transmission on the full rate GSM voic
RR-94-006
853 1
Documentation Eurecom

Institut EURECOM
2229, route des Crêtes
Sophia Antipolis
06560, VALBONNE
FRANCE

Research Report N° 94-006

Data Transmission on the Full Rate GSM Voice Channel

Jean Varaldi Pierre A. Humblet Dirk T.M. Slock

April 26, 1994

Telephone:	+33 93 00 26 26	E-mail:
Jean Varaldi:	+33 93 00 26 03	varaldi@eurecom.fr
Pierre A. Humblet:	+33 93 00 26 35	humblet@eurecom.fr
Dirk T.M. Slock:	+33 93 00 26 06	slock@eurecom.fr
Fax:	+33 93 00 26 27	

Abstract

In this report, we examine the possibility of transmitting data on the GSM full rate voice channel. We conclude that the goal is achievable in principle, but that it requires too much processing power. We also examine the transmission of V.23 signals through the encoder-decoder chain and show experimentally that they cannot be distinguished reliably.

Contents

Abstract i

1 Introduction 1

2 The GSM full rate vocoder 1

3 Experiments with V.23 signals 3

References 5

1 Introduction

In the GSM system voice is encoded by a vocoder before transmission on the air interface. The transmission rate is 13 kb/s in the Full Rate vocoder, before addition of redundancy for channel coding. The GSM standard also specifies data transmission at rates up to 9.6 kb/s. However operators have not yet deployed this capability, and the subscription to the service will cost extra when it will be available. Thus it would be useful to transmit data on the voice channel, either by using specially designed modems, or preferrably by using existing V-series standards. The second option would make it easy to communicate from a GSM mobile to the huge installed modem base in the fixed telephone network.

This paper examines the feasibility of these two options. It concludes that transmission at rates close to 13 kb/s is possible, but that it requires outlandish processing ressources.

It also examines the use of a V.23 interface (FSK modulation) at 1200 b/s in the direction to the mobile, with a reverse channel at 75 b/s. The interest of this configuration is that it requires only the development of the receive section of the mobile modems. No change at all would be needed in the modems already installed in the fixed network. The speeds are sufficient for useful interactive data transmissions. Unfortunately we conclude that this approach would also require excessive computing resources.

According to section A1.3.4 of [1], V.21 transmission at 300 b/s is not subject to any significant degradation in absence of uncorrected transmission errors on the air interface. This appears to be the only way to transmit data on the GSM full rate voice channel.

Two parts follow this introduction. In the first we describe the Full Rate GSM vocoder, keeping in mind data transmission at 1200 b/s. In the second we describe some experiments with V.23 signals transmitted on the GSM full rate voice channel.

2 The GSM full rate vocoder

The GSM full rate vocoder [1] uses Linear Predictive Coding. Voice frames of 20 ms or 160 samples (24 bits at 1200 b/s) are analyzed and encoded into 260 bits. 182 of those bits are protected by an error correcting code, 78 are not.

The first significant operation consists in the estimation of the autocorrelation function of the signal in a frame, and of filter coefficients representing the autocorrelation. This representation uses 36 of the 260 output bits, 24 of them are protected.

The residuals are then evaluated, using coefficients resulting from a time varying combination of the filter coefficients from the current and from the

previous frame (*). The 160 residual samples are then divided in 4 segments of 40 samples (5 ms, or 6 bits at 1200 b/s). The cross-correlations are then evaluated between each segment of 40 samples and 40 previous samples, with a lag varying between 40 and 120 (thus cross-correlation with samples anywhere in the previous 3 segments). For each segment the peak of the cross-correlation is identified, and it is normalized to produce a gain coefficient. The lag of the peak and the gain coefficient are coded using 7 and 2 bits respectively, for a total of 36 bits, all protected.

The residuals are then subjected to a first order prediction (using the lag and the gain just measured), and the residuals' residuals are evaluated in turn. They are low pass filtered and separated in 3 subsegments (of 14, 13 and 13 samples) by down sampling by a factor 3. The subsegment with maximum energy is selected, its index is encoded (2 bits) as well as its elements (45 bits), for a total of 188 bits (122 protected).

From a data transmission at R b/s point of view, it is clear that the whole operation can be represented by a Markov chain. The state of the chain is the sequence of $R*$ 20 ms bits transmitted in a frame (2^{24} states at 1200 b/s). Transitions between any pairs of states are possible. If a continuous-time signal is assigned to each of the transmitted sequences, then to each transition in the chain corresponds a unique continuous time waveform reproduced in a deterministic fashion at the receiver (in absence of uncorrected error).

The decomposition of each frame into 4 sequences may lead to some structure in the transition matrix, but not to a reduction of the number of states. However if the linear combination (*) above is ignored, the states can be defined as the $R*$ 15 ms bits transmitted over 3 sequences (2^{18} states at 1200 b/s), each such state having only $2^{R*\ 5\ \mathrm{ms}}$ transitions (64 at 1200 b/s). The resulting Markov chain is periodic with a period of 4.

If we had the technology to process such huge trellises, we believe that reliable transmission would be possible. The only design part lies in the selection of the transmitted signals. As the encoder presumably makes good use of all the encoded bits, it must be possible to find enough transmitted signals corrresponding to well separated received signals, even at rates approaching 13 kb/s (at the cost of 10^{250} states). The receiver would simply consist of correlators and of a Viterbi decoder.

On the other hand the previous description makes it clear that the GSM vocoder relies on long term features that are abundant in speech, but not present in the signals produced by today's modems. In particular the second evaluation of residuals is completely inappropriate. Even if there was only one evaluation of residuals (or if the gain of the second evaluation was always zero), the down sampling by a factor of 3 would cause much damage. This leads us to the final part of this paper.

3 Experiments with V.23 signals

The Mobile Communication Group has developed a C language implementation of the GSM full rate speech coder and decoder. For this project we have also coded an implementation of a V.23 transmitter producing a continuous phase FSK signal at 1300 or 2100 Hz, and we have linked the two packages to run on a workstation. We passed the FSK signal through the encoder/decoder chain and processed the output. The FSK signal was synchronized to start on GSM frame boundaries, an idealization that we would eventually remove.

In a first series of experiments we approximated the signal by a Markov chain, each state consisting of the most recent n bits. This type of approximation can be quite accurate when the distortion occurs in a time-invariant manner. This is not the case here, the frame structure of the GSM coder makes it clearly time variant.

The "states" of our chain actually represents a collection of true "Markov states" and they are better labelled "reduced states". A set of output waveforms can be expected in each reduced state for the transmission of a "0", and another set for the transmission of a "1". For each reduced state we calculated the autocorrelation function of the waveforms in these sets (each waveform consists of only 6 samples at 8000 samples/sec), and performed a Karuhnen-Loeve expansion. For each set in each reduced state we kept the two eigenvectors with the most energy. The reason for this choice is that we did not want to consider synchronous reception. Thus we expected two significant eigenvalues. This expectation turned out to be relatively accurate.

We used a Decision Feedback structure, ignoring decision errors in our trials (we used transmitted bits and not decoded bits). In each state we projected the received signal on the space spanned by the two selected eigenvectors in the "0" set, and on those in the "1" set, and we used envelope detection to make decisions.

The system performed exceedingly well in absence of the coder/decoder (figures 1 and 2), but badly when they were included (figures 3 to 6). The error probability remained well above 10^{-2} irrespective of the number of states (up to 128) and of the degree of causality considered.

In a second series of experiments, we calculated the complex envelope and attempted to make decisions based on the instantaneous frequency. Three sets of figures are presented, (7-10), (11-14) and (15-18). In each set, the signal in the time domain is displayed (figure 7) along with its normalized version (figure 8). The normalization was performed on each waveform corresponding to one bit. Another interesting graph is the instantaneous phase of the signal (figure 9). The data sequence used and the one demodulated are also given (figure 10). No errors occur in the first set.

The visual inspection of the signal makes clear that the encoding and decoding processes can actually produce output waveforms at 2100 Hz from

input signals at 1300 Hz (time 710 in figure 11). The instantaneous phase of the signal (figure 13) shows this switchover.

The reverse phenomenon where a waveform at 1300 Hz is produced from an input signal at 2100 Hz is present in the third set (time 450 in figure 15).

Furthermore, the same local context (reduced state), i.e. 6 previous identical bits, exists at time 450 in the third set, and at time 2480 in the first set, although the waveforms are very different and an error occurs in one case and not in the other. Consequently, the local context cannot help us in our decision algorithm. One should consider using much larger Markov chains and gaining knowledge of the GSM frame structure.

Note also that one would expect errors when the coding/decoding process results in signals with low energy. However, a data sequence can be retrieved from such a signal (figure 16) and a high energy signal can lead to an error (figure 12).

References

[1] ETSI. *Recommendation GSM 06.10*, February 1992.

