ACTA IMEKO
ISSN: 2221-870X
December 2025, Volume 14, Number 4, 1 - 8

Unexpected sine-fitting residual RMS bias on additive noise-corrupted data points

Francisco Alegria[1]

[1] *Instituto de Telecomunicações and Instituto Superior Técnico, University of Lisbon, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal*

ABSTRACT
This work presents a novel result in signal processing theory, addressing the statistical properties of least squares estimation when fitting sinusoidal models to noise-corrupted data—a fundamental operation in numerous signal processing applications. We rigorously demonstrate the previously unrecognized estimation bias in the root mean square (RMS) of residuals when processing signals with additive Gaussian white noise, even when the sinusoidal frequency is known. Our theoretical framework derives a closed-form expression for this bias. The analytical derivations are validated through comprehensive Monte Carlo simulations. This work contributes to current trends in robust signal parameter estimation, uncertainty quantification, and performance analysis of signal processing algorithms under non-ideal conditions—essential considerations for applications in communications, radar, sonar, audio, biomedical signal processing, and measurement.

Section: RESEARCH PAPER

Keywords: additive noise; estimation bias; estimation uncertainty; Monte Carlo methods; numerical validation; sinusoidal parameter estimation

Citation: F. Alegria, Unexpected sine-fitting residual RMS bias on additive noise corrupted data points, Acta IMEKO, vol. 14 (2025) no. 4, pp. 1-8. DOI: 10.21014/actaimeko.v14i4.2139

Section Editor: Francesco Lamonaca, University of Calabria, Italy

Received June 28, 2025; **In final form** December 2, 2025; **Published** December 2025

Funding: This work is funded by FCT/MECI through national funds and when applicable co-funded EU funds under UID/50008: Instituto de Telecomunicações.

Corresponding author: Francisco Alegria, e-mail: falegria@lx.it.pt

1. INTRODUCTION

Signal parameter estimation in the presence of noise represents one of the most fundamental challenges in modern signal processing, with far-reaching implications across numerous scientific and engineering disciplines. Within this domain, sinusoidal parameter estimation stands as a cornerstone technique, underpinning critical applications from high-precision metrology to advanced sensing systems, communications, and biomedical signal analysis. Despite decades of research, some theoretical gaps remain in our understanding of how noise affects estimation performance—gaps that this work addresses through rigorous mathematical analysis and comprehensive validation.

The root mean square (RMS) value of residuals obtained from sine fitting serves as a critical metric in signal processing and metrology, functioning as an essential quantitative measure for assessing measurement quality, characterizing system performance, and detecting underlying signal distortions. This has been repeatedly emphasized in ADC testing and calibration studies [1]–[3]. Its importance extends across multiple high-impact domains, including analogue-to-digital converter (ADC) testing, sensor calibration, vibration analysis, biomedical signal processing, and communications systems.

Understanding the expected value of the RMS of residuals is crucial for several compelling reasons that directly impact signal processing applications. First, it provides a statistically sound benchmark against which the actual RMS of residuals from a fit can be compared, enabling robust assessment of goodness-of-fit metrics. When the observed RMS deviates significantly from its expected value, it reliably indicates critical issues such as model misspecification, non-Gaussian noise characteristics, or incorrect assumptions about noise variance—insights that can prevent costly errors in system design and implementation. Furthermore, this expected value is intrinsically linked to the known noise level in the data, offering valuable insights into how effectively a model can extract the underlying signal in challenging noise environments.

The residual RMS captures the deviation between the measured signals and the ideal fitted sine waves, providing a direct measure of noise and nonlinear distortions in the system.

Previous works have shown its central role in determining signal-to-noise-and-distortion ratio (SINAD) and effective number of bits (ENOB) [1], [2], and in addressing subtle estimation biases [3], [4]. This capability is essential for ADC performance evaluation, where residual RMS is used to compute SINAD and ENOB, which are key metrics codified in international standards like IEEE Std. 1241 for ADC testing [5], for example.

The applications of our theoretical framework extend to numerous sensor technologies that benefit from precise residual RMS estimation. In microelectromechanical systems (MEMS) accelerometers and gyroscopes, sinusoidal excitation is the standard approach for calibration and performance characterization. The accurate estimation of noise floors and nonlinearities in these sensors is critical for applications ranging from consumer electronics to precision navigation systems. When these sensors undergo sinusoidal motion, the RMS of residuals between the measured output and the fitted sinusoidal models provides crucial information about sensor noise density—directly impacting resolution capability—and cross-axis sensitivity that manifests as additional sinusoidal components in the residuals. In MEMS gyroscopes deployed in autonomous vehicles and drones, accurate characterization of noise through RMS residual analysis enables more reliable attitude estimation and motion control [6].

Pressure sensors and microphones frequently encounter sinusoidal pressure variations in applications ranging from industrial process monitoring to acoustic measurements. The analysis of sinusoidal fitting residuals in these contexts enables precise characterization of sensor bandwidth limitations, which appear as amplitude and phase distortions at higher frequencies. Our theoretical framework enhances the identification of resonance effects that create frequency-dependent deviations [7] from expected responses and improves the quantification of environmental noise contributions to measurement uncertainty—critical factors in designing robust sensing systems.

In underwater acoustic sensors and hydrophones, where sinusoidal calibration signals are standard practice, the expected RMS residuals derived from our work serve as a reference point for detecting biofouling or physical damage that alters sensor response characteristics. Similarly, in industrial pressure sensors subjected to cyclic loading, changes in the RMS residuals over time can indicate membrane fatigue or other degradation mechanisms before catastrophic failure occurs [8]—a capability with significant implications for predictive maintenance and system reliability.

