International Journal of Pure and Applied Mathematics

Volume 20 No. 1 2005, 109-140

SKEW-NORMAL PROCESSES AS MODELS FOR RANDOM SIGNALS CORRUPTED BY GAUSSIAN NOISE

A.F. Gualtierotti

Institut de Hautes Études en Administration Publique (IDHEAP)
21 Route de la Maladière
CH-1022 Chavannes-près-Renens, SWITZERLAND
e-mail: antonio.gualtierotti@idheap.unil.ch

Abstract: A family of skew-normal processes is defined and their properties listed, especially those that relate to the detection of random signals corrupted by additive Gaussian noise.

AMS Subject Classification: 60G99, 60G30, 60G25, 60H30, 94A13
Key Words: skew-normal process, equivalence, likelihood, detection, linear space, prediction, representation, Cramér-Hida

1. Introduction

Detection of a signal S corrupted by a noise N modulating a received waveform X, that is testing for the presence of the signal and rejecting H_0, the "noise only" assumption, is one of the basic problems of statistical communication theory (Balakrishnan [13]). The detection model used (choice of X and N) is valid only in case the law of the received waveform, P_X, is absolutely continuous with respect to that of the noise, P_N, otherwise one would achieve error free detection, a feature indicative of poor modeling (Slepian [71]). Furthermore, when absolute continuity obtains, one often cannot do better than implement the likelihood ratio as a detector (Helstrom [36]). So the two basic mathematical problems one must solve for a detection problem, given X and N, are those of assessing absolute continuity of the law of X with respect to that of N and of computing the resulting Radon-Nikodým derivative. One wants next to

Received: March 20, 2005 © 2005, Academic Publications Ltd.

estimate the probabilities of false alarm (deciding that $X \neq N$ when $X = N$) and false rejection (deciding $X = N$ when $X \neq N$), usually from finite dimensional approximations.

These problems have a long history (Gikhman et al [30], Kailath et al [46]) and have been partially solved in the following cases which are not necessarily mutually exclusive:

1. P_X and P_N belong to the same family of laws, notably the Gaussian (Feldman [28], Hajek [34], Jørsboe [43], Parzen [61, 62], Rao et al [64]), but there are also many results for laws of processes with independent increments, infinitely divisible marginals and Markov processes (Briggs [15], Brockett et al [16, 17, 18], Hudson et al [40], Mémin et al [57], Newman [58, 59], Sato [66, 67], Skorokhod [69, 70]).

The methods used for working with this category of models involve typically a limiting argument applied to the likelihoods of marginal distributions.

2. P_N is the law of a martingale, typically the Wiener process, and $X = S + N$, with S in the reproducing kernel Hilbert space (RKHS: Aronszajn [3], Povzner [63]) of N (Kailath [44, 45], Liptser et al [52], Mémin [56]).

The methods used for working with this category of models involve typically martingale techniques and in particular Girsanov's Theorem (Girsanov [31]).

3. N is a Gaussian process continuous in mean square, or a mixture of such, or even a linear process with Wiener and Poisson components, and $X = S + N$, with S in the RKHS of N (Baker et al [9, 10, 11, 12], Climescu et al [20]).

The methods used for working with this category of models are based on martingale techniques and the Cramér-Hida representation (Cramér [22], Hida [37]).

4. There are also quite general theorems (Jacod et al [41, 42]) that could be used were one able to compute explicitly the conditional expectation of the Radon-Nykodým derivative (see, as an example, Proposition 3 below) or the local characteristics of related martingales. But these techniques, though they contribute a new, structured and universal understanding of the problem, seem to yield little in terms of specifics as they rely mostly on "existence" results.

5. Finally, there are results for $X = S + N$ with S and N independent and these are usually obtained by reducing the problem to one of detection of a sure (nonrandom) signal (Baker [8], Duncan [24]).

All the solutions that have just been enumerated have practical drawbacks. In many cases X is not Gaussian (Baker et al [9, 12], Dwyer [26, 27], Huang et al [39], Schwartz et al [68], Wilson et al [73]), nor does it have independent increments (Baker et al [9, 12]). Except in rather controlled situations signal and noise are not independent (Horton [38]). In physical problems, N is seldom a martingale (Baker et al [9, 12]) and the Cramér-Hida representation is an existence result that can seldom be computed explicitly. Also, in Case 2 and Case 3, the law of the signal being unknown, the probability of false rejection cannot be computed.

In the present paper another track is chosen in search of new solutions to the same problems: it is attempted to construct the law of X directly from that of N with the idea that the law of X, equivalence and likelihood should emerge simultaneously from the calculation. This is done in the restricted context of skew-normal processes (to be defined below in Proposition 1) whose marginals have a signal-plus-noise representation. This seems to be an unavoidable requirement as there are results that say that there is such a representation whenever absolute continuity prevails, that is $P_X \ll P_N$ (Baker et al [10], Fernique [29], Mel'ničenko [55]).

Skew-normal distributions have been the subject of much interest in statistics in past years (Arellano-Valle et al [2], Azzalini et al [4, 5, 6, 7], Capitanio et al [19]) and, though they were devised to extend the class of analytically tractable models beyond the Gaussian, they have been shown to be richer in features than one could have expected given the simplicity of the extension.

The basic assumption will be that the signal can be expressed as a random series in the RKHS of the noise (which is in line with standard signal analysis practice – Allen et al [1] – and the non singular detection requirement – Slepian [71]) and that it interacts explicitly with the noise at the level of the marginal distributions of X. One is thus led to consider a continuous covariance C with associated RKHS $H(C)$ and to choose arbitrarily in $H(C)$ a fixed family of functions, $\{h_n,\ n \in I\!N\}$. These "basic signals" will be "modulated" by jointly independent random variables, say $\left\{X_n^{(0)},\ n \in I\!N\right\}$, which have mean zero and (absolutely continuous) symmetric density. One will assume that for each $\omega \in \Omega$,

$$S(\omega, \cdot) = \sum_n X_n^{(0)} h_n \quad \text{and} \quad |S|(\omega, \cdot) = \sum_n \left| X_n^{(0)} \right| h_n$$

are elements of $H(C)$. One will also assume that convergence takes place in mean of order two. One needs of course to know when such assumptions are legitimate and the appendix contains some remarks about these convergence issues.

2. Skew-Normal Distributions

The skew-normal distributions used in the sequel are built as follows. Suppose $\underline{X}_0 \in I\!R^m$ is a vector of random variables which are independent, symmetric with mean zero, and $\underline{X} \in I\!R^n$ is a vector of independent standard normal random variables. The components of $\underline{X}_0$ being $X_1^{(0)}, \ldots, X_m^{(0)}$, $|\underline{X}_0|$ will have components $\left|X_1^{(0)}\right|, \ldots, \left|X_m^{(0)}\right|$. One has the following lemma, where $A_{n,m}$ and $B_{n,n}$ are matrices:

Lemma 1. *Suppose $\underline{X}_0$ has a density $f_{\underline{X}_0}$ and the required assumptions of Arellano-Valle et al [2] are satisfied. Then, provided $B_{n,n}B^\star_{n,n}$ is invertible,*

$$\underline{U} = A_{n,m}\,|\underline{X}_0| + B_{n,n}\underline{X}$$

has density

$$f_{\underline{U}}(\underline{u}) = 2^m f_{\mathcal{N}\left(\underline{0},B_{n,n}B^\star_{n,n}\right)}(\underline{u}) \int_{I\!R^m_+} e^{Q(\underline{u},\underline{v})} f_{\underline{X}_0}(\underline{v})\, d\underline{v},$$

with

$$Q(\underline{u},\underline{v}) = \langle \left(B_{n,n}B^\star_{n,n}\right)^{-1}\underline{u}, A_{n,m}\underline{v}\rangle_{I\!R^n} - \frac{1}{2}\langle \left(B_{n,n}B^\star_{n,n}\right)^{-1}A_{n,m}\underline{v}, A_{n,m}\underline{v}\rangle_{I\!R^n}.$$

Proof. Let $\underline{V} = A_{n,m}\underline{X}_0 + B_{n,n}\underline{X}$. According to Arellano-Valle et al [2], the law of $\underline{U}$ is that of $\underline{V}$ conditional on $\underline{X}_0 > \underline{0}$ which has density

$$2^m f_{\underline{V}}(\underline{v})\, P\left(\underline{X}_0 > \underline{0} \mid \underline{V} = \underline{v}\right).$$

Now, for Borel sets $E \in I\!R^m$ and $F \in I\!R^n$,

$$P\left(\underline{X}_0 \in E, \underline{V} \in F\right) = \int_E d\underline{u}\; f_{\underline{X}_0}(\underline{u})\, P\left(\underline{V} \in F \mid \underline{X}_0 = \underline{u}\right).$$

But the conditional law of $\underline{V}$ given that $\underline{X}_0 = \underline{u}$ is Gaussian with mean $A_{n,m}\underline{u}$ and covariance $B_{n,n}B^\star_{n,n}$. Consequently one has that

$$f_{\underline{X}_0,\underline{V}}(\underline{u},\underline{v}) = f_{\underline{X}_0}(\underline{u})\, f_{\mathcal{N}\left(A_{n,m}\underline{u},B_{n,n}B^\star_{n,n}\right)}(\underline{v}).$$

Now $f_{\mathcal{N}\left(A_{n,m}\underline{u},B_{n,n}B^\star_{n,n}\right)}(\underline{v})$ can conveniently be written as

$$e^{Q(\underline{u},\underline{v})} f_{\mathcal{N}\left(\underline{0},B_{n,n}B^\star_{n,n}\right)}(\underline{v}),$$

with $Q(\underline{u}, \underline{v})$ as in the statement of the lemma. Thus

$$f_{\underline{X}_0, \underline{V}}(\underline{u}, \underline{v}) = e^{Q(\underline{u}, \underline{v})} f_{\underline{X}_0}(\underline{u}) f_{\mathcal{N}\left(\underline{0}, B_{n,n} B^{\star}_{n,n}\right)}(\underline{v}) .$$

Consequently

$$f_{\underline{V}}(\underline{v}) = f_{\mathcal{N}\left(\underline{0}, B_{n,n} B^{\star}_{n,n}\right)}(\underline{v}) \int_{\mathbb{R}^m} d\underline{u}\ f_{\underline{X}_0}(\underline{u}) e^{Q(\underline{u}, \underline{v})}$$

and

$$f_{\underline{X}_0 | \underline{V} = \underline{v}}(\underline{u}) = \frac{e^{Q(\underline{u}, \underline{v})} f_{\underline{X}_0}(\underline{u})}{\int_{\mathbb{R}^m} d\underline{u}\ f_{\underline{X}_0}(\underline{u}) e^{Q(\underline{u}, \underline{v})}}.$$

