IEEE standard 1057, Cramér-Rao bound and the parsimony principle

Tomas Andersson and Peter Händel
KTH Signals, Sensors and Systems, Royal Institute of Technology
SE–100 44 Stockholm, Sweden, Fax: +46 8 790 7260
E-mail: tomas.andersson@s3.kth.se and peter.handel@s3.kth.se

Abstract— The objective of this paper is to present the statistical properties of the IEEE-STD-1057/IEEE-STD-1241 sine-wave fitting algorithm. The proper Cramér-Rao bound is derived for both the three-parameter and four-parameter fitting algorithms. Further, we investigate the sine-wave fitting criterion, and derive its statistical properties on both the three- and four-parameter fitting cases. In a setup where the frequency is *partially* known, an analytical expression of the mean of the residual is derived. From the analysis, a simple formula is obtained indicating when to use one over the other. Finally some numerical evaluations confirming our analytical findings are presented.

Index Terms— sine-wave fitting, ADC testing, parsimony principle.

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] and an investigation of their performance has been the main topic in [5], [6]. In this paper tone frequency estimation is studied in general, and the performance of the standardized three- and four-parameter algorithms in particular.

Assume that the data record contains the sequence of measurement samples $\mathbf{x} = [x_1 \dots x_N]^T$ taken at time instants $\{t_1, \dots, t_N\}$. It is further assumed that data can be modeled by

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

where A, B, C and ω are (known or unknown) constants. Throughout this paper the *generic parameter vector* $\boldsymbol{\vartheta}$ represents either the set (A, B, C) or (A, B, C, ω) depending on if the frequency ω is known, or not. The sine-wave fit problem is solved by minimizing the sum-squared-error

$$V(\boldsymbol{\vartheta}) = \frac{1}{N}\sum_{n=1}^{N}(x_n - s_n(\boldsymbol{\vartheta}))^2 \tag{2}$$

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

$$x_n = s_n(\boldsymbol{\vartheta}) + w_n. \tag{3}$$

The process w_n describes the modeling error, noise, etc, and is assumed to be zero-mean white Gaussian with variance σ^2. Consider the parameter vector $\boldsymbol{\vartheta} = \boldsymbol{\theta}$, where

$$\boldsymbol{\theta} = [\theta^T \ \omega]^T, \quad \text{and} \quad \theta = \begin{bmatrix} A & B & C \end{bmatrix}^T. \tag{4}$$

Let

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

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

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

The algorithms in [3], [4] find the estimates by minimizing (6). When the frequency ω is known and $\mathbf{D}(\omega)$ has full column rank, (6) is minimized in least-squares sense by

$$\hat{\theta} = \left[\mathbf{D}(\omega)^T\mathbf{D}(\omega)\right]^{-1}\mathbf{D}(\omega)^T\mathbf{x}. \tag{7}$$

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,

$$V(\omega) = \frac{1}{N}\left[\mathbf{x}^T\mathbf{x} - g(\omega)\right] \tag{8}$$

where

$$g(\omega) = \mathbf{x}^T\,\mathbf{D}(\omega)\ \left[\mathbf{D}(\omega)^T\,\mathbf{D}(\omega)\right]^{-1}\mathbf{D}(\omega)^T\,\mathbf{x}. \tag{9}$$

The unknown frequency ω can be found by a one-dimensional search for the maximum of (9) [7]. The dependency of (9) on ω is non-trivial. Although, efficient algorithms exist for this class of non-linear least-squares problems. Once (9) 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), replacing ω in (7) with $\hat{\omega}$.

II. Cramér-Rao Bound

A lower bound on the accuracy (covariance) of any unbiased estimator is given by the CRB, that is $\mathrm{Cov}(\hat{\boldsymbol{\vartheta}}) \geq \mathrm{CRB}(\boldsymbol{\vartheta})$ where $\geq$ is to be interpreted as that the difference $\mathrm{Cov}(\hat{\boldsymbol{\vartheta}}) - \mathrm{CRB}(\boldsymbol{\vartheta})$ is positive semidefinite. The CRB is given by the inverse of the Fisher information matrix $\mathbf{J}(\boldsymbol{\vartheta})$, that is

$$\mathrm{CRB}(\boldsymbol{\vartheta}) = \mathbf{J}(\boldsymbol{\vartheta})^{-1} \tag{10}$$

