XVI IMEKO World Congress
Measurement - Supports Science - Improves Technology - Protects Environment ... and Provides Employment - Now and in the Future
Vienna, AUSTRIA, 2000, September 25-28

ERRORS IN PARAMETER ESTIMATION OF QUANTIZED SINUSOIDS

M. Fonseca da Silva, A. Cruz Serra,
Lab. Medidas Eléctricas, DEEC,IST/IT, Lisbon Technical University, Portugal
Phone: 351-218418474, Fax: 351-218417672,

Abstract: Signal-processing applications frequently use least-square sine-fitting algorithms. The parameter estimation provided by these algorithms is exposed to errors due to different causes. Good results in amplitude estimation may be achieved when the frequency is unknown applying the new method presented in this paper.
Keywords: sine fitting, parameter estimation.

1 INTRODUCTION

New generation of multipurpose measurement equipment is transforming the role of computers in instrumentation. The new features involve usually analog-to-digital and digital-to-analog converters substituting typical discrete instruments like multimeters, oscilloscopes, spectrum analyzers, etc. Frequently they imply the use of sophisticated signal processing algorithms. Among them are the sine fitting algorithms. They are used in different situations like automatic impedance measurement [1] or in the characterization of analog-to-digital converters.

In this paper we will analyze some aspects related with the sine-fitting algorithms suggested in the IEEE Standard for Digitizing Waveform Recorders [2] which estimate the parameters of a sine wave defined as $A_0 \cos(2p f_0 t_n) + B_0 \sin(2p f_0 t_n) + C_0$ that may fit a set of M samples, $y_1,...,y_M$, acquired at a frequency $f_s = 1/T_s$.

The residuals, r_n,of the fit are given by

$$r_n = y_n - A_0 \cos(2p f_0 t_n) - B_0 \sin(2p f_0 t_n) - C_0 \qquad (1)$$

In the graphics and tables we will refer the error relative to the amplitude of the AC component $e_A = e_{rms} / \sqrt{A_0^2 + B_0^2}$, where $e_{rms} = \sqrt{\sum_{n=1}^{M} r_n^2 / M}$ is the rms error.

2 THREE-PARAMETER SINE FITTING ALGORITHM (KNOWN FREQUENCY)

The knowledge of the frequency makes things easier, but normally one has to determine it and some precautions must be taken, otherwise errors will appear. Furthermore, one is frequently dealing with a small amount of samples and periods of the sine wave that may also be degraded by noise and/or distortion.

The importance of accurate frequency estimations may be shown in simple cases involving pure sine waves. As it is shown in Figure 1(a), a small error in frequency evaluation has produced small errors in the amplitude. On the other hand, the phase in the central sample is close to the correct value but the phase errors at the beginning or at the end of the interval are very high. Increasing the number of samples is not a solution when the frequency is not accurately estimated. The errors will also increase, as shown in Figure 1(b): the error reaches the maximum of 100% at approximately a number of samples $N = f_s / \Delta f_0$, where Δf_0 is the absolute error of the estimation of the frequency f_0.

The influence of both Δf_0 and the number of samples may be shown in the example in Figure 2, for the frequency $f_0 = 0.04 f_s$ ($T_0 \equiv 25$ samples). When the estimated frequency diverts from the correct frequency, local minimums are found. Their location depends on the number of samples, M:

– Increasing M increases the number of minimums, decreasing the gap between minimums.

– If the sampling interval contains more than one period and $\Delta f_0 = p f_s$, the maximums to the left and to the right of the correct frequency, correspond approximately to $f = f_0 \pm kpf_s / M$, with k a positive integer. Increasing the number of samples, the first left hand-side and right hand-side minimums are closer to the correct frequency, as also shown in Table 1.

