Bandwidth Estimation in the Sampling Theorems

Tibor Pogány¹

Abstract

The spectral representations of the Shannon-Kotelnikov sampling cardinal series expansion (SCSE) of non - bandlimited (NBL) stochastic signals were presented in some earlier papers by the author. This result gives the main tool in deriving the optimal value of the bandwidth in the approximation of the NBL stochastic signal by a sampling series of a BL signal such that it possesses the same spectral process.

In this paper the Brown aliasing error upper bound is extended to the NBL multidimensional stochastic processes and to the NBL homogeneous random fields. The magnitude of the derived bound is ordered under some smoothness condition upon the random field. Finally some statistical interpretations and methods are presented in the Shannon-Kotelnikov discretization procedure for the stochastic signals.

Keywords: Weakly stationary scalar and multidimensional processes; Homogeneous random fields; Sampling cardinal series expansion; Spectral representation; Aliasing error; Bandlimited and non-band-limited signals; Brown's upper bound.

AMS Subj.Class.(1991): 60 G 10, 60 F 25.

1 Preliminaries, some definitions

Let us define a weakly stationary (WS) random signal $\{S(t)|t \in \mathbb{R}^m\}$, ES(t) = 0, $DS(t) := E|S(t)|^2 < \infty$, on the fixed probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The following cases of these signals shall be considered:

- 1. weakly stationary scalar random processes (WSSP),
- 2. WS vectorial random processes (WSVP) and
- 3. homogeneous random fields (HRF).

¹Faculty of Maritime Studies, University of Rijeka, Studenstska 2, 51000 Rijeka, Croatia

As a main characteristic of the WS signals the correlation function shall be introduced. Namely, the correlation function of the WSSP $\{X(t)|t \in \mathbb{R}\}$ is $K_X(t) := EX(t)X^*(0)$, where * denotes the complex conjugate. The role of the correlation function for the *p*-dimensional WSVP $\mathbf{X}(t) = (X_1(t), \ldots, X_p(t))$ plays the so-called correlation matrix $\mathbf{K}(t) = (K_{jm}(t))_{p \times p}$ where $K_{jm}(t) = EX_j(t)X_m^*(0)$ is the crosscorrelation function of the WSSPs $X_j(t)$ and $X_m(t)$ for all $1 \leq j \neq m \leq p$, moreover j = m gives the correlation (or auto-correlation) function of the considered coordinate in the WSVP $\mathbf{X}(t)$. In fact, if $K_{jm}(t) \neq 0$ then $X_j(t)$ and $X_m(t)$ are stationarily correlated processes. Finally, the correlation function of the homogeneous random field $\{\xi(x)|x \in \mathbb{R}^q\}$ has been defined as $\mathcal{K}(x) := \mathbb{E}\xi(x)\xi^*(0)$.

According to the well-known Bochner-Khintchine theorem there exists the Fourier-Stieltjes integral representation:

$$K(t) = \int_{\mathbf{R}} e^{it\lambda} \mathrm{d}F(\lambda). \tag{1}$$

The WSSP process which correlation function possesses the above representation is non-band-limited (NBL), i.e. there is no interval of a positive Lebesgue measure such as $dF(\lambda) \equiv 0$ on it.

By using the Bochner - Khintchine theorem

$$K_{jm}(t) = \int_{\mathbb{R}} e^{it\lambda} dF_{jm}(\lambda), \qquad (2)$$

$$\mathcal{K}(x) = \int_{\mathbb{R}^q} e^{i(\lambda,x)} \mathrm{d}\mathcal{F}(\lambda)$$
 (3)

we can get the Fourier-Stieltjes representations of the cross-correlation function of WSSPs $X_j(t)$ and $X_m(t)$ and the correlation function of the HRF $\xi(x)$. Here(a, b) denotes the usual inner product $\sum_{i=1}^{q} a_i b_i$ of the q-dimensional vectors a, b. In all these integrals F, \mathcal{F} denote the spectral distribution functions of the considered WS signals.