The Fisher information matrix is given by [7]

$$\mathbf{J}(\boldsymbol{\vartheta}) = \mathrm{E}\left\{\left(\frac{\partial \ln p(\mathbf{x};\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right)\left(\frac{\partial \ln p(\mathbf{x};\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right)^T\right\} \quad (11)$$

where $p(\mathbf{x};\boldsymbol{\vartheta})$ denotes the probability density function (pdf), and where the derivative is evaluated at the true parameters. The pdf for $\mathbf{x}$ is given by

$$p(\mathbf{x};\boldsymbol{\vartheta}) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-(\mathbf{x}-\mathbf{s}(\boldsymbol{\vartheta}))^T(\mathbf{x}-\mathbf{s}(\boldsymbol{\vartheta}))/(2\sigma^2)} \quad (12)$$

where $\mathbf{s}(\vartheta) = [s_1(\vartheta)\ldots s_N(\vartheta)]$. The derivative of a function $y(\boldsymbol{\vartheta})$ with respect to $\boldsymbol{\vartheta}$ is defined as the vector

$$\frac{\partial y(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \begin{bmatrix} \frac{\partial y(\boldsymbol{\vartheta})}{\partial \vartheta_1} & \ldots & \frac{\partial y(\boldsymbol{\vartheta})}{\partial \vartheta_p}\end{bmatrix}^T \quad (13)$$

where $p = \dim(\boldsymbol{\vartheta})$ and $\{\vartheta_1 \ldots \vartheta_p\}$ are the elements in $\boldsymbol{\vartheta}$. Now, inserting (12)-(13) into (11) yields [7]

$$\mathbf{J}(\boldsymbol{\vartheta}) = \frac{1}{\sigma^2}\sum_{n=1}^{N} \psi_n \psi_n^T \quad (14)$$

where,

$$\psi_n = \frac{\partial s_n(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \begin{bmatrix} \frac{\partial s_n(\boldsymbol{\vartheta})}{\partial \vartheta_1} & \ldots & \frac{\partial s_n(\boldsymbol{\vartheta})}{\partial \vartheta_p}\end{bmatrix}^T . \quad (15)$$

In the forthcoming discussion the formula describing the elements of the Fisher information matrix $J(\boldsymbol{\vartheta})$ is useful.

$$[\mathbf{J}(\boldsymbol{\vartheta})]_{\ell,r} = \frac{1}{\sigma^2}\sum_{n=1}^{N} [\psi_n]_\ell [\psi_n]_r \quad (16)$$

In (16), $[\cdot]_{\ell,r}$ denotes the ℓ, r-th element of the matrix within the parentheses and $[\cdot]_\ell$ denotes the ℓ-th element in the vector. With $\boldsymbol{\vartheta} = \boldsymbol{\theta}$ according to (4), one has to calculate

$$\mathbf{J}(\boldsymbol{\theta}) = \begin{bmatrix} J_{AA} & J_{AB} & J_{AC} & J_{A\omega} \\ J_{AB} & J_{BB} & J_{BC} & J_{B\omega} \\ J_{AC} & J_{BC} & J_{CC} & J_{C\omega} \\ J_{A\omega} & J_{B\omega} & J_{C\omega} & J_{\omega\omega} \end{bmatrix} \quad (17)$$

where $J_{AA} = [\mathbf{J}(\boldsymbol{\theta})]_{1,1}$ according to (16), etc. The derivatives in (16) are straightforward to calculate [6]. The resulting expression is suitable for numerical evaluation of the information matrix, and the CRB follows by numerical inversion of the result; See [6] for a comparision on the performance of the frequency estimate using IEEE standard 1057 [3] and the exact CRB.