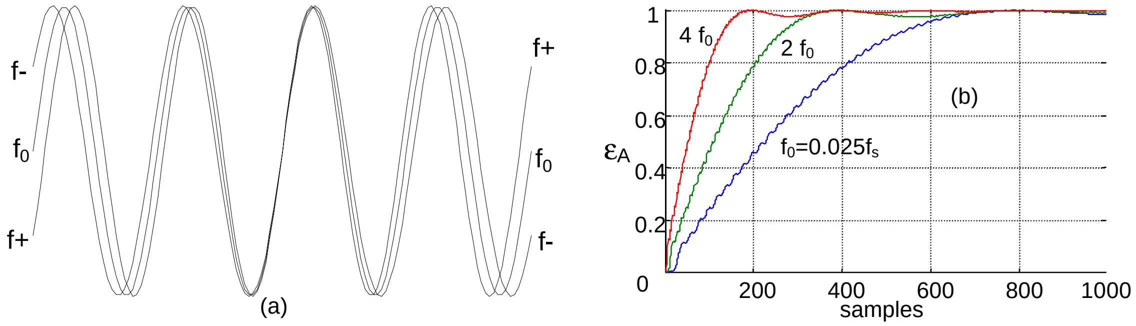

Figure 1 – (a) Errors in the estimated frequency produce initial negative or positive phases if the estimation is by excess or default, respectively. (b) Three-parameter algorithm error e_A as a function of the number of samples for three different signal frequencies, with a frequency error $\Delta f_0 = 5\% f_0$.

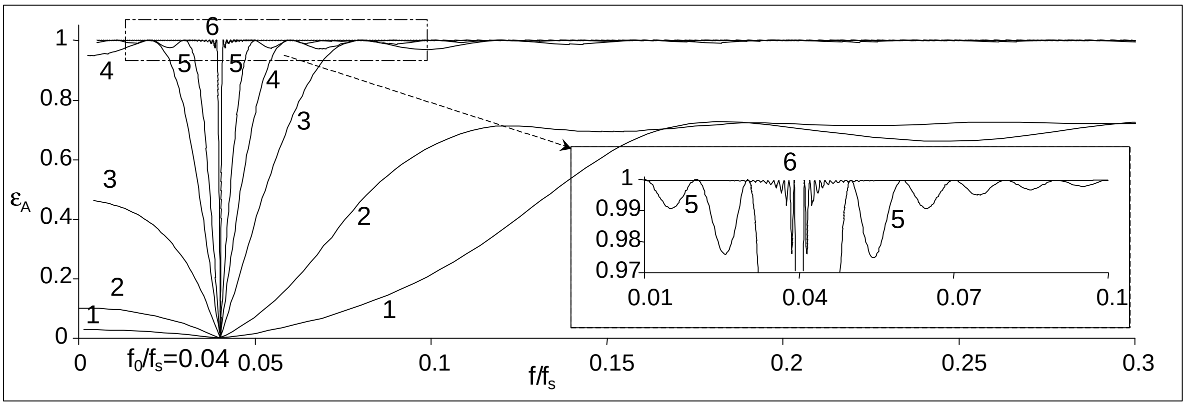

Figure 2 – Three-parameter algorithm error e_A as a function of estimated frequency, *f*, for $f_0 = 0.04 f_s$ and different number of samples. The correspondence between curves and the number of samples is: curves 1≡8 samples, 2≡15, 3≡25, 4≡50, 5≡100 and 6≡1000.

Table 1 – 1st minimums close to the correct frequency, $f_0 = 0.04 f_s$, and corresponding relative errors.

number of samples	1st left hand-side minimum $(\times f_s)$	e_A (%)	1st right hand-side minimum $(\times f_s)$	e_A (%)
5000	0.03971	97.6	0.04029	97.63
2000	0.03928	97.57	0.04071	97.65
1000	0.03857	97.53	0.04143	97.69
500	0.03713	97.45	0.04285	97.76
200	0.03278	97.15	0.0471	97.97
100	0.02535	96.51	0.05412	98.24
50	none	92.13	0.06802	98.65

If the number of samples is small containing only one period or less (curves 1 to 3 in Figure 2), there are no minimums apart the main one at f_0, or the few that may exist, occur for values very far from f_0.

In the vicinity of f_0, the sensitivity of the algorithm increases with the number of samples.

The main conclusion from the previous considerations is that if the frequency is imperfectly known one risks to get very bad results from the three-parameter algorithm. Under these circumstances, it is recommended to use the four-parameter algorithm.