If, on the other hand, $dF(\lambda) \equiv 0$ ($dF(\lambda) \equiv 0$) outside of the interval [-a, a], then the considered process is band-limited (BL) to the given bandwidth a > 0 and [-a, a] is the sampling support interval. More precisely,

$$a = \begin{cases} w & \text{WSSP} \\ w_0 = \min(w_j, w_m) & \text{WSVP} \\ W = (w_1, \dots, w_q) & \text{HRF} \end{cases}$$

Consequently the sampling interval [-a, a] becomes a q-dimensional rectangle in the HRF case. So the Fourier-Stieltjes representations of the correlation function for the BL signals are

$$K(t) = \int_{-w}^{w} e^{it\lambda} \mathrm{d}F(\lambda) \tag{4}$$

$$K_{jm}(t) = \int_{-w_0}^{w_0} e^{it\lambda} \mathrm{d}F_{jm}(\lambda)$$
 (5)

$$\mathcal{K}(x) = \int_{\sum_{j=1}^{q} [-w_j, w_j]} e^{i(\lambda, x)} \mathrm{d}\mathcal{F}(\lambda)$$
(6)

The important consequences of these facts are the following spectral representations of the NBL signals, namely

$$X(t) = \int_{\mathbb{R}} e^{it\lambda} dZ(\lambda)$$
(7)

$$\mathbf{X}(t) = \int_{\mathbf{R}} e^{it\lambda} d\mathbf{Z}(\lambda)$$
 (8)

$$\xi(x) = \int_{\mathbf{R}^q} e^{i(\lambda,x)} \mathrm{d}\mathcal{Z}(\lambda) \tag{9}$$

where $Z(\lambda)$ is the spectral process of the WSSP X(t); $Z(\lambda) = (Z_1(\lambda), \ldots, Z_p(\lambda))$ and $Z(\lambda) \equiv Z(\lambda_1, \ldots, \lambda_q)$ is the spectral field of the HRF $\xi(x)$. All spectral signals are with orthogonal increments.

From (4-6) it follows that

$$X(t) = \int_{-w}^{w} e^{it\lambda} dZ(\lambda)$$
(10)

$$\mathbf{X}(t) = \left(\int_{-w_1}^{w_1} e^{it\lambda} \mathrm{d}Z_1(\lambda), \dots, \int_{-w_p}^{w_p} e^{it\lambda} \mathrm{d}Z_p(\lambda) \right)$$
(11)

$$\xi(x) = \int_{\underset{j=1}{\overset{q}{\xrightarrow{[-w_j,w_j]}}}} e^{i(\lambda,x)} \mathrm{d}\mathcal{Z}(\lambda), \qquad (12)$$

for BL signals and $Z(\lambda)$, $Z(\lambda)$ has the same meaning as in the relation (7-9).

The connection between the stochastic signals X(t), X(t) and $\xi(x)$ and their deterministic correlation functions K(t), K(t) and $\mathcal{K}(x)$ are given by

$$dF(\lambda) = E|dZ(\lambda)|^2$$
(13)

$$\mathrm{d}F_{jm}(\lambda) = \mathrm{E}\mathrm{d}Z_j(\lambda)\mathrm{d}Z_m^*(\lambda) \tag{14}$$

$$d\mathcal{F}(\lambda) = E|d\mathcal{Z}(\lambda)|^2.$$
(15)

2 Sampling cardinal series expansions

Consider an NBL WSSP $\{X(t)|t \in \mathbb{R}\}$ and the given bandwidth w > 0. The sampling cardinal series expansion (SCSE) of the signal X(t) with respect to the bandwidth w reads as follows:

$$X_{w}(t) := \sum_{-\infty}^{+\infty} X(\frac{n\pi}{w}) \operatorname{sinc}(wt - n\pi)$$
(16)

where

$$\operatorname{sinc}(u) := \begin{cases} u^{-1} \sin(u) & u \neq 0 \\ 1 & u = 0 \end{cases}$$
(17)

Accordingly, the vectorial SCSE $X_w(t)$ of the *p*-dimensional WSVP $\{X(t) = (X_1(t), \dots, X_p(t)) | t \in \mathbb{R}\}$ is introduced in the paper Pogány, Peruničic (1992) as

$$\mathbf{X}_{\mathbf{w}}(t) := \left(X_{w_1}^{(1)}(t), \dots, X_{w_p}^{(p)}(t) \right),$$
(18)

where $X_{w_k}^{(k)}(t)(t)$ denotes the SCSE to w_k of $X_k(t)$, $k = \overline{1, p}$, i.e.

