Francisco André Corrêa Alegria*

Number of samples to use in estimating sinewave amplitude in the presence of noise

https://doi.org/10.1515/teme-2023-0118
Received August 20, 2023; accepted October 9, 2023;
published online November 1, 2023

Abstract: The number of samples required to estimate the amplitude of a digitized sinewave depends on the amount of additive noise present, more specifically, on the precision of the estimator used which is depends directly on the additive noise standard deviation. Here an analytical approximate expression for this precision is derived and then used to derive an analytical expression useful in computing the minimum of number of samples that should be acquired to guarantee a given bound on the prevision of the sinewave amplitude estimates.

Keywords: sinewave fitting; uncertainty; ADC; noise; number of samples

1 Introduction

There are different non-ideal factors that influence the precision of a sinewave amplitude estimation using least squares sine-fitting procedures, namely, voltage noise [1, 2], quantization error [3–5], phase noise [6], jitter [7–9], frequency error [10] and harmonic distortion [11] just to mention a few. Arguably one of the most important and most studied factors is the amplitude noise, also called voltage noise or additive noise. The focus here is on this particular type of non-ideality and the effect it has on the precision of sinewave amplitude estimation. Naturally this type of noise has an influence on the precision of all other types of estimators and test procedures like the histogram test of analog-to-digital converters [12–14] or even the tests that determine the amount of noise itself present in a test setup [15, 16].

The goal is thus to derive an expression for the computation of the number of samples required to achieve a precision better than a chosen bound on the precision of the amplitude estimation given the amount of additive noise present. Ideally the expression should be simple and accurate. Since an exact expression for the standard deviation of the estimated amplitude is not known we will settle for an approximate expression which is sufficiently accurate for most practical cases. As it will be seen at the end of this paper, a very simple expression is proposed to be used for the calculation of the minimum number of samples. Monte Carlo simulations were used to validate the accuracy of the expression. This work, however, does not consider other types of noise like phase noise or noise coming from the power supply [17].

There are, however, other non-idealities like phase noise and jitter [11], power supply noise [17, 18] or frequency error [19] that affect the estimation results or even other types of stimulus signals employed [20, 21]. The effect of these is not the subject of this paper. Naturally this non-ideal phenomena are not exclusive to the estimation of sinewave amplitude but occur in most engineering domains like estimating the flow velocity [22, 23] or measuring distance using acoustic waves [24], just to mention a few.

2 Sinewave fitting

Sinewave fitting is a procedure used to estimate three parameters of a sinewave, namely the amplitude, initial phase and average value that best fit, in a least-squares error sense, a set of supplied data points. There are more sophisticated algorithms that are even able to estimate the frequency of the signal if it is unknown (four-parameter sine fitting).

Consider M data points $z_1, z_2, \ldots, z_M$ given by

$$z_i = C + A \cdot \cos\left(\omega_x t_i + \varphi\right). \quad (1)$$

where φ is the initial phase and ω_x is the angular frequency ($2\pi f_x$). We consider the phase φ to be a random variable uniformly distributed in an interval with length 2π.

This data is affected by additive voltage white Gaussian noise, d_i, with null mean and standard deviation σ_v:

$$y_i = z_i + d_i = C + A \cdot \cos\left(\omega_x t_i + \varphi\right) + d_i. \quad (2)$$

***Corresponding author: Francisco André Corrêa Alegria**, Instituto de Telecomunicações and Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal, E-mail: falegria@lx.it.pt. https://orcid.org/0000-0003-0854-489X