3 FOUR-PARAMETER ALGORITHM SINE FITTING (UNKNOWN FREQUENCY)

With the four-parameter algorithm [2] each iteration produces a new set of values A_i, B_i, C_i and a correction Δf_i to the frequency to be used in the next iteration. The main problem of the algorithm is that it is highly dependent on the number of samples and especially on the initial estimated values, including naturally the frequency [3][4]. It presents local minimums in the most unpredictable places as shown in Figure 3. There are local minimums sometimes in different positions from the others for different initial estimated values A_0, B_0 and C_0 and even the initial phase. Figure 3 (a) depicts simulation results with a fixed initial phase and distinct initial values of the parameters A_0 and B_0. Relatively to their correct values (*A*, *B* and *C*) we have used: *A* and *B* (curve 1), *A* and *0.25B* (2), *0.25A* and *B* (3) and *0.25A* and *0.25B* (4). Similarly Figure 3 (b) depicts simulation results with the parameters A_0 and B_0, fixed in the values *0.9A* and *0.9B* and different phases: 30º(5), 120º(6), 210º(7) and 300º (8). The DC component C_0 has a negligible influence in the behavior of the algorithm and has been kept constant and equal to its correct value in all the simulations.

The algorithm looks for the solutions of a nonlinear system of equations, which must be solved in an iterative way. Occasionally the algorithm produces corrections to the frequency that must be interpreted having in mind some particularities presented in Table 2.

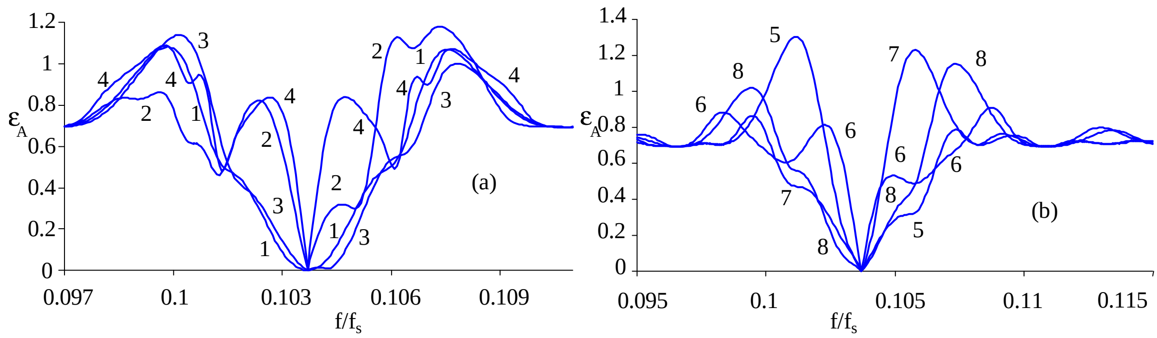

Figure 3 – Error e_A of the four-parameter sine fitting algorithm as a function of estimated frequency for 200 samples, $f_0 = 0.1037 f_s$ and in the conditions of the text: (a) different initial estimated amplitude values and fixed initial phase; (b) different initial phases and fixed initial algorithm estimated values.

Table 2 – Corrections to be made to the frequency resulting from the algorithm in especial cases.

frequency resulting from the algorithm	frequency to be used in the next iteration
$(f_0 + \Delta f_0)/f_s > 0.5$ (Nyquist frequency)	$f_0'/f_s = 1 - (f_0 + \Delta f_0)/f_s$
$(f_0 + \Delta f_0)/f_s > 1$	$f_0'/f_s = 1 - (f_0 + \Delta f_0)/f_s$
$(f_0 + \Delta f_0)/f_s > 0$	$f_0'/f_s = \lvert f_0 + \Delta f_0 \rvert / f_s$

Considering pure sine waves, the algorithm has a good performance when the number of samples per period is relatively high. However diminishing this number close to 2 samples per period increases the conditions to a convergence to local minimums that are associated to incorrect values. It is extremely critical a good estimate of the frequency as well as the other parameters. The DC component, C_0, is the exception, because its influence is minimal and not critical.