Recent experimental results demonstrate that adaptive filtering based on least mean square error algorithms can extract higher-order sinusoidal harmonic signals from strong background noise along optical fibres, enhancing the signal-to-noise ratio of distributed optical fibre vibration sensors based on coherent optical time domain reflectometry [9].

By integrating sine fitting with mean square error analysis, researchers have achieved substantial reductions in calibration uncertainty, enhancing the reliability of vibration measurements in challenging noise environments. This methodology is particularly valuable for applications requiring high-precision sensor calibration, such as structural health monitoring and aerospace engineering [10]. Our work extends these capabilities by providing the theoretical underpinnings that enable more accurate uncertainty quantification.

Despite the wide use of sine fitting in metrology [11]–[13] and uncertainty analysis [14]–[16], no prior work has provided a closed-form expression for the residual RMS bias. The novel contribution of this work lies in providing a rigorous theoretical foundation for the expected value of RMS residuals in sinusoidal fitting with additive noise—a foundation that has been lacking despite the widespread use of these techniques. We derive closed-form expressions that predict this bias under various conditions, enabling more accurate interpretation of measurement results and improved system design across multiple domains.

We validate our proposed expressions using comprehensive numerical simulations, employing Monte Carlo procedures. By systematically varying the parameters and assessing the correctness of our analytical derivations across different value ranges, we demonstrate the robustness and broad applicability of our theoretical framework. This validation approach ensures that our findings can be confidently applied across the diverse signal processing applications discussed above.

This paper concludes with insights into the implications of our findings and outlines promising directions for future research related to sinusoidal fitting and parameter estimation uncertainty in the presence of various non-ideal phenomena. These future directions highlight the broader impact of our work on the signal processing community and its potential to inspire new research trajectories.

2. STATE OF THE ART

Sine fitting is a fundamental technique in metrology, signal processing, and instrumentation, used to estimate the parameters (amplitude, frequency, phase, and offset) of a sinusoidal signal from noise-corrupted measurements. The residuals obtained from sine fitting—the differences between the measured data points and the fitted sine wave—play a crucial role in assessing signal quality, estimating noise levels, and computing key performance metrics, such as the signal-to-noise ratio (SNR).

The IEEE Standard 1241-2010 for Analog-to-Digital Converter (ADC) testing [5] relies heavily on sine-fitting algorithms to characterize ADC performance. Several Acta IMEKO papers have also addressed this issue, from algorithmic improvements [2] to amplitude estimation in demanding setups, such as the Planck-Balance [3]. The residuals are used to compute total harmonic distortion (THD), spurious-free dynamic range (SFDR), and SNR. Alegria & Serra [17] demonstrate that residual analysis in sine fitting is critical for correcting bias errors in amplitude and phase estimation, particularly in low-SNR scenarios.

The RMS of sine-fitting residuals is directly related to the noise power, enabling accurate SNR computation. This has been further confirmed in random sampling and interpolation studies [4], [11].

In optical metrology, sine-fitting residuals are used to assess phase noise and jitter in interferometric measurements. Dändliker et al. [18] apply sine-fitting residuals to heterodyne interferometry, improving displacement measurement accuracy.

Sine fitting is used to extract oscillatory components in biomedical signals, where residuals help identify artifacts and noise sources. Stoica et al. [19] discuss frequency-selective residual analysis for detecting pathological patterns in EEG signals. The unexpected bias in residual RMS when fitting noisy sinusoidal data has significant implications. If residuals are biased, SNR calculations become unreliable, leading to incorrect system performance assessments. In metrology, biased residuals can lead to wrong uncertainty estimations, affecting traceability. Systematic residual biases may indicate nonlinearities or unmodeled signal components that require correction.

If a 12-bit ADC is tested using sine fitting, a biased residual RMS could lead to an overestimated SNR, falsely suggesting better performance than reality. This could result in the misclassification of ADC quality in critical applications, like medical imaging or aerospace systems. The analysis of sine-fitting residuals is crucial across multiple disciplines, from metrology to biomedical engineering. Recent works on accelerometer calibration [12], ADC noise testing [13], and power quality uncertainty [14], [15] show that even small biases in RMS values can propagate into uncertainty budgets, underscoring the practical significance of the present contribution. Understanding and correcting residual RMS bias is essential for accurate SNR estimation, distortion analysis, and system calibration. This study addresses a gap in the literature by quantifying and explaining unexpected residual biases, leading to more reliable signal parameter estimation.

The RMS of sine-fitting residuals is not only used for SNR estimation but also serves as the basis for several other critical performance metrics in signal analysis and metrology. SINAD is a widely used metric in ADC testing and communication systems. The noise + distortion power is directly estimated from the residual RMS after sine fitting. IEEE Std 1241-2010 [5] explicitly recommends using sine-fitting residuals for SINAD computation in ADC testing. ENOB is derived from SINAD and indicates the effective resolution of an ADC under dynamic conditions. Vucijak et al. use sine-fitting residuals RMS for electric power calibrations [20]. Kollar proposes an improved residual analysis in ADC testing [21]. Davies and Levenson describe a method to arrive at an unbiased estimate of the RMS and a means to determine the measurement uncertainty, which is highly relevant to understanding the uncertainty of RMS values derived from measurements, including residuals from fitting processes [22]. DeVoe discusses sinusoidal fitting and mentions residuals in the context of measuring Allan variance. While the primary focus is Allan variance, the methods of sinusoidal fitting and the analysis of residuals are pertinent to understanding their uncertainties [23].