The formula of the lemma follows readily. □

3. Skew-Normal Processes: A Construction

3.1. Marginal Distributions

Let

$$S_m(\omega, \cdot) = \sum_{j=1}^{m} X_j^{(0)} h_j \quad \text{and} \quad |S_m|(\omega, \cdot) = \sum_{j=1}^{m} \left| X_j^{(0)} \right| h_j .$$

Fix also $0 \leq t_1 < t_2 < t_3 < \cdots < t_n \leq T < \infty$. Then the vector $\underline{S}^{(m,n)}$ with components

$$S_m(\omega, t_i), \ 1 \leq i \leq n$$

can be given the representation

$$\underline{S}^{(m,n)} = A_{n,m} \underline{X}_0$$

with $A_{n,m}$ having entries

$$a_{i,j} = h_j(t_i), \quad 1 \leq i \leq n, \quad 1 \leq j \leq m ,$$

and $\underline{X}_0$ having entries $X_j^{(0)}$, $1 \leq j \leq m$. One defines $\left|\underline{S}^{(m,n)}\right|$ similarly. If one now chooses for $B_{n,n}$ the square root of Σ_n, the latter having entries

$$\sigma_{i,j} = C(t_i, t_j), \quad 1 \leq i \leq n, \quad 1 \leq j \leq n,$$

the vectors $\underline{U}$ and $\underline{V}$ of Lemma 1 have the respective "signal–plus–independent noise" representations

$$\underline{U}^{(m,n)} = \left|\underline{S}^{(m,n)}\right| + \underline{N}_n \quad \text{and} \quad \underline{V}^{(m,n)} = \underline{S}^{(m,n)} + \underline{N}_n ,$$

with $\underline{N}_n = B_{n,n}\underline{X}$, $\underline{X}$ having components $X_i,\ 1 \le i \le n$. From Lemma 1 $\underline{U}^{(m,n)}$ has density

$$f_{\underline{U}^{(m,n)}}(\underline{u}) = 2^m f_{\underline{N}_n}(\underline{u}) \int_{\mathbb{R}_+^m} e^{Q(\underline{u},\underline{v})} f_{\underline{X}_0}(\underline{v})\, d\underline{v}\,,$$

with

$$Q(\underline{u},\underline{v}) = \langle \Sigma_n^{-1}\underline{u}, A_{n,m}\underline{v}\rangle_{\mathbb{R}^n} - \frac{1}{2}\langle \Sigma_n^{-1}A_{n,m}\underline{v}, A_{n,m}\underline{v}\rangle_{\mathbb{R}^n}.$$

Define $\pi_n : \mathbb{R}^\infty \longrightarrow \mathbb{R}^n$ by $\pi_n(\underline{x}) = \underline{x}_n$, the components of the latter being the first n components of $\underline{x}$. $\mathcal{E}(f,t)) = f(t)$ defines the evaluation map and $\mathcal{E}(f,\underline{t}_n)$ is the vector with components $\mathcal{E}(f,t_i),\ 1 \le i \le n$. The map $J_n : \mathbb{R}^n \longrightarrow H(C)$ is defined by

$$J_n(\underline{x}_n) = \sum_{i=1}^{n} x_i h_i$$

and the map $J : \mathbb{R}^\infty \longrightarrow H(C)$ by

$$J(\underline{x}) = \sum_{i=1}^{\infty} x_i h_i,$$

whenever the latter makes sense. In particular, $J(\underline{X}_0)$ and $J(|\underline{X}_0|)$ make sense, $\underline{X}_0$ having components $X_n^{(0)},\ n \in \mathbb{N}$. Define then ν_i as the measure with density

$$2I_{]0,\infty[} f_{X_i^{(0)}}$$

(I_A being the indicator function of the set A), ν^n as the product of $\nu_1, \ldots, \nu_n$ and ν^∞ as the product of the ν_i's, $i \in \mathbb{N}$. Finally denote $P_{H_n(C)}$ the projection in $H(C)$ whose range is the subspace generated by $\{C(\cdot,t_i),\ 1 \le i \le n\}$. Q may then be interpreted as follows. Firstly one has that $A_{n,m}\underline{v}$ has components

$$\langle J_m(\underline{v}), C(\cdot,t_i)\rangle_{H(C)},\ 1 \le i \le n.$$

Thus

$$\langle \Sigma_n^{-1}A_{n,m}\underline{v}, A_{n,m}\underline{v}\rangle_{\mathbb{R}^n} = \left\| P_{H_n(C)} J_m(\underline{v})\right\|_{H(C)}^2.$$

Secondly,

$$\langle \Sigma_n^{-1}\underline{u}, A_{n,m}\underline{v}\rangle_{\mathbb{R}^n} = \langle \underline{u}, \Sigma_n^{-1}\mathcal{E}(J_m(\underline{v}), \underline{t}_n)\rangle_{\mathbb{R}^n},$$

and thirdly

$$2^m \int_{\mathbb{R}_+^m} e^{Q(\underline{u},\underline{v})} f_{\underline{X}_0}(\underline{v})\, d\underline{v} = \int_{\mathbb{R}_+^\infty} \nu^\infty(d\underline{v})\, g_m(\underline{v} \mid \underline{u})\ ,$$

with

$$g_m\left(\underline{v}\mid\underline{u}\right)=I_{\mathbb{R}_+^m}\left(\pi_m\left[\underline{v}\right]\right)e^{\langle\underline{u},\Sigma_n^{-1}\mathcal{E}(J_m(\pi_m[\underline{v}]),\underline{t}_n)\rangle_{\mathbb{R}^n}-\frac{1}{2}\left\|P_{H_n(C)}J_m(\pi_m[\underline{v}])\right\|_{H(C)}^2}.$$

One then has the following lemma.

Lemma 2. *Let the vector* $\underline{S}^{(n)}$ *have components* $S\left(\omega,t_i\right),\ 1\le i\le n$. *The vector* $\left|\underline{S}^{(n)}\right|$ *is defined similarly. Let*

$$\underline{U}^{(n)}=\left|\underline{S}^{(n)}\right|+\underline{N}_n\,,$$

with $\underline{N}_n=\Sigma_n^{\frac{1}{2}}\underline{X}$, $\underline{X}$ *having components* $X_i,\ 1\le i\le n$. $\underline{U}^{(n)}$ *has then a density of the form*

$$f_{\underline{U}^{(n)}}\left(\underline{u}\right)=f_{\mathcal{N}(\underline{0},\Sigma_n)}\left(\underline{u}\right)\int_{\mathbb{R}_+^\infty}\nu^\infty\left(d\underline{v}\right)g\left(\underline{v}\mid\underline{u}\right)\ ,$$

with

$$g\left(\underline{v}\mid\underline{u}\right)=e^{\langle\underline{u},\Sigma_n^{-1}\mathcal{E}(J[\underline{v}],\underline{t}_n)\rangle_{\mathbb{R}^n}-\frac{1}{2}\left\|P_{H_n(C)}J[\underline{v}]\right\|_{H(C)}^2}.$$

Proof. For any vectors $\underline{a}$, $\underline{x}$, and any symmetric and strictly positive definite matrix A, all in $\mathbb{R}^n$, one has that

$$\begin{aligned}\langle\underline{a},\underline{x}\rangle_{\mathbb{R}^n}&-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle_{\mathbb{R}^n}\\&=\frac{1}{2}\langle A^{-1}\underline{a},\underline{a}\rangle_{\mathbb{R}^n}-\frac{1}{2}\langle A\left(\underline{x}-A^{-1}\underline{a}\right),\left(\underline{x}-A^{-1}\underline{a}\right)\rangle_{\mathbb{R}^n},\end{aligned}$$

so that

$$\langle\underline{a},\underline{x}\rangle_{\mathbb{R}^n}-\frac{1}{2}\langle A\underline{x},\underline{x}\rangle_{\mathbb{R}^n}\le\frac{1}{2}\langle A^{-1}\underline{a},\underline{a}\rangle_{\mathbb{R}^n}.$$

Letting

$$A=A_{n,m}^\star\Sigma_n^{-1}A_{n,m}\quad\text{and}\quad\underline{a}=A_{n,m}^\star\Sigma_n^{-1}\underline{u},$$

one has firstly that (Harville [35])

$$A^{-1}=A_{n,m}^-\Sigma_n\left(A_{n,m}^-\right)^\star,$$

where $A_{n,m}^-$ is a generalized inverse of $A_{n,m}$. Secondly $\langle A^{-1}\underline{a},\underline{a}\rangle_{\mathbb{R}^n}$ then becomes

$$\langle A_{n,m}A_{n,m}^-\Sigma_n\left(A_{n,m}A_{n,m}^-\right)^\star\Sigma_n^{-1}\underline{u},\Sigma_n^{-1}\underline{u}\rangle_{\mathbb{R}^n}.$$

But, and thirdly, $A_{n,m}A_{n,m}^{-} = P_A$, a projection (Harville [35]). Consequently

$$\langle A^{-1}\underline{a},\underline{a}\rangle_{\mathbb{R}^n} = \left\|\Sigma_n^{\frac{1}{2}} P_A \Sigma_n^{-1}\underline{u}\right\|_{\mathbb{R}^n}^2 \leq \left\|\Sigma_n^{\frac{1}{2}}\right\| \left\|\Sigma_n^{-1}\underline{u}\right\|_{\mathbb{R}^n}^2 .$$

Thus, finally,

$$g_m\left(\underline{v} \mid \underline{u}\right) \leq e^{\frac{1}{2}\left\|\Sigma_n^{\frac{1}{2}}\right\| \left\|\Sigma_n^{-1}\underline{u}\right\|_{\mathbb{R}^n}^2},$$

which is independent of $\underline{v}$ and the dimension m. The definition of S as a convergent series and dominated convergence then yield that

$$\lim_m \int_{\mathbb{R}_+^\infty} \nu^\infty\left(d\underline{v}\right) g_m\left(\underline{v} \mid \underline{u}\right) = \int_{\mathbb{R}_+^\infty} \nu^\infty\left(d\underline{v}\right) g\left(\underline{v} \mid \underline{u}\right).$$

An application of Fubini's Theorem shows furthermore that

$$f_{\mathcal{N}(\underline{0},\Sigma_n)}\left(\underline{u}\right) \int_{\mathbb{R}_+^\infty} \nu^\infty\left(d\underline{v}\right) g\left(\underline{v} \mid \underline{u}\right)$$

is a density. Consequently, by Scheffé's Theorem (Billingsley [14]) for every Borel set B,

$$\lim_m \int_B d\underline{u}\ f_{\mathcal{N}(\underline{0},\Sigma_n)}\left(\underline{u}\right) \int_{\mathbb{R}_+^\infty} \nu^\infty\left(d\underline{v}\right) g_m\left(\underline{v} \mid \underline{u}\right)$$
$$= \int_B d\underline{u}\ f_{\mathcal{N}(\underline{0},\Sigma_n)}\left(\underline{u}\right) \int_{\mathbb{R}_+^\infty} \nu^\infty\left(d\underline{v}\right) g\left(\underline{v} \mid \underline{u}\right).$$

On the other hand, $\underline{U}^{(m,n)}$ converges almost surely to $\underline{U}^{(n)}$, so that, for every continuity set B (Billingsley [14]),

$$\lim_m \int_B f_{\underline{U}^{(m,n)}} = \int_B f_{\underline{U}^{(n)}}.$$

So the lemma is proved. □

Remark. A common regression model (Gu [33], Wahba [72]) has the form

$$Y_i = f\left(x_i\right) + \epsilon_i,\ 1 \leq i \leq n,\ \epsilon_i \sim \mathcal{N}\left(0,\sigma^2\right),\ \text{IID},\ f \text{ in some RKHS.}$$

The laws of Lemma 2 provide a model for the case when f is random and the ϵ_i's have "any" Gaussian distribution. Such models are of interest in the Bayesian approach to support vector machines (Law et al [50]).

