FAST ONE-CYCLE FREQUENCY ESTIMATION OF A SINGLE SINUSOID IN NOISE USING DOWNSAMPLED LINEAR PREDICTION MODEL

Krzysztof Duda1), Tomasz P. Zieliński2)

1) AGH University of Science and Technology, Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Department of Measurement and Electronics, al. Mickiewicza 30, 30-059 Kraków, Poland (kluda@agh.edu.pl, +48 126 172 841)
2) AGH University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Institute of Telecommunications, al. Mickiewicza 30, 30-059 Kraków, Poland (tzielin@agh.edu.pl)

Abstract

A new solution to the problem of frequency estimation of a single sinusoid embedded in the white Gaussian noise is presented. It exploits, approximately, only one signal cycle, and is based on the well-known 2nd order autoregressive difference equation into which a downsampling is introduced. The proposed method is a generalization of the linear prediction based Prony method for the case of a single undamped sinusoid. It is shown that, thanks to the proposed downsampling in the linear prediction signal model, the overall variance of the least squares solution of frequency estimation is decreased, when compared to the Prony method, and locally it is even close to the Cramér–Rao Lower Bound, which is a significant improvement. The frequency estimation variance of the proposed solution is comparable with, computationally more complex, the Matrix Pencil and the Steiglitz–McBride methods. It is shown that application of the proposed downsampling to the popular smart DFT frequency estimation method also significantly reduces the method variance and makes it even better than the least squares smart DFT. The noise immunity of the proposed solution is achieved simultaneously with the reduction of computational complexity at the cost of narrowing the range of measured frequencies, i.e. a sinusoidal signal must be sufficiently oversampled to apply the proposed downsampling in the autoregressive model. The case of 64 samples per period with downsampling up to 16, i.e. 1/4th of the cycle, is presented in detail, but other sampling scenarios, from 16 to 512 samples per period, are considered as well.

Keywords: frequency estimation, linear prediction, Prony method, smart DFT.

© 2021 Polish Academy of Sciences. All rights reserved

1. Introduction

Frequency estimation of a sinusoidal signal disturbed by additive zero-mean white Gaussian noise (WGN) is a fundamental signal processing task [1–7]. Optimal, i.e. minimum variance and unbiased, frequency estimator has a form of a nonlinear least squares (LS) solution [1] which must be computed via optimization. Linear LS solution can only be obtained for \( \cos(\omega_0) \),
and not for frequency $\omega_0$ alone, as e.g. in following methods: the Prony, the Pisarenko, the Steiglitz–McBride, the Kumaresan-Tufts, the Matrix Pencil, etc. [2–5,8–13].

The Discrete Fourier Transform (DFT) is widely used for frequency estimation because it has low computational complexity and good immunity to sinusoidal disturbances. Unfortunately, the DFT frequency estimation is biased by the spectral leakage and the picket-fence effect [7]. The spectral leakage is mitigated by the use of time windows with high attenuation of spectral sidelobes or by iterative compensation, e.g. [7,14,15], and the picket-fence effect is reduced by the interpolated DFT (IpDFT) algorithms, e.g. [4–6,16,17].

One-cycle, leakage-free frequency estimation can be obtained by the smart DFT (SDFT) [18–21] which gained popularity for P-class phasor estimation in electric power systems. The frequency in the SDFT is estimated based on three DFT bins, computed for three consecutive signal fragments selected by the rectangular observation window shifted by one sample.

Other, recently published one-cycle frequency estimators include the two complex bandpass filters method [22], and the Discrete-Time Frequency-Gain Transducer [23].

In this paper we propose a conceptually new solution for frequency estimation of an approximately one cycle sinusoid which can be applied alternatively in the time domain or in the frequency domain. In both cases we introduce downsampling into the well-known sinusoidal autoregressive model. Thanks to this, we extend the linear prediction based Prony method to a single undamped sinusoid, defined in the time domain, and also enhance the SDFT method, defined in the frequency domain. In both cases a significant variance reduction is obtained in comparison with existing counterparts, i.e. the Prony method and the SDFT method [18,19]. The proposed solution outperforms also the least-squares SDFT (LS-SDFT) version [20] and locally, in some frequency intervals, it performs comparably to the best known methods, i.e. the Matrix Pencil and the Steiglitz–McBride [4,5,11–13] which are significantly more computationally complex.