This may be brought to evidence by a simple example of a pure sine wave with a frequency $f_0 = 0.093 f_s$ (approximately 10.75 samples per period); for certain number of samples, the convergence occurs in local minimums as it is shown in Table 3.

Table 3 – Some simulations have converged to local minimums. The correct values should be: $f_0/f_s = 0.093$, $A_0 = 0.9301$, $B_0 = -0.3673$ $C_0 = 0$.

number of samples	number of periods	initial estimate				after 10th iteration	
		A_0	B_0	C_0	f_0/f_s (DFT)	f_0/f_s	e_A
48	4.464	0.928	-0.531	0.0021	0.08333	0.1715	1.0033
59	5.487	0.932	0.0285	-0.0025	0.10169	0.0262	0.9997
199	18.507	0.932	0.0278	-0.0018	0.09548	0.1017	1.0517
500	46.5	0.931	0.0266	-0.0006	0.094	0.0877	1.0027

In Figure 4 it is shown the evolution of the error obtained in the application of the four-parameter algorithm with the iteration number for a given situation. The frequency for the 1st iteration was estimated by the application of a DFT. The initial values of A_1, B_1 and C_1 were obtained using the three-parameter algorithm with the previous estimated frequency. The successive corrections made to the frequency led to a convergence in a local minimum with a high error e_A and A_0, B_0 and the final frequency very incorrect (iterations 12 to 14 have been omitted because they don't differ significantly from 11 and 15). It must be said that increasing the number of samples is not always a solution as it was verified with this signal when 500 samples were used.

In these examples a pure sine wave was used but quantification was not considered. In fact considering the effects of quantification, noise, harmonic distortion or saturation the probability of convergence to local minimums becomes higher.

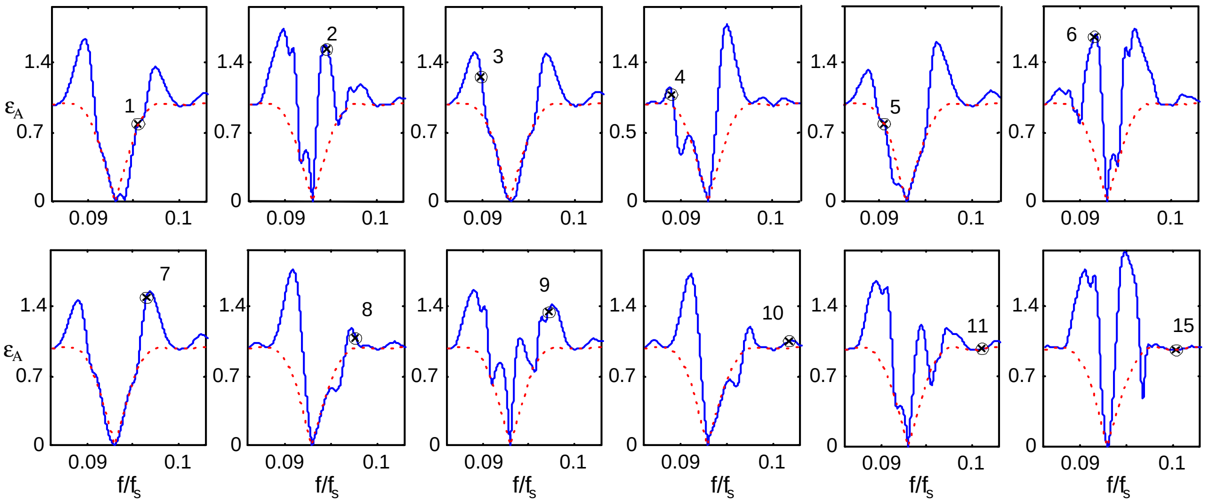

Figure 4 – Full line curves represent the error e_A as a function of f/f_s, for $f_0 = 0.093 f_s$, 199 samples and the values of A_i, B_i and C_i used in iteration i in the four-parameter algorithm. The mark ⊗ near the iteration number represents the situation correspondent to the frequency used by the algorithm in each iteration. The error of the three-parameter algorithm is represented in dots.