3.2. Moments and Characteristic Function of Marginal Distributions

They all follow directly from the explicit "signal–plus–noise" representation of the "marginal random variables" $\underline{U}^{(n)}$ of Lemma 2.

Let indeed

$$\mu_n = E\left[|X_n|\right] \quad \text{and} \quad \sigma_n^2 = V\left[|X_n|\right],$$

$$\mu = \sum_{n=1}^{\infty} \mu_n h_n \quad \text{and} \quad T = \sum_{n=1}^{\infty} \sigma_n^2 h_n \otimes h_n, \quad h_n \otimes h_n\left(x\right) = \langle x, h_n \rangle_{H(C)} h_n.$$

T is the covariance operator (Parthasarathy [60]) of the law of $|S|$ and has thus a finite trace which is smaller than or equal to $\sum_{j=1}^{\infty} \sigma_j^2 \|h_j\|_{H(C)}^2$ (Conway [21]). Also

$$\mu\left(u\right) = E\left[|S|\left(\cdot, u\right)\right] = \langle \mu, C\left(\cdot, u\right)\rangle_{H(C)} = \sum_{n=1}^{\infty} \mu_n h_n\left(u\right),$$

and

$$\Gamma\left(u, v\right) = \operatorname{cov}\left[|S|\left(\cdot, u\right), |S|\left(\cdot, v\right)\right]$$
$$= \langle TC\left(\cdot, u\right), C\left(\cdot, v\right)\rangle_{H(C)} = \sum_{n=1}^{\infty} \sigma_n^2 h_n\left(u\right) h_n\left(v\right).$$

Then

$$E\left[\underline{U}^{(n)}\right] = \mathcal{E}\left(\mu, \underline{t}_n\right)$$

and $V\left[\underline{U}^{(n)}\right]$ has components of the form

$$\Gamma\left(t_i, t_j\right) + C\left(t_i, t_j\right) = \langle \left[I_{H(C)} + T\right] C\left(\cdot, t_i\right), C\left(\cdot, t_j\right)\rangle_{H(C)}.$$

Finally, the characteristic function is

$$\begin{aligned}
\varphi_{\underline{U}^{(n)}}\left(\underline{\theta}\right) &= e^{-\frac{1}{2}\langle \Sigma_n \underline{\theta}, \underline{\theta}\rangle_{I\!R^n}} \prod_{i=1}^{\infty} \varphi_{\nu_i}\left(\langle \underline{\theta}, \mathcal{E}\left(h_i, \underline{t}_n\right)\rangle_{I\!R^n}\right) \\
&= e^{-\frac{1}{2}\langle \Sigma_n \underline{\theta}, \underline{\theta}\rangle_{I\!R^n}} \prod_{i=1}^{\infty} \varphi_{\nu_i}\left(\langle h_i, \sum_{j=i}^{n} \theta_j C\left(\cdot, t_j\right)\rangle_{H(C)}\right).
\end{aligned}$$

Expliciter forms occur of course whenever a specification for ν_i is given. Thus, for ν_i's from IID standard normals, one has that

$$\varphi_{\underline{U}^{(n)}}(\underline{\theta}) = e^{-\frac{1}{2}Q_n(\underline{\theta})} \prod_{j=1}^{\infty} \left\{ 1 + i\sqrt{\frac{2}{\pi}} \int_0^{L_{n,j}(\underline{\theta})} e^{\frac{x^2}{2}} dx \right\},$$

where

$$L_{n,j}(\underline{\theta}) = \langle h_j, \sum_{i=1}^{n} \theta_i C(\cdot\ , t_i) \rangle_{H(C)},$$

$$Q_n(\underline{\theta}) = \langle \Sigma_n \underline{\theta}, \underline{\theta} \rangle_{\mathbb{R}^n} + \langle T \left[\sum_{i=1}^{n} \theta_i C(\cdot\ , t_i) \right], \left[\sum_{i=1}^{n} \theta_i C(\cdot\ , t_i) \right] \rangle_{H(C)}.$$

3.3. Consistency of Marginal Distributions

Let $\Pi_{i_1,\ldots,i_p}$, $p < n$, be the matrix whose rows are $\underline{e}^{\star}_{i_1}, \ \ldots\ , \underline{e}^{\star}_{i_p}$, an arbitrary choice of standard basis vectors of $\mathbb{R}^n$. Then, *mutatis mutandis,*

$$\underline{U}^{(p)} = \Pi_{i_1,\ldots,i_p} \underline{U}^{(n)}$$

and

$$\begin{aligned} \varphi_{\underline{U}^{(p)}}(\underline{\theta}_p) &= \varphi_{\underline{U}^{(n)}}\left(\Pi^{\star}_{i_1,\ldots,i_p} \underline{\theta}_p\right) \\ &= e^{-\frac{1}{2}\langle \Pi_{i_1,\ldots,i_p} \Sigma_n \Pi^{\star}_{i_1,\ldots,i_p} \underline{\theta}_p, \underline{\theta}_p \rangle_{\mathbb{R}^p}} \Pi_{i=1}^{\infty} \varphi_{\nu_i}\left(\langle \Pi^{\star}_{i_1,\ldots,i_p} \underline{\theta}_p, \mathcal{E}(h_i, \underline{t}_n) \rangle_{\mathbb{R}^n} \right). \end{aligned}$$

But $\Pi_{i_1,\ldots,i_p} \Sigma_n \Pi^{\star}_{i_1,\ldots,i_p} = \Sigma_p$ with entries $C(t_{i_k}, t_{i_l})$, $\ 1 \leq k, l \leq p$ and

$$\langle \Pi^{\star}_{i_1,\ldots,i_p} \underline{\theta}_p, \mathcal{E}(h_i, \underline{t}_n) \rangle_{\mathbb{R}^n} = \langle \underline{\theta}_p, \mathcal{E}\left(h_i, \Pi_{i_1,\ldots,i_p} \underline{t}_n\right) \rangle_{\mathbb{R}^n}.$$

One may thus state the following result.

Proposition 1. *The densities of Lemma 2 define the marginal distributions of a stochastic process.*

The process so constructed will be called a skew-normal process and be denoted X, and any zero-mean Gaussain process N with covariance C will be called the normal process underlying X. From Section 3.2 one has the following corollary.

Corollary 1. *The first two moments of a skew-normal process X are given by the following expressions: for the expectation,*

$$\mu_X(t) = \langle \sum_{j=1}^{\infty} \mu_j h_j, C(\cdot\ , t) \rangle_{H(C)},$$

and, for the covariance function,

$$C_X(s,t) = \langle \left[I_{H(C)} + T \right] C(\cdot\,,s)\,, C(\cdot\,,t) \rangle_{H(C)}.$$

4. Skew-Normal Processes: Properties

The main fact about skew-normal processes is that their law is equivalent to that of their underlying normal and that the ensuing Radon-Nikodým derivative has a nice and simple form. Thus many if not most of the qualitative properties such as path smoothness of a skew-normal process are those of its underlying normal.

4.1. Skew-Normal Processes are Equivalent to their Underlying Normals

To be specific, let $\mathcal{B}\left[\mathbb{R}^{[0,T]}\right]$ denote the Borel sets of $\mathbb{R}^{[0,T]}$, P_N, the law induced on the Borel sets by N, and $\mathcal{L}_2[N]$, the subspace of

$$L_2\left[\left(\mathbb{R}^{[0,T]},\ \mathcal{B}\left[\mathbb{R}^{[0,T]}\right],\ P_N\right)\right]$$

generated by the finite linear combinations of the form

$$\sum_{k=1}^{n} \alpha_k \mathcal{E}(\cdot\,,t_k)\,.$$

Let also the map

$$U_N : H(C) \to \mathcal{L}_2[N] \quad \text{be defined by} \quad U_N[C(\cdot\,,t)] = \mathcal{E}(\cdot\,,t)\,.$$

U_N is an isometry. Set $N_j = U_N[h_j]$. Then the following theorem holds true.

Theorem 1. *P_X and P_N are mutually absolutely continuous and one has that*

$$\frac{dP_X}{dP_N} = \int_{\mathbb{R}_+^\infty} e^{U_N[J(\underline{v})] - \frac{1}{2}\|J(\underline{v})\|^2_{H(C)}}\ \nu_\infty(d\underline{v})\,,$$

where

$$U_N[J(\underline{v})] = \sum_{j=1}^{\infty} v_j N_j$$

and, with respect to P_N, *the set* $\{N_j,\ j \in \mathbb{N}\}$ *is made of zero-mean normal random variables with covariances* $\langle h_j, h_k\rangle_{H(C)},\ (j,k) \in \mathbb{N} \times \mathbb{N}$.

The likelihood is thus a mixture of exponentials of Gaussians "modulated by the components of the harmonics of the signal."