We wish to estimate the sine wave that best fits, in a least square error sense, to these M points. The estimates of the sine wave are obtained, in a matrix form, with [16]

$$\begin{bmatrix} \widehat{A_I} \\ \widehat{A_Q} \\ \hat{C} \end{bmatrix} = \left(D^T D\right)^{-1} D^T \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_M \end{bmatrix} \tag{3}$$

with

$$D = \begin{bmatrix} \cos(\omega_a t_1) & \sin(\omega_a t_1) & 1 \\ \cos(\omega_a t_2) & \sin(\omega_a t_2) & 1 \\ \cdots & \cdots & \cdots \\ \cos(\omega_a t_M) & \sin(\omega_a t_M) & 1 \end{bmatrix}, \tag{4}$$

where ω_a is the angular frequency of the sinusoid we are trying to adjust to the data, $\widehat{A_I}$ and $\widehat{A_Q}$ are the in-phase and in-quadrature amplitudes and $\hat{C}$ is the estimated average value (offset) The amplitude of the sinewave can be obtained with

$$\hat{A} = \sqrt{A_I^2 + A_Q^2}. \tag{5}$$

In the case of coherent sampling (known frequency) one has

$$\hat{A} = \sqrt{\frac{4}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} y_i y_j \cos\left[\omega_a (t_1 - t_j)\right]}. \tag{6}$$

In this work we focus on the standard deviation of the estimated amplitude, $\sigma_{\hat{A}}$.

3 Precision of amplitude estimation

In Ref. [25] the expected value of the estimated amplitude was determined using a 3rd order Taylor series approximation to the square root function found in (6). The nonlinearity of this function is what makes the derivation of an exact analytical expression difficult. The result obtained in Ref. [25] was thus an approximate expression, specifically

$$\mu_{\hat{A}} \approx \sqrt{A^2 + \frac{4}{M}\sigma_v^2} - \frac{\frac{16}{M^2}\sigma_v^4 + \frac{8}{M}\sigma_v^2 A^2}{8\left(A^2 + \frac{4}{M}\sigma_v^2\right)^{3/2}}. \tag{7}$$

Here a similar procedure will be used to derive an expression for the standard deviation of the estimate. From [26] (p. 113) one can write

$$\sigma_{\hat{A}}^2 \approx \left(\left| \frac{\partial \hat{A}}{\partial \widehat{A^2}} \right|_{\widehat{A^2} = \mu_{\widehat{A^2}}} \right)^2 \cdot \sigma_{\widehat{A^2}}^2. \tag{8}$$

The derivative in (8) is the derivative of the square root function which is

$$\left. \frac{\partial \hat{A}}{\partial \widehat{A^2}} \right|_{\widehat{A^2} = \mu_{\widehat{A^2}}} = \frac{1}{2\sqrt{\mu_{\widehat{A^2}}}}. \tag{9}$$

The expected value of the square amplitude is, from [25] (Eq. (27)),

$$\mu_{\widehat{A^2}} = A^2 + \frac{4}{M}\sigma_v^2. \tag{10}$$

Using this and the expression for the variance of the square amplitude estimation, determined in Ref. [25] (Eq. (52)), which was

$$\sigma_{\widehat{A^2}}^2 = \frac{16}{M^2}\sigma_v^4 + \frac{8}{M}\sigma_v^2 A^2, \tag{11}$$

one gets

$$\sigma_{\hat{A}}^2 \approx \left(\frac{1}{2\sqrt{A^2 + \frac{4}{M}\sigma_v^2}} \right)^2 \left(\frac{16}{M^2}\sigma_v^4 + \frac{8}{M}\sigma_v^2 A^2 \right). \tag{12}$$

This expression can be rewritten in a more appealing form as

$$\sigma_{\hat{A}}^2 \approx \frac{\frac{16}{M^2}\sigma_v^4 + \frac{8}{M}\sigma_v^2 A^2}{4A^2 + \frac{16}{M}\sigma_v^2}. \tag{13}$$

The standard deviation is thus, after some algebraic manipulation given by,

$$\sigma_{\hat{A}} \approx \frac{\sqrt{2}\sigma_v}{\sqrt{M}} \sqrt{\frac{1 + \frac{1}{M}\left(\frac{\sqrt{2}\sigma_v}{A}\right)^2}{1 + \frac{2}{M}\left(\frac{\sqrt{2}\sigma_v}{A}\right)^2}}. \tag{14}$$