$$X_{w_k}^{(k)}(t) := \sum_{-\infty}^{+\infty} X(\frac{n\pi}{w_k}) \operatorname{sinc}(w_k t - n\pi).$$

Finally the SCSE $\xi_{\mathcal{W}}(x)$ of the NBL HRF $\xi(x)$ with respect to the vectorial bandwidth \mathcal{W} is given by

$$\xi_{\mathcal{W}}(x) := \sum_{j=1}^{q} \sum_{-\infty}^{+\infty} \xi(\frac{n_1 \pi}{w_1}, \dots, \frac{n_q \pi}{w_q}) \times \prod_{j=1}^{q} \operatorname{sinc}(w_j x_j - n_j \pi).$$
(19)

In fact, if the WS signal is BL to the bandwidth w > 0 then it is also BL to any larger bandwidth $w_1 > w$, Belyaev (1959). So it makes no sense to consider the SCSE of a BL signal with respect to any greater bandwidth then the initial one. (Of course it is possible to introduce an SCSE for a BL process (with tighter bandwidth then the initial one), but we are not interested in this point of view here). Now if we consider BL WS signals when the given initial bandwidth coincides with the SCSE-bandwidth, then the following classical results hold:

$$X(t) = X_{\omega}(t) \tag{20}$$

$$\mathbf{X}(t) = \mathbf{X}_{\mathbf{w}}(t) \tag{21}$$

$$\xi(x) = \xi_{\mathcal{W}}(x) \tag{22}$$

where the equalities are used in the mean-square sense. But these results are valid only under some continuity conditions upon $Z(\lambda)$, $Z(\lambda)$ and $Z(\lambda)$ at the end points (or vertices) of the sampling interval [-a, a], consult Balakrishnan (1957), Beutler (1961), Pogány (1989; 1991 a; 1991 b) and (Wong 1971:105-106) for more details.

3 The problem

Consider an NBL WS random signal. There are only few realistic problems in the measurement procedures in the frequency domain as well as in the time domain. More precisely speaking the concept of the band-limitation and the durationlimitation cannot happen simultaneously. Therefore it is of importance to approximate NBL signals with BL signals on the same probability space, i.e. with the signal possessing the same spectral signal that the nominal one, Slepian (1976). Therefore we have the following

WSSP-PROBLEM. Derive the bandwidth w > 0 in the mean-square approximation

$$X(t) = \int_{\mathbb{R}} e^{it\lambda} dZ(\lambda) \stackrel{\text{m.s.}}{\approx} \int_{-w}^{w} e^{it\lambda} dZ(\lambda) = X_{w}(t)$$

so that the so-called aliasing error (AE) $\mathcal{E}_{w}^{X}(t)$ satisfies:

$$\mathcal{E}_{w}^{X}(t) := \mathsf{E}|X(t) - X_{w}(t)|^{2} \le \varepsilon^{2}$$
(23)

where $\varepsilon > 0$ is the given approximation error level.

WSVP-PROBLEM. Derive the bandwidth w > 0 in the meansquare approximation $\mathbf{X}(t) \approx \mathbf{X}_{\mathbf{w}}(t)$ so that the cross-aliasing error satisfies

$$|\mathcal{E}_{w}^{X_{j},X_{k}}|(t) := |\mathsf{E}(X_{j}(t) - X_{w}^{(j)}(t))(X_{k}(t) - X_{w}^{(k)}(t)^{*}| \le \varepsilon^{2},$$
(24)

for the given $\varepsilon > 0$, $1 \leq j$, $k \leq p$.

HRF-PROBLEM. Derive $W = (w_1, \ldots, w_q)$ in the mean-square approximation $\xi(x) \approx \xi_W(x)$, so that

$$\mathcal{E}_{\mathcal{W}}^{\ell}(t) := \mathsf{E}|\xi(x) - \xi_{\mathcal{W}}(x)|^2 \le \varepsilon^2, \tag{25}$$

for the given $\varepsilon > 0$.

There will be no difficulties in recognizing when $(\cdot)_w(t)$ denotes the BL WS process, vectorial process or field and when it denotes its sampling cardinal series expansion to the bandwidth w, in the further exposition of the matter (in fact these quantities are equal to each other, see (20-22).

4 Some earlier results

Before solving our principal problems, we shall give some earlier results in the meansquare AE upper bound ordering. Therefore these results will be listed in brief manner.