3. MATHEMATICAL MODEL

The sinusoidal model under consideration here can be expressed as

$$z(t) = C + A \cdot \cos(2\pi f \cdot t + \varphi) + n(t)\,, \tag{1}$$

where C is the sinusoidal offset, A is the amplitude, f is the frequency and φ is the initial phase. These data points are affected by additive noise described by $n(\mathrm{t})$. We will also consider that this signal is sampled at a constant rate f_S, such that the sampling instants are given by

$$t_i = \frac{i}{f_s}, i = 0, 1, \dots, M-1\,, \tag{2}$$

where M is the number of samples. The sampled values of our signal are thus given by

$$z_i = C + A \cdot \cos(2\pi f \cdot t_i + \varphi) + n_i. \tag{3}$$

Note that the contribution of the additive noise to each sample is represented by n_i.

The model parameters, excluding frequency, can be estimated using closed-form expressions. In order to achieve this, however, it is more convenient to express the sinusoidal model, which is non-linear in φ, as

$$z_i = C + A_\mathrm{I} \cdot \cos(2\pi f \cdot t_i) + A_\mathrm{Q} \cdot \sin(2\pi f \cdot t_i) + n_i, \tag{4}$$

where these two new parameters, A_I and A_Q are related to the sine wave amplitude and initial phase by

$$A_\mathrm{I} = A \cdot \cos(\varphi) \tag{5}$$

and

$$A_\mathrm{Q} = A \cdot \sin(\varphi). \tag{6}$$

Our model thus becomes linear in the unknown parameters C, A_I and A_Q, which can be estimated from the data points, z_i using

$$\widehat{A_\mathrm{I}} = \frac{2}{M}\sum_{i=0}^{M-1} z_i \cdot \cos(\omega \cdot t_i)\,, \tag{7}$$

$$\widehat{A_\mathrm{Q}} = \frac{2}{M}\sum_{i=0}^{M-1} z_i \cdot \sin(\omega \cdot t_i), \tag{8}$$

and

$$\hat{C} = \frac{1}{M}\sum_{i=0}^{M-1} z_i\,. \tag{9}$$

From these two amplitude estimates, usually called in-phase ($\widehat{A_\mathrm{I}}$) and in-quadrature ($\widehat{A_\mathrm{Q}}$) amplitudes, one can obtain the sinusoidal amplitude estimative using

$$\hat{A} = \sqrt{\widehat{A_\mathrm{I}}^2 + \widehat{A_\mathrm{Q}}^2}, \tag{10}$$

and the initial phase estimative using

$$\hat{\varphi} = \mathrm{atan}\left(\frac{\widehat{A_\mathrm{Q}}}{\widehat{A_\mathrm{I}}}\right). \tag{11}$$

Note the hat over the symbols to indicate that they are estimated values, which will necessarily be different from the real values due to the presence of noise in the data points.

The difference between the real values, given by (4), and the estimated values, given by

$$\hat{z}_i = \hat{C} + \hat{A} \cdot \cos(2\pi f \cdot t_i + \hat{\varphi}), \tag{12}$$

are called the residuals,

$$e_i = z_i - \hat{z}_i \tag{13}$$

Inserting (3) and (12) leads to

$$\begin{aligned} e_i = {} & [C + A \cdot \cos(2\pi f \cdot t_i + \varphi) + n_i] \\ & - [\hat{C} + \hat{A} \cdot \cos(2\pi f \cdot t_i + \hat{\varphi})]. \end{aligned} \tag{14}$$

The square root of the sum of their squares is called, naturally, the "root mean square" value:

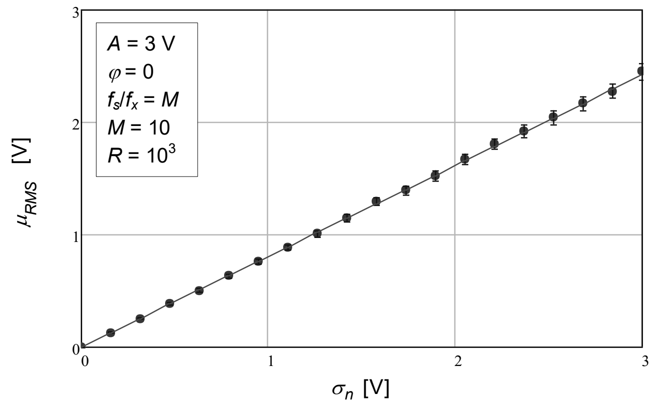

Figure 1. Expected value of the RMS value of the sine-fitting residuals as a function of additive noise standard deviation, determined using numerical simulations. The circles represent the values obtained with the Monte Carlo analysis. The confidence intervals for a confidence level of 99.9 % are represented by the vertical bars. The solid line represents the value given by the theoretical expression (39).

$$\widehat{RMS}_{\text{residuals}} = \sqrt{\frac{1}{M}\sum_{i=0}^{M-1}(z_i - \hat{z}_i)^2}. \tag{15}$$

In the following, we will focus on the bias of this estimated value due to the presence of additive noise.

