Standardized sinus-wave fitting algorithms, extensions and applications

Tomas Andersson and Peter Händel
E-mail: tomas.andersson@s3.kth.se and peter.handel@s3.kth.se
Department of Signals, Sensors and Systems
Royal Institute of Technology
SE–100 44 Stockholm, Sweden

***Abstract*— The objective of this paper is to present the statistical properties of the IEEE-STD-1057/IEEE-STD-1241 sine-wave fitting algorithms. The proper Cramér-Rao bound is derived for both the three-parameter and four-parameter algorithms. Further, we investigate the residual, and derive its statistical properties, in both the three- and four-parameter fitting case. In a practical setup were the frequency is *almost* known, we derive an analytical expression for the mean of the residual. By analysis we compare the two methods and state in which case one should use one over the other. Finally, we present some numerical evaluations confirming our analytical findings.**

***Index Terms*— IEEE-STD-1057, IEEE-STD-1241, Cramér-Rao Bound, three-parameter sine-wave fitting, four-parameter sine-wave fitting, ADC testing, parsimony principle, multi sine-wave fitting.**

I. Introduction

Tone frequency estimation has been extensively studied in the literature. An early paper is [1], and a comprehensive list of references can be found in [2].

In testing digital waveform recorders and analog-to-digital converters (ADCs), an important part is to fit a sinusoidal model to the recorded data, as well as to calculate the parameters that in least-squares result in the best fit. Algorithms have been standardized in IEEE Standard 1057 and IEEE Standard 1241 [3], [4]. For easy reference, the three- and four-parameter sine-wave fit algorithms of [4, Sect. 4.1.4] are hereafter denoted as the three- and four-parameter algorithm, respectively. Software implementations of the four-parameter algorithm can be found in [5], [6], and investigations of its performance is the main topic in [7]–[9].

The aim of this paper is to study tone frequency estimation in general, and the performance of the standardized three- and four-parameter algorithms in particular.

II. Signal model and problem set-up

Assume that the data record contains the sequence of samples $\mathbf{y} = (y_1, \ldots, y_N)^T$ taken at time instants $t_1, \ldots, t_N$. It is further assumed that data can be modeled by

$$y_n[\boldsymbol{\vartheta}] = A\cos(\omega t_n) + B\sin(\omega t_n) + C \tag{1}$$

where A, B, C and ω are (known or unknown) constants. Stressing the dependence of $y_n[\boldsymbol{\vartheta}]$ on the generic parameter vector $\boldsymbol{\vartheta}$ turns out to be convenient for the following discussion, where the unknown parameters are gathered in $\boldsymbol{\vartheta}$. Throughout this paper $\boldsymbol{\vartheta}$ represents either the set of three parameters (A, B, C), or the set of four (A, B, C, ω), depending if the frequency ω is known or not. The sine-wave fit problem is solved by minimizing the sum-squared-error

$$V(\boldsymbol{\vartheta}) = \frac{1}{N}\sum_{n=1}^{N}(y_n - y_n[\boldsymbol{\vartheta}])^2 \tag{2}$$

with respect to the unknown parameters $\boldsymbol{\vartheta}$. Consider the signal model where the measurements are described by

$$y_n = y_n[\boldsymbol{\vartheta}_0] + e_n \tag{3}$$

where y_n is the observation, $y_n[\boldsymbol{\vartheta}_0]$ the underlying sine-wave (1) described by the true parameter vector $\boldsymbol{\vartheta}_0$. The process e_n describes the modeling error, noise, etc, and is assumed to be zero-mean white Gaussian with variance σ^2.