Proof. In what follows, $\mathcal{L}_n[N]$ will denote the subspace of $\mathcal{L}_2[N]$ generated by the set of random variables

$$\mathcal{E}(\cdot,t_1),\ \ldots\ ,\mathcal{E}(\cdot,t_n).$$

The projection onto $\mathcal{L}_n[N]$ in $L_2[N]$ will be represented as $P_{\mathcal{L}_n[N]}$. The density of Lemma 2 can then be expressed as

$$\frac{dP_X^{(n)}}{dP_N^{(n)}}[f] = \int_{\mathbb{R}_+^\infty} e^{Q(\mathcal{E}(f,\underline{t}_n),\underline{v})}\nu_\infty(d\underline{v}),\quad f \in \mathbb{R}^{[0,T]},$$

where $P_X^{(n)} = P_X \circ \mathcal{E}(\cdot,\underline{t}_n)^{-1}$, $P_N^{(n)} = P_N \circ \mathcal{E}(\cdot,\underline{t}_n)^{-1}$ and

$$Q(\mathcal{E}(f,\underline{t}_n),\underline{v}) = \langle \mathcal{E}(f,\underline{t}_n), \Sigma_n^{-1}\mathcal{E}(J[\underline{v}],\underline{t}_n)\rangle_{\mathbb{R}^n} - \frac{1}{2}\left\|P_{H_n(C)}J[\underline{v}]\right\|_{H(C)}^2.$$

Consider now a sequence $\{\mathcal{T}_n \subseteq [0,T],\ n \in \mathbb{N}\}$ such that $n \leq p$ implies $\mathcal{T}_n \subseteq \mathcal{T}_p$ and $\bigcup_{n=1}^\infty \mathcal{T}_n$ is dense in $[0,T]$. $\underline{t}_n$ will henceforth be another symbol for $\mathcal{T}_n$. Then

$$\lim_n \left\|P_{H_n(C)}J(\underline{v})\right\|_{H(C)}^2 = \|J(\underline{v})\|_{H(C)}^2,$$

and

$$\langle \mathcal{E}(f,\underline{t}_n), \Sigma_n^{-1}\mathcal{E}((J(\underline{v})),\underline{t}_n)\rangle_{\mathbb{R}^n} = P_{\mathcal{L}_n[N]}U_N[J(\underline{v})].$$

Fix $\underline{v}$ and let

$$X_n(f,\underline{v}) = P_{\mathcal{L}_n[N]}U_N[J(\underline{v})] \quad \text{and} \quad X(f,\underline{v}) = U_N[J(\underline{v})].$$

One has there two Gaussian random variables with mean equal to zero and respective variances

$$\sigma_n^2(\underline{v}) = \left\|P_{H_n(C)}J(\underline{v})\right\|_{H(C)}^2 \quad \text{and} \quad \sigma^2(\underline{v}) = \|J(\underline{v})\|_{H(C)}^2.$$

$X_n(\cdot\ ,\underline{v})$ converges thus in probability to $X(\cdot\ ,\underline{v})$ and consequently so do their exponentials. If now $Z \sim \mathcal{N}\left(0,\sigma^2\right)$,

$$E\left[e^{Z-\frac{1}{2}\sigma^2}\right] = 1,$$

so that convergence is in $L_1[P_N]$. Furthermore, if $a \geq 0$ and $b \geq 0$, $|a-b| \leq a+b$. Thus

$$\lim_n E_{P_N}\left[\left|\int_{I\!R_+^\infty} e^{X_n(f,\underline{v})-\frac{1}{2}\sigma_n^2(\underline{v})}\nu_\infty(d\underline{v}) - \int_{I\!R_+^\infty} e^{X(f,\underline{v})-\frac{1}{2}\sigma^2(\underline{v})}\nu_\infty(d\underline{v})\right|\right]$$
$$\leq \lim_n \int_{I\!R_+^\infty} E_{P_N}\left[\left|e^{X_n(f,\underline{v})-\frac{1}{2}\sigma_n^2(\underline{v})} - e^{X(f,\underline{v})-\frac{1}{2}\sigma^2(\underline{v})}\right|\right]\nu_\infty(d\underline{v}) = 0. \quad \square$$

A form of the theorem that is easier to "read" is obtained by choosing for the h_j's "simple" functions of $H(C)$ and for the $X_n^{(0)}$, IID standard normals. A calculation thus yields the following theorem.

Corollary 2. *Let the $X_n^{(0)}$ be IID standard normal and $\{e_j,\ j \in \mathcal{J}\} \subseteq H(C)$ be an orthonormal set. Suppose $h_j = \sqrt{\lambda_j}e_j$, $\sum_{j\in\mathcal{J}}\sqrt{\lambda_j} < \infty$. Then*

$$\frac{dP_X}{dP_N} = \frac{e^{\frac{1}{2}\sum_{j\in\mathcal{J}}\frac{\lambda_j N_j^2}{1+\lambda_j}}}{\prod_{j\in\mathcal{J}}\sqrt{1+\lambda_j}}\prod_{j\in\mathcal{J}} 2\Phi\left(\sqrt{\frac{\lambda_j}{1+\lambda_j}}\,N_j\right),$$

where the family $\{N_j,\ j \in \mathcal{J}\}$ is made of IID standard normal random variables with respect to P_N.

4.2. False Rejection

Probabilistic statements with respect to P_X about the likelihood of Theorem 1 require only that the law with respect to P_X of the variables $N_j,\ j \in I\!N$, be known. Now

$$E_{P_X}\left[e^{\langle\underline{\theta},\underline{N}_p\rangle_{I\!R^p}}\right] = \int_{I\!R_+^\infty}\nu(d\underline{v})\,E_{P_N}\left[e^{\langle\underline{\theta},\underline{N}_p\rangle_{I\!R^p}+U_N[J(\underline{v})]-\frac{1}{2}\|J(\underline{v})\|^2_{H(C)}}\right].$$

Let

$$\tilde{v}_j = \begin{cases} \theta_j + v_j & \text{if} \quad 1 \leq j \leq p \\ v_j & \text{if} \quad j > p \end{cases}.$$

Then

$$E_{P_N}\left[e^{\langle\underline{\theta},\underline{N}_p\rangle_{I\!R^p}+U_N[J(\underline{v})]-\frac{1}{2}\|J(\underline{v})\|^2_{H(C)}}\right] = e^{\frac{1}{2}\|J(\underline{\tilde{v}})\|^2_{H(C)}-\frac{1}{2}\|J(\underline{v})\|^2_{H(C)}}.$$

But

$$\|J(\underline{\tilde{v}})\|^2_{H(C)} - \|J(\underline{v})\|^2_{H(C)} = \left\|\sum_{j=1}^p \theta_j h_j\right\|^2_{H(C)} - 2\langle\sum_{j=1}^p \theta_j h_j, \sum_{j=1}^\infty v_j h_j\rangle_{H(C)},$$

so that (see Corollary 4 below)

$$E_{P_X}\left[e^{\langle \underline{\theta}, \underline{N}_p\rangle_{I\!R^p}}\right] = e^{\frac{1}{2}\left\|\sum_{j=1}^p \theta_j h_j\right\|^2_{H(C)}} E\left[e^{\langle |S|, \sum_{j=1}^p \theta_j h_j\rangle_{H(C)}}\right].$$

One thus has the following result.

Proposition 2. *The random variables* $N_j,\ j \in I\!N$, *have, with respect to* P_X, *a skew-normal law.*

4.3. Representation of Skew-Normal Processes as Signal-Plus-Noise

It would be useful for modeling to know how skew-normal processes relate to their underlying normals in terms of random variables and processes rather than in law. But, since $P_X \ll P_N$, one has, as already stated and with adequate assumptions (Baker et al [10], Fernique [29], Melničenko [55]), that $X = S + \widetilde{N}$, $\widetilde{N}$ a process with the same law as N, that is, skew-normal processes are of the form signal–plus–noise. However it is usually hard to obtain S. When N is the Wiener process however, one can identify S and $\widetilde{N}$ as follows. Let ($[f]$ represents the equivalence class of f)

$$h_j(t) = \int_0^t \dot{h}_j(x)\, dx, \quad \left[\dot{h}_j\right] \in L_2[0,T],$$

and

$$f(\underline{v}, x) = \sum_{j=1}^{\infty} v_j \dot{h}_j(x).$$

Then ($\langle X\rangle$ represents the quadratic variation of the continuous martingale X)

$$U_N J(\underline{v}) = \int_0^T f(\underline{v}, x)\, \mathcal{E}(\cdot\ , dx)$$

and

$$\langle \int_0^{\cdot} f(\underline{v}, x)\, \mathcal{E}(\cdot\ , dx)\rangle(\cdot\ , T) = \int_0^T f^2(\underline{v}, x)\, dx = \|J(\underline{v})\|^2_{H(C)}.$$

Define then successively

$$\ln\{\Gamma(\underline{v}, c, t)\} = \int_0^t f(\underline{v}, x)\, \mathcal{E}(c, dx) - \frac{1}{2}\int_0^t f^2(\underline{v}, x)\, dx,$$

$$\xi(c,t) = \int_{I\!R_+^\infty} \Gamma(\underline{v}, c, t)\, f(\underline{v}, t)\, \nu_\infty(d\underline{v}),$$

$$\eta(c,t) = 1 + \int_0^t \xi(c,x)\,\mathcal{E}(c,dx),$$

$$\alpha(c,t) = \frac{\xi(c,t)}{\eta(c,t)}. \qquad \square$$

One then has the following result.

Proposition 3. *$X(\omega,t) = \int_0^t \alpha(X(\omega,\cdot),\ x)\,dx + W^X(\omega,t)$, where W^X is a Wiener process for the filtration $\underline{\sigma}(X)$.*

Proof. From Theorem 1, one has that

$$\frac{dP_X}{dP_N}[c] = \int_{I\!R_+^\infty} \Gamma(\underline{v},c,T)\,\nu_\infty(d\underline{v})$$

and thus that

$$E\left[\frac{dP_X}{dP_N} \mid \sigma_t(\mathcal{E})\right][c] = \int_{I\!R_+^\infty} \Gamma(\underline{v},c,t)\,\nu_\infty(d\underline{v}),$$

since $\Gamma(\underline{v},c,\cdot)$ is a martingale. But Γ is also such that

$$\Gamma(\underline{v},c,t) = 1 + \int_0^t \Gamma(\underline{v},c,x)\,f(\underline{v},x)\,\mathcal{E}(c,dx),$$

and, consequently,

$$E\left[\frac{dP_X}{dP_N} \mid \sigma_t(\mathcal{E})\right][c] = 1 + \int_{I\!R_+^\infty} \left\{\int_0^t \Gamma(\underline{v},c,x)\,f(\underline{v},x)\,\mathcal{E}(c,dx)\right\}\nu_\infty(d\underline{v}).$$

By the stochastic Fubini theorem (Durret [25]) one may then write

$$E\left[\frac{dP_X}{dP_N} \mid \sigma_t(\mathcal{E})\right][c] = 1 + \int_0^t \left\{\int_{I\!R_+^\infty} \Gamma(\underline{v},c,x)\,f(\underline{v},x)\,\nu_\infty(d\underline{v})\right\}\mathcal{E}(c,dx).$$

One finishes by applying the general theory (Mémin [56]). $\square$

Remark. If one assumes that all the functions $\dot{h}_j$ have a continuous derivative, then, as

$$\int_0^t \dot{h}_j(x)\,\mathcal{E}(c,dx) = \dot{h}_j(t)\,c(t) - \int_0^t c(x)\,\frac{d}{dx}\left[\dot{h}_j\right](x)\,dx,$$

α is “easily” computable.