4. ANALYTICAL DERIVATION

The method of least squares aims to find the parameter estimates that minimize the sum of the squared differences between the observed data points, z_i, and the values predicted by the model $\hat{z}_i$. The key assumptions for the standard statistical results to hold are:

- The model is linear in the parameters being estimated (which we have established for the known frequency case);
- The errors n_i are independent and identically distributed (i.i.d.);
- The errors n_i follow a Gaussian (normal) distribution with a mean of zero;
- The errors n_i have a constant and known variance σ_n^2.

Under these assumptions, the statistical properties of the residuals and their sum of squares can be determined. The sum of squared residuals (SSR), also often denoted as RSS or SSE, is given by

$$\widehat{SSR} = \sum_{i=0}^{M-1}(z_i - \hat{z}_i)^2. \tag{16}$$

For a linear regression model fitting p parameters to M data points, where the errors are Gaussian with mean zero and known variance σ_n^2, a fundamental result from statistical theory states that the quantity SSR/σ_n^2 follows a chi-squared (χ^2) distribution with $M - p$ degrees of freedom:

$$\frac{\widehat{SSR}}{\sigma_n^2} \sim \chi^2(M - p). \tag{17}$$

In our specific case of fitting a sinusoidal model with known frequency by estimating amplitude, phase, and offset, we have $p = 3$ parameters. Therefore,

$$\frac{\widehat{SSR}}{\sigma_n^2} \sim \chi^2(M - 3). \tag{18}$$

The root mean square of the residuals is defined as the square root of the mean squared error (MSE). The MSE is the SSR divided by the number of data points M:

$$\widehat{RMS}_{\text{residuals}} = \sqrt{\frac{\widehat{SSR}}{M}}. \tag{19}$$

We are interested in the expected value of $\widehat{RMS}_{\text{residuals}}$, which, from (19) is

$$\mathrm{E}\{\widehat{RMS}_{\text{residuals}}\} = \sqrt{\frac{1}{M}} \cdot \mathrm{E}\left\{\sqrt{\widehat{SSR}}\right\}. \tag{20}$$

Since

$$\widehat{SSR} = \sigma_n^2 \cdot \frac{\widehat{SSR}}{\sigma_n^2}, \tag{21}$$

where SSR/σ_n^2 is a random variable following the $\chi^2(M - 3)$ distribution, we can write

$$\mathrm{E}\left\{\sqrt{\widehat{SSR}}\right\} = \mathrm{E}\left\{\sqrt{\sigma_n^2 \cdot \chi^2(k)}\right\} = \sigma_n \cdot \mathrm{E}\left\{\sqrt{\chi^2(k)}\right\}, \tag{22}$$

where $k = M - 3$. Introducing it into (20) leads to

$$\mathrm{E}\{\widehat{RMS}_{\text{residuals}}\} = \sigma_n \cdot \sqrt{\frac{1}{M}} \cdot \mathrm{E}\left\{\sqrt{\chi^2(k)}\right\}. \tag{23}$$

The expected value of the square root of a chi-squared random variable with k degrees of freedom, $X \sim \chi^2(k)$, is given by the formula

$$\mathrm{E}\left\{\sqrt{\chi^2}\right\} = \sqrt{2} \cdot \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}, \tag{24}$$

where $\Gamma(x)$ is the Gamma function.

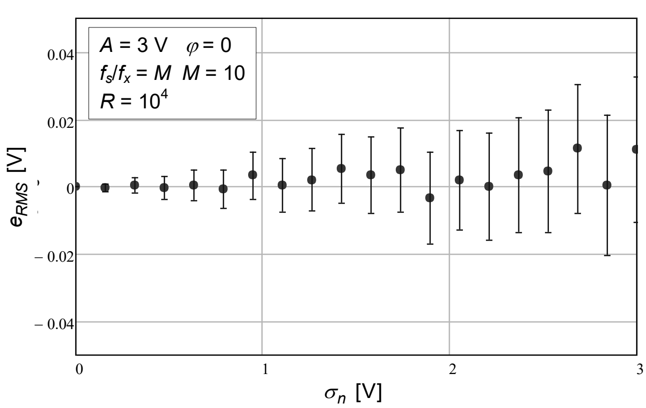

Figure 2. Error of the expected value of the RMS value of the sine-fitting residuals computed as in (41) as a function of additive noise standard deviation, determined using numerical simulations. The circles represent the values obtained with the Monte Carlo analysis. The confidence intervals for a confidence level of 99.9 % are represented by the vertical bars.

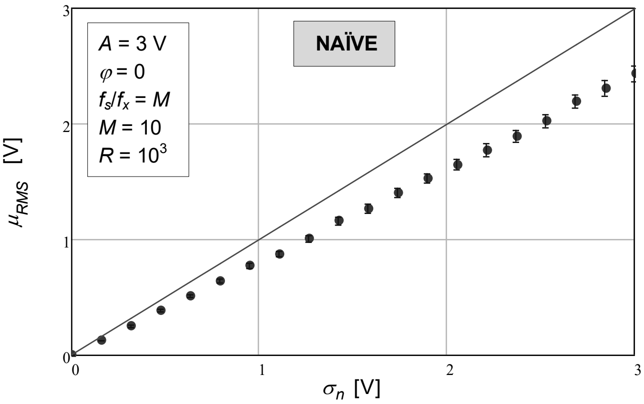

Figure 3. Expected value of the RMS value of the sine-fitting residuals as a function of additive noise standard deviation, determined using numerical simulations. The circles represent the values obtained with the Monte Carlo analysis. The confidence intervals for a confidence level of 99.9 % are represented by the vertical bars. The solid line represents the value given by the theoretical expression (39).