By employing a vector notation where $\boldsymbol{\vartheta}^T = (\theta^T, \omega)$, with

$$\theta = \begin{bmatrix} A & B & C \end{bmatrix}^T \tag{4}$$

and

$$\mathbf{D}(\omega) = \begin{bmatrix} \cos\omega t_1 & \sin\omega t_1 & 1 \\ \vdots & \vdots & \vdots \\ \cos\omega t_N & \sin\omega t_N & 1 \end{bmatrix} \tag{5}$$

the sum-squared-error (2) can be written as

$$V(\omega, \theta) = \frac{1}{N}(\mathbf{y} - \mathbf{D}(\omega)\,\theta)^T(\mathbf{y} - \mathbf{D}(\omega)\,\theta). \tag{6}$$

When the frequency ω is known, (6) is minimized in least-square sense by solving the set of linear equations $\mathbf{D}(\omega)\theta = \mathbf{y}$, giving the solution

$$\hat{\theta} = (\mathbf{D}(\omega)^T\mathbf{D}(\omega))^{-1}\mathbf{D}(\omega)^T\mathbf{y}. \tag{7}$$

When the frequency is unknown, the criterion (6) can be concentrated with respect to θ by plugging in the least-squares solution (7) into (6). Thus,

$$V(\omega) = \frac{1}{N}\left(\mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{\Pi}(\omega)\mathbf{y}\right) \tag{8}$$

where

$$\mathbf{\Pi}(\omega) = \mathbf{D}(\omega)(\mathbf{D}(\omega)^T\,\mathbf{D}(\omega))^{-1}\mathbf{D}(\omega)^T. \tag{9}$$

It is straightforward to show that ω can be found by a one-dimensional search for the maximum of [9]

$$g(\omega) = \mathbf{y}^T\,\mathbf{D}(\omega)\,(\mathbf{D}(\omega)^T\,\mathbf{D}(\omega))^{-1}\mathbf{D}(\omega)^T\,\mathbf{y} \tag{10}$$

The dependency of (10) on ω is non-trivial. Although, efficient algorithms exists for this class of non-linear least squares problems.

Once (10) has been maximized and the corresponding argument (say $\hat{\omega}$) has been determined, the unknowns in θ are obtained by a least-squares fit (7).

III. Cramér-Rao bound

A lower bound on the accuracy (covariance) of any unbiased estimator is given by the CRB. The covariance of any unbiased estimator of the parameters is bounded by the CRB, that is

$$\mathrm{Cov}(\hat{\boldsymbol{\vartheta}}) \geq \mathrm{CRB}(\boldsymbol{\vartheta}) \tag{11}$$

where $\geq$ is to be interpreted as the difference $\mathrm{Cov}(\hat{\boldsymbol{\vartheta}}) - \mathrm{CRB}(\boldsymbol{\vartheta})$ is positive semidefinite. In the full paper, we derive the asymptotic ($N >> 1$) CRB for the estimated parameters. This will be performed for both the three- and four-parameter model, respectively. These results are an extension (by adding the constant C in the signal model) of the well known results of [1]. By a re-parameterization of the sine-wave model, i.e

$$y_n[\alpha, \phi, C, \omega] = \alpha \sin(\omega t_n + \phi) + C \tag{12}$$

we will also show that the CRB on ϕ is lower for a three-parameter model (known frequency) than for the four-parameter model. The parameter accuracy effects the sine-wave fitting performance. Further discussions on the topic is made in Section V, as well as in the full paper.

IV. The parsimony principle

Consider the criterion (2) for the signal (3). Then

$$\mathrm{E}[V(\boldsymbol{\vartheta}_0)] = \frac{1}{N}\sum_{n=1}^{N} \mathrm{E}[(y_n - y_n[\boldsymbol{\vartheta}_0])^2] = \sigma^2 \tag{13}$$

where $\mathrm{E}[\cdot]$ denotes statistical expectation. Thus, when the estimate (say, $\hat{\boldsymbol{\vartheta}}$) is $\hat{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta}_0$, the residual is white noise and has minimum variance. In a practical scenario $\boldsymbol{\vartheta}_0$ is replaced by an estimate $\hat{\boldsymbol{\vartheta}}$. By the parsimony principle it is possible to show that the residual mean value evaluated with the estimated parameter $\hat{\boldsymbol{\vartheta}}$ is given by [10]

$$\mathrm{E}[V(\hat{\boldsymbol{\vartheta}})] \quad \simeq \quad \sigma^2\left(1 + \frac{p}{N}\right) \tag{14}$$

where $p = \dim[\boldsymbol{\vartheta}]$. The result (14) holds in a large sample scenario ($N >> 1$). A more thorough derivation will be given in the full paper.