3.1 Application of the algorithm with increasing number of samples

As it was mentioned in 2 and shown in Figure 2 and Figure 3, if the number of samples is very small containing 1 period or less, the maximum of e_A closest to f_0 is located very far from f_0. As a consequence, we have adopted the following procedure to improve convergence:

1) Estimate roughly the frequency from a DFT, considering all the samples, M (or FFT if M is a power of 2)
2) Apply this frequency to the three-parameter algorithm to estimate the values of A_0, B_0 and C_0 to be used in the first iteration.
3) Consider a reduced number of samples, a subset $N_1 < M$ containing roughly 1 period, based on the value obtained in 1) for the frequency.
4) Apply the estimates of the frequency and the other parameters to the four-parameter algorithm with the subset of samples N_1, allowing only frequency corrections between iterations much smaller than f_s / N_1, until the resulting error is smaller than a predefined level.
5) Increase the number of the subset N_1 considered by the algorithm, attending to the error and the correction to the frequency obtained in the previous iteration.
6) Repeat steps 4) and 5) until the total number of samples M is reached.

This procedure assures that the frequency is not "over-corrected" but, on the contrary, it is modified towards the absolute minimum. With exactly the same signal mentioned in Figure 4 it has been possible to obtain the convergence to the absolute minimum as is shown in Figure 5, employing this procedure. The estimated frequency was $f_0 = 0.095 f_s$ and only 9 samples were used in the first iteration. The error obtained has determined a very small increase to 10 samples in the second iteration having resulted a smaller error with a still considerable correction of frequency; then the number of samples were increased just a little bit more, to 16. In the following iterations the number of samples were progressively increased to 59 and finally to 199 samples, since the errors and corrections of frequency became smaller.