$$\mathrm{E}\left\{\sqrt{\chi^2(M-3)}\right\} = \sqrt{2}\cdot\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}. \tag{25}$$

Therefore, the exact expected value of the RMS of the residuals is

$$\mathrm{E}\left\{\widehat{RMS}_{\text{residuals}}\right\} = \sigma_n\cdot\sqrt{\frac{2}{M}}\cdot\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}. \tag{26}$$

Focusing on the square of this, we have

$$\mathrm{E}^2\left\{\widehat{RMS}_{\text{residuals}}\right\} = \sigma_n^2\cdot\frac{2}{M}\cdot\frac{\Gamma^2\left(\frac{k+1}{2}\right)}{\Gamma^2\left(\frac{k}{2}\right)}. \tag{27}$$

A well-known asymptotic expansion for the ratio of Gamma functions, for large values of the argument x, is

$$\frac{\Gamma(x+a)}{\Gamma(x)} = x^a\cdot\frac{1+a(a-1)}{2\,x} + \frac{a(a-1)(a-2)(3a-1)}{24\,x^2} + \cdots. \tag{28}$$

In our case, $x = k/2$ and $a = 1/2$. Using the first term of this expansion, we have

$$\frac{\Gamma(x+a)}{\Gamma(x)} \approx x^a\cdot\left[1+\frac{a(a-1)}{2\,x}\right]. \tag{29}$$

This leads to

$$\frac{\Gamma\left(\frac{k}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \approx \sqrt{\frac{k}{2}}\cdot\left[1+\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)}{k}\right], \tag{30}$$

which simplifies to

$$\frac{\Gamma\left(\frac{k}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \approx \sqrt{\frac{k}{2}}\cdot\left(1-\frac{1}{4\,k}\right). \tag{31}$$

Taking the square of this approximation leads to

$$\frac{\Gamma^2\left(\frac{k}{2}+\frac{1}{2}\right)}{\Gamma^2\left(\frac{k}{2}\right)} \approx \frac{k}{2}\cdot\left(1-\frac{1}{4\,k}\right)^2. \tag{32}$$

Simplifying leads to

$$\frac{\Gamma^2\left(\frac{k}{2}+\frac{1}{2}\right)}{\Gamma^2\left(\frac{k}{2}\right)} \approx \frac{k}{2}\cdot\left(1-\frac{1}{2\,k}+\frac{1}{16\,k^2}\right) = \frac{k}{2}-\frac{1}{4}+\frac{1}{32\,k}. \tag{33}$$

Inserting this into (27) leads to

$$\mathrm{E}^2\{RMS_{\text{residuals}}\} = \sigma_n^2\cdot\frac{2}{M}\cdot\left(\frac{k}{2}-\frac{1}{4}+\frac{1}{32\,k}\right) = \frac{\sigma_n^2}{M}\cdot\left(k-\frac{1}{2}+\frac{1}{16\,k}\right). \tag{34}$$

For our case, the degrees of freedom are $k = M-3$. Substituting this into the formula and simplifying leads to

$$\mathrm{E}^2\left\{\widehat{RMS}_{\text{residuals}}\right\} = \sigma_n^2\cdot\left(\frac{M-3}{M}-\frac{1}{2\,M}+\frac{1}{16\,M\,(M-3)}\right). \tag{35}$$

Rewriting the first term in parenthesis leads to

$$\mathrm{E}^2\left\{\widehat{RMS}_{\text{residuals}}\right\} = \sigma_n^2\cdot\left(1-\frac{3}{M}-\frac{1}{2\,M}+\frac{1}{16\,M\,(M-3)}\right). \tag{36}$$

Neglecting the last term in the parenthesis, that goes as $1/M^2$, we have

$$\mathrm{E}^2\left\{\widehat{RMS}_{\text{residuals}}\right\} = \sigma_n^2\cdot\left(1-\frac{3}{M}-\frac{1}{2\,M}\right). \tag{37}$$

Simplifying leads to

$$\mathrm{E}^2\left\{\widehat{RMS}_{\text{residuals}}\right\} = \sigma_n^2\cdot\left(1-\frac{3.5}{M}\right). \tag{38}$$

Finally, taking the square root leads to

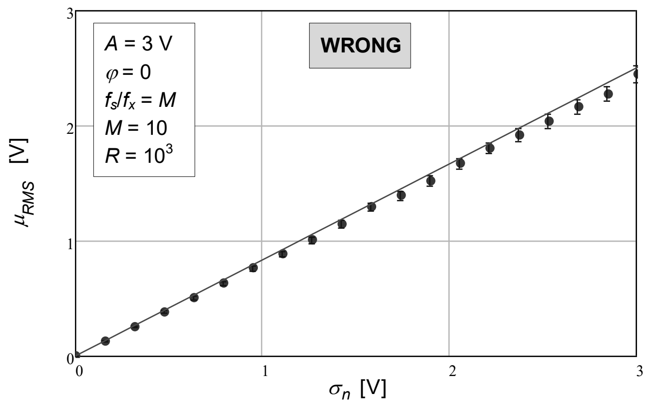

Figure 4. Expected value of the RMS value of the sine-fitting residuals as a function of additive noise standard deviation, determined using numerical simulations. The circles represent the values obtained with the Monte Carlo analysis. The confidence intervals for a confidence level of 99.9 % are represented by the vertical bars. The solid line represents the value given by the theoretical expression (43).