In order to get an analytical insight, an asymptotic expression of $\mathbf{J}(\boldsymbol{\theta})$ (valid for large N, that is $N \gg 1/\omega$), is given below. The details are presented in [8]. The result reads

$$\mathrm{CRB}(\boldsymbol{\theta}) \simeq 2\sigma^2 \begin{bmatrix} N & 0 & 0 & -\frac{BN^2}{2} \\ 0 & N & 0 & \frac{AN^2}{2} \\ 0 & 0 & 2N & 0 \\ -\frac{BN^2}{2} & \frac{AN^2}{2} & 0 & \frac{(A^2+B^2)N^3}{3} \end{bmatrix}^{-1} \quad (18)$$

where $\simeq$ denotes an approximative equality in which only the asymptotically (for $N \gg 1$) relevant terms have been retained.

A. Three Parameter Model

For a known frequency, the corresponding information matrix is given by the upper left 3×3 sub-matrix of (17). For large N, the sub-matrix of (18) is diagonal, and its inverse determines the lower bound on the variance of the estimates as

$$\mathrm{Var}(\hat{A}) = \mathrm{Var}(\hat{B}) \geq \mathrm{CRB}(A) \simeq \frac{2\sigma^2}{N} \quad (19)$$

$$\mathrm{Var}(\hat{C}) \geq \mathrm{CRB}(C) \simeq \frac{\sigma^2}{N} \quad (20)$$

Since the upper left 3×3 sub-matrix in (18) is diagonal, the lower bound on estimation accuracy of A and B does not depend if the offset is assumed known or is estimated.

B. Four Parameter Models

The four parameter model corresponds to $\boldsymbol{\theta} = \begin{bmatrix}\theta^T & \omega\end{bmatrix}^T$ as in (4). Calculations of the CRB yields [8]

$$\mathrm{Var}(\hat{A}) \geq \mathrm{CRB}(A) \simeq \frac{2\sigma^2}{N}\left(1 + \frac{3B^2}{\alpha^2}\right) \quad (21)$$

$$\mathrm{Var}(\hat{B}) \geq \mathrm{CRB}(B) \simeq \frac{2\sigma^2}{N}\left(1 + \frac{3A^2}{\alpha^2}\right) \quad (22)$$

$$\mathrm{Var}(\hat{C}) \geq \mathrm{CRB}(C) \simeq \frac{\sigma^2}{N} \quad (23)$$

$$\mathrm{Var}(\hat{\omega}) \geq \mathrm{CRB}(\omega) \simeq \frac{24\sigma^2}{\alpha^2 N^3} \quad (24)$$

where $\alpha^2 = A^2 + B^2$, that is the signal amplitude squared. As shown in (19)-(20) above, the result of A and B is independent of C. Also, $\mathrm{CRB}(\omega)$ is asymptotically independent of the fact if C is estimated or assumed known. One may note from (21)-(22) that the CRBs for the A and B parameters are 1 to 4 times the corresponding CRB in (19). In order to further analyze the above CRBs, a re-parameterized model is considered. Consider

$$s_n(\boldsymbol{\Psi}) = \alpha \sin(\omega t_n + \phi) + C \quad (25)$$

where $A = \alpha \sin \phi$ and $B = \alpha \cos \phi$. Further, let $\boldsymbol{\Psi} = \begin{bmatrix}\alpha & \phi & C & \omega\end{bmatrix}^T$. Then the $\mathrm{CRB}(\boldsymbol{\Psi})$ follows from the general relationship [7]

$$\mathrm{CRB}(\boldsymbol{\Psi}) = \left[\frac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\theta}}\right] \mathrm{CRB}(\boldsymbol{\theta}) \left[\frac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\theta}}\right]^T \quad (26)$$

where the k, r-th element of $[\partial\boldsymbol{\Psi}/\partial\boldsymbol{\theta}]$ is $\partial\boldsymbol{\Psi}_k/\partial\boldsymbol{\theta}_r$. The CRB for the re-parameterized model is given by

$$\mathrm{Var}(\hat{\alpha}) \geq \mathrm{CRB}(\alpha) \simeq \frac{2\sigma^2}{N} \quad (27)$$

$$\mathrm{Var}(\hat{\phi}) \geq \mathrm{CRB}(\phi) \simeq \begin{cases} \frac{2\sigma^2}{N\alpha^2} & \text{three parameter model} \\ \frac{8\sigma^2}{N\alpha^2} & \text{four parameter model} \end{cases} \quad (28)$$

$$\mathrm{Var}(\hat{C}) \geq \mathrm{CRB}(C) \simeq \frac{\sigma^2}{N} \tag{29}$$

$$\mathrm{Var}(\hat{\omega}) \geq \mathrm{CRB}(\omega) \simeq \frac{24\sigma^2}{\alpha^2 N^3} \quad \text{four parameter model} \tag{30}$$

One may conclude from (27)-(30), that the uncertainty in the phase ϕ is dependent on whether ω is known or not, while the uncertainty in the amplitude α and DC-level C is independent of this fact.