One of the foreseen applications of the proposed methods is phasor estimation in P-class phasor measurement units (PMU) [24].

Section 2 describes the proposed autoregressive downsampling sinusoidal signal model. Section 3 presents the proposed time domain LS solution and the frequency domain DFT-based solution, both based on the proposed downsampling signal model. It also discusses noise propagation and computational complexity of the proposed methods. The results of conducted numerical experiments are given in Section 4, and Section 5 concludes the paper.

2. Autoregressive downsampling sinusoidal signal model

We assume the following model of a sinusoidal signal:

$$x_n = A_0 \cos (\omega_0 n + \varphi_0), \quad n = 0, 1, 2, \ldots, N-1,$$

where $A_0 > 0$ is the amplitude, $\varphi_0$ is the phase in radians, $n$ is integer index and $N$ is the number of samples, $\omega_0$ is the angular frequency in radians, and $\omega_0 = \pi \text{ rad}$ corresponds to the half of the sampling frequency in Hz. The frequency can be expressed as:

$$\omega_0 = \frac{2\pi}{N} k_0,$$

where a real value $k_0$ is the number of sinusoidal signal cycles in the observation interval.

The sinusoidal signal of interest (1) is observed in the presence of additive disturbance:

$$s_n = x_n + e_n,$$

where $s_n$ is the measured signal, and $e_n$ denotes the zero-mean WGN.
The sinusoidal signal (1) fulfills the following well-known autoregressive second order linear difference equation

\[ x_n = 2 \cos(\omega_0) x_{n-1} - x_{n-2}, \quad n = 2, 3, \ldots, N-1 \] (4)

which is used in many linear prediction based frequency estimation methods like Prony, Pisarenko, etc. see e.g. [2–5]. According to (4) the current value of a sinusoid is a linear combination of two previous values. We extend this self-predictive formula by introducing downsampling denoted by \( L = 1, 2, \ldots \):

\[ x_n = 2 \cos(\omega_0L) x_{n-L} - x_{n-2L}, \quad n = 2L, 2L + 1, \ldots, N-1. \] (5)

Thus, (5) is a downsampled version of (4). The sinusoidal signal of interest (1) must be \( L \) times oversampled for correct frequency estimation based on (5), i.e. it must contain more than \( 2L \) samples in one period. Equation (5) is next used for extension of the Prony method and the SDFT method for a single undamped sinusoid.

3. Proposed frequency estimation methods

3.1. Time domain LS solution – extension of the Prony method

Based on (5), an overdetermined set of linear equations is built and solved in the LS sense. The equations are:

\[ \mathbf{E} \mathbf{c} = \mathbf{b}, \] (6a)

where:

\[ \mathbf{E} = \begin{bmatrix} s_L \\ \vdots \\ s_{N-1-L} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} s_0 + s_{2L} \\ \vdots \\ s_{N-1-2L} + s_{N-1} \end{bmatrix}, \quad c = 2 \cos(\omega_0L) \] (6b)

and the LS solution is:

\[ c = (\mathbf{E}^T \mathbf{E})^{-1} \mathbf{E}^T \mathbf{b}. \] (7)

Finally, the frequency is given by:

\[ \omega_0 = \frac{L}{c} \cos^{-1} \left( \frac{c}{2} \right). \] (8)

For the considered undamped single sinusoid (1) for \( L = 1 \) the LS solution (7), (8) is the same as in the Prony method and in the Pisarenko method. The downsampling reduces the number of equations in (6), thus the computational burden decreases along with the increase of the \( L \) value. We also note that although the number of equations in (6) decreases with the increase of \( L \) still all measured samples \( s_n \) are used.

3.2. Frequency domain solution – extension of the SDFT method

By calculating the DFT bins with index \( k = 1 \) for vectors \( \mathbf{E} \) and \( \mathbf{b} \) in (6) the frequency is given by:

\[ \omega_0 = \frac{L}{2} \cos^{-1} \left( \frac{\text{DFT}_k=1(\mathbf{b})}{2 \text{DFT}_k=1(\mathbf{E})} \right). \] (9)

For \( L = 1 \) (9) defines the SDFT [18, 19].
3.3. Noise propagation

In this section approximate analysis of noise propagation through successive stages of the proposed algorithm is outlined and the noise suppression capability of the new downsampled LS method is justified.

