SECOND-ORDER BASED CYCLIC FREQUENCY ESTIMATES: THE CASE OF DIGITAL COMMUNICATION SIGNALS

Pierre Jallon, Antoine Chevreuil

Equipe signal et communications
Institut Gaspard Monge - UMR 8049
Université de Marne la Vallée
5, boulevard Descartes
77454 Marne La Vallée Cedex 2

ABSTRACT

Both the standard estimate [1] of the second-order cyclic frequency of a digital communication signal and its improved versions (based on a weighted criterion [2] [3][4] or a denoising of the cyclic periodogram [5]) do not take into account a key property of the signal. Indeed, the support of the cyclo-spectrum at the true cyclic frequency is a narrow interval centered around half this cyclic frequency. This fact is taken into account in this contribution. A theoretical fact explains the improvement over the standard method. Simulations confirm the expectation.

1. INTRODUCTION

The estimation of the second-order cyclic frequencies of a signal is an old signal processing problem involved in various applications. In particular, in a digital communication context, the knowledge of the cyclic frequencies is essential prior to synchronization, equalization [6], blind source separation [7, 8]. Both military (passive listening with automatic classification of modulations) and cooperative (blind synchronization of high-speed distributed networks) applications are concerned.

In the context of linear modulations, the complex envelope of the signal received may be written as

$$x_a(t) = \sum_{n\in\mathbb{Z}} s_n c_a(t - nT_s) + b_a(t) \qquad (1)$$

where T_s is the symbol period, $(s_n)_{n\in\mathbb{Z}}$ a zero-mean, white i.i.d. sequence of symbols to be transmitted, c_a the unknown impulse response of the composite filter resulting from the shaping and the effect of a linear time-invariant channel; the transmit filter concentrate the energy of the signal in a band of frequencies. Specifically, the Fourier transform $\hat{c}_a$ of c_a has a support included in $\left[-\frac{1+\gamma}{2T_s}\ ,\ \frac{1+\gamma}{2T_s}\right]$, where γ is the bandwith excess and is assumed to be lower than 1. Lastly, in (1), $b_a(t)$ is an additive stationary noise. In the literature, almost every demodulation scheme (blind or not) implicitly assume known (or at least correctly estimated) the symbol period T_s in order to sample x_a at the rate $1/T_s$ (or a multiple). In order to achieve the estimation of T_s, it has been noticed by Gardner [9] that x_a is cyclostationary. Focussing in particular on the auto-correlation function $R_{x_a}(t,\tau) \stackrel{\triangle}{=} \mathbb{E}x_a(t+\tau)x_a(t)^*$, for $t, \tau \in \mathbb{R}$, it can be shown that

$$R_{x_a}(t,\tau) = \sum_{k=-1,0,1} R_{x_a}^{(k/T_s)}(\tau) e^{i2\pi kt/T_s}. \qquad (2)$$

The fact that the expansion (2) shows only 3 terms is due to the bandwidth limitation of signal x_a. As a consequence, the estimation of T_s reduces, for the second-order point of view, to the estimation of the periodicity of the functions $t \mapsto R_{x_a}(t,\tau)$. This problematic can easily be shifted to a discrete-time context. Indeed, let us choose T_e any sampling period, and set $x(n) \stackrel{\triangle}{=} x_a(nT_e)$. Hence $R_x(n,\delta) \stackrel{\triangle}{=} \mathbb{E}x(n+\delta)x(n)^*$ is an almost periodic function, since $R_x(n,\delta) = R_{x_a}(nT_e, \delta T_e)$. Its cyclic frequencies are $0, \pm\alpha_0$ with $\alpha_0 \stackrel{\triangle}{=} T_e/T_s$. In the sequel, we assume that T_e is neither equal to T_s nor to a multiple of T_s, so that $n \mapsto R_x(n,\delta)$ is effectively almost periodic (not constant) and can be written

$$R_x(n,\delta) = \sum_{k=-1,0,1} R_x^{(k\alpha_0)}(\delta) e^{i2\pi nk\alpha_0}$$

with $R_x^{(k\alpha_0)}(\delta) = R_{x_a}^{(k/T_s)}(\delta T_e)$. We may introduce the function

