A BAYESIAN METHOD FOR OSCILLATOR CHARACTERIZATION

F. Vernotte
Observatoire de Besançon
41 bis av. de l'Observatoire, BP 1615, F-25010 Besançon, France
tel +33-3-81.66.69.22 fax +33-3-81.66.69.44
e-mail francois@obs-besancon.fr

G. Zalamansky
Université de Metz
Campus Bridoux, F-57000 Metz, France

Abstract

The power spectral density of frequency fluctuations of an oscillator is generally modeled as a sum of power laws with integer exponents (from -2 to +2). However, a power law with a fractional exponent may exist. We propose a method for measuring such a noise level and determining the probability density of the exponent. This yields a criterion for compatibility with an integer exponent. This method is based upon a Bayesian approach called the reference analysis of Bernardo-Berger. The application to a sequence of frequency measurement from a quartz oscillator illustrates this paper.

INTRODUCTION

It is commonly assumed that $S_y(f)$, the power spectral density (PSD) of frequency deviation of an oscillator, may be modeled as the sum of 5 power laws, defining 5 types of noise:

$$S_y(f) = \sum_{\alpha=-2}^{+2} h_\alpha f^\alpha \tag{1}$$

where h_α is the level of the f^α noise. But it may be noticed that models with non-integer exponents are occasionally used.

The estimation of the noise levels is mainly achieved by using the Allan variance [1], which is defined versus the integration time τ as:

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle \left(\overline{y}_{k+1} - \overline{y}_k\right)^2 \right\rangle. \tag{2}$$

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
DEC 1998		**00-00-1998 to 00-00-1998**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Bayesian Method for Oscillator Characterization	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Observatoire de Besanson,41 bis av. de l'observatoire,BP 1615, F-25010 Besanson, France,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADA415578. 30th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, VA, 1-3 Dec 1998

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**9**	

In the frequency domain, the Allan variance may be considered as a filter. If the Allan variance versus the integration time τ is plotted, the graph exhibits different slopes, each slope corresponding to a type of noise:

$$\sigma_y^2(\tau) = C_\mu \tau^\mu \qquad \Leftrightarrow \qquad S_y(f) = h_\alpha f^\alpha \qquad \text{and} \qquad \alpha = -\mu - 1. \tag{3}$$

The estimation of C_μ yields an estimation of the noise level h_α.

However, this curve may exhibit an exponent μ which seems to be non-integer. Does this mean that the corresponding PSD is not compatible with the 5 power law model? In this paper, we propose a method for estimating the most probable value of this exponent in order to solve this ambiguity. This method is applied to an example of stability measurement.

CLASSICAL STABILITY ANALYSIS OF AN OSCILLATOR

Sequence of frequency measures

Figure 1 shows average frequency measures $\overline{\nu}_k$ of a 10 MHz quartz oscillator compared to a cesium clock. The sampling rate is 10 s and the integration time of each frequency measure is also 10 s (sampling without dead time).

In order to obtain dimensionless $\overline{y}_k$ samples, we must subtract the nominal frequency ν_0 (10 MHz) from the frequency measures and normalize by ν_0:

$$\overline{y}_k = \frac{\overline{\nu}_k - \nu_0}{\nu_0}. \tag{4}$$

Variance analysis

Figure 2 is a log-log plot of the Allan deviation of the quartz $\overline{y}_k$ samples versus the integration time τ. A least squares fit of these variance measures (solid line), weighted by their uncertainties, detects only two types of noise: a white noise and an f^{-2} noise. The corresponding noise level estimations are:

$$\begin{aligned} h_0 &= (2.2 \pm 0.4) \cdot 10^{-5}\text{s} && \text{at } 1\sigma \text{ (68\% confidence)} \\ h_{-2} &= (2.3 \pm 0.6) \cdot 10^{-12}\text{s}^{-1} && \text{at } 1\sigma \text{ (68\% confidence)} \end{aligned}$$

(for the assessment of the h_α noise levels and their uncertainties, we used the multi-variance method described in [2]).

However, for large τ values (corresponding to low frequencies), the variance measures move away from the fitted curve. Two explanations are possible:

- instead of an f^{-2} noise, there is a noise whose non-integer exponent is contained between -2 and -3 ;
- since the uncertainty domains of the variance measures contain the fitted curve, this apparent divergence may be due to a statistical effect.