<X,V1>^2 + <X,V2>^2 ou X est le signal avant Coder-Decoder

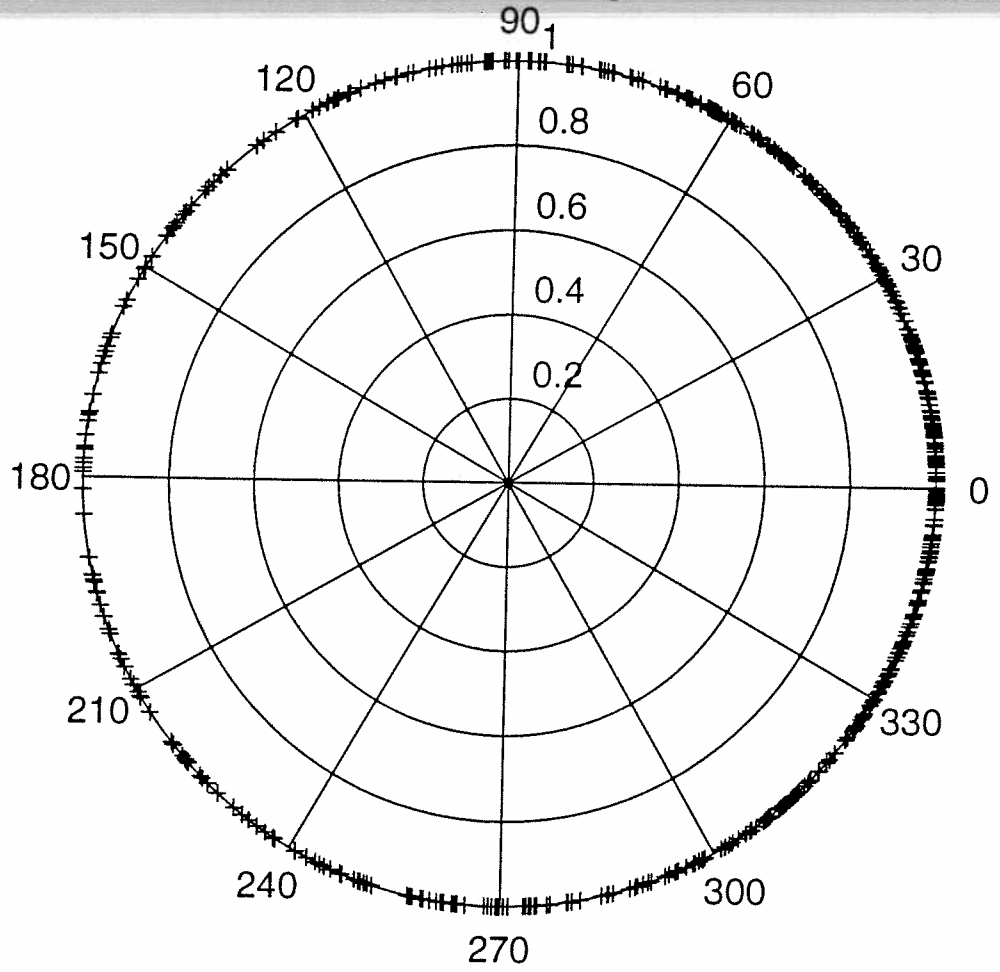


Figure 1

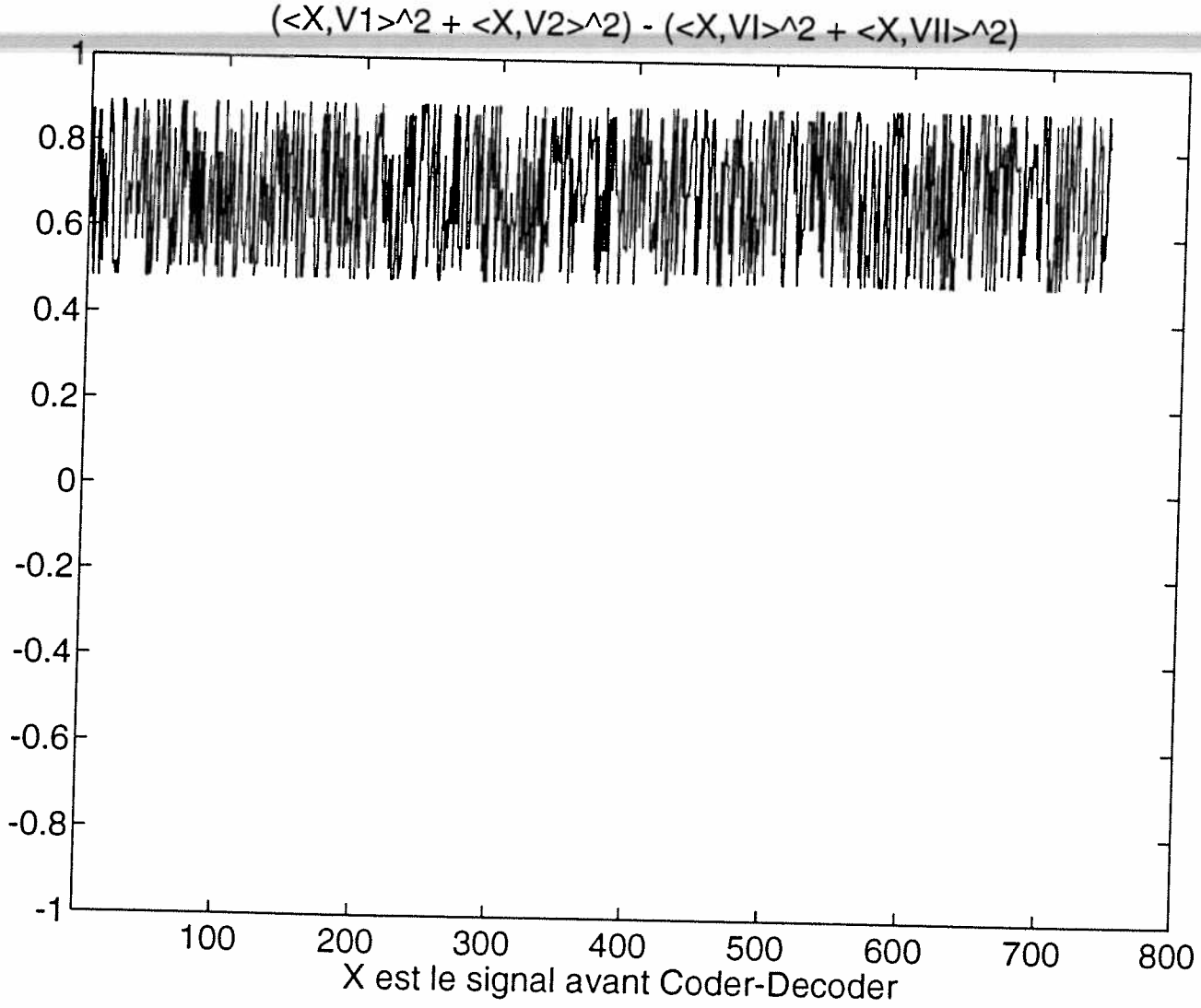
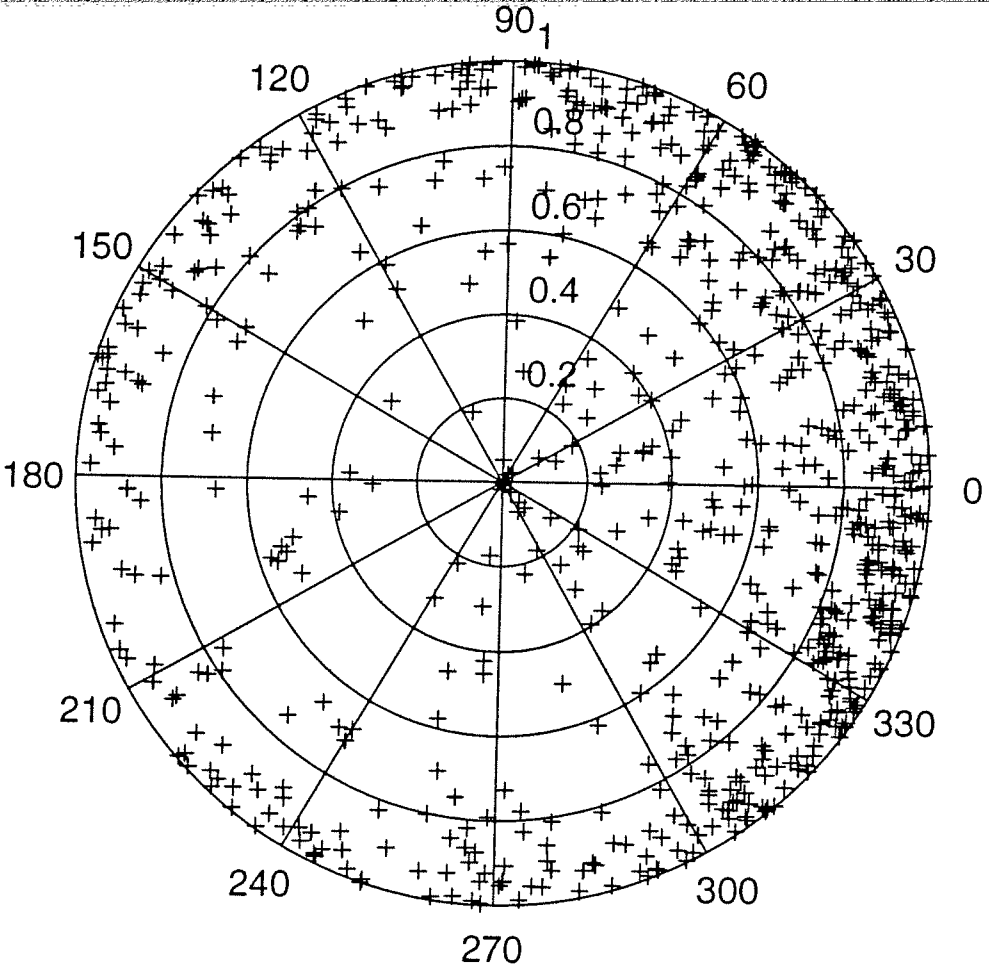

(<X,V1>^2 + <X,V2>^2) - (<X,VI>^2 + <X,VII>^2)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
100
200
300
400
500
600
700
800
X est le signal avant Coder-Decoder