Fig. 1 illustrates noise propagation through the first stage of the method, i.e. the LS computation (7), depending on the downsampling factor \( L \). The left plot depicts the standard deviation (STD) of coefficient \( c \) computed by (7) as a function of the STD of WGN (3) while the right plot shows the STD for a given \( L \) normalized by the STD for \( L = 1 \). It is observed that the best noise suppression and the lowest STD is obtained for \( L = 1 \), i.e., as expected, for the highest number of equations, namely \( N - 2L \). Thus at this stage the proposed downsampling is in fact counterproductive. Now, we show that this loss of noise immunity is compensated in excess by the second computational stage in which inverse cosine is calculated and division by \( L \) is done (8).

![Fig. 1. Noise propagation by the LS computation (7) in independence of the downsampling factor \( L \) (left), \( Ne \) in the legend denotes the number of equations solved in (7). The curves move up with the increase of \( L \). The ratio of the STD for \( L = 1, \ldots, 16 \) and the STD for \( L = 1 \) (right).](image)

The Taylor expansion of \( \cos^{-1} \left( \frac{c}{2} \right) \) in (8) is given by the formula:

\[
\cos^{-1} \left( \frac{c}{2} \right) = \frac{\pi}{2} - \frac{c}{2} - \frac{1}{6} \left( \frac{c}{2} \right)^3 - \frac{3}{40} \left( \frac{c}{2} \right)^5 - \frac{5}{112} \left( \frac{c}{2} \right)^7 \ldots.
\]

(10)

For the sinusoidal signal (1) disturbed by additive WGN the coefficient \( c \) estimated by the LS method (7) is also disturbed by additive noise, i.e.:

\[
c = c_{\text{true}} + e_c,
\]

(11)

where \( c_{\text{true}} \) denotes the true value and \( e_c \) is the noise. It results from (10) that the lowest STD of frequency estimation based on (8) is obtained for \( c \) close to 0. Fig. 2 (left) illustrates this observation by presenting the STD of the estimated frequency as a function of \( c \) for the STD of \( e_c \) equal to 0.05, 0.1, and 0.2. It is observed that the values of \( c \) close to \( \pm 2 \) should be avoided.

The coefficient \( c \) is close to 0 for \( \omega_0 L \) close to \( \pi/2 \) thus the value of \( L \) should be close to

\[
L = \frac{\pi}{2\omega_0} = \frac{\pi}{2 \frac{2\pi}{N} k_0} = \frac{N}{4k_0}.
\]

(12)
Fig. 2. STD of frequency estimation (8) as a function of \(c\) for three different STD values of noise \(e_c\) in (11) and \(L = 1\) (left). The ratio of the STD for \(L = 1\) and the STD for \(L = 1, \ldots, 16\) for STD of \(e_c\) equal to 0.1 (right). The curves move up with the increase of \(L\).

For example, for one-cycle sinusoid \(k_0 = 1\) and \(N = 64\) samples, the preferred value of \(L\) is 16, which is 1/4th of the cycle.

In (8) the value of \(\cos^{-1}(c/2)\) is divided by \(L\) which additionally suppresses the STD value. This effect is illustrated in Fig. 2 (right) for \(L = 1, \ldots, 16\) where the STD of frequency estimation for \(L = 1\) divided by the STD for \(L = 1, \ldots, 16\) is depicted. From this characteristic the STD suppression resulting from the introduced downsampling can be read out. For example, for \(L = 16\) the STD of frequency estimation is around 16 times smaller than for \(L = 1\). We also note, that the STD ratio in Fig. 2 (right) is computed for a fixed value of \(c\), that is for a fixed frequency in radians, but the proposed downsampling increases the signal frequency in radians, and even higher STD suppression can be obtained for a one-cycle signal, which is further illustrated by means of numeric simulations.