As an immediate consequence of Proposition 3 one has the following representation result for local martingales with respect to the filtration $\underline{\sigma}(X)$. Indeed (see Lenglart [51]), we have the following corollary.

Corollary 3. *Let M be a local martingale with respect to $\underline{\sigma}(X)$. It has a continuous version in the form of*

$$M(\cdot,t) = \mu_M + \int_0^t f(\cdot,x)\, X(\cdot,dx) - \int_0^t f(\cdot,x)\, \alpha(X(\cdot,\cdot),x)\, dx,$$

where f is predictable and $\int_0^t f^2(\cdot,x)\, dx < \infty$, almost surely.

4.4. The Linear Space of a Skew-Normal Process

Interest in this space stems from the fact that it is the space in which linear estimation is performed: as the linear space of a Gaussian process is made of Gaussian random variables, that of a skew-normal process is made of skew-normal ones. It is however as one will see more difficult to "locate" independent random variables in it.

Consider the map $J_{N,X}.H(C) \to H(C_X)$ defined by

$$J_{N,X}[C(\cdot,\ t)] = C_X(\cdot,\ t).$$

A calculation yields that

$$J_{N,X}^{\star} J_{N,X} = I_{H(C)} + T.$$

$J_{N,X}$ is thus linear, bounded with bounded inverse so that $H(C)$ and $H(C_X)$ contain the same (continuous) functions. Let

$$U_X : C_X(\cdot,\ t) \mapsto \widetilde{X}(\cdot,t) = [X(\cdot,\ t) - E[X(\cdot,\ t)]]$$

denote the usual isometry. For $f \in H(C)$, set

$$\widetilde{X}[f] = U_X J_{N,X}[f] \quad \text{and} \quad X[f] = \widetilde{X}[f] + E[X[f]].$$

One has immediately that

$$E[X[f]] = \langle \mu_X, f \rangle_{H(C)} \quad \text{and} \quad E\left[\widetilde{X}[f]\,\widetilde{X}[g]\right] = \langle \left(I_{H(C)} + T\right) f, g \rangle_{H(C)}$$

as well as the following proposition.

Proposition 4. *Let $\{f_1,\ldots,f_p\} \subseteq H(C)$ and $f(\underline{\alpha}) = \sum_{k=1}^{p} \alpha_k f_k$. Then*

$$E\left[e^{i\sum_{k=1}^{p}\alpha_k X[f_k]}\right] = e^{-\frac{1}{2}\|f(\underline{\alpha})\|^2_{H(C)}} \prod_{j=1}^{\infty} \varphi_{\nu_j}\left(\langle h_j, f\left(\underline{\alpha}\right)\rangle_{H(C)}\right).$$

Proof. One has, using the notation met in the proof of Theorem 1,

$$\sum_{k=1}^{p} \alpha_k \widetilde{X}\left[f_k\right] = \lim_n U_X J_{N,X}\left(P_{H_n(C)}\left[f\left(\underline{\alpha}\right)\right]\right).$$

But, as Σ_n is the matrix with entries $C\left(t_k, t_l\right),\ 1 \leq k, l \leq n$,

$$P_{H_n(C)}\left[f\left(\underline{\alpha}\right)\right](t) = \langle \Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right),\ \mathcal{E}\left(C\left(t,\cdot\right), \underline{t}_n\right)\rangle_{I\!R^n},$$

so that

$$U_X J_{N,X} P_{H_n(C)}\left[f\left(\underline{\alpha}\right)\right] = \langle \Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right),\ \mathcal{E}\left(\widetilde{X}\left(\cdot,\cdot\right), \underline{t}_n\right)\rangle_{I\!R^n},$$

and

$$\sum_{k=1}^{p} \alpha_k X\left[f_k\right] = \lim_n \langle \Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right),\ \mathcal{E}\left(X\left(\cdot,\cdot\right), \underline{t}_n\right)\rangle_{I\!R^n}.$$

Thus

$$E\left[e^{i\sum_{k=1}^{p}\alpha_k X[f_k]}\right] = \lim_n E\left[e^{i\langle \Sigma_n^{-1}\mathcal{E}(f(\underline{\alpha}),\underline{t}_n),\ \mathcal{E}(X(\cdot,\cdot),\underline{t}_n)\rangle_{I\!R^n}}\right].$$

One must now use the formula for the characteristic function given in Section 3.2 with

$$\underline{\theta} = \Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right)$$

to obtain that

$$\langle \Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right)\mathcal{E}\left(h_j, \underline{t}_n\right)\rangle_{I\!R^n} = \langle h_j, P_{H_n(C)}\left[f\left(\underline{\alpha}\right)\right]\rangle_{H(C)}$$

and that

$$\langle \Sigma_n\left\{\Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right)\right\}, \Sigma_n^{-1}\mathcal{E}\left(f\left(\underline{\alpha}\right), \underline{t}_n\right)\rangle_{I\!R^n} = \left\|P_{H_n(C)}\left[f\left(\underline{\alpha}\right)\right]\right\|^2_{H(C)}.$$

Going to the limit, one has the result. □

Corollary 4. *Whenever the substitution "$\alpha_k \to -i\beta_k$" is legitimate,*

$$E\left[e^{\sum_{k=1}^{p}\beta_k X[f_k]}\right] = e^{\frac{1}{2}\left\|\sum_{k=1}^{p}\beta_k f_k\right\|^2_{H(C)}} E\left[e^{\langle |S|, \sum_{k=1}^{p}\beta_k f_k\rangle_{H(C)}}\right].$$

A sufficient condition (Kwapień [48]) will be $E\left[e^{\|S\|_{H(C)}}\right] < \infty$ and the latter obtains in particular (Kwapień [48]) whenever there exists $\alpha > 0$ such that

$$\sum_{n=1}^{\infty} E\left[I_{\left\{|X_n|>\frac{\alpha}{\|h_n\|_{H(C)}}\right\}} e^{|X_n|\|h_n\|_{H(C)}}\right] < \infty.$$

This condition becomes for:

1. double exponential random variables with parameter $\gamma_n > \|h_n\|_{H(C)}$ and notation $\theta_n = \frac{\gamma_n}{\|h_n\|_{H(C)}}$,

$$\sum_{n=1}^{\infty} \frac{e^{-\alpha\theta_n\left(1-\frac{1}{\theta_n}\right)}}{1-\frac{1}{\theta_n}} < \infty,$$

2. and for Gaussian random variables with variance σ_n^2 and notation $\theta_n = \sigma_n \|h_n\|_{H(C)}$,

$$\sum_{n=1}^{\infty} e^{\frac{1}{2}\theta_n^2}\left[1-\Phi\left(\frac{\alpha-\theta_n^2}{\theta_n}\right)\right] < \infty$$

Remarks. 1. Let H_0 be the closed subspace spanned in $H(C)$ by $\{h_n,\ n \in I\!N\}$. If $H_0^{\perp}$ is not the zero subspace, the linear space of X will contain Gaussian as well as skew-normal random variables.

2. When the ν_n stem from IID standard normal random variables, the formula of Proposition 4 becomes

$$E\left[e^{i\sum_{k=1}^{p}\alpha_k X[f_k]}\right]$$
$$= e^{-\frac{1}{2}\langle[I_{H(C)}+T]f(\underline{\alpha}),f(\underline{\alpha})\rangle_{H(C)}} \prod_{j=1}^{\infty}\left\{1+i\sqrt{\frac{2}{\pi}}\int_0^{\langle h_j,f(\underline{\alpha})\rangle_{H(C)}} e^{\frac{x^2}{2}}\,dx\right\},$$

highlighting the role of the covariance operator.

3. The product term in the characteristic function of Proposition 4 is rather unwieldy. It simplifies considerably when dealing with frames (Allen et al [1], Duffin et al [23], Ruskai et al [65]).

A sequence $\{h_n,\ n \in I\!N\} \subseteq H(C)$ is a frame for $H(C)$ if there exists constants $A,\ B > 0$ such that, for all $h \in H(C)$,

$$A\|h\|_{H(C)}^2 \le \sum_{n=1}^{\infty} \left|\langle h, h_n\rangle_{H(C)}\right|^2 \le B\|h\|_{H(C)}^2.$$

Let U be the operator that sends $h \in H(C)$ to the sequence in l_2 with entries $\langle h, h_n\rangle_{H(C)}$. The frame operator is the bounded linear operator with bounded inverse $S = U^{\star}U$ with the property that

$$Sh = \sum_{n=1}^{\infty} \langle h, h_n\rangle_{H(C)} h_n.$$

Whenever the frame is exact, that is, when no element of the frame sequence can be deleted from it without the sequence no longer being a frame, then the sequences

$$\{h_n,\ n \in I\!N\} \quad \text{and} \quad \left\{S^{-1}h_n,\ n \in I\!N\right\}$$

are biorthogonal and $\langle h_n, S^{-1}h_n\rangle_{H(C)} = 1$. Furthermore

$$\sup_n \|h\|^2_{H(C)} \leq B \quad \text{and} \quad \inf_n \|h\|^2_{H(C)} \geq A.$$

As it is seen this will impose constraints on ν^{∞}. But whenever these are satisfied one has that, for $f_k = S^{-1}h_k,\ 1 \leq k \leq q$,

$$\prod_{j=1}^{\infty} \varphi_{\nu_j}\left(\langle h_j, f\left(\underline{\alpha}\right)\rangle_{H(C)}\right) = \prod_{k=1}^{q} \varphi_{\nu_j}\left(\alpha_k\right).$$

4. If $\{h_n,\ n \in I\!N\} \subseteq H(C)$ is an orthogonal family and $f_k = h_k,\ 1 \leq k \leq p$, the characteristic function becomes

$$e^{-\frac{1}{2}\sum_{k=1}^{q} \alpha_k^2 \|f_k\|^2_{H(C)}} \prod_{k=1}^{q} \varphi_{\nu_k}\left(\alpha_k \|f_k\|^2_{H(C)}\right),$$

that is one has independent random variables.