In order to choose between these two explanations, we decided to estimate the probability density of the exponent with a Bayesian approach.

BAYESIAN APPROACH

Principle

The goal of all measurement is the estimation of an unknown quantity θ from measures ξ, i.e. determining $p(\theta|\xi)$, the density of probability of the quantity θ knowing the measures ξ. The Bayesian theory is based on the following equality [3]:

$$p(\theta|\xi) \propto p(\xi|\theta)\pi(\theta) \tag{5}$$

where $p(\xi|\theta)$ is the distribution of the measures ξ for a fixed value of the quantity θ and $\pi(\theta)$ is the *a priori* density of probability of the quantity θ, i.e. before performing any measurement.

The determination of this a priori density, called the *prior*, is generally one of the main difficulties of this approach (particularly in the case of total lack of knowledge!). In this paper, we use the Jeffrey's prior which ensures properties such as invariance [3].

Spectral density and covariance matrix

Let us define the vector y whose components are the N $\overline{y}_k$ samples. We assume that y is a Gaussian vector. The probability distribution of y is:

$$p(y) = \frac{\exp\left(-\frac{y^t C^{-1} y}{2}\right)}{(2\pi)^{\frac{N}{2}}\sqrt{|C|}} \tag{6}$$

where C is the covariance matrix. Since $S_y(f)$ is the Fourier transform of $\mathcal{R}_y(\tau)$, the autocorrelation function of the frequency deviation, the general term of C is:

$$C_{ij} = 2\int S_y(f)\cos\left(2\pi f(t_i - t_j)\right) df. \tag{7}$$

Equation (7) reveals the key role played by the spectral density of the noise in the expected fluctuation. We will present a general method for estimating the parameters of the model for $S_y(f)$.

Assumed model for the spectral density

We assume that the sequence of frequency measures is composed of a white noise y_w, whose variance (i.e. the level) is unitary, and of a red noise y_r whose level is unknown, multiplied by the real standard deviation of the white noise σ_w (σ_w is easily estimated from high sampling rate frequency measurement):

$$y = (y_w + y_r)\sigma_w. \tag{8}$$

This yields the following model for $S_y(f)$:

$$S_y(f) = h_0 + h_\alpha \cdot f^\alpha \qquad \text{with } -3 \leq \alpha \leq -1 \tag{9}$$

where $h_0 = 2.2 \cdot 10^{-5}$s, h_α and α are the unknown parameters.

Let us denote y_n, the normalized vector:

$$y_n = y_w + y_r. \tag{10}$$

The corresponding normalized PSD $S_n(f)$ is:

$$S_n(f) = 1 + H \cdot u_\alpha \cdot f^\alpha \tag{11}$$

where u_α is an amplitude factor whose meaning will be explained below (see equation (23)).

Statistical model

The part of the spectral density due to the red noise y_r may be written:

$$S_r(f) = H \cdot u_\alpha \cdot f^\alpha. \tag{12}$$

We used the Bernardo-Berger analysis [3, 4] for estimating the unknown parameter $\theta = (\alpha, H)$.

- ***Construction of the estimators:***

Let us introduce the orthonormal basis of $\Re^N$, $\{p_0, \ldots, p_j, \ldots, p_{N-1}\}$, defined such as the i^{th} component of p_j is:

$$p_{ij} = \widetilde{p}_j(t_i) \tag{13}$$

where t_i is the date of the i^{th} frequency measure and $\widetilde{p}_j(t)$ is a polynomial of degree j, satisfying the orthonormality condition [5]:

$$\sum_{i=0}^{N-1} \widetilde{p}_j(t_i) \cdot \widetilde{p}_k(t_i) = \delta_{jk}. \tag{14}$$

It can be shown than the scalar product of a vector p_j by the noise vector y is an estimate of the noise spectrum for a given frequency f_j [6]. Let us denote ξ_j such an estimate applied to the normalized noise:

$$\xi_j = p_j \cdot y_n. \tag{15}$$

Practically, we limited to 16 the number of estimators p_j (from degrees 0 to 15) for limiting the computation and because the high degrees, estimating the high frequencies, are less informative for a red noise.