Figure 2

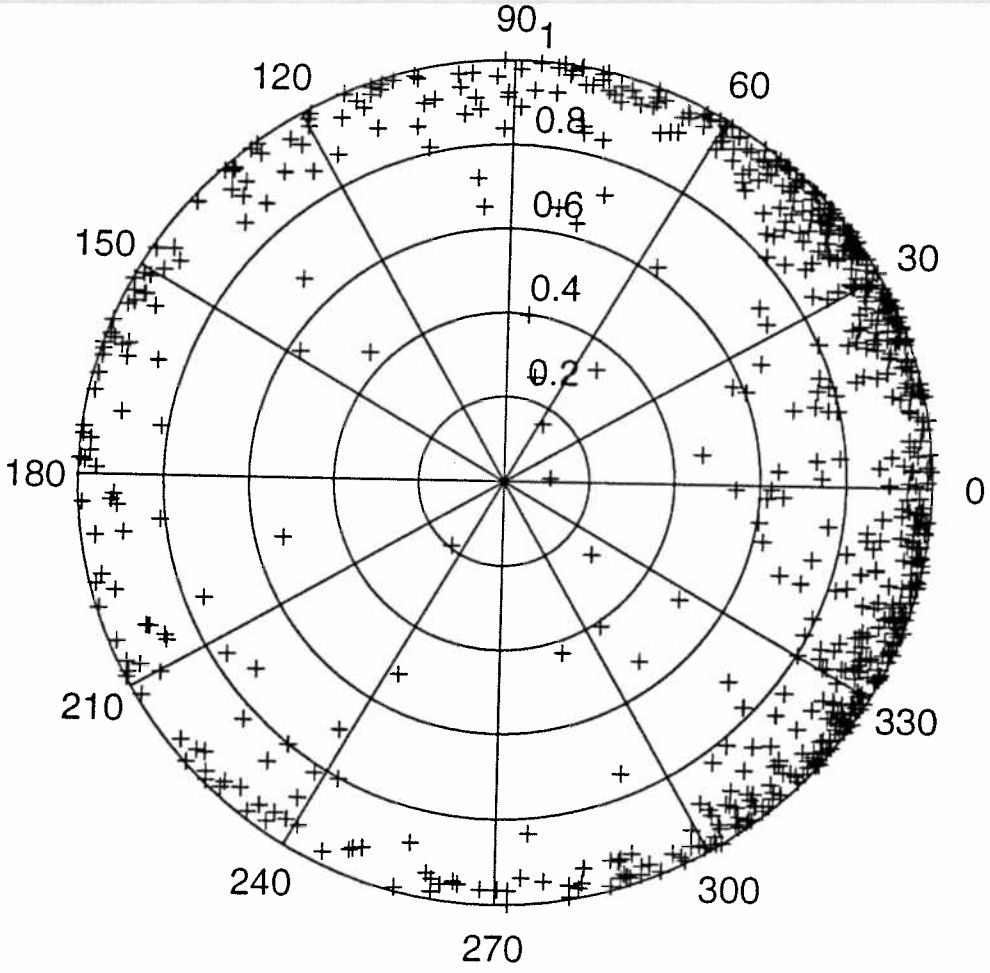
<X,V1>^2 + <X,v2>^2 ou X est le signal, V1 et V2 les vecteurs de b ases



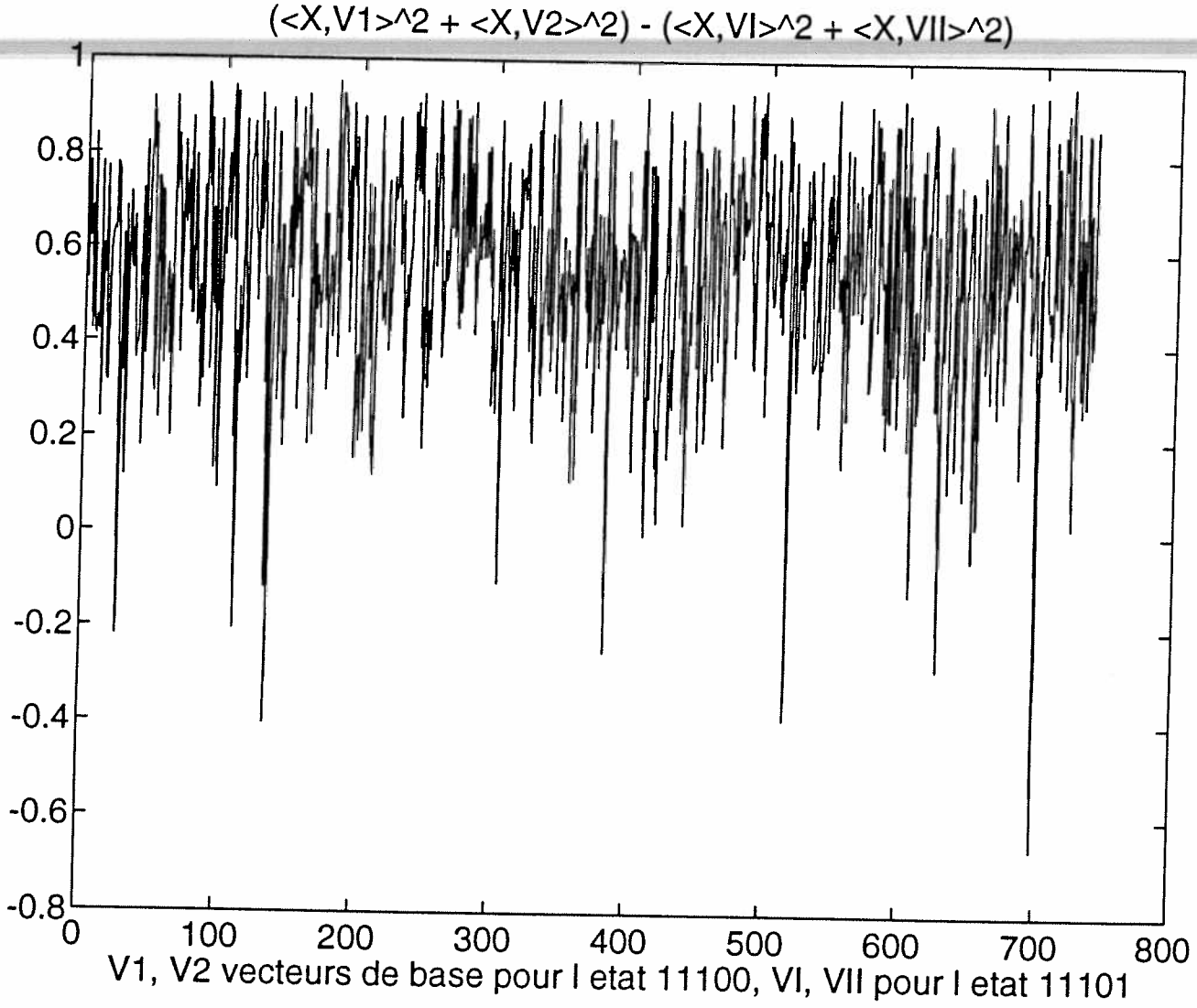
Etat ---> 00001

figure 3

<X,V1>^2 + <X,v2>^2 ou X est le signal, V1 et V2 les vecteurs de b ase



Etat ---> 11110


(<X,V1>^2 + <X,V2>^2) - (<X,VI>^2 + <X,VII>^2)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
100
200
300
400
500
600
700
800
V1, V2 vecteurs de base pour l etat 11100, VI, VII pour l etat 11101

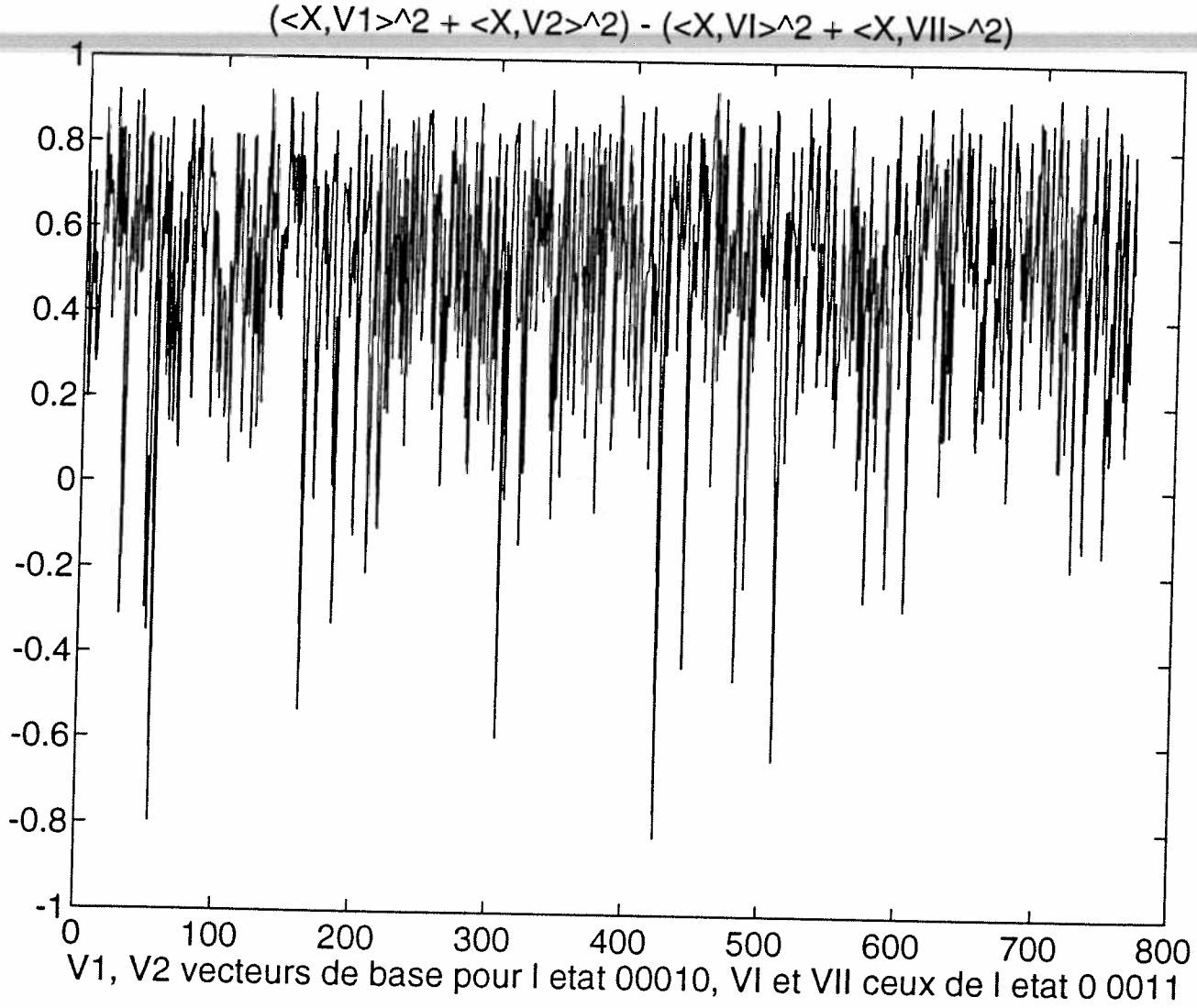

(<X,V1>^2 + <X,V2>^2) - (<X,VI>^2 + <X,VII>^2)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
100
200
300
400
500
600
700
800
V1, V2 vecteurs de base pour l etat 00010, VI et VII ceux de l etat 0 0011