5. Remarks 1, 3 and 4 show that the linear space of a skew-normal process is complex as it is rather difficult, if one does not restrict the signal space, to find f_k's that are simultaneously orthogonal and orthogonal to the h_n's.

6. Suppose that Corollary 4 obtains. A calculation yields then that

$$E\left[\left(e^{X[f]} - e^{X[g]}\right)^2\right]$$

$$= \left(E^{\frac{1}{2}}\left[e^{2\langle |S|, f\rangle_{H(C)}}\right] e^{\|f\|^2_{H(C)}} - E^{\frac{1}{2}}\left[e^{2\langle |S|, g\rangle_{H(C)}}\right] e^{\|g\|^2_{H(C)}}\right)^2$$

$$+ \quad 2\, e^{\frac{1}{2}\|f+g\|^2_{H(C)}} \times$$

$$\left(E^{\frac{1}{2}}\left[e^{2\langle |S|, f\rangle_{H(C)}}\right] E^{\frac{1}{2}}\left[e^{2\langle |S|, g\rangle_{H(C)}}\right] e^{\frac{1}{2}\|f-g\|^2_{H(C)}} - E\left[e^{\langle |S|, f+g\rangle_{H(C)}}\right]\right).$$

Consequently

$$\lim_{f\to g} E\left[\left(e^{X[f]}-e^{X[g]}\right)^2\right]=0.$$

Remark 6 yields the following result.

Corollary 5. *When Corollary 4 obtains, the map* $\Psi:\mathcal{L}_2[X]\to L_2[\sigma(X)]$ *defined by* $\Psi[X]=e^X$ *is continuous and the variables* $\left\{e^X,\ X\in\mathcal{L}_2[X]\right\}$ *generate* $L_2[\sigma(X)]$.

4.5. A Prediction Formula

One is able to compute straightforwardly the following prediction formula which shows that "skewing the normal" adds to the linear term of the Gaussian prediction component a non linear term, the "signal," which has however, with time, a diminishing impact as it contains a projection onto the orthogonal complement of "present time."

Proposition 5. *Define* $X_t:H(C)\to\mathcal{L}_2[X]$ *by*

$$X_t[h]=U_XJ_{N,X}P_{H_t(C)}[h]+\langle P_{H_t(C)}\mu_X,\ h\rangle_{H(C)}.$$

Let

$$\mu_t(d\underline{v}\mid X)=\frac{e^{X_t[J(\underline{v})]-\frac{1}{2}\left\|P_{H_t(C)}J(\underline{v})\right\|^2_{H(C)}}}{E_{\nu_\infty}\left[e^{X_t[J(\underline{v})]-\frac{1}{2}\left\|P_{H_t(C)}J(\underline{v})\right\|^2_{H(C)}}\right]}\,\nu_\infty(d\underline{v})$$

and $\theta>t$. *Then*

$$\begin{aligned}E[X_\theta\mid\sigma_t(X)]&=X_t[C(\cdot,\ \theta)]\\&\quad+\int_{\mathbb{R}_+^\infty}\langle P^\perp_{H_t(C)}C(\cdot,\ \theta),J(\underline{v})\rangle_{H(C)}\ \mu_t(d\underline{v}\mid X).\end{aligned}$$

Proof. One has that

$$E[X_\theta\mid\sigma_t(X)]=\lim_n E\left[X_\theta\mid\sigma_t\left(\mathcal{E}\left(X,\underline{t}^{(n)}\right)\right)\right]$$

and that

$$E\left[X_\theta\mid\mathcal{E}\left(X,\underline{t}^{(n)}\right)=\underline{u}_n\right]=\int_{\mathbb{R}}x\,\frac{f_{X_\theta,\mathcal{E}\left(X,\underline{t}^{(n)}\right)}(x,\underline{u}_n)}{f_{\mathcal{E}\left(X,\underline{t}^{(n)}\right)}(\underline{u}_n)}\,dx.$$

One then factors $f_{X_\theta,\mathcal{E}\left(X,\underline{t}^{(n)}\right)}$ into

$$\varphi_{[\mu_n(\theta,\underline{u}_n),\sigma_n^2(\theta)]}\left(x\right)\varphi_{[\underline{0}_n,\Sigma_n]}\left(\underline{u}_n\right)\int_{I\!R^\infty} e^{F(\theta,x,\underline{u}_n,\underline{v})} g\left(\underline{v}\mid\underline{u}_n\right)\nu_\infty\left(d\underline{v}\right),$$

where

$$\begin{aligned}
\mu_n\left(\theta,\underline{u}_n\right) &= \langle \Sigma_n^{-1}\mathcal{E}\left(C\left(\cdot,\ \theta\right),\underline{t}_n\right),\ \underline{u}_n\rangle_{I\!R^n},\\
\sigma_n^2\left(\theta\right) &= \left\| P^{\perp}_{H_n(C)}C\left(\cdot,\ \theta\right)\right\|^2_{H(C)},\\
F\left(\theta,x,\underline{u}_n,\underline{v}\right) &= L\left(\theta,x,\underline{u}_n,\underline{v}\right)-\frac{1}{2}\,\frac{\langle P^{\perp}_{H_n(C)}C\left(\cdot,\ \theta\right),\ J\left(\underline{v}\right)\rangle^2_{H(C)}}{\sigma_n^2\left(\theta\right)},\\
L\left(\theta,x,\underline{u}_n,\underline{v}\right) &= \frac{\langle P^{\perp}_{H_n(C)}C\left(\cdot,\ \theta\right),\ J\left(\underline{v}\right)\rangle_{H(C)}}{\sigma_n^2\left(\theta\right)}\ \left\{x-\mu_n\left(\theta,\underline{u}_n\right)\right\},
\end{aligned}$$

and recognizes again that, for $h\in H\left(C\right)$,

$$\langle \Sigma_n^{-1}\mathcal{E}\left(h,\underline{t}_n\right),\mathcal{E}\left(X,\underline{t}_n\right)\rangle_{I\!R^n} = U_X J_{N,X} P_{H_n(C)}h + \langle P_{H_n(C)}\mu_X, h\rangle_{H(C)}.$$

Integrating then with respect to x and then going to the limit yields the result. □

4.6. Representation of Skew-Normal Processes in Terms of the Underlying Normal: Their Cramér-Hida Representation

The Cramér-Hida representation (Cramér [22], Hida [37]) is the tool that allows one to keep using martingale methods when working with noises that are not martingales (Baker et al [9, 10, 11, 12], Climescu et al [20]). One sees below that knowing the Cramér-Hida representation of the underlying Gaussian process is enough to obtain that of the associated skew-symmetric one. The reason for such a result is the particular structure of the covariance operator as the identity plus an operator with finite trace (Corollary 1).

A calculation yields:

Lemma 3. *Define*

$$J_X^C : H\left(C\right)\rightarrow H\left(C\right)\quad by\quad J_X^C\left[C\left(\cdot,\ t\right)\right] = C_X\left(\cdot,\ t\right).$$

Then $J_X^C = I_{H(C)} + T$.

The (proper) Cramér-Hida representation (Cramér [22], Hida [37]) says that there exist

$$\{N_1, \ \dots \ , N_m\} \subseteq \mathcal{L}_2 [N] , \ m \leq \infty,$$

such that, defining

$$N_k (\cdot, t) = P_{\mathcal{L}_t[N]} N_k,$$

one obtains mutually independent Gaussian additive processes with variance functions $\mu_k (t)$ such that

$$\mu_m \ll \mu_{m-1} \ll \ \cdots \ \ll \mu_2 \ll \mu_1.$$

Furthermore

$$N (\cdot, t) = \sum_{k=1}^{m} \int_0^t F_k (t, x) N_k (\cdot, dx) ,$$

where

$$F_k (t, x) = 0 \text{ for } x > t \text{ and } \sum_{k=1}^{m} \int_0^T [F_k (t, x)]^2 \mu_k (dx) < \infty.$$

Let now G be a zero-mean Gaussian process whose law is equivalent to that of N and let $\mathcal{E}$ denote the evaluation process on $I\!R^{[0,T]}$. $(\mathcal{E}, P_G)$ has a nonanticipative representation with respect to $(\mathcal{E}, P_N)$ in the following sense (Kallianpur et al [47]): there exists a process $(\mathcal{F}, P_N)$ such that

$$E_{P_N} [\mathcal{F} (\cdot, \ s) \mathcal{F} (\cdot, \ t)] = E_{P_G} [\mathcal{E} (\cdot, \ s) \mathcal{E} (\cdot, \ t)]$$

and

$$[\mathcal{F} (\cdot, \ t)] \in \mathcal{L}_t [(\mathcal{E}, P_N)] ,$$

where $\mathcal{L}_t [(\mathcal{E}, P_N)]$ is the closed linear subspace of $L_2 [(\mathcal{E}, P_N)]$ generated by $\{[\mathcal{E} (\cdot, \ s)] , \ s \leq t\}$. $\mathcal{F}$ is obtained as follows. The Gohberg-Krein factorization (Gohberg et al [32]) says that in $H (C)$

$$J_Y^C = \left(I_{H(C)} + W_-\right) \Delta \left(I_{H(C)} + W_+\right) .$$

Given an operator R on $H (C)$, one has an operator $\widetilde{R}$ on $L_2 [(\mathcal{E}, P_N)]$ defined by $U_{\mathcal{E}} R U_{\mathcal{E}}^{\star}$, where $U_{\mathcal{E}} [C (\cdot, \ t)] = [\mathcal{E} (\cdot, \ t)]$. One must then choose (Kallianpur [47])

$$[\mathcal{F} (\cdot, \ t)] = \widetilde{\Delta}^{\frac{1}{2}} \left(I_{L_2[(\mathcal{E}, P_N)]} + \widetilde{W}_+\right) [\mathcal{E} (\cdot, \ t)] .$$

Let $\mathcal{E}_k$ play, in the present context, the role of N_k above. One then has

$$\mathcal{F}(\cdot,t) = \sum_{k=1}^{m} \int_0^t \left\{ F_k(t,x) + \sum_{l=1}^{m} \int_0^t F_{k,l}(x,y)\, F_l(t,y)\, \mu_l(dy) \right\} \mathcal{E}_k(\cdot,dx),$$

where

$$\sum_{k=1}^{m}\sum_{l=1}^{m} \int_0^T \int_0^T [F_{k,l}(x,y)]^2 \mu_k(dx)\, \mu_l(dy) < \infty.$$

Proposition 6. *Let $V = U_X U_N^{\star}$. X has then the representation*

$$X(\cdot,t) = \sum_{k=1}^{m} \int_0^t \left\{ F_k(t,x) + \sum_{l=1}^{m} \int_0^t F_{k,l}(x,y)\, F_l(t,y)\, \mu_l(dy) \right\} (V[\mathcal{E}_k])(\cdot,dx).$$

4.7. A Conclusion in the Form of a Question

The construction of skew-normal processes that has just been explained uses most of the facts that are known about the problem that was considered and the random objects that were involved. Indeed, as told in Introduction, all the available evidence says that for detection with respect to a Gaussian noise to be non singular the signal must be pathwise in the RKHS of the noise and that requirement is equivalent to the trace-class requirement of the construction (Lukic et al [54]). Furthermore the Karhunen-Loève representation (Karhunen [49], Loève [53]) says that second order processes can be obtained as random series. There are however two restrictions imposed by the construction that was used. The trace-class operator involved in it is positive whereas the Gaussian case allows for negative eigenvalues (Rao et al [64]) and the random elements in the Karhunen-Loève representation are uncorrelated rather than independent. These considerations lead to the question of how "thick" or "thin" the set of laws of skew-normal processes is in the class of laws P_X for which $P_X \ll P_N$. An answer to that question would delimit the modeling scope of skew-normal processes.