V. Three vs. Four parameter model

In Section III we outlined that the CRB is lower for the parameters in a three-parameter model than for the ones in a four-parameter model. Further, from eq (14) it is clear that the three-parameter fit results (in mean) in a smaller sum-squared-error (2) than the four-parameter model, i.e a better fit of the sine-wave is obtained. However, a three-parameter method requires that the frequency ω is known. In ADC measurements a sine-wave generator is used to excite the device under study. In this case the frequency may not be completely known, i.e it may slightly differ from the preset value. However, if the frequency error is small we may assume it to be known in favor of using the four-parameter model. In the full paper, it is shown how the mean value of the residual (14) for a three-parameter model depends on the frequency error. From this study we can answer the question whether to use a three-parameter fit instead of a four-parameter fit, when fitting a sine-wave to a given set of data. That is, an answer to the question how large deviation from the nominal frequency is allowed, before the four-parameter method outperforms the tree-parameter method.

VI. Simulations

The mean value of the residual using both the three- and four-parameter model is evaluated for some data sets. In the three-parameter case, the frequency used in the sine-wave fit is varied around the nominal frequency. This to illustrate the behavior of the residual value when the frequency knowledge is not perfect, as well as to show the agreement between the theoretical results and simulations.

VII. Summary of Contributions

The contributions in this paper are:

- Derivation of the asymptotic Cramér-Rao bound on the variance of the parameters of the three- and four-parameter sine-wave models.
- Analysis of the statistics of the sine-wave fit criterion, resulting in an analytic expression for the residual in the three- and four-parameter cases.
- The answer to the question: *"When to use a three-parameter fit in favor of a four parameter fit?"*

References

[1] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, 1974.

[2] P. Stoica, "List of references on spectral line analysis," *Signal Processing*, no. 31, pp. 329–340, 1993.

[3] *IEEE Standard for Digitizing Waveform recorders*. IEEE Std. 1057-1994.

[4] *IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters*. IEEE Std. 1241.

[5] J. Márkus and I. Kollár, "Standard framework for IEEE-STD-1241 in Matlab," in *IEEE Instrumentation and Measurement Technology Conference*, (Budapest, Hungary), pp. 1847–52, IMTC/2001, May 21-23 2001.

[6] J. Blair, "Sine-fitting software for IEEE standards 1057 and 1241," in *the 16th IEEE Instr. and Meas. Technology Conference*, vol. 3, (Venice, Italy), pp. 1504–06, IMTC/99, May 1999.

[7] P. Arpaia, A. C. Serra, P. Daponte, and C. L. Monteiro, "A critical note to IEEE 1057-94 standard on hysteretic ADC dynamic testing," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, pp. 941–948, August 2001.

[8] T. Bilau, T. Megyeri, J. M. A.S̃árhegyi, and I.Kollár, "Four parameter fitting of sine wave testing results: iteration and convergence," in *4-th International Conference on Advanced A/D and D/A Conversion Techniques and their Applications & 7-th European Workshop on ADC Modelling and Testing*, (Prague, Czech Republic), pp. 185–189, June 26-28 2002.

[9] P. Händel, "Properties of the IEEE-STD-1057 four parameter sine wave fit algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, pp. 1189–1193, Dec. 2000.

[10] T. Söderström and P. Stoica, *System Identification*. Prentice Hall International, 1989.