Figure 5

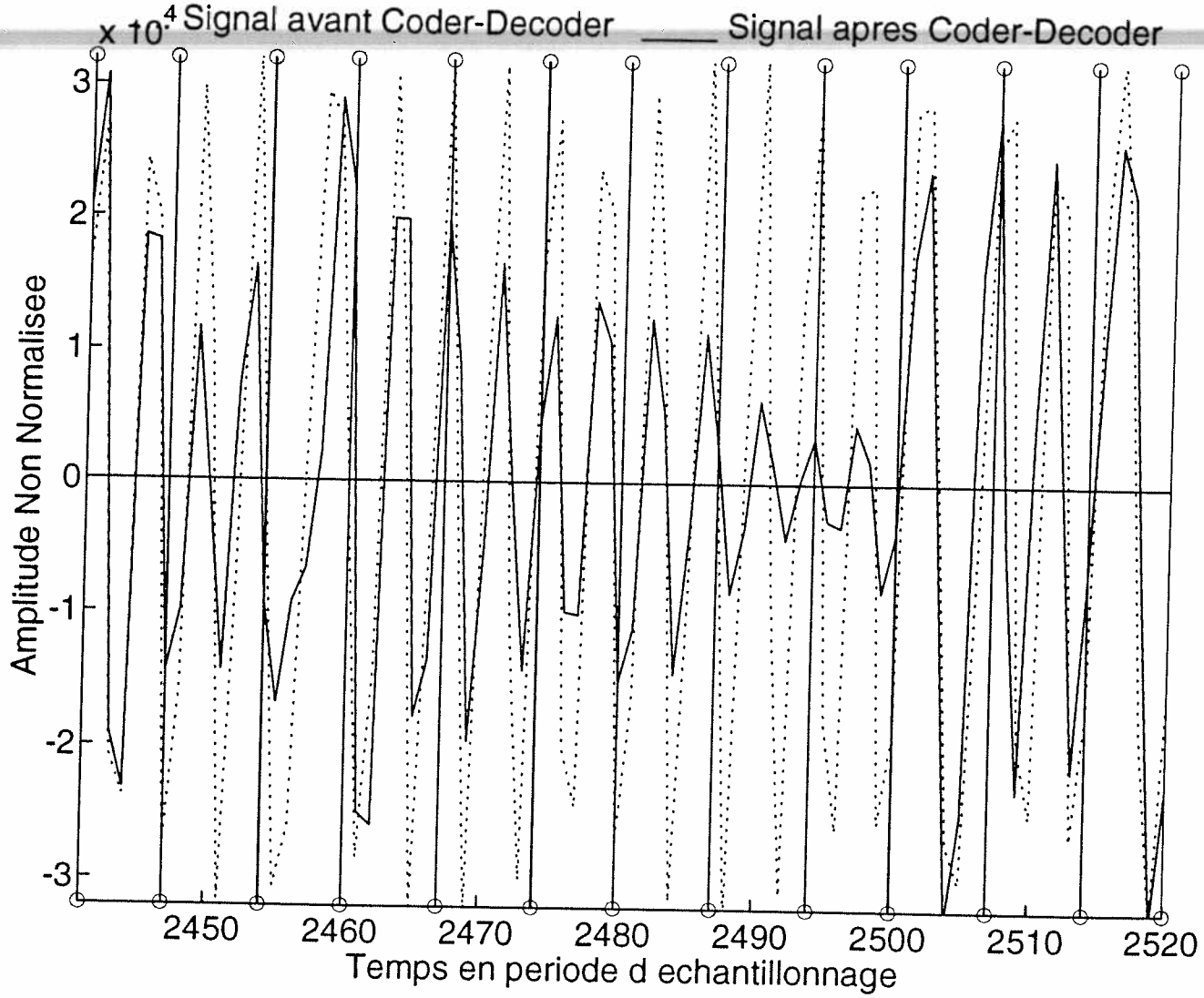
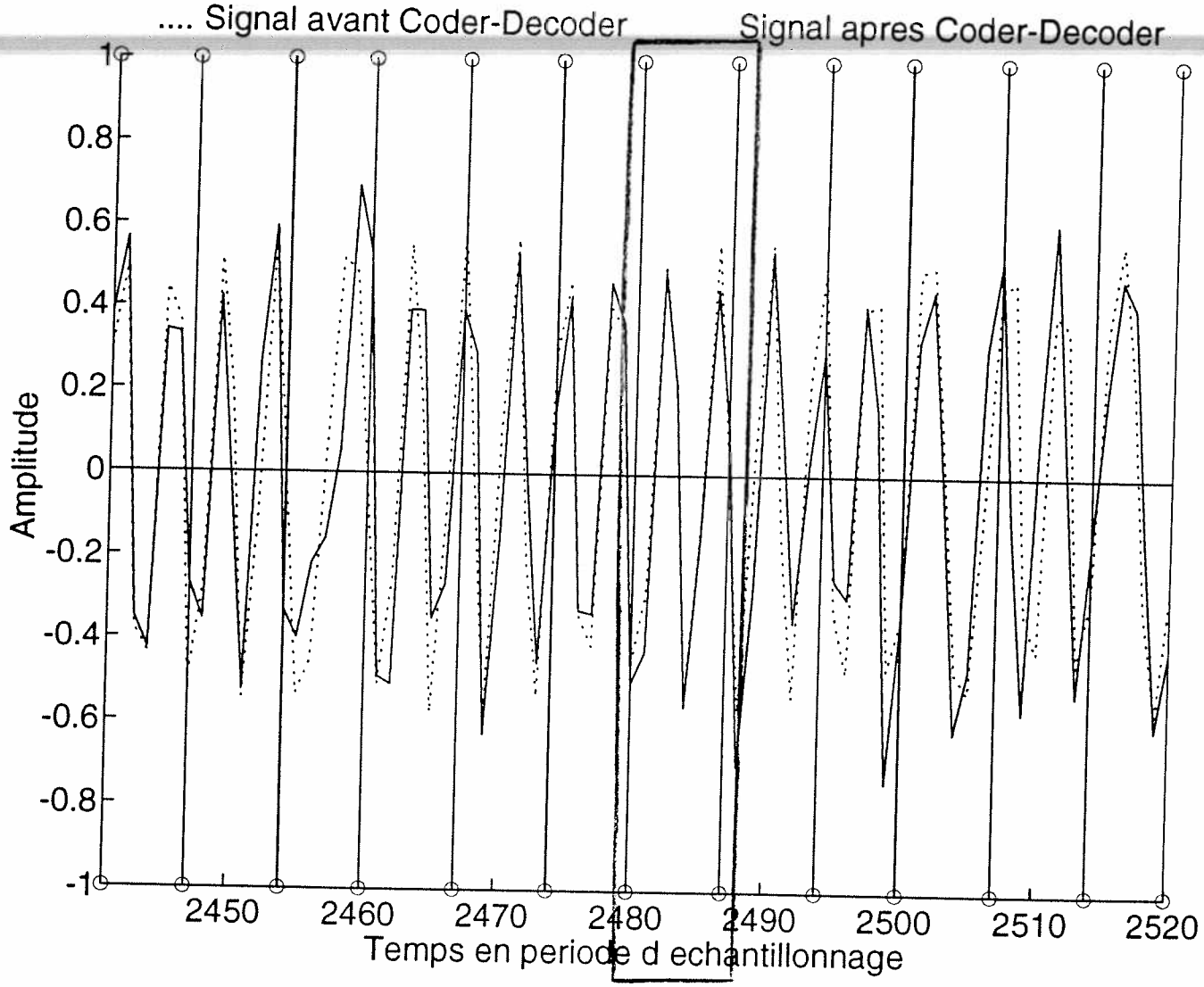

x 10^4 Signal avant Coder-Decoder
Signal apres Coder-Decoder
Amplitude Non Normalisee
3
2
1
0
-1
-2
-3
2450
2460
2470
2480
2490
2500
2510
2520
Temps en periode d echantillonnage

Figure 7


.... Signal avant Coder-Decoder
Signal apres Coder-Decoder
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
Amplitude
2450
2460
2470
2480
2490
2500
2510
2520
Temps en periode d echantillonnage