We see that the standard deviation of the estimated amplitude depends in a complicated way on the additive noise standard deviation, σ_v, the number of acquired samples, M, and the actual signal amplitude A. Defining, for convenience, the signal to noise ratio as

$$SNR = \frac{A}{\sqrt{2}\sigma_v}, \tag{15}$$

we can write the expression for the standard deviation of the estimated sinewave amplitude as

$$\sigma_{\hat{A}} \approx \frac{A}{SNR\sqrt{M}} \sqrt{\frac{1 + M \cdot SNR^2}{2 + M \cdot SNR^2}}. \tag{16}$$

This is an approximate expression derived from the 3rd order Taylor series approximation to a square root function. One can further simplify this expression for the case where the product of number of samples and squared signal to noise ratio is much greater than 2, which is the situation found in most practical cases, leading to

$$\sigma_{\hat{A}} \approx \frac{A}{SNR\sqrt{M}}, \tag{17}$$

which is a fairly simple expression.

To facilitate comparison, we present here also the equivalent expression using the noise standard deviation instead of the SNR. Making use of (15) leads to

$$\sigma_{\hat{A}} \approx \sigma_v \sqrt{\frac{2}{M}}. \tag{18}$$

We have reached an expression that is identical to the Cramér–Rao Lower Bound as given in Section 3.1 of [27] (Eq. (3.41)). Note that this estimator is, however, biased as seen in Ref. [25].

4 Validation

To validate these two expressions presented here, a Monte Carlo Analysis was carried out. A set of M data points sampled from a sinewave with amplitude 2 V and a chosen frequency (f_x) at a specific rate (f_s) such that

$$\frac{f_s}{f_x} = M, \tag{19}$$

so that the data covered exactly one period of the signal, that is, guaranteeing coherent sampling. To each data point was added a random value taken from a normal distribution with a standard deviation that ranged from 0 (no noise) to 4 V (double the sinusoidal amplitude).

A three-parameter sine fitting least squares estimation of the sinewave amplitude, $\hat{A}$, was made as described earlier. This was repeated 10,000 times (R) and the standard deviation of those amplitude estimations was computed. The result can be seen in Figure 1 together with the values given by first theoretical expression, Eq. (16), using a solid line, and those given by the second theoretical expression, (17), using a dashed line.

From the results one concludes that both expressions give approximate values, as expected, and that expression (16), the more complex one, is in fact a better approximation, with values very close to the real values which are inside the vertical bars represented. Those bars have a with corresponding to an interval with a 99.9 % confidence level in the case of normally distributes values. The values given by (17) give a worst approximation in general but provide a very good approximation for the case of small additive noise standard deviation (smaller than 1 V in the simulation represented in Figure 1).

In the previous section we mentioned that this estimator is biased as shown in Ref. [25]. It is interesting to show, for such a low number of samples as the one used in Figure 1, just 3, the behavior of the estimation amplitude bias as a function of the additive noise standard deviation. This is done in Figure 2. For small amounts of additive noise (left side of chart) the bias is not significant. For very high amounts of additive noise as the ones shown in the chart (with a standard deviation up to twice the signal amplitude) the bias is considerable.

In the simulation in Figure 1 the extreme case of the number of samples equal to 3 was considered. This is the minimum number of samples possible and the worst case in terms of accuracy of the approximation made. In Figure 3 a more usual case encountered in practice was considered, namely, the use of 50 samples. Has can be seem, the approximation given by (16) has practically no error (the solid line in inside the vertical bars) while the approximation given by (17) improved considerably. Naturally both approximations begin to fail more severely if the amount of additive noise standard deviation would increase considerable beyond the