Moreover, in order to ensure convergence for very low frequencies (even if the low cut-off frequency tends towards 0), the polynomials must satisfy the moment condition [5, 6]: the minimum degree j_{min} of a polynomial to ensure convergence up to an exponent α is:

$$j_{min} \geq \frac{-\alpha - 1}{2}. \tag{16}$$

Since we have assumed $\alpha \leq -3$, the first 2 estimators (p_0 and p_1) must be removed. Thus we have $n = 14$ estimators $\{p_2, \ldots, p_{15}\}$ and $n = 14$ estimates $\{\xi_2, \ldots, \xi_{15}\}$.

- ***Construction of the priors:***

The covariance matrix defined in relationship (7) is an ensemble average of the different estimate products over an infinite number of realizations of this process:

$$C = \langle \xi \cdot \xi^t \rangle \tag{17}$$

$$C_{ij} = \langle \xi_i \cdot \xi_j \rangle. \tag{18}$$

As for the noise vector y_n, the estimate vector ξ may be split into two terms, according to equations (10) and (15):

$$\xi = \xi_w + \xi_r. \tag{19}$$

The covariance matrix may also be split:

$$C = \langle \xi_w \xi_w^t \rangle + \langle \xi_r \xi_r^t \rangle = I_n + H \cdot u_\alpha \cdot V(\alpha) \tag{20}$$

where I_n is the identity matrix in $\Re^n$ and the general term of the matrix $V(\alpha)$ is:

$$[V(\alpha)]_{ij} = 2 \int_{1/T}^{f_h} f^\alpha \cos\left(2\pi f(t_i - t_j)\right) df. \tag{21}$$

The high cut-off frequency f_h in (21) is the Nyquist frequency and T is the total duration of the sequence.

Let $e_i(\alpha)$ denote the i^{th} eigenvector of $V(\alpha)$ and $\gamma_i(\alpha)$ its i^{th} eigenvalue ($i \in \{0, \ldots, n-1\}$). The averaged quadratic norm of the estimate vector ξ is:

$$\left\langle \|\xi\|^2 \right\rangle = n + H \cdot u_\alpha \sum_{i=0}^{n-1} \gamma_i(\alpha) = \left\langle \|\xi_w\|^2 \right\rangle + \left\langle \|\xi_r\|^2 \right\rangle. \tag{22}$$

The factor u_α is chosen in such a way that, for $H = 1$, the averaged quadratic norms $\left\langle \|\xi_w\|^2 \right\rangle$ and $\left\langle \|\xi_r\|^2 \right\rangle$ are equal:

$$u_\alpha = \frac{n}{\sum_{i=0}^{n-1} \gamma_i(\alpha)}. \tag{23}$$

The direct problem is now solved since ξ is a vector of $\Re^n$ with a probability distribution given the parameter θ equal to:

$$p(\xi|\theta) = \frac{1}{(2\pi)^{n/2}\sqrt{|C|}} \exp(-\frac{1}{2}\xi^t C^{-1}\xi). \tag{24}$$

Denoting "$\mathrm{Tr}(M)$" the trace of a matrix M and X the matrix defined as:

$$X = u_\alpha \cdot V(\alpha) \tag{25}$$

the Fisher information matrix $I(\theta)$ is (see [4]):

$$I(\theta) = \frac{1}{2} \begin{pmatrix} H^2 \mathrm{Tr}\left(\mathrm{C}^{-1}\frac{\mathrm{dX}}{\mathrm{d}\alpha}\mathrm{C}^{-1}\frac{\mathrm{dX}}{\mathrm{d}\alpha}\right) & H\mathrm{Tr}\left(\mathrm{C}^{-1}\mathrm{X}\mathrm{C}^{-1}\frac{\mathrm{dX}}{\mathrm{d}\alpha}\right) \\ H\mathrm{Tr}\left(\mathrm{C}^{-1}\mathrm{X}\mathrm{C}^{-1}\frac{\mathrm{dX}}{\mathrm{d}\alpha}\right) & \mathrm{Tr}\left(\mathrm{C}^{-1}\mathrm{X}\mathrm{C}^{-1}\mathrm{X}\right) \end{pmatrix}. \tag{26}$$

The Jeffrey's prior $\pi(\theta)$ is defined as:

$$\pi(\theta) = \sqrt{|I(\theta)|}. \tag{27}$$

The parameter θ is a two-dimensional parameter composed of the exponent parameter α and of the amplitude parameter H. Since we are mostly interested in α, H is called a nuisance parameter.