In wave-form fitting, the quality of the parameter estimates is not as important as the quality of the wave-form fit, that is the value of (2) for the given estimate. Clearly, with a number of parameters p that equals the number of data sampels N, the residual (2) can be forced to zero. On the other hand, increasing the number of parameters implies that the estimator fits the parameters to the noise, and not to the signal. Thus, one should strive to use as few parameters as possible, but still have a model flexible enough to describe the behavior of the signal.

In the forthcoming section, the problem when to use the three- and four-parameter models is studied by aid of the parsimony principle.

III. The Parsimony Principle

Consider the criterion (2) for the signal (3) described by a vector $\boldsymbol{\vartheta}_0$, that is, the set of true but unknown parameters. Then the expected value of (2) is given by

$$\mathrm{E}[V(\boldsymbol{\vartheta}_0)] = \frac{1}{N}\sum_{n=1}^{N} \mathrm{E}[(x_n - s_n(\boldsymbol{\vartheta}_0))^2] = \sigma^2 \tag{31}$$

Here the expectation is with respect to the measurement noise. Thus, when the estimate (say, $\hat{\boldsymbol{\vartheta}}$) is exact $\hat{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta}_0$, the residual is white noise and has minimum variance. Now, assuming an estimate $\hat{\boldsymbol{\vartheta}}$ based on some *past* data, then a Taylor series expansion of $\varepsilon_n(\boldsymbol{\vartheta}) = x_n - s_n(\boldsymbol{\vartheta})$ around $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$ gives

$$\mathrm{E}[V(\hat{\boldsymbol{\vartheta}})] = \frac{1}{N}\sum_{n=1}^{N} \mathrm{E}\left[\left(\varepsilon_n(\boldsymbol{\vartheta}_0) + \left.\frac{\partial \varepsilon_n(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right|^{T}_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)\right)^2\right] \tag{32}$$

Then from (15) with $\psi_n(\boldsymbol{\vartheta}_0)$ equal to ψ_n evaluated at the point $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$, it follows that

$$\begin{aligned}
\mathrm{E}[V(\hat{\boldsymbol{\vartheta}})] &= \frac{1}{N}\sum_{n=1}^{N} \mathrm{E}\left[\left(\varepsilon_n(\boldsymbol{\vartheta}_0) - \psi_n^T(\boldsymbol{\vartheta}_0)(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)\right)^2\right] \\
&= \sigma^2 + \mathrm{E}[\psi_n^T(\boldsymbol{\vartheta}_0)(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^T \psi_n(\boldsymbol{\vartheta}_0)] \\
&= \sigma^2 + \mathrm{Tr}\left\{\mathrm{E}_{\hat{\boldsymbol{\vartheta}}}[(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^T] \right. \\
&\qquad \left. \mathrm{E}_{\mathbf{w}}[\psi_n(\boldsymbol{\vartheta}_0)\psi_n^T(\boldsymbol{\vartheta}_0)]\right\}
\end{aligned} \tag{33}$$

where in the second equality it is assumed that the estimate is unbiased. In the third equality the expectations can be separated due to the fact that $\hat{\boldsymbol{\vartheta}}$ depends on past data and $\psi_n(\boldsymbol{\vartheta}_0)$ on future data. Also in the third equality the trace (Tr) operator and the expectation have changed order. Further, if an unbiased statistically efficient estimator is employed, the mean-square error will be equal to the CRB [9]. That is,

$$\mathrm{E}_{\hat{\boldsymbol{\vartheta}}}\Big[(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0)^T\Big] \simeq \frac{\sigma^2}{N}\mathrm{E}_{\mathbf{w}}\left[\psi_n(\boldsymbol{\vartheta}_0)\psi_n^T(\boldsymbol{\vartheta}_0)\right]^{-1} \tag{34}$$

In (34), the left hand side is the mean-square error and the right hand side is the CRB($\boldsymbol{\vartheta}$) (14). Inserting (34) into (33) yields an identity matrix of dimension p within the Tr operator. Thus,

$$\mathrm{E}[V(\hat{\boldsymbol{\vartheta}})] \simeq \sigma^2\left(1 + \frac{p}{N}\right). \tag{35}$$

where $p = \dim(\boldsymbol{\vartheta})$. According to (35), the residual increases with an increasing model order. This is a known result and a general derivation can be found in [9]. From (35) it is clear that the four-parameter fit results in a larger sum-squared-error (2) than the three-parameter method, which however requires the frequency to be known.

In most practical cases the frequency is known up to some uncertainty quantity. Here such a quantity is denoted by ω_δ, meaning that the true frequency ω_0 is within the range of the *assumed* frequency $\tilde{\omega}$, as $|\tilde{\omega} - \omega_0| \leq \omega_\delta$. The four-parameter method results in a constant mean residual independent of ω_δ, as seen in (35), whereas in the three-parameter case the expectation of the residual depends on ω. Let $\mathcal{V}(\omega)$ denote the expectation (w.r.t the noise w_n) of (8) as $\mathcal{V}(\omega) = \mathrm{E}\left[V(\omega)\right]$. A Taylor series expansion of $\mathcal{V}(\omega)$ around the frequency $\omega = \omega_0$ gives

$$\begin{aligned}
\mathcal{V}(\omega) \simeq \mathcal{V}(\omega_0) + \left.\frac{\partial \mathcal{V}(\omega)}{\partial \omega}\right|_{\omega=\omega_0} (\omega - \omega_0) + \\
\frac{1}{2}\left.\frac{\partial^2 \mathcal{V}(\omega)}{\partial \omega^2}\right|_{\omega=\omega_0} (\omega - \omega_0)^2
\end{aligned} \tag{36}$$

The first two derivatives of $\mathcal{V}(\omega)$ with respect to ω are given below, that is

$$\frac{\partial \mathcal{V}(\omega)}{\partial \omega} = \mathrm{E}\left\{-\frac{2}{N}[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta})]^T \frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \omega}\right\} \tag{37}$$

$$\begin{aligned}
\frac{\partial^2 \mathcal{V}(\omega)}{\partial \omega^2} = \mathrm{E}\left\{\frac{2}{N}\frac{\partial \mathbf{s}^T(\boldsymbol{\theta})}{\partial \omega}\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \omega}\right\} - \\
\mathrm{E}\left\{\frac{2}{N}\left[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta})\right]^T \frac{\partial^2 \mathbf{s}(\boldsymbol{\theta})}{\partial \omega^2}\right\}.
\end{aligned} \tag{38}$$

The used estimator is unbiased which implies that the expectation of the difference in (37) evaluated for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ equals the zero vector. Therefore, the first derivative of $\mathcal{V}(\omega)$ with respect to ω in (36) equals zero. The same argument holds in (38), that is putting the second term to zero as well. In the reconstruction of $s_n(\boldsymbol{\theta})$, the estimate of the three parameters in θ depends on the assumed

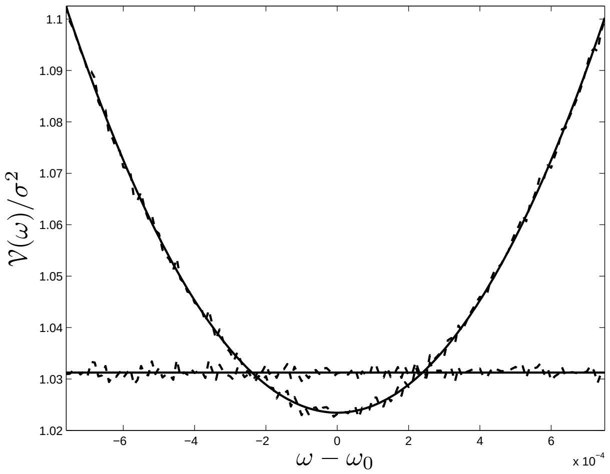

Fig. 1. The expected value of the residual $\mathcal{V}(\omega)$ for a three-(parabola) and four-parameter(straight line) model. The dashed lines are numerical evaluations of the residual using the different models. The curves have been obtained for $N = 128$ and SNR $= 20$dB.

frequency $\tilde{\omega}$, see (7). Replacing θ with its least-square solution (7) and taking the derivative with respect to ω using the chain rule yields

$$\begin{aligned}\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \omega} =& \mathbf{D}_\omega[\mathbf{D}^T\mathbf{D}]^{-1}\mathbf{D}^T\mathbf{x} + \mathbf{D}[\mathbf{D}^T\mathbf{D}]^{-1}\mathbf{D}_\omega^T\mathbf{x} \\ &- 2\mathbf{D}\mathbf{D}^T\mathbf{D}\left[\mathbf{D}_\omega^T\mathbf{D}\right]^{-1}\mathbf{D}^T\mathbf{D}\mathbf{D}^T\mathbf{x} \qquad (39)\end{aligned}$$

where the subscript $[\cdot]_\omega$ denotes an element-wise derivative with respect to ω. The derivation of (39) is straightforward, but tedious. A thorough derivation can be found in [8]. Evaluated at $\omega = \omega_0$, the second derivative (38) can be expressed as

$$\left.\frac{\partial^2 \mathcal{V}(\omega)}{\partial \omega^2}\right|_{\omega=\omega_0} = \frac{(A^2+B^2)N^2}{12} \qquad (40)$$

Inserting (40) into the Taylor series expansion (36), the expected value of the residual using a three-parameter model is described by a second order parabola

$$\mathcal{V}(\omega) = \sigma^2(1+\frac{3}{N}) + \frac{(A^2+B^2)N^2}{12}(\omega-\omega_0)^2. \quad (41)$$

In Figure 1, the expected value of the residual (2) is shown, both the theoretical results (35) and (41), respectively; and experimental evaluations thereof. In the evaluation the true frequency ω_0 has been randomly drawn in the region $[(2\pi)/N, \pi-(2\pi)/N]$. The residual have then been evaluated using a frequency with a variable displacement from the true frequency for the three-parameter method, and an ML estimate of the true frequency for the four-parameter fit, respectively [7]. The expectation has been approximated with the sample mean, using $2.5\cdot 10^5$ Monte-Carlo simulations at each frequency point. In Figure 1, it is seen that for small deviations of ω from ω_0 a sine-wave fit using a three-parameter model performs better than a four-parameter model in terms of a smaller residual. The *break even* frequency displacement from the true frequency ω_0 is obtained by setting (35) equal to (41) for $p = 4$ and solve for $|\omega-\omega_0|$. The result yields

$$|\omega-\omega_0| = \sqrt{\frac{24\sigma^2}{\alpha^2 N^3}} = \sqrt{\mathrm{CRB}(\omega)}, \qquad (42)$$

where the $\mathrm{CRB}(\omega)$ is given by (24). Hence, if the uncertainty of the true frequency ω_0 is less than (42) a three-parameter model should be used when fitting a sine-wave to a set of measurement data.

IV. Conclusions

In this paper, the IEEE standard 1057 for tone frequency estimation and signal recovery has been considered. A performance analysis employing the Cramér-Rao bound and the parsimony principle has been performed. Further, the criterion function used in the sine-wave fitting part of IEEE standard 1057 has been studied in some detail. The difference in performance between the three- and four-parameter model has also been studied. In particular, an expression (42) is derived that provides a *rule-of-thumb* for the selection of the proper number of free parameters in the sine-wave model.

References

[1] D. Rife and R. Boorstyn, “Single tone parameter estimation from discrete-time observations,” *IEEE Transaction on Information Theory*, vol. 20, 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*, 2000.

[5] 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 Transaction on Instrumentation and Measurement.*, vol. 50, pp. 941–948, August 2001.

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

[7] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol 1. Prentice Hall, 1993.

[8] T. Andersson, “Selected topics in frequency estimation,” Licentiate Thesis TRITA–S3–SB–0323, Royal Institute of Technology, May 2003. www.s3.kth.se

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