References

[1] R.L. Allen, D.W. Mills, *Signal Analysis,* IEEE Press, Piscataway, New Jersey (2004).

[2] R.B. Arellano-Valle, G. del Pino, E. San Martín, Definition and probabilistic properties of skew-distributions, *Statistics and Probability Letters,* **58** (2002), 111-121.

[3] N. Aronszajn, Theory of reproducing kernels, *Transactions of the American Mathematical Society,* **68** (1950), 337-404.

[4] A. Azzalini, A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics,* **12** (1985) 171-178.

[5] A. Azzalini, Further results on a class of distributions which includes the normal ones, *Statistica,* **46** (1986), 199-208.

[6] A. Azzalini, A. Dalla Valle, The multivariate skew-normal distribution, *Biometrika,* **83** (1996), 715-726.

[7] A. Azzalini, A. Capitanio, Statistical applications of the multivariate skew-normal distribution, *Journal of the Royal Statistical Society,* **61 B** (1999), 579-602.

[8] C.R. Baker, On equivalence of probability measures, *Annals of Probability,* **1** (1973), 690-698.

[9] C.R. Baker, Ed., *Stochastic Processes in Underwater Acoustics,* Lecture Notes in Control and Information Sciences **85**, Springer-Verlag, Berlin (1986).

[10] C.R. Baker, A.F. Gualtierotti, Discrimination with respect to a Gaussian process, *Probability Theory and Related Fields,* **71** (1986), 159-182.

[11] C.R. Baker, A.F. Gualtierotti, Absolute continuity and mutual information for Gaussian mixtures, *Stochastics and Stochastic Reports,* **39** (1992), 139-157.

[12] C.R. Baker, A.F. Gualtierotti, Likelihood ratio detection of stochastic signals, in *Advances in Statistical Signal Processing* **2** (Ed-s: V. Poor, J.B. Thomas), JAI Press, Greenwhich, Connecticut (1993), 1-34.

[13] A.V. Balakrishnan, *Communication Theory,* McGraw-Hill, New York (1968).

[14] P. Billingsley, *Convergence of Probability Measures,* Wiley, New York (1968).

[15] V.D. Briggs, Densities for infinitely divisible random processes, *Journal of Multivariate Analysis,* **5** (1975), 178-205.

[16] P.L. Brockett, The likelihood ratio detector for non-Gaussian infinitely divisible, and linear stochastic processes, *Annals of Statistics,* **12** (1984), 737-744.

[17] P.L. Brockett, H.G. Tucker, A conditional dichotomy theorem for stochastic processes with independent increments, *Journal of Multivariate Analysis,* **7** (1977), 13-27.

[18] P.L. Brockett, W.N. Hudson, H.G. Tucker, The distribution of the likelihood ratio for additive processes, *Journal of Multivariate Analysis,* **8** (1978), 233-243.

[19] A. Capitanio, A. Azzalini, E. Stanghellini, Graphical models for skew-normal variates, *Scandinavian Journal of Statistics,* **30** (2002), 129-144.

[20] A. Climescu-Haulica, A.F. Gualtierotti, Likelihood ratio detection of random signals: the case of causally filtered and weighted Wiener and Poisson noises, *IJPAM,* **2** (2002), 155-217.

[21] J.B. Conway, *A Course in Operator Theory,* American Mathematical Society, Providence, Rhode Isaland (2000).

[22] H. Cramér, On some classes of non-stationary stochastic processes, In: *Proc. of the 4th Berkeley Symposium in Math. Statist. and Prob.,* Part II (Ed-s: L. Le Cam, J. Neyman, E.L. Scott), University of California Press, Berkeley, California (1961), 57-77.

[23] R. Duffin, A. Schaeffer, A class of nonharmonic Fourier series, *Transactions of the American Mathematical Society,* **72** (1952), 341-366.

[24] T.E. Duncan, Likelihood functions for stochastic signals in white noise, *Information, Control,* **16** (1970), 303-310.

[25] R. Durret, *Stochastic Calculus, A Practical Intreoduction,* CRC Press, Boca Raton, Florida (1996).

[26] R.F. Dwyer, A statistical frequency domain signal processing method, In: *Statistical Signal Processing*, (Ed-s: E.J. Wegman, J.G. Smith), Dekker, New York (1984), 79-90.

[27] R.F. Dwyer, A technique for improving detection and estimation of signals contaminated by under ice noise, in *Statistical Signal Processing* (Ed-s: E.J. Wegman, J.G. Smith), Dekker, New York (1984), 153-165.

[28] J. Feldman, Equivalence and perpendicularity of Gaussain processes, *Pacific Journal of Mathematics,* **8** (1958), 699-708.

[29] X. Fernique, Extension du théorème de Cameron-Martin aux translations aléatoires, *Comptes Rendus de l'Académie des Sciences de Paris, I* **335** (2002), 65-68.

[30] I.I. Gikhman, A.V. Skorokhod, On densities of probability measures on function spaces, *Russian Mathematical Surveys*, **21** (1966), 83-156.

[31] J.V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theory of Probability and its Applications,* **5** (1960), 285-301.

[32] I.C. Gohberg, M.G. Krein, *Theory and Applications of Volterra Operators in Hilbert Space* (Eng. Trans.), American Mathematical Society, Providence, Rhode Island (1970).

[33] C. Gu, *Smoothing Spline ANOVA Models,* Springer Verlag, New York (2002).

[34] J. Hajek, On a property of normal distributions of any stochastic process, *AMS/IMS Selected Translations Math. Stat. Prob.*, **1** (1961), 245-252.

[35] D.A. Harville, *Matrix Algebra from a Statistician's Perspective,* Springer Verlag, New York (1997).

[36] C.W. Helstrom, *Statistical Theory of Signal Detection,* Second Edition, Pergamon, Oxford (1968).

[37] T. Hida, Canonical representations of Gaussian processes and their applications, *Memoirs of the College of Science,* University of Kyoto, **A** (1960), 33(1).

[38] C.W. Horton, Sr., *Signal Processing of Underwater Acoustic Waves,* U.S. Government Printing Office, Washington, D.C. (1969).

[39] Y.F. Huang, J.B. Thomas, Detection of constant signals in skewed noise, In: *Proceedings Nineteenth Annual Allerton Conference on Communications, Control and Computing* (1981), 607-616.

[40] W.N. Hudson, H.G. Tucker, Equivalence of infinitely divisible distributions, *Annals of Probability,* **3** (1975), 70-79.

[41] J. Jacod, J. Mémin, Caractéristiques locales, et conditions de continuité absolue pour les semi-martingales, *Zeit. Wahrsch. VerW. Gebiete,* **22** (1972), 1-36.

[42] J. Jacod, A.N. Shiryayev, *Limit Theorems for Stochastic Processes,* Springer Verlag, Berlin (1987).

[43] O.G. Jørsboe, *Equivalence or Singularity of Gaussian Measures on Function Spaces,* Various Publications Series **4**, Matematisk Institut, Aarhus Universitet, Aarhus, Denmark (1968).

[44] T. Kailath, On measures equivalent to Wiener measure, *Annals of Mathematical Statistics,* **38** (1967), 261-263.

[45] T. Kailath, The structure of Radon-Nikodým derivatives with respect to Wiener and related measures, *Annals of Mathematical Statistics,* **42** (1971), 1054-1067.

[46] T. Kailath, V.H. Poor, Detection of stochastic processes, *IEEE-IT,* **44** (1998), 2230-2259.

[47] G. Kallianpur, H. Oodaira, Non-anticipative representation of equivalent Gaussian processes, *Annals of Probability,* **1** (1973), 104-122.

[48] S. Kwapień, W.A. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple,* Birkhäuser, Boston (1992).

[49] K. Karhunen, Zur Spektraltheorie Stochastischer Prozesse, *Ann. Acad. Sci. Fennicae,* **37** (1946), 3-79 .

[50] M.H. Law, J.T. Kwok, Bayesian support vector regression, In: *Proceedings of the Eighth International Workshop on Artificial Intelligence and Statistics,* Key West, Florida (2001), 239-244.

[51] E. Lenglart, *Sur Quelques Points Remarquables de la Théorie des Martingales Locales et Applications,* Thèse, Université de Rouen (1976).

[52] R.S. Liptser, A.N. Shiryayev, On absolute continuity of measures corresponding to diffusion type processes, *Izv. AN SSSR, Ser. Matem.,* **36** (1972), 874-889.

[53] M. Loève, Fonctions aléatoires du second ordre, *Rev. Sci.,* **84** (1946), 195-206.

[54] M.N. Lukic, J.H. Beder, Stochastic processes with sample paths in reproducing kernel Hilbert space, *Transactions of the American Mathematical Society,* **353** (2001), 3945-3969.

[55] G.A. Mel'ničenko, The structure of processes that are absolutely continuous with respect to a Gaussian process, *Usp. Mat. Nauk.,* **32** (1977), 197-198.

[56] J. Mémin, *Sur Quelques Problèmes Fondamentaux de la Théorie du Filtrage,* Thèse, Université de Rennes (1974).

[57] J. Mémin, A.N. Shiryayev, Distance de Hellinger-Kakutani des lois correspondant à deux processus à accroissements indépendants, *Zeit. Wahrsch. VerW. Gebiete,* **70** (1985), 67-89.