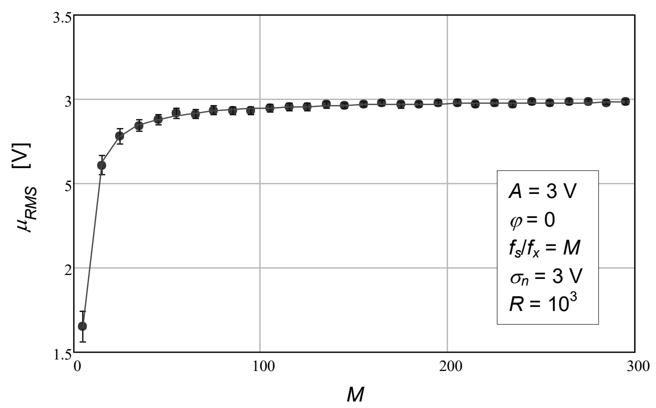

Figure 5. Expected value of the RMS value of the sine-fitting residuals as a function of the number of data points, determined using numerical simulations. The circles represent the values obtained with the Monte Carlo analysis. The confidence intervals for a confidence level of 99.9 % are represented by the vertical bars. The solid line represents the value given by the theoretical expression (39).

$$\mathrm{E}\{\widehat{RMS}_{\text{residuals}}\} = \sigma_n \cdot \sqrt{1 - \frac{3.5}{M}}. \tag{39}$$

This is the unexpected result that motivated this publication. That the expected value of the root mean square of the residuals is proportional to the additive noise standard deviation is expected. That there is a multiplicative correction to that value is also expected, since the RMS computation divides the summation by the number of data points and not by that number minus the number of the degrees of fit, eliminated by the least-squares fitting procedure, which are 3 (amplitude, offset, and initial phase). What is not expected is that we find the numeric constant “3.5” and not “3” (the degrees of freedom) in eq. (39). Naturally, for a large number of samples this discrepancy does not make a difference, since this term is much smaller than 1.

5. NUMERICAL VALIDATION

The results of the analytical derivation require some validation. Here we employed a Monte Carlo procedure to numerically simulate a large number of data points obtained from a sinusoid, where we could control the amount of additive noise introduced. The root mean square value of the residuals was determined and the trials were repeated a large number of times in order to compute the expected value of the determined RMS number. This was then plotted on a chart with error bars indicating the confidence intervals for a confidence level of 99.9 %. Finally, the values corresponding to the theoretical expression determined before, namely,

$$\mu_{\widehat{RMS}_T} = \sigma_n \cdot \sqrt{1 - \frac{3.5}{M}}, \tag{40}$$

were plotted. One can observe the results in Figure 1.

$$e_{RMS} = \mu_{\widehat{RMS}_S} - \mu_{\widehat{RMS}_T}. \tag{41}$$

The result is plotted in Figure 2. As observed, all error bars are around 0.

For comparison purposes, we can plot what a naïve expression that does not account for the reduced degrees of freedom due to sine fitting results in. Equation (42) is plotted in Figure 3.

$$\mu_{\widehat{RMS}_{\text{NAÏVE}}} = \sigma_n. \tag{42}$$

One can also observe in Figure 4 what a wrong expression, like the one in equation (43), results if one simply uses the constant “3” in that equation instead of the correct value of “3.5”.

$$\mu_{\widehat{RMS}_{\text{WRONG}}} = \sigma_n \cdot \sqrt{1 - \frac{3}{M}}. \tag{43}$$

Naturally, the difference is not large and becomes smaller and smaller as the number of points increases. The case represented in Figure 4 is for just 10 points, and even in this case, the difference is very small. In real scenarios, the number of points is usually much larger than this.

One can also observe that the expected value of the RMS changes with the number of points in Figure 5. As this number increases, the expected value of the residuals approaches the correct value, which is 3 V in this example.

In Figure 6, one can see, as expected, by looking at equation (40), that the expected value of the estimated root mean square does not depend on the sinusoidal amplitude.

These numerical simulations show that the analytical expression given in (40) is indeed valid and accounts for the dependence of the expected value of the root mean square on the number of samples and additive noise standard deviation.

6. CONCLUSIONS

In this work, we have derived a comprehensive analytical expression for the expected value of the root mean square (RMS) of sine-fitting residuals as a function of additive noise standard deviation. The closed-form solution presented in equation (39) represents a significant theoretical advancement in signal processing, providing a rigorous mathematical foundation for understanding estimation bias in the presence of noise—a phenomenon with far-reaching implications across numerous application domains.

Our analytical framework has been extensively validated through numerical simulations employing Monte Carlo

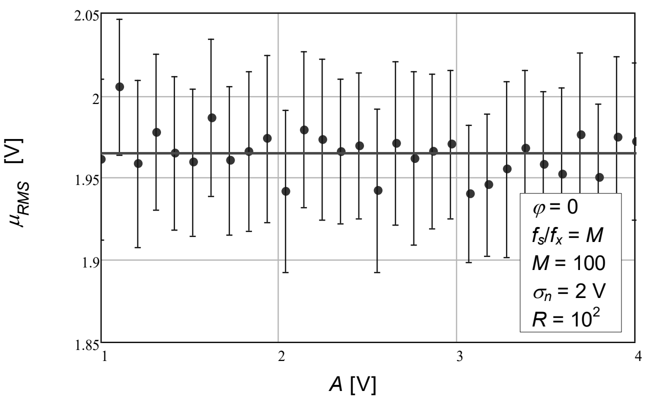

Figure 6. Expected value of the RMS value of the sine-fitting residuals as a function of the sinusoidal amplitude, determined using numerical simulations. The circles represent the values obtained with the Monte Carlo analysis. The confidence intervals for a confidence level of 99.9 % are represented by the vertical bars. The solid line represents the value given by the theoretical expression (39).