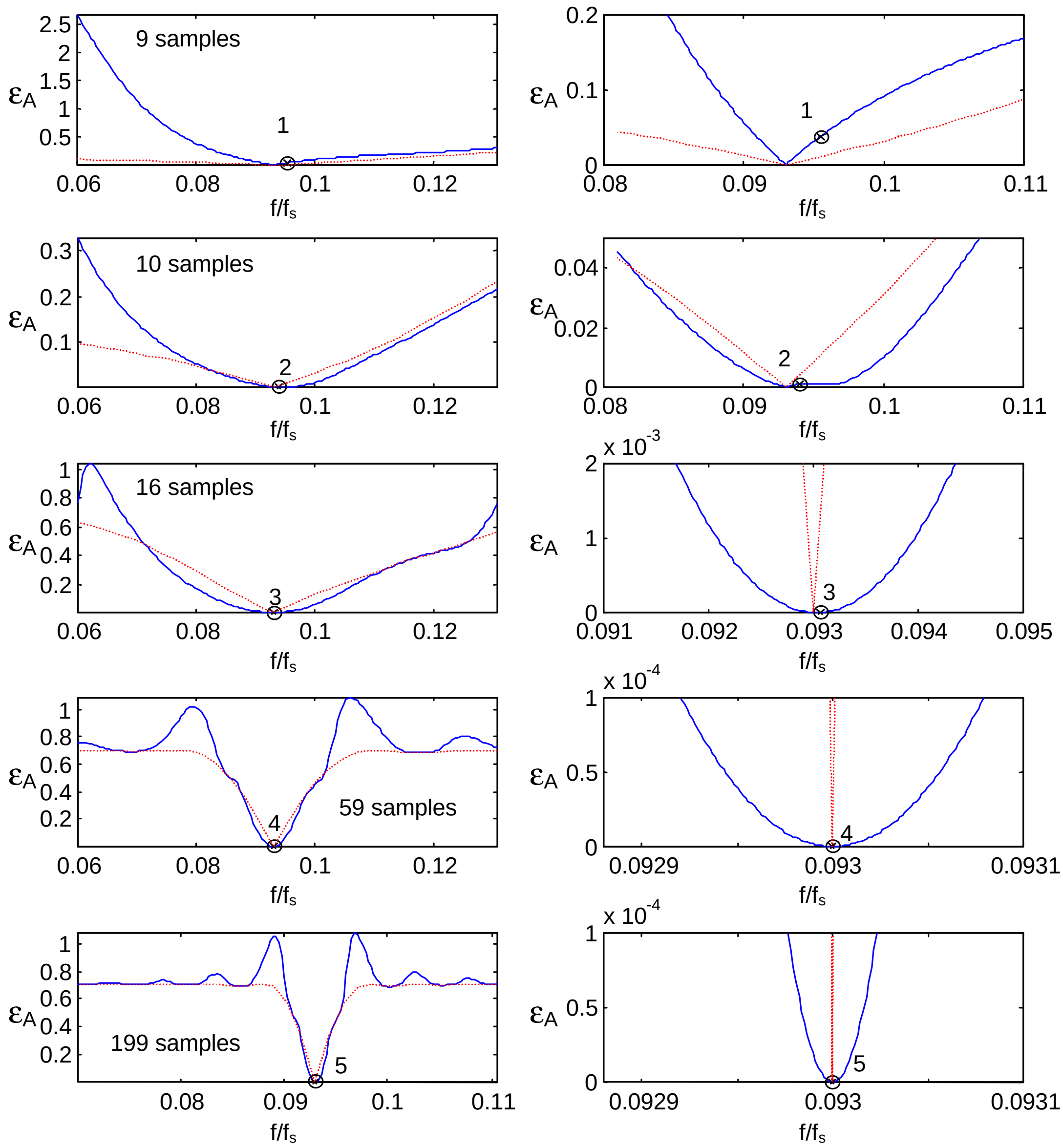

Figure 5 – Error e_A as a function of f/f_s for the four-parameter algorithm, with $f_0 = 0.093 f_s$ and 199 samples as in Figure 4, but applying a progressive number of samples. The region near the correct function is enlarged in the graphics on the right.

3.2 Seeking a good estimate of the frequency

As we have stated the estimate of the frequency restricts the behavior of the four-parameter algorithm because it also conditions the values of the other parameters. So in order to find a good frequency to start the four-parameter algorithm in association with good estimates to the other parameters we have implemented another procedure to get the better frequency to be used:

1) Estimate roughly the frequency from a DFT, considering all the samples, which will be named f_1.
2) Evaluate the frequency implementing a software Schmitt-trigger, with hysteresis and the reference values are based on the maximum and minimum values of the sampled data, dividing the gap in three equal intervals: $V_{ref1} = (2V_{\max} + V_{\min})/3$ and $V_{ref2} = (V_{\max} + 2V_{\min})/3$.

   The result is a function u(t) with a frequency f_2 close to the one we are determining. This method is very useful if the set of samples contains an odd and very small number of semi-periods.
3) Apply a moving average to smooth the sampled data. This is very important if there is noise in the signal. The new set of data constructed averaging three consecutive samples:

   $$z_n = (y_{n-1} + y_n + y_{n+1})/3. \qquad (2)$$

   is applied to steps 1) and 2) determining two more frequencies f_3 and f_4.

4) Apply these 4 frequencies to the three-parameter algorithm; the frequency that leads to the smaller error is the one that has more chances to be close to the right value of frequency.
5) Based on this selected frequency the amount of samples is reduced in order to get approximately an integer number of periods; with this reduced number of samples a new DFT is calculated, getting another value to the frequency, f_5. Apply this frequency f_5 to the three-parameter algorithm. If the error obtained is the smallest, this is the elected frequency; otherwise, the one obtained in 4) will be chosen to the next step.
6) Having in mind that the closest maximums are approximately at $f_0 \pm f_s / M$ (see 2) we only allow a frequency change of $\pm f_s / M$. This will be enough to get closer to the absolute minimum and run the algorithm with good results.

3.3 Experimental results

An example of the application of this method is depicted in Figure 6. An ADC acquired a signal, with an overdrive higher than 70% of full scale and approximately 27 samples per period. The samples corresponding to the saturation of the ADC are ignored when the algorithms are applied. Using the conventional four-parameter algorithm the convergence occurs in a local minimum as shown in Figure 6(a); on the contrary, with our proposed method, with progressive subsets of samples (27, 38, 53 and 97) a successful convergence has been achieved as shown in Figure 6(b).

