Positive Definite and Unimodal Gauss-Hermite Expansion of Probability Density Functions

Wolfgang Freude

Summary · Zusammenfassung

The approximation of Gaussian-like probability density functions (p.d.f.) by Gauss-Hermite series of Gram-Charlier and Edgeworth type, or by a Cornish-Fisher expansion frequently violates two constraints: the curves should be positive definite, and unimodal. A new standardization allows the choice of the basic Gauss-Hermite system such that these diffieulties are avoided. If the p.d.f. is given, the expansion coefficients of the approximating series may be computed by a relative least-squares error fit; usually, the two p.d.f. constraints are not violated. If the moments are known, the expansion coefficients are computed by a matrix inversion, where an iterative procedure adjusts the standardizing parameters according to the two p.d.f. constraints. The method is compared to the approximation by a Pearson function.

Positiv-definite und unimodale Gauß-Hermite-Entwicklung von Wahrscheinlichkeitsdichte-Funktionen

Die Näherung von gaußähnlichen Wahrscheinlichkeitsdichte-Funktionen (p.d.f.) mit Gauß-Hermite-Reihen vom Gram-Charlier- und Edgeworth-Typ oder durch eine Cornish-Fisher-Entwicklung verletzt häufig zwei Nebenbedingungen: die Kurven sollten positiv-definit und unimodal sein. Eine neue Standardisierung vermeidet diese Schwierigkeit durch eine geeignete Wahl des Gauß-Hermite-Entwicklungssystems. Ist die p.d.f. bekannt, so kann man die Entwicklungskoeffizienten der Reihe durch eine Anpassung mit der Methode der kleinsten Quadrate berechnen; für gewöhnlich werden die Nebenbedingungen nicht verletzt. Sind die Momente bekannt, dann können die Entwicklungskoeffizienten durch eine Matrix-Inversion berechnet werden, wobei ein Iterationsprozeß die Standardisierungsparameter so anpaßt, daß die Nebenbedingungen erfüllt werden. Das Verfahren wird mit der Näherung durch eine Pearson-Funktion verglichen.

1. Introduction

When simulating coherent optical receivers numerically, it is often the case that in calculating the bit error probability $BER = \int p_x(x)\,dx$ either the probability density function (p.d.f.) $p_x(x)$ of the received signal x is known as a numerical table, or the moments of the signal may be estimated. In both cases it is useful to reconstruct an analytical formula for the p.d.f. with the constraints that it must be non-negative and with only one maximum, i.e., positive definite and unimodal.

Received October 30, 1989.

Priv.-Doz. Dr.-Ing. W. Freude, Institut für Hochfrequenztechnik und Quantenelektronik, Universität Karlsruhe, Postfach 69 80, D-7500 Karlsruhe 1.

AEÜ, Vol. 45
(1991), No. 2

In the past 40 years there have been numerous attempts to approximate p.d.f.'s in terms of the moments of the process. Special attention was paid to Gaussian-like p.d.f., which where either represented by Gauss-Hermite series of the Gram-Charlier and Edgeworth type, or, for the Cornish-Fisher expansion, by approximating the standardized argument z with normal distribution function $P_g(z)$ by a Hermite series of the non-normally distributed argument x of the actual distribution $P_x(x)$. As a reference, see, for example, [1, Sections 17.6–17.7], [2], [3, Sections 26.2.47–26.2.48, 26.9 example 6], [4]. Closely related to the Gram-Charlier and Edgeworth expansions is the saddlepoint approximation [5], which in effect differs in the standardization of the random variable.

For all of these approximations the fundamental (Gauss-)Hermite function system was uniquely defined by the very moments of the process, so it happened not infrequently that the two basic constraints of a p.d.f. were violated. As no degree of freedom was left, one had to take or to leave whatever the method presented for the chosen expansion type and the given process. Especially for long-tailed p.d.f., or for the calculation of very low error probabilities, well-behaved representations of p.d.f. are requested.

The importance of the problem has been recognized long since in connection with the expansions of Gram-Charlier and Edgeworth. An early remark is found in [6]. First in [7], later on in [8], [9] moment regions were calculated (for the first four moments only) in which the constraints above were violated; also, there are notes in [10], [11].

A different approach is to approximate a p.d.f. by the family of the p.d.f.'s associated with the Pearson distributions [12, Section 6]. e.g., the beta-density function. The parameters of these p.d.f. may be calculated from the moments. For a specified region of moments the functions are positive definite and unimodal. The relations are published for a set of four moments.

No further attempts to tackle the problem are known to the author. The reason for this deficiency might be that the statistics community lacked interest in esoteric Gaussian-like p.d.f. tail regions where for optical coherent receivers [13] symbol transmission error probabilities as low as 10^{-9} are to be calculated.

The present paper discusses a Gauss-Hermite approximation, and proposes a different standardization of the random variable with two or three process-independent parameters, which may be varied in an iterative manner to find a positive definite and unimodal representation out of the infinite manifold of density functions with a given finite number of identical moments.

For analytically or numerically given p.d.f. a relative least-squares fit (r.l.s.f.) is applied to compute the expansion coefficients of a Gauss-Hermite series, a method which allows a simple weighting of the available data, and needs not the problematic direct calculation of moments [10], [11]. If the moments of the process can be estimated, a method is proposed to reconstruct the p.d.f. of the process as a Gauss-Hermite series by a simple matrix inversion. The techniques described are demonstrated for non-Gaussian p.d.f. and compared to existing methods.

2. Relative Least-Squares Error Fit

A statistical process with p.d.f. $p_x(x)$ has an outcome which is described by the random variable x. If μ_1' and σ are mean and standard deviation of the process, the corresponding standardized random variable X' is usually defined by

$$X' = (x - \mu_1')/\sigma \,. \tag{1}$$

Here, a different standardization is chosen as was already done in characterizing single-mode optical waveguides [14], [15], where the (not yet determined) paramcters x_0 and χ resemble μ_1' and σ^{-1}, respectively,

$$X = \chi(x - x_0) \,. \tag{2}$$

Consequently, a series expansion of $p_x(x)$ in terms of the Gauss-Hermite function system R_n with the parameters x_0, χ, and the Hermitc functions H_n [3], may be written as

$$p_x(x) \approx \sum_{n=1}^{N} c_n R_{n-1}(X) \Big/ \int_{-\infty}^{+\infty} \sum_{n=1}^{N} c_n R_{n-1}(X)\, \mathrm{d}x \,,$$
$$R_n(X) = \exp(-X^2/2)\, H_n(X) \,. \tag{3}$$

Aside from the new standardization, eq. (3) may equally well represent the Gram-Charlier, Edgeworth, or saddlepoint expansion, if the coefficients c_n are computed properly. It is interesting to note that the structure of the corresponding characteristic function $C_x(\xi)$ with the standardized variable Ξ is invariant with respect to a Fourier transformation,

$$C_x(\xi) = \overline{\exp(-\mathrm{i}2\pi\xi x)} = \int_{-\infty}^{+\infty} p_x(x) \exp(-\mathrm{i}2\pi\xi x)\, \mathrm{d}x$$
$$= \left(\sqrt{2\pi}/\chi\right) \exp(-\mathrm{i}\chi\Xi x_0) \sum_{n=1}^{N} (-\mathrm{i})^{n-1} c_n R_{n-1}(\Xi),$$
$$\Xi = 2\pi\xi/\chi \,. \tag{4}$$

If the p.d.f. $p_x(x)$ is given analytically or numerically, the basic parameters x_0 and χ may be found either by computing the first ordinary and the second central moment, or, as it is done here, by least-squares fitting a Gaussian $R_0(X)$ to the values $p_x(x)$ with $x = x_k$, $X = X_k$ inside a reduced range $K_1 \le k \le K_2$ of the main lobe,

$$\sum_{k=K_1}^{K_2} [p_x(x_k) - A_0 R_0(X_k)]^2 = \min \,,$$
$$\ln\{A_0 R_0[\chi(x_k - x_0)]\} = A_1 + A_2 x_k + A_3 x_k^2 \,. \tag{5}$$

With the abbreviations $\boldsymbol{B}$, $\boldsymbol{A}$ for column vectors with the elements B_l, $A_{l>0}$, and $\boldsymbol{S}$ for Hermitian matrices with the matrix elements S_{lt}, one has to solve the

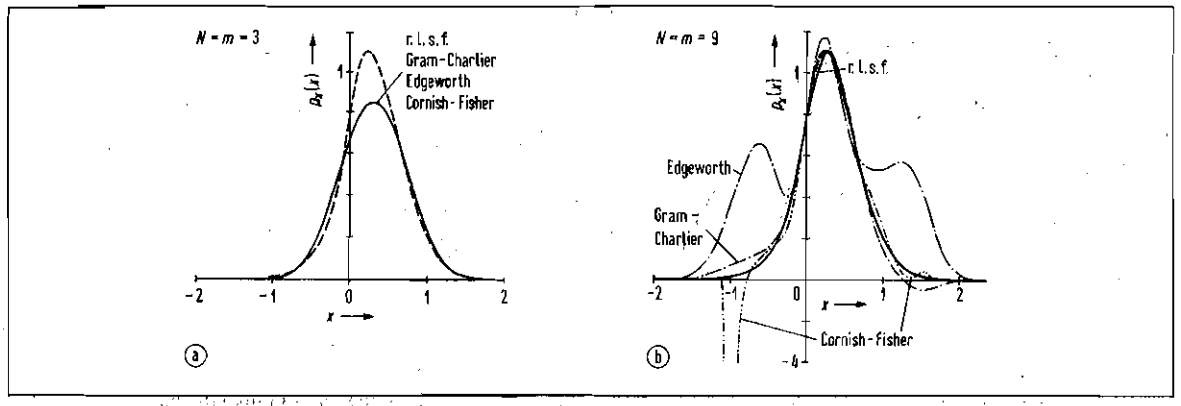

Fig. 1. Comparison between various approximations for the non-Gaussian p.d.f. eq. (13). The fitting range is $-1 \le x \le 1.5$.
(a) Maximum order of moments is $m=3$.
(b) Maximum order of moments is $m=9$; for r.l.s.f. the r.r.m.s. error is 20%.

system of equations for A,

$$\boldsymbol{B}=\boldsymbol{S}\boldsymbol{A}\,,\quad \boldsymbol{A}=\boldsymbol{S}^{-1}\boldsymbol{B}\,,$$
$$B_l=\sum_{k=K_1}^{K_2} x_k^{l-1}\ln p_x(x_k)\,, \tag{6}$$
$$S_{li}=\sum_{k=K_1}^{K_2} x_k^{l+i-2}=S_{il}\,.$$

For the parameters A_0, x_0 and χ one finds

$$A_0=\exp[A_1-A_2^2/(4A_3)]\,,$$
$$x_0=-A_2/(2A_3)\,,\quad \chi=(-2A_3)^{1/2}\,. \tag{7}$$

The expansion coefficients for $p_x(x)$ in the chosen function system are computed by minimizing the sum of the squared *relative* errors for all values $p_x(x)$ with $x=x_k$, $X=X_k$, $1\le k\le K$,

$$\sum_{k=1}^{K} W_k^2\left[p_x(x_k)-\sum_{n=1}^{N} c_n R_{n-1}(X_k)\right]^2=\min\,,$$
$$W_k=1/p_x(x_k)\,. \tag{8}$$

For the coefficient vector $\boldsymbol{C}$ with the elements c_n one has to solve the system of equations

$$\boldsymbol{B}=\boldsymbol{S}\boldsymbol{C}\,,\quad \boldsymbol{C}=\boldsymbol{S}^{-1}\boldsymbol{B}\,,$$
$$B_l=\sum_{k=1}^{K} W_k^2\,p_x(x_k)\,R_{l-1}[\chi(x_k-x_0)]\,, \tag{9}$$
$$S_{ln}=\sum_{k=1}^{K} W_k^2\,R_{l-1}[\chi(x_k-x_0)]\,R_{n-1}[\chi(x_k-x_0)]=S_{nl}\,.$$

The normalization integral in eq. (3) modifies in effect the c_n computed from eq. (9) by a constant factor, which usually deviates only by small amounts from one.

The moments μ'_m of order m about the origin are

$$\mu'_m(x_0,\chi)=\int_{-\infty}^{+\infty} x^m\,p_x(x)\,\mathrm{d}x=\sum_{n=1}^{N} d_{mn}(x_0,\chi)\,c_n\,, \tag{10}$$

$$d_{mn}(x_0,\chi)=(\sqrt{2}/\chi)\int_{-\infty}^{+\infty}(\sqrt{2}\,u/\chi+x_0)^m\cdot$$
$$\cdot\exp(-u^2)\,H_{n-1}(\sqrt{2}\,u)\,\mathrm{d}u\,. \tag{11}$$

The integration variable u is chosen to facilitate the application of the Hermite integration technique, a numerically very fast formula of Gaussian type [3, Section 25, Formula 25.4.46, Table 25.10] for u_l and w_l,

$$\int_{-\infty}^{+\infty}\exp(-u^2)\,f(u)\,\mathrm{d}u= \tag{12}$$
$$=\begin{cases}\displaystyle\sum_{l=1}^{I} w_l\,[f(u_l)+f(-u_l)] & I \text{ even}\,,\\ \displaystyle w_1\,f(u_1=0)+\sum_{l=2}^{I} w_l\,[f(u_l)+f(-u_l)] & I \text{ odd}\,.\end{cases}$$

The result is exact, if (as it is the case here) $f(u)$ is a polynomial of maximum degree $n-1+m$, and if $I=[(n-1+m)/2]+1$ is valid ([z] stands for the integer part of z).

The advantages of the proposed r.l.s.f. method over the Gram-Charlier, Edgeworth and Cornish-Fisher expansions will be demonstrated for the p.d.f.

$$p_x(x)=\frac{1}{(2\pi\,0.5^2)^{1/2}}\cdot$$
$$\cdot\exp[-x^2/(2\cdot 0.5^2)]\left[1+\frac{2}{\pi}\arctan(5x)\right]. \tag{13}$$

Fig. 1 shows this function with dashed lines. Fig. 1 a displays as a heavy line the r.l.s.f. approximation with $N=3$ coefficients in eq. (3), i.e., for moments up to a maximum order $m=3$. The equivalent curves for the approximations Gram-Charlier, Edgeworth, and Cornish-Fisher are taken from [16], [17] (both for Fig. 1 a and Fig. 1 b), and are for $m=3$ not to be discriminated from the r.l.s.f. graph. As to be expected from the low number of expansion terms, the quality of the approximations is poor.

The various expansions differ much more, if nine-term series $N=9$ with moments up to the maximum order $m=9$ are evaluated, Fig. 1 b. For the r.l.s.f curve the relative root mean square (r.r.m.s.) error for the fitting range $-1\le x\le 1.5$ in comparison to the ideal curve eq. (13) amounts to 20%; the deviation is hardly to be seen left of the peak. The three other expansions exhibit large errors, especially negative p.d.f. ranges.

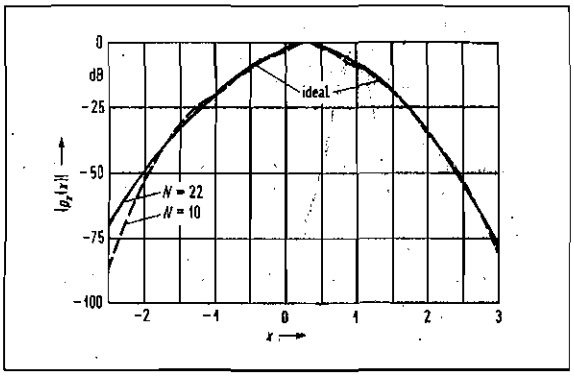

Fig. 2. Non-Gaussian p.d.f. eq. (13) for $N = 10$ (r.r.m.s. error 41%, see Fig. 1b) and $N = 22$ (r.r.m.s. error 50%) coefficients. The fitting range is $-4.6 \leq x \leq 5.62$.

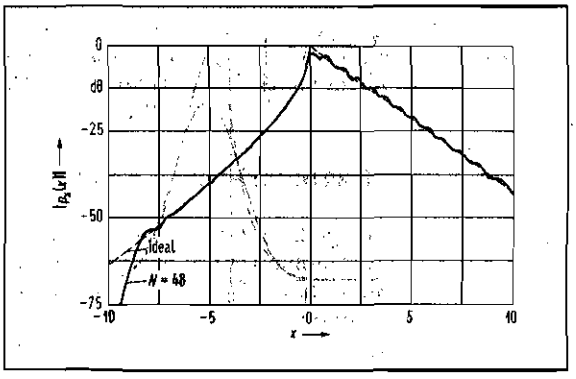

Fig. 3. Slowly decaying non-Gaussian p.d.f. eq. (14) for $N = 48$. The r.r.m.s. error is 31% in a fitting range $-9.58 \leq x \leq 10.74$.

These irregularities increase with increasing maximum order of the moments, as opposed to the r.l.s.f. technique. This is partly due to the fact that in performing the moment integral eq. (10) a sum of Gaussian-like decaying terms has to be evaluated, and that, therefore, computers with finite accuracy evaluate the tail contributions erroneously [14], but partly the cause for this behaviour lies also in the fixed Gauss-Hermite function system, which is defined by the moments themselves.

The r.l.s.f. method guarantees that the p.d.f. tails are as much taken into account as the central parts. No time-consuming and error-prone numerical integration is required, only the solution of a system of equations where fast and accurate standard algorithms are available. Therefore, it is unlikely (but not impossible!) that negative p.d.f. regions occur. As an example, Fig. 2 shows the ideal (eq. (13)) and the r.l.s.f. curves for $N = 10$ (as in Fig. 1b) and $N = 22$ coefficients in a semilogarithmic display, where $|p_x(x)|$, measured in dB (decibel), is given by $10 \lg(|p_x(x)|)$. The approximations are positive definite and unimodal all over the real x-axis. The r.r.m.s. errors 41% and 50% are larger than in Fig. 1b because of the larger fitting range $-4.6 \leq x \leq 5.62$. For the curve $N = 22$ of Fig. 2 no significant deviation from the ideal is to be seen, but it becomes lower by 45 dB on the left hand side of the fitting range at $x = -4.6$, where the ideal value would have been about -200 dB. On the right hand side for $x = 5.62$ the level (ideally about -270 dB!) becomes lower by only 22 dB. Obviously, the number of coefficients should be increased for such demanding requirements.

An especially critical example of a long-tailed distribution is

$$p_x(x) = \frac{1}{4 \cdot 0.5} \exp[-|x|/(2 \cdot 0.5)] \left[1 + \frac{2}{\pi} \arctan(5x)\right], \tag{14}$$

differing from eq. (13) by its slow exponential decay and its sharp edge at $x = 0$. Fig. 3 gives the approximation for $N = 48$ coefficients. The r.r.m.s. error is 31% in a fitting range of $-9.58 \leq x \leq 10.74$. As in Fig. 1 and Fig. 2, the approximation is positive definite, and unimodal. On the left and right hand side of the fitting range the curve is lower by about 25 dB and 21 dB, respectively, as compared to eq. (14).

3. Iteration Procedure to Satisfy p.d.f. Constraints

If with x_0, χ a certain function system is chosen, the elements d_{mn}, eq. (11), of a matrix $\boldsymbol{D}$ are determined independently of the elements μ'_m, eq. (10), which constitute the moment vector $\boldsymbol{M}'$. Consequently, it makes sense to write eq. (10) in matrix form,

$$\boldsymbol{M}' = \boldsymbol{D}\boldsymbol{C}, \quad \boldsymbol{C} = \boldsymbol{D}^{-1}\boldsymbol{M}', \tag{15}$$

so that the elements c_n of the coefficient vector $\boldsymbol{C}$ in the p.d.f. expansion eq. (1) are to be computed by a simple matrix inversion ($\boldsymbol{D}$ must be square for that purpose, i.e., the maximum number of moments must equal the number N of coefficients).

In the conventional case of the Gram-Charlier or Edgeworth series, where $x_0 = \mu'_1$, $\chi = 1/\sigma = (\mu'_2 - \mu'^2_1)^{-1/2}$ are valid by definition, $\boldsymbol{C}$ is completely fixed by $\boldsymbol{M}'$, and if the resulting p.d.f. eq. (1) derived from, say, the first ten moments is not positive definite by chance (which is very unlikely for an infinite range of the random variable x), the expansion is physically meaningless.

With the present method, the system parameters x_0 and χ may be adjusted starting from ther initial values above, so that the p.d.f. obeys the constraints. If for a given $\boldsymbol{M}'$ this cannot be achieved, the number of relevant moments (and coefficients!) may be reduced by one, and the procedure repeated.

For adjusting the parameters x_0 and χ, an objective function $g(x_0, \chi)$ is to be minimized. Most important, the positive-definiteness of $p_x(x)$ has to be tested. This may be done efficiently by computing the real zeros of the (for $N \leq 10$ explicitly programmed) polynomial $S = \sum_{n=1}^{N} c_n H_{n-1}(X)$ in eq. (1). The ranges X_{neg} for a negative p.d.f. are checked, and g is defined as the maximum of the dominant exponential for the region

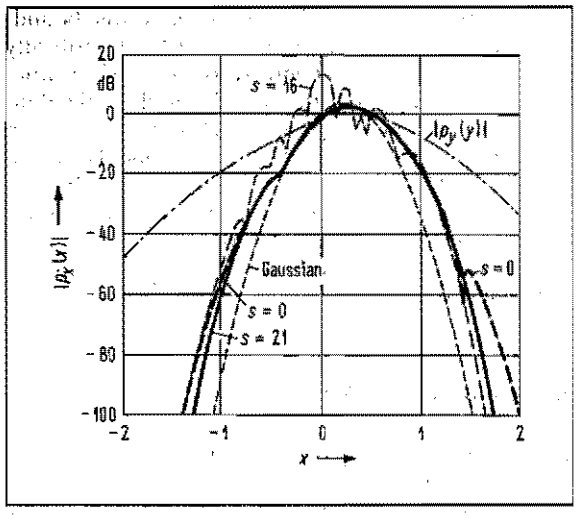

Fig. 4. Filter input p.d.f. $|p_y(y)|$; filter output p.d.f.'s $|p_x(x)|$: Gaussian, initial and final (——), intermediate (sth iteration).

X_{neg}, first line of eq. (16),

$$g = \begin{cases} \max[\exp(-X_{neg}^2/2)]/E \geq 0 \\ \qquad \text{if ranges } X_{neg} \text{ exist,} \\ -(N-Z)E \leq 0 \quad \text{if no ranges } X_{neg} \text{ exist.} \end{cases} \tag{16}$$

E is a normalization factor, e.g., $E = 10^{-9}$, as an estimate of the wanted probability to be computed from $p_x(x)$. If after a proper variation of x_0 and χ no ranges X_{neg} exist any more, $p_x(x)$ is positive definite. When this is true, the objective function is switched to a different definition, second line of eq. (16), for improving the smoothness of $p_x(x)$ by checking the derivative S' of S. Z is the number of real zeros ($1 \leq Z \leq N$) of the (for $N \leq 10$ explicitly programmed) derivative S'; this secondary objective function g counts Z, yielding a measure of the undulation in $p_x(x)$. A standard simplex procedure, see, e.g., [18, Section 2.2], may search for a minimum g while varying the set x_0, χ. For each set the corresponding $\boldsymbol{D}$, and, by matrix inversion, the elements of $\boldsymbol{C}$ are computed to determine the polynomials S, S'.

Fig. 4 shows some numerically reconstructed p.d.f. in a semi-logarithmic display. The distorted Gaussian-like p.d.f. $p_y(y)$ eq. (13) of the new random variable y is represented by a series eq. (1); the c_n were found by the r.l.s.f. technique described above, ($x_0 = 0.31$, $\chi = 2.6$). The random variable y be statistically independent, and its moments transformed by a digital transversal lowpass filter of second order with 100 coefficients and a ratio 'corner frequency to reciprocal sampling interval' of 0.1. The initial p.d.f. $p_x(x)$ at the filter output, reconstructed from the inversion eq. (15) with $x_0 = \mu_1$, $\chi = \sigma^{-1}$ (μ_1, σ being the first two moments at the filter output), is designated by ($s = 0$: $x_0 = 0.31$, $\chi = 4.6$); the p.d.f. defined by only the first two moments $s = 0$ is named 'Gaussian'. x_0, χ and therefore $\boldsymbol{D}$ and $\boldsymbol{C}$ for the fixed moments at the filter output are iterated from the initial state $s = 0$ through a very bad intermediate state ($s = 16$: $x_0 = 0.14$, $\chi = 5.7$) to the final, smooth state ($s = 21$: $x_0 = 0.22$, $\chi = 5.2$). Each kink in the curves $s = 0$, $s = 16$ indicates a sign change of the p.d.f. It is remarkable that the coarse structure of the p.d.f. remains unchanged, while their details vary strongly with small changes in x_0, χ. The computing time for the iteration using FORTRAN 77 and a PC/AT DOS 3.30 machine with a coprocessor clock frequeney of 12 MHz was only 40 s.

4. Double-Gauss-Hermite Function System

So far, the basic function system has a Gaussian decay factor which is identieal for both the $+X$ and the $-X$ directions in eq. (3). As a consequence, only asymptotically symmetric p.d.f. will be approximated with low errors. If this symmetry is strongly disturbed as for phase noise [19], [20, Section 3.3.4] $x = \cos z$ with a Gaussian density of z, the convergence properties of the iteration procedure become worse.

Therefore, $R_n(X)$ eq. (3) was generalized using different Gaussian functions for different signs of X, namely

$$X_1 = \chi_1(x - x_0), \quad X_2 = \chi_2(x - x_0),$$
$$R_n(X) = \begin{cases} R_n(X_1) & \text{for } x \geq x_0 \\ R_n(X_2) & \text{for } x \leq x_0 \end{cases} \tag{17}$$

Now, the basic function system is characterized by the three parameters x_0, χ_1, and χ_2. As before, the integrals in eq. (11) (this time of the type $\int_0^\infty u^m \exp(-u^2)\,du$) may be solved analytically, so d_{mn} can be calculated efficiently by summing up precomputed factors which are weighted by powers of x_0, χ_1, and χ_2. As a consequence, the computational effort does not increase compared with the single-Gauss-Hermite expansion.

The techniques described above were tested with output data x from the simulation of an optical DPSK heterodyne receiver with dominant phase noise (DPSK: differential phase shift keying). For computing the *BER* the p.d.f. $p_x(x)$ of the time sequence $x = x_1, x_2, \ldots, x_t$ of output data in front of the decision eircuit has to be reconstructed from an estimate of the associated moments. Details of the simulation and of this estimation process are of no importance for the present discussion. For a rough estimate, the p.d.f. was reconstructed directly by computing the span of the sorted sequence $x_{i-1} < x_i$ for a fixed subscript interval $i \ldots i+J$, i.e., $p_x[(1/2)(x_{i+J} + x_i)] \sim J/(x_{i+J} - x_i)$ (assuming that the probability of identieal values x_i in the subscript interval J is negligible). This technique [21, Section 13.3] is far superior to the usual method, where the frequency of data for a fixed interval of x is counted. Next, the Pearson method for a beta-p.d.f. using four (and only four) moments was employed with $p_x(x) = (1 + x/a_1)^p (1 - x/a_2)^q$; the parameters a_1, a_2, p and q are uniquely determined by the moments of the process. Finally, the Gauss-Hermite and the

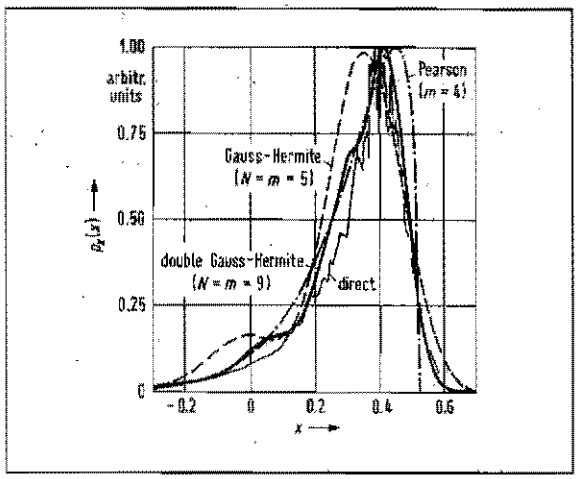

Fig. 5. Reconstruction of a p.d.f. from phase noise sample data: direct method, Pearson (beta-p.d.f., 4 moments), Gauss-Hermite (5 moments), double-Gauss-Hermite (9 moments).

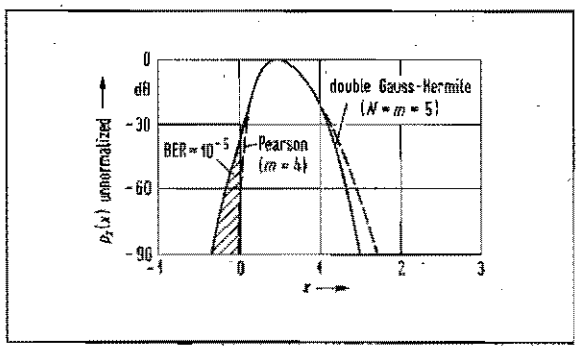

Fig. 6. Reconstruction of a p.d.f. from Gaussian noise sample data: Pearson (beta-p.d.f., 4 moments), double-Gauss-Hermite (9 moments).

double-Gauss-Hermite approximations were calculated; the maximum useful number of moments to reach convergence in the sense of the minimizing criteria eq. (16) for varying x_0, χ and x_0, χ_1, χ_2 was $N=m=5$ and $N=m=9$. The results are shown in Fig. 5. Obviously, the double-Gauss-Hermite technique is closest (in mode and curve details) to the direct reconstruction. As is to be expected, the simple Gauss-Hermite series reacts 'unwillingly' in approximating an asymptotically asymmetric p.d.f. The linear display of Fig. 5 is not very informative if low *BER* are to be estimated.

In a different simulation run only additive Gaussian noise was assumed to disturb the DPSK receiver. The model parameters were chosen such that a simple theory [22] yielded $BER = \int_{-\infty}^{0} p_x(x)\,dx \approx 10^{-5}$. The semi-logarithmic display Fig. 6 compares the p.d.f.'s from the Pearson and the double-Gauss-Hermite reconstruction; both curves reproduce *per definitionem* identical moments up to the order four. A strong discrepancy is to be seen. First, the limited definition range of the Pearson function with $a_1 \le x \le a_2$, $a_1 = -0.05$, $a_2 = 3.1$ shows up for the lower boundary. Secondly, the Pearson function is asymptotically strongly asymmetric in contrast to the physical situation. The evaluation of the *BER* integral for the double-Gauss-Hermite curve yields $BER \approx 0.9 \cdot 10^{-5}$ in close agreement with the theoretical value. However, for $x \le 0$ the order of magnitude of the Pearson function is seen to be smaller by a factor of 1000, and extended over a narrower range, so with $BER \approx 4 \cdot 10^{-11}$ the error probability is grossly underestimated.

5. Further Improvements

In reconstructing a p.d.f. from moments it was assumed that the true moments were known. However, usually the moments are only estimated from a data sample, and are therefore inaccurate. If moment error bounds can be computed, the iteration process minimizing the objective function eq. (16) can be improved by varying not only the parameters of the function system, but also the moments themselves inside their error bounds. This should add enough degrees of freedom to fulfill the p.d.f. constraints. Work is in progress to develop error bounds for various moment estimation techniques, and to add the moment variation to the iteration algorithm.

6. Conclusion

Probability density functions may be expanded in Gauss-Hermite series by a relative least-squares error fit, thereby avoiding the problems of negative regions and non-unimodality usually associated with the standard series expansions derived from moments. The moments of a statistically independent random variable may be efficiently computed, then, and transformed by linear filters in explicit form.

To reconstruct a p.d.f. from given moments the proposed iteration algorithm is well suited. Its main numerical advantages are the formulation for a fast numerical integration, the determination of the expansion coefficients by matrix inversion, and the direct search for zeros of the basic polynomials.

Acknowledgements

Thanks are due to Ir. F. Libbrecht, Laboratorium voor Electromagnetisme en Acustica, Rijksuniversiteit Gent, Belgium, for referring to the Pearson distributions. Further, the work of A. Schönfelder is acknowledged who contributed to Figs. 5 and 6.

References

[1] Cramér, H.: Mathematical methods of statistics. 10th printing. Princeton: Univ. Press, 1966.

[2] Cramèr, H.: Studies in the history of probability and statistics. XXVIII. On the history of certain expansions

used in mathematical statistics. Biometrika **59** (1972), 205–207.

[3] Abramowitz, M.; Stegun, I. A.: Handbook of mathematical functions. 9th printing. New York: Dover Publ., 1970.

[4] Fisher, R. A.; Cornish, E. A.: The percentile points of distributions having known cumulants. Technometrics **2** (1960), 209–225.

[5] Daniels, H. E.: Exact saddlepoint approximations. Biometrika **67** (1980), 59–63.

[6] Shenton, L. R.: Efficiency of the method of moments and the Gram-Charlier type A distribution. Biometrika **38** (1951), 58–73.

[7] Barton, D. E.; Dennis, K. E.: The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal. Biometrika **39** (1952), 425–427.

[8] Pearson, E. S.: Note on Mr Quenouilles Edgeworth type A transformation. Biometrika **46** (1959), 203.

[9] Draper, N. E.; Tierney, D. E.: Regions of positive and unimodal series expansion of the Edgeworth and Gram-Charlier approximations. Biometrika **59** (1972), 463–465.

[10] Pearson, E. S.: Some problems arising in approximating to probability distributions, using moments. Biometrika **50** (1963), 95–111.

[11] Brunk, H. D.: Univariate density estimation by orthogonal series. Biometrika **65** (1978), 521–528.

[12] Kendall, M. G.; Stuart, A.: The advanced theory of statistics. Vol. I. London: Charles Griffin, 1969.

[13] Corvini, P. J.; Koch, L. T.: Computer simulation of high-bit-rate optical fiber transmission using single-frequency lasers. IEEE J. Lightwave Technol. LT-**5** (1987), 1591–1595.

[14] Freude, W.; Sharma, A.: Refractive-index profile and modal dispersion prediction for a single-mode optical waveguide from its far-field radiation pattern. IEEE J. Lightwave Technol. LT-**3** (1985), 628–634. Printers correction: *ibid.* LT-**4** (1986), 375. Note the following additional typesetting errors: Two equations above (18), $\cdots = \sum_{\mu=1}^{M} c_\mu \chi^2 \cdots$ should read $\cdots = -\sum_{\mu=1}^{M} c_\mu \chi^2 \cdots$; in (19) replace $8\pi^2$ by $4\pi^2$.

[15] Freude, W.; Richter, H.: Refractive-index profile determination of single-mode fibres by far-field power measurements at 1300 nm. Electron. Lett. **22** (1986), 945–947.

[16] Cygan, D.: Berechnung der Wahrscheinlichkeitsdichtefunktion am Ausgang eines Filters bei beliebiger Eingangsverteilung und beliebiger Autokorrelationsfunktion. Diploma Thesis, Technische Universität München, 1986.

[17] Cygan, D.; Franz, J.; Söder, G.: Einfluß eines Filters auf nicht-gaußverteilte Zufallsprozesse. AEÜ **40** (1986), 377–384.

[18] Schwarz, H. R.: Numerische Mathematik. Stuttgart: Teubner, 1986.

[19] Franz, J.; Rapp, C.; Söder, G.: Influence of baseband filtering on laser phase noise in coherent optical transmission systems. J. Opt. Commun. **1** (1986), 15–20.

[20] Franz, J.: Optische Übertragungssysteme mit Überlagerungsempfang. Berlin: Springer, 1988.

[21] Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetteling, W. T.: Numerical Recipes. Cambridge: University Press, 1988.

[22] Fritzsche, C.: Einfluß des Laserrauschens auf die Empfangseigenschaften kohärenter optischer Nachrichtensysteme am Beispiel eines 140 Mbit/s FSK-Heterodynempfängers. Thesis, Universität Karlsruhe, 1987.