A similar reasoning can be conducted for the frequency domain solution, i.e. for the proposed downsampled SDFT (9). An increase of \(L\) decreases the length of the DFT which increases the STD, but next the inverse cosine close to \(\pi/2\) is computed and divided by \(L\) which significantly reduces the STD.

To summarize, the proposed downsampling reduces the variance of frequency estimation by:
– shifting the signal’s frequency closer to \(\pi/2\) and
– dividing the frequency estimate by downsampling factor \(L\).

3.4. Computational complexity of proposed methods

In general, a pseudoinverse matrix \((E^TE)^{-1}E^T\) in the LS solution (7) is computed with well-established methods as e.g. the singular value decomposition [25]. However, in the considered case of a single sinusoid, \(E\) is a vector, and therefore only two inner products of vectors having \(N - 2L\) elements plus one division are required in (7). In total we have \(4N - 8L + 1\) real value floating point operations, including division by \(L\) and by 2 and calculation of one inverse cosine.

In (9) two inner products of vectors having \(N - 2L\) elements are required also, but the DFT basis is complex valued which results in approximately two times higher computational load than the time domain LS solution (8).

Execution times of Matlab implementations of both methods are reported in the next section.
4. Results

In the following analysis we assume a sinusoidal signal containing up to 2 cycles which is disturbed by the additive WGN. Short-cycle frequency estimation is a typical protective and control application in electric power systems [24]. For a sinusoid with \( N = 64 \) samples and for \( L = 1 \) (i.e. the Prony method) a set of \( N - 2L = 62 \) equations (6) is solved, whereas for \( L = 16 \) there are only \( N - 2L = 32 \) equations (6) (see also the legend in Fig. 1).

4.1. Noise

Figs. 3, 4 (left) present the root-mean-square error (RMSE) for the proposed time domain and the frequency domain downsamplers in comparison with the Prony method (Fig. 3) and the SDFT (Fig. 4) in respect to the Cramér–Rao Lower Bound (CRLB) [1] plotted in dotted line. Figs. 3 and 4 (right) depict the RMSE ratio for \( L = 1 \) and \( L = 1, \ldots, 16 \), that is the RMSE reduction for \( L \) greater than 1. It is observed that by increasing the value of downsampling factor \( L \) the RMSE is significantly reduced especially for \( L \) below 8, i.e. \( 8/64 = 1/8 \)th of the cycle. For \( L \) above 8 the RMSE begins to increase for \( k \) close to 2, but the local minima are very close to the CRLB. It is observed that the time domain LS solution (Fig. 3) performs better than the frequency domain DFT based method (Fig. 4) as it is closer to the CRLB in longer frequency intervals, and thus it offers broader frequency measurement range. In both cases the counterparts methods i.e. the Prony and the SDFT (the case of \( L = 1 \)) are significantly improved with the increase of \( L \). Figs. 3, 4 (right) can be compared with Fig. 2 (right) after taking into account different frequency ranges in these figures. The range of \( k \in (0.5, 2) \) for \( N = 64 \) corresponds to frequency range \( \omega_0 \in (0.0491, 0.1963) \) rad. Because \( c = 2 \cos (\omega_0 L) \), the range of \( c \) for \( L = 1 \) is \( c \in (1.9976, 1.9616) \), while for \( L = 16 \) it is \( c \in (1.4142, -2) \). For \( L = 16 \) there is \( c = 0 \) for \( k = 1 \) and at this frequency the RMSE in Fig. 3 (left) is the lowest.

Figs. 5, 6 depict the RMSE of both proposed solutions for different signal lengths \( N = \{16, 32, 64, 126, 256, 512\} \) for \( L = \{1, N/8, N/4\} \). It is observed that a significant reduction of RMSE is obtained in comparison to the counterparts methods i.e. the Prony and the SDFT (left subplots, \( L = 1 \)). The proposed methods are scalable with the number of samples in a single cycle. For each considered case they reach the CRLB for similar frequency \( k \).
Fig. 4. RMSE of the proposed frequency domain DFT based method for different values of $L$ and fixed signal length $N = 64$ (left). $L = 1$ (solid black line) is the SDFT method. $SNR = 40$ dB. Black dotted line depicts the CRLB. The ratio of RMSE for $L = 1$ and $L = 1, \ldots, 16$ (right).