In presence of nuisance parameter, Bernardo and Berger suggested that α should first be fixed and the conditional prior $\pi(H|\alpha)$ computed for that value. The full prior is then:

$$\pi(\theta) = \pi(H|\alpha) \cdot \pi(\alpha). \tag{28}$$

The conditional prior $\pi(H|\alpha)$ is given by:

$$\pi(H|\alpha) = \sqrt{|[I(\theta)]_{22}|} \tag{29}$$

where $[I(\theta)]_{11}$, $[I(\theta)]_{12} = [I(\theta)]_{21}$ and $[I(\theta)]_{22}$ are the elements of the Fisher information matrix $I(\theta)$.

The prior for α may be computed as:

$$\pi(\alpha) = c \cdot \exp\left(\int \pi(H|\alpha) \ln |k(\alpha, H)|^{1/2} \, d\lambda\right) \tag{30}$$

where c is a normalization coefficient ensuring that $\int \pi(\alpha) d\alpha = 1$ and:

$$k(\alpha, H) = [I(\theta)]_{11} - \frac{[I(\theta)]_{12}^2}{[I(\theta)]_{11}}. \tag{31}$$

This prior is plotted in Figure 3.

- ***Construction of the posteriors:***

According to the Bayes theorem, the posterior probability distribution is given by:

$$p(\theta|\xi) = \frac{p(\xi|\theta)\pi(\theta)}{\int p(\xi|\theta')\pi(\theta')d\theta'}. \tag{32}$$

The posterior probability distribution for α is then given by:

$$p(\alpha|\xi) = \frac{\int p(\xi|\alpha, H)\pi(H|\alpha)\pi(\alpha)dH}{\int\int p(\xi|\alpha', H')\pi(H'|\alpha')\pi(\alpha')dH'd\alpha'}. \tag{33}$$

RESULTS AND DISCUSSION

Compatibility with an integer exponent

Figure 4 shows the posterior probability distribution for the exponent α of the red noise using the Bernardo-Berger prior.

The exponent value obtained for the maximum of likelihood, just as for the maximum of the distribution, is $\alpha = -2.2$.

However, $\alpha = -2$ is fully compatible with this prior distribution. Thus, we may conclude that the apparent divergence between the variance measures and the fitted curve in Figure 2 is probably due to a statistical bias of the data. The spectral density $S_y(f)$ is then compatible with the following model:

$$S_y(f) = h_0 + h_{-2}f^{-2}. \tag{34}$$

Noise level estimation

Selecting an exponent value $\alpha = -2$, we obtained the posterior probability distribution plotted in Figure 5. As in the variance analysis, we chose a confidence interval of 68% (16% probability that h_{-2} is smaller than the low bound and 16% probability that h_{-2} is greater than the high bound):

$$h_{-2} = \left(2.3 \begin{array}{c} + \\ - \end{array} \begin{array}{c} 2 \\ 0.8 \end{array}\right) \cdot 10^{-12}\mathrm{s}^{-1} \qquad \text{at } 1\sigma \text{ (68\% confidence)}$$

The difference between the maximum likelihood value ($h_{-2} = 2.2991 \cdot 10^{-12}\mathrm{s}^{-1}$) and the variance analysis value ($h_{-2} = 2.2949 \cdot 10^{-12}\mathrm{s}^{-1}$) is only 0.18%.

However, the confidence intervals given by these two methods are quiet different. The main difference concerns the symmetry of the variance analysis interval: in this case, we don't take into account the fact that the noise levels are positive, whereas the prior of the Bayesian approach is null for negative values of h_{-2}.

Moreover, the variance analysis interval seems to be a bit underestimated.

CONCLUSION

The variance analysis is an useful tool for a quick estimation of the noise levels in the output signal of an oscillator. However, a negative estimate of a noise level may occur. Generally, in this case, this value is rejected and the corresponding noise level is assumed to be null. On the other hand, although the Bayesian method is a bit heavier, it takes into account properly the a priori information, and gives a more reliable estimation of these noise levels and especially of their confidence intervals.

However, the main advantage of the Bayesian method concerns the verification of the validity of the power law model of spectral density. Each time the model is suspected, such an approach should be used in order to estimate the exponent of the power law. In particular, this method should be very interesting for the study of the f^{-1} and f^{+1} noise, whose origin remains mysterious [7].

REFERENCES

[1] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leesson, T. E. McCuningal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. Vessot, and G. M. R. Winkler, *"Characterization of frequency stability,"* **IEEE Transactions on Instrumentation and Measurement, IM-20**, 105–120.