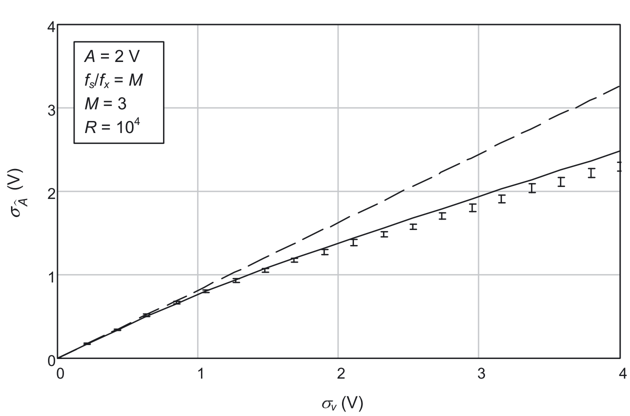

Figure 1: Standard deviation estimated sine wave amplitude as a function of the additive noise standard deviation when 3 samples are acquired. The vertical bars represent the values obtained with the Monte Carlo analysis using a confidence level of 99.9 %. The solid line represents the theoretical value given by (16) and the dashed line represents the theoretical value given by (17).

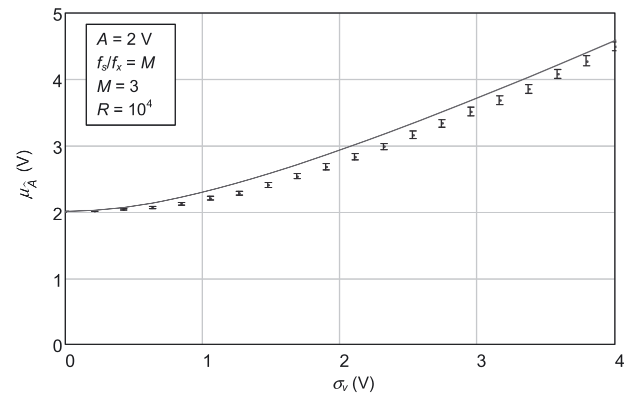

Figure 2: Estimated sine wave amplitude as a function of the additive noise standard deviation when 3 samples are acquired. The vertical bars represent the values obtained with the Monte Carlo analysis using a confidence level of 99.9 %. The solid line represents the theoretical values given by (7).

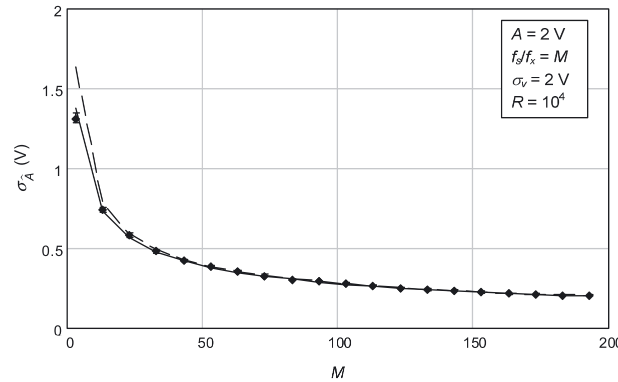

Figure 4: Standard deviation estimated sine wave amplitude as a function of the number of samples. The vertical bars represent the values obtained with the Monte Carlo analysis using a confidence level of 99.9 %. The solid line represents the theoretical value given by (16) and the dashed line represents the theoretical value given by (17).

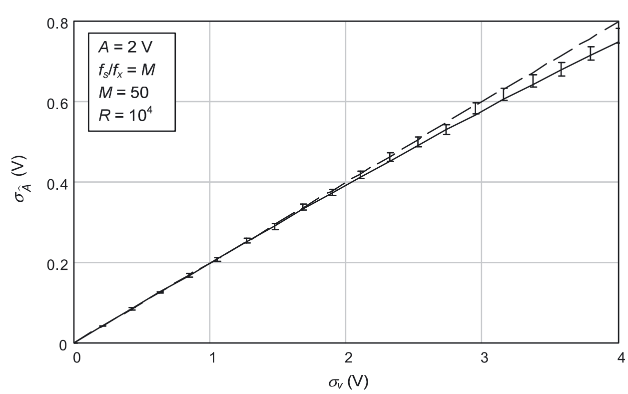

Figure 3: Standard deviation estimated sine wave amplitude as a function of the additive noise standard deviation when 50 samples are acquired. The vertical bars represent the values obtained with the Monte Carlo analysis using a confidence level of 99.9 %. The solid line represents the theoretical value given by (16) and the dashed line represents the theoretical value given by (17).

values depicted in the figure which are themselves already quite large going up to twice the signal amplitude.