$$R_x^{(\alpha)}(\delta) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}x(n+\delta)x(n)^* e^{-2i\pi n\alpha}, \qquad (3)$$

which has the property that for any δ, $R_x^{(\alpha)}(\delta)$ vanishes if $\alpha \notin \{0, \pm\alpha_0\}$; conversely, if $\alpha = \pm\alpha_0$, there are indices δ such that $R_x^{(\alpha)}(\delta) \neq 0$. This remark justifies the definition of the function

$$J(\alpha) \stackrel{\triangle}{=} \sum_{\tau=-M}^{M} \left| R_x^{(\alpha)}(\tau) \right|^2 \qquad (4)$$

which is a strictly positive function that vanishes if $\alpha \neq 0$ and $\alpha \neq \pm\alpha_0$ and hence allows one to know α_0, hence T_s. In practice, $J(\alpha)$ is estimated as

$$\hat{J}(\alpha) = \sum_{\delta=-M}^{M} \left|\hat{R}_x^{(\alpha)}(\delta)\right|^2 \tag{5}$$

where $\hat{R}_x^{(\alpha)}(\delta)$ is the unbiased empirical estimate of $R_x^{(\alpha)}(\delta)$ defined as

$$\hat{R}_x^{(\alpha)}(\delta) = \frac{1}{N}\sum_{n=0}^{N-1} x(n+\delta)x(n)^* e^{-2i\pi n\alpha}$$

with N the number of available data. It is a well-known fact (see [10][3]) that $\hat{R}_x^{(\alpha)}(\delta)$ is a consistent estimate of $R_x^{(\alpha)}(\delta)$. As a consequence, if[1] the condition $T_e < T_s/2$ holds, the standard estimate of α_0 is defined as $\hat{\alpha}_0 = \text{Argmax}_{\alpha\in(0,\frac{1}{2})}\hat{J}(\alpha)$.

Such a definition allows one to see the estimation of α_0 as the estimation of the frequency of a sinusoide in a cyclo-stationary noise [11]; in this latter reference, it is proven that $\hat{\alpha}_0$ is a consistent estimate of α_0, hence $N^{3/2}(\hat{\alpha}_0 - \alpha_0)$ converges to a zero-mean Gaussian random variable. The rate of convergence ($N^{3/2}$) makes the algorithm efficient in a local domain around the true parameter α_0. However, it has long been noticed that, for short data, the function $\hat{J}(\alpha)$ is likely to be above $\hat{J}(\alpha_0)$, hence making erroneous the estimate $\hat{\alpha}_0$. In short, this is due to the fact $J(\alpha_0)$ tends to 0 if $\gamma \to 0$. Hence for small excess bandwidth factors γ, the value of $J(\alpha_0)$ is small.

This was first reported by Gini *et al.* [2]. The authors remedy the problem in considering a weighted version of $\hat{J}(\alpha)$; Mazet *et al.* [4], in this direction, prove that taking into account a weighting independent of the (unknown) α_0, thus suboptimal, does not impact the performance. In short, the weighting makes the (asymptotic) mean of $\hat{J}(\alpha), \alpha \notin \{0\,,\, \pm\alpha_0\}$ be constant, which drastically improves the performance. Though theoretically appealing, the approach is computationally involved and suffers from serious numerical problems due to the calculation of the weighting matrix. In an other direction, Touati et al. [5] adopt another point of view in order to improve the performance of the estimate $\hat{\alpha}_0$. As

$$R_x^{(\alpha)}(\delta) = \int_0^1 S_x^{(\alpha)}(e^{i2\pi\nu})e^{i2\pi\delta\nu}d\nu,$$

where the cyclo-spectrum $S_x^{(\alpha)}$ is defined as

$$S_x^{(\alpha)}(e^{i2\pi\nu}) \stackrel{\triangle}{=} \sum_{\delta} R_x^{(\alpha)}(\delta)e^{i2\pi\nu\delta},$$

an idea consists in providing an efficient estimate of $S_x^{(\alpha)}$ for every α, i.e. to apply denoising techniques of the standard cyclo-periodogram.