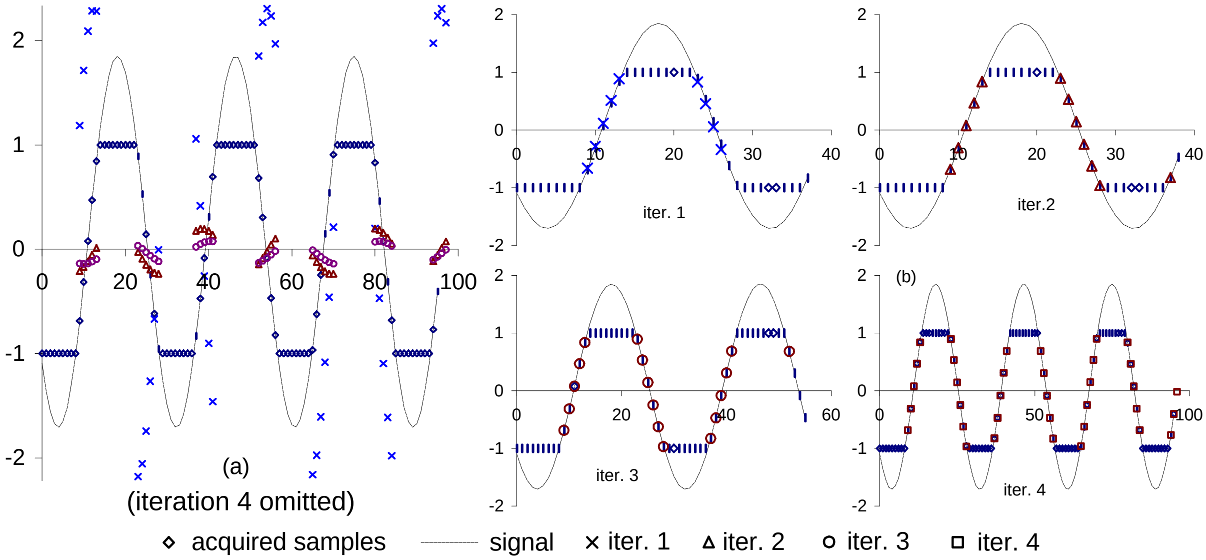

Figure 6 – (a) The use of the conventional four-parameter algorithm produced a bad convergence with the entire number of non-saturated samples. (b) With progressive subsets of samples it was possible to fit a sinewave to the samples, even in presence of “saturation” and with a small number of samples and periods. Note that the 1st iteration with just 27 samples (9 non-saturated) has a very small error.

4 CONCLUSIONS

The three and four-parameter algorithms [2] can perform good results in adverse cases like small number of samples, small number of samples per period, noise and distortion.

However they must be used with some precautions. The importance of accurate frequency estimations and the parameters have been discussed. Even in cases where the implementation of the DFT (or FFT) is difficult (for example, due to extremely long time calculations), one can achieve the convergence starting the algorithms with rough estimations of the initial parameters, using the suggested method, in order to avoid local minimums.

REFERENCES

[1] Fonseca da Silva, M, Serra, A: “A “Study of the sensitivity in an automatic capacitance measurement system”, IEEE IMTC/97, Ottawa, Canada, pp 329-334, 1997.
[2] IEEE Standard for Digitizing Waveform Recorders, IEEE Std 1057-1994.
[3] Haddadi, D., Dallet, D., Marchegay, P., “A Globally Convergent Newton Method for Characterizing A/D Converters”, 4th Workshop on ADC Modeling and Testing, Bordeaux, pp 72-77, 1999.
[4] Project Dynad –SMT4-CT98-2214, Draft standard v1.0, pp 65-74.

AUTHOR(S): Fonseca da Silva, A. Cruz Serra, Lab. Medidas Eléctricas, DEEC,IST/IT, Lisbon Technical University, Portugal, fonseca.silva@lx.it.pt, acserra@alfa.ist.utl.pt