Fig. 5. RMSE of the proposed time domain LS based method for different signal lengths $N = \{16, 32, 64, 126, 256, 512\}$ and fixed $L = 1$, the Prony method (left), $L = N/8$ (middle), and $L = N/4$ (right). $SNR = 40$ dB. Dotted lines depict the CRLBs.

Fig. 7 depicts the RMSE to CRLB ratio for selected configurations of the proposed solution, i.e. LS with $L = \{8, 11, 16\}$ and DFT with $L = \{8, 11, 16\}$ in comparison with the LS-SDFT [20] with $R = 16$ equations and the Matrix Pencil and the Steiglitz–McBride methods with implementations taken from [5]. It is seen that the best results are obtained by the Steiglitz–McBride method. The proposed time domain LS solution is very close to the second best, the Matrix Pencil method, in a wide frequency interval, and the proposed DFT based solution is close to the Matrix Pencil method in a significantly shorter interval. The LS-SDFT algorithm performs the worst.

The shorter measurement interval of the proposed solutions is not necessarily a serious practical drawback because the measurement range can be selected at the local RMSE minimum in applications where the frequency deviation is measured around the nominal frequency value.
Fig. 6. RMSE of the proposed frequency domain DFT based method for different signal lengths \( N = \{16, 32, 64, 126, 256, 512\} \) and fixed \( L = 1 \), the SDFT (left), \( L = N/8 \) (middle), and \( L = N/4 \) (right). \( SNR = 40 \) dB. Dotted lines depict the CRLBs.

Fig. 7. RMSE to CRLB ratio as a function of frequency for \( N = 64 \) and \( SNR = 40 \) dB. The proposed time domain LS solution (left) and the proposed frequency domain solution (right), both for \( L = \{8, 11, 16\} \), \( R = 16 \) denotes the LS-SDFT method from [20], MatPen is the Matrix Pencil and STMB is the Steiglitz–McBride method.

For example, in electric power system for the nominal 50 Hz frequency the ±5 Hz, that is 10%, frequency deviation is assumed [24].

In Fig. 7 the proposed time domain LS solution \( L = 16 \) in the frequency interval \( k \in (0.9, 1.1) \), i.e. for nominal frequency \( k = 1 \) with 10% deviation, performs approximately 1.03 times better than the Matrix Pencil method and approximately 1.1 times worse than the Steiglitz–McBride method. For nominal frequency \( k = 1.5 \) with 10% deviation, i.e. \( k \in (1.35, 1.65) \), the performance of the proposed time domain LS solution \( L = 11 \) is practically the same as for the Matrix Pencil method and only approximately 1.05 times worse than the for Steiglitz–McBride method. The proposed frequency domain DFT based solution performs similarly for \( L = 11 \) but in the shorter frequency interval \( k \in (1.4, 1.6) \).
Fig. 8 depicts the comparison of the RMSE to CRLB ratio, for the methods considered, as a function of signal to noise ratio (SNR) for 10% frequency deviation intervals $k \in (0.9, 1.1)$ Fig. 8 (left) and $k \in (1.35, 1.65)$ Fig. 8 (right). The RMSE was computed for signal frequency uniformly distributed in these two intervals. For one-cycle estimation the proposed LS, $L = 16$ method performs better than the Matrix Pencil method and for SNR below approximately 30 dB even better than the Steiglitz–McBride method. For 1.5 cycle estimation both proposed methods with $L = 16$ perform comparably to the Matrix Pencil method.

The decrease of the RMSE to CRLB ratio observed in Fig. 8 for the proposed LS method for SNR close to zero is caused by the implementation and does not mean that the frequency estimation is more accurate. The frequency is computed in Matlab as $w_{LS} = \text{real}(\cos(c/2))/L$ thus the range of its values is limited to interval $[0, \pi/L]$, which saturates the error of frequency estimation. For low SNR the CRLB increases but the frequency estimation error saturates which results in the observed decrease of the RMSE to CRLB ratio.