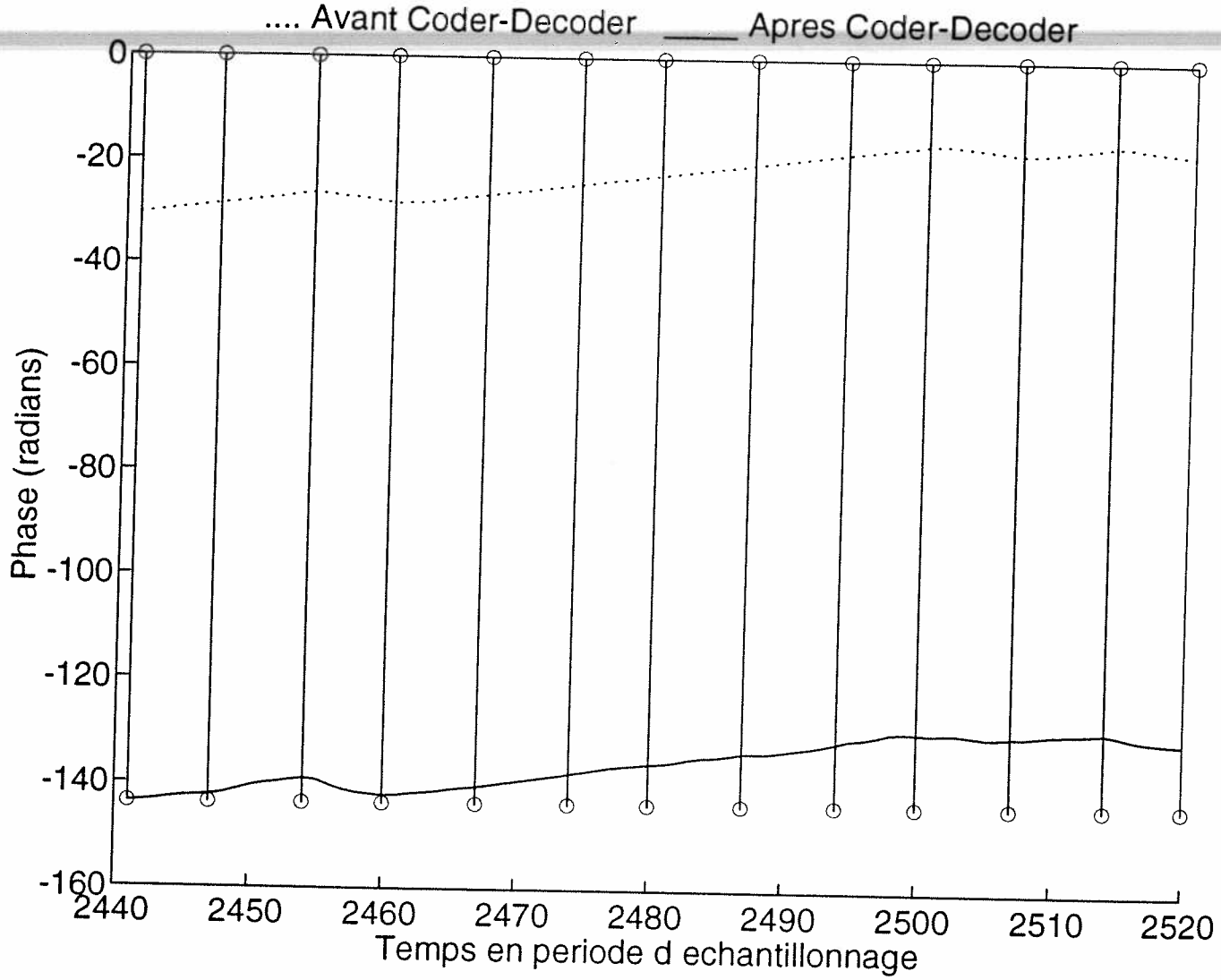
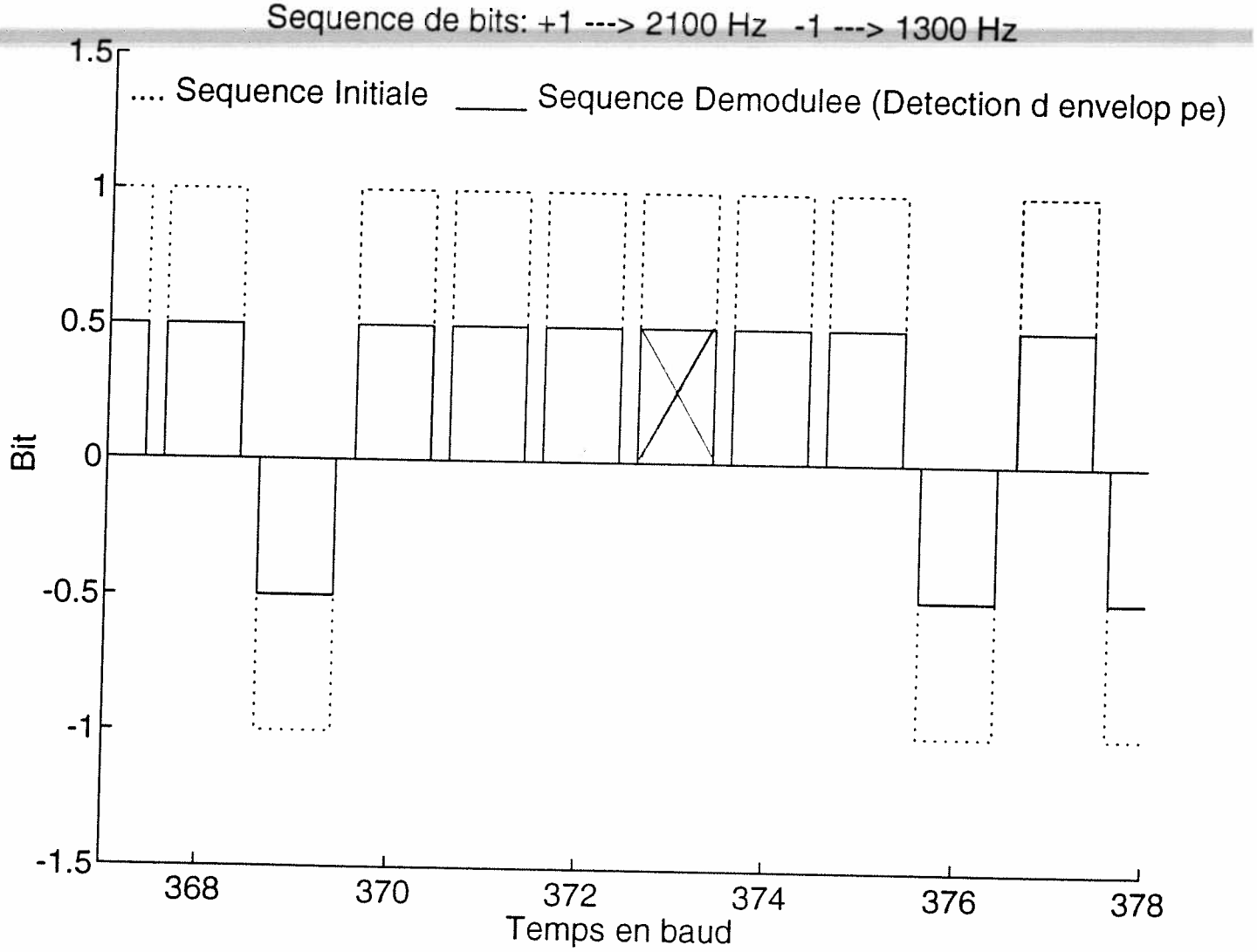

.... Avant Coder-Decoder
____ Apres Coder-Decoder
0
-20
-40
-60
-80
-100
-120
-140
-160
Phase (radians)
2440
2450
2460
2470
2480
2490
2500
2510
2520
Temps en periode d echantillonnage

Figure 3


Sequence de bits: +1 ---> 2100 Hz -1 ---> 1300 Hz
.... Sequence Initiale
Sequence Demodulee (Detection d envelop pe)
1.5
1
0.5
0
-0.5
-1
-1.5
Bit
368
370
372
374
376
378
Temps en baud

Page 10

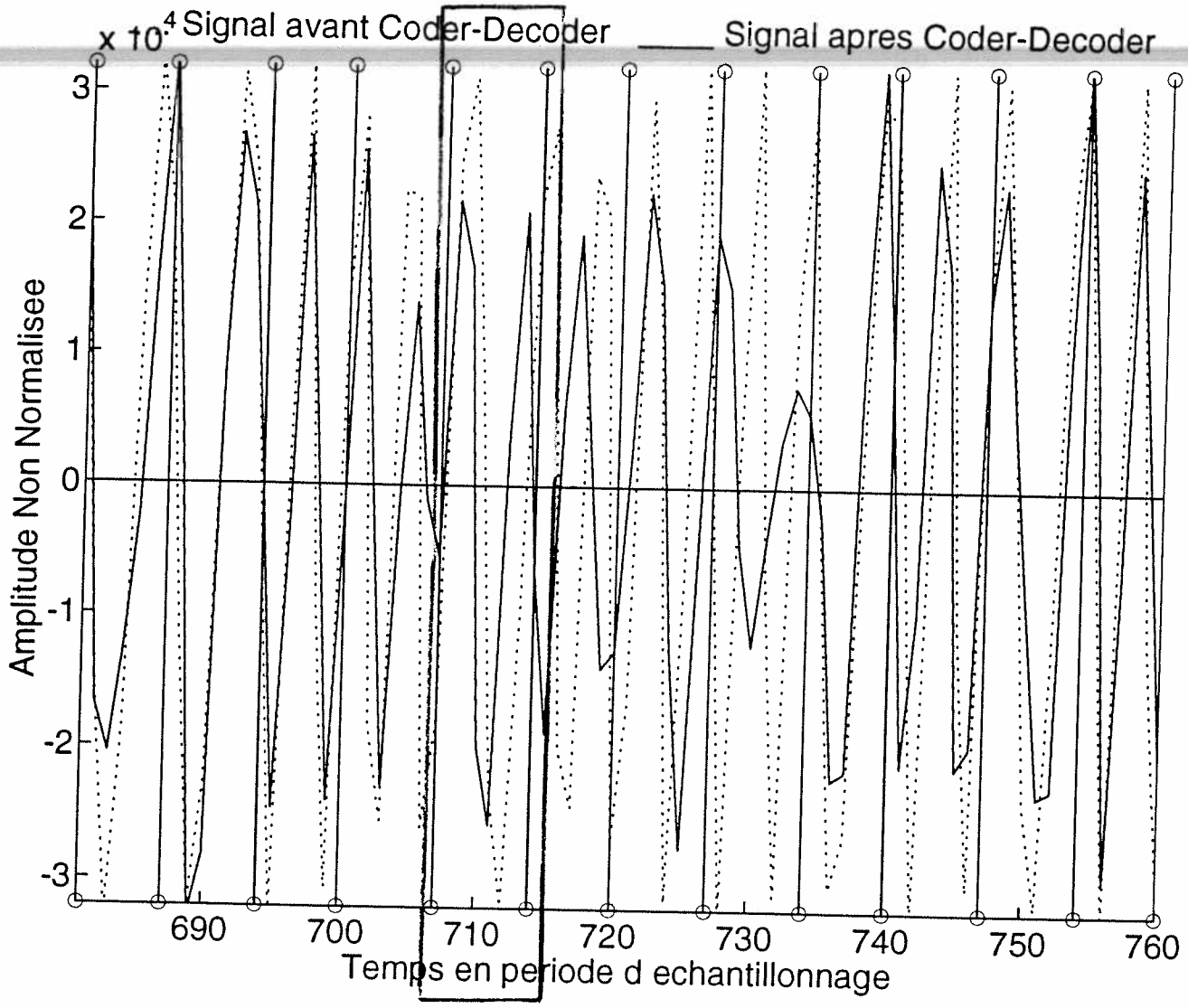

x 10^4 Signal avant Coder-Decoder
Signal apres Coder-Decoder
Amplitude Non Normalisee
3
2
1
0
-1
-2
-3
690
700
710
720
730
740
750
760
Temps en periode d echantillonnage