In Figure 4 the value of amplitude estimation standard deviation is represented as a function of the number of samples. It can be seen that it decreases when the number of samples increase, which was expected by observing expression (17), and that both approximations presented here become very good for higher number of samples. This data corresponds to a situation where the noise standard deviation is equal to the signal amplitude: $\sigma_v = A$.

This analysis as shown that the approximate analytical expression (17) derived is a good approximation when the number of samples is not excessively low and for a fairly large range of values of noise standard deviation, justifying its use.

5 Minimum number of samples

Having obtained a useful analytical expression for the standard deviation of the amplitude estimator it is now possible to derive a way to compute the minimum number of samples required to guarantee that the standard deviation of the estimated sine wave amplitude is lower than a given bound B (still considering coherent sampling):

$$\sigma_{\hat{A}} \leq B_{\sigma_{\hat{A}}}. \tag{20}$$

Inserting (14) leads to

$$\frac{\sqrt{2}\sigma_v}{\sqrt{M}} \sqrt{\frac{1 + \frac{1}{M}\left(\frac{\sqrt{2}\sigma_v}{A}\right)^2}{1 + \frac{2}{M}\left(\frac{\sqrt{2}\sigma_v}{A}\right)^2}} \leq B_{\sigma_{\hat{A}}}. \tag{21}$$

Solving in order to M leads to

$$M \geq \frac{\sigma_v^2}{A^2} \times \frac{1 - 2\left(\frac{B_{\sigma_{\hat{A}}}}{A}\right)^2 + \sqrt{1 + 4\left(\frac{B_{\sigma_{\hat{A}}}}{A}\right)^4}}{\left(\frac{B_{\sigma_{\hat{A}}}}{A}\right)^2}. \tag{22}$$

There is another solution that leads to negative values of M which thus is not relevant.

Looking at (22) in order to simplify it, one may consider the case where the value of the bound $B_{\sigma_{\hat{A}}}$ is much smaller than the amplitude of the sine wave (A) which is what happens usually. In those condition (22) becomes

$$M \geq 2\frac{\sigma_v^2}{B_{\sigma_{\hat{A}}}^2}. \tag{23}$$

From (23) one concludes, as expected, that the number of samples required increases with the amount of additive noise present and when the bound which was set for the amplitude estimation precision becomes more stringent (lower values of $B_{\sigma_{\hat{A}}}$).

For example, for a case where the sinewave amplitude is 2 V, the noise standard deviation is 0.5 V, and the bound in the amplitude estimation standard deviation is 0.01 V, expression (22) gives a value for the minimum number of samples of 4999.875 while expression (23) gives a value of 5000 which, since the number of samples has to be integer, is essentially the same result. Other examples could be given but the conclusion would be the same. There is no added value in using (22) instead of (23).

Since sometimes the amount of noise present in a given circuit is expressed using the signal-to-noise ratio (SNR), as was done in Section 3, we can also supply, for convenience, an expression for the minimum number of samples as a function of SNR. Making use of (15) in (22) leads to

$$M \geq \left(\frac{A}{SNR \cdot B_{\sigma_{\hat{A}}}}\right)^2. \tag{24}$$

The minimum number of samples to use to guarantee a maximum standard deviation of the sinewave amplitude estimation of $B_{\sigma_{\hat{A}}}$ is directly proportional to the square of the signal amplitude and inversely proportional to the square of signal-to-noise ratio and the square of the desired bound.

The goal set out in the beginning of this paper – to derive an expression for the minimum number of samples necessary – has thus been achieved. The expression proposed is (23), as a function of noise standard deviation or (23) as a function of signal-to-noise ratio.

6 Conclusions

An expression is proposed for the computation of the minimum number of samples required to guarantee that the standard deviation of sinewave amplitude estimation using least squares three-parameter sine fitting algorithms is lower than a given bound. This knowledge leads to efficient data acquisition since a higher than required number of samples leads to more time wasted in signal processing and higher sampling rates.