4.2. Execution time

All known reference algorithms used for comparison are significantly more complex than the proposed two solutions. For $N = 64$ the Steiglitz–McBride method first solves the set of 62 equations for two unknown parameters, then filters the signal with a 3 tap IIR filter, solves the set of 64 equations for four unknown parameters, and computes the roots of the second order polynomial. In the Matrix Pencil method singular value decomposition of $43 \times 21$ matrix is required.

The LS-SDFT is also more computationally complex than the proposed solution in the frequency domain because it requires a higher number of DFT bins.

Fig. 9 depicts normalized Matlab execution time of the proposed downsampled methods, i.e. LS and DFT both with $L = \{1, N/8, N/4\}$ in comparison with the LS-SDFT with $R = \{N/8, N/4\}$ equations, and the Matrix Pencil as well as the Steiglitz–McBride methods. It is observed that the proposed LS solution is the fastest one. The second simplest is the proposed DFT based solution. In both cases a slight speed-up is observed with the increase of $L$. For $N = 16$ the LS solution is approximately 1.5 times faster than the DFT, and for $N = 512$ it is around 5 times faster. The
Matrix Pencil is the slowest and the Steiglitz–McBride is the second slowest method. On average, the proposed LS solution executes around 600 times faster than the Matrix Pencil and around 30 times faster than the Steiglitz–McBride.

Fig. 9. Normalized execution time in Matlab for different signal lengths $N = \{16, 32, 64, 128, 256, 512\}$.

5. Conclusions

We showed that downsampling introduced into the well-known autoregressive sinusoidal signal model can be used to increase the noise immunity of linear prediction based frequency estimation methods. The noise immunity is gained simultaneously with the reduction of computational complexity. The advantage over the counterpart methods i.e. the Prony and the SDFT is very significant. The variance of frequency estimation of the proposed solution is comparable with the best and computationally more complex methods: the Matrix Pencil and the Steiglitz–McBride.

The price for improved noise immunity and reduced computational complexity is narrowed frequency estimation range. Based on autoregressive sinusoidal signal model, the frequency can be estimated in the whole frequency range from to $\pi$ rad. By introducing downsampling by $L$ the measurement range is narrowed up to $\pi/L$ rad. Thus the signal must contain over $2L$ samples in one period.

However, narrowing the frequency estimation interval is not always a limitation. For example, the proposed frequency estimation methods fit very well to low cost implementation in standalone embedded systems, e.g. for $P$ class phasor estimation. By proper configuration of the proposed methods a close to optimal frequency estimator can be obtained for the frequency range of interest. Usually, comparable results require significantly higher computational burden.

In specific implementations the disturbances other than the WGN must be rejected by a proper signal preprocessing, i.e. the signal must be enhanced to be a single sinusoid in noise. Such a strategy, i.e. the prefilter followed by a general purpose low-latency frequency estimator was already verified by us for phasor estimation in [23].

References


**Krzysztof Duda** received the M.S. degree in Automatics and Metrology and the Ph.D. degree in Electronics from the AGH University of Science and Technology, Kraków, Poland, in 1998 and 2002, respectively. He has been with the Department of Measurement and Electronics, AGH University of Science and Technology, where, from 2002 to 2019, he was Assistant Professor, and since 2019, he has been Professor. His current research interests include the development and applications of digital signal processing and analysis.

**Tomasz P. Zieliński** received the M.S. degree in Electronics and the D.Sc. degree in Electrical Engineering from the AGH University of Science and Technology (AGH-UST), Kraków, Poland, in 1982 and 1996, respectively, and the Ph.D. degree in electrical engineering from the Bulgarian Academy of Sciences, Sofia, Bulgaria, in 1988. He was with the Department of Measurement and Instrumentation and the Department of Telecommunications, AGH-UST, in 1982, where he became Adjunct in 1989, Assistant Professor in 1996, Associate Professor in 2000, Professor of Telecommunications in 2003, and Full Professor in 2006. He has authored or co-authored over 150 scientific journal and conference papers. He has also authored three monographs (all in Polish) entitled Time-Frequency and Time-Scale Representations of Non-Stationary Signals (1996), From Theory to Digital Signal Processing (2002, 2004), and Digital Signal Processing: From Theory to Applications (2005, 2007, 2009, and 2014). His current research interests include advanced digital signal processing in telecommunication and biomedical systems, in particular, time-frequency signal analysis.