The WSSP X(t) possesses the mean-square derivative of the order r if its 2th moment $M_{2r}(X)$ is finite, i.e.

$$M_{2r}(X) := \int_{\mathbb{R}} \lambda^{2r} \mathrm{d}F(\lambda) \equiv |K^{(2r)}(0)| < \infty.$$
(26)

It is well-known that the Cauchy-Schwarz inequality applied to the cross-spectral distribution function $F_{jk}(\lambda)$ of the j th and k th coordinates in WSVP $\mathbf{X}(t)$ results in

$$|\mathrm{d}F_{jk}(\lambda)|^2 \le \mathrm{d}F_{jj}(\lambda)\mathrm{d}F_{kk}(\lambda). \tag{27}$$

Using this inequality we can show that the existence of the derivatives of coordinate WSSPs in $\mathbf{X}(t)$ is sufficient for the existence of the cross-moment

$$|M_{r,s}^{X_j,X_k}(t)| := |\int_{\mathbb{R}} \lambda^r \lambda^s \mathrm{d}F_{j,k}(\lambda)| < \infty.$$

Similarly to the above definition we introduce the mean-square mixed derivative of the order $|r| = \sum_{j=1}^{q} r_j$ of a HRF $\xi(x)$. This derivative exists if the 2|r|th mixed moment of the considered field is finite, in other words if

$$\mathcal{M}_{2|\mathbf{r}|}(\xi) := \int_{\mathbf{R}^q} \lambda^{2|\mathbf{r}|} \mathrm{d}\mathcal{F}(\lambda) \equiv \left| \frac{\partial^{2|\mathbf{r}|}}{\partial x^{2|\mathbf{r}|}} \mathcal{K}(0) \right| < \infty.$$
(28)

THEOREM 1 (Brown, 1978) Let X(t) be a NBL WSSP, w > 0. Then we have

$$\mathcal{E}_{w}^{x}(t) \leq 4 \int_{|\lambda| > w} \mathrm{d}F(\lambda),$$
 (29)

where the constant 4 is sharp.

THEOREM 2 (Pogány, 1993) Let X(t) be an r-fold differentiable NBL WSSP. Then we have

$$\mathcal{E}_{w}^{x}(t) \leq 4w^{-2r} |K^{(2r)}(0)|. \tag{30}$$

THEOREM 3 (Pogány-Peruničic, 1992) Let X(t), Y(t) be stationarily correlated NBL WSSPs. Then $X_{w'}(t)$ and $Y_{w''}(t)$ are also stationarily correlated iff $w' \equiv w''$. It also holds

$$|\mathcal{E}_{w}^{x,y}(t)|^{2} \leq \mathcal{E}_{w}^{x}(t)\mathcal{E}_{w}^{y}(t).$$
(31)

THEOREM 4 (Pogány, 1993) Let $\xi(x)$ be a NBL HRF, W > I. Then if $\xi(x)$ is $|s| = \sum_{i=1}^{q} s_i$ -fold differentiable HRF we have

$$\mathcal{E}_{\mathcal{W}}^{\xi}(x) \leq 4\mathcal{W}^{-2|s|} \left| \frac{\partial^{2|r|}}{\partial x^{2|r|}} \mathcal{K}(0) \right|, \tag{32}$$

where $a^{|\gamma|s|} \stackrel{\Delta}{=} a_1^{\gamma s_1} \cdots a_q^{\gamma s_q}$.

5 Bandwidth estimation

In solving the **PROBLEM** we use the results of chapter 4. The approach will be robust, but very clear and simple. At first, by means of the Brown's theorem we can formulate for the given mean-square approximation error level ε the following

SOLUTION 1. Let X(t) be an r-fold differentiable NBL WSSP, and let $\varepsilon > 0$ be the given error level in the sense of the relation (23). Then

$$w \ge (2/\varepsilon)^{1/\tau} |K^{(2\tau)}(0)|^{1/2\tau}$$
(33)

With the help of the foregoing evaluation (33) and the variant of the Cauchy-Schwarz inequality (28) let us consider two stationarily correlated r-, s-fold differentiable NBL WSSPs X(t) and Y(t) respectively. For both of these processes we have the evaluation (33). Also $w_x = w_y \equiv w$.