procedures, confirming its accuracy across a wide range of operating conditions. The results demonstrate that our expression correctly predicts the complex dependencies between residual RMS, additive noise standard deviation, number of samples, and sinusoidal amplitude. Notably, even for a low number of samples—a common constraint in practical applications—the analytical expression accurately computes the expected value of RMS residuals, making it immediately applicable to real-world signal processing challenges.

The theoretical contributions of this work extend beyond the specific case of sinusoidal parameter estimation, offering insights into the fundamental nature of least squares estimation in the presence of noise. By quantifying the expected bias in residual RMS, we provide a statistical benchmark that enables more accurate assessment of model fit quality and more reliable detection of non-ideal conditions, such as non-Gaussian noise or model misspecification.

Based on our findings, we offer the following practical recommendations for signal processing practitioners:

1) Bias compensation: When using RMS residuals to assess noise levels in sinusoidal signals, practitioners should apply the correction factor derived in this work to avoid systematic underestimation of noise variance, particularly in high-precision applications.
2) Sample size selection: Our analysis reveals that, while the bias effect is present regardless of sample size, its relative impact varies with the number of samples. For applications requiring high accuracy, we recommend using sample sizes that minimize this effect based on the equations provided.
3) Uncertainty quantification: For metrology and calibration applications, our expressions enable more accurate uncertainty budgets by properly accounting for the statistical properties of residuals, leading to more reliable measurement results.
4) Algorithm design: Signal processing algorithm designers should incorporate our findings when developing adaptive filtering or parameter estimation techniques that rely on residual statistics as performance metrics or convergence criteria.

Other digitizer nonidealities, such as integral nonlinearity (INL), differential nonlinearity (DNL), and aperture jitter, can also contribute to residual RMS. In many ADC testing setups, these dominate over the bias we quantify here. Nevertheless, for high-resolution converters (≥ 16 bits) or in metrological applications where linearity is corrected, the residual RMS bias described in this paper may become a non-negligible contributor to the total uncertainty budget.

This study establishes a foundation for several promising research directions. As our immediate next step, we plan to investigate the uncertainty of signal-to-noise ratio estimation in the presence of additive noise, building directly on the theoretical framework established here. Future work will extend these results to more complex signal models, including multi-tone signals, damped sinusoids, and non-stationary signals with time-varying parameters. While the current analysis focuses on Gaussian white noise, extending the framework to handle coloured noise, impulsive noise, and other non-Gaussian distributions represents an important future direction. Investigating how these findings can improve adaptive algorithms for parameter tracking in non-stationary environments could yield significant performance improvements in applications such as biomedical signal processing and communications.

While this study focused on least-squares estimation, alternative approaches, such as maximum likelihood estimation (MLE) [24], provide improved efficiency in the presence of quantization or non-Gaussian noise. The closed-form bias derived here is specific to the LS case, but the methodology could be extended to investigate whether MLE or robust estimators reduce or exacerbate the residual RMS bias under digitizer-relevant noise distributions, such as uniform (quantization) or Laplacian (heavy-tailed). This represents a promising direction for future work.

The analytical framework developed in this paper provides a solid theoretical foundation that will benefit numerous signal processing applications requiring precise parameter estimation in noisy environments. By quantifying and explaining the previously unrecognized bias in residual RMS, this work contributes to more accurate signal characterization, more reliable system performance evaluation, and, ultimately, more robust signal processing systems across scientific and engineering disciplines.

ACKNOWLEDGEMENT

The author wishes to thank the Instituto de Telecomunicações - Lisbon for the facilities provided and the Fundação para a Ciência e a Tecnologia for the funding awarded.

REFERENCES

[1] V. Pálfi, T. Virosztek, I. Kollár, Full information ADC test procedures using sinusoidal excitation, implemented in MATLAB and LabVIEW, Acta IMEKO vol. 4 (2015) no. 3, pp. 4–13.
DOI: 10.21014/acta_imeko.v4i3.257
[2] A. Baccigalupi, G. d'Alessandro, M. D'Arco, R. Schiano Lo Moriello, Least square procedures to improve the results of the three-parameter sine fitting algorithm, Acta IMEKO vol. 4 (2015) no. 2, pp. 100–106.
DOI: 10.21014/acta_imeko.v4i2.173
[3] S. Lin, C. Rothleitner, N. Rogge, T. Fröhlich, Influences on amplitude estimation using the three-parameter sine fitting algorithm in the velocity mode of the Planck-Balance, Acta IMEKO vol. 9 (2020) no. 3, pp. 124–131.
DOI: 10.21014/acta_imeko.v9i3.781
[4] D. Agrež, D. Ilic, J. Drnovsek, Signal and zero padding to improve parameters estimations of sinusoidal signals in the frequency domain, Acta IMEKO vol. 5 (2016) no. 3, pp. 47–54.
DOI: 10.21014/acta_imeko.v5i3.385
[5] IEEE Std 1241-2010: IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters.
DOI: 10.1109/IEEESTD.2011.5692956
[6] S. A. El-Hakem Hegazy, A. M. Kamel, I. I. Arafa, Y. Z. Elhalwagy, MEMS Gyro Noise Estimation and Modeling for Precise Navigation Simulation, 8th Int. Japan-Africa Conf. on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 14-15 December 2020, pp. 84–89.
DOI: 10.1109/JAC-ECC51597.2020.9355936
[7] C. Wang, X. Yi, J. Lv, Q. Peng, The Mechanism of How a Cavity Affects the Fluctuating Pressure Measurement Results of a Static Pressure Probe, Dynamics 5 (2025) 1.
DOI: 10.3390/dynamics5010001
[8] A. Malarkodi, G. Latha, Underwater Electroacoustic Transducer Calibration, in Handbook of Vibroacoustics, Noise and Harshness, N. Garg, C. Gautam, S. Rab, M. Wan, R. Agarwal,