The mathematical approximations made in order to arrive at simple, easy to use, expressions for the computation of the minimum number of samples to use, were numerically validated using a Monte Carlo procedure.

As expected, the greater the amount of noise present, the more samples need to be acquired. What might not be expected is that this dependence is quadratic. In other ADC test methods, like the histogram test method, for example, that dependence is linear [28].

Only the effect of additive noise was considered. Other non-ideal effects also lead to constrains in the minimum of samples that should be used. The consideration of those effects is thus important and should be done in the future.

Research ethics: Not applicable.
Author contributions: The author have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The author states no conflict of interest.
Research funding: This work was supported in by Fundação para Ciência e a Tecnologia under the research projects UIDB/50008/2020 and FCT.CPCA.2022.01.
Data availability: Not applicable.

References

[1] I. Kollar and J. J. Blair, “Improved determination of the best fitting sine wave in ADC testing,” *IEEE Trans. Instrum. Meas.*, vol. 54, no. 5, pp. 1978—1983, 2005.
[2] F. C. Alegria and A. C. Serra, “Uncertainty of the estimates of sine wave fitting of digital data in the presence of additive noise,” in *2006 IEEE Instrumentation and Measurement Technology Conference Proceedings*, 2006.
[3] N. Björsell and P. Händel, “Truncated Gaussian noise in ADC histogram tests,” *Measurement*, vol. 40, no. 1, pp. 36—42, 2007.
[4] F. C. Alegria, P. Girão, V. Haasz, and A. C. Serra, “Performance of data acquisition systems from the user’s point of view”, *IEEE Trans. Instrum. Meas.*, vol. 53, no. 4, pp. 907—914, 2004.
[5] F. C. Alegria, P. Girão, V. Haasz, and A. Serra, “Performance of data acquisition systems from user’s point of view,” in *IEEE Instrumentation and Measurement Technology Conference*, Vail, Colorado, USA, 2003, pp. 940—945.
[6] D. B. Leeson, “Oscillator phase noise: a 50-year review,” *IEEE Trans. Ultrason. Ferroelectrics Freq. Control*, vol. 63, no. 8, pp. 1208—1225, 2016.
[7] F. C. Alegria and A. C. Serra, “Gaussian jitter-induced bias of sine wave amplitude estimation using three-parameter sine fitting,” *IEEE Trans. Instrum. Meas.*, vol. 59, no. 9, pp. 2328—2333, 2010.