[58] C.M. Newman, The inner product of path space measures corresponding to random processes with independent increments, *Bull. Amer. Math. Soc.,* **78** (1972), 268-271.

[59] C.M. Newman, On the orthogonality of independent increment processes, In: *Topics in Probability Theory* (Ed-s: D.W. Stroock, S.R.S. Varadhan), Courant Inst. Math. Sci., New York University, New York (1973), 93-111.

[60] K.R. Parthasarathy, *Probability Measures on Metric Spaces,* Academic Press, New York (1967).

[61] E. Parzen, Probability density functionals and reproducing kernel Hilbert spaces, In: *Proc. Symp. Time Series Analysis,* Wiley, New York (1963).

[62] E. Parzen, *Time Series Analysis Papers,* Holden-Day, San Francisco (1967) .

[63] A. Povzner, On a class of Hilbert spaces of functions, *Doklady Akad. Nauk. S.S.S.R. (N.S.),* **68** (1949), 817-820.

[64] C.R. Rao, V.S. Varadarajan, Discrimination of Gaussian processes, *Sankhya A,* **25** (1963), 303-330.

[65] M.R. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, L. Raphael, eds., *Wavelets and Their Applications,* Jones and Bartlett, Boston (1992).

[66] K.I. Sato, *Lévy Processes and Infinitely Divisible Distributions,* Cambridge University Press, Cambridge, UK (1999).

[67] K.I. Sato, *Density Transformation in Lévy Processes,* Lectures Notes for the Concentrated Advanced Course on Lévy Processes, MaPhySto, University of Aarhus (2000).

[68] S.C. Schwartz, J.B. Thomas, Detection in a non-Gaussian environment, In: *Statistical Signal Processing* (Ed-s: E.J. Wegman, J.G. Smith), Dekker, New York (1984), 93-105.

[69] A.V. Skorokhod, On the differentiability of measures which correspond to stochastic processes, I. Processes with independent increments, *Theory of Probability and its Applications,* **2** (1957), 407-432.

[70] A.V. Skorokhod, On the differentiability of measures which correspond to stochastic processes, II. Markov processes, *Theory of Probability and its Applications,* **5** (1960), 40-49.

[71] D. Slepian, Some comments on the detection of Gaussian signals in Gaussian noise, *IEEE-IT,* **4** (1958), 65-69.

[72] G. Wahba, *Spline Models for Observational Data,* CBMS-NSF Regional Conference Series in Applied Mathematics, **59**, SIAM, Philadelphia (1990).

[73] G.R. Wilson, D.R. Powell, Experimental and modeled density estimates of underwater acoustic returns, In: *Statistical Signal Processing* (Ed-s: E.J. Wegman, J.G. Smith), Dekker, New York (1984), 223-239.

Appendix: Convergence of Some Random Elements in $H(C)$

All the statements to follow are based on reference Kwapién et al [48, Chapter 2, Section 2.2]. Given a random element R in a normed space and $\alpha > 0$, $R^{(\alpha)}$ is the random element $RI_{\{\|R\| \leq \alpha\}}$.

Let $\{X_n,\ n \in I\!N\}$ be a sequence of independent random variables. Then

$$H_n = X_n h_n$$

is a random element in $H(C)$ and, by the three series theorem, $\sum_{n=1}^{\infty} H_n$ converges almost surely if and only if there exists an $\alpha > 0$ such that, simultaneously,

1. $\sum_{n=1}^{\infty} P\left(\|H_n\| > \alpha\right) < \infty$,
2. $\sum_{n=1}^{\infty} E\left[H_n^{(\alpha)}\right]$ is convergent,
3. $\sum_{n=1}^{\infty} V\left[H_n^{(\alpha)}\right]$ is convergent.

A sufficient condition for convergence in p-th mean is that there be some $\alpha > 0$ such that

$$\sum_{n=1}^{\infty} E\left[\|H_n\|_{H(C)}^p I_{\{\|H_n\|_{H(C)}>\alpha\}}\right] < \infty.$$

Given the definition of H_n, letting $\alpha_n = \frac{\alpha}{\|h_n\|_{H(C)}}$, these conditions translate into

1. $\sum_{n=1}^{\infty} P\left(|X_n| > \alpha_n\right) < \infty$,
2. $\sum_{n=1}^{\infty} E\left[X_n^{(\alpha_n)}\right] h_n$ is convergent,
3. $\sum_{n=1}^{\infty} V\left[X_n^{(\alpha_n)}\right] \|h_n\|_{H(C)}^2$ is convergent,
4. $\sum_{n=1}^{\infty} E\left[|X_n|^p I_{\{|X_n|>\alpha_n\}}\right] \|h_n\|_{H(C)}^p$ is convergent.

Particular cases (sufficient conditions) may be obtained as follows. Fix $p > 0$ and assume that there are strictly positive α, β and λ such that

$$\begin{aligned} P\left(|X_n| > \alpha\right) &\geq \beta, \\ &\text{and, for all } t > \alpha, \\ E\left[|X_n|^p I_{\{|X_n|>t\}}\right] &\leq \lambda t^p P\left(|X_n| > t\right), \end{aligned}$$

then, for each sequence $\{h_n \in H(C),\ n \in I\!N\}$,

$$\sum_{n=1}^{\infty} X_n h_n$$

converges almost surely if and only if it converges in p-th mean.

These conditions are illustrated below.

Double Exponential Random Variables

1. *Sufficient Conditions:* Suppose X_n has a law with density $\frac{\gamma_n}{2} e^{-\gamma_n |x|}$. Then

$$P\left(|X_n| > \alpha\right) = e^{-\gamma_n \alpha}$$

and $P(|X_n| > \alpha) \geq \beta$ implies that $\gamma_n \leq \frac{1}{\alpha} \ln \frac{1}{\beta}$. Furthermore

$$E\left[|X_n|^p I_{\{|X_n|>t\}}\right] = P(|X_n| > t) \sum_{i=0}^{p} \frac{p!}{i!} \frac{t^p}{\gamma_n^{p-i}}.$$

Consequently, for $p = 2$, one has that

$$E\left[|X_n|^2 I_{\{|X_n|>t\}}\right] = 2\left\{1 + \frac{1}{\gamma_n t} + \frac{1}{\gamma_n^2 t^2}\right\} t^2 P(|X_n| > t)$$

and the required condition obtains whenever the family $\{\gamma_n,\ n \in I\!N\}$ is bounded away from 0 and from ∞. As

$$E\left[\left\|\sum_{i=p}^{q} X_n h_n\right\|_{H(C)}^2\right] = 0$$

and

$$E\left[\left\|\sum_{i=p}^{q} |X_n|\, h_n\right\|_{H(C)}^2\right] = \sum_{i=p}^{q}\sum_{j=p}^{q} \frac{1}{\gamma_i\gamma_j}\langle h_i, h_j\rangle_{H(C)} = \left\|\sum_{i=p}^{q} \frac{h_i}{\gamma_i}\right\|_{H(C)}^2,$$

convergence almost surely will result from convergence of $\sum_{i=1}^{\infty} h_n$.

2. *Necessary and Sufficient Conditions:* The first condition requires that $\sum_{i=1}^{\infty} e^{-\alpha\gamma_i} < \infty$. For the second, one has that

$$E\left[|X_n|\, I_{\{|X_n|\leq\alpha_n\}}\right] = \left\{1 - \frac{1 + \alpha\frac{\gamma_n}{\|h_n\|_{H(C)}}}{e^{\alpha\frac{\gamma_n}{\|h_n\|_{H(C)}}}}\right\}\frac{1}{\gamma_n},$$

and, for the third,

$$E\left[X_n^2 I_{\{|X_n|\leq\alpha_n\}}\right] = \left\{1 - \frac{1 + \alpha\frac{\gamma_n^2}{\|h_n\|_{H(C)}^2} + \alpha^2\frac{\gamma_n}{\|h_n\|_{H(C)}}}{e^{\alpha\frac{\gamma_n}{\|h_n\|_{H(C)}}}}\right\}\frac{2}{\gamma_n^2}.$$

Thus γ_n must become large fast enough and adequately "control" the norm of h_n.

Normal Random Variables

1. *Sufficient Conditions:* Suppose $X_n \sim \mathcal{N}\left(0, \sigma_n^2\right)$. Then, using the inequalities

$$\frac{x}{1+x^2} e^{-\frac{x^2}{2}} \leq \int_x^{\infty} e^{-\frac{y^2}{2}} dy \leq \frac{1}{x} e^{-\frac{x^2}{2}},$$

$$P\left(|X_n| > \alpha\right) = 2\left\{1 - \Phi\left(\frac{\alpha}{\sigma_n}\right)\right\} \geq 2\frac{\alpha\sigma_n}{\alpha^2 + \sigma_n^2}\varphi\left(\frac{\alpha}{\sigma_n}\right)$$

and again the family $\{\sigma_n,\ n \in I\!N\}$ must be bounded away from 0 and ∞. Furthermore

$$E\left[|X_n|^2 I_{\{|X_n|>t\}}\right] \leq \sqrt{\frac{2}{\pi}}\sigma_n\left(1 + \sqrt{2\pi}\sigma_n\right) t^2 \left\{1 - \Phi\left(\frac{\alpha}{\sigma_n}\right)\right\},$$

so that mean-square convergence will obtain under conditions analogous to those that prevail in the double exponential case.

2. *Necessary and Sufficient Conditions:* With $\theta_n = \frac{\alpha}{\sigma_n \|h_n\|_{H(C)}}$, one has, for the first condition,

$$\sum_{n=1}^{\infty} P\left(|X_n| > \alpha_n\right) \leq \sum_{n=1}^{\infty} \frac{1}{\theta_n} e^{-\frac{1}{2}\theta_n^2}.$$

The second and third conditions require respectively the following formulae:

$$E\left[|X_n| I_{\{|X_n| \leq \alpha_n\}}\right] = \sqrt{\frac{2}{\pi}}\left\{1 - e^{-\frac{\theta_n^2}{2}}\right\}\frac{\alpha}{\theta_n \|h_n\|_{H(C)}}$$

and

$$E\left[X_n^2 I_{\{|X_n| \leq \alpha_n\}}\right] = 2\left\{\left(\Phi\left(\theta_n\right) - \frac{1}{2}\right) - \sqrt{\frac{2}{\pi}}\theta_n e^{-\frac{\theta_n^2}{2}}\right\}\frac{\alpha^2}{\theta_n^2 \|h_n\|_{H(C)}^2}.$$

Thus it suffices that θ_n becomes large fast enough.