Digital Object Identifier of the paper: 10.1109/TIM.2016.2600998

Efficient Implementation of Least Squares Sine Fitting Algorithms

B. Renczes, I. Kollár, T. Dabóczi
Budapest University of Technology and Economics,
Department of Measurement and Information Systems, Budapest, Hungary

***Abstract*—In this paper, three- and four-parameter least squares (LS) sine fitting algorithms are investigated. It is pointed out that the three-parameter fitting is well conditioned in its standard form, both for short and long records. Then, the conditioning of the four-parameter fitting (4PF) is investigated. A scaling factor is derived in order to ensure good conditioning of the equations. A Monte Carlo analysis is carried out to demonstrate that in practical cases, using this scaling factor ensures good conditioning for the four-parameter system. It is also shown that parameters can be determined precisely using direct pseudoinverse calculation for both methods. Hence, in this case, there is no need to use the computationally more demanding decomposition methods, although these are generally recommended for the solution of LS problems. In addition, data centering for time instants is introduced in order to further improve the numerical properties of the 4PF. It is shown that with this method, the fourparameter problem can be approximated with a diagonal matrix. Finally, an evaluation method is presented to significantly decrease roundoff errors of the widely used LS methods.**

Keywords—Least squares methods, sine fitting, four-parameter fitting, numerical stability, condition number, analog-digital conversion.

I. INTRODUCTION

Digital computers offer a fast and efficient way to process analog signals. However, they cannot deal with the analog input directly. Hence, an analog-to-digital conversion is needed prior to running digital signal processing algorithms. For the characterization of analog-to-digital converters (ADCs) it is crucial that we can determine the quality of the conversion in a standardized way. In IEEE Standard 1241-2010 methods are prescribed to test ADCs [1]. Due to the fact that a sinusoidal waveform can be generated with high purity, ADCs are mostly tested with sine wave excitation. Such excitation is also needed to evaluate the histogram test of ADCs [2].

Since the parameters of the generated analog input signal are not known precisely, a sine wave has to be fitted to the measured noisy data. The difference between the measured and the fitted sine wave characterizes the quality of the conversion. The most widely used method is the Least Squares (LS) sine wave fitting algorithm [1]. It attempts to fit a sine to the measured data, minimizing the Mean Square Error (MSE)

$$y_k = A \cdot \cos(2\pi f t_k) + B \cdot \sin(2\pi f t_k) + \mathrm{C}, \tag{1}$$

$$MSE = \sum_{k=0}^{N-1} (x_k - y_k)^2, \tag{2}$$

where y_k and x_k denote the k^{th} sample in the fitted sine wave and the measured data set, respectively, N is the record length, A, B and C are the amplitudes of the cosinusoidal and sinusoidal components and the offset, respectively. The signal frequency is denoted by f. Furthermore, t_k is the time instant, at which x_k was sampled [1]. For regular sampling time instants are given by

$$t_k = k/f_s\,, \qquad k = 0, \dots, N-1, \tag{3}$$

where f_s denotes the sampling frequency. If the sampling is not equidistant, k can assume non-integer values. In this paper equidistant sampling is assumed. Since the signals are processed digitally, the sampling frequency is not necessarily known. The parameter that is needed to describe the sine wave is the ratio of the signal frequency to the sampling frequency f/f_s.

In case of the three-parameter fitting (3PF) f/f_s is assumed to be known. This problem is linear in the parameters. Thus, parameters can be calculated in one step, without iteration [1]. However, if the frequency ratio is unknown, the problem becomes non-linear. It is to be mentioned here that the unknown frequency is still assumed to be constant, that is, the frequency drift of the generator, [3], is not considered. There are two different approaches for the solution. In the first one, the frequency ratio is estimated in advance. Then, a 3PF is executed. The frequency ratio can be estimated, for instance, by interpolated FFT (see e.g. [4] or an overview in [5]). It was pointed out in [6] that if the frequency is known within an interval, the 3PF outperforms the four-parameter fitting (4PF), which is the other approach. In the 4PF, all the four parameters are estimated at the same time.

The properties of the LS fitting methods have been widely investigated. It was shown in [7] that harmonic distortion and noise affect the result of the 4PF. In particular, these distortions result in a biased frequency estimate.

In addition, the 3PF was also investigated in point of bias. [8] and [9] pointed out that although the amplitude estimate of this method is also biased, it is asymptotically unbiased. A more rigorous analysis has been executed, taking also quantization into consideration. It was shown that the amplitude and offset estimates are biased if the quantization cannot be modelled as additive, uniformly distributed noise [10], [11]. Besides, convergence of the 4PF has been analyzed, [12].

Certainly, there are other possible algorithms to fit a sine wave to noisy measured data, such as the Maximum Likelihood (ML) estimator [13] or the Quantile Based Estimator [10]. However, ML is computationally more demanding compared to the LS [14], while the use of QBE is limited to coherent sampling. In fact, if error sequence $x_k - y_k$ is random, white and of Gaussian distribution, the LS estimate coincides with the ML estimate [15]. Furthermore, in practice the LS method yields good results.

This paper focuses on the 3PF and 4PF algorithms, aiming to improve numerical stability and computational time of these methods.

In Section II the conditioning of the 3PF and the 4PF is investigated. It is proved that the 3PF is well-conditioned, provided that at least 4 periods are sampled. Contrarily, the 4PF is shown to become ill-conditioned for high amplitude values and long records. In Section II-C, a scaling factor is suggested for the 4PF in order to ensure well-conditioning. This scaling factor is shown to enhance conditioning significantly. In Section III data centering for time instants is introduced. This technique is shown to improve numerical properties of both the 3PF and the 4PF. Finally, in Section IV proposals are given in order to evaluate Least Squares algorithms in a numerically efficient way.

II. INVESTIGATION OF THE CONDITIONING OF THE STANDARD METHODS

In this section, the conditioning of the 3PF and the 4PF is investigated, assuming they are evaluated as it is prescribed in [1].