Figure 11

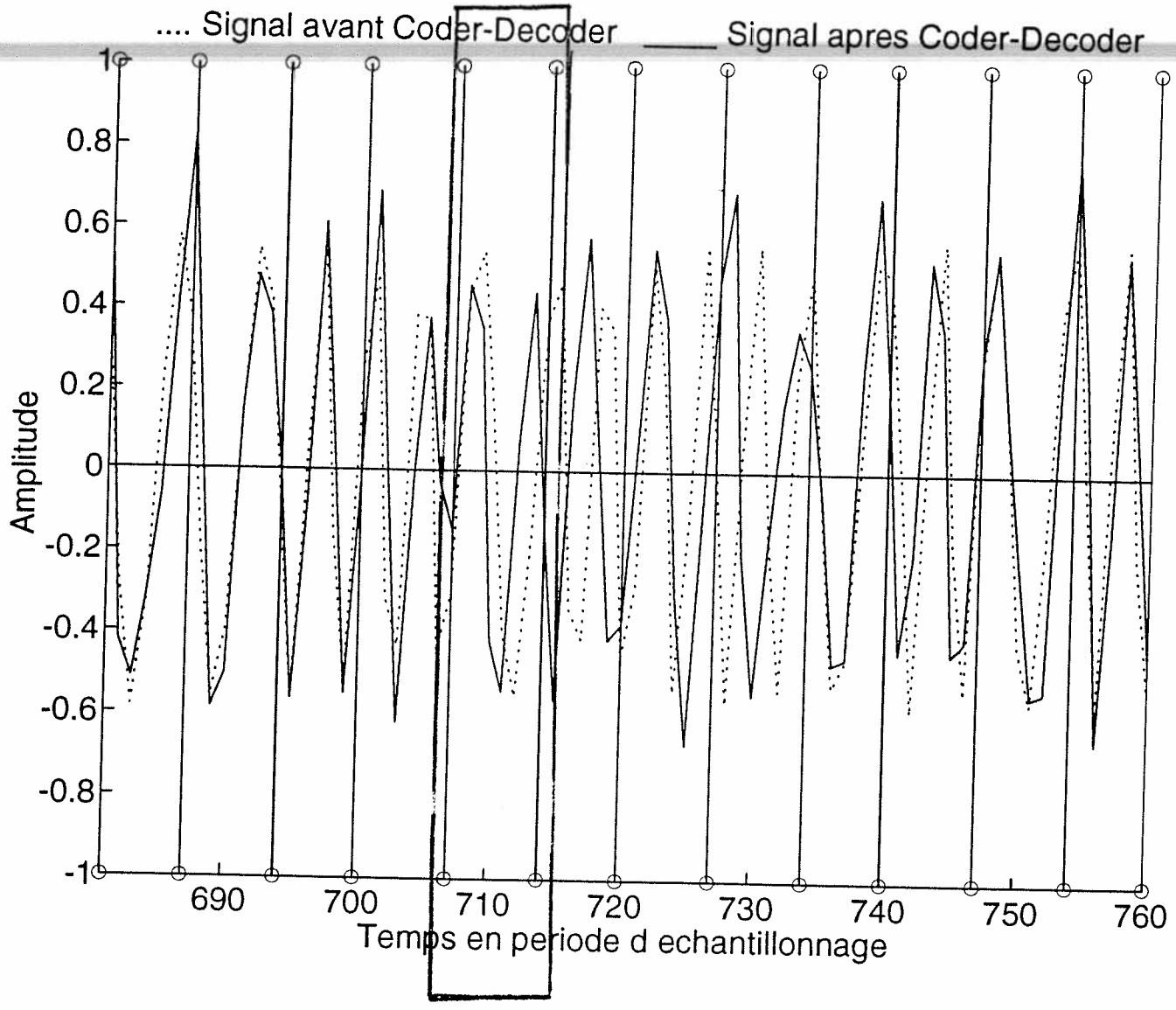

.... Signal avant Coder-Decoder
____ Signal apres Coder-Decoder
Amplitude
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
690
700
710
720
730
740
750
760
Temps en periode d echantillonnage

Figure 12

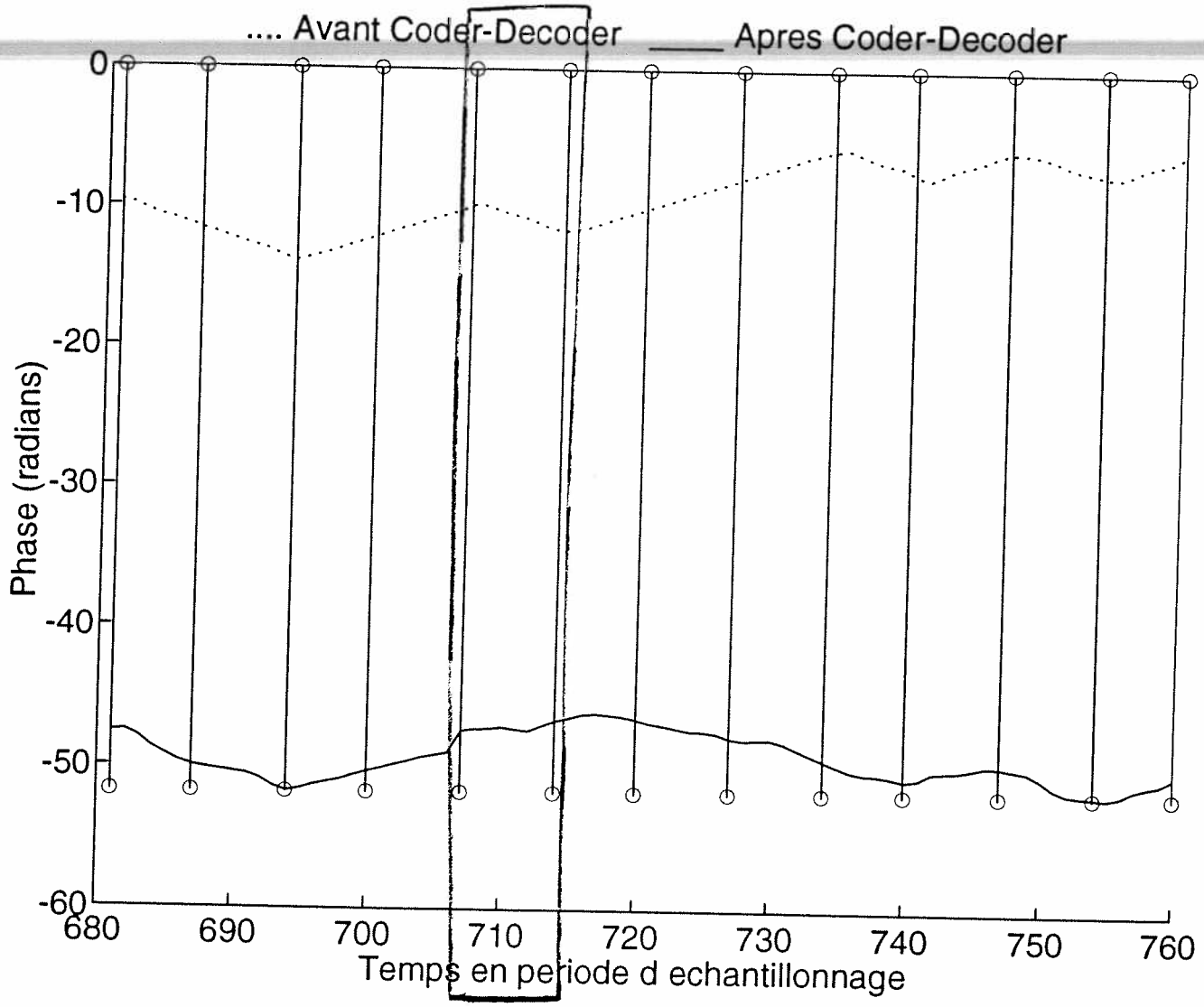

.... Avant Coder-Decoder
____ Apres Coder-Decoder
0
-10
-20
-30
-40
-50
-60
Phase (radians)
680
690
700
710
720
730
740
750
760
Temps en periode d echantillonnage