[8] E. Balestrieri, F. Picariello, S. Rapuano, and I. Tudosa, “Review on jitter terminology and definitions,” *Measurement*, vol. 145, pp. 264—273, 2019.
[9] S. Shariat-Panahi, F. C. Alegria, A. Mànuel, and A. C. Serra, “IEEE 1057 jitter test of waveform recorders,” *IEEE Trans. Instrum. Meas.*, vol. 58, no. 7, pp. 2234—2244, 2009.
[10] F. C. Alegria and A. C. Serra, “Influence of frequency errors in the variance of the cumulative histogram,” *IEEE Trans. Instrum. Meas.*, vol. 50, no. 2, pp. 461—464, 2001.
[11] F. C. Alegria, A. Moschitta, P. Carbone, A. C. Serra, and D. Petri, “Effective ADC linearity testing using sinewaves,” *IEEE Trans. Circuits Syst. I: Regul. Pap.*, vol. 52, no. 7, pp. 1267—1275, 2005.
[12] F. C. Alegria and A. C. Serra, “Standard histogram test precision of ADC gain and offset error estimation,” *IEEE Trans. Instrum. Meas.*, vol. 56, no. 5, pp. 1527—1531, 2007.
[13] R. S. Gamad and D. K. Mishra, “Gain error, offset error and ENOB estimation of an A/D converter using histogram technique,” *Measurement*, vol. 42, no. 4, pp. 570—576, 2009.
[14] A. C. Serra, F. C. Alegria, L. Michaeli, P. Michalko, and J. Saliga, “Fast ADC testing by repetitive histogram analysis,” in *2006 IEEE Instrumentation and Measurement Technology Conference Proceedings*, 2006.
[15] F. C. Alegria and A. C. Serra, “Uncertainty of ADC random noise estimates obtained with the IEEE 1057 standard test,” *IEEE Trans. Instrum. Meas.*, vol. 54, no. 1, pp. 110—116, 2005.
[16] IEEE, “IEEE standard for digitizing waveform recorders,” *IEEE Std.*, pp. 1057—2017, 2018. https://doi.org/10.1109/IEEESTD.2018.8291741.
[17] E. Martinho, F. C. Alegria, A. Dionisio, C. Grangeia, and F. Almeida, “3D resistivity imaging and distribution of water-soluble salts in Portuguese renaissance stone bas-reliefs,” *Eng. Geol.*, vol. 141, pp. 33—44, 2012.
[18] M. Tehranipoor and K. M. Butler, “Power supply noise: a survey on effects and research,” *IEEE Des. Test Comput.*, vol. 27, no. 2, pp. 51—67, 2010.
[19] F. A. C. Alegria and A. M. da Cruz Serra, “Influence of frequency errors in the variance of the cumulative histogram,” *IEEE Trans. Instrum. Meas.*, vol. 50, no. 2, pp. 461—464, 2001.
[20] A. C. Serra, F. C. Alegria, R. Martins, and M. F. da Silva, “Analog to digital converters testing — new proposals,” *Comput. Stand. Interfac.*, vol. 26, no. 1, pp. 3—13, 2004.
[21] F. C. Alegria, P. Arpaia, P. Daponte, and A. C. Serra, “ADC histogram test using small-amplitude input waves,” in *XVI IMEKO World Congress*, vol. X, Austria, 2000, pp. 9—14.
[22] O. B. Canelas, R. M. L. Ferreira, S. Guillén-Ludeña, F. C. Alegria, and A. H. Cardoso, “Three-dimensional flow structure at fixed 70 open-channel confluence with bed discordance,” *J. Hydraul. Res.*, vol. 58, no. 3, pp. 434—446, 2020.
[23] O. Birjukova, S. G. Ludeña, F. Alegria, and A. H. Cardoso, “Three dimensional flow field at confluent fixed-bed open channels,” in *Proceedings of the Riverflow 2014*, 2014, pp. 1007—1014.
[24] F. Gueuning, M. Varlan, C. Eugene, and P. Dupuis, “Accurate distance measurement by an autonomous ultrasonic system combining time-of-flight and phase-shift methods,” in *Quality Measurement: The Indispensable Bridge Between Theory and Reality (No Measurements? No Science! Joint Conference — 1996: IEEE Instrumentation and Measurement Technology Conference and IMEKO Tec*, vol. 1, Brussels, Belgium, 1996, pp. 399—404.
[25] F. C. Alegria, “Bias of amplitude estimation using three-parameter sine fitting in the presence of additive noise,” *Measurement*, vol. 42, no. 5, pp. 748—756, 2009.
[26] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York, USA, McGraw-Hill, 1991.
[27] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, New Jersey, USA, Prentice Hall, 1998.
[28] F. C. Alegria and A. C. Serra, “Uncertainty in the ADC transition voltages determined with the histogram method,” in *Proceedings of the 6th Workshop on ADC Modeling and Testing*, Lisbon, 2001.

Bionotes

Francisco André Corrêa Alegria
Instituto de Telecomunicações and Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal
falegria@lx.it.pt
https://orcid.org/0000-0003-0854-489X

Francisco André Corrêa Alegria was born in Lisbon, Portugal, on July 2, 1972. He received the Diploma, M.S., and Ph.D. degrees in electrical engineering and computers from the Instituto Superior Técnico (IST), Technical University of Lisbon, in 1995, 1997, and 2002, respectively. Since 1994, he has been a member of the instrumentation and measurement research line at the Instituto de Telecomunicações, Technical University of Lisbon. Since 1997, he has been a member of the teaching and research staff of IST where he is now Associate Professor with Habilitation. His current research interests include analog-to-digital-converter characterization techniques, automatic measurement systems, and computer vision.