[2] F. Vernotte, E. Lantz, J. Groslambert, and J. Gagnepain, *"Oscillator noise analysis: multivariance measurement,"* **IEEE Trans. on Inst. and Meas., IM-42**, 342–350.

[3] J. Bernardo and A. Smith, **Bayesian theory**, Wiley & Son, 1993.

[4] G. Zalamansky and C. Robert, *"Bayesian analysis with application to the timing residuals of a pulsar,"* in Proceedings of MaxEnt'98, July 1998, Garching, Germany.

[5] J. E. Deeter and P. E. Boynton, *"Techniques for the estimation of red power spectra. I. context and methodology,"* **The Astrophysical Journal, 261**, 337–350.

[6] F. Vernotte, G. Zalamansky, M. McHugh, and E. Lantz, *"Cut-off frequencies and noise power law model of spectral density: adaptation of the multi-variance method for irregularly spaced timing data using the lowest mode estimator approach.,"* **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 43**, (3), 403–409.

[7] M. Planat, V. Giordano, G. Marianneau, F. Vernotte, M. Mourey, C. Eckert, and J. Miehe, *"Is the frequency noise of an oscillator of deterministic origin ?,"* **IEEE Trans. on U. F. F. C. , 43**, (2), 326–330.

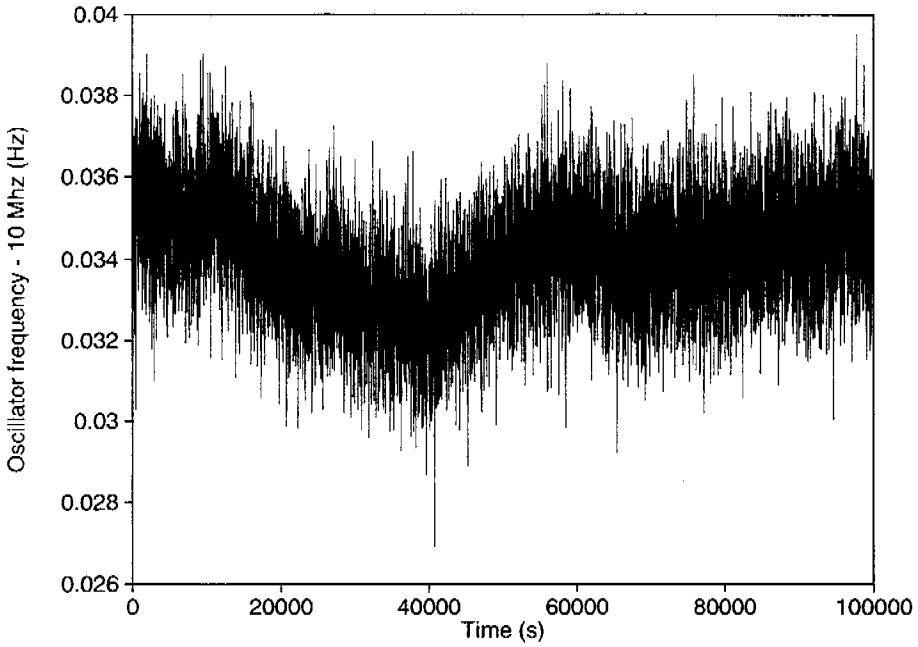

Figure 1: Sequence of frequency measures

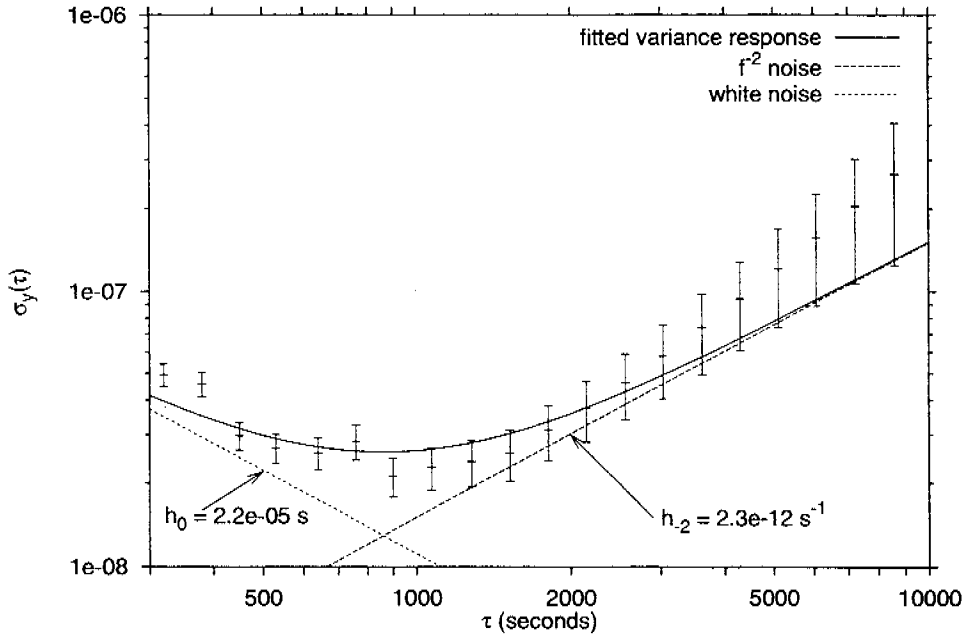

Figure 2: Allan deviation of the sequence of frequency measures

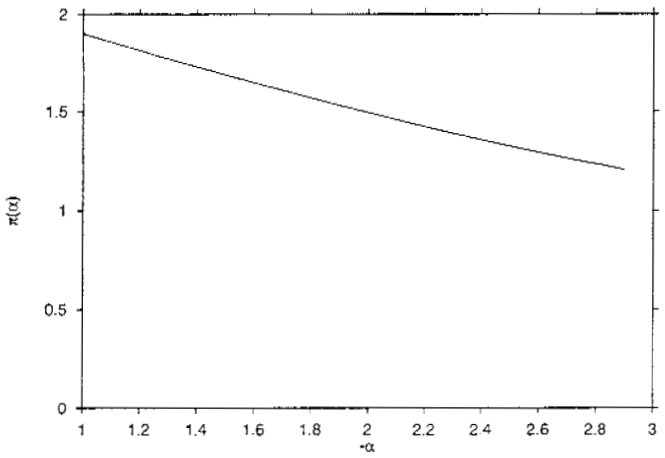

Figure 3: Reference prior for the power α

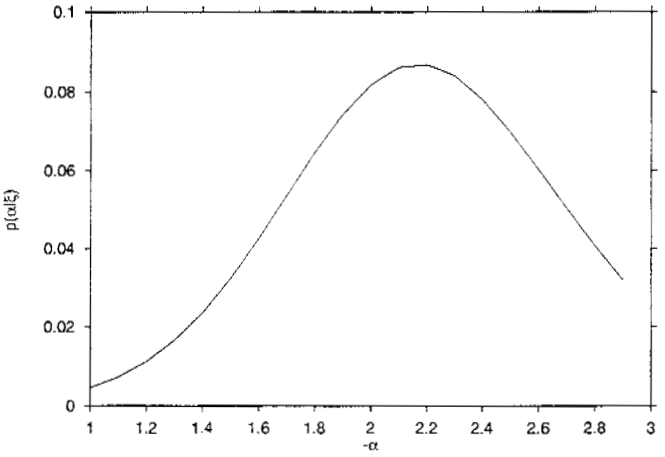

Figure 4: Posterior probability density for the power α

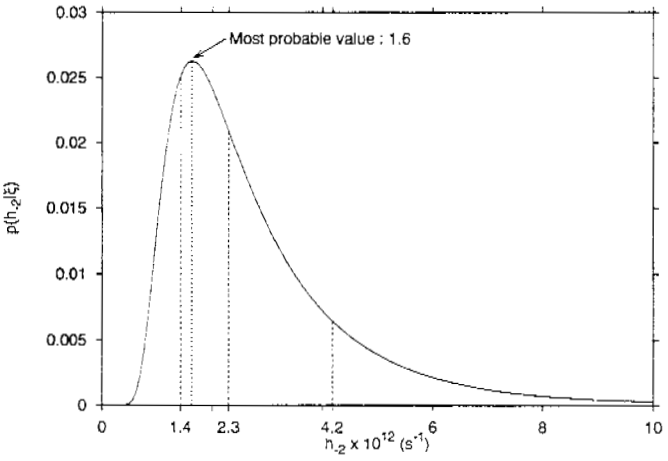

Figure 5: Posterior probability density for the noise level h_{-2}