Figure 13

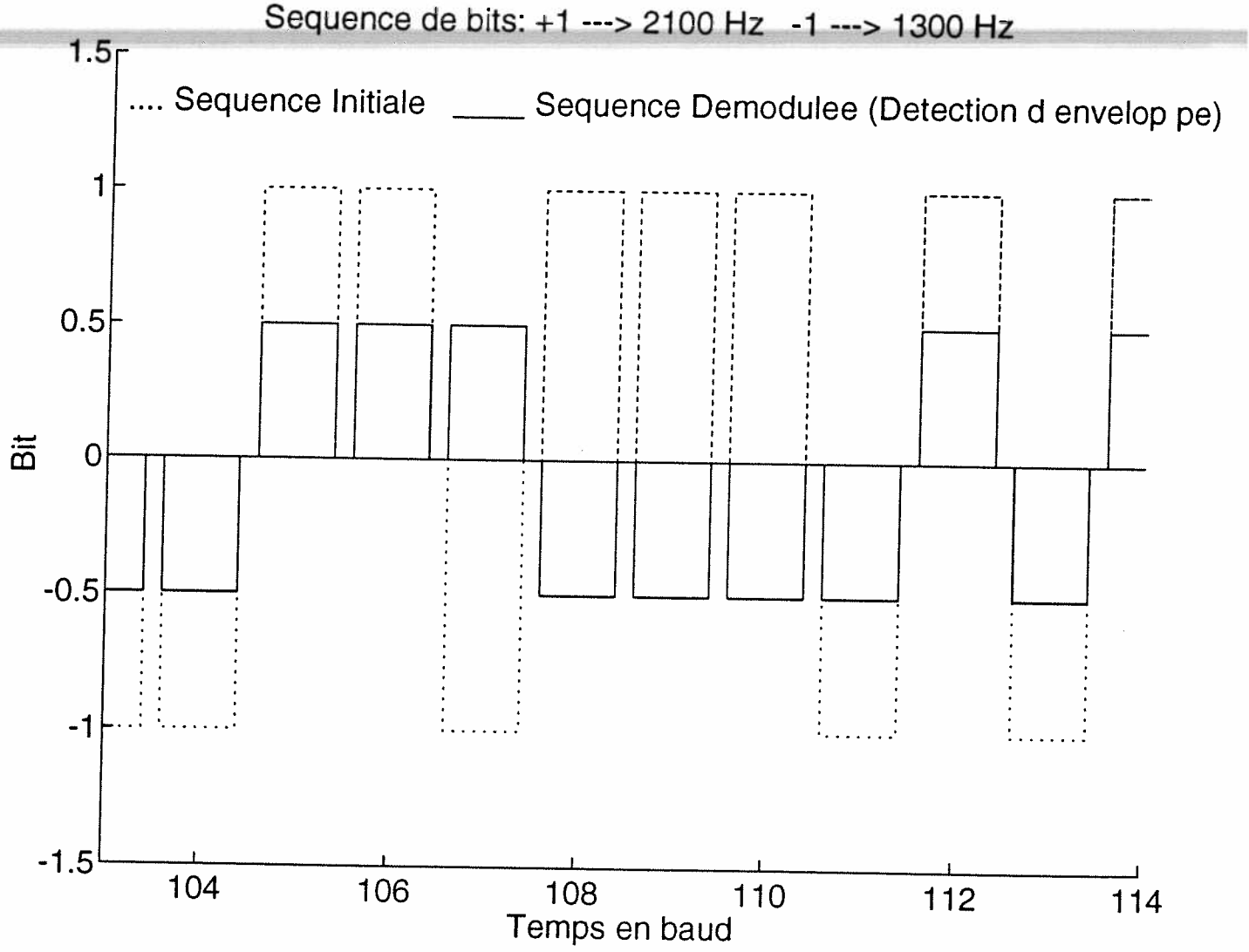

Sequence de bits: +1 ---> 2100 Hz -1 ---> 1300 Hz
.... Sequence Initiale
____ Sequence Demodulee (Detection d envelop pe)
1.5
1
0.5
0
-0.5
-1
-1.5
Bit
104
106
108
110
112
114
Temps en baud

Figure 14

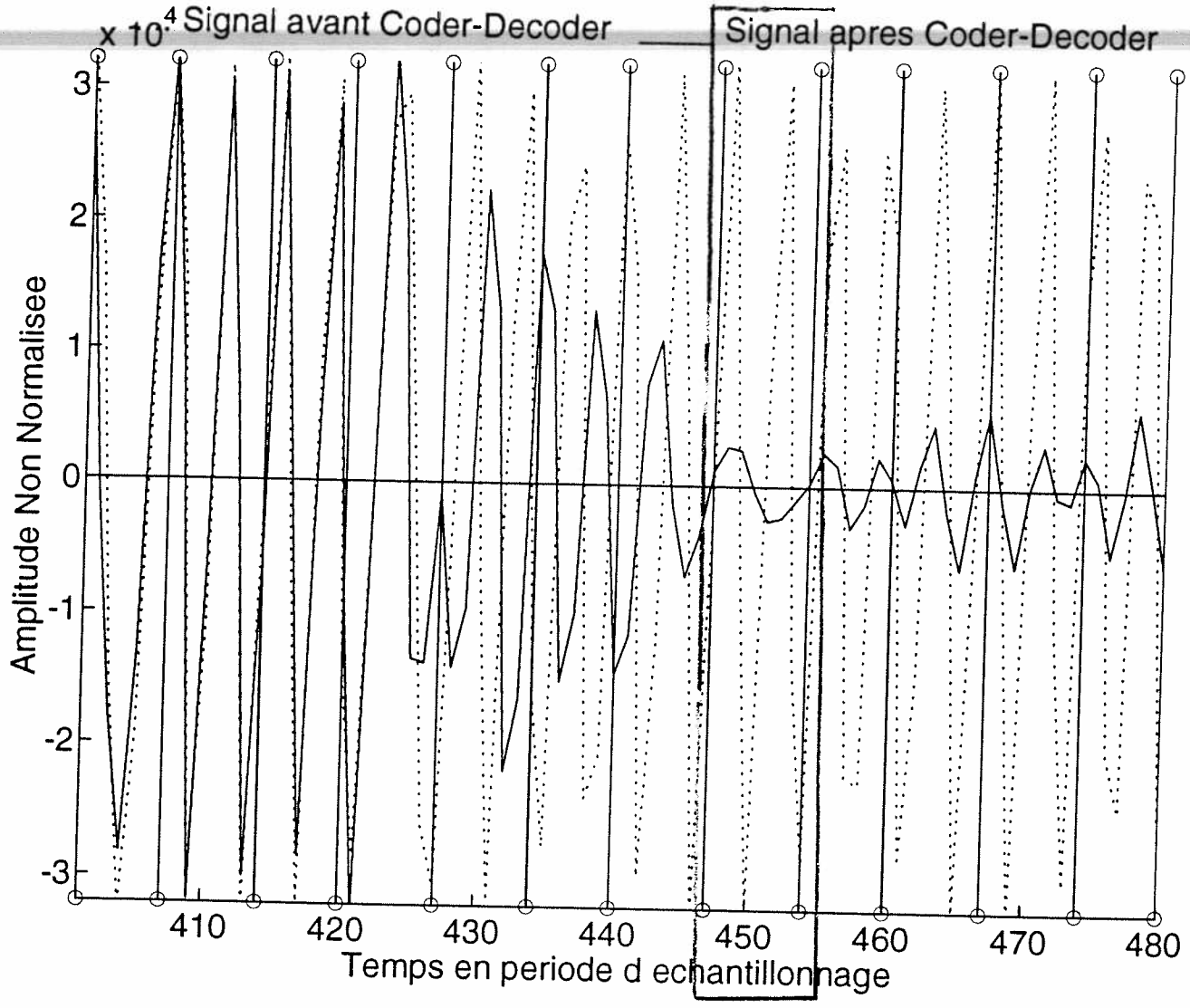

x 10^4 Signal avant Coder-Decoder
Signal apres Coder-Decoder
Amplitude Non Normalisee
Temps en periode d echantillonnage
3
2
1
0
-1
-2
-3
410
420
430
440
450
460
470
480