SOLUTION 2. Let $\mathbf{X}(t) = (X_1(t), \dots, X_p(t))$ be a NBL WSVP and $w_1 = \dots = w_p \equiv w$ and ε be the given error level. Then

$$w \geq \max_{1 \geq j,k \geq p} \left[\frac{4}{\varepsilon^2} |K_{jj}^{(2s_j)}(0) K_{kk}^{(2s_k)}(0)|^{1/2} \right]^{1/(s_j+s_k)}$$
(34)

if $X_j(t)$ possesses s_j mean-square derivatives, $j = \overline{1, p}$.

According to the result of the theorem 4. we clearly get the bandwidth estimate for the NBL HRF case. However, the WSSP and WSVP cases, the general solution is not unique:

SOLUTION 3. Let $\xi(x)$ be a NBL, |s|-fold differentiable HRF, $\varepsilon > 0$. Then we have

$$\prod_{j=1}^{q} w_j^{2s_j} \ge (4/\varepsilon^2) \left| \frac{\partial^{2|s|}}{\partial x^{2|s|}} \mathcal{K}(0) \right|.$$
(35)

REMARK. We need some additional information about the relations between the coordinate-bandwidths in \mathcal{B} to order an optimal value of \mathcal{B} . For instance, let us know all of the coefficients $\{\beta_j\}_{j=1}^q$ in the equality sequence

$$w_1 = \beta_2 w_2 = \ldots = \beta_q w_q \in \mathbb{R}_+. \tag{36}$$

Then (35) becomes

$$w_{1} \geq \left(\frac{2}{\varepsilon} \prod_{j=2}^{q} \beta_{j}^{s_{j}}\right)^{1/|s|} \left(\left| \frac{\partial^{2|s|}}{\partial x^{2|s|}} \mathcal{K}(0) \right| \right)^{1/2|s|}.$$
(37)

Now, substituting (37) into (36) we easily get

$$w_{k} \geq \beta_{k}^{-1} \left(\frac{2}{\varepsilon} \prod_{j=2}^{q} \beta_{j}^{s_{j}} \right)^{1/|s|} \left(\left| \frac{\partial^{2|s|}}{\partial x^{2|s|}} \mathcal{K}(0) \right| \right)^{1/2|s|}.$$
(38)

where $k = \overline{1, q}, \beta_1 \equiv 1.$

Few applications of the derived formulae to the wellknown Shannon-Kotelnikov formula will be discussed in the next chapter.

6 Some statistical interpretations

It is more convenient, because of the applicational goals, to describe a random signal $\{S(t)|t \in \mathbb{R}^m\}$ in a discrete form, for example, with a sequence $\{s_j\}_{j \in \mathbb{Z}}$ of random variables, or by a vector. Of course, this description must be 'sufficiently good in some manner' to recognize the initial signal S(t) with the given accuracy, (Gulyás, 1986).

One of the discretization techniques applied to the continuous time signals is the Shannon-Kotelnikov formula or more precisely its variant for weakly stationary BL random signals, proved by Balakrishnan (1957) for WSSP; and by Parzen (1956), for BL HRF. We must outline here that Parzen did not give an exact proof, his approach is mainly heuristic. (To avoid some continuity difficulties, closely connected to the sampling sum convergence questions consult Pogány (1991 a). For NBL random signals the paper Splettstößer (1981) give many informations.

The sampling discretization of a random signal S(t) means a sum representation

$$S(t) = \sum_{n \in \mathbb{Z}} \mathbf{s}_n \varphi_n(t), \quad t \in \mathbb{R}^m.$$
(39)

The system of functions $\{\varphi_n(t)|t \in \mathbb{R}^m\}$ one choose according to a given value set $\{s_n = S(t_n)|t_n \in T \subseteq \mathbb{R}^m\}$, where T is known. Otherwise, the coefficients s_n could be computed directly. Therefore the observed signal value set $\{s_n\}$ is the object of statistical investigations and interpretations, since it is a sequence of random variables.