and S. Yadav, Eds., Singapore: Springer, 2025, pp. 1205–1224. DOI: 10.1007/978-981-97-8100-3_67
[9] Y. Wang, J. Zou, Y. Xu, Y. Chen, X. Liu, Q. Bai, B. Jin, Optical Fiber Vibration Sensor Using Least Mean Square Error Algorithm, Sensors 20 (2020) 2000. DOI: 10.3390/s20072000
[10] T. Shimoda, W. Kokuyama, H. Nozato, Precise sinusoidal signal extraction from noisy waveform in vibration calibration, Metrologia 59 (2022) 3, 035010. DOI: 10.1088/1681-7575/ac6cba
[11] D. Agrež, Estimation of the signal component from random equivalent and non-coherent sampling measurements, Acta IMEKO vol. 6 (2017) no. 4, pp. 44–50. DOI:
[12] G. D'Emilia, E. Natale, E. P. Tomasini, Dynamic calibration uncertainty of three-axis low frequency accelerometers, Acta IMEKO vol. 4 (2015) no. 4, pp. 92–98. DOI: 10.21014/acta_imeko.v4i4.239
[13] F. Alegria, Random noise test of analog-to-digital converters, Acta IMEKO vol. 12 (2023) no. 1, pp. 1–7. DOI: 10.21014/actaimeko.v12i2.1502
[14] K. Demerdziev, V. Dimchev, Active power measurement uncertainty modelling in harmonically distorted signals, Acta IMEKO vol. 12 (2023) no. 3, pp. 1–8. DOI: 10.21014/actaimeko.v12i3.1463
[15] B. Velkovski, K. Kogoj, D. Agrež, Evaluating the uncertainty of a virtual power quality disturbance generator and its use in power quality classifier evaluation, Acta IMEKO vol. 12 (2023) no. 3, pp. 1–7. DOI: 10.21014/actaimeko.v12i3.1473
[16] A. Ferrero, H. Vardhana Jetti, S. Ronaghi, S. Salicone, A general Monte-Carlo approach to consider a maximum admissible risk in decision-making procedures based on measurement results, Acta IMEKO vol. 12 (2023) no. 4, pp. 1–7. DOI: 10.21014/actaimeko.v12i4.1602
[17] F. Alegria, A. Cruz Serra, Uncertainty of the Estimates of Sine Wave Fitting of Digital Data in the Presence of Additive Noise, IEEE Instrumentation and Measurement Technology Conference, Sorrento, Italy, 24-27 April 2006, pp. 1643–1647. DOI: 10.1109/IMTC.2006.328187
[18] R. Dändliker, R. Thalmann, D. Prongué, Two-wavelength laser interferometry using superheterodyne detection, Optics Letters vol. 13 (1988) issue 5, pp. 339–341. DOI: 10.1364/OL.13.000339
[19] P. Stoica, J. Li, H. He, Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram, IEEE Transactions on Signal Processing vol. 57 (2009) issue 3, pp. 843–858. DOI: 10.1109/TSP.2008.2008973
[20] N. M. Vucijak, N. M. Radojevic, Three, Four and Seven Parameters Sine-fitting Algorithms Applied in the Electric Power Calibrations, EUROCON 2005 - The Int. Conf. on "Computer as a Tool", Belgrade, Serbia, 21-24 November 2005, pp. 1148–1150. DOI: 10.1109/EURCON.2005.1630156
[21] I. Kollár, Improved Residual Analysis in ADC Testing, 13th Int. Symp. on Measurements for Research and Industry Applications, 9th Workshop on ADC Modelling and Testing, Athens, Greece, 29 September- 1 October 2004, pp. 869–874. Online [Accessed 15 December 2025] https://www.imeko.org/publications/iwadc-2004/IMEKO-IWADC-2004-009.pdf
[22] A. Davies, M. S. Levenson, Estimating the root mean square of a wave front and its uncertainty, Applied Optics vol. 40 (2001) no. 34, pp. 6203–6209. DOI: 10.1364/AO.40.006203
[23] R. G. De Voe, Measuring the Allan variance by sinusoidal fitting, Review of Scientific Instruments vol. 89 (2018) no. 2, 024702. DOI: 10.1063/1.5010140
[24] J. Šaliga, I. Kollár, L. Michaeli, J. Buša, J. Lipták, T. Virosztek, A Comparison of Least Squares and Maximum Likelihood Methods Using Sine Fittings in ADC Testing, Measurement vol. 46 (2013) 10, pp. 4362–4368. DOI: 10.1016/j.measurement.2013.05.004