[1]This is, in practice not restrictive, since a raw spectral analysis of the received signal may be performed prior to the choice of the sampling period. This condition holds along the paper

However, in all these approaches, a crucial fact is not taken into account. Namely, the cyclo-spectrum at the unknown cyclic frequency α_0, can be written, since condition $T_e < T_s/2$ is assumed to hold: $\forall\nu \in [-1/2\,,\, 1/2)$,

$$S_x^{(\alpha_0)}(e^{i2\pi\nu}) = \frac{1}{T_sT_e}\hat{c}_a\left(\frac{\nu}{T_e}\right)\hat{c}_a\left(\frac{\nu}{T_e} - \frac{1}{T_s}\right)^*$$

hence the support of $S_x^{(\alpha_0)}$ is included in $\mathcal{B}_{\alpha_0,\gamma}$ with

$$\mathcal{B}_{\alpha_0,\gamma} \stackrel{\triangle}{=} \left[(1-\gamma)\frac{\alpha_0}{2}\,,\, (1+\gamma)\frac{\alpha_0}{2}\right]$$

(see [9] for instance). We present a new estimate of α_0 relying on this fact. In this respect, we consider the new statistic

$$\hat{J}_f(\alpha) \stackrel{\triangle}{=} \sum_{\delta=-M}^{M} \left|\hat{R}^{(\alpha)}_{[f(z)]x(n)}(\delta)\right|^2 \tag{6}$$

where $f(z)$ is a certain filter depending on the parameter α. Of course $\hat{J}_f(\alpha)$ is merely the natural (consistent) estimate of $J_f(\alpha) \stackrel{\triangle}{=} \sum_{\delta=-M}^{M}\left|R^{(\alpha)}_{[f(z)]x(n)}(\delta)\right|^2$, which shares with $J(\alpha)$ the key property that $J_f(\alpha)$ vanishes if $\alpha \neq 0$ or $\alpha \neq \pm\alpha_0$. A set of filters $f(z)$ is set forth in Section 2; a theoretical argument proves that the new estimate based on the maximization of (6) for this class of filters is expected to be better than $\hat{\alpha}_0$. Section 3 focuses on numerical aspects; simulations show the clear improvement brought by the approach. In particular, the estimate has a high level of confidence in quite difficult contexts (short data, small excess bandwidth factor).

2. THEORETICAL APPROACH

We propose to work out the basic properties of the random variable $\hat{J}_f(\alpha)$ for a certain filter $f(z)$. In this respect, we may set

$$y(n) \stackrel{\triangle}{=} [f(z)]x(n).$$

2.1. Analysis for $\alpha = \alpha_0$

The cyclo-spectrum of $y(n)$ is directly expressed as

$$S_y^{(\alpha_0)}(e^{i2\pi\nu}) = f(e^{i2\pi\nu})f(e^{i2\pi(\nu-\alpha_0)})^* S_x^{(\alpha_0)}(e^{i2\pi\nu}). \tag{7}$$

Now the support of $S_x^{(\alpha_0)}$ is $\mathcal{B}_{\alpha_0,\gamma}$, hence $J_f(\alpha_0)$ expresses as

$$\sum_{\delta=-M}^{M}\left|\int_{\mathcal{B}_{\alpha_0,\gamma}} f(e^{i2\pi\nu})f(e^{i2\pi(\nu-\alpha_0)})^* S_x^{(\alpha_0)}(e^{i2\pi\nu})e^{i2\pi\delta\nu}d\nu\right|^2.$$

As a consequence

$$J_f(\alpha_0) = J(\alpha_0). \tag{8}$$

when $f(z)$ is a filter such that $f(e^{i2\pi\nu}) = 1$ for all $\nu \in \mathcal{B}_{\alpha_0,\gamma}\bigcup(\mathcal{B}_{-\alpha_0,\gamma})$.

2.2. Analysis for $\alpha \neq \alpha_0$

In this case $\hat{J}_f(\alpha)$ tends to zero in probability. More precisely, we may consider the vector $\hat{\mathbf{R}}_y^{(\alpha)} \triangleq [\hat{R}_y^{(\alpha)}(-M), \cdots, \hat{R}_y^{(\alpha)}(M)]^T$. Standard statistical tools show that $\sqrt{N}\left(\hat{\mathbf{R}}_y^{(\alpha)}\right)$ converges to a zero-mean Gaussian vector with covariance $\Gamma_y(\alpha)$ the expression of which may be found in [11]. As $\hat{J}_f(\alpha) = \hat{\mathbf{R}}_y^{(\alpha)H}\hat{\mathbf{R}}_y^{(\alpha)}$, we can deduce that

$$\phi_f(\alpha) \triangleq \lim_{N\to\infty} N\mathbb{E}\hat{J}_f(\alpha) = \phi_f^{(1)}(\alpha) + \phi_f^{(2)}(\alpha) \qquad (9)$$

with

$$\begin{aligned}
\phi_f^{(1)}(\alpha) &= (2M+1)\int_0^1 S_y^{(0)}(e^{2i\pi\nu})S_y^{(0)}(e^{2i\pi(\nu-\alpha)})d\nu \\
\phi_f^{(2)}(\alpha) &= 2Re\lambda\int_0^1 S_y^{(\alpha_0)}(e^{2i\pi\nu})S_y^{(\alpha_0)}(e^{2i\pi(\nu-\alpha)})d\nu
\end{aligned}$$

with $\lambda \triangleq \sum_{k=-M}^{M} e^{i2\pi k\alpha_0}$. In the latter formula, $S_y^{(\alpha_0)}$ is given by (7) and

$$S_y^{(0)}(e^{i2\pi\nu}) = |f(e^{i2\pi\nu})|^2 S_x^{(0)}(e^{i2\pi\nu}). \qquad (10)$$

We emphasize that the study of the asymptotic mean (9) is quite meaningful; indeed, if $\phi_f(\alpha)$ is "large", this means that $\hat{J}_f(\alpha)$ is likely to be above the value $\hat{J}_f(\alpha_0)$: this makes the algorithm be fooled. Of course, the analysis of the mean of $\hat{J}_f(\alpha)$ is not sufficient; one should require to study the asymptotic variance. The computation of this latter being too demanding (it involves the computation of 8th order statistics of signal y), we restrict our analysis to the asymptotic mean.

As a consequence, one may expect a better performance of the estimator of α_0 for filters $f(z)$ that verify

1. $\phi_f(\alpha)$ is small for every α
2. Eq. (8) holds.

Let us denote by $\hat{\gamma}$ a certain parameter in $(0,1)$. We consider, for every α a filter $f(z)$, denoted by $f_{\alpha,\hat{\gamma}}(z)$ defined by

$$f_{\alpha,\hat{\gamma}}(e^{2i\pi\nu}) = \begin{cases} 1 & \text{if} \quad \nu \in \mathcal{B}_{\alpha,\hat{\gamma}} \bigcup (\mathcal{B}_{-\alpha,\hat{\gamma}}) \\ 0 & \text{if not} \end{cases} \qquad (11)$$

Clearly, (8) with $f = f_{\alpha_0,\hat{\gamma}}$ holds when

$$\hat{\gamma} \geq \gamma. \qquad (12)$$

It remains to show that such filters $f_{\alpha,\hat{\gamma}}(z)$ have the effect of decreasing the asymptotic mean $\phi_{f_{\alpha,\hat{\gamma}}}(\alpha)$ for every α. It was first noticed by Mazet that for $\alpha > \alpha_0\gamma$, $\phi_f^{(2)}(\alpha) = 0$ with $f(z) = 1$. Thanks to the definition of $f_{\alpha,\hat{\gamma}}(z)$, we have a stronger result:

$$\phi^{(2)}_{f_{\alpha,\hat{\gamma}}}(\alpha) = 0$$

when $\hat{\gamma} < \frac{1}{\gamma}$ This is proved by elementary algebra. Notice this latter condition is not restrictive at all: since $\gamma < 1$, it suffices to take any $\hat{\gamma} < 1$.

As far as $\phi^{(1)}_{f_{\alpha,\hat{\gamma}}}(\alpha)$ is concerned, we obviously have for every $\alpha \neq \alpha_0$:

$$\phi^{(1)}_{f_{\alpha,\hat{\gamma}}}(\alpha) < \phi^{(1)}_{f(z)=1}(\alpha)$$

This is due to the fact that the effect of $f_{\alpha,\hat{\gamma}}(z)$ in the expression of $\phi^{(1)}_{f_{\alpha,\hat{\gamma}}}(\alpha)$ is simply to restrict the integral of a positive function to a smaller interval. This means that whatever $\hat{\gamma}$ may be, the asymptotic mean of our criterion is below this of the standard criterion. More precisely, we may state that if $\gamma_1 < \gamma_2$, then for every $\alpha \neq \alpha_0$, we have

$$\phi^{(1)}_{f_{\alpha,\gamma_1}}(\alpha) < \phi^{(1)}_{f_{\alpha,\gamma_2}}(\alpha).$$

This says that the asymptotic mean is all the smaller as the parameter $\hat{\gamma}$ of the filter $f_{\alpha,\hat{\gamma}}$ is small. On the other hand, (12) should be satisfied in order to keep (8) true for $f(z) = f_{\alpha_0,\hat{\gamma}}(z)$, indicating a tradeoff. In practice, the choice of the parameter has been noticed not to be difficult (see next section).

3. SIMULATIONS

3.1. Illustrations and results

We have considered QPSK modulated signals. The sampling period is fixed to $T_e = 1$. We generate independent data sets according to the following principles: $T_s = 4$ making the positive cyclic frequency to be estimated $\alpha_0 = 1/4$; the noise is white and Gaussian; the channel stems from 1) a raised-cosine shaping filter with an excess bandwidth factor $\gamma = 0.1$ and 2) multi-path effects (3 random paths with random delays in $(0\ ,\ 2T_s)$ and random attenuations).

We fixed $M = 10$, relying on the fact that, for the channels considered, $R_y^{(\alpha_0)}(\delta)$ is numerically weak if $|\delta| > 10$. As far as the choice of parameter $\hat{\gamma}$ is concerned, we notice that a good choice should be $\hat{\gamma} = \gamma$; as γ is not available, a compromise has to be made. We have simply chosen, in all the simulations $\hat{\gamma} = 0.2$.

The results presented below show that even in the present context (with $\gamma = 0.1$), this choice for $\hat{\gamma}$, though not optimal, allows our algorithm to overpass the performance of the standard algorithm.

The maximization of function $\alpha \mapsto \hat{J}_f(\alpha)$ cannot be achieved thanks to a standard steepest descent algorithm. This is due to the fact that the function in question cannot be expressed simply as a function of the variable α. The parameter that maximizes the function is found after an exhaustive search. The set the search is perform on is a uniform grid of the (wide) interval $(0.05\ ,\ 0.45)$ with a step of 2^{-10}. The true parameter lies in the grid. Lastly, the duration of observation is $1000T_s$.

In Figure 1, we have plotted for a given trial the shapes of $\alpha \mapsto \hat{J}(\alpha)$ and $\alpha \mapsto \hat{J}_f(\alpha)$ for a SNR of $15dB$. The plain lines represent the asymptotic mean. As was expected in section 2, the level of the estimation noise for $\alpha \neq \alpha_0$ is much more favorable in our approach.

3.2. Improvement

It is perfectly explained in [2][4] why a *whitening* of the statistic $\alpha \mapsto \hat{J}(\alpha)$ may improve the estimation of α_0. In these contributions, the asymptotic whitening at all orders is achieved after transforming the vector $\hat{\mathbf{R}}_x^{(\alpha)}$ into $\Gamma_x^{-1/2}(\alpha)\hat{\mathbf{R}}_x^{(\alpha)}$. The matrix $\Gamma(\alpha)$ is ill-conditionned, hence a tricky procedure of selection of subspace has to be considered. The performance of the estimate of α_0 has been reported to be very sensitive to this procedure. In order to overcome these difficulties, an idea consists in considering the whitening only at the first-order, *i.e.* consider the criterion $\hat{J}(\alpha)$ normalized by the asymptotic mean $\lim_{N\to\infty} N\mathbb{E}\hat{J}(\alpha)$.

Of course, the same procedure may apply to our estimate. We theoretically justify in a forthcoming paper the benefit to this procedure.

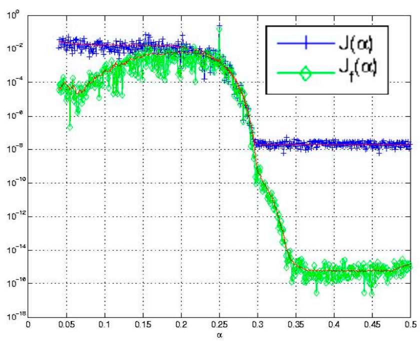

Fig. 1. $\hat{J}(\alpha)$ and $\hat{J}_f$

In order to measure the performance of the different algorithms, we give in Figure 2 the histograms of the functions

$$a_1(\alpha) = \max_{\alpha,\alpha\neq\alpha_0} \frac{\hat{J}_f(\alpha)}{\hat{J}_f(\alpha_0)} \text{ and } a_2(\alpha) = \max_{\alpha,\alpha\neq\alpha_0} \frac{\hat{J}(\alpha)}{\hat{J}(\alpha_0)}$$

Of course, if, for a realization, $a_1(\alpha)$ for instance, is above 1, the associated estimate of α_0 is wrong. The SNR is $15dB$. Both the standard and our estimate are improved by the first-order whitening; however, we reach 100% of correct estimation with the our filter-based method.

4. CONCLUSION

A digital communication signal shows a unique non-null second-order cyclic frequency; besides, the associated cyclo-spectrum has the feature that its support is narrow around half the cyclic-frequency. This is a precious piece of information that allows us to introduce a novel estimate of this cyclo-frequency. Besides the theoretical justification, the simulations clearly indicate the superiority of this estimate as compared to the standard estimate even in a difficult scenario.

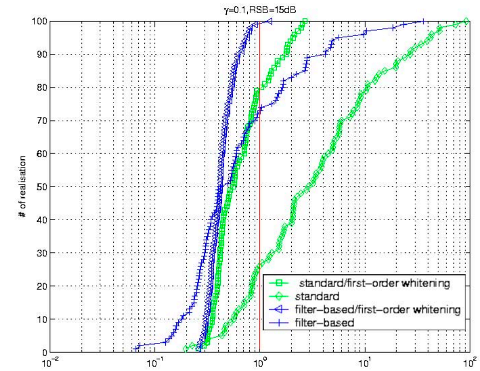

Fig. 2. Histograms of $\hat{J}(\alpha)$, $\hat{J}_f$, $a_1(\alpha)$ and $a_2(\alpha)$

5. REFERENCES

[1] W.A. Gardner, "Signal interception, a unifying theorical framework for feature detection framework for feature detection,," *IEEE Transactions on Communications*, vol. 36, no. 8, 1988.

[2] F. Gini and G.B. Giannakis, "Frequency offset and symbol timing recovery in flat-fading channels: a cyclostationary approach," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 400–411, 1998.

[3] A.V. Dandawatté and G.B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, 1994.

[4] L. Mazet and P. Loubaton, "Cyclic correlation based symbol rate estimation," in *ASILOMAR*, 1999, pp. 1008–1012.

[5] A. Touati and J-C. Pesquet, "Wavelet estimation of cyclospectra," in *Proceedings of ICASSP*, 2003, vol. 6, pp. 289–292.

[6] I. Bradaric, A.P. Petropulu, and K.I. Diamantaras, "On blind identifiability of fir-mimo systems with cyclostationary inputs using second order statistics," *IEEE trans. on signal processing*, 2003.

[7] Abed-Meraim K., Yong Xiang, J.H. Manton, and Yingbo Hua, "Blind source-separation using second-order cyclostationary statistics," *IEEE Trans on signal processing*, 2001.

[8] Pierre Jallon, Antoine Chevreuil, Philippe Loubaton, and Pascal Chevalier, "Separation of convolutive mixtures of cyclostationary sources: a contrast function based approach," in *ICA*, Grenade, Spain, September 2004.

[9] S.V. Schell, "An overview of sensor-array processing for cyclostationary signals," in *Cyclostationarity in Communication and Signal Processing*. 1994, pp. 168–239, W.A. Gardner ed., IEEE PRESS.

[10] D. Dehay and H.L. Hurd, "Representation and estimation for periodically and almost periodically correlated random processes'," *Cyclostationarity in Communications and Signal Processing, W.A. Gardner, Ed., IEEE Press*, 1993.

[11] P. Ciblat, P. Loubaton E. Serpedin, and G.B. Giannakis, "Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1922–1934, july 2002.