One of the possible formulations of the Shannon-Kotelnikov sampling expansion to w > 0 of a band - limited weakly stationary scalar, continuous time random process $\{X(t)|r \in \mathbb{R}\}$ gives the following result:

$$\lim_{n \to \infty} \left(\sum_{n = -\infty}^{+\infty} X(\frac{n\pi}{w}) \operatorname{sinc}(wt - n\pi) - X(t) \right) = (F^+ + F^-) \sin^2(wt), \quad (40)$$

where F^{\pm} are the masses of the spectral distribution function $F(\lambda)$ at the end points $\pm w$ of the sampling support interval [-w,w], Pog any (1991 b). If the values F^{\pm} are known by $\{X(\frac{n\pi}{w})\}_{n=-\infty}^{\infty}$ one easily recognizes the nominal BL WSSP X(t).

The coefficients $X(\frac{n\pi}{w})$ are random variables (i.e. the split-values of the process $X(t;\omega)$ at $t = \frac{n\pi}{w}$). Before X(t) will be measured at $\frac{n\pi}{w}$, $n \in \{-N, \ldots, -1, 0, 1, \ldots, M\}$, $N, M \in \mathbb{N}$ are given, we must order the bandwidth w, if X(t) is a BL WSSP. The truncation error such as it originates from the duration - limited measurements has very extended literature listed for example in Butzer, Splettstößer (1977), Jerri (1977), so the discussion omits this side of the problem. By the reasons explained in detailed above, for the given mean-square approximation error level $e_i 0$, at first we order w using the relation (33).

In the case of the WSVPs, (34) would be applied.

Finally, for NBL HRFs, (35) is convenient together with some additional restrictions upon the connections between the coordinate - bandwidths in W.

For the observation error in measuring the signal values $S(\frac{n\pi}{w})$ consult Butzer, Splettstößer (1977), Jerri (1977).

References

- Balakrishnan, A.V. (1957): A note on the sampling principle for continuous signals. IRE Trans. Inform. Theory, IT-3, 143-146.
- [2] Belyaev, Yu.K. (1959): Analitičestkie slučajnie processy. Teorija Verojat.Primenen., 4, 437-444.
- [3] Beutler, F.J. (1961): Sampling theorems and bases in a Hilbert space. Inform. Control, 4, 97-117.
- [4] Brown Jr., J.L. (1978): On mean-square aliasing error in the cardinal series expansion of random processes. IEEE Trans. Inform. Theory, IT-24, 254-256.
- [5] Butzer, P.L. and Splettstößer, W. (1977): A sampling theorem for durationlimited functions with error estimates. Inform. Control, 29, 55-65.

- [6] Butzer, P.L., Splettstößer, W. and Stens, R.L. (1988): The sampling theorem and linear prediction in signal analysis. Jahresber. Deutsch. Math.- Verein., 90, 1-70.
- [7] Gulyás, O.(1986): Folytonos paraméterű folyamatok diszkretizálási módszerei. In: G. Tusnádi and M. Ziermann (Eds.): *Idösorok analizise*. Budapest: Műszaki Könyvkiadó.
- [8] Jerri, A.J. (1977): The Shannon sampling theorem its various extensions and applications: A tutorial review. Proc. IEEE, 65, 1565-1596.
- [9] Parzen, E. (1956): A simple proof and some extensions of sampling theorem. Technical Reports, 7, Stanford: Stanford University.
- [10] Pogány, T. (1988): On the discretization methods of continuous time stochastic processes. Statistička Revija, 18, 42-56 (in Serbian).
- [11] Pogány, T. (1989): An approach to the sampling theorem for continuous time processes. Austral.J.Statist., 31, 427-432.
- [12] Pogány, T. (1991 a): Almost sure sampling reconstruction of non-band-limited homogeneous random fields. Publ.Inst. Math. 49, 221-232.
- [13] Pogány, T. (1991 b). On a very tight truncation error bound for stationary stochastic processes. IEEE Trans. Signal Process, SP-39, 1918-1919.
- [14] Pogány, T. (1993): On the Brown aliasing error upper bound for homogeneous random fields. Signal Processing, 34 (to appear).
- [15] Pogány, T. and Perunicic, P. (1992): On the spectral representation of the sampling cardinal series expansion of weakly stationary stochastic processes. *Stochastica*, 13, 89-100.
- [16] Slepian, D. (1976): On bandwidth. Proc.IEEE, 64, 292-300.
- [17] Splettstößer, W. (1981): On the approximation of random processes by convolution processes. Z.Angew.Math.Mech., 61, 235-241.
- [18] Wong, E. (1971): Stochastic Processes in Information and Dynamical Systems. New York: McGraw-Hill.