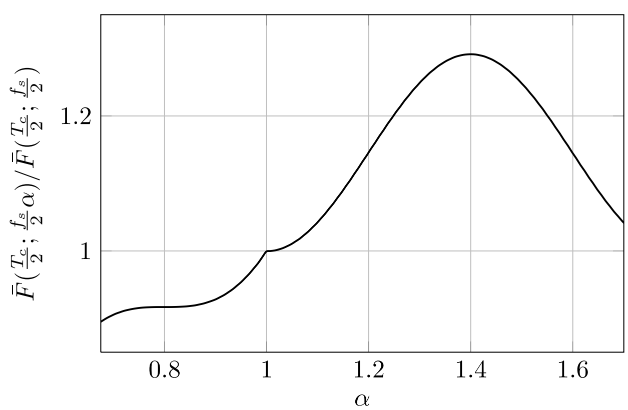

Fig. 2. Estimation performance versus filter bandwidth B_r

6. CONCLUSION

Investigation of the behavior of detection and estimation performance under sub-Nyquist sampling has shown that the performance of signal processing systems with a sampling rate constraint can be significantly increased by adjusting the analog pre-filter $h(t;B_r)$. Interestingly, if a broadband transmit signal of bandwidth B_t is received under a sampling rate constraint $2B_t > f_s$, the optimum bandwidth of the analog receive filter B_r lies above the aliasing-free region $2B_r < f_s$.

7. REFERENCES

[1] R. H. Walden, ”Analog-to-digital converter survey and analysis,” *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[2] H. Nyquist, ”Certain topics in telegraph transmission theory”, *Trans. AIEE*, vol. 47, pp. 617644, Apr. 1928.

[3] C. E. Shannon, Communication in the Presence of Noise, *Proc. IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[4] H. D. Luke, ”The origins of the sampling theorem,” *IEEE Commun. Mag.*, vol. 37, no. 4, pp.106–108, Apr. 1999.

[5] E. J. Candes, J. Romberg, T. Tao, ”Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[6] D. L. Donoho, ”Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[7] M. A. Davenport, P.T. Boufounos, M. B. Wakin, R. G. Baraniuk, ”Signal Processing With Compressive Measurements,” *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 445–460, Apr. 2010.

[8] C. Yuxin, Y.C. Eldar, A. J. Goldsmith , ”Shannon Meets Nyquist: Capacity of Sampled Gaussian Channels,” *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4889–4914, Aug. 2013

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddler River, NJ: Pretice Hall, 1998.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddler River, NJ: Pretice Hall, 1993.

[11] U. Grenander, G. Szegö, *Toeplitz forms and their applications*. Berkeley, CA: Univ. of California Press, 1958.