Figure 15

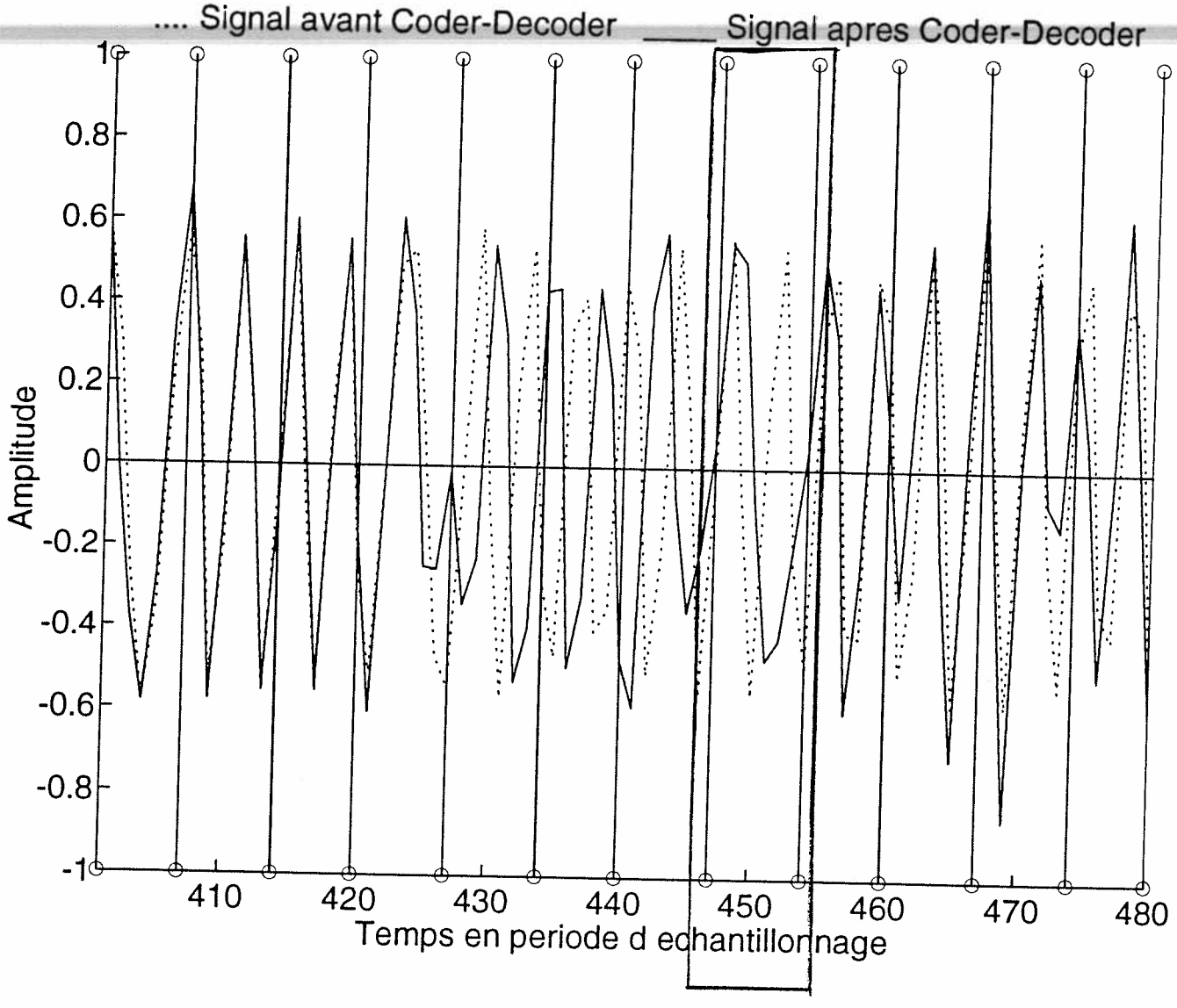

.... Signal avant Coder-Decoder
____ Signal apres Coder-Decoder
Amplitude
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
410
420
430
440
450
460
470
480
Temps en periode d echantillonnage

Figure 6

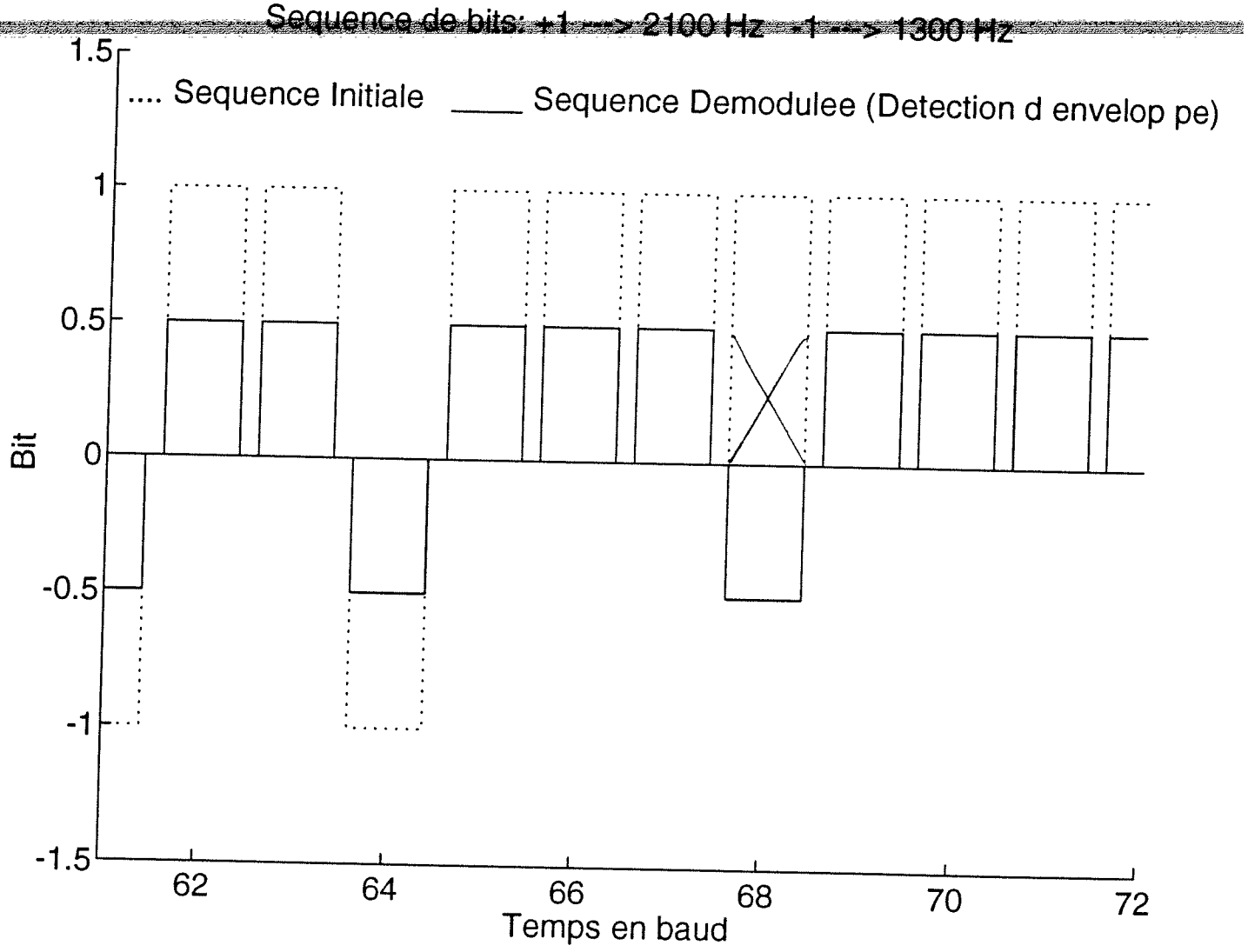

Sequence de bits: +1 ---> 2100 Hz -1 ---> 1300 Hz
.... Sequence Initiale
____ Sequence Demodulee (Detection d envelop pe)
1.5
1
0.5
0
-0.5
-1
-1.5
Bit
62
64
66
68
70
72
Temps en baud

Figure 17

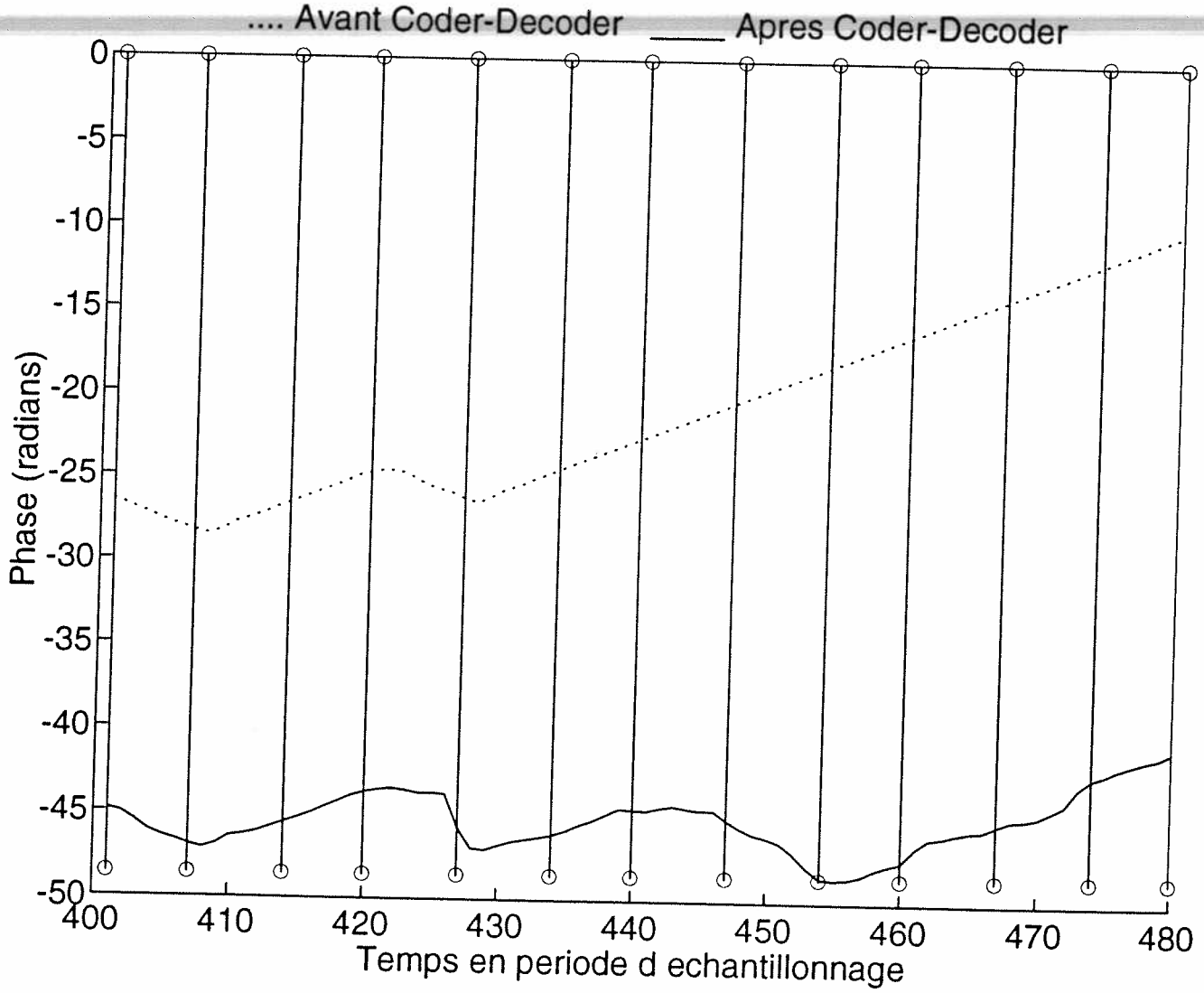

.... Avant Coder-Decoder
____ Apres Coder-Decoder
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
Phase (radians)
400
410
420
430
440
450
460
470
480
Temps en periode d echantillonnage

Figure 19