A. The three-parameter case

In case of the three-parameter fitting (3PF) f/f_s is assumed to be known. For this problem the following equation system has to be solved in LS sense

$$\mathbf{x} = \mathbf{D}_0\mathbf{s}_0 + \mathbf{e}. \tag{4}$$

where $\mathbf{D}_0$ is the system matrix of the 3PF

$$\mathbf{D}_0 = \begin{pmatrix} \cos\varphi_0 & \sin\varphi_0 & 1 \\ \cos\varphi_1 & \sin\varphi_1 & 1 \\ \vdots & \vdots & \vdots \\ \cos\varphi_{N-1} & \sin\varphi_{N-1} & 1 \end{pmatrix}, \qquad \varphi_k = \frac{2\pi f_0}{f_s}k, \tag{5}$$

$\mathbf{s}_0$ contains the amplitudes of the cosine and the sine waves, and the unit offset and $\mathbf{e}$ is the error sequence to be minimized in LS sense. Finally, f_0 is the frequency estimate. The problem is linear in the parameters. Thus, parameters can be calculated in one step, without iteration. The solution of (4) is given using the pseudo-inverse of $\mathbf{D}_0$ [16]

$$\hat{\mathbf{s}} = \mathbf{D}_0^{+}\mathbf{x}. \tag{6}$$

where $\hat{\mathbf{s}}$ contains the parameter estimates and $\mathbf{D}_0^{+}$ is the pseudo-inverse of $\mathbf{D}_0$. There are several methods to calculate the pseudo-inverse in (6). It can be *directly evaluated* using formula

$$\mathbf{D}_0^{+} = (\mathbf{D}_0^{\mathbf{T}}\mathbf{D}_0)^{-1}\mathbf{D}_0^{\mathbf{T}}. \tag{7}$$

However, in practical cases this formula is rarely used due to numerical considerations [17]. If $\mathbf{D}_0$ is ill-conditioned (we will see in Section II-B that in standard 4PF this can happen), i.e., the ratio of its maximal singular value to its minimal one is high, then solving (6) is numerically instable. The ill-conditioned problem results in high sensitivity of the estimated parameters to small perturbations of $\mathbf{x}$. In fact, the condition number (CN) of the system matrix $\mathbf{D}_0$ provides an *upper bound* for the error of the solution. The smaller the CN is (this can be modified e.g., by applying a linear transformation to (4)), the smaller the upper bound of the error of the solution will be, in proportion to the CN.

Notice that ill-conditioning is independent of the measured signal. Therefore, the additive noise on $\mathbf{x}$, which is known to increase the variance of the estimate, does not influence the numerical accuracy of the solution. Thus, numerical accuracy and the effect of input noise can be investigated separately from each other. This paper investigates the former problem.

The pseudo-inverse calculation of ill-conditioned problems is also numerically instable if it is calculated by (7). In this case, the CN of $\mathbf{D}_0$ is squared at the calculation of $\mathbf{D}_0^{\mathbf{T}}\mathbf{D}_0$. To avoid the latter problem, numerically stable methods are used in practice, like singular value decomposition (SVD) or QR-decomposition [16]. However, decomposition methods can be computationally demanding especially for long observations. If $\mathbf{D}_0^{\mathbf{T}}\mathbf{D}_0$ is well-conditioned, the calculation method of (7) can be used without numerical issues.

In the following the conditioning of the 3PF, that is, the conditioning of $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ will be investigated. This matrix can be calculated by

$$\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0 = \mathbf{H} = \begin{pmatrix} \sum_{k=0}^{N-1} cos^2\varphi_k & \sum_{k=0}^{N-1} \cos\varphi_k \sin\varphi_k & \sum_{k=0}^{N-1} \cos\varphi_k \\ \sum_{k=0}^{N-1} \cos\varphi_k \sin\varphi_k & \sum_{k=0}^{N-1} sin^2\varphi_k & \sum_{k=0}^{N-1} \sin\varphi_k \\ \sum_{k=0}^{N-1} \cos\varphi_k & \sum_{k=0}^{N-1} \sin\varphi_k & \sum_{k=0}^{N-1} 1 \end{pmatrix}. \tag{8}$$

This matrix can be given as

$$\mathbf{H} = N\left(\widetilde{\mathbf{H}} + \mathbf{E}\right), \tag{9}$$

where

$$\widetilde{\mathbf{H}} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{10}$$

and $\mathbf{E}$ contains the error terms [26]. From an analytical point of view, $\mathbf{H}$ can be described by N times matrix $\widetilde{\mathbf{H}}$, and a perturbation matrix $\mathbf{E}$. Notice that the multiplication with N does not change the CN. Thus, in the following the CN of $\widetilde{\mathbf{H}} + \mathbf{E}$ will be investigated.

Matrix entries of $\mathbf{E}$ can be proved to be bounded with the following elements [26]

$$\mathbf{E}_b = \frac{1}{J}\begin{pmatrix} \pm\frac{1}{8} & \pm\frac{1}{8} & \pm\frac{1}{2\sqrt{2}} \\ \pm\frac{1}{8} & \pm\frac{1}{8} & \pm\frac{1}{2\sqrt{2}} \\ \pm\frac{1}{2\sqrt{2}} & \pm\frac{1}{2\sqrt{2}} & 0 \end{pmatrix}, \tag{11}$$

provided that

$$\frac{J}{N} = \frac{f_0}{f_s} \leq 4, \tag{12}$$

where $\mathbf{E}_b$ denotes the bound elements of perturbation matrix $\mathbf{E}$, and J denotes the number of sampled periods. The condition in (12) prescribes that at least 4 samples should be sampled from one period. It can be observed that the elements of $\mathbf{E}_b$ decrease with increasing J. Thus, $\mathbf{H}$ can arbitrarily approach a diagonal matrix. Furthermore, for coherent sampling $\mathbf{E}$ is a null-matrix and $\mathbf{H}$ is diagonal.

Knowing the bounds on the elements of $\mathbf{E}$, eigenvalues of $\widetilde{\mathbf{H}} + \mathbf{E}$ can be estimated. Let us notate the eigenvalues of $\widetilde{\mathbf{H}}$ with λ_i and the eigenvalues of $\widetilde{\mathbf{H}} + \mathbf{E}$ with λ_i'. From matrix perturbation theory, the following limits can be given

$$|\lambda_i - \lambda_i'| \leq \|\mathbf{E}_b\|_F \tag{13}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, [21]. However, as described in Section I, for the calculation of the CN the singular values are needed. Let us notice that $\mathbf{H} = \mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ is symmetric and positive semidefinite [16]. It follows that its eigenvalues equal to its singular values. Thus,

$$|\sigma_i - \sigma_i'| \leq \|\mathbf{E}_b\|_F = \frac{0.75}{J} \tag{14}$$

holds, where the singular values are denoted by σ_i and σ_i'. It is obvious from (10) that the largest σ_i' equals to 1 and the smallest σ_i' equals to 1/2. Thus, the following limitation can be given for conditioning of the 3PF

$$cond(\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0) = cond\left(\tilde{\mathbf{H}} + \mathbf{E}\right) = \frac{\max(\sigma_i)}{\min(\sigma_i)} \leq \frac{\max(\sigma_i') + \|\mathbf{E}_b\|_F}{\min(\sigma_i') - \|\mathbf{E}_b\|_F}, \tag{15}$$

that is

$$cond(\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0) \leq \frac{1 + \frac{0.75}{J}}{0.5 - \frac{0.75}{J}} \quad \text{if} \quad J > 1.5. \tag{16}$$

The condition on *J* is needed in order to ensure that the denominator is greater than 0. It follows that the CN is lower than 11, provided that the number of sampled periods is greater than 2. If *J* is increased beyond 4, the CN drops under 3.8. Furthermore, inequality (16) can be approximated by

$$cond(\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0) \leq 2\frac{1 + \frac{0.75}{J}}{1 - \frac{1.5}{J}} \approx 2 + \frac{4.5}{J} \quad \text{if} \quad J \text{ is large.} \tag{17}$$

It can be seen that the condition number is small. Consequently, it is not important to use computationally more demanding decomposition methods.

The calculated maximum for the CN of $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ is a worst case limit. In practical situations, the CN is usually smaller than this limit. This is shown for different record lengths, if $J/N = 1/1000$ in Fig. 1. It can be observed that if at least 2 periods are sampled, the CN is under 2.25 and the CN asymptotically approaches 2. Furthermore, for coherent sampling, the CN is exactly 2, because in this case $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ is diagonal.

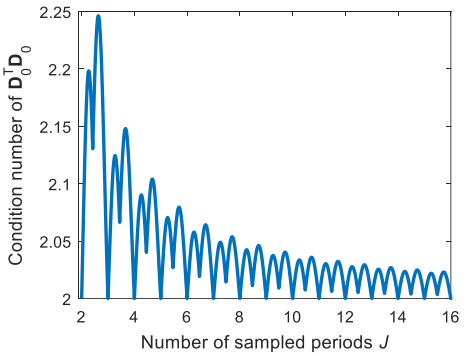

Fig. 1. CN of $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ for different record lengths for the given example

B. The four-parameter case

In the 4PF, all the four parameters are estimated at the same time. To solve this problem, an iterative solution is needed. In [1] Taylor series expansion around the frequency estimate in iteration step *i* is suggested for frequency estimation. The problem can be described as

$$\mathbf{x} = \mathbf{D}_i\mathbf{s}_i + \mathbf{e}, \tag{18}$$

$$\mathbf{D}_i = \begin{pmatrix} \cos\varphi_0 & \sin\varphi_0 & 1 & D_{i,14} \\ \cos\varphi_1 & \sin\varphi_1 & 1 & D_{i,24} \\ \vdots & \vdots & \vdots & \vdots \\ \cos\varphi_{N-1} & \sin\varphi_{N-1} & 1 & D_{i,N4} \end{pmatrix},$$

$$D_{i,k4} = 2\pi(-A(k-1)\ sin\,\varphi_{k-1} + B(k-1)\,cos\,\varphi_{k-1}) \tag{19}$$

$$\varphi_k = k2\pi f_i/f_s$$

$$\mathbf{s_i^T} = \left(A_i \quad B_i \quad C_i \quad \Delta\left(\frac{f_i}{f_s}\right)\right), \tag{20}$$

The matrix equation can be simplified, if $\Delta\omega_i/f_s$ is searched instead of $\Delta(f_i/f_s)$, where ω is the angular frequency, and $\omega = 2\pi f$. In this case $\boldsymbol{s}_{i,4} = \Delta(\omega_i/f_s)$ and

$$D_{i,k4} = -A\,(k-1)sin\,\varphi_{k-1} + B(k-1)cos\,\varphi_{k-1}. \tag{21}$$

After the solution of (18) the frequency ratio has to be updated

$$\omega_{i+1}/f_s = \omega_i/f_s + \Delta(\omega_i/f_s). \tag{22}$$

Then, $\mathbf{D}_{i+1}$ has to be constructed with the updated frequency ratio.

The 4PF is more complex than the 3PF. The conditioning of the 4PF system matrix (19) is much worse than that of the 3PF system matrix, since

$$\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i = \mathbf{H} = N\left(\widetilde{\mathbf{H}} + \mathbf{E}\right), \tag{23}$$

where

$$\widetilde{\mathbf{H}} = \begin{pmatrix} 1/2 & 0 & 0 & BN/4 \\ 0 & 1/2 & 0 & -AN/4 \\ 0 & 0 & 1 & 0 \\ BN/4 & -AN/4 & 0 & R^2N^2/6 \end{pmatrix}, \tag{24}$$

R is the aggregated amplitude: $R = \sqrt{A^2 + B^2}$. Again, with increasing J, the elements of $\mathbf{E}$ approach zero. Furthermore, if sampling is coherent, that is, if J is an integer number, $\mathbf{E}$ is a null-matrix. However, for the 4PF, $\mathbf{H}$ cannot be approximated with a diagonal $\widetilde{\mathbf{H}}$, as it was done for the 3PF.

It is easy to construct ill-conditioned examples, like a condition number of 10^{20}, with the following parameter setting:

$$y_k = s_0(1)\cdot\cos\varphi_k + s_0(2)\cdot\sin\varphi_k + s_0(3) \quad k = 0\ldots N-1, N = 10^6$$

$$\mathbf{s}_0^{\mathrm{T}} = (20000 \quad 25000 \quad 2^{15}) \tag{25}$$

and $J/N = 0.001$. Since the sampling is coherent, $\mathbf{E} = \mathbf{0}$. With the given data, the CN is

$$\mathrm{cond}\left(\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i\right) = 1.37\cdot 10^{21}, \tag{26}$$

so the problem is ill-conditioned. In addition, the CN is even larger than $1/eps_d = 4.5\cdot 10^{15}$, where eps_d is the resolution of double precision number representation [22]. This implies that the pseudo inverse should not be calculated as in (7).

However, it is visible from (24) that the CN of the 4PF depends on the amplitude of the signal and the record length. The higher the amplitude or the longer the record is, the worse the conditioning of $\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i$ will be. Furthermore, it is clear that $\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i$ is not diagonal, even for coherent sampling. This is caused by the added fourth parameter. To reach diagonality,

e.g., the Gram-Schmidt orthogonalization could be used to ensure robustness [18]. In Section III another method will be described, with the help of which $\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i$ can be approximated with a diagonal matrix.

C. Improvement of conditioning by scaling for the four-parameter case

In Section II-B it was pointed out that the CN of the 4PF depends on the amplitude of the sine and on the record length. The problem is caused by the fact that while parameters *A*, *B* and *C* are connected to quantities in the same dimension (to a voltage value on the analog side), f/f_s is connected to the signal/sampling frequency. The fourth parameter is the needed frequency change. Thus, it is connected to the derivative of the signal model with respect to f/f_s, see (19). It follows that the fourth column is proportional to sampling instant *k*. Since *k* goes from 0 to *N*, it can cover several orders of magnitude. Furthermore, the derivative is also proportional to *A* and *B*. These parameters can also assume values in a wide range. This explains the well-conditioning of 3PF and the ill-conditioning of 4PF.

The problem can be solved by proper scaling of (19), [16]. In [19] it was shown that the condition number of $\mathbf{D}_i^{\mathrm{T}}$ for long records can be optimized to approximately 3.7. It follows that the CN of $\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i$ can be optimized to approximately 14. Furthermore, it was pointed out in [20] that scaling the fourth column results in a much better-conditioned matrix. This means that the fourth column in (19) is divided by a scaling factor γ. The division affects the last row and the last column of $\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i$

$$\left(\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i\right)_{sc} = N\left(\widetilde{\mathbf{H}}_{sc} + \mathbf{E}_{sc}\right), \tag{27}$$

where

$$\widetilde{\mathbf{H}}_{sc} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{BN}{4\gamma} \\ 0 & \frac{1}{2} & 0 & -\frac{AN}{4\gamma} \\ 0 & 0 & 1 & 0 \\ \frac{BN}{4\gamma} & -\frac{AN}{4\gamma} & 0 & \frac{R^2N^2}{6\gamma^2} \end{pmatrix}, \tag{28}$$

$\mathbf{E}_{sc}$ contains the scaled values of the fourth row and column of **E** in (23). In order to solve the same equation system, the fourth parameter in (20) has also to be changed

$$\mathbf{s}_{i,sc}^{\mathbf{T}} = \begin{pmatrix} A_i & B_i & C_i & \gamma \cdot \Delta(\omega_i/f_s) \end{pmatrix}. \tag{29}$$

Now, the scaling factor has to be found, for which the condition number of $\left(\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i\right)_{sc}$ is minimal. The CN of $\left(\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i\right)_{sc}$ equals to that of $\widetilde{\mathbf{H}}_{sc} + \mathbf{E}_{sc}$. In the following it is assumed that $\mathbf{E}_{sc} = \mathbf{0}$. Thus, the CN of $\widetilde{\mathbf{H}}_{sc}$ has to be investigated. To this aim, singular values σ_i of (28) have to be determined as a function of γ. In Section II-A it was shown that $\sigma_i = \lambda_i$. Eigenvalues λ_i are given by the characteristic polynomial of (28)

$$\begin{aligned} C(\lambda) = [0.5-\lambda]\left\{(0.5-\lambda)(1-\lambda)\left(\frac{R^2}{6\gamma^2}N^2-\lambda\right) - \frac{AN}{4\gamma}\frac{AN}{4\gamma}(1-\lambda)\right\} \\ -\frac{BN}{4\gamma}(0.5-\lambda)\frac{BN}{4\gamma}(1-\lambda) = \\ =(0.5-\lambda)(1-\lambda)\left\{(0.5-\lambda)\left(\frac{R^2}{6\gamma}N^2-\lambda\right) - \frac{A^2N^2}{16\gamma^2} - \frac{B^2N^2}{16\gamma^2}\right\}= \end{aligned} \tag{30}$$

$$= C(\lambda) = (0.5-\lambda)(1-\lambda)\left\{(0.5-\lambda)\left(\frac{R^2}{6\gamma^2}N^2-\lambda\right)-\frac{R^2}{16\gamma^2}N^2\right\}.$$

It follows that $\sigma_1 = 1$ and $\sigma_2 = 0.5$ are always singular values. From the third term the other two singular values can be determined

$$C_2(\lambda) = (0.5-\lambda)\left(\frac{R^2}{6\gamma^2}N^2-\lambda\right)-\frac{R^2}{16\gamma^2}N^2 = \frac{R^2N^2}{12\gamma^2} - 0.5\lambda - \frac{R^2N^2}{6\gamma^2}\lambda + \lambda^2 - \frac{R^2N^2}{16\gamma^2} = \\ = \lambda^2 - \left(\frac{R^2N^2}{6\gamma^2}+0.5\right)\lambda + \frac{R^2N^2}{48\gamma^2} \tag{31}$$

The third and fourth roots of the characteristic equation are:

$$\sigma_{3,4} = \frac{\frac{R^2N^2}{6\gamma^2}+0.5 \pm \sqrt{\left(\frac{R^2N^2}{6\gamma^2}+0.5\right)^2 - \frac{R^2N^2}{12\gamma^2}}}{2}. \tag{32}$$

Let us use notation $z = \frac{R^2N^2}{\gamma^2}$. After simplifications we get:

$$\sigma_{3,4} = \frac{\frac{z}{6}+0.5 \pm \sqrt{\frac{z^2}{36}+\frac{z}{12}+0.25}}{2}. \tag{33}$$

If z is close to zero, $\sigma_3 \approx 0.5$, and σ_4 get close to 0, too. Thus, the problem becomes ill-conditioned because of σ_4. Contrarily, if z is large, $\sigma_3 \approx z/6$, and $\sigma_4 \approx 0.25$. Thus, the conditioning of the problem becomes ill with increasing z, too. The minimal condition number can be achieved at z=3.429, see Fig. 2. The optimal scaling factor that minimizes the CN of the problem is therefore:

$$\gamma_{opt} = RN/\sqrt{3.429} = RN/1.852, \tag{34}$$

and the CN of $\widetilde{\mathbf{H}}_{sc}$ is 14.0 in this case. This is a major improvement, compared to (26). The result shows that the minimal CN given in [19] can be reached. As expected, the scaling factor depends on the record length and the aggregated amplitude. It should be noted the result coincides with [19], except (34) also contains record length *N*. However, this difference is significant.

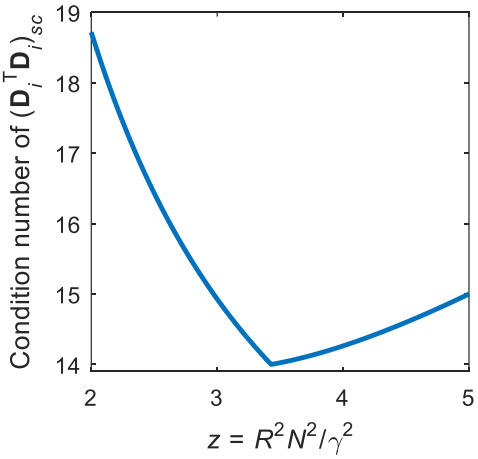

Fig. 2. Condition number of $(\mathbf{D}_i^T\mathbf{D}_i)_{sc}$ as a function of the scaling factor

Notice that γ_{opt} is given for the case when $\mathbf{s}_{i,4} = \Delta\,(\omega_i/f_s)$. If $\boldsymbol{s}_{i,4} = \Delta\,(f_i/f_s)$, the optimal scaling factor is

$$\gamma_{opt,freq.} = \gamma_{opt}/(2\pi) = RN/11.62. \tag{35}$$

However, this derivation holds only if $\mathbf{E}_{sc} = \mathbf{0}$. In order to show the gained enhancement in case $\mathbf{E}_{sc} \neq \mathbf{0}$, a Monte Carlo simulation was carried out with the following parameters: J/N was uniformly distributed in $[10^{-3}; 0.25]$, A and B were uniformly distributed in $[0; 20000]$. The simulation was run for 10^5 times for two different cases: first, the number of sampled periods was uniformly distributed in $[4; 5]$. For the second time, it was in the interval $[100; 101]$. Results can be seen in Fig. 3.

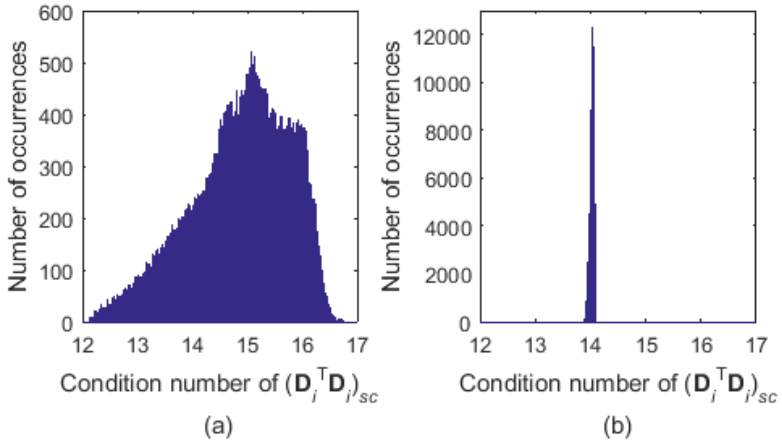

Fig. 3. Histogram of the CN of $\left(\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i\right)_{sc}$ if (a) $4 \leq J \leq 5$ and (b) $100 \leq J \leq 101$

It can be observed that if at least 4 periods are sampled, the CN of $\left(\mathbf{D}_i^{\mathbf{T}}\mathbf{D}_i\right)_{sc}$ is less than 17. For increasing record length, as expected, the conditions of the approximation are fulfilled better. For $100 \leq J \leq 101$, the CN is between 13.9 and 14.1.

III. Centering time instants

A. *Description of the method*

In Section II it was pointed out that for the 3PF matrix $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ can be approximated with a diagonal matrix. In the following a new technique is introduced in order to be able to approximate the 4PF matrix $\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i$ with a diagonal one, too.

As described in (3), in practical applications mostly uniform sampling occurs and for the measurement k goes from 0 to $N-1$. However, [1] does not define time instants t_k. Since the offset of time is usually not important in time invariant systems, the starting value of k can be chosen arbitrarily. This way, an offset is given to the time axis, similarly to data centering for the case of polynomial fitting [23]. Thus, time instants of the measurement can be transformed so that they become symmetrical to zero. By this, for coherent sampling the 4^{th} column of $\mathbf{D}_i$ will be shown to be orthogonal to its other columns. Thus, we can expect improvement on the condition number.

With data centering for time instants, $t=0$ can be shifted to the middle of the data set. Generally, the needed time offset l can be calculated as:

$$l = (N-1)/2. \tag{36}$$

After this data centering, the following parameters will be determined:

$$\left(\mathbf{s}_i^{\mathrm{T}}\right)' = \left(A_i' \;\; B_i' \;\; C_i \;\; \Delta\left(\omega_i/f_s\right)\right), \tag{37}$$

that is, the offset and the fine tuning of the frequency remain unchanged, while A_i' and B_i' are the amplitudes at *t=0*. With the new parameters, the time domain signal can be written as:

$$y_k = A' \cdot \cos(\varphi_{k-l}) + B' \cdot \sin(\varphi_{k-l}) + \mathrm{C}. \tag{38}$$

Notice that the index of *y* is unchanged, since data centering does not influence the fitted sine wave as a time domain signal. The original parameters can be calculated with the new parameters by

$$A = A'\cos\left(2\pi\frac{f}{f_s}l\right) - B'\sin\left(2\pi\frac{f}{f_s}l\right), \;\; B = A'\sin\left(2\pi\frac{f}{f_s}l\right) + B'\cos\left(2\pi\frac{f}{f_s}l\right), \tag{39}$$

Naturally, the aggregated amplitude remains unchanged:

$$R = \sqrt{A'^2 + B'^2} = \sqrt{A^2 + B^2}. \tag{40}$$

This technique becomes advantageous if every data point is used, i.e., there is no discarded sample. Namely, in this case it can be exploited that due to symmetry, the sum of odd functions is exactly 0.

Both three- and four-parameter LS matrices contain sums of functions of sine and cosine values. Consequently, every sum contains elements of even or odd functions. Data centering sets time parameters so that the sampling instants are symmetrical to zero. For odd functions, for instance, for $\sin(\phi)$, the following equation holds:

$$\sum_{k=0}^{N-1} \sin(\varphi_{k-l}) = 0. \tag{41}$$

Notice that the sum is exactly zero. Thus, there is no need to calculate it. Similarly:

$$\sum_{k=0}^{N-1} \sin(\varphi_{k-l})\cos(\varphi_{k-l}) = \sum_{k=0}^{N-1} \frac{1}{2}\sin(2\varphi_{k-l}) = 0. \tag{42}$$

After describing data centering for time instants, its effect on both the 3PF and 4PF will be demonstrated. In order to further increase numerical stability, the original algorithm is also slightly modified at another point. It can be seen from (10) that the conditioning of $\widetilde{\mathbf{H}}$ is not optimal. It could be improved if the offset parameter were scaled, similarly to the case of the 4PF in Section II-C. Now, the third column of $\mathbf{D'}_0$ and $\mathbf{D'}_i$ should be divided by $\sqrt{2}$ to ensure that the diagonal elements in $\widetilde{\mathbf{H}}'$ are equal to each other. By this means, its condition number can be decreased to 1. Certainly, the parameter vector is also modified

$$\left(\mathbf{s}_i^{\mathrm{T}}\right)' = \left(A_i' \;\; B_i' \;\; C_i\sqrt{2} \;\; \Delta\left(f_i/f_s\right)\right), \tag{43}$$

After this modification, let us observe the effect of data centering.

B. Condition number enhancement for the 3PF

The effect of data centering and dividing the third column of $\mathbf{D}_0$ by$\sqrt{2}$ on the 3PF is the following. Similarly to the original problem, the system matrix can be described by

$$\mathbf{H}' = N\left(\widetilde{\mathbf{H}}' + \mathbf{E}'\right), \tag{44}$$

where

$$\widetilde{\mathbf{H}}' = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}, \tag{45}$$

and $\mathbf{E}'$ can be bounded by

$$\mathbf{E}'_b = \frac{1}{J}\begin{pmatrix} \pm 1/8 & 0 & \pm 1/4 \\ 0 & \pm 1/8 & 0 \\ \pm 1/4 & 0 & 0 \end{pmatrix}, \tag{46}$$

provided that (12) holds, see [26]. According to the matrix perturbation theory (14) we have

$$\text{cond}\left\{(\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0)'\right\} = \text{cond}\left\{\left(\widetilde{\mathbf{H}} + \mathbf{E}\right)'\right\} \leq \frac{0.5 + \|\mathbf{E}'_b\|_F}{0.5 - \|\mathbf{E}'_b\|_F} = \frac{0.5 + \dfrac{0.4}{J}}{0.5 - \dfrac{0.4}{J}}, \quad \text{if } J > 0.8 \tag{47}$$

This can be approximated as

$$\text{cond}\left\{(\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0)'\right\} \leq \frac{0.5 + \dfrac{0.4}{J}}{0.5 - \dfrac{0.4}{J}} \approx 1 + \frac{1.6}{J}, \quad \text{if } J \text{ is large} \tag{48}$$

Comparing (17) to (48), the upper bound of the condition number has been decreased. In fact, with data centering and scaling the third column of $\mathbf{D}'_0$, the condition number is always lower than 1.5, if at least 4 periods are sampled, and at least 4 samples are sampled from one period.

C. Condition number enhancement for the 4PF

Similarly to the 3PF, the effect of time offsetting and scaling the third column of $\mathbf{D}'_i$ can be investigated. For the 4PF we have

$$\mathbf{H}' = N\left(\widetilde{\mathbf{H}}' + \mathbf{E}'\right), \tag{49}$$

where

$$\widetilde{\mathbf{H}}' = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & \dfrac{R^2(N^2-1)}{24} \end{pmatrix}, \tag{50}$$

and $\mathbf{E}'$ contains the error terms of $\widetilde{\mathbf{H}}'$, [26]. It is obvious that with increasing record length or for large amplitude values, $\widetilde{\mathbf{H}}'$ becomes ill-conditioned. Thus, the fourth column of $\mathbf{H}'$ should be divided by

$$\gamma' = \sqrt{\frac{R^2(N^2-1)}{12}}. \tag{51}$$

After this scaling we have

$$\mathbf{H}'_{\mathrm{sc}} = N\left(\widetilde{\mathbf{H}}'_{sc} + \mathbf{E}'_{sc}\right), \tag{52}$$

where

$$\widetilde{\mathbf{H}}'_{sc} = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{pmatrix}, \tag{53}$$

and it can be shown that $\mathbf{E}'_{sc}$ can be bounded by matrix $\mathbf{E}'_{sc,b}$, and

$$\left\|\mathbf{E}'_{sc,b}\right\|_F \leq \frac{0.96}{J}, \text{ if } J \geq 4 \tag{54}$$

provided that (12) holds, see [26]. Notice here that this bound is valid only if at least 4 periods are sampled. According to the matrix perturbation theory, [21], we have

$$\mathrm{cond}(\mathbf{H}'_{sc}) = \mathrm{cond}\left(\left(\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i\right)'_{sc}\right) \leq \frac{0.5 + \frac{0.96}{J}}{0.5 - \frac{0.96}{J}}, \text{ if } J \geq 4. \tag{55}$$

This can be approximated as

$$\mathrm{cond}(\mathbf{H}'_{sc}) \leq \frac{0.5 + \frac{0.96}{J}}{0.5 - \frac{0.96}{J}} \approx 1 + \frac{3.84}{J}, \text{ if } J \text{ is large.} \tag{56}$$

This bound means that the condition number is guaranteed to be under 2.85, if at least 4 periods are sampled, and at least 4 samples are sampled from one period. Notice that for coherent sampling the condition number can be decreased to 1, that is, to the smallest possible value.

IV. Proposals on the evaluation of Least Squares algorithms

In this section, the evaluation of 3PF and 4PF is investigated. Since in the previous sections it was pointed out that the algorithms are well-conditioned, at least after proper scaling for the 4PF, the direct evaluation of the pseudo-inverse calculation can be utilized. For this evaluation, $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ or $\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i$ has to be calculated. The proposals on the evaluation is given for the 3PF, but they also hold for the 4PF. Thus, in the following $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$ will be investigated

$$\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0 = \mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{12} & h_{22} & h_{23} \\ h_{13} & h_{23} & h_{33} \end{pmatrix}. \tag{57}$$

First, it should be noticed that this matrix is symmetric. It follows that 6 elements have to be calculated instead of 9. Furthermore, h_{33} equals to N. Thus, there is no need to calculate it. Finally, using trigonometric identities

$$h_{11} = N - h_{22}, \tag{58}$$

that is, only one of these elements has to be computed. To conclude, only 4 elements from 9 has to be determined in order to be able to construct $\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0$. For the 4PF 8 elements from 16 have to be calculated to construct $\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i$.

If data centering is also used, h_{12} and h_{23} are exactly 0. Thus, for the 3PF 2 elements, for the 4PF 6 elements have to be determined. From a numerical point of view, it is beneficial if it is known that the value of an element is exactly 0. In this case, the result is exact and it is not distorted by the roundoff errors that are accumulating when summations are performed [24], or the roundoff error due to imprecise phase evaluation cannot be neglected [25].

Furthermore, roundoff errors can be decreased for all matrix entries in (57), considering the following. Every element is a sum of cosinusoidal and sinusoidal functions that are equally sampled. (The only exception is h_{33}, but as it was discussed, it is always equal to *N*.) For these sums, a closed formula can be derived [26]. For example:

$$h_{11} = \sum_{k=0}^{N-1} \cos^2 k\varphi_1 = \frac{N}{2} + \frac{\cos(N-1)\varphi_1 \sin N\varphi_1}{2\sin\varphi_1}. \tag{59}$$

By this means, there is no need for summation. The sums can be determined with a few multiplications and additions. Thus, the result is not distorted by the roundoff error of the summation. In addition, roundoff errors due to imprecise phase evaluation can be eliminated, provided phase information $(N-1)\varphi_1$ is determined precisely. A detailed description for this error and a method for calculating precise phase information can be found in [25].

It follows that accuracy is also improved. This is especially important for large values of *N* and/or for single precision evaluation, [20].

To conclude, results in Sections III and IV can be summed up in the following steps:

- Use data centering for time instants, that is, shift $t = 0$ to the middle of the data set.
- Divide the third column of $\mathbf{D'}_0$ or $\mathbf{D'}_i$, that is, system matrices after data centering, by $\sqrt{2}$.
- In case of the 4PF, divide the fourth column of $\mathbf{D'}_i$ by γ' in (51).
- For the calculation of the sums in $(\mathbf{D}_0^{\mathrm{T}}\mathbf{D}_0)'$ or $\left(\mathbf{D}_i^{\mathrm{T}}\mathbf{D}_i\right)'$ use closed formulas, and also take into consideration that these matrices are symmetric. Closed formulas can be found in [26].
- Calculate the pseudo-inverse directly.
- Calculate the parameter vector.
- To get the original parameter vector, divide the offset by $\sqrt{2}$ and in case of the 4PF, divide the frequency change by γ'. To get original parameters *A* and *B*, use (39).

CONCLUSIONS

In this paper three- and four-parameter Least Squares sine fitting algorithms were investigated. It was proved that the three-parameter fitting is well-conditioned, if at least four periods are sampled, and at least four samples are sampled in a period. Then for the conditioning of the four-parameter fitting, a scaling factor was given which minimizes the condition number under some conditions. A Monte Carlo analysis was carried out to show the effectiveness of the scaling. Data centering for time instants was introduced. It was shown that with this technique, the 4PF problem can be approximated with a diagonal matrix. Furthermore, it was proved that the condition number of the 3PF is lower than 1.5, and the condition number of the 4PF is lower than 3, if at least 4 periods are sampled and at least 4 parameters are sampled from a period. For long records, the condition number of both methods was shown to approach 1. Furthermore, for coherent sampling, the condition number of 1 is reached. This value is the smallest possible condition number. Thus, both algorithms can be evaluated in a numerically stable way. Finally, an evaluation method was given for both the three- and four-parameter

methods. Proof for the condition number calculations can be found in [26]. Furthermore, source files of the described algorithms implemented in MATLAB are available in [27].

ACKNOWLEDGEMENT

This work was partially supported by the Hungarian Research Fund (OTKA) under Grant K–115820. The authors thank the anonymous reviewers and Tadeusz Dobrowiecki for their suggestions which helped to improve the paper.

REFERENCES

[1] Standard IEEE-1241-2010, "IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters" (2011). doi: 10.1109/IEEESTD.2011.5692956

[2] J. Blair. "Histogram measurement of ADC nonlinearities using sine waves", *IEEE Trans. Instrum. Meas.*, vol. 43, no. 3, pp.:373-383, 1994. doi: 10.1109/19.293454

[3] T. Dabóczi, "ADC Testing Using a Resonator-Based Observer: Processing Very Long Time Records and/or Testing Systems With Limited Stability", *IEEE Trans. Instrum. Meas.*, vol. 62, no. 5, pp. 1166-1173, Dec. 2012, doi:10.1109/TIM.2012.2224276

[4] N. Giaquinto, A. Trotta, "Fast and Accurate ADC Testing Via an Enhanced Sine Wave Fitting Algorithm", *IEEE Trans. Instrum. Meas.*, vol. 46, no. 4., pp. 1020-1025, 1997, doi: 10.1109/19.650820

[5] J. Schoukens, R. Pintelon, H. Van Hamme, "The interpolated fast fourier transform: A comparative study", *IEEE Trans. Instrum. Meas.*, vol. 41, no. 2, pp 226-232, 1992. doi: 10.1109/19.137352

[6] T. Andersson, P. Händel, "IEEE Standard 1057, Cramér-Rao bound and the parsimony principle", *IEEE Trans. Instrum. Meas.*, vol. 55, no. 1, pp 44-53, 2006. 10.1109/TIM.2005.861497

[7] J. Deyst; T. Sounders; O. Solomon, „Bounds on least-squares four-parameter sine-fit errors due to harmonic distortion and noise", *IEEE Trans Instrum. Meas.*, Vol. 44, No. 3, pp. 637-642, Jun. 1995, doi: 10.1109/19.387298

[8] 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, doi: 10.1016/j.measurement.2008.12.006

[9] P. Händel, "Amplitude estimation using IEEE-STD-1057 three-parameter sine wave fit: Statistical distribution, bias and variance," *Measurement*, vol. 43, no. 6, pp. 766–770, Jul. 2010, doi: 10.1016/j.measurement.2010.02.007

[10] A. Moschitta, J. Schoukens, P. Carbone, "Parametric System Identification Using Quantized Data", *IEEE Trans. Instrum. Meas*., vol. 64, no. 8, pp. 2312-2322, Aug. 2015, doi: 10.1109/TIM.2015.2390833

[11] P. Carbone, J. Schoukens, I. Kollár, A. Moschitta, "Accurate Sine-Wave Amplitude Measurements", *IEEE Trans. Instrum. Meas.*, vol. 64, no. 12, pp. 3201-3208, Dec. 2015, doi: 10.1109/TIM.2015.2463331

[12] Z. Bilau, T. Megyeri, A. Sárhegyi, J. Márkus, I. Kollár, „Four-parameter fitting of sine wave testing result: iteration and convergence", *Computer Standards & Interfaces*, vol. *26, no.* 1, pp. 51-56, 2004, doi: 10.1016/S0920-5489(03)00062-X

[13] L. Balogh, I. Kollár, A. Sárhegyi. "Maximum likelihood estimation of ADC parameters." Proceedings of *IEEE Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 24-29, Austin, USA, 2010, doi: 10.1109/IMTC.2010.5488286

[14] J. Šaliga, I. Kollár, L. Michaeli, J. Buša, J. Lipták, T. Virosztek "A comparison of least squares and maximum likelihood methods using sine fitting in ADC testing." *Measurement* Vol. 46, No. 10, pp. 4362-4368, 2013, doi: 10.1016/j.measurement.2013.05.004

[15] I. Kollár, 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, Oct. 2005. doi: 10.1109/IMTC.2004.1351190

[16] G. Allaire, S. Kaber, "Numerical Linear Algebra", *Springer*, New York, USA, 2008

[17] W. Stewart., *"*An Introduction to Matrix Computations*", Academic Press*, New York 1973.

[18] K. Chen, Y. Xue, "Four-parameter sine wave fitting by Gram–Schmidt orthogonalization", Measurement, vol. 41, pp. 76-87, 2008. doi:10.1016/j.measurement.2006.11.016

[19] K. Chen, Y. Xue, "Improving four-parameter sine wave fitting by normalization", Computer Standards and Interfaces, vol. 29, pp. 184-190, 2007. doi:10.1016/j.csi.2006.05.005

[20] P. Ramos, T. Radil, F. Janeiro, "Implementation of sine-fitting algorithms in systems with 32-bit floating point representation", Measurement 45, pp. 155-163, 2012. doi:10.1016/j.measurement.2011.05.011

[21] Ren-Cang Li, "Relative perturbation theory: I. Eigenvalue and singular value variations." *SIAM Journal on Matrix Analysis and Applications* vol. 19, No. 4, pp. 956-982, 1998, doi: 10.1137/S089547989629849X

[22] B. Widrow, I. Kollár, "Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications", *Cambridge University Press*, Cambridge, UK, 2008. DOI: 10.1017/CBO9780511754661

[23] S. Kay, "Fundamentals of Statistical Signal Processing", Prentice-Hall PTR, Englewood Cliffs, NJ, 1993

[24] N. J. Higham, “The accuracy of floating point summation”, SIAM Journal on Scientific Computing 14 (4): pp. 783–799, 1993. doi: 10.1137/0914050

[25] B. Renczes, I. Kollár, A. Moschitta, P. Carbone, „Numerical Optimization Problems of Sine-Wave Fitting Algorithms in the Presence of Roundoff Errors”, *IEEE Trans. Instrum. Meas.*, , Vol. 65, No. 8 pp. 1785-1795, 2016, doi: 10.1109/TIM.2016.2562218

[26] B. Renczes, I. Kollár, “Derivation for upper bounds of condition numbers for three- and four-parameter Least Squares Sine Fitting Algorithms”, http://home.mit.bme.hu/~renczes/pubinfo/TIM2016

[27] B. Renczes, I. Kollár, “Source files for Efficient Implementation of Least Squares Sine Fitting Algorithms”, 2016. URL: https://github.com/renczes/TIM2016_Efficient_LS/