BAYESIAN ESTIMATION OF CLIPPED GAUSSIAN PROCESSES WITH APPLICATION TO OFDM

Paolo Banelli [1], *Geert Leus* [2*], and *Georgios B. Giannakis* [3†]
[1] University of Perugia, Dept. of Elec. and Inform. Engr., Perugia, Italy
[2] Katholieke Universiteit Leuven, Dept. of Elec. Engr., Leuven, Belgium
[3] University of Minnesota, Dept. of Elec. and Comp. Engr., Minneapolis, MN, USA

ABSTRACT

Several engineering applications ranging from control to communications have to deal with a clipped Gaussian process, observed in the presence of Additive White Gaussian Noise (AWGN). For such a scenario, we derive in this paper a closed form expression of a Bayesian estimator, which recovers the original undistorted Gaussian process by minimizing the mean square estimation error. In addition, we use the obtained closed form expression to show that the Bayesian estimator results in a Bit-Error Rate (BER) improvement compared to existing receivers for an Orthogonal Frequency Division Multiplexing (OFDM) system in an AWGN channel that is impaired by a clipping device at the transmitter.

1 INTRODUCTION

Estimating a Gaussian process $x(t)$ that passes through a clipping device $f(x)$, and is corrupted by Additive White Gaussian Noise (AWGN) $n(t)$, is a problem encountered in several engineering applications ranging from control to communications (see Fig. 1). The goal of this paper is to derive a closed form expression of a Bayesian estimator, which recovers the original undistorted Gaussian process by minimizing the mean square estimation error [1].

The paper focuses on the general derivation of the result, rather than applying it to a specific system. However, it is clear that the system model under consideration applies to frequency-flat communication systems, when the transmitted signal is a superposition of many independent signals, which thanks to the central limit theorem can be approximated by a Gaussian process, and when the D/A converters and amplifiers at the transmitter introduce a clipping of the transmitted signal. Examples of such communication systems are Orthogonal Frequency Division Multiplexing (OFDM) and downlink Code Division Multiple Access (CDMA), where the Probability Density Function (PDF) of the transmitted signal is well approximated by a Gaussian if the number of active carriers (for OFDM) or active users (for downlink CDMA) is sufficiently high.

* Postdoctoral Fellow of the FWO-Vlaanderen (Belgium).
† This work was supported in part by the NSF Grant 9979443.

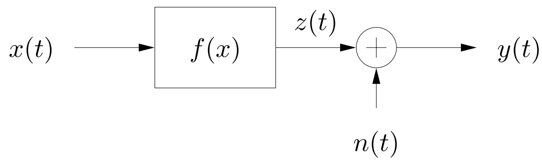

Figure 1: System model.

2 BAYESIAN ESTIMATOR

We consider the system depicted in Fig. 1. The output $y(t)$ of this system can be written as

$$y\left(t\right) = f\left[x(t)\right] + n(t) = z\left(t\right) + n(t), \tag{1}$$

where $x(t)$ is a zero-mean stationary Gaussian process, $f(x)$ models a clipping device:

$$f\left(x\right) = \begin{cases} x & \text{if } |x| < x_o \\ x_o \text{sign}(x) & \text{if } |x| > x_o \end{cases}, \tag{2}$$

$z(t)$ is the output of the clipping device, and $n(t)$ is zero-mean AWGN. We wish to estimate $x(t)$ from the knowledge of $y(t)$. The goal is to minimize both the nonlinear distortion introduced by $f(x)$, and the additive distortion introduced by $n(t)$.

Estimating $x(t)$ from the knowledge of $y(t)$ is a classical a posteriori estimation problem. From the Bayesian estimation theory [1] it is well known that the instantaneous optimum estimator, in the Minimum Mean-Square Error (MMSE) sense, of $x(t)$ given $y(t)$ is expressed by the conditional expectation of $x(t)$ given $y(t)$. In other words, for each t,

$$\begin{aligned}\hat{x}_{opt} = E\left\{x|y\right\} &= \int_{-\infty}^{+\infty} x p_{x|y}\left(x,y\right) dx \\ &= \frac{1}{p_y\left(y\right)} \int_{-\infty}^{+\infty} x p_{y|x}\left(x,y\right) p_x\left(x\right) dx,\end{aligned} \tag{3}$$

where $p_x(x)$ and $p_y(y)$ are the PDFs of $x(t)$ and $y(t)$, respectively; and $p_{y|x}(x,y)$ is the conditional PDF of

$y(t)$ given $x(t)$. The PDF of $x(t)$ can be written as

$$p_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{x^2}{2\sigma_x^2}}, \tag{4}$$

where σ_x^2 is the signal power. When $x(t)$ is clipped by $f(x)$ (see (2)), it is easy to show that the PDF of the output $z(t)$ can be written as

$$p_z(z) = p_x(z)\, u_{2x_o}(z) + A\delta(z + x_o) + A\delta(z - x_o), \tag{5}$$

where $\delta(z)$ is the Dirac impulse, $u_{2\Delta}(z)$ is the rectangular function given by

$$u_{2\Delta}(z) = \begin{cases} 1 & \text{if } |z| < \Delta \\ 0 & \text{if } |z| > \Delta \end{cases},$$

and A is the probability that x is larger than x_o:

$$A = \mathrm{Prob}\{x > x_o\} = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{x_o}{\sqrt{2\sigma_x^2}}\right)\right],$$

with $\mathrm{erf}(\cdot)$ being the error function. The PDF of $n(t)$ can be written as

$$p_n(n) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{n^2}{2\sigma_n^2}}, \tag{6}$$

where σ_n^2 is the noise power. Since we may assume that $x(t)$ is independent of $n(t)$, $y(t)$ is the sum of two independent random variables and its PDF is

$$p_y(y) = \int_{-\infty}^{+\infty} p_z(\tau)\, p_n(y - \tau)\, d\tau. \tag{7}$$

Substituting (5) and (6) into (7), we obtain

$$\begin{aligned} p_y(y) &= \frac{A}{\sqrt{2\pi\sigma_n^2}}\left[e^{-\frac{(y-x_o)^2}{2\sigma_n^2}} + e^{-\frac{(y+x_o)^2}{2\sigma_n^2}}\right] \\ &+ \frac{1}{2\sqrt{2\pi(\sigma_n^2+\sigma_x^2)}} e^{-\frac{y^2}{2(\sigma_n^2+\sigma_x^2)}} \left[\mathrm{erf}(\alpha_1) - \mathrm{erf}(\alpha_2)\right], \end{aligned} \tag{8}$$

where

$$\alpha_1 = \frac{\sigma_x^2(x_o - y) + \sigma_n^2 x_o}{\sigma_x\sigma_n\sqrt{2(\sigma_x^2+\sigma_n^2)}},\ \alpha_2 = \frac{-\sigma_x^2(x_o + y) - \sigma_n^2 x_o}{\sigma_x\sigma_n\sqrt{2(\sigma_x^2+\sigma_n^2)}}.$$

From (1), (2), and (6), it is also easy to check that

$$\begin{aligned} p_{y|x}(x,y) &= p_n(y - f(x)) \\ &= \frac{1}{\sqrt{2\pi\sigma_n^2}} \cdot \begin{cases} e^{-\frac{(y-x)^2}{2\sigma_n^2}} & \text{if } |x| < x_o \\ e^{-\frac{(y-x_o)^2}{2\sigma_n^2}} & \text{if } x > x_o \\ e^{-\frac{(y+x_o)^2}{2\sigma_n^2}} & \text{if } x < -x_o \end{cases}. \end{aligned} \tag{9}$$

Substituting (4), (8), and (9) into (3), results in the following closed form expression of the Bayesian estimator

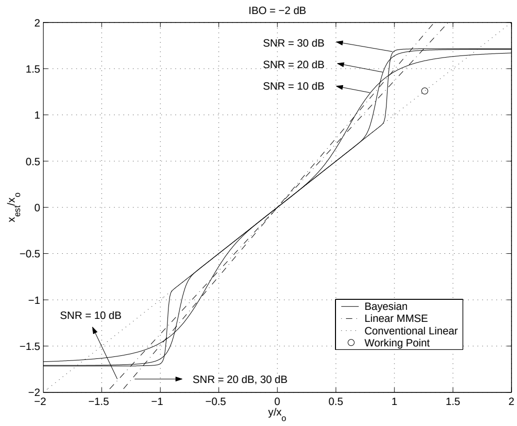

Figure 2: Bayesian vs. linear ($IBO = -2$ dB).

of $x(t)$ given $y(t)$:

$$\begin{aligned} \hat{x}_{opt} &= \frac{1}{p_y(y)}\left\{\frac{\sigma_x}{2\pi\sigma_n} e^{-\frac{x_o^2}{2\sigma_x^2}}\left(e^{-\frac{(y-x_o)^2}{2\sigma_n^2}} - e^{-\frac{(y+x_o)^2}{2\sigma_n^2}}\right)\right. \\ &+ \frac{\sigma_x^2}{2\sqrt{2\pi(\sigma_n^2+\sigma_x^2)}^3} y e^{-\frac{y^2}{2(\sigma_n^2+\sigma_x^2)}} \left[\mathrm{erf}(\alpha_1) - \mathrm{erf}(\alpha_2)\right] \\ &+ \left.\frac{\sigma_x\sigma_n}{2\pi(\sigma_n^2+\sigma_x^2)} e^{-\frac{y^2}{2(\sigma_n^2+\sigma_x^2)}}\left(e^{-\alpha_2^2} - e^{-\alpha_1^2}\right)\right\}. \end{aligned} \tag{10}$$

As expected, the above expression shows that the Bayesian estimator depends on the signal power σ_x^2, the noise power σ_n^2, and the clipping level x_o. Let the Input Back-Off (IBO) be the ratio between the input power corresponding to the clipping level x_o and the signal power σ_x^2. Notice that the IBO determines the working point of the system. And let the Signal-to-Noise Ratio (SNR) be the ratio between the power of $z(t)$ and the noise power σ_n^2. We can then express $\hat{x}_{opt}$ as a function of x_o^2, IBO, and SNR by making the following substitutions in (10):

$$\begin{aligned} \sigma_x^2 &= \frac{x_o^2}{IBO}, \\ \sigma_n^2 &= \frac{E\{z^2\}}{SNR} = \frac{\gamma\sigma_x^2}{SNR}. \end{aligned} \tag{11}$$

Notice that γ is the ratio between the power of $z(t)$ and the signal power σ_x^2. It only depends on the IBO. For the clipping function $f(x)$, γ can be written as [2]

$$\begin{aligned} \gamma = \frac{E\{z^2\}}{\sigma_x^2} &= 1 - e^{-IBO/2}\sqrt{2IBO/\pi} \\ &+ (IBO - 1)\,\mathrm{erfc}\left(\sqrt{IBO/2}\right), \end{aligned} \tag{12}$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function.

3 ALTERNATIVE ESTIMATORS

In this section, we review three alternative approaches to solving the same estimation problem. The first one

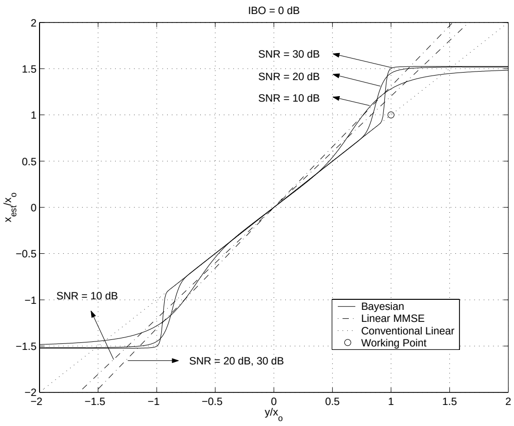

Figure 3: Bayesian vs. linear ($IBO = 0$ dB).

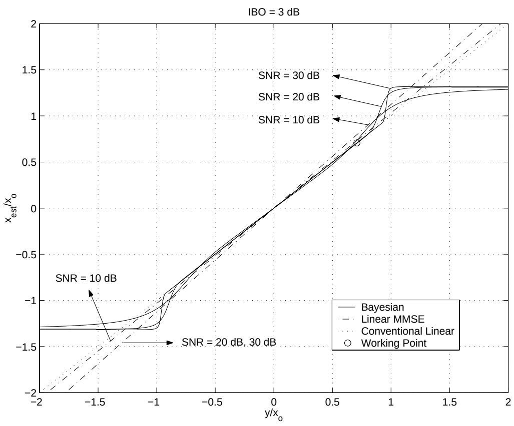

Figure 4: Bayesian vs. linear ($IBO = 3$ dB).

is the conventional linear estimator:

$$\hat{x}_{lin}^{conv} = y. \tag{13}$$

The second one is the linear MMSE estimator. It relies on the Bussgang theorem, which states that a zero-mean stationary Gaussian process distorted by a memoryless non-linearity can always be written as the sum of two uncorrelated stationary processes, one being a scaled replica of the undistorted zero-mean stationary Gaussian process [2]. Therefore, we can write $y(t)$ as

$$y(t) = z(t) + n(t) = \alpha x(t) + n_d(t) + n(t), \tag{14}$$

where $n_d(t)$, which we refer to as the non-linear distortion noise, is a stationary process uncorrelated with $x(t)$. The scale α only depends on the IBO. For the clipping function $f(x)$, α can be written as [2]

$$\alpha = \frac{E\left\{xf(x)\right\}}{\sigma_x^2} = \operatorname{erf}\left(\sqrt{IBO/2}\right). \tag{15}$$

Since $n_d(t)$ is uncorrelated with $x(t)$, its power is

$$E\{n_d^2\} = E\{z^2\} - \alpha^2\sigma_x^2 = (\gamma - \alpha^2)\sigma_x^2, \tag{16}$$

where α is given in (15) and γ is given in (12). Using (16) and (11), the linear MMSE estimator can now be expressed as

$$\begin{aligned} \hat{x}_{lin}^{MMSE} &= \frac{\alpha\sigma_x^2}{\alpha^2\sigma_x^2 + E\{n_d^2\} + \sigma_n^2} y \\ &= \frac{\alpha SNR}{\gamma(1 + SNR)} y. \end{aligned} \tag{17}$$

The third estimator is the cubic MMSE one:

$$\hat{x}_{cub}^{MMSE} = ay^3 + by, \tag{18}$$

where a and b are chosen to minimize the mean-square estimation error:

$$(a, b) = \min_{(a',b') \in \mathbb{R}^2} E\left\{|a'y^3 + b'y - x|^2\right\}.$$

Since it is difficult to derive a and b in closed form, we will instead compute them using a training sequence. Note that the Bayesian estimator in Section 2 can be viewed as an MMSE estimator based on an infinite order polynomial.

Figs. 2, 3, and 4 show the normalized (by x_o) input-output of the Bayesian estimator for several SNRs and IBOs, and compare it with the conventional linear and linear MMSE alternatives. To explain the particular shape of the Bayesian estimator, let us focus on the region of positive received signals $y(t)$. It is clear that when $y(t) \ll x_o$, it is with high probability equal to $x(t) + n(t)$, at least when $\sigma_n^2 \ll \sigma_x^2$. In this case, the conventional linear estimator is optimal. On the other hand, when $y(t) \gg x_o$, it is with high probability equal to $x_o + n(t)$, at least when $\sigma_n^2 \ll \sigma_x^2$. The optimal estimator now corresponds to a decision device. The saturation level of this decision device can be shown to be

$$\hat{x}_{sat} = \lim_{y \to \infty} \hat{x}_{opt}(y) = \frac{1}{A\sqrt{2\pi}} \frac{\sigma_x^3}{\sigma_x^2 + \sigma_n^2} e^{-\frac{x_o^2}{2\sigma_x^2}},$$

which clearly depends on the signal power σ_x^2, the noise power σ_n^2, and the clipping level x_o. The region in between is the region where the problem shows its non-linear nature, and the width of this region increases with decreasing SNR.

4 APPLICATION TO OFDM

In an OFDM system [3] [6], an N-point IFFT is applied on a block of N data symbols $\{a_m[n]\}_{n=0}^{N-1}$ drawn from a finite alphabet. The resulting block of N symbols $\{x_m[k]\}_{k=0}^{N-1}$ is then transmitted over the channel. Denoting the duration of a block by $T_b = NT_s$, the transmitted signal is

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{k=0}^{N-1} p(t - mT_b - kT_s) x_m[k], \tag{19}$$

where $p(t)$ represents the pulse shaper. In a baseband system such as xDSL, $x(t)$ is always real. This is obtained by imposing the condition of a conjugate symmetric spectrum on the block of N data symbols $\{a_m[n]\}_{n=0}^{N-1}$ [5]. In a Radio Frequency (RF) system, $x(t)$ is generally complex and its real and imaginary parts are modulated on respectively a cosine and a sine with frequency equal to the carrier frequency. Either way, if the data symbols are random, $x(t)$ (for baseband) or its I-Q components (for RF) can be viewed as zero-mean Gaussian processes if the number of subcarriers N is sufficiently high (let's say $N \geq 64$). Moreover, if $p(t)$ is rectangular or symmetric bandlimited, then $x(t)$ (for baseband) or its I-Q components (for RF) are also stationary. Hence, if a clipping device is used on $x(t)$ (for baseband) or its I-Q components (for RF) to limit the dynamic range, the closed forms presented in this paper are applicable. Suppose clipping is the only distortion introduced by the transmitter, and the channel is AWGN, i.e., the received OFDM signal is $y(t) = f[x(t)] + n(t)$, with $x(t)$ as in (19). The block of N data symbols $\{a_m[n]\}_{n=0}^{N-1}$ can then be estimated at the receiver by applying an N-point FFT on $\{\hat{x}_m[k]\}_{k=0}^{N-1}$, where $\hat{x}_m[k] = \hat{x}(t - mT_b - kT_s)$, with $\hat{x}(t)$ representing an estimate of $x(t)$, obtained as in (10), (13), (17), or (18).

Fig. 5 shows the BER performance of a baseband OFDM system for different estimators, and different IBO values. For simplicity, we use BPSK modulation. From the Bussgang theorem (see (14)), we see that clipping introduces two undesirable effects that degrade the performance of the linear MMSE estimator. The first is the power attenuation by a factor α^2, whereas the second is the non-linear distortion noise $n_d(t)$. The first effect is responsible for an SNR penalty, which manifests itself as a translation of the BER curve to the right for decreasing IBO values. The second effect causes a BER floor at high SNR, where the non-linear distortion noise $n_d(t)$ masks $n(t)$ (see also [4] for further details on this issue). It is clear that the effect of the power attenuation by a factor α^2 can not be reduced, because it results from the power loss that is introduced by clipping, which is an irreversible operation. On the other hand, the effect of the non-linear distortion noise $n_d(t)$ can be reduced by exploiting the a priori knowledge of its distribution or, equivalently, the distribution of the clipped signal, as our Bayesian approach does. Keeping these considerations in mind, it is easy to understand why our Bayesian estimator results in a BER performance improvement only at high SNR. If $n(t)$ is the dominant noise term, then $n_d(t)$ may be ignored and we basically obtain a linear scenario, where the useful signal $x(t)$ is attenuated by a factor α. Consequently, the performance of the linear MMSE estimator is very close to the optimum performance at low SNR. Note that for any PSK modulation, the conventional linear estimator has the same performance as the linear MMSE estimator, since the information is conveyed in the phase of the transmitted signal. For multi-amplitude modula-

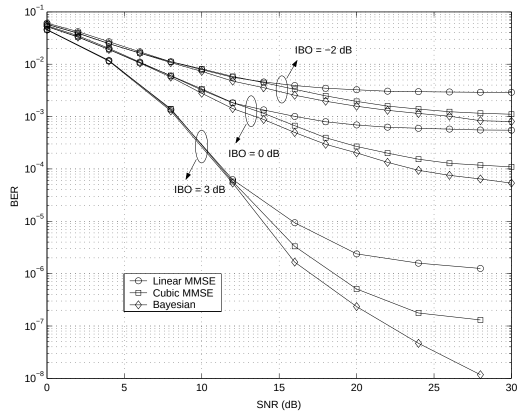

Figure 5: BER performance of a baseband BPSK-OFDM system.

tions, on the other hand, the linear MMSE estimator will always outperform the conventional linear estimator. Finally, we observe that the performance of the cubic MMSE fit lies between the performance of the linear MMSE and the Bayesian estimator.

5 CONCLUSIONS

We have derived in closed form the Bayesian estimator of a clipped Gaussian process in AWGN. The result can be applied to many engineering systems ranging from control to communications. As an illustration, we have used the obtained closed form expression to show that the Bayesian estimator results in a BER improvement compared to existing receivers for an OFDM system in an AWGN channel that is impaired by a clipping device at the transmitter. Future work includes the extension to frequency-selective channels that follow the clipping device.

REFERENCES

[1] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall Intl. Ed., 1993.

[2] A. Papoulis. *Probability, Random Variables and Stochastic Processes.* 3ed. McGraw-Hill Inc., 1991.

[3] J. A. C. Bingham. Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come. *IEEE Commun. Mag.*, pp. 5–14, May 1990.

[4] P. Banelli and S. Cacopardi. Theoretical Analysis and Performance of OFDM Signals in Nonlinear AWGN Channels. *IEEE Trans. on Commun.*, pp. 430–441, March 2000.

[5] J. M. Cioffi, V. Oksman, J.-J. Werner, T. Pollet, P. M. P. Spruyt, J. S. Chow, and K. S. Jacobsen. Very-High-Speed Digital Subscriber Lines. *IEEE Commun. Mag.*, pp. 72–79, April 1999.

[6] Z. Wang and G. B. Giannakis. Wireless Multicarrier Communications. *IEEE Signal Proc. Mag.*, pp